R

Random Graphs, A Whirlwind Tour of

FAN CHUNG
University of California, San Diego, USA

Article Outline

Glossary

Random graph a graph which is chosen from a certain probability distribution over the set of all graphs satisfying some given set of constraints. Statements about random graphs are thus statements about "typical" graphs for the given constraints.

Graph diameter the maximum distance between any pair of vertices in a graph. Here the distance between two vertices is the length of any shortest path joining the vertices.

Node degree distribution n_k the number of vertices having degree k.

Power law distribution a node degree distribution which (at least approximately) obeys $n_k \propto k^{-\beta}$ with fixed positive exponent β.

Erdős–Rényi random graph $G(n, p)$ a graph with n vertices, in which each possible edge between the $\binom{n}{2}$ pairs of vertices is present with probability p. The fact that the edges are chosen independently makes it relatively easy to compute properties of $G(n, p)$.

Random graph model $G(\mathbf{w})$ for expected degree sequence $\mathbf{w}$:] a model which generates random graphs which, *on average*, will have a prescribed degree sequence $\mathbf{w} = (w_1, w_2, \ldots, w_n)$.

On-line random graph a graph which grows in size over time, according to given probabilistic rules, starting from a start graph G_0. One can make statements about these graphs in the limit of long time (and hence large vertex number n, as n approaches infinity).

Introduction

Nowadays we are surrounded by assorted large information networks. For example, the phone network has all users as vertices which are interconnected by phone calls from one user to another. The Web can be viewed as a network with webpages as vertices which are then linked to other webpages. There are various biological networks arising from numerous databases, such as the gene network which represents the regulatory effect among genes. Of interest are many social networks expressing various types of social interactions. Some noted examples include the Collaboration graph (denoting coauthorship among mathematicians) and the Hollywood graph (consisting of actors/actresses and their joint appearances in feature films), among others.

How are these networks formed? What are basic structures of such large networks? How do they evolve? What are the underlying principles that dictate their behaviors?

To answer these questions, graph theory comes into play. Random graphs have a similar flavor as these large information networks in a natural way. For example, the phone network is formed by making random phone calls while a random graph results from adding a random edge one at a time. Although the classical random graphs can not directly be used to model real networks and seem to exhibit different 'shapes', the methods and approaches in random graph theory provides useful tools for the modeling and analysis of these information networks.

In this article, we will start with some basic graph theory in Sect. "Some Basic Graph Theory". We then introduce the main themes of random graphs in Sect. "Random

Graphs in a Nutshell". Then we consider the classical random graph theory in Sect. "Classical Random Graphs" before we proceed to describe some general random graph models with given degree distributions, in particular, the power law graphs in Sect. "Random Power Law Graphs". In Sect. "On-Line Random Graphs", we will cover two types of "on-line" graph models, including the model of preferential attachment and the duplication model.

Although random graphs can be used to analyze various aspects of realistic networks, we wish to point out that there is no silver bullet to answer all the difficult problems about these large complex networks. In the last section we will put things in perspective by clarifying what random graphs can and can not do.

Some Basic Graph Theory

All the information networks that we have mentioned can be formulated in terms of graphs. A graph G consists of a vertex set, denoted by $V = V(G)$ (which contains all the objects that we wish to deal with) and an edge set $E = E(G)$ which consists of of specified pairwise relations between vertices. For example, a friendship graph has the vertex set consisting of people of interest and the edge set denoting the pairs of people who are friends. In Table 1 we list a number of graphs associated with various networks.

As an introduction to graph theory, we describe the so-called *party problem*:

> *Among six people in a party, show that there are at least three people who know each other or there are three people who do not know each other.*

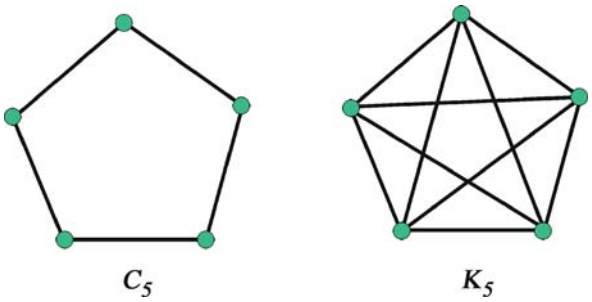
This can be said in graph-theoretical terms:

> *Any graph on 6 vertices must contain a triangle or contain three independent vertices with no edge among them.*

Indeed, 6 is the smallest number for this to occur since there is a graph on 5 vertices that contain neither a triangle nor three independent vertices. Such a graph is a cycle on 5 vertices, denoted by C_5, as seen in Fig. 1.

Random Graphs, A Whirlwind Tour of, Table 1
Graph models for several networks

Graph	Vertices	Edges
Flight schedule graph	Cities	Flights
Phone graph	Telephone numbers	Phone calls
Collaboration graph	Authors in Math Review	Coauthorship
Web graph	Webpages	Links
Biological graph	Genes	Regulatory effects



Random Graphs, A Whirlwind Tour of, Figure 1
A five cycle C_5 and a complete graph K_5

Let K_n denote a complete graph on n vertices which has all $\binom{n}{2}$ edges. For example, a triangle is K_3 which turns out is also C_3. The above party problem is a toy case of the so-called *Ramsey theory* which deals with unavoidable patterns in large graphs. In 1930, Ramsey [40] showed the following:

> *For any two positive integers k and l, there is an associated number R(k, l) such that any graph on $n \geq R(k, l)$ vertices must contain either K_k as a subgraph or contain l independent vertices.*

For example, $R(3, 3) = 6$ as stated in the party problem. It is not too difficult to show that $R(4, 4) = 17$. However, the value of $R(5, 5)$ is not yet determined (in spite of the huge computational power we have today). All that is known is $43 \leq R(5, 5) \leq 49$ (see [25] and [36]). Relatively few exact Ramsey numbers $R(k, l)$ are determined. For an extensive survey on this topic, the reader is referred to the dynamic survey in the Electronic Journal of Combinatorics at http://www.combinatorics.org/.

In 1947, Erdős wrote an important paper [21] that helped start two areas including combinatorial probabilistic methods and Ramsey theory. He established the following lower bound for the Ramsey number $R(k, k)$ by proving

$$R(k, k) \geq 2^{\frac{k}{2}}, \tag{1}$$

the argument is quite simple and elegant:

Suppose we wish to find a graph on n vertices that does not contain K_k or an independent subset of k vertices. How large can n be? For a fixed integer n, there are all together $2^{\binom{n}{2}}$ possible graphs on n vertices. We say a graph is *bad* if it contains K_k or an independent subset of k vertices. How many bad graphs can there be? There are $\binom{n}{k}$

ways to choose k out of n vertices. So, there are at most $2\binom{n}{k}2^{\binom{n}{2}-\binom{k}{2}}$ bad graphs. Therefore there is a graph on n vertices that is not bad if

$$2^{\binom{n}{2}} \geq 2\binom{n}{k}2^{\binom{n}{2}-\binom{k}{2}}.$$

So, for $n \geq 2^{k/2}$, there must be a graph on n vertices that is not bad, which implies (1).

We note that for the upper bound there is an inductive proof to show that $R(k,k) \leq \binom{2k-2}{k-2}$ which is about 4^k. In the previous five decades, there have been some improvements only by a factor of a lower order for both the upper and lower bounds [19,42]. It remains unsettled (with Erdős award unclaimed) to determine if $\lim_{k\to\infty}(R(k,k))^{1/k}$ exists or what value it should be.

A basic notion in graph theory is "adjacency". A vertex u is said to be *adjacent* to another vertex v if $\{u,v\}$ is an edge. Or, we say u is a *neighbor* of v. Equivalently, v is a neighbor of u. The *degree* of a vertex u is the number of edges containing u. If we restrict ourselves to *simple* graphs (i. e., at most one edge between any pairs of vertices), then the degree of u is just the number of neighbors that u has. Suppose that in a graph G the vertex v_i has degree d_i for $1 \leq i \leq n$. Then $(d_1, d_2, \ldots, d_n)$ forms a degree sequence for G. Sometimes, we organize the degree sequence so that $d_1 \geq d_2 \geq \cdots \geq d_n$. Here comes a natural question on graph realization: For what values d_i is the sequence $(d_1, d_2, \ldots, d_n)$ a degree sequence of some graph?

To answer this question, first we observe that the sum of all d_i's must be even since that is exactly twice the number of edges. This is the folklore "Handshake Theorem".

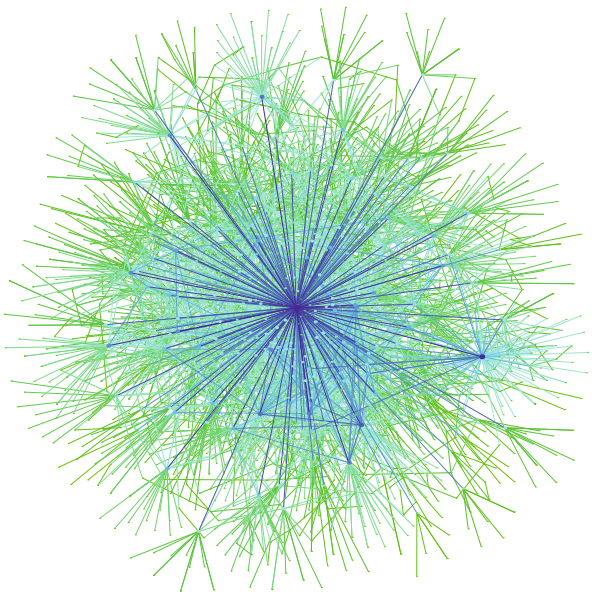
In a 1961 paper, Erdős and Gallai [24] answered the above question. They gave a necessary and sufficient condition by showing that a sequence $(d_1, d_2, \ldots, d_n)$, where $d_i \geq d_{i+1}$, is a degree sequence of some graph if and only if the sum of d_i's is even and for each integer $r \leq n-1$,

$$\sum_{i=1}^{r} d_i \leq r(r-1) + \sum_{i=r+1}^{n} \min\{r, d_i\}.$$

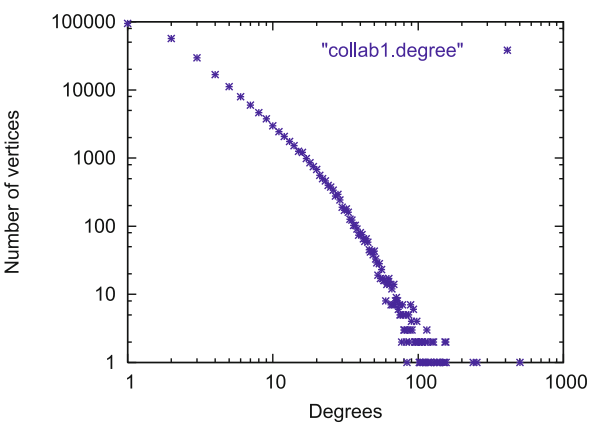
Another way to keep track of the degrees of a graph is to consider the degree distribution as follows: Let n_k denote the number of vertices having degree k. Instead of writing down the degree sequence (which consists of n numbers and n can be a very large number), we just use n_k. Therefore, the number of values that we need to keep does not exceed the maximum degree. If all degrees are the same value, we say the graph is regular. In this case, only one of the n_k's is nonzero.

Many real-world graphs have degree distribution satisfying the so-called "*power law*". Namely, the number n_k of vertices of degree k is proportional to $k^{-\beta}$ for some fixed positive value β. For example, the Collaboration graph, as illustrated in Fig. 2, can be approximated by a power law with exponent $\beta = 2.46$. The degree distribution of the Collaboration graph is included in Fig. 3 in log-log scale.

In a graph G, a *path* is a sequence of vertices v_0, $v_1, \ldots, v_k$ such that v_{i-1} is adjacent to v_i for $i = 1, \ldots, k$. The length of a path is the number of edges in the path. For example, the above mentioned path has length k join-



Random Graphs, A Whirlwind Tour of, Figure 2
The collaboration graph



Random Graphs, A Whirlwind Tour of, Figure 3
The number of vertices for each possible degree for the collaboration graph

ing v_0 and v_k. If $v_0 = v_k$, the path is said to be a *cycle*. A graph which contains no cycle is called a *tree*. A graph is *connected* if any two vertices can be joined by a path. For a graph G, a maximum subset of vertices each pair of which can be joined by paths is called a *connected component*. Thus, a graph is connected if there is only one connected component. In a connected graph, the *distance* between two vertices u and v is the length of a shortest path joining u and v. The maximum distance among all pairs of vertices is called the *diameter* of a graph.

In 1967, the psychologist Stanley Milgram [37] conducted a series of experiments which indicated that any two strangers are connected by a chain of intermediate acquaintances with the average chain length about six. Since then, the so-called "*small world phenomenon*" has long been a subject of anecdotal observation and folklore. Recent studies have suggested that the phenomenon is pervasive in numerous networks arising in nature and technology, and in particular, in the structural evolution of the World Wide Web [5,33,44].

In addition to "six degrees of separation", various numbers have emerged with many networks. In 1999, Barabási et al. [5] estimated that any two webpages are at most 19 clicks away from one another (in certain models of the Internet). Broder et al. [11] set up crawlers on a webgraph of 200 million nodes and 1.5 billion links and reported that the average path length is about 16. A mathematician who has written joint papers is likely to have Erdős number at most eight [29] (i. e., with a chain of coauthors with length at most 8 connecting to Erdős). The majority of actors or actresses have the so-called "Kevin Bacon number" two or three.

Before we make sense of these numbers, some clarification is in order: There are in fact two different interpretations of 'short distance' in a network. One notion is the diameter of the graph. Another notion is the *average distance* (which might be closer to what was meant by these experiments). We will discuss the small world phenomenon further in a later section.

Random Graphs in a Nutshell

What does a "random graph" mean? Before proceeding to describe random graphs, some clarification for "random" is in order. According to the Cambridge Dictionary, "random" means "happening, done or chosen by chance rather than according to a plan". Quite contrary to this explanation, our random graphs have precise meanings and can be clearly defined. Using the terminology in probability, a random graph is a random variable defined in a probability space with a probability distribution. In layman's terms, we first put all graphs on n vertices in a lottery box and then the graph we pick out of the box is a random graph. (In this case, all graphs are chosen with equal probability.)

What do we want from our random graphs? Well, we would like to say that a random graph (in some given model) has certain properties (e. g., having small diameter). Such a statement means that with probability close to 1 (as the number n of vertices approaches infinity), the random graph we pick out of the lottery box satisfies the property that we specified. In other words, a random graph has a specified property means that almost all graphs of interest have the desired property. Note that this is quite a strong implication! Any statement about a random graph is really about almost all graphs! The beauty of random graphs lies in being able to use relatively few parameters in the model to capture the behavior of almost all graphs of interest (which can be quite numerous and complex).

In the early days of the subject, Erdős and Rényi introduced two random graph models. The first one is a random graph $\mathcal{F}(n, m)$ defined on all graphs with n vertices and m edges each of which is chosen with equal probability. The second is the celebrated Erdős–Rényi random graph $\mathcal{G}(n, p)$ defined on all graphs on n vertices and each edge is chosen independently with probability p. Consequently, a graph on n vertices and x edges is chosen with probability $p^x(1-p)^{\binom{n}{2}-x}$ in $\mathcal{G}(n, p)$.

The advantage of the Erdős–Rényi model is the independence of choices for the edges (i. e., each pair of vertices has its own dice for determining being chosen as an edge). Since the probability of two independent events is the product of probabilities of two events, we can compute with ease. For example, the probability of a random graph in $\mathcal{G}(n, p)$ containing a fixed triangle is 1/8 for $p = 1/2$. It is possible to compute such a probability for a random graph in $\mathcal{F}(n, m)$, e. g., $\binom{\binom{n}{2}-3}{m-3}/\binom{\binom{n}{2}}{m}$, which is a more complicated expression. For many problems, such as the diameter problem, it can be quite nontrivial for $\mathcal{F}(n, m)$ because the dependency among edges is getting in the way.

To model real graphs, there are some obvious difficulties. For example, the random graph $\mathcal{G}(n, p)$ has all degrees very close to pn if the graph is not so sparse, (i. e., $p \geq \log n/n$). The distribution of the degrees follows the same bell curve for every vertex. As we know, many real-world graphs satisfy the power law which is very different from the degree distribution of $\mathcal{G}(n, p)$. In order to model real-world networks, it is imperative to consider random graphs with general degree distribution and, in particular, the power law distribution.

There are basically two types of random graph models for general degree distributions. The *configuration model* is a take-off from random regular graphs [6]. The way to define random regular graphs $\mathcal{G}_k$ of degree k on n vertices is to consider all possible matchings in a complete graph K_{kn}. Note that a matching is a maximum set of vertex-disjoint edges. Each matching is chosen with equal probability. We then get a random k-regular graph by partitioning the vertices of k into subsets. Each k-subset then is associated with a vertex in a random regular graph $\mathcal{G}_k$. Although such a random regular graph might contain loops (i. e., an edge having both endpoints the same vertex), the probability of such an event is of a lower order and can be controlled. It is then obvious to define random graphs with general degrees. Instead of partitioning the vertex set of the large graph into equal parts, we choose a random matching of a complete graph on $\sum_i d_i$ vertices which are partitioned into subsets of sizes $d_1, d_2, \ldots, d_n$. Then we form the random graph by associating each edge in the matching with an edge between associated vertices.

In the configuration model, there are nontrivial dependencies among the edges. There is another random graph model for general expected degrees as a generalization of the Erdős–Rényi model. Let $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ denote the specified degrees. The $\mathcal{G}(\mathbf{w})$ model yields random graphs with expected degrees $\mathbf{w}$. The edge between v_i and v_j is independently chosen with probability $w_i w_j / W$ where $W = \sum_i w_i$. In other words, each pair of vertices has its own dice with probability assigned so that the expected degree at vertex v_i is exactly w_i. The Erdős–Rényi model is just the case with all w_i's equal to pn. Since the $\mathcal{G}(\mathbf{w})$ model inherits the robustness and independence of the Erdős–Rényi model, many strong properties can be derived. We will discuss some of these further in Sect. "Parameters for Modeling Power Law Graphs", especially when $\mathbf{w}$ satisfies power laws.

All the random graph models mentioned above are *off-line* models. Since real-world graphs are dynamically changing—both in adding and deleting vertices and edges—there are several *on-line random graph models* in which the probability spaces are changing at the tick of the clock. In fact, in the study of complex real-world graphs, the on-line model came to attention first.

There are a large number of research papers, surveys and books on random graphs, mostly about the Erdős–Rényi model $\mathcal{G}(n, p)$. After the year 2000, the study of real-world graphs has led to interesting directions and new methods for analyzing random graphs with general degree distributions. Many on-line models have been proposed and published. Here we will only be able to cover the main ones—the preferential attachment schemes and the duplication model later.

Classical Random Graphs

In early 60's, Erdős and Rényi wrote a series of influential papers on random graphs. Their modeling and analysis are thorough and elegant. Their approaches and methods are powerful and have had enormous impact up to this day. In this section, we will give a brief overview. First we will describe the classical results on the evolution of random graphs $\mathcal{G}(n, p)$ of the Erdős–Rényi model. Then we will discuss the diameter of $\mathcal{G}(n, p)$ as the edge density ranges from 0 to 1.

The Evolution of the Erdős–Rényi Graph

What does a random graph in $\mathcal{G}(n, p)$ look like? Erdős and Rényi [23] gave a full answer for the edge density p ranging from 0 to 1.

At the start, there is no edge and the edge density is 0. We have isolated vertices.

As p increases, the expected number $p\binom{n}{2}$ of edges gets larger. When there are about $\sqrt{n}$ edges, how many connected components are there and what sizes and structures are they? For $0 < p \ll 1/n$, Erdős and Rényi [23] showed that the random graph G is a disjoint union of trees. Furthermore, they gave a beautiful formula. For $p = cn^{-k/(k-1)}$, the probability that j is the number of connected components in G formed by trees on k vertices is $\lambda^j e^{-\lambda}/j!$ where $\lambda = (2c)^{k-1}k^{k-2}/k!$. For example, when we have about $\sqrt{n}$ edges, the probability that the random graph G contains j trees on 3 vertices is close to $2^j/j!$ (if n is large enough).

As we have more edges, cycles start to appear. When the graph has a linear number of edges, i. e., $p = c/n$, with $c < 1$, almost all vertices are in connected components of trees and there are only a small number of cycles. Namely, the expected number of cycles is $1/2 \log 1/(1-c) - c/2 - (c^2)/4$.

When a random graph has edges ranging from slightly below $n/2$ to slightly over $n/2$ edges, i. e., $p = (1 + o(1))/n$, there is an unusual phenomenon, called "*double jumps*". What are double jumps and why is it so unusual? In the study of "threshold function" or, "phase transition" that happens in natural or evolving systems, it is of interest to identify the critical point, below which the behavior is dramatically different from what is above. Erdős and Rényi [23] found that as p is smaller than $1/n$, the largest component in G has size $O(\log n)$ and all components are either trees or unicyclic (i. e., each component contains at most one cycle). If p is $(1 + \mu)/n$ and $\mu > 0$, then the

Random Graphs, A Whirlwind Tour of, Table 2
The diameter of random graphs $\mathcal{G}(n, p)$

Range	diam($\mathcal{G}(n,p)$)	Reference
$\frac{np}{\log n} \to \infty$	Concentrated on at most 2 values	[9]
$\frac{np}{\log n} = c > 8$	Concentrated on at most 2 values	[12]
$8 \geq \frac{np}{\log n} = c > 2$	Concentrated on at most 3 values	[12]
$2 \geq \frac{np}{\log n} = c > 1$	Concentrated on at most 4 values	[10]
$1 \geq \frac{np}{\log n} = c > c_0$	Concentrated on at most $2\lfloor \frac{1}{c_0} \rfloor + 4$ values	[12]
$\log n > np \to \infty$	$\text{diam}(\mathcal{G}(n,p)) = (1 + o(1)) \frac{\log n}{\log(np)}$	[12]
$np \geq c > 1$	The ratio $\text{diam}(\mathcal{G}_{n,p}) / \left(\frac{\log n}{\log(np)} \right)$ is finite (between 1 and $f(c)$)	[12]
$np < 1$	diam($\mathcal{G}(n,p)$) equals the diameter of a tree component if $(1 - np)n^{1/3} \to \infty$	[35]

giant component emerges. However, when $p = 1/n$, the largest component is of size $O(n^{2/3})$. There has been detailed analysis examining this tricky transition in details (see [7] and [31]).

When $p = c/n$ for $c > 1$, the random graph G has one giant component and all others are quite small, of size $O(\log n)$. Also, Erdős and Rényi [23] determined the number of vertices in the giant connected component to be $f(c)n$ where

$$f(c) = 1 - \frac{1}{c} \sum_{k=1}^{\infty} \frac{k^{k-1}}{k!} \left(ce^{-c} \right)^k . \tag{2}$$

Finally when $p = c \log n/n$ and $c > 1$, the random graph G is almost always connected. When c gets larger, G is not only connected but gets closer to a regular graph. Namely, all vertices have degrees close to pn.

The Diameter of the Erdős–Rényi Graph

We consider the diameter of a random graph G in $\mathcal{G}(n, p)$ for all ranges of p including the range for which $\mathcal{G}(n, p)$ is not connected. For a disconnected graph G, the diameter of G is defined to be the diameter of its largest connected component.

Roughly speaking, the diameter of a random graph in $\mathcal{G}(n, p)$ is of order $\log n/(\log(np))$ if the expected degree np is at least 1. Note that this is best possible in the following sense. For any graph with degrees at most d, the number of vertices that can be reached within distance k is at most $1 + d + d(d-1) + d(d-1)^2 + \cdots + d(d-1)^{k-1}$. This sum should be at least n if k is the diameter. Therefore we known that the diameter, denoted by diam(G) is at least $(\log n)/\log(d-1)$.

To be precise, it can be shown the diameter of a random graph G in $\mathcal{G}(n, p)$ is $(1 + o(1))(\log n)/(\log np)$ if the expected degree np goes to infinity as n approaches infinity. When $np \geq c > 1$, the diameter diam(G) is within a constant factor of $(\log n)/(\log np)$ where the constant depends only on c and is independent of n. When $np = c < 1$, the random graph is surely disconnected and diam(G) is equal to the diameter of a tree component.

In fact, the diameter of a graph G in $\mathcal{G}(n, p)$ is quite predictable as follows. The values for the diameter of $\mathcal{G}(n, p)$ is almost surely concentrated on at most two values around $(\log n)/(\log np)$ if $(np)/(\log n) = c > 8$. When $(np)/(\log n) = c > 2$, the diameter of $\mathcal{G}(n, p)$ is almost surely concentrated on at most three values. For the range $2 \geq (np)/(\log n) = c > 1$, the diameter of $\mathcal{G}(n, p)$ is almost surely concentrated on at most four values.

It is of particular interest to consider random graphs $\mathcal{G}(n, p)$ for the range of $np > 1$ and $np \leq c \log n$ for some constant c since this range includes the emergence of the unique giant component. Because of a phase transition in connectivity at $p = \log n/n$, the problem of determining the diameter of $\mathcal{G}(n, p)$ and its concentration seems to be difficult for certain ranges of p. If $(np)/(\log n) = c > c_0$ for any (small) constant c and c_0, then the diameter of $\mathcal{G}(n, p)$ is almost surely concentrated on finitely many values, namely, no more than $2\lfloor 1/(c_0) \rfloor + 4$ values.

These facts are summarized in Table 2 with references listed. As we can see from the table, numerous questions remain.

Random Power Law Graphs

Parameters for Modeling Power Law Graphs

A large realistic network usually has a huge number of parameters with complicated descriptions. By "modeling a realistic network", we mean cutting down the number of parameters to relatively few and still capture a good part of the character of the network.

To choose the parameters for modeling a real network, the exponent β of the power law is relatively easy to select.

We can plot the log-degree versus log-frequency table and choose a good approximation of the slope.

In a graph G, suppose that there are y vertices of degree x. Then G is considered to be a power law graph if x and y satisfy (or can be approximated by) the following equation:

$$\log y = \alpha - \beta \log x \,. \tag{3}$$

In other words, we have

$$|\,\{v|\deg(v) = x\}\,| \approx y = \frac{e^{\alpha}}{x^{\beta}} \,.$$

Basically, α is the logarithm of the volume of the graph and β can be regarded as the log-log growth rate of the graph.

To take a closer look of the degree distribution of a typical realistic graph, several impediments obviously exist.

(a) When we fit the power law model, there are discrepancies especially when the degree is very small or very large. There is almost always a heavy tail distribution at the upper range and there seems to be scattering at the lower range. For example, for the collaboration graph, should we or shouldn't we include the data point for isolated vertices (an author with no coauthors)? Should we stay with the largest component or include all small components (including the isolated vertices)?

(b) The power law states that the number of vertices of degree k is proportional to $k^{-\beta}$. We can approximate the number of vertices of degree k by the function $f(k) = ck^{-\beta}$ for some constant c. However, $f(k)$ is usually not an integer. By taking either the ceiling or floor of $f(k)$, some errors are inevitable. In fact, such errors are acute when k or $f(k)$ is small.

(c) The power law model is usually a better fit in the middle range (than at either end). Still, in many examples, there is a visible slight "hump" in the curve instead of the straight line representing the power law in the log-log table.

Among the above three points, (c) is mainly due to first-order approximations. The straight line with slope β is a linear approximation of the actual plotted data. Thus the power law model is an important and necessary step for more complicated real cases. Here, we will first discuss (b) and then (a).

Item (b) concerns rounding errors which can be checked by the following basic calculations about the power law graphs as described by Eq. (3).

(1) The maximum degree of the graph is at most $e^{\alpha/\beta}$. Note that $0 \le \log y = \alpha - \beta \log x$.

(2) The number of vertices n can be computed as follows (under the assumption that the maximum degree is $e^{\alpha/\beta}$). By summing $y(x)$ for x from 1 to $e^{\alpha/\beta}$, we have

$$n = \sum_{x=1}^{e^{\frac{\alpha}{\beta}}} \frac{e^{\alpha}}{x^{\beta}} \approx \begin{cases} \zeta(\beta)e^{\alpha} & \text{if } \beta > 1\,, \\ \alpha e^{\alpha} & \text{if } \beta = 1\,, \\ \frac{e^{\frac{\alpha}{\beta}}}{1-\beta} & \text{if } 0 < \beta < 1\,, \end{cases}$$

where $\zeta(t) = \sum_{n=1}^{\infty} 1/(n^t)$ is the Riemann Zeta function.

(3) The number of edges E can be computed as follows:

$$E = \frac{1}{2}\sum_{x=1}^{e^{\frac{\alpha}{\beta}}} x\frac{e^{\alpha}}{x^{\beta}} \approx \begin{cases} \frac{1}{2}\zeta(\beta-1)e^{\alpha} & \text{if } \beta > 2\,, \\ \frac{1}{4}\alpha e^{\alpha} & \text{if } \beta = 2\,, \\ \frac{1}{2}\frac{e^{\frac{2\alpha}{\beta}}}{2-\beta} & \text{if } 0 < \beta < 2\,. \end{cases}$$

(4) The differences of the real numbers in (1)–(3) and their integer parts can be estimated as follows: For the number n of vertices, the error term is at most $e^{\alpha/\beta}$. For $\beta \ge 1$, it is $o(n)$, which is a lower order term. For $0 < \beta < 1$, the error term for n is relatively large. In this case, we have

$$n \le \frac{e^{\frac{\alpha}{\beta}}}{1-\beta} - e^{\frac{\alpha}{\beta}} = \frac{\beta e^{\frac{\alpha}{\beta}}}{1-\beta} \,.$$

As can be seen, n can have the same magnitude as $(e^{\alpha/\beta})/(1-\beta)$. Therefore the rounding error can be of the same order of magnitude. For the number E of edges, similar situations occur. For $\beta \ge 2$, the rounding error term of E is $o(E)$, a lower order term. For $0 < \beta < 2$, the error of E has the same magnitude as in the formula of item (3). Thus, one is advised to exercise caution when dealing with the case $0 < \beta < 2$.

To deal with the concerns mentioned above in (a), we need additional parameters.

- The average degree w is a useful parameter.
- The second order average degree $\tilde{w} = \sum_i w_i^2 / \sum_i w_i$.
- The maximum degree $m = d_{\max}$ and also the minimum degree $d_{\min}$ denote the range that the power law degree distribution fits (within acceptable approximation). In other words, the maximum degree $m = d_{\max}$ and the minimum degree $d_{\min}$ are meant to be the largest and the least degrees in a power law subgraph of G. Often, $d_{\min}$ is taken to be 1 unless otherwise specified.

With these parameters, we are ready to define a random power law graph. For random graphs with given expected degree sequences satisfying a power law distribution with exponent β, we may assume that the expected degrees

are $w_i = ci^{-\frac{1}{\beta-1}}$ for i satisfying $i_0 \le i < n + i_0$. Here c depends on the average degree and i_0 depends on the maximum degree m, namely, $c = (\beta-2)/(\beta-1)\, wn^{1/(\beta-1)}$, $i_0 = n((w(\beta-2))/(m(\beta-1)))^{\beta-1}$.

The power law graphs with exponent $\beta > 3$ are quite different from those with exponent $\beta < 3$ as evidenced by the value of $\tilde{w}$ (assuming $m \gg w$).

$$\tilde{w} = \begin{cases} (1+o(1))w\frac{(\beta-2)^2}{(\beta-1)(\beta-3)} & \text{if } \beta > 3\,, \\ (1+o(1))\frac{1}{2}w\ln\frac{2m}{w} & \text{if } \beta = 3\,, \\ (1+o(1))d^{\beta-2}\frac{(\beta-2)^{\beta-1}m^{3-\beta}}{(\beta-1)^{\beta-2}(3-\beta)} & \text{if } 2 < \beta < 3\,. \end{cases}$$

The above values of $\tilde{w}$ are quite useful in the study of average distance and diameter of random graphs.

The Evolution of Random Power Law Graphs

A natural question concerning the configuration model is how the random graphs evolve for power law distributions. Can we mimic the classical analysis as in the Erdős–Rényi random graph model?

Here we consider a configuration model with degree distribution as in the (α, β)-graph. As it turns out, the evolution only depends on β and not on α as follows.

1. When $\beta > \beta_0 = 3.47875\ldots$, the random graph almost surely has no giant component where the value $\beta_0 = 3.47875\ldots$ is a solution to

   $$\zeta(\beta-2) - 2\zeta\beta - 1) = 0\,.$$

   When $\beta < \beta_0 = 3.47875\ldots$, there is almost surely a unique giant component.
2. When $2 < \beta < \beta_0 = 3.47875\ldots$, the second largest component is almost surely of size $\Theta(\log n)$. For any $2 \le x < \Theta(\log n)$, there is almost surely a component of size x.
3. When $\beta = 2$, almost surely the second largest component is of size $\Theta(\log n/(\log\log n))$. For any $2 \le x < \Theta(\log n/(\log\log n))$, there is almost surely a component of size x.
4. When $1 < \beta < 2$, the second largest component is almost surely of size $\Theta(1)$. The graph is almost surely not connected.
5. When $0 < \beta < 1$, the graph is almost surely connected.
6. When $\beta = \beta_0 = 3.47875\ldots$, the situation is complicated. It is similar to the double jump of the random graph $G(n,p)$ with $p = 1/n$. For $\beta = 1$, there is a nontrivial probability for either case that the graph is connected or disconnected.

A useful tool in configuration model is a result of Molloy and Reed [38,39]:

For a random graph with $(\gamma_i + o(1))n$ vertices of degree i, where γ_i are nonnegative values which sum to 1 and n is the number of vertices, the giant component emerges when $Q = \sum_{i\ge1} i(i-2)\gamma_i > 0$, provided that the maximum degree is less than $n^{1/4-\epsilon}$ and some "smoothness" conditions are satisfied. Also, there is almost surely no giant component when $Q = \sum_{i\ge1} i(i-2)\gamma_i < 0$ and the maximum degree is less than $n^{1/8-\epsilon}$.

Let us consider Q for our (α, β)-graphs with $\beta > 3$.

$$\begin{aligned} Q &= \frac{1}{n}\sum_{x=1}^{e^{\frac{\alpha}{\beta}}} x(x-2)\lfloor\frac{e^\alpha}{x^\beta}\rfloor \\ &\approx \frac{1}{\zeta(\beta)}\left(\sum_{x=1}^{e^{\frac{\alpha}{\beta}}}\frac{1}{x^{\beta-2}} - 2\sum_{x=1}^{e^{\frac{\alpha}{\beta}}}\frac{1}{x^{\beta-1}}\right) \\ &\approx \frac{\zeta(\beta-2)-2\zeta(\beta-1)}{\zeta(\beta)} \end{aligned}$$

Hence, we consider the value $\beta_0 = 3.47875\ldots$, which we recall is a solution to $\zeta(\beta-2) - 2\zeta(\beta-1) = 0$. If $\beta > \beta_0$, we have

$$\sum_{x=1}^{e^{\frac{\alpha}{\beta}}} x(x-2)\lfloor\frac{e^\alpha}{x^\beta}\rfloor < 0\,.$$

We remark that for $\beta > 8$, Molloy and Reed's result immediately implies that almost surely there is no giant component. When $\beta \le 8$, additional analysis is needed to deal with the degree constraints [2].

It can be shown that the second largest component almost surely has size $\Theta(\log n)$. Furthermore, the second largest component has size at least $\Theta(\log n)$.

$G(\mathbf{w})$ Model for Power Law Graphs

In the Erdős–Rényi model $G(n,p)$, the threshold function for the phase transition of the giant component is at $p = 1/n$. Namely, when the average degree pn is less than 1, all connected components are small (of size $O(\log n)$) and there is no giant component. When the average degree is more than 1, the giant component emerges in full swing. (There is a "double jump" which takes place when the average degree is close to 1 as discussed in Sect. "Random Graphs in a Nutshell".)

For the random graph model $G(\mathbf{w})$, with given expected degrees $\mathbf{w}$, it is natural to ask the same question:

What parameter in which range will trigger the (sudden) emergence of the giant component?

In addition to w, the expected average degree, we have scores of parameters, e. g., $\tilde{w}$ and higher order average degrees. Which parameter w, $\tilde{w}$ or others is critical for the rise of the giant component?

These questions were answered in [13]:

Suppose that G is a random graph in $\mathcal{G}(\mathbf{w})$ with expected degree sequence $\mathbf{w}$. *If the expected average degree w is strictly greater than 1, then the following holds:*

(1) *Almost surely G has a unique giant component. Furthermore, the volume of the giant component is at least* $(1 - \frac{2}{\sqrt{w\,\mathrm{e}}} + o(1))\mathrm{Vol}(G)$ *if* $w \geq 4/\mathrm{e} = 1.4715\ldots$, *and is at least* $(1 - (1 + \log w)/w + o(1))\mathrm{Vol}(G)$ *if* $w < 2$.

(2) *The second largest component almost surely has size at most* $(1 + o(1))\mu(w)\log n$, *where*

$$\mu(w) = \begin{cases} \frac{1}{1+\log w - \log 4} & \text{if } w > \frac{4}{\mathrm{e}}\,; \\ \frac{1}{w-1-\log w} & \text{if } 1 < w < 2\,. \end{cases}$$

Moreover, with probability at least $1 - n^{-k}$, the second largest component has size at most $(k + 1 + o(1))\mu(w)\log n$, *for any* $k \geq 1$.

There is a sharp asymptotic estimate for the volume of the giant component for a random graph in $\mathcal{G}(\mathbf{w})$. In [15], it was proved that if the expected average degree is strictly greater than 1, then almost surely the giant component in a graph G in $\mathcal{G}(\mathbf{w})$ has volume $\lambda_0\mathrm{Vol}(G) + O(\sqrt{n}\log^{3.5} n)$, where λ_0 is the unique nonzero root of the following equation:

$$\sum_{i=1}^{n} w_i \mathrm{e}^{-w_i\lambda} = (1-\lambda)\sum_{i=1}^{n} w_i\,. \tag{4}$$

Because of the robustness of the $\mathcal{G}(\mathbf{w})$ model, many properties can be derived for appropriate degree distributions, including power law graphs.

Average Distance and the Diameter A random graph G in $\mathcal{G}(\mathbf{w})$ has average distance almost surely $(1 + o(1)) \cdot (\log n)/(\log \tilde{w})$, if $\mathbf{w}$ satisfies certain conditions (called admissible conditions in [14]). The diameter is almost surely $\Theta((\log n)/(\log \tilde{w}))$. In addition to studying the average distance and diameter, the structure of a random power law graph is very interesting, especially for the range $2 < \beta < 3$ where the power law exponents β for numerous real networks reside. In this range, the power law graph can be roughly described as an "octopus" with a dense subgraph having small diameter $O(\log\log n)$, as the core while the overall diameter is $O(\log n)$ and the average distance is $O(\log\log n)$. When $\beta > 3$ and the average degree w is strictly greater than 1, almost surely the average distance is $(1 + o(1))(\log n)/(\log \tilde{w})$ and the diameter is $\Theta(\log n)$. A phase transition occurs at $\beta = 3$ and then the graph has diameter almost surely $\Theta(\log n)$ and average distance $\Theta(\log n/\log\log n)$.

Eigenvalues Eigenvalues of graphs are useful for controlling many graph properties and consequently have numerous algorithmic applications including clustering algorithms, low rank approximations, information retrieval and computer vision. In the study of the spectra of power law graphs, there are basically two competing approaches. One is to prove analogues of Wigner's semi-circle law (such as for $\mathcal{G}(n, p)$) while the other predicts that the eigenvalues follow a power law distribution [27]. Although the semi-circle law and the power law have nothing in common, both approaches are essentially correct if one considers the appropriate matrices. there are in fact several ways to associate a matrix to a graph. The usual adjacency matrix A associated with a (simple) graph has eigenvalues quite sensitive to the maximum degree (which is a *local* property). The combinatorial Laplacian $D - A$ with D denoting the diagonal degree matrix is a major tool for enumerating spanning trees and has numerous applications. Another matrix associated with a graph is the (normalized) Laplacian $L = I - D^{-1/2}AD^{-1/2}$ which controls the expansion/isoperimetrical properties (which are *global*) and essentially determines the mixing rate of a random walk on the graph. The traditional random matrices and random graphs are regular or almost regular so the spectra of all the above three matrices are basically the same (with possibly a scaling factor or a linear shift). However, for graphs with uneven degrees, the above three matrices can have very different distributions.

Here we state bounds for eigenvalues for random graphs in $\mathcal{G}(\mathbf{w})$ with a general degree distribution from which the results on random power law graphs then follow [18].

1. The largest eigenvalue of the adjacency matrix of a random graph with a given expected degree sequence is determined by m, the maximum degree, and $\tilde{w}$, the weighted average of the squares of the expected degrees. In this case the largest eigenvalue of the adjacency matrix is almost surely $(1 + o(1))\max\{\tilde{w}, \sqrt{m}\}$ provided some minor conditions are satisfied. In addition, if the kth largest expected degree m_k is significantly larger than $\tilde{w}^2$, then the kth largest eigenvalue of the adjacency matrix is almost surely $(1 + o(1))\sqrt{m_k}$.

2. For a random power law graph with exponent $\beta > 2.5$, the largest eigenvalue of a random power law graph is almost surely $(1 + o(1))\sqrt{m}$ where m is the maximum degree. Moreover, the k largest eigenvalues of a random power law graph with exponent β have power law distribution with exponent $2\beta - 1$ if the maximum degree is sufficiently large and k is bounded above by a function depending on β, m and w, the average degree. When $2 < \beta < 2.5$, the largest eigenvalue is heavily concentrated at $cm^{3-\beta}$ for some constant c depending on β and the average degree.
3. The eigenvalues of the Laplacian satisfy the semi-circle law under the condition that the minimum expected degree is relatively large ($\gg$ the square root of the expected average degree). This condition contains the basic case when all degrees are equal (the Erdős–Rényi model). If we weaken the condition on the minimum expected degree, we can still have the following strong bound for the eigenvalues of the Laplacian which implies strong expansion rates for rapid mixing,

$$\max_{i \neq 0} |1 - \lambda_i| \leq (1 + o(1))\frac{4}{\sqrt{w}} + \frac{g(n)\log^2 n}{w_{\min}}$$

where w is the expected average degree, $w_{\min}$ is the minimum expected degree and $g(n)$ is any slow growing function of n.

On-Line Random Graphs

Preferential Attachment Schemes

The preferential attachment scheme is often attributed to Herbert Simon. In his paper [41] of 1955, he gave a model for word distribution using the preferential attachment scheme and derived *Zipf's law* (i. e., the probability of a word having occurred exactly i times is proportional to $1/i$).

The basic setup for the preferential attachment scheme is a simple *local* growth rule which leads to a *global* consequence—a power law distribution. Since this local growth rule gives preferences to vertices with large degrees, the scheme is often described by "*the rich get richer*". Of interest is to determine the exponent of the power law from the parameters of the local growth rule.

There are two parameters for the preferential attachment model:

- A probability p, where $0 \leq p \leq 1$.
- An initial graph G_0, that we have at time 0.

Usually, G_0 is taken to be the graph formed by one vertex having one loop. (We consider the degree of this vertex to be 1, and in general a loop adds 1 to the degree of a vertex.) Note, in this model multiple edges and loops are allowed.

We also have two operations we can do on a graph:

Vertex-step Add a new vertex v, and add an edge $\{u, v\}$ from v by randomly and independently choosing u in proportion to the degree of u in the current graph.

Edge-step Add a new edge $\{r, s\}$ by independently choosing vertices r and s with probability proportional to their degrees.

Note that for the edge-step, r and s could be the same vertex. Thus loops could be created. However, as the graph gets large, the probability of adding a loop can be well bounded and is quite small.

The random graph model $G(p, G_0)$ is defined as follows:

Begin with the initial graph G_0.
For $t > 0$, at time t, the graph G_t is formed by
modifying G_{t-1} as follows:
with probability p, take a vertex-step,
otherwise, take an edge-step.

When G_0 is the graph consisting of a single loop, we will simplify the notation and write $G(p) = G(p, G_0)$.

There were quite a number of papers analyzing the preferential attachment model $G(p)$, usually having similar conclusions of power law degree distribution. However, many of these analyses are heuristics without specifying the ranges for the power law to hold. Heuristics often run into the danger of incorrect deductions and incomplete conclusions. It is quite essential to use rigorous proofs which help specify the appropriate conditions and ranges for the power law. The following statement was proved in [16].

For the preferential attachment model $G(p)$, almost surely the number of vertices with degree k at time t is

$$M_k t + O(2\sqrt{k^3 t \ln(t)}) .$$

where $M_1 = (2p)/(4 - p)$ and $M_k = (2p)/(4 - p)(\Gamma(k)\Gamma(1 + 2/(2 - p)))/(\Gamma(k + 1 + 2/(2 - p))) = O(k^{-(2+p/(2-p))})$, for $k \geq 2$. In other words, almost surely the graphs generated by $G(p)$ have the power law degree distribution with the exponent $\beta = 2 + p/(2 - p)$.

Duplication Models

Networks of interactions are present in all biological systems. The interactions among species in ecosystems, between cells in an organism and among molecules in a cell all lead to complex biological networks. Using current technological advances, extensive data of such interactions has been acquired. To find the underlying structure in

Random Graphs, A Whirlwind Tour of, Table 3
Power law exponents for biological and nonbiological networks

Biological networks	Exponent β	References
Yeast protein -protein net	1.6, 1.7	[20,43]
E. Coli metabolic net	1.7, 2.2	[3,28]
Yeast gene expression net	1.4–1.7	[20]
Gene functional interaction	1.6	[30]
Nonbiological networks		
Internet graph	2.2 (indegree), 2.6 (outdegree)	[4,27,34]
Phone call graph	2.1 –2.3	[1,2]
Collaboration graph	2.4	[29]
Hollywood graph	2.3	[4]

these databases, it is of great importance to understand the basic principles of various genetic and metabolic networks.

It has been observed that many biological networks have power law graphs with exponents β less than 2. The ranges for the exponents of the power law for biological networks are quite different from the ranges for nonbiological networks. Various examples, such as the WWW-graphs, call graphs, and various social networks, among others, are power law graphs with the exponent β between 2 and 3. Table 3 lists the exponents of a variety of biological and nonbiological networks with associated references. As we saw in Sect. "Preferential Attachment Schemes", the preferential attachment model generates graphs with power law degree distribution with exponents β between 2 and 3. Therefore there is a need to consider alternative models for biological networks.

The duplication of the information in the genome – genes and their controlling elements – is a driving force in evolution and a determinative factor of biological networks. The process of duplication is quite different from the preferential attachment process that is regarded by many as the basic growth rule for most nonbiological networks.

Here we consider a duplication model. If we only allow pure duplication, the resulting graph depends heavily on the initial graph and does not satisfy the power law. So we consider a duplication model that allows randomness within the duplication step as defined below. We will see that this duplication model generates power law graphs with exponents in the range including the interval between 1 and 2 and therefore is more suitable for modeling complex biological networks.

There are two basic parameters for the duplication model:

- A selection probability p, where $0 \leq p \leq 1$.
- An initial graph G_0, that we have at time 0.

Usually, G_0 is taken to be the graph formed by one vertex. However, G_0 can be taken to be any finite simple connected graph. Unlike the preferential attachment model, in this model the generated random graph is always a simple graph.

There is one basic operation:

Duplication step: A *sample* vertex u is selected randomly and uniformly from the current graph. A new vertex v and edge $\{u, v\}$ is added to the graph. For each neighbor w of u, with probability p, $\{v, w\}$ is added as a new edge.

The edge $\{u, v\}$ in the duplication step is called a *basic* edge. The vertex u is said to be the *parent* of v and v is called a *child* of u. We note that a vertex can have several children or no child at all and that each vertex not in the initial graph G_0 has a parent. All basic edges from children to parents form a forest where the vertices in G_0 are roots of component trees. All terms like "leafs" and "descendants", if not defined, refer to this forest.

The duplication step can be further decomposed into two parts—vertex-duplication and edge-duplication as follows:

Vertex-duplication: At time t, randomly select a sample vertex u and add a new vertex v and an edge $\{u, v\}$.

Edge-duplication: At time t, for the vertex v created, its parent u and each neighbor w of u, with probability p, add an edge $\{v, w\}$ to w.

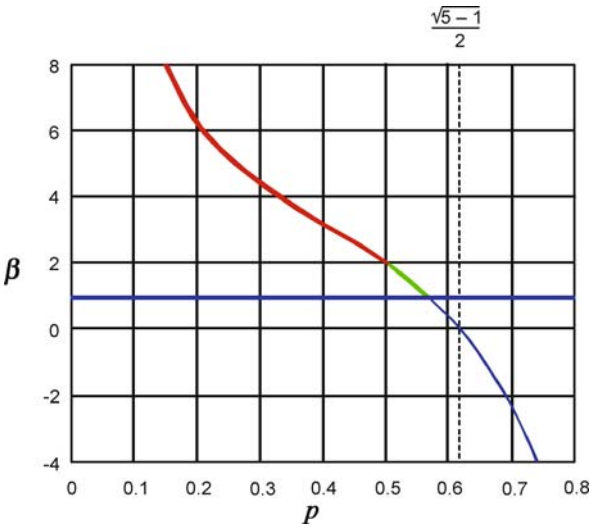
For any vertex v, a *descendant* of v can only be connected to descendants of v's neighbors (including v itself). An edge $\{x, y\}$ is said to be a descendant of an edge $\{u, v\}$, if "x is a descendant of u and y is a descendant of v" or "x is a descendant of v and y is a descendant of u".

We remark that having the basic edges $\{u, v\}$ makes the graph G always connected. This helps avoid degenerate cases such as having mostly isolated vertices.

For the above duplication model, it can be shown [17] that its degree distribution obeys a power law with the exponent β of the power law satisfying the following equation:

$$1 + p = p\beta + p^{\beta-1} . \tag{5}$$

We remark that the solutions for (5) that are illustrated in Fig. 4 consist of two parts. One is the line $\beta = 1$. The other is a curve which is a monotonically decreasing function of p. The two curves intersect at $(x, 1)$ where $x = 0.56714329\ldots$ the solution of $x = -\log x$. One very interesting range for β is when p is near 1/2. To get a power-law with exponent 1.5, for example, one should choose $p = 0.535898\ldots$ Also we see that the second curve intersects zero at $p = (\sqrt{5} - 1)/2$, an intriguing number (the "golden mean"). At $p = 1/2$, one solution for β is 2. Although there are two solutions for each p, the stable



Random Graphs, A Whirlwind Tour of, Figure 4
The value of β as a function of p. The *green* section of the curve shows the range of β values which are typically found for biological networks

solutions are on the curve when $p < 0.56714329\ldots$ and $\beta = 1$ for $p > 0.56714329\ldots$

Remarks

The small world phenomenon, that occurs ubiquitously in numerous existing networks, refers to two similar but different properties:

Small distance Between any pair of nodes, there is a short path.

The clustering effect Two nodes are more likely to be adjacent if they share a common neighbor.

There have been various approaches to model networks that exhibit the small world phenomenon. In particular, the aspect of small distances can be well explained by using random graphs with general degree distributions which include the power law distribution. However, the other feature concerning the clustering effect seems much harder to model.

To model the clustering effect, a typical approach is to add random edges to a grid graph or the like [26,32,44]. Such grid-based models are quite restrictive and far from satisfactory for modeling biological networks or collaboration graphs, for example. On the other hand, random power law graphs are good for modeling small distance, but fail miserably for modeling the clustering effect. In a way, the aspect of small distances is about neighborhood expansion while the aspect of the clustering effect is about neighborhood density. The related graph-theoretical parameters seem to be of an entirely different scale. For example, while the clustering effect is quite sensitive to average degree, the small distance effect is not.

The heart of the problem can be quite simply stated: For a given network, what is its true geometry? How can we capture the geometry of the network (without invoking too many parameters)?

Bibliography

1. Abello J, Buchsbaum A, Westbrook J (1998) A functional approach to external graph algorithms. In: Proc. 6th European Symposium on Algorithms. Springer, Berlin, pp 332–343
2. Aiello W, Chung F, Lu L (2000) A random graph model for massive graphs. In: Proc. of the Thirty-Second Annual ACM Symposium on Theory of Computing. ACM Press, New York, pp 171–180
3. Albert R, Barabási A-L (2002) Statistical mechanics of complex networks. Rev Mod Phys 74:47–97
4. Barabási A-L, Albert R (1999) Emergence of scaling in random networks. Science 286:509–512
5. Barabási A-L, Albert R, Jeong H (2000) Scale-free characteristics of random networks: the topology of the world-wide web. Physica A 281:69–77
6. Bender EA, Canfield ER (1978) The asymptotic number of labeled graphs with given degree sequences. J Comb Theory Ser A 24:296–307
7. Bollobás B (1984) The evolution of random graphs. Trans Amer Math Soc 286:257–274
8. Bollobás B (1981) The diameter of random graphs. Trans Amer Math Soc 267:41–52
9. Bollobás B (1985) Random graphs. Academic Press, New York, pp xvi+447
10. Bollobás B (1984) The evolution of sparse graphs. In: Graph theory and combinatorics. Academic Press, London, New York, pp 35–57, (Cambridge 1983)
11. Broder A, Kumar R, Maghoul F, Raghavan P, Rajagopalan S, Stata R, Tomkins A, Wiener J (2000) Graph structure in the web. Proc. of the WWW9 Conference, Amsterdam, May 2000, Paper version appeared in Computer Networks 33, pp 309–320
12. Chung F, Lu L (2001) The diameter of sparse random graphs. Adv Appl Math 26:257–279
13. Chung F, Lu L (2002) Connected components in random graphs with given expected degree sequences. Ann Comb 6:125–145
14. Chung F, Lu L (2002) The average distances in random graphs with given expected degrees. Proc Nat Acad Sci 99:15879–15882
15. Chung F, Lu L (2006) The volume of the giant component of a random graph with given expected degrees. SIAM J Discret Math 20:395–411
16. Chung F, Lu L (2006) Complex graphs and networks. In: CBMS Lecture Series, vol 107. AMS Publications, Providence, pp vii+264
17. Chung F, Lu L, Dewey G, Galas DJ (2003) Duplication models for biological networks. J Comput Biol 10(5):677–687
18. Chung F, Lu L, Vu V (2003) The spectra of random graphs with given expected degrees. Proc Nat Acad Sci 100(11):6313–6318

19. Conlon D A new upper bound for diagonal Ramsey numbers. Ann Math (in press)
20. Dorogovtsev SN, Mendes JFF (2001) Scaling properties of scale-free evolving networks: continuous approach. Phys Rev E 63(056125):19
21. Erdős P (1947) Some remarks on the theory of graphs. Bull Amer Math Soc 53:292–294
22. Erdős P, Rényi A (1959) On random graphs, vol I. Publ Math Debrecen 6:290–297
23. Erdős P, Rényi A (1960) On the evolution of random graphs. Magyar Tud Akad Mat Kutató Int Közl 5:17–61
24. Erdős P, Gallai T (1961) Gráfok előírt fokú pontokkal (Graphs with points of prescribed degrees, in Hungarian). Mat Lapok 11:264–274
25. Exoo G (1989) A lower bound for $R(5, 5)$. J Graph Theory 13:97–98
26. Fabrikant A, Koutsoupias E, Papadimitriou CH (2002) Heuristically optimized trade-offs: a new paradigm for power laws in the internet. In: Automata, languages and programming. Lecture Notes in Computer Science, vol 2380. Springer, Berlin, pp 110–122
27. Faloutsos M, Faloutsos P, Faloutsos C (1999) On power-law relationships of the Internet topology. In: Proc. of the ACM SIGCOM Conference. ACM Press, New York, pp 251–262
28. Friedman R, Hughes A (2001) Gene duplications and the structure of eukaryotic genomes. Genome Res 11:373–381
29. Grossman J, Ion P, De Castro R (2007) The Erdős number project. http://www.oakland.edu/enp
30. Gu Z, Cavalcanti A, Chen F-C, Bouman P, Li W-H (2002) Extent of gene duplication in the genomes of drosophila, nematode, and yeast. Mol Biol Evol 19:256–262
31. Janson S, Knuth DE, Łuczak T, Pittel B (1993) The birth of the giant component. Rand Struct Algorithms 4:231–358
32. Kleinberg J (2000) The small-world phenomenon: an algorithmic perspective. Proc. 32nd ACM Symposium on Theory of Computing. ACM, New York, pp 163–170
33. Kleinberg J, Kumar R, Raphavan P, Rajagopalan S, Tomkins A (1999) The web as a graph: measurements, models and methods. In: Proc. of the International Conference on Combinatorics and Computing. Lecture Notes in Computer Science, vol 1627. Springer, Berlin, pp 1–17
34. Kumar R, Raghavan P, Rajagopalan S, Tomkins A (1999) Trawling the web for emerging cyber communities. In: Proc. of the 8th World Wide Web Conference. Toronto, 1999
35. Łuczak T (1998) Random trees and random graphs. Rand Struct Algorithms 13:485–500
36. McKay BD, Radziszowski SP (1994) Subgraph counting identities and Ramsey numbers. J Comb Theory (B) 61:125–132
37. Milgram S (1967) The small world problem. Psychol Today 2:60–67
38. Molloy M, Reed B (1995) A critical point for random graphs with a given degree sequence. Rand Struct Algorithms 6:161–179
39. Molloy M, Reed B (1998) The size of the giant component of a random graph with a given degree sequence. Combin Probab Comput 7:295–305
40. Ramsey FP (1930) On a problem of formal logic. Proc London Math Soc 30:264–286
41. Simon HA (1955) On a class of skew distribution functions. Biometrika 42:425–440
42. Spencer J (1975) Ramsey's theorem—a new lower bound. J Comb Theory (A) 18:108–115
43. Stubbs L (2002) Genome comparison techniques. In: Galas D, McCormack S (eds) Genomic technologies: present and future. Caister Academic Press
44. Watts DJ, Strogatz SH (1998) Collective dynamics of 'small world' networks. Nature 393:440–442

Random Matrix Theory

Güler Ergün
Department of Mathematical Sciences,
University of Bath, Bath, UK

Article Outline

Glossary

Random matrices Large matrices with randomly distributed elements obeying the given probability laws and symmetry classes.

Orthogonal ensembles Real symmetric random matrix ensembles which are invariant under all orthogonal transformations. Majority of practical systems are described by these ensembles.

Unitary ensembles Hermitian random matrix ensembles which are invariant under all unitary transformations.

Symplectic ensembles Hermitian self-dual random matrix ensembles which are invariant under all symplectic transformations.

Definition of the Subject

Random Matrix Theory (RMT) is a method of studying the statistical behavior of large complex systems, by defining an ensemble which considers all possible laws of interactions within the system. The important question addressed by random matrix theory is: Given a random matrix ensemble what are the probability laws which govern its eigenvalues or eigenvectors. This question is pertinent

to many areas in physics and mathematics, for instance statistical behavior of compound nucleus, conductivity in disordered metals, behavior of chaotic systems or zeros of the Riemann zeta function. The success of random matrices lies in the universality regime of the eigenvalue statistics. There is compelling evidence that when the size of the matrix is very large then the eigenvalue distribution tends, in a certain sense, towards a limiting distribution. This only depends on the symmetry properties of the matrix and is independent of the initial probability law imposed on the matrix entries. For example, in many-body systems, the random matrix models are very successful in describing the spectral fluctuation properties of complex atoms, molecules and atomic nuclei, despite the fact that the interactions between the constitutes of these systems are very different, and not at all random: The nuclei are bound by short-range nuclear forces whereas atoms and molecules are governed by the long-range Coulomb forces, but the Hamiltonian of these models only considers the symmetry properties without the detailed knowledge of the system under consideration. The reason for such universality is not so clear but may be an outcome of a law of large numbers operating in the background. In the proceeding sections after giving the history and development of the theory, we will explore the random matrix ensembles, eigenvalue statistics as well as their pertinence to various complex systems.

Introduction

Although, the Random Matrix Theory was initiated in the early 1900s by statisticians in their studies of the product moment distribution [59], it made its progress in the hands of physicists. It was Eugene Wigner who introduced the random matrix idea to the theoretical physics community in the mid 1950s. During the 1950s, due to the development of nuclear weapons and nuclear power stations, physicists of this era were involved in measuring fission resonances in various nuclei. Wigner was predominantly studying [53] the statistical properties of nuclear spectra of heavy nuclei, but the large number of resonances was making his analysis very difficult. It was clear that the resonance spectra were far too complex to be analyzed by existing models like the atomic shell model. This prompted Wigner to look for statistical methods which would deal with such complexity. Later, the realization that the statistical distribution of the nuclear resonance energies shared the same properties as the eigenvalues of random matrices, led him to study a random matrix model [54], in which an infinite real symmetric matrix with identically distributed elements having the values zero in the diagonal and + 1 or − 1 in the off diagonal was used, so that all sign interactions were equally probable. In this model he showed that the distribution of the characteristic values of the matrix i. e., the density of states (DOS) was a semicircle, and it was also obvious that the characteristic values were repelling each other.[1]

During a number of conferences [55,56,57] in 1956–1957, on the basis of eigenvalues of a 2×2 real symmetric matrix, Wigner gave an assumed formula for the level spacing distribution, and he advocated Wishart's work [59], which was giving the joint probability distribution of eigenvalues, to confirm his findings. At about the same time analysis of experimental data by Gurevich and Pevsner [29] was also confirming that the nuclear levels were repelling each other and the distribution of consecutive level spacings was not totally random, but showed some regularities.

The interest in random matrix hypothesis was becoming wide spread, in [44] Porter and Rosenzweig summarized the developments on the statistical properties of atomic and nuclear spectra, by including experimental, theoretical, and numerical findings, which helped to strengthen the applicability of such statistical laws to nuclear reactions. It was time to make these statistical theories more mathematically rigorous; Wigner used the method of moments to prove that the density of eigenvalues for large symmetric matrices distributed as a semi circle [58]. Motivated by this, in 1960 Mehta and Gaudin obtained the same distribution by integrating joint probability density function of eigenvalues, and simultaneously Mehta analytically showed that the Wigner surmise was a very good approximation for the large random matrices. Soon after that, Gaudin gave the exact expression for the spacing distribution [26].

Dyson considered these statistical theories of random matrices as a new kind of Statistical Mechanics [10] such that one ignores the exact knowledge of the nature of the system. That is to say, a compound nucleus can be viewed as a *black box*, and one would need to define an ensemble of systems in which all possible laws of interactions are equally probable. In an attempt to explore this so-called "new kind of statistical mechanics" Dyson presented a series of papers [10,11,12], and in these papers he showed that all random matrix ensembles fall into three universality classes called the *orthogonal ensemble* (OE), *unitary ensemble* (UE) and *symplectic ensemble* (SE). These ensembles account for the symmetry properties of the Hamiltonians as well as the mathematical rigor of all possible interactions being equally likely. For matrices with Gaus-

[1] This was already known to Wigner and von Neuman [51].

sian distributed entries they are referred to as GOE, GUE and GSE, respectively. The random matrix theory (RMT) was developed all through the 1960s by Wigner, Mehta, Gaudin and Dyson. Most of the early work up to 1965 on random matrices is collected in *Statistical Theories of spectra* by Porter.

The RMT took a new dimension in the 1980s when Bohigas, Giannoni and Schmit (BGS) [5] conjectured that it could be applied to describing the spectra of all chaotic systems. This BGS statement not only renewed the interest in random matrices but also started a surge of research in different subject areas such as quantum chaos, quantum graphs, mesoscopic systems etc., and now there exists an overwhelming evidence that the BGS proposition may be true. There are several review articles which summarize the RMT applied research up to 1980s by Brody and up to more recent times by Bohigas, Guhr. A comprehensive book in this subject area is *Random Matrices* by Mehta [38], however it is aimed at the more advanced reader. As for the new comers, the pedagogical derivation of most of the basic principles of RMT may be found in *Quantum Signatures of Chaos* by Haake [30].

Nowadays, RMT is at the center of several areas of mathematics, particularly number theory, combinatorics, diffusion processes, probability and statistics. At the same time it is used as a tool to understand many complex systems from biology, quantum chaos, wireless communication and finance.

In the following sections, first the most commonly used random matrix ensembles, namely the Gaussian ensembles, will be presented. Then some of their central statistical properties such as the density of states and spacing distribution will be discussed.

Random Matrix Ensemble Classifications

Because of the broad use of RMT a variety of random matrix ensembles have been studied. These ensembles can be loosely divided into three types, Gaussian, circular and hyperbolic. In the next sections, although the emphasis will be on the Gaussian ensembles, we also give very brief descriptions of the circular and the hyperbolic ensembles.

Gaussian Ensembles

These ensembles are categorized into three universality classes, which only take into account: (1) the time-reversal invariance (more often referred to as the time-reversal symmetry) and (2) the strong spin-orbit coupling of the Hamiltonian of an underlying quantum system.

In Gaussian ensembles, the type of transformation gives the classifying name to the ensemble. For quick reference these ensembles are grouped in the Table 1 with consideration of the space-time symmetries and the values of the angular momentum.

Circular Ensembles

Introduced by Dyson [10] with an assumption that the system is not characterized by its Hamiltonian but by a unitary matrix S. The matrix is considered to be a scattering matrix (S-Matrix), whose elements give the transition probabilities between various states and its eigenvalues belong to the unit circle in the complex plane. Because of this compactness of the ensembles the universal distribution can be obtained with a uniform probability law. From the point of view of the fundamental symmetries, these ensembles can be categorized into three types as: Circular orthogonal ensembles (COE), circular unitary ensembles (CUE), circular symplectic ensembles (CSE). In the large N limit, the statistical properties of their eigenvalues are identical to those of the Gaussian ensembles.

Hyperbolic Ensembles

These ensembles are appropriately described in terms of Brownian motion in some matrix spaces. A seminal study of such a system is Dyson's Brownian motion model [9], which applies the notion of the RMT to a new type of Coulomb gas model so that the "Coulomb gas" could be interpreted as a dynamical system.

The matrix defined is a transfer matrix (M-matrix), which is a matrix used to give the relationship between the fluxes at one edge of a conductor to fluxes at the other edge (i. e., the number of ingoing electrons minus the outgoing electrons should be the same on each side of a conductor). Unlike the S-matrix, the M-matrix does not belong to a compact group. Also, it is not possible to define a normalizable probability law, but instead a diffusion equation as a function of a fictitious time. The term "time" can be any property of the model such as the length of a conductor.

Construction of Gaussian Ensembles

In fixing the elements of random matrices we are not completely free, since, first of all the matrix elements must obey the class restrictions given in the previous section. In the GOE case the Hamiltonian matrix is considered to be real symmetric. This means that for an $N \times N$ real symmetric matrix, there are $N(N+1)/2$ elements i. e., H_{ij} with $i, j = 1, \ldots, N$ can be chosen independently and the rest of them are determined by the symmetry. The ultimate aim is to define a joint probability density $P(\mathcal{H})$ for the matrix elements satisfying the conditions:

Random Matrix Theory, Table 1
The grouping of the quantum systems by the symmetry properties and the value of angular momentum. β is an index of the classification

Gaussian Ensemble Classification					
Symmetries	Total angular mom.	Hamiltonian	Canonical group	Ensemble	β
Time-reversal	Integer	Real symmetric	Orthogonal	GOE	1
Time-reversal, Rotational	Half-Integer				
Time-reversal, No-rotational	Half-Integer	Quaternion real	Symplectic	GSE	4
None	Any	Hermitian	Unitary	GUE	2

1. The probability $P(\mathcal{H})$ must be invariant under any transformation i. e. $\mathcal{H} \to \mathcal{W}^{-1}\mathcal{H}'\mathcal{W}$ with $\mathcal{W}$ is either orthogonal ($\beta = 1$), unitary ($\beta = 2$), or symplectic ($\beta = 4$) matrix. Then, e. g. for $\beta = 1$, when $\mathcal{W} = \mathcal{O}$ is any real orthogonal matrix then we have

$$P(\mathcal{H})\,\mathrm{d}\mathcal{H} = P(\mathcal{H}')\,\mathrm{d}\mathcal{H}'\,, \qquad \mathcal{O}^{-1} = \mathcal{O}^{\mathrm{T}}, \quad \text{and} \quad \mathcal{O}\mathcal{O}^{\mathrm{T}} = 1\,. \tag{1}$$

2. The matrix elements H_{ij} are statistically independent. Thus, the joint probability density function must be the product of the densities of these elements,

$$P(\mathcal{H}) = \prod_{ij} P_{ij}(H_{ij})\,. \tag{2}$$

The first condition is absolutely essential, $P(\mathcal{H})$ must be invariant under any transformation, because these transforms determine the eigenvalues, but the second one is only chosen to simplify the calculations. However, as will be more apparent later in the derivations, the consideration of these conditions necessitates the distribution of the random matrix elements to be Gaussian.

To find $P(\mathcal{H})$ for the real symmetric Hamiltonian matrices, we will consider a 2×2 real symmetric matrix $\mathcal{H}$ and an orthogonal transformation matrix

$$\mathcal{O} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \tag{3}$$

which is the two-dimensional rotation through θ. The rotation $\mathcal{O}$ acts as a transformation $\mathcal{H} = \mathcal{O}^{\mathrm{T}}\mathcal{H}'\mathcal{O}$, from which we get the relations

$$\begin{aligned} H_{11} &= \frac{H'_{11}+H'_{22}}{2} + \frac{H'_{11}-H'_{22}}{2}\cos 2\theta - H'_{12}\sin 2\theta \\ H_{12} &= \frac{H'_{11}-H'_{22}}{2}\sin 2\theta + H'_{12}\cos 2\theta \\ H_{22} &= \frac{H'_{11}+H'_{22}}{2} - \frac{H'_{11}-H'_{22}}{2}\cos 2\theta + H'_{12}\sin 2\theta\,. \end{aligned} \tag{4}$$

The joint probability density function is simply the product of the three densities

$$P(\mathcal{H}) = P_{11}(H_{11})P_{12}(H_{12})P_{22}(H_{22})\,, \tag{5}$$

and the orthogonal invariance of $P(\mathcal{H})$ implies that $\mathrm{d}P(\mathcal{H})\,\mathrm{d}\theta = 0$. Therefore we have

$$\frac{\mathrm{d}P_{11}(H_{11})}{P_{11}(H_{11})}\frac{\mathrm{d}H_{11}}{\mathrm{d}\theta}P(\mathcal{H}) + \frac{\mathrm{d}P_{12}(H_{12})}{P_{12}(H_{12})}\frac{\mathrm{d}H_{12}}{\mathrm{d}\theta}P(\mathcal{H}) + \frac{\mathrm{d}P_{22}(H_{22})}{P_{22}(H_{22})}\frac{\mathrm{d}H_{22}}{\mathrm{d}\theta}P(\mathcal{H}) = 0\,. \tag{6}$$

From the Eqs. (4) we can find equivalent expressions for the three differentials of $\mathrm{d}\mathcal{H}/\mathrm{d}\theta$ as:

$$\frac{\mathrm{d}H_{11}}{\mathrm{d}\theta} = -2H_{12}, \quad \frac{\mathrm{d}H_{12}}{\mathrm{d}\theta} = H_{11}-H_{22}, \quad \frac{\mathrm{d}H_{22}}{\mathrm{d}\theta} = 2H_{12}. \tag{7}$$

Substituting these later relations into Eq. (6), and dividing by $-H_{12}(H_{11} - H_{22})$ we get

$$\frac{\mathrm{d}\ln P_{11}(H_{11})}{\mathrm{d}H_{11}}\frac{2}{H_{11}-H_{22}} - \frac{\mathrm{d}\ln P_{22}(H_{22})}{\mathrm{d}H_{22}}\cdot\frac{2}{H_{11}-H_{22}} = \frac{\mathrm{d}\ln P_{12}(H_{12})}{\mathrm{d}H_{12}}\frac{1}{H_{12}} = -2a\,. \tag{8}$$

The two sides of this equation depend on different variables, which means that each side must be equal to a constant, hence the constant a is introduced. From Eq. (8) we first find a solution for $P_{12}(H_{12})$ by simply integrating

$$\begin{aligned} \mathrm{d}\ln P_{12}(H_{12}) &= -2aH_{12}\mathrm{d}H_{12} \\ P_{12}(H_{12}) &= C_{12}\exp\left(-aH_{12}^2\right)\,, \end{aligned} \tag{9}$$

in which the constant coefficient C_{12} is a consequence of the integration. Then re-writing Eq. (8) as

$$\frac{\mathrm{d}\ln P_{11}(H_{11})}{\mathrm{d}H_{11}} + aH_{11} = \frac{\mathrm{d}\ln P_{22}(H_{22})}{\mathrm{d}H_{22}} + aH_{22} = b \quad (10)$$

where another constant b is introduced as the equation depends on a different variable on each side, and following the same sequence of steps as before, we get solutions of P_{11} and P_{22},

$$\begin{aligned} P_{11}(H_{11}) &= C_{11}\exp\left[(-a/2)H_{11}^2 + bH_{11}\right] \\ P_{22}(H_{22}) &= C_{22}\exp\left[(-a/2)H_{22}^2 + bH_{22}\right] . \end{aligned} \quad (11)$$

These solutions have Gaussian distribution form, and the product of them gives

$$P(\mathcal{H}) = C\exp\left[(-a/2)(H_{11}^2 + H_{22}^2 + 2H_{12}^2) + b(H_{11} + H_{22})\right] . \quad (12)$$

If we fix the mean values of the diagonal elements to be zero i. e., $\langle H_{\mathrm{ii}}\rangle = 0$ then the second constant $b = 0$. The overall constant C is determined by the normalization

$$\int_{-\infty}^{+\infty} \mathrm{d}H_{11}\mathrm{d}H_{22}\mathrm{d}H_{12}P(\mathcal{H}) = 1 . \quad (13)$$

A simple calculation gives $C = 1/2(a/\pi)^{3/2}$. For the convergence of Eq. (13) the constant a must be positive, also it is related to the widths of the distributions of the diagonal and off-diagonal elements. The variances of the diagonal matrix elements

$$\begin{aligned} \langle H_{22}^2\rangle &= \langle H_{11}^2\rangle \\ &= \sqrt{a/2\pi}\int H_{11}^2\exp\left[(-a/2)H_{11}^2\right]\mathrm{d}H_{11} \\ &= 1/a , \end{aligned} \quad (14)$$

whereas for the off-diagonal elements

$$\langle H_{12}^2\rangle = \sqrt{a/\pi}\int H_{12}^2\exp(-aH_{12}^2)\mathrm{d}H_{12} = 1/2a . \quad (15)$$

By taking σ^2 as the variance of the off diagonal matrix elements, we obtain $a = 1/2\sigma^2$ for a 2×2 matrix.

We make a remark here that for large dimensional matrices the off-diagonal elements greatly outnumber the diagonal ones so, for the purpose of numerical simulations it does not matter much if the variances of the diagonal and the off-diagonal elements are not the same.

Although we have only considered the 2×2 Hamiltonian matrix in finding $P(\mathcal{H})$, this can easily be extended to $N\times N$ dimension with a suitable transformation matrix. In this way we find that a real symmetric random matrix having a probability density

$$P(\mathcal{H}) = \left(\frac{1}{4\sigma^2\pi}\right)^{N/2}\left(\frac{1}{2\sigma^2\pi}\right)^{N(N-1)/4} \cdot\exp\left\{-\frac{1}{4\sigma^2}\sum_{i,j}H_{ij}^2\right\} , \quad (16)$$

defines the *Gaussian orthogonal ensemble*.

Analogously, we obtain the *Gaussian unitary ensemble* (GUE) and the *Gaussian symplectic ensemble* (GSE), with the consideration that the joint probability distribution of the matrix elements is invariant under unitary and symplectic transformations, respectively. Thus, the general form of the probability density of a random Hermitian matrix is given by:

$$P(\mathcal{H}) = \left(\frac{1}{4\sigma^2\pi}\right)^{N/2}\left(\frac{1}{2\sigma^2\pi}\right)^{N(N-1)/2} \cdot\exp\left\{-\frac{1}{4\sigma^2}\sum_{i,j}\left[(H_R)_{ij}^2 + (H_I)_{ij}^2\right]\right\} , \quad (17)$$

where $(H_{\mathrm{R}})_{ij}^2$ and $(H_{\mathrm{I}})_{ij}^2$ are the real and the imaginary parts of the off-diagonal matrix elements H_{ij}, respectively. And the probability density of a random quaternion matrix is:

$$P(\mathcal{H}) = \left(\frac{1}{4\sigma^2\pi}\right)^{N/2}\left(\frac{1}{2\sigma^2\pi}\right)^{N(N-1)} \cdot\exp\left\{-\frac{1}{4\sigma^2}\sum_{i,j}\left[(H_0)_{ij}^2 + (H_1)_{ij}^2 + (H_2)_{ij}^2 + (H_3)_{ij}^2)\right]\right\} , \quad (18)$$

where $(H_0)_{ij}$, $(H_1)_{ij}$, $(H_2)_{ij}$ and $(H_3)_{ij}$ are the quaternionic components of $\mathbf{H}_{ij}$.

Without loss of generality, the probability density $P(\mathcal{H})$ with $\mathcal{H}$ of an arbitrary $N\times N$ dimension can be written in a form common to all three ensembles as

$$P_{N\beta}(\mathcal{H}) = C_{N\beta}\mathrm{e}^{-\frac{N\beta}{4\sigma^2}\mathrm{Tr}[\mathcal{H}^2]} , \quad (19)$$

where Tr stands for the *matrix trace*: $\mathrm{Tr}\,\mathcal{H} = \sum_i H_{\mathrm{ii}}$, $\beta = 1, 2, 4$ is used as an index for the ensemble classification, with σ^2 being the variance of the off diagonal matrix elements and the factor N ensuring the moments of the eigenvalues λ_i of $\mathcal{H}$ to be finite in the limit $N\to\infty$, while $\langle\lambda_i^2\rangle = 1$.

Eigenvalues of Gaussian Ensembles

Using the joint probability distribution of the matrix entries in Eq. (19) we will write the joint probability distribution of the eigenvalues for all three Gaussian ensembles. Again, we consider the simplest case of a real symmetric 2×2 Hamiltonian matrix H, and the orthogonal transformation matrix $\mathcal{O}$, as given in Eq. (3), such that

$$\mathcal{H} = \mathcal{O}^{\mathrm{T}} \mathcal{H}_{\mathrm{D}} \mathcal{O} \,, \tag{20}$$

where $\mathcal{H}_{\mathrm{D}}$ is the diagonal matrix containing the eigenvalues of $\mathcal{H}$. In this case there are only two eigenvalues, which are given by

$$\lambda_{1,2} = \frac{1}{2}(H_{11} + H_{22}) \pm \sqrt{(H_{11} - H_{22})^2 + 4H_{12}^2} \,. \tag{21}$$

Writing out Eq. (20) explicitly

$$\begin{aligned} H_{11} &= \lambda_1 \cos^2\theta + \lambda_2 \sin^2\theta \\ H_{22} &= \lambda_1 \sin^2\theta + \lambda_2 \cos^2\theta \\ H_{12} &= (\lambda_1 - \lambda_2)\cos\theta \sin\theta \,, \end{aligned} \tag{22}$$

we can see that the elements of $\mathcal{H}$ are linear functions of the eigenvalues and a single angle θ parameter which specifies the set of eigenvectors. The Jacobian of this transformation is simply

$$J = \det\left[\frac{\partial(H_{11}, H_{22}, H_{12})}{\partial(\lambda_1, \lambda_2, \theta)}\right] = (\lambda_1 - \lambda_2) \,, \tag{23}$$

and since $\mathrm{Tr}[\mathcal{H}^2] = (\lambda_1^2 + \lambda_2^2)$, we can just replace the matrix elements H_{ij} with the eigenvalues of $\mathcal{H}$ and integrate over θ to get

$$P(\lambda_1, \lambda_2) = C|\lambda_1 - \lambda_2| \mathrm{e}^{-\frac{N\beta}{4\sigma^2}(\lambda_1^2 + \lambda_2^2)} \,, \tag{24}$$

which is the joint probability density of the eigenvalues for the GOE of 2×2 dimension.

We follow the same procedure as above to find the joint probability density of the eigenvalues for the GUE case. Here we consider a 2×2 complex Hermitian matrix $\mathcal{H}$, and a simple 2×2 unitary transformation matrix $\mathcal{U}$ to diagonalize it, where

$$\mathcal{U} = \begin{pmatrix} \cos\theta & -\mathrm{e}^{-i\phi}\sin\theta \\ \mathrm{e}^{i\phi}\sin\theta & \cos\theta \end{pmatrix} \,. \tag{25}$$

However, we note that this unitary matrix is not the most general 2×2 unitary transformation matrix, but rather an element of the coset space $\mathcal{V}(2)/V(1)V(1)$. This is to ensure a one-to-one correspondence $\mathcal{H} \to (\lambda_1, \lambda_2, \mathcal{U})$ after the ordering of the eigenvalues, which are given by

$$\lambda_{1,2} = \frac{1}{2}(H_{11} + H_{22}) \pm \sqrt{(H_{11} - H_{22})^2 + 4H_{12}^* H_{12}} \,. \tag{26}$$

This expression is essentially the same as the one in Eq. (21) with the only difference being $H_{12}^2 \to |H_{12}|^2$. In complete analogy to Eq. (20) we write

$$\mathcal{H} = \mathcal{U}^{\dagger} \mathcal{H}_{\mathrm{D}} \mathcal{U} \tag{27}$$

to represent the matrix elements of $\mathcal{H}$ in terms of the two eigenvalues $\lambda_{1,2}$ and the two angles θ, ϕ

$$\begin{aligned} H_{11} &= \lambda_1 \cos^2\theta + \lambda_2 \sin^2\theta \\ H_{22} &= \lambda_1 \sin^2\theta + \lambda_2 \cos^2\theta \\ H_{12} &= (\lambda_1 - \lambda_2)[\cos\theta \sin\theta \cos\phi - i \cos\theta \sin\theta \sin\phi] \,, \end{aligned} \tag{28}$$

from which the Jacobian is

$$J = \det\left[\frac{\partial(H_{11}, H_{22}, H_{12}^R, H_{12}^I)}{\partial(\lambda_1, \lambda_2, \theta, \phi)}\right] = (\lambda_1 - \lambda_2)^2 \cos\theta \sin\theta \,. \tag{29}$$

Again, $\mathrm{Tr}[\mathcal{H}^2] = \lambda_1^2 + \lambda_2^2$ thus Eq. (19) can be written for GUE, in terms of the eigenvalues and the two parameters namely angles as

$$P(\lambda_1, \lambda_2, \theta, \phi) = C|\lambda_1 - \lambda_2|^2 \mathrm{e}^{-\frac{N\beta}{4\sigma^2}(\lambda_1^2 + \lambda_2^2)} \cos\theta \sin\theta \,, \tag{30}$$

integrating this over the angles θ and ϕ we get

$$P(\lambda_1, \lambda_2) = C|\lambda_1 - \lambda_2|^2 \mathrm{e}^{-\frac{N\beta}{4\sigma^2}(\lambda_1^2 + \lambda_2^2)} \,. \tag{31}$$

Although the constant C has been modified as a result of the integration, we still call it C to avoid unnecessary complication.

Now we move on to the GSE case, in which the smallest Hamiltonian matrix $\mathcal{H}$ is 4×4 but the procedure for finding the joint probability density of the eigenvalues is completely analogous to the preceding GOE and GUE cases.

The most convenient way to write $\mathcal{H}$ is the quaternion notation, reducing the $2N \times 2N$ matrix to an $N \times N$ by representing each H_{ij} with a block of 2×2, superposition of unity and the quaternion matrices. Then the degenerate eigenvalues of quaternion real $\mathcal{H}$ are

$$\lambda_{1,2} = \frac{1}{2}\left(H_{11}^0 - H_{22}^0\right) \pm \left[\left(H_{11}^0 - H_{22}^0\right)^2 + 4\sum_{\mu=0}^{3}\left(H_{12}^{\mu}\right)^2\right]^{1/2} \,, \tag{32}$$

each of these λs are double, so that we have 4 eigenvalues since the matrix is 4×4. The total number of freely

picked real entries in a $2N \times 2N$ quaternion real matrix $\mathcal{H}$ is given by $N(2N-1)$, thus we have 6 real entries in this case. To diagonalize $\mathcal{H}$, not the most general, but a convenient choice is a 2×2 symplectic matrix

$$S = \begin{pmatrix} e^{-\xi \cdot \tau} \cos\theta & \sin\theta \\ -\sin\theta & e^{\xi \cdot \tau} \cos\theta \end{pmatrix}, \tag{33}$$

where θ, ξ are the four real parameters provided by S. As before using the notation $\mathcal{H} = S^{\dagger} \mathcal{H}_{\mathrm{D}} S$, we write the elements of $\mathcal{H}$ as a function of the eigenvalues and the parameters θ, ξ,

$$\begin{aligned} H^0_{11} &= \lambda_1 \cos^2\theta + \lambda_2 \sin^2\theta \\ H^0_{22} &= \lambda_1 \sin^2\theta + \lambda_2 \cos^2\theta \\ H^0_{12} &= -(\lambda_1 - \lambda_2) \cos\theta \sin\theta \cos\xi \end{aligned} \tag{34}$$

$$\mathbf{H}^{\mu}_{12} = \xi(\lambda_1 - \lambda_2) \cos\theta \sin\theta \sin\xi \,. \tag{35}$$

The Jacobian of this transformation is

$$J = \det\left[\frac{\partial(H^0_{11}, H^0_{22}, H^{\mu}_{12})}{\partial(\lambda_1, \lambda_2, \theta, \xi)}\right] \propto (\lambda_1 - \lambda_2)^4, \tag{36}$$

where the coefficient of the proportionality is not specified since it is independent of the eigenvalues, and will only contribute to the constant C after integration over all four parameters θ and ξ. Thus, the joint probability distribution of the eigenvalues for the smallest GSE is

$$P(\lambda_1, \lambda_2) = C|\lambda_1 - \lambda_2|^4 e^{-\frac{N\beta}{4\sigma^2}(\lambda_1^2 + \lambda_2^2)} \,. \tag{37}$$

From the general form of the Eqs. (24,31,37), we can guess the generic joint probability distribution of the eigenvalues for all three Gaussian ensembles to be

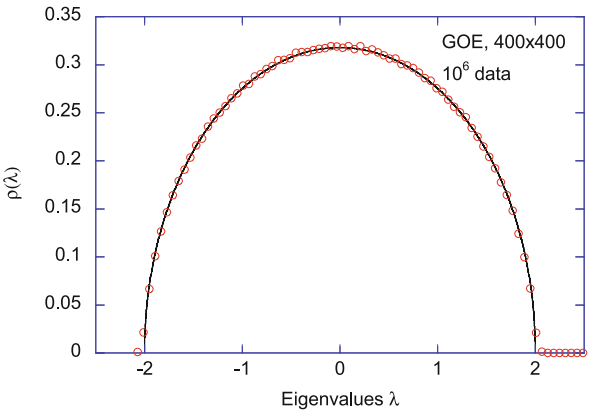
$$\begin{aligned} &P_{N\beta}(\lambda_1, \ldots, \lambda_N) \\ &\quad = C_{N\beta} \prod_{i<j} |\lambda_i - \lambda_j|^{\beta} \exp\left(-\frac{N\beta}{4\sigma^2} \sum_{k=1}^{N} \lambda_k^2\right), \end{aligned} \tag{38}$$

in which the constant $C_{N\beta}$ is an ensemble dependent constant, and it is chosen such that the $P_{N\beta}$ is normalized to unity.

Density of States: Semicircle Law

Having at our disposal the generalized joint probability density of eigenvalues, now we can derive the famous *semicircle law* for the mean eigenvalue densities (see Fig. 1). The essential step is integrating Eq. (38) in the previous section i. e.,

$$N \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P_{N\beta}(\lambda_1, \lambda_2, \ldots, \lambda_N) \, \mathrm{d}\lambda_2, \ldots, \mathrm{d}\lambda_N \tag{39}$$



Random Matrix Theory, Figure 1
Averaged density of states (DOS) for GOE (with a suitable scaling, all three ensembles have the same form of DOS distribution). The *solid line* is the theoretical result superimposed over numerical data values (*circles*)

over all variables accept one to determine the ensemble-averaged eigenvalue or level density $\rho(\lambda)$, and the semicircle law arises asymptotically as $N \to \infty$. This integration is not at all trivial, and was first carried out by Mehta and Gaudin [39], using some remarkable techniques. However, there are many ways to derive the semicircle law, but the one that is perhaps the most elegant and up-to-date is the representation of determinants by Gaussian integrals over Grassmaninan variables, we will use this technique in the following derivations of the semicircle law.

General Formalism

Consider an $N \times N$ Hermitian matrix $\mathcal{H}$, which has ordered eigenvalues $\lambda_k (k = 1, \ldots, N)$. The normalized spectrum of $\mathcal{H}$ can be expressed as sums of delta function, which essentially picks out the positions of the eigenvalues,

$$\rho(\lambda) = \frac{1}{N} \sum_{k=1}^{N} \delta(\lambda - \lambda_k) \,. \tag{40}$$

By representing the delta function as a Lorentzian curve with vanishing widths.

$$\delta(\lambda - \lambda_k) = \lim_{\varepsilon \to 0} \frac{\varepsilon}{\pi} \frac{1}{(\lambda - \lambda_k)^2 + \varepsilon^2} \,, \tag{41}$$

and substituting this latter relation to the former we get

$$\rho(\lambda) = \frac{1}{N\pi} \mathrm{Im}\,\mathrm{Tr}\left(\frac{1}{\lambda \mathcal{I} - \mathcal{H}}\right), \tag{42}$$

where $\mathcal{I}$ is identity matrix and λ is assumed to have a small imaginary part but we will not write it explicitly for the

sake of less cumbersome expressions. To proceed with the supersymmetric calculations we need to represent this last relation in terms of ratios of the determinants, for which we introduce the following generating function

$$\mathcal{Z}(\lambda,\mu) \equiv \frac{\det(\lambda\mathcal{I}-\mathcal{H})}{\det(\mu\mathcal{I}-\mathcal{H})}\,, \tag{43}$$

where $\mu \equiv \lambda \pm ix/2N$. Thus, differentiating the generating function with respect x, i. e.,

$$\frac{\partial}{\partial x}\mathcal{Z}(\lambda,\mu)|_{x\to 0} = \frac{1}{N}\mathrm{Im}\,\mathrm{Tr}\left(\frac{1}{\lambda\mathcal{I}-\mathcal{H}}\right)\,, \tag{44}$$

and when $x \to 0$ we recover the expression for the density of the eigenvalues we have defined previously. So that our starting point will be

$$\frac{\det(\lambda\mathcal{I}-\mathcal{H})}{\det(\mu\mathcal{I}-\mathcal{H})}\,. \tag{45}$$

We are interested in averaging the expression in Eq. (45) over all realization of $\mathcal{H}$. For which we will use the "supersymmetrization" method. In this method, the denominator of the expression is represented by a general Gaussian integral over a complex N dimensional vector $\mathbf{S} = (S_1, \ldots, S_N)^{\mathrm{T}}$, where T stands for the vector transpose, and S_i are commuting variables

$$(2\pi)\det{}^{-1}(\mu\mathcal{I}-\mathcal{H}) = \frac{1}{i^N}\int \mathrm{d}\mathbf{S}^\dagger\,\mathrm{d}\mathbf{S}\mathrm{e}^{\frac{i}{2}\mu\mathbf{S}^\dagger\mathbf{S}-\frac{i}{2}\mathbf{S}^\dagger\mathcal{H}\mathbf{S}}, \tag{46}$$

with the assumption that μ has a small positive imaginary part i. e. we choose $\mu = \lambda + ix/N$ so that the integral is convergent. For the determinant in the numerator, we use Gaussian integrals over anticommuting (Grassmannian) N components vectors χ, $\chi^\dagger$, which gives

$$(2\pi)^{-1}\det(\lambda\mathcal{I}-\mathcal{H}) = \frac{1}{i^N}\int \mathrm{d}\chi^\dagger\,\mathrm{d}\chi\mathrm{e}^{\frac{i}{2}\lambda\chi^\dagger\chi-\frac{i}{2}\chi^\dagger\mathcal{H}\chi}\,. \tag{47}$$

Thus, we write the expression in Eq. (45) as

$$\langle\cdots\rangle_{\mathcal{H}} = \int \mathrm{d}^2\chi \cdot \int \mathrm{d}^2\mathbf{S}\mathrm{e}^{\frac{i}{2}(\lambda\chi^\dagger\chi+\mu\mathbf{S}^\dagger\mathbf{S})}\left\langle \mathrm{e}^{-\frac{i}{2}(\mathbf{S}^\dagger\mathcal{H}\mathbf{S}+\chi^\dagger\mathcal{H}\chi)}\right\rangle_{\mathcal{H}}\,. \tag{48}$$

The commuting and the anti commuting variables appear symmetrically in this expression, which in fact gives the name "supersymmetry" to the method. To perform the average over the GUE matrix $\mathcal{H}$ we can decompose elements of $\mathcal{H}$ into its real and imaginary parts, $H_{kl} = H^R_{kl} + iH^I_{kl}$ and write

$$\left\langle \mathrm{e}^{-\frac{i}{2}\sum_{kl}H_{kl}A_{kl}}\right\rangle_{\mathcal{H}} = \left\langle \mathrm{e}^{\sum_k -\frac{i}{2}A_{kk}H^R_{kk}}\right\rangle \cdot \left\langle \mathrm{e}^{\sum_{k>l}-\frac{i}{2}(A_{kl}+A_{lk})H^R_{kl}}\right\rangle \cdot \left\langle \mathrm{e}^{\sum_{k>l}\frac{1}{2}(A_{kl}-A_{lk})H^I_{kl}}\right\rangle\,, \tag{49}$$

where $A_{kl} = S^*_k S_l + \chi^*_k\chi_l$. The explicit meaning of angle brackets is the average over the joint probability density for GUE, here we take

$$P(\mathcal{H}) = C_N\mathrm{e}^{-\frac{N}{2}\mathrm{Tr}[\mathcal{H}^2]}\,\mathrm{d}\mathcal{H}\,, \tag{50}$$

where C_N is a constant normalization factor and the probability distribution has zero mean and unity variance. By means of this probability distribution we get

$$\begin{aligned}\left\langle \mathrm{e}^{-\frac{i}{2}A_{kk}H^R_{kk}}\right\rangle &= \exp\left[-\frac{A^2_{kk}}{8N}\right]\,,\\ \left\langle \mathrm{e}^{-\frac{i}{2}(A_{kl}+A_{lk})H^R_{kl}}\right\rangle &= \exp\left[-\frac{(A_{kl}+A_{lk})^2}{16N}\right]\,,\\ \left\langle \mathrm{e}^{\frac{1}{2}(A_{kl}-A_{lk})H^I_{kl}}\right\rangle &= \exp\left[\frac{(A_{kl}-A_{lk})^2}{16N}\right]\,.\end{aligned} \tag{51}$$

Substitution of these equations into (49) yields

$$\left\langle \mathrm{e}^{-\frac{i}{2}\sum_{kl}H_{kl}A_{kl}}\right\rangle_{\mathcal{H}} \propto \mathrm{e}^{-\frac{1}{8N}\sum_{kl}A_{kl}A_{lk}}\,, \tag{52}$$

it is easy to show that $\sum_{kl}H_{kl}A_{lk} = \mathrm{Tr}[\mathcal{H}\mathcal{A}]$ and $\sum_{kl}A_{kl}A_{lk} = \mathrm{Tr}[\mathcal{A}^2]$ then the expression in Eq. (52) is just the identity

$$\left\langle \mathrm{e}^{-\frac{i}{2}\mathrm{Tr}[\mathcal{H}\mathcal{A}]}\right\rangle_{\mathrm{GUE}} \propto \mathrm{e}^{-\frac{1}{8N}\mathrm{Tr}[\mathcal{A}^2]}\,. \tag{53}$$

The far right exponent in Eq. (48) have the relations $\chi^\dagger\mathcal{H}\chi = -\mathrm{Tr}[\mathcal{H}\chi\otimes\chi^\dagger]$ and $\mathbf{S}^\dagger\mathcal{H}\mathbf{S} = \mathrm{Tr}[\mathcal{H}\mathbf{S}\otimes\mathbf{S}^\dagger]$. These relations imply that the right hand side exponent of the identity is:

$$\mathrm{Tr}[\mathcal{A}^2] = \mathrm{Tr}[\mathbf{S}\otimes\mathbf{S}^\dagger - \chi\otimes\chi^\dagger]^2. \tag{54}$$

Expanding the latter we get

$$\mathrm{Tr}[\mathcal{A}^2] = (\mathbf{S}^\dagger\mathbf{S})^2 - 2(\mathbf{S}^\dagger\chi)(\chi^\dagger\mathbf{S}) - (\chi^\dagger\chi)^2. \tag{55}$$

The quartic term in the far right of Eq. (55) needs to be "decoupled" before we can carry on with the integration, we can do so by using the usual Hubbard–Stratonovich transformation

$$\mathrm{e}^{\frac{1}{8N}(\chi^\dagger\chi)^2} = \int_{-\infty}^{\infty}\frac{\mathrm{d}q}{\sqrt{2\pi}}\mathrm{e}^{-\frac{q^2}{2}-\frac{q}{\sqrt{2N}}\chi^\dagger\chi}. \tag{56}$$

Thus, re-writing Eq. (48) with these new relations we have

$$\langle\cdots\rangle_{\mathcal{H}} = \int \mathrm{d}^2\mathbf{S}\exp\left\{-\frac{1}{8N}(\mathbf{S}^\dagger\mathbf{S})^2 + \frac{i}{2}\mu\mathbf{S}^\dagger\mathbf{S}\right\} \cdot \int_{-\infty}^{\infty}\frac{\mathrm{d}q}{\sqrt{2\pi}}\mathrm{e}^{-q^2/2}\int \mathrm{d}\chi^\dagger\,\mathrm{d}\chi \cdot \exp\left\{\chi^\dagger\left[1/2(i\lambda - q/N)\mathcal{I} - \mathbf{S}\otimes\mathbf{S}^\dagger/4N\right]\chi\right\}. \quad (57)$$

The last integral over the Grassmannian variables is the familiar Gaussian integral in Eq. (47), which leads to

$$\langle\cdots\rangle_{\mathcal{H}} = \int \mathrm{d}^2\mathbf{S}\mathrm{e}^{-\frac{1}{8N}(\mathbf{S}^\dagger\mathbf{S})^2+\frac{i}{2}\mu\mathbf{S}^\dagger\mathbf{S}} \cdot \int_{-\infty}^{\infty}\frac{\mathrm{d}q}{\sqrt{2\pi}}\mathrm{e}^{-\frac{q^2}{2}}\det\left[\frac{1}{2}\left(i\lambda - \frac{q}{N}\right)\mathcal{I} - \frac{\mathbf{S}\otimes\mathbf{S}^\dagger}{4N}\right]. \quad (58)$$

Further manipulation can be made by introducing the variable $q_F = i\lambda - q/\sqrt{N}$ and shifting the contour of integration in such a way that, the integral over q_F goes along the real axis, thus, we rewrite the above expression as

$$\langle\cdots\rangle_{\mathcal{H}} \propto \int_{-\infty}^{\infty}\frac{\mathrm{d}q_F}{\sqrt{2\pi}}\mathrm{e}^{-\frac{N}{2}(q_F-i\lambda)^2} \cdot \int \mathrm{d}^2\mathbf{S}\mathrm{e}^{\frac{1}{8N}(\mathbf{S}^\dagger\mathbf{S})^2+\frac{i}{2}\mu\mathbf{S}^\dagger\mathbf{S}}\det\left[q_F\mathcal{I} - \frac{1}{2N}\mathbf{S}\otimes\mathbf{S}^\dagger\right], \quad (59)$$

where we shifted the contour for $q_F \in (-\infty, \infty)$ to be real.

Further simplification can be made by noticing that the $N\times N$ matrix $\mathbf{S}\otimes\mathbf{S}^\dagger$ is of rank unity, i. e., it has $(N-1)$ zero eigenvalues, and only one nonzero eigenvalue equal to $(\mathbf{S}\mathbf{S}^\dagger)$. Then the determinant in the previous expression is

$$\det\left[q_F\mathcal{I} - \frac{1}{2N}\mathbf{S}\otimes\mathbf{S}^\dagger\right] \equiv q_F^{N-1}\left(q_F - \frac{1}{2N}\mathbf{S}^\dagger\mathbf{S}\right). \quad (60)$$

Our next step is to introduce polar coordinates: $\mathbf{S} = r\mathbf{n}$ with $\mathbf{n}^\dagger\mathbf{n} = 1$ and $\int \mathrm{d}^2\mathbf{S} = r^{2N-1}\mathrm{d}r\,\mathrm{d}\mathbf{n}$, where $\int \mathrm{d}\mathbf{n} = \Omega_N$ corresponds to the area of a $2N$ dimensional unit sphere, which only produces a constant factor. Further introducing $p = r^2$ and changing $p \to 2Np$ and then following with the obvious manipulations we get:

$$\left\langle\frac{\det(\lambda\mathcal{I}-\mathcal{H})}{\det(\mu\mathcal{I}-\mathcal{H})}\right\rangle_{\mathcal{H}} = C_N\mathrm{e}^{\frac{N}{2}\lambda^2}\int_{-\infty}^{\infty}\frac{\mathrm{d}q_F}{\sqrt{2\pi}\,q_F} \cdot \exp\left\{-\frac{N}{2}\left(q_F^2 - 2i\lambda q_F - 2\ln q_F\right)\right\} \cdot \int_0^{\infty}\frac{\mathrm{d}p(q_F-p)}{\sqrt{2\pi}\,p}\exp\left\{-\frac{N}{2}\left(p^2 - 2i\mu p - 2\ln p\right)\right\}. \quad (61)$$

In this expression C_N stands for the accumulated constant factors such that when $\mu = \lambda$ the right hand side must yield unity identically.

Up to now all the expressions were valid for finite-size matrices, but for $N\to\infty$ with appropriate scaling we expect the results to be universal. This means that in the asymptotic limit the results will be broadly insensitive to the details of the distribution of random matrices, thus it can be applicable to complex or quantum chaotic systems. More over in such a limit the integrals in Eq. (61) can be evaluated by the saddle-point method, or Laplace's method. To use the latter method we write the terms in the exponent in Eq. (61) as

$$\mathcal{L}(q_F) = [q_F^2 - 2i\lambda q_F - 2\ln q_F] \quad \text{and} \quad \mathcal{L}(p) = [p^2 - 2i\mu p - 2\ln p]. \quad (62)$$

The first derivatives of these equations yield the saddle-points: $q_F^{s.p} = (i\lambda \pm \sqrt{4-\lambda^2})/2$ and $p^{s.p} = (i\mu + \sqrt{4-\mu^2})/2$ since $Re(p) \geq 0$, for $\lambda \leq \sqrt{4}$ the so called *bulk* of the spectrum. The imaginary part of μ here denoted as ix/N does not contribute, when $N\to\infty$ also for consistency we should put $\mu = \lambda$. Because of the presence of the factor $(q_F - p)$ in the integrand, it is easy to see that only the choice $q_F^{s.p} = (i\lambda - \sqrt{4-\lambda^2})/2$ yields the leading-order contribution. Substituting this choice into the integrand in Eq. (61) and evaluating the Gaussian fluctuations around the saddle-point values, finally yields

$$\left\langle\frac{\det(\lambda\mathcal{I}-\mathcal{H})}{\det[(\lambda+ix/N)\mathcal{I}-\mathcal{H}]}\right\rangle_{\mathcal{H}}\Bigg|_{\substack{x>0\\N\to\infty}} = \exp\left\{-\frac{x}{2}[i\lambda + \sqrt{4-\lambda^2}]\right\}, \quad (63)$$

where we re-substituted $\mu = \lambda + ix/N$. To recover averaged density of states $\langle\rho(\lambda)\rangle$ from the expression above we simply need to differentiate both sides with respect to x, and let $x\to 0$, for which we write

$$\frac{\partial}{\partial x}\left\langle\det(\lambda-\mathcal{H})\exp\{-\mathrm{Tr}\ln[(\lambda+ix/N)-\mathcal{H}]\}\right\rangle = \frac{\partial}{\partial x}\exp\left\{-\frac{x}{2}[i\lambda + \sqrt{4-\lambda^2}]\right\}. \quad (64)$$

Performing the differentiation and letting $x \to 0$ we get

$$-i\left\langle \frac{1}{N}\,\mathrm{Tr}\left(\frac{1}{\lambda\mathcal{I}-\mathcal{H}}\right)\right\rangle_{\mathcal{H}}\Bigg|_{\substack{x\to 0\\ N\to\infty}} = -\frac{i\lambda}{2}-\frac{1}{2}\sqrt{4-\lambda^2}\Bigg|_{x\to 0}\,. \quad (65)$$

Comparing this last relation with Eq. (42) we see that the averaged density of states is given by

$$\langle\rho(\lambda)\rangle = \frac{1}{2\pi}\sqrt{4-\lambda^2}\,. \quad (66)$$

It is easy to repeat all the calculations for $x < 0$ and find that for any real value of x the result in Eq. (63) can be written as

$$\left\langle \frac{\det(\lambda\mathcal{I}-\mathcal{H})}{\det[(\lambda-ix/N)\mathcal{I}-\mathcal{H}]}\right\rangle_{\mathcal{H}}\Bigg|_{\substack{x<0\\ N\to\infty}} = \exp\left\{-\frac{i\lambda x}{2}-|x|\pi\rho(\lambda)\right\}\,. \quad (67)$$

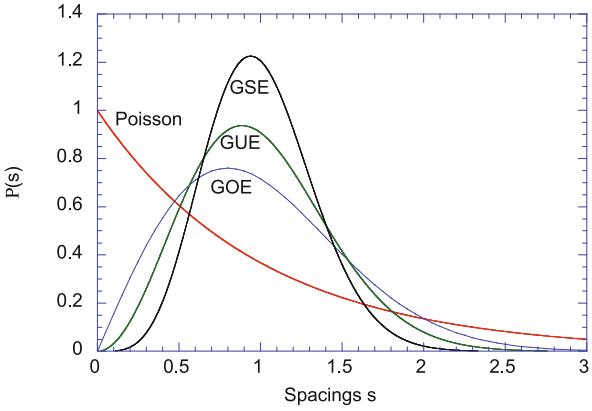
Spacing Distribution: Wigner Surmise

The *level repulsion* phenomenon that the small spacings between two neighboring levels are unlikely can be seen from the expression $|\lambda_i-\lambda_j|^\beta$, which appeared as a result of the Jacobian of the transformation in all three cases of the Gaussian ensembles. Wigner in [56] conjectured that the distribution of spacings between adjacent eigenvalues may be given by

$$P(s) \approx Cs^\beta \exp(-As^2)\,. \quad (68)$$

He simply guessed this probability law, and demonstrated his argument by using a 2×2 real symmetric matrix. Through the joint probability distribution of the eigenvalues, and utilizing many novel techniques Gaudin obtained an exact expression for the spacing distribution. However, the derivations are rather lengthy and beyond the scope of this introductory section, hence we refer the reader to [26]. Here, only the simplest case of $N = 2$ dimensional ensembles of random matrices will be considered by taking the distribution of eigenvalues for $N = 2$. The spacing distribution for a single pair of levels can be written as

$$P(s) = C\int_{-\infty}^{\infty} \mathrm{d}\lambda_1 \cdot \int_{-\infty}^{\infty} \mathrm{d}\lambda_2\,\delta(s-|\lambda_1-\lambda_2|)|\lambda_1-\lambda_2|^\beta \mathrm{e}^{-\frac{A}{2}(\lambda_1^2+\lambda_2^2)}\,. \quad (69)$$



Random Matrix Theory, Figure 2
***Wigner Surmise*; Nearest neighbor spacing distributions for Poissonian random process, the orthogonal, unitary and symplectic ensembles**

Changing the variables as $\lambda_1 = (\sqrt{2}/2)(u+v)$ and $\lambda_2 = (\sqrt{2}/2)(u-v)$, makes the integration easy to perform. The constants A and C are fixed by the following normalization conditions:

$$\int_0^\infty P(s)\mathrm{d}s = 1 \quad \text{and} \quad \langle s\rangle \equiv \int_0^\infty sP(s)\mathrm{d}s = 1\,, \quad (70)$$

the latter implies that the units of eigenvalues are such that the mean spacing is unity. Performing the elementary integrations for $\beta = 1, 2$ and 4 gives

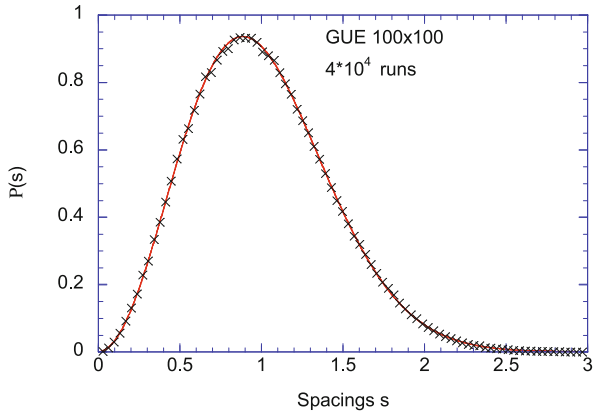
$$P(s) = \begin{cases} \frac{\pi}{2}s\mathrm{e}^{-s^2\pi/4} & \beta = 1 \quad \text{(GOE)} \\ \frac{32}{\pi^2}s^2\mathrm{e}^{-s^2 4/\pi} & \beta = 2 \quad \text{(GUE)} \\ \frac{2^{18}}{3^6\pi^3 2}s^4\mathrm{e}^{-s^2 64/9\pi} & \beta = 4 \quad \text{(GSE)} \end{cases} \quad (71)$$

which are referred to as the *Wigner Surmise* (see Fig. 2) in the literature, and these are not very different from the asymptotic i. e., $N \to \infty$ derivations (see Fig. 3).

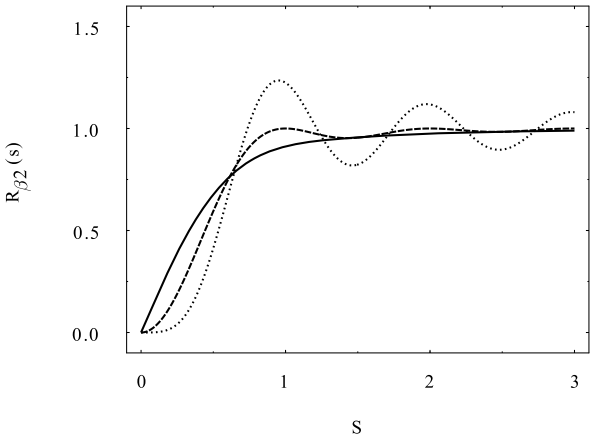
Two-Level Correlation Function

The two-level correlation function is a highly investigated topic in RMT, because it is very useful for spectral data analysis. A number of spectral correlations such as the *number variance* (for historical reasons also called Δ_3 statistics) are expressed in terms of it. It is defined as the probability density to find two eigenvalues (or levels) λ_1 and λ_2 at two given positions regardless of the positions of all the other eigenvalues. The more general form of this definition is given by the n-level correlation function

$$R_n(\lambda_1,\,\ldots,\lambda_n) = \frac{N!}{(N-n)!} \cdot \int P_N(\lambda_1,\,\ldots,\lambda_N)\mathrm{d}\lambda_{n+1}\,\ldots\,\mathrm{d}\lambda_N\,, \quad (72)$$



Random Matrix Theory, Figure 3
Comparison of nearest neighbor spacing distributions for 2 × 2 Wigner surmise (*solid line*) and numerical simulation (*crosses*) of 100 × 100 GUE



Random Matrix Theory, Figure 4
The two-level correlation functions $R_{\beta 2}(s)$ on the unfolded energy scale. The *solid line* is the GOE, the *dashed line* is the GUE and the *dotted line* is the GSE, which overshoots the value one due to strong oscillations

where the coefficient on the right hand side takes into account all combinations of selecting n eigenvalues out of the total number N. The integration is taken over the joint probability density P_N of eigenvalues. First systematic derivations of n-level correlation function were given by Dyson [12] and since then a number of other methods have been developed, for example the method of orthogonal polynomials and the supersymmetry approach. The detailed derivations of the n-level correlation and cluster functions for all three Gaussian ensembles can be found in Mehta's book, where the essential step of derivations involves the idea of expressing the joint probability density of eigenvalues as a determinant. However, the calculations are too lengthly, in particular the GOE case is rather complicated, and for our purpose it is sufficient to give only the following results of the two-level correlation functions $R_{\beta,2}(s)$ (see Fig. 4) for all Gaussian ensembles:

$$R_{1,2}(s) = 1 - \left[\left(\frac{\sin \pi s}{\pi s}\right)^2 + \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\sin \pi s}{\pi s}\right)\int_s^\infty \frac{\sin \pi t}{\pi t}\mathrm{d}t\right],$$

$$R_{2,2}(s) = 1 - \left(\frac{\sin \pi s}{\pi s}\right)^2,$$

$$R_{4,2}(s) = 1 - \left[\left(\frac{\sin 2\pi s}{2\pi s}\right)^2 - \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\sin 2\pi s}{2\pi s}\right)\int_0^s \frac{\sin 2\pi t}{2\pi t}\mathrm{d}t\right]. \quad (73)$$

In these summarized results s is measured in the units of mean spacing between neighboring eigenvalues.

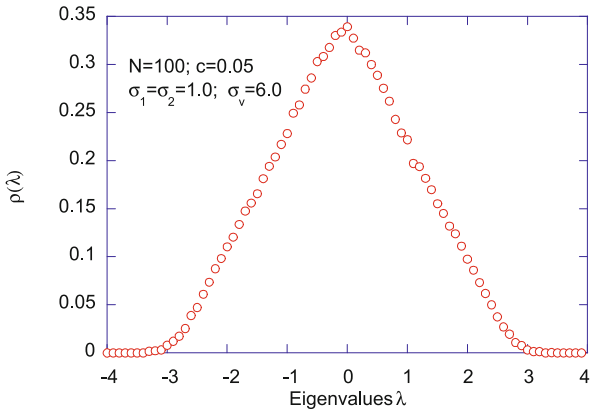
Random Matrix Theory and Complex Systems

In this section we will present several examples where RMT is pertinent to complex systems ranging from complex networks to finance.

Spectra of Graphs

Many complex systems whether technological, social or biological have a network structure. The dynamics and the growth of these networks play a major role in our daily life. Spread of disease, information, traffic flow, electricity distribution; all depend on the structure of their related networks. Understandably, the complex network studies are of great interest to researchers in a diverse range of disciplines. From a mathematical point of view complex networks are modeled as graphs [4,7,16,20,21,33] which are sets of many interacting components namely vertices (points) connected by edges (lines).

The eigenvalues of the adjacency matrices of graphs are related to their topological structures and transport properties. Random Matrix Theory is popularly employed to predict the spectral properties of various graphs. Recent numerical studies show that the eigenvalue spectrum of real-world networks, graphs with scale free links distribution, deviates from those of uncorrelated random graphs. In the large N (number of vertices) limit spectral density of random graphs (not sparse [47]) converge to a semi-



Random Matrix Theory, Figure 5
Two block diagonal random matrices with intra system interaction σ_1 and σ_2 coupled via very sparse (c = 0.05 connections) but strong σ_v couplings strength gives a triangular spectrum for a 100 × 100 system size

circle distribution, while the scale free networks develop a triangle density distribution with a power-law tail, and the small-world networks have a complex spectral density with several sharp peaks [22,23,28]. In a block structured random matrix model introduced by Ergun [17] to study coupled systems or a system with modules, it was shown that introduction of a few strong inter-couplings between sub-systems gives a triangular shape spectrum (see Fig. 5). This model can be thought of as a network with sub-networks with weighted links within and between subnetworks.

Composite Materials

Electrical and optical properties of metal-dielectric mixtures (composites) are frequently investigated by employing randomly connected Resistor-Capacitor (R-C) networks or Inductor (in series with a resistor)-Capacitor (RL-C) networks [1,2,6,8,34]. In such models one considers a binary inhomogeneous media with a random network of metallic bonds (conductance σ_0, concentration p) and dielectric bonds (conductance σ_1, concentration $1-p$). If the reactance network is large enough then there will always be circuits of resonance type with purely imaginary conductance showing poles at some frequencies. These poles are a particular example of eigenvalues of *matrix pencils* $\mathcal{H}-\lambda\mathcal{W}$ where $\mathcal{W}$ is a positive definite and $\mathcal{H}$ is a random real symmetric matrix.

Spectral studies of Kirchhoff or popularly known as Laplacian [50] matrices and random banded matrices [41] can be considered as the theoretical framework of the spectra of such networks.

Protein–Protein Interactions

Protein–protein interactions (PPIs) are essential for every living cell, information about these interactions helps to understand diseases (e. g. cancer), and to develop effective methodologies to treat them. There exists a vast amount of data on human PPIs and more are being mined every day. Currently, the use of network analogies [31,43] on these data suggests that PPIs form an interacting network of networks and have much in common with real world networks such as the World Wide Web and the Internet. In particular, the distribution of PPIs is scaled with power laws; such a scaling is often associated with robustness to random attacks. The underlying principles of power laws in networks with human attributes are identified to be: (1) continuous growth, and (2) preferential attachment processes. These considerations cannot hold at a microscopic level and therefore one must consider some other mechanisms such as spatial constraints and chemical reactions to account for the hierarchical organizations at the cellular level. In an empirical study of the yeast core protein interaction network Luo et al. [35] showed that the Wigner Surmise for the GOE case is a good candidate to see the spectral correlation of the yeast interaction matrix. Their study is essentially confirming that the yeast core protein network is a network of sub networks, which share the spectral properties of block-structured random networks.

To put their observation onto a theoretical footing consider the following: If $\mathcal{M}$ is a block diagonal matrix with $\mathcal{H}_1$ and $\mathcal{H}_2$ being its two blocks of $N\times N$ matrices with the same statistics then the eigenvalues of the sub matrices will form two uncorrelated sub sequences leading to a spacing distribution with $\rho(s)=1/2$ at $s=0$.

$$\mathcal{M}=\begin{pmatrix}\mathcal{H}_1 & \oslash\\ \oslash & \mathcal{H}_2\end{pmatrix}$$

$$\Longrightarrow\begin{pmatrix}\lambda_1^1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_N^1 & & & & \\ & & & \lambda_1^2 & & & \\ & & & & \ddots & & \\ & & & & & \lambda_N^2 & \end{pmatrix}.$$

Further increasing the number of non-interacting sub matrices eventually leads to a Poisson spacing distribution where all eigenvalues are uncorrelated. Conversely, if we introduce interactions between $\mathcal{H}_1$ and $\mathcal{H}_2$ via an off diagonal block matrix then the two sub sequences of eigenvalues will merge to a correlated single sequence getting back to Wigner surmise spacing statistics.

Ecological Webs

In general, interactions between ecological species are highly non-linear and may obey a generalized Lotka–Volterra competition relation:

$$\dot{x}_i = x_i \left(1 - \sum_{j=1}^{N} a_{ij} x_j \right) ,$$

where x_i is a function of time giving the density of the ith species, a_{ij} represents the effect of inter-specific (if $i \neq j$), intra-specific (if $i = j$) interactions and a_{ij} forms the $N \times N$ interaction matrix $\mathcal{A}$. The stability of the possible equilibrium or time independent configurations of such systems may be studied by linearization of this equation around equilibrium. The linearized model can be written as

$$\dot{\mathbf{x}} = \mathcal{J}\mathbf{x} ,$$

where $\mathcal{J}$ is the Jacobian matrix having $J_{ij} = x_i^* a_{ij}$ elements with x_i^* equilibrium value of species i, and bold face $\mathbf{x}$ is the column vector of N species. In this formalism an equilibrium is stable if the eigenvalues, of the Jacobian matrix have negative real parts.

In the 1970s Gardner and Ashby [25] introduced a similar linear model

$$\dot{\mathbf{x}} = \mathcal{A}\mathbf{x}$$

to study how the stability of a system would change with its size or if the connections within the system were incomplete, i. e., the elements of the matrix $\mathcal{A}$ are taken from a specified distribution with probability C "connectance" if there is connection (or interaction) between x_i and x_j and set to zero with probability $(1 - C)$ if no-interaction. Their Monte Carlo-type study showed that large complex systems which are assembled randomly would be stable up to a critical level of connectance. May (1972) [36] furthered these investigations by introducing a random matrix model where the matrix

$$\mathcal{A} = \mathcal{B} - \mathcal{I} ,$$

$\mathcal{I}$ is a unit matrix and $\mathcal{B}$ is a random matrix. The elements of $\mathcal{B}$ are picked independently and identically from a probability distribution. From his analysis he linked the stability of a complex system to its spectral radius namely, the spectral radius of a random $N \times N$ matrix $\mathcal{A}$ with connectance C and average interaction strength α is given by

$$\rho(\mathcal{A}) = \alpha(NC)^{1/2} .$$

These investigations started a surge of research interest to employ appropriate random matrices as ecological interaction matrices (sometimes referred to as the community matrices) to study species stability and variability in ecosystems.

Quasi-Species

The macromolecular evolution model was introduced in the 1970s by Manfred Eigen [14] to describe the process of selection in a collection of self-reproducing macromolecular species by employing a set of non-linear equations. Jones et al. [32] showed that in equilibrium, one can determine the solutions of these equations via the spectrum of a real positive matrix which represents the replication rates and mutation probabilities of the macromolecules. The largest real eigenvalue of this matrix gives the average replication rates and mutation probabilities. The corresponding eigenvector, also called Perron vector, describes densities of the quasi-species and hence which can only take real and positive values. This condition limits the number of equilibrium solutions thus a particular interest is to determine what values of a positive matrix guarantee this selective behavior, see for instance the works of Rumschitzki [48] and McCaskill [37].

Financial Correlation Matrices

The study of correlation (or covariance) matrices has a long history in finance dating back to the Markowitz's theory of optimal portfolios. The fundamental question is: Given a set of financial assets characterized by their average return and risk, what is the optimal weight of each asset, such that the overall portfolio provides the best return for a fixed level of risk, or conversely, the smallest risk for a given overall return? This question becomes even more difficult when the data is noisy or incomplete. Potters et al. [46] has suggested using RMT statistics to clean the empirical correlation matrices to minimize possible biases in the estimation of future risk. They also showed that the largest eigenvalue of the correlation matrices has a financial interpretation of being the "market mode".

Summary

We have introduced a number of key concepts used in random matrix theory and given several examples where RMT is utilized. To summarize, this theory was developed by physicists to describe the spectra of complex nuclei. But, nowadays, it is applied to many diverse research areas from quantum chaos, disordered systems to finance.

Below we give a short list of some of the frequently used results which were covered in this article.

The Gaussian ensembles are categorized into three universality classes: (1) Orthogonal ensembles, systems that have time-reversal invariance; the majority of practical systems are described by these ensembles, (2) unitary ensembles, which are used when the time-reversal invariance is violated, and (3) symplectic ensembles apply only to odd-spin systems without rotational symmetry.

The joint probability distribution for all Gaussian ensembles with $N \times N$ random matrix H is given by the general expression

$$P_{N\beta}(H) = C_{N\beta} e^{-\frac{N\beta}{4\sigma^2}\mathrm{Tr}[\mathcal{H}^2]} ,$$

where Tr stands for the *trace*, and β is used as an index for the ensemble classification. Then the eigenvalues of the Gaussian ensembles are distributed according to

$$P_{N\beta}(\lambda_1, \ldots, \lambda_N) = C_{N\beta} \prod_{i<j} |\lambda_i - \lambda_j|^\beta \exp\left(-\frac{N\beta}{4\sigma^2} \sum_{k=1}^{N} \lambda_k^2\right) ,$$

where the constant $C_{N\beta}$ is an ensemble dependent constant, and it is chosen such that the $P_{N\beta}$ is normalized to unity. This expression is common for all Gaussian ensembles and derived from the joint probability distribution given above. Integrating this expression over all eigenvalues except one gives the semicircle law asymptotically,

$$\langle\rho(\lambda)\rangle = \begin{cases} \frac{1}{2\pi\beta\sigma^2}(4N\beta\sigma^2 - \lambda^2)^{1/2} , & |\lambda| < 2\sqrt{\beta\sigma^2 N} \\ 0 , & |\lambda| > 2\sqrt{\beta\sigma^2 N} \end{cases} ,$$

where $\langle\rho(\lambda)\rangle$ is normalized to N, and σ^2 is the variance of the matrix elements.

The three Wigner laws (referred to as the *Wigner Surmise*) for the nearest neighbor spacing distributions for 2×2 matrices are

$$P(s) = \begin{cases} \frac{\pi}{2} s e^{-s^2\pi/4} & \beta = 1 \quad \text{(GOE)} \\ \frac{32}{\pi^2} s^2 e^{-s^2 4/\pi} & \beta = 2 \quad \text{(GUE)} \\ \frac{2^{18}}{3^6\pi^3 2} s^4 e^{-s^2 64/9\pi} & \beta = 4 \quad \text{(GSE)} \end{cases} .$$

They are also found to be very good approximations in the asymptotic limit i. e., $N \to \infty$ derivations.

Future Directions

The use of random matrices in sub-atomic level interactions has led to many great mathematical theorems and shed new light on the understanding of nuclear spectra. Other branches of physical and mathematical sciences equally benefited from the resultant theorems yet biological sciences haven't had the same share despite an important link between random matrices and the field of biological complexity that was established in the 1970s by Robert May. With recent technological advances massive amounts of DNA data sets are being produced, however, the understanding of protein–protein or gene interactions still lags behind. Studies of structured sparse random matrices in the context of DNA data sets may produce fruitful results. Another area where random matrix theory is not fully consulted is the interpretation of the spectra of the adjacency matrices of various graphs and the spectra of the Laplacian of these graphs. There is no doubt that Random Matrix Theory is an effective tool to study a variety of complex systems; one would also hope to see its usage in the theory of machine learning and climate modeling.

Bibliography

Primary Literature

1. Almond DP, Bowen CR (2004) Anomalous power law dispersions in AC conductivity and permittivity shown to be characteristics of microstructural electrical networks. Phy Rev Lett 92:1576011–1576014
2. Almond DP, Vainas B (1999) The dielectric properties of random R-C networks as an explanation of the universal power law dielectric response of solids. J Phys: Condensed Matter 11:9081–9093
3. Bak P, Sneppen K (1993) Punctuated equilibrium and criticality in a simple model of evolution. Phys Rev Lett 71:4083
4. Barabási A-L, Albert R (1999) Emergence of scaling in random networks. Science 286:509
5. Bohigas O, Giannoni M-J, Schmit C (1984) Characterization of chaotic quantum spectra and universality of level fluctuation laws. Phys Rev Lett 52:1
6. Clerc P, Giraud G, Laugier JM, Luck JM (1990) The electrical conductivity of binary disordered systems, percolation clusters, fractals and related models. Adv Phys 39:191–309
7. Dorogovtsev SN, Mendes JFF, Samukhin AN (2000) Structure of Growing Networks: Exact Solution of the Barabási–Albert's Model. Phys Rev Lett 85:4633
8. Dykhne AM (1970) Anomalous resistance of a plasma in a strong magnetic field. Zh Eksp Teor Fiz 59:641–647
9. Dyson FJ (1962) A Brownian-motion model for the eigenvalues of a random matrix. J Math Phys 3:1191
10. Dyson FJ (1962) Statistical theory of the energy levels of complex systems, I. J Math Phys 3:140
11. Dyson FJ (1962) Statistical theory of the energy levels of complex systems, II. J Math Phys 3:157
12. Dyson FJ (1962) Statistical theory of the energy levels of complex systems, III. J Math Phys 3:166
13. Efetov KB (1983) Supersymmetry and theory of disordered metals. Adv Phys 32:53

14. Eigen M (1971) Selforganisation of sequence space and tensor products of representation spaces. Naturwissenschaften 58:465–523
15. Erdős P, Rényi A (1959) On random graphs I. Publ Math 6:290
16. Ergün G (2002) Human sexual contact network as a bipartite graph. Physica A 308:483
17. Ergün G, Semicircle to triangular distribution of density of states: Supersymmetric Approach (submitted)
18. Ergün G, Separation of a large eigenvalue from the bulk of the spectrum (submitted)
19. Ergün G, Fyodorov YV (2003) Level Curvature distribution in a model of two uncoupled chaotic subsystems. Phys Rev E 68:046124
20. Ergün G, Rodgers GJ (2002) Growing Random Networks with Fitness. Physica A 303:261
21. Ergün G, Zheng D (2003) Coupled Growing Networks. Adv Comp Syst 6:4
22. Farkas IJ, Derényi I, Barabási A-L, Vicsek T (2001) Spectra of real-world graphs: Beyond the semi-circle law. Phys Rev E 64:026704
23. Farkas IJ, Derényi I, Jeong H, Neda Z, Oltvai ZN, Ravasz E, Schubert A, Barabási A-L, Vicsek T (2002) Networks in life: Scaling properties and eigenvalue spectra. Physica A 314:25
24. Fyodorov YV, Chubykalo OA, Izrailev FM, Casati G (1996) Wigner random banded matrices with sparse structure: Local spectral density of states. Phys Rev Lett 76:1603
25. Gardner MR, Ashby WR (1970) Connectance of large daynamic systems: Critical values for stability. Nature 228:784
26. Gaudin M (1961) Sur la loi limite de l'espacement des valeurs propres d'une matrice aléatoire. Nucl Phys 25:447
27. Gaudin M, Mehta ML (1960) On the density of eigenvalues of a random matrix. Nucl Phys 18:420
28. Goh K-I, Kahng B, Kim D (2001) Spectra and eigenvectors of scale-free networks. Phys Rev E 64:051903
29. Gurevich II, Pevsner MI (1957) Repulsion of nuclear levels. Nucl Phys 2:575
30. Haake F (2000) Quantum Signatures of Chaos, 2nd edn. Springer, Berlin
31. Jeong H, Mason SP, Barabási A-L, Oltvai ZN (2001) Lethality and centrality in protein networks. Nature 411:41–42
32. Jones BL, Enns RH, Rangnekar SS (1976) On the theory of selection of coupled macromolecular systems. Bull Math Biol 38:15–28
33. Krapivsky PL, Redner S, Leyvraz F (2000) Connectivity of growing random networks. Phys Rev Lett 85:4629
34. Luck JM, Jonckheere T (1998) Dielectric resonances of binary random networks. J Phys A: Math Gen 31:3687–3717
35. Luo F, Zhongb J, Yang Y, Scheuermann RH, Zhou J (2006) Application of random matrix theory to biological networks. Phys Lett A 357:420–423
36. May RL (1972) Will a large complex system be stable? Nature 238:413
37. McCaskill JS (1984) Localisation threshold for macromolecular quasispecies from continuously distributed replication rates. J Chem Phys 80:5194–5202
38. Mehta ML (1991) Random Matrices, 2nd edn. Academic, San Diego
39. Mehta ML, Gaudin M (1960) On the density of eigenvalues of a random matrix. Nuc Phys 18:420
40. Mirlin AD (2000) Statistics of energy levels and eigenfunctions in disordered systems. Phys Rep 326:259
41. Mirlin AD, Fyodorov YV (1991) Universality of level correlation function of sparse random matrics. J Phys A 24:2273–2286
42. Newman MEJ, Strogatz SH, Watts DJ (2001) Random graphs with arbitrary degree distributions and their applications. Phys Rev E 64:026118
43. Pellegrini M, Haynor D, Johnson JM (2004) Protein interaction networks. Expert Rev Proteomics 1:239–249
44. Porter CE, Rosenzweig N (1960) Statistical properties of atomic and nuclear spectra. Suomalaisen Tiedeakatemian Toimituksia (Ann. Acad. Sci. Fennicae) AVI Phys 44:166
45. Porter CE, Thomas RG (1956) Fluctuations of nuclear reaction widths. Phys Rev 104:483
46. Potters M, Bouchaud J-P, Laloux L (2005) Financial applications of random matrix theory: Old laces and new pieces. physics/0512090
47. Rodgers GJ, Bray AJ (1988) Density of states of a sparse random matrix. Phys Rev B 37:3557
48. Rumschitzki DS (1987) Spectral properties of Eigen evolution matrices. J Math Biol 24:667–680
50. Säring J, Mehlig B, Fyodorov YV, Luck JM (2003) On the random symmetric matrices with a constraint: The spectral density of random impedence networks. Phys Rev E 67:047101
51. von Neuman J, Wigner E (1929). Phys Z 30:467
52. Watts DJ, Strogatz SH (1998) Collective dynamics of "small-world" networks. Nature 393:440
53. Wigner EP (1951) On the statistical distribution of the widths and spacing of nuclear resonance levels. Phil Soc 62:548
54. Wigner EP (1955) Characteristic vectors of bordered matrices with infinite dimensions. Ann Math 62:548
55. Wigner EP (1957) Distribution of neutron resonance levels. In: International conference on the neuron interactions with the nucleus, Columbia University, New York, 9–13 September 1957. Columbia Univ. Rept. CU-175 TID-8547
56. Wigner EP (1957) Results and theory of resonance absorption (Conference on Neutron Physics by Time-of-Flight, Gatlinburg, Tenessee, November 1–2, 1956). Oak Ridge Natl Lab Rept ORNL-2309:59
57. Wigner EP (1957) Statistical Properties of Real Symmetric Matrices with Many dimensions. In: Can. Math. Congr. Proc., 174. Univ of Toronto Press, Toronto
58. Wigner EP (1958) On the distribution of the roots of certain symmetric matrices. Ann Math 67:2
59. Wishart J (1928) The generalised product moment distribution in samples from a normal multivariate population. Biometrica 20A:32

Books and Reviews

Abramowitz M, Stegun I (1972) Handbook of mathematical fucntions, 10th edn. Dover, New York
Albert R, Barabási A-L (2002) Statistical mechanics of complex networks. Rev Mod Phys 74:47
Bak P (1996) How Nature Works: The science of self-organised criticality. Copernicus, New York
Barabási A-L (2002) Linked: How everything is connected to everything else and what it means. Perseus, Cambridge
Bellman R (1997) Introduction to Matrix Analysis, 2nd edn. SIAM, Philadelphia
Berezin FA (1987) Introduction to Superanalysis. Reidel, Dordrecht
Bohigas O (1991) Random matrix theories and chaotic dynamics, Session LII, 1989. Chaos and Quantum Physics, Les Houches

Bollobás B (2001) Random Graphs, 2nd edn. Cambridge Univ Press, Cambridge
Brody TA, Flores J, French JB, Mello PA, Pandey A, Wong SSM (1981) Random-matrix physics: spectrum and strength fluctuations. Rev Mod Phys 53:385
Cvetković D, Domb M, Sachs H (1995) Spectra of Graphs: Theory and Applications. Johann Ambrosius Barth, Heidelberg
Cvetković D, Rowlinson P, Simić S (1997) Eigenspaces of Graphs. Cambridge Univ Press, Cambridge
Dorogovtsev SN, Mendes JFF (2002) Evolution of networks. Adv Phys 51:1079
Efetov KB (1997) Supersymmetry in Disorder and Chaos. Cambridge Univ Press, Cambridge
Faloutsos M, Faloutsos P, Faloutsos C (1999) On power-law relationships of the internet topology. Proc ACM SIGCOMM Comp Comm Rev 29:251
Fyodorov YV (1995) Basic features of Efetov's supersymmetry approach. Mesoscopic Quantum Physics, Les Houches, Session LXI, 1994
Guhr T, Müller-Groeling A, Weidenmüller HA (1998) Random matrix theories in quantum physics: common concepts. Phys Rep 299:189
Hinch EJ (1991) Perturbation Methods. Cambridge Univ Press, Cambridge
Lieb EH, Mattis DC (1966) Mathematical Physics in One Dimension. Academic, New York
Markowitz H (1959) Portfolio Selection: Efficient Diversification of Investments. Wiley, New York
Merzbacher E (1970) Quantum Mechanics, 2nd edn. Wiley, London
Porter CE (1965) Statistical Theories of Spectra: fluctuations. Academic, New York
Stöckmann H-J (2000) Quantum Chaos: an introduction. Cambridge Univ Press, Cambridge
Watts DJ (1999) Small Worlds. Princeton Univ Press, Princeton
Wilks S (1972) Mathematical Statistics. Wiley, Japan
Wilson RJ (1996) Graph Theory. Longman, Edingburg

Random Walks in Random Environment

Ofer Zeitouni
Department of Mathematics, University of Minnesota, Minneapolis, USA

Article Outline

Glossary

Aging A stochastic process X_t exhibits aging if for some function $f(x, y)$, the limit of $Ef(X_t, X_{t+s_t})$ as $t \to \infty$ exists and is not trivial for a function $s_t \to_{t\to\infty} \infty$.

Annealed law Average of the quenched law, over all possible environments.

Brownian motion A continuous process $\{X_t\}$ with continuous sample paths, $X_0 = 0$, with increments over non-overlapping intervals independent and normally distributed with zero mean and variance equal to the length of the interval.

Convex A function $f\colon X \to \mathbb{R}$, where X is a linear space, is convex if for $\alpha \in [0,1]$ and $x, y \in X$, $f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha)f(y)$.

Ergodic law If P is a law on a collection of random variables $\{X - z\}_{z\in\mathbb{Z}^d}$, then it is ergodic if for any event A that is invariant under a transformation $X_z \mapsto X_{z+e}$, with $e \in \mathbb{Z}^d$, $|e| = 1$, it holds that $P(A) \in \{0, 1\}$.

Environment Collection of transition probabilities indexed by the sites of the lattice $\mathbb{Z}^d$, that is a collection of vectors belonging to the $2d$-simplex.

i.i.d. Independent identically distributed.

Large deviations principle (LDP) A sequence of probability measures P_n on a common (topological) space X satisfies the large deviations principle if for some non-negative function $I\colon X \to \mathbb{R}_+$ with closed level sets $\{x \in X\colon I(x) \le a\}$, for any $A \subset X$,

$$-\inf_{x\in A^o} I(x) \le \liminf n^{-1} \log P_n(A) \le \limsup n^{-1} \log P_n(A) \le -\inf_{x\in\bar{A}} I(x)$$

where $\bar{A}$ is the closure of A and A^o is its interior.

Law Probability distribution. When dealing with a stochastic process $\{X_t\}$, this is the collection of distribution functions $P(X_{t_1} \le x_1, \ldots, X_{t_k} \le x_k)$.

Law of large numbers (LLN) A sequence X_n satisfies the LLN if X_n/n converges to a deterministic limit.

Lipschitz norm Given a function on a normed space X with norm $\|\cdot\|$, its Lipschitz norm is defined as $\sup_{x\neq y} |f(x) - f(y)|/\|x - y\|$.

Martingale Let $\mathcal{F}_n$ denote a filtration, that is an increasing sequence of σ-algebras. A process $\{X_n\}$ is a (discrete time) martingale with respect to $\mathcal{F}_n$ if $E(X_{n+1}|\mathcal{F}_n) = X_n$.

Mixing environment An environment whose law satisfies a *mixing condition*, that is if A, B are events that depend on a finite number of sites then $P(A \cap T^z B) \to_{|z|\to\infty} P(A)P(B)$ where T^z is the shift of B by $z \in \mathbb{Z}^d$. For $d = 1$, the environment is strongly mix-

ing if the convergence above is uniform in the choice of A, B of the same local dependence.

P-a.s. A property holds P-a.s. if the probability of the event that the property does not hold (under the law P) vanishes.

Quenched law Law of the random walk, with the environment given.

Variational distance For two probability distributions ν, μ, the variational distance is defined as $\sup_A |\nu(A) - \mu(A)|$.

2d-Simplex Collection of vectors of length $2d$ with nonnegative entries that sum to 1.

Definition of the Subject

The term random walks in random environments (abbreviated RWRE) refers to a class of Markov chains evolving in $\mathbb{Z}^d$, where the transition mechanism itself is random, and forms a stationary random field indexed by $\mathbb{Z}^d$. They first appeared in the literature in the early 1970s, for $d = 1$, modeling in a natural way transport properties of a tagged particle in an inhomogeneous medium, as well as biological systems. They can be naturally used to model complex transport phenomena in any dimension.

The analysis of the case $d = 1$ during the 1970s and early 1980s, revealed that the behavior of such random walks can differ drastically from that of ordinary random walks (that correspond to a homogeneous environment). The extension to higher dimensions ($d \geq 2$) was proposed and studied in [51], but it is only in the last decade, starting with the work [102], that rapid progress has been made, although several basic questions remain open.

Introduction

We recall that random walks and their scaling limits, diffusion processes, provide a simple yet powerful description of random processes, and are fundamental in the description of many fields, from biology through economics, engineering, and statistical mechanics. A large body of work has accumulated concerning the properties of such processes, and very detailed information is available. We refer to [61] and [93] for background on random walks and diffusion processes.

In many situations, the medium in which the process evolves is highly irregular. Without further modeling, this results with spatially inhomogeneous Markov processes, and not much can be said. Things are however different if some degree of homogeneity is assumed on the *law* of the environment. When the underlying state space on which the walk moves with nearest neighbor steps is the lattice $\mathbb{Z}^d$, $d \geq 1$, and the law of the environment is assumed stationary, we call the resulting random walk a *random walks in random environment* (RWRE).

Informally, one starts with an *environment*, which is a configuration of transition probabilities (one transition probability for each site in $\mathbb{Z}^d$). One then starts a random walk that, when at location $z \in \mathbb{Z}^d$, moves at the next time step according to the transition probability associated with z. In the one dimensional case, this model has been thoroughly analyzed, and it is known that the behavior can differ dramatically from the behavior of ordinary random walk (in particular, the standard central limit theorem can fail). In higher dimensions, it is expected that the RWRE behaves more like ordinary random walk, and much of the available analysis points in this direction; however, gaps in the understanding of the model still remain, and the law of large numbers has not yet been proved in full generality.

A precise formulation of the RWRE model is as follows. Let S denote the 2d-dimensional simplex, set $\Omega = S^{Z^d}$, and let $\omega(z, \cdot) = \{\omega(z, z + e)\}_{e \in Z^d, |e|=1}$ denote the coordinate of $\omega \in \Omega$ corresponding to $z \in Z^d$. ω is an "environment" for an inhomogeneous nearest neighbor random walk (RWRE) started at x with *quenched* transition probabilities $P_\omega(X_{n+1} = z + e | X_n = z) = \omega(x, x + e)$ ($e \in \mathbb{Z}^d, |e| = 1$), whose law is denoted P^x_ω. We write E^x_ω for expectations with respect to the law P^x_ω, and write P_ω and E_ω for P^0_ω and E^0_ω. In the RWRE model, the environment is random, of law P, which is always assumed stationary and ergodic. We often assume that the environment is *uniformly elliptic*, that is there exists an $\epsilon > 0$ such that P-a.s., $\omega(x, x + e) \geq \epsilon$ for all $x, e \in \mathbb{Z}^d, |e| = 1$. Finally, we denote by $\mathbb{P}$ the *annealed* law of the RWRE started at 0, that is the law of $\{X_n\}$ under the measure $P \times P^0_\omega$, and again we write $\mathbb{E}$ for expectations with respect to $\mathbb{P}$ and E for expectations with respect to P. Note that under the law P_ω, the RWRE is a (space inhomogeneous) Markov process, whereas under $\mathbb{P}$, its law is space homogeneous but not Markovian, since the knowledge of the path of the RWRE up to time n modifies the a-priori law on the environment.

To get an idea of some of the unusual features of the RWRE model, we begin by discussing the one dimensional case. (A further motivation is that the study of certain reinforced random walks can be reduced to that of one dimensional RWRE, a connection first described in [73]). This model is fairly well understood, and we review the results in Sect. "One Dimensional RWRE", emphasizing the points in which there are differences from walks in a homogeneous environment. We then turn in Sect. "Multi Dimensional RWRE – non Perturbative Regime" to the multidimensional case, while Sect. "Multi Dimensional RWRE – the Perturbative Regime" is devoted to the per-

turbative regime. Sect. "Diffusions in Random Environments" quickly reviews the available results for the related model of (non reversible) diffusions in random environments. In Sect. "Topics Left Out and Future Directions", we collect some information about related models. This article borrows heavily from [101,107] and [108].

One Dimensional RWRE

When $d = 1$, we write $\omega_x = \omega(x, x+1)$, $\rho_x = (1-\omega_x)/\omega_x$, and $u = E\log\rho_0$. Because explicit recursions and computations are possible in the one dimensional case, the understanding of the RWRE is rather complete. We summarize the main features below.

Ergodic Behavior, Limit Laws, and Traps

Recall that for a homogeneous environment (that is, when the stationary measure P has a marginal which charges a single value: $\omega_i = \bar{\omega}$ for all i), we have $X_n/n \to v_{\bar{\omega}} := 2\bar{\omega} - 1$ and $(X_n - nv_{\bar{\omega}})/\sqrt{4\bar{\omega}(1-\bar{\omega})n}$ converges in distribution to a standard Gaussian. We describe next the corresponding results for the RWRE model, emphasizing the dramatic difference in behavior.

As it turns out, the sign of u determines the direction of escape of the RWRE, while the limiting behavior depends on an explicit function of the law of the environment. The following theorem is essentially due to [92], see also [1,107].

Theorem 1 (Transience, recurrence, limit speed, $d = 1$) *(a) If $u < 0$ then $X_n \to_{n\to\infty} \infty$, $\mathbb{P}$-a.s. If $u > 0$ then $X_n \to -\infty$, $\mathbb{P}$-a.s. Finally, if $u = 0$ then the RWRE oscillates, that is, $\mathbb{P}$-a.s.,*

$$\limsup_{n\to\infty} X_n = \infty\,, \quad \liminf_{n\to\infty} X_n = -\infty\,.$$

Further, there is a deterministic v such that

$$\lim_{n\to\infty} \frac{X_n}{n} = v\,,\ \mathbb{P} - a.s.\,. \tag{1}$$

If P is a product measure, then

$$v = \begin{cases} (1 - E(\rho_0))/(1 + E(\rho_0))\,, & E(\rho_0) < 1\,, \\ -(1 - E(\rho_0^{-1}))/(1 + E(\rho_0^{-1}))\,, & E(\rho_0^{-1}) < 1\,, \\ 0\,, & else\,. \end{cases} \tag{2}$$

The statement (1) that X_n/n converges to a deterministic limit (under both the quenched and annealed measures) is referred to as a *law of large numbers* (LLN).

Remark 2 The surprising features of the RWRE model can be appreciated if one notes the following facts, all for a product measure P.

- Suppose $u < 0$, that is $X_n \to \infty$, $\mathbb{P}$-a.s. By Jensen's inequality, $\log E\rho_0 \geq E\log\rho_0$, but it is quite possible that $E\rho_0 > 1$. Thus, it is possible to construct i.i.d. environments in which the RWRE is transient, but the speed v vanishes.
- Suppose $\bar{v} = 2E\omega_0 - 1$ denotes the speed of a (biased) simple random walk with probability of jump to the right equal, at any site, to $E\omega_0$. It is easy to construct examples with $\bar{v} > 0$ but $u > 0$, which means that $X_n \to -\infty$ even if the static speed $\bar{v}$ points to the right. However, by Jensen's inequality, $v < 0$ implies that $\bar{v} < 0$. Thus, if the static speed $\bar{v}$ is positive, the RWRE may be transient to the left but if so, only with zero speed. We come back to this point in Subsubsect. "RWRE with Deterministic Components", where we show that the last property is not necessarily true in high dimension.
- Another application of Jensen's inequality reveals that $|v| \leq |\bar{v}|$, with examples of strict inequality readily available. Thus, the effect of the random environment is to force a *slowdown* with respect to the (averaged, deterministic) environment.

Another aspect in which the RWRE differs from the standard model of random walk is in its fluctuations. Consider product measures P with $u := E(\log\rho_0) < 0$ (i. e., RWRE is transient to $+\infty$). Set $s = \sup\{r\colon E(\rho_0^r) < 1\}$ and note that because $u < 0$, necessarily $s \in (0, \infty]$. When $s > 2$, the behavior is similar to standard random walk, and one has a central limit theorem of the following form. For some deterministic strictly positive constant σ, the random variable $W_n := (X_n - nv)/\sigma\sqrt{n}$ converges, under the annealed law, to a standard Gaussian random variable, that is

$$\mathbb{P}(W_n > x) \to_{n\to\infty} \frac{1}{\sqrt{2\pi}} \int_x^\infty \mathrm{e}^{-y^2/2}\mathrm{d}y\,,$$

see [46,55,75,107] for this statement and a similar one under the quenched law (with random centering). On the other hand, when $s < 2$, due to the existence of localized pockets of environments ("traps") where the walk spends a large time, one gets stable limit laws. In particular, for $s \in (0, 1)$, under mild assumptions on the law of ρ_0, the random variable X_n/n^s converges, under the annealed law, to a stable random variable with parameters (s, b), where b is a deterministic constant, whose value has been identified

in [39]. (A stable law with parameters (s, b) is the distribution of a random variable S with characteristic function

$$E(e^{itS}) = \exp\left(-b|t|^s\left(1 - i\frac{t}{|t|}\tan(\pi s/2)\right)\right), \ t \in \mathbb{R}.)$$

We refer to the regime where $s \in (0, 2]$, in which the fluctuations are not in the CLT scale, [55], as the *sub-diffusive* regime. It is interesting to note that these statements do *not* possess a quenched counterpart, and in fact a quenched limit law is not possible for $s < 2$, see [75,76].

Remark 3 When P is a strongly mixing environment, the parameter s has to be defined differently, by means of the large deviations rate function for the variable $n^{-1}\sum_{i=1}^{n}\log\rho_i$. The Gaussian limit laws apply in such situations when $s > 2$, see [20,75,107]. The stable limit laws are more delicate, and are not known for general ergodic environments with good mixing properties. For a class of Markovian environments, such results are contained in [67].

Limit Laws and Aging, Recurrent RWRE: Sinai's Walk

When $E(\log\rho_0) = 0$, the traps alluded to in the previous section stop being local, and the whole environment becomes a diffused trap. The walk spends most of its time "at the bottom of the trap", and as time evolves it is harder and harder for the RWRE to move. This is the phenomenon of *aging*, captured in the following theorem:

Theorem 4 *There exists a random variable B^n, depending on the environment only, such that for any $\eta > 0$,*

$$\mathbb{P}\left(\left|\frac{X_n}{(\log n)^2} - B^n\right| > \eta\right) \underset{n\to\infty}{\to} 0 .$$

Further, for $h > 1$,

$$\lim_{\eta\to 0}\lim_{n\to\infty}\mathbb{P}\left(\frac{|X_{n^h} - X_n|}{(\log n)^2} < \eta\right) = \frac{1}{h^2}\left[\frac{5}{3} - \frac{2}{3}e^{-(h-1)}\right]. \tag{3}$$

The first part of Theorem 4 is due to Sinai [91], with Kesten [52] providing the evaluation of the limiting law of B^n, see also [47]. It is actually not hard to understand the anomalous scaling $(\log n)^2$: indeed, the time for the particle to overcome a stretch of the environment of length $c_1 \log n$ in which the drift points "backwards" is exponential in $c_1 \log n$, i. e. an appropriate c_1 can be chosen such that this time is of order n. Hence, the range of the RWRE at time n cannot be larger than the distance in which there exists such a stretch. Due to the scaling properties of random walk, this distance is of order $(\log n)^2$.

The second part of Theorem 4 is implicit in [47], and also follows from the analysis of the time spent by the RWRE at "bottom of traps". We refer to [62] for a detailed study of aging in the Sinai model by renormalization techniques, and to [24,30,107] for rigorous proofs that avoid renormalization arguments. For information concerning the time spent by the walk at the most visited site (which can be of order n in the Sinai model), see [28,49,89], and [43] for the transient case.

Tail Estimates and Large deviations

Another question of interest relates to the probability of seeing a-typical behavior of the RWRE. These probabilities turn out to depend on the measure discussed, that is whether one considers the quenched or annealed measures.

Following Varadhan [105], we say that the sequence of random variables X_n/n satisfies the large deviations principle (LDP) with convex, continuous rate function I on a compact set K, if for any measurable subset of K,

$$\lim_{n\to\infty}\frac{1}{n}\log P(X_n/n \in A) = -\inf_{x\in A} I(x) .$$

Cramér's theorem (Theorem 2.2.3 in [29]) states that rescaled random walk X_n/n in a homogeneous environment with $\omega_i = \bar{\omega}$ for all i satisfies the LDP with a strictly convex rate function $I(x)$ that vanishes only on $v_{\bar{\omega}} = 2\bar{\omega} - 1$. The situation is different for the RWRE. The following theorem combines results from [25,32,48].

Theorem 5 *For P-a.e. realization of the environment ω, the sequence X_n/n satisfies, under P^0_ω, a LDP on $[-1, 1]$ with a deterministic convex rate function $I_P(\cdot)$. If P is an i.i.d. measure, then under the annealed measure $\mathbb{P}$, the same sequence satisfies a LDP with convex rate function*

$$I(x) = \inf_{Q\in\mathcal{M}_1^e}(h(Q|P) + I_Q(x)) , \tag{4}$$

where $h(Q|P)$ is the specific entropy of Q with respect to P and $\mathcal{M}_1^e$ denotes the space of stationary ergodic measures on Ω. Always, $I(x) \le I_P(x)$, and both I and I_P may vanish for $x \in [0, v]$, and only for such x. In particular, neither I nor I_P need be strictly convex.

The rate function for the RWRE thus differs from the case of homogeneous environments in two important aspects: it may vanish on the whole segment $[0, v]$, indicating subexponential behavior for the probability of slowdown, and further the rate function is in general not strictly convex. When $I(x)$ vanishes for $x \in [0, v]$, it means that the probability of seeing an a-typical slowdown of the random walk

decays at a speed less than exponentially. A precise characterization of these is available in [33,44,77,78,107].

The *specific entropy* $h(Q|P)$ (see [29] for definition) appearing in Theorem 5 measures the rate of decay of the probability that a block of the environment will look as if it came from Q, when it was generated by P. Thus, one may interpret Theorem 5 as follows: to create an annealed large deviation, one may first "modify" the environment (at a certain exponential cost measured by the specific entropy h) and then apply the quenched LDP in the new environment.

Multi Dimensional RWRE – non Perturbative Regime

We turn our attention to RWRE in the lattice $\mathbb{Z}^d$ with $d > 1$. Unless stated otherwise explicitly, we only consider in the sequel measures P that are i.i.d. and uniformly elliptic.

Ergodic Properties and a 0–1 Law

A natural starting point for the discussion of ergodic properties of the RWRE (X_n) would have been an analogue of Theorem 1. Unfortunately, obtaining such a statement has been a major challenge since the early 1980's, and is still open. To explain the challenge, we need to digress and introduce a certain conjectured 0–1 law.

Fix $\ell \in S^{d-1}$, i. e. ℓ is a unit vector in $\mathbb{R}^d$. Define the events

$$A_\ell^+ = \left\{ \lim_{n\to\infty} X_n \cdot \ell = \infty \right\} ,$$
$$A_\ell^- = \left\{ \lim_{n\to\infty} X_n \cdot \ell = -\infty \right\} .$$

The following proposition is due to Kalikow [51].

Proposition 6 *Assume P is i.i.d. and elliptic, i. e. $P(\omega(0, e) > 0) = 1$ for all e with $|e| = 1$, and that $\ell \in S^{d-1}$. Then, $\mathbb{P}(A_\ell^+ \cup A_\ell^-) \in \{0, 1\}$.*

Note that for $d = 1$, Theorem 1 implies that $\mathbb{P}(A_\ell^+) \in \{0, 1\}$. If one ever hopes to obtain a LLN, then one should be able to prove the following.

Conjecture 7 (Kalikow) *Assume P is i.i.d. and uniformly elliptic, and that $\ell \in S^{d-1}$. Then, $\mathbb{P}(A_\ell^+) \in \{0, 1\}$.*

Efforts to prove Conjecture 7 are ongoing. The following summarizes its status at the current time, and combines results from [19,111].

Theorem 8

(a) Conjecture 7 holds for $d = 1, 2$ and elliptic i.i.d. environments.

(b) There exist ergodic environments that are elliptic (for $d = 2$) and even uniformly elliptic and mixing (for $d \geq 3$), for which a deterministic direction $\ell \in S^{d-1}$ exists such that $P_\omega^0(A_\ell^+) \in (0, 1)$, for P-almost every ω.

As mentioned above, part (a) of Theorem 8 for $d = 1$ is a direct consequence of the LLN, Theorem 1.

As it turns out, the validity of Conjecture 7 is the only obstruction to a LLN, as demonstrated in the following theorem, which combines results from [4,102,106,110].

Theorem 9 *Assume P is i.i.d. and uniformly elliptic.*

(a) Fix $\ell \in S^{d-1}$. Then,

$$\lim_{n\to\infty} \frac{X_n \cdot \ell}{n} = \nu_+ \mathbf{1}_{A_\ell^+} + \nu_- \mathbf{1}_{A_\ell^-} , \mathbb{P} - a.s. . \quad (5)$$

In particular, when $d = 2$ the LLN holds true.

(b) $\mathbb{P}$-almost surely, there are at most two possible limit points, denoted ν_1, ν_2, for the sequence X_n/n. Further, ν_1, ν_2 are deterministic, and if $\nu_1 \neq \nu_2$ then there exists a constant $a \geq 0$ such that $\nu_2 = -a\nu_1$.

(c) When $d \geq 5$, if $\nu_1 \neq \nu_2$ then at least one of ν_1 and ν_2 equals 0.

Part (a) of the theorem implies that if Conjecture 7 is true, then the LLN holds for P i.i.d. and uniformly elliptic.

The proof of Theorem 9, and of many of the other results in this section, uses the machinery of regeneration times, introduced in [102]. Roughly, a random time k is a regeneration time relative to a direction $\ell \in S^{d-1}$ if $X_k \cdot \ell \geq X_n \cdot \ell$ for all $k \geq n$ but $X_k \cdot \ell < X_n \cdot \ell$ for all $k < n$ (i. e., $X_n \cdot \ell$ sets a record at time k, and never moves backward from that record). It turns out that the sequence of inter-regeneration times and inter-regeneration distances is an i.i.d. sequence under the annealed measure $\mathbb{P}$, if P is i.i.d. Once such an i.i.d. sequence has been identified, ergodic arguments yield the LLN, and the (annealed) CLT involves studying tail behavior of the regeneration times. In some ballistic cases, one may translate an annealed CLT to a quenched one, see [7] and [82]. We note in passing that regeneration techniques have also been useful for certain non-i.i.d. environments. We refer the interested reader to [26,27,79].

So far, there is no known criterion that allows one to decide the question of transience or recurrence for RWRE in dimension $d \geq 2$, although one certainly expect transience as soon as $d \geq 3$.

Ballistic Behavior and Sznitman's Conditions

Lacking an explicit expression for the speed of the RWRE for $d \geq 2$, a natural goal is to identify a large family of

models for which $X_n/n \to v \neq 0$. RWRE's that satisfy such a relation are called *ballistic*. As we saw in Theorem 1, when $d = 1$ and $X_n \to \infty$, and the environment is i.i.d., the RWRE is ballistic if and only if $E\rho_0 < 1$.

Define $d_0 := \sum[\omega(0, e_i) - \omega(0, -e_i)]e_i$ as the drift at the origin. If there exists a direction $\ell \in S^{d-1}$ such that $d_0 \cdot \ell > 0$ for P-a.e. environment, a simple martingale argument shows that $X_n/n \to v$ with $v \cdot \ell > 0$. Following Zerner [109], we call such environments *non-nestling*. We will be mainly interested in *nestling* environments, that is environments in which the origin belongs to the closed convex hull of the support of d_0. (The source for the name lies in the fact that when the walk is nestling, it is possible to construct localized regions, called traps, to which the walk return many times, leading to the mental picture of a bird that keeps returning to a nest. We note that as we show in Remark 18 below, one should not confuse the condition $Ed_0 \neq 0$ with ballistic behavior, as it does not guarantee a limiting non-zero speed.)

Traps tend to slow down the particle. However, unlike $d = 1$, all attempts to build explicit traps that slow down the particle to a sub-diffusive scale quickly fail. One thus suspects that a good control of trapping properties relates to an analysis of the RWRE. With this motivation in mind, Sznitman introduced conditions on the environment that eventually lead to a good understanding of the ballistic regime. Fix a direction $\ell \in S^{d-1}$, and for $b > 0$, define the region $U_{\ell,b,L} = \{x \in \mathbb{Z}^d : x \cdot \ell \in (-bL, L)\}$. Let $T_{\ell,b,L} = \min\{n > 0 : X_n \notin U_{\ell,b,L}\}$.

Definition 10 Let $\gamma \in (0, 1)$ be given. Then, P satisfies condition T_γ relative to ℓ if for all ℓ' in some neighborhood of ℓ, and all $b > 0$,

$$\limsup_{L\to\infty} \frac{1}{L^\gamma} \log \mathbb{P}\left(X_{T_{\ell,b,L}} \cdot \ell < 0\right) < 0 \,. \tag{6}$$

P satisfies condition T' relative to ℓ if it satisfies condition T_γ relative to ℓ for all $\gamma \in (0, 1)$. It satisfies condition T relative to ℓ if it satisfies condition T_1 relative to ℓ.

In words, condition T relative to ℓ holds if the exit from a slab that is contained between two hyperplanes perpendicular to ℓ, located respectively at distance $+L$ in the ℓ direction and $-bL$ in the opposite direction, occurs through the "backward" direction with probability that is exponentially small in L. Condition T' relaxes the exponential decay to "almost" exponential decay (there is an alternative description of condition T' in terms of regeneration distances, see Proposition 12 below). The power of condition T' is the following.

Theorem 11 (Sznitman) *Assume P is i.i.d. and uniformly elliptic, and that condition T' relative to some direction ℓ holds. Then, the process (X_n) is ballistic, i. e. $X_n/n \to v \neq 0$ for some deterministic v with $v \cdot \ell > 0$, and there is a deterministic $\sigma^2 > 0$ such that, under the annealed measure $\mathbb{P}$, $(X_n - nv)/\sigma\sqrt{n}$ converges in distribution to a standard Gaussian random variable.*

(The convergence in distribution in Theorem 11 actually extends to an invariance principle; The results in [7] yield a quenched CLT as soon as $d \geq 4$.) The key to the usefulness of Condition T' is in the following result from [99], where τ_1 denotes the first regeneration time.

Proposition 12 *Let $d \geq 2$, $\ell \in S^{d-1}$, and $\gamma \in (0, 1]$. The following are equivalent:*

(a) Condition T_γ holds.
(b) $\mathbb{P}(A_\ell^+) = 1$ and, with $X^* := \sup_{0\leq n\leq\tau_1} |X_n|$, there exists a $c > 0$ such that

$$\mathbb{E}\left(\exp(c(X^*)^\gamma)\right) < \infty \,.$$

The proof is detailed in [99], see also the exposition in [101]. From part (b) of Proposition 12, tail estimates on τ_1 follow. We omit further details.

It can be checked (by a martingale argument) that condition T (and hence T') holds for a certain direction ℓ when the environment is non-nestling. The first example of a nestling environment that satisfies condition T was provided by Sznitman, who showed that the class of environments satisfying the *Kalikow condition* from [51] also satisfy condition T (Kalikow himself had showed that his condition implies the 0–1 law, and it was shown in [102] to imply ballistic behavior). However, the verification of condition T' seems a-priori not obvious. It is thus extraordinary that an effective criterion for checking it exists, see [98]. This criterion is used in [99] to construct an example of a ballistic RWRE that does not satisfy Kalikow's condition but does satisfy T', relative to some ℓ.

Sznitman's Conjecture As noted in Proposition 12, condition T' is equivalent to certain exponential moments on the maximal distance from the origin the RWRE has achieved before time τ_1. In [99], Sznitman actually proves that condition T_γ relative to ℓ with any $\gamma \in (1/2, 1)$ implies condition T' relative to the same ℓ. This led him to the following conjecture, see [98]:

Conjecture 13 (Sznitman) *Assume P is uniformly elliptic and i.i.d. Then, condition T relative to ℓ is implied by condition T_γ relative to ℓ for any $\gamma \in (0, 1)$.*

It is also reasonable to expect ("plausible", in the language of [98]) that in addition, ballistic behavior with speed v implies condition T relative to $\ell = v/|v|$, for $d > 1$.

For $d = 1$, and i.i.d. environment, all the conditions T_γ with respect to the direction $\ell = 1$ are equivalent to $E \log \rho_0 < 0$, see [96]. Hence, Conjecture 13 holds when $d = 1$ (note that this is *not* the case for the conclusions concerning ballistic behavior, which do not hold true for $d = 1$).

In the ballistic situation, some information on the environment viewed from the point of view of the particle can be deduced. We refer to [12] and [79] for details.

Large Deviations, Quenched and Annealed

In dimension $d = 1$, the large deviations for the sequence X_n/n were obtained by considering hitting times. While this approach can be partially extended to obtain quenched LDP's for some RWRE's, see [109], its scope is limited, and in particular it does not apply to all i.i.d. measures P, nor to an annealed LDP.

A different approach was taken by Varadhan [106], who obtained the following.

Theorem 14 *Assume $d \geq 2$.*

(a) Assume P is a uniformly elliptic, ergodic measure. Then, for P-a.e. environment ω, the sequence of variables X_n/n under P^0_ω satisfies the (quenched) LDP (on $[-1,1]^d$) with speed n and deterministic, convex rate function I.

(b) Assume further that P is i.i.d. Then, the sequence of random variables X_n/n satisfies, under $\mathbb{P}$, the (annealed) LDP with speed n and convex rate function $\mathcal{I}$.

(c) The rate functions I and $\mathcal{I}$ possess the same zero set. Further, this (convex) set is either a single point or a segment of a line.

An alternative description of the quenched rate function, that is more instructive than the sub-additivity argument, has been developed for the related model of diffusions in random environments in [54]. Part (b) of Theorem 14 was extended to certain mixing environments in [80].

As for dimension $d = 1$, both I and $\mathcal{I}$ are in general not strictly convex. The quenched statement is an application of the ergodic subadditive theorem [64]. The annealed LDP is obtained by noting that the process of histories of the walk is a Markov chain, and applying the general large deviations theory for such chains.

Remark 15 In the multi-dimensional case, a formula like (4), with its intuitive description of the way an annealed deviation is obtained, is not available, since the modification of big chunks of the environment has probability which decays exponentially in volume order, i. e. n^d, instead of n.

As for $d = 1$, it is natural to study slowdown estimates in the region where the rate functions vanish, and in particular to study the probability of slowdown. This study is closely related to the analysis of Condition T', and we refer to [96,97] for details.

Non-ballistic Results

The analysis of RWRE for environments that do not exhibit ballistic behavior is still limited. Still, two important classes of models have been identified, for which the analysis could be carried out. We sketch those below.

Balanced Environment Balanced environments satisfy the constraint $\omega(0, e_i) = \omega(0, -e_i)$ for all i, in which case the local drift vanishes everywhere. In that case, X_n itself is a martingale with bounded increments, and thus $X_n/n \to 0$, $\mathbb{P}$-a.s. In fact, an invariance principle also holds.

Theorem 16 *Assume P is stationary and ergodic, balanced, and uniformly elliptic Then $X_n/n \to 0$, $\mathbb{P}$-a.s., and there exists a deterministic $\sigma^2 > 0$ such that $X_n/\sigma\sqrt{n}$ converges in distribution (under the annealed measure $\mathbb{P}$) to a Gaussian random variables. Further, X_n is recurrent if $d = 2$ and transient if $d \geq 3$.*

Theorem 16 is essentially due to [59] (the recurrence statement is due to Kesten, and can be found in [107]). It is one of the few instances where "classical" homogenization can be applied to the study of multi-dimensional RWRE.

RWRE with Deterministic Components A key to the analysis of the ballistic case is the existence of certain regeneration times. Those were used to create an i.i.d. sequence under the measure $\mathbb{P}$.

In the non-ballistic case, regeneration times as defined above do not exist. However, if the dimension of the space is large enough and some of the components are deterministic, an alternative to regeneration times can be found, based on *cut times* for simple random walk. For a simple random walk $\{S_n\}$, a cut time k is such that $\{S_n, n \leq k\} \cap \{S_n, n > k\} = \emptyset$. Such cut times exist for $d \geq 4$ [40]. By considering the cut times induced by the components of the RWRE evolving in the deterministic directions, [14] proved the following.

Theorem 17 *Assume $d = d_1 + d_2$ with $d_1 \geq 5$. Assume P is a uniformly elliptic i.i.d. measure, with $\omega(x, x + e) = q(e)$ for $e = \pm e_i$, $i = 1, \ldots, d_1$ and a deterministic q. Then, there exists a deterministic constant v such that $X_n/n \to v$, $\mathbb{P}$-a.s.. Further, if $d_1 \geq 13$, then the quenched CLT holds, i. e. there exists a deterministic $\sigma^2 > 0$ such that,*

for P-almost every ω, $(X_n - nv)/\sigma\sqrt{n}$ converges in distribution, under P^0_ω to a standard Gaussian variable.

Remark 18

(a) The convergence in distribution in Theorem 17 extends to a full invariance principle.
(b) An amusing consequence of Theorem 17 is that, for $d > 5$, one may construct P i.i.d. and uniformly elliptic such that $E(d_0 \cdot \ell) < 0$ but the resulting RWRE is ballistic with $v \cdot \ell > 0$. Recall that this is impossible in dimension $d = 1$, see Remark 2b). Also, for $d > 6$, one may construct for every $\epsilon > 0$ a P i.i.d. and uniformly elliptic such that $|\omega(x, x + e) - 1/2d| < \epsilon$, $E(d_0) \neq 0$, but $X_n/n \to 0$, $\mathbb{P}$-a.s., or such that $E(d_0) = 0$ but the walk is ballistic. We refer to [14] for the construction.

Multi Dimensional RWRE – the Perturbative Regime

We discuss in this section the perturbative analysis of the RWRE. By P being a small perturbation from a kernel q we mean that $q(\pm e_i) \geq 0$, $\sum_i [q(e_i) + q(-e_i)] = 1$, and for some ϵ small, $|\omega(x, x + e) - q(e)| < \epsilon$ for $e \in \{\pm e_i\}$. When $q(e) = 1/2d$ for $e = \pm e_i$, we say that P is a small perturbation from simple random walk.

We already observed, see Remark 18, that in the perturbative regime for simple random walk, the RWRE can exhibit behavior which is very different from the behavior of simple random walk.

Ballistic Walks

Sznitman's criterion [99] for condition T' to hold together with an renormalization analysis allowed him to give in [99] sufficient conditions for ballistic behavior when ϵ is small. Set $\rho_0(3) = 5/2$ and $\rho_0(d) = 3$ for $d \geq 4$.

Theorem 17 *Let $d \geq 3$ and $\rho < \rho_0(d)$. Then there exists an $\epsilon_0 = \epsilon_0(d, \rho) > 0$ such that if P is i.i.d. and an ϵ perturbation from simple random walk, and $Ed_0 \cdot e_1 > \epsilon^\rho$, then the T' condition relative to e_1 holds.*

Contrasting Theorem 19 with the examples in Remark 18 shows that some condition on the strength of the averaged drift Ed_0 as function of ϵ is necessary for ballistic behavior. Also, $\rho_0(d) > 2$ is used in constructing the examples in [99] that show that Kalikow's condition is strictly included in condition T'. We note that the case $d = 2$ is still open.

In another direction, if one writes $\omega(x, x + e) = q(e) + \epsilon\xi(x, x + e)$ with ξ i.i.d., and either $\sum eq(e) \neq 0$ or $\sum eq(e) = 0$ but $\sum eE\xi(0, e) \neq 0$, then for ϵ small enough, Kalikow's condition holds. Expansions in ϵ of the speed of the RWRE are provided in [83].

Balanced Walks

Recall the balanced walks introduced in Subsect. "Non-Ballistic Results", c.f. Theorem 16. The existence of an invariance measure viewed from the point of view of the particle, and the control achieved on this measure by approximations with periodized environments, allow one to get an expansion of the diffusivity matrix in terms of the strength of the perturbation from simple random walk. We refer the reader to [60] for details.

Isotropic RWRE

The existence of sub-diffusive behavior for the RWRE model in $d = 1$ immediately raises the question as to whether such sub-diffusive behavior is present in higher dimension. As implied by Theorem 11, this is not the case when the environment satisfies condition T'. Since it may be expected that condition T' characterizes ballistic behavior for $d > 1$, it is reasonable to expect (but not proved!) that for P i.i.d. and uniformly elliptic, and $d > 1$, no sub-diffusive behavior is possible when the walk is transient in direction ℓ (and further, in the ballistic regime, when re-centering around the limiting velocity v, one expects fluctuations in the diffusive scale).

Outside the ballistic regime, rigorous results are few. Early attempts to address the question of existence of a diffusive regime appeared in [35,41], using a formal renormalization group analysis in the small perturbation regime, with the conclusion that no sub-diffusive behavior exists at $d \geq 3$ in the perturbative regime, and that at most logarithmic corrections to diffusive behavior exist at $d = 2$. While this conclusion certainly conforms with what one would expect, soon after it was pointed out that counter-examples can be constructed (albeit not with i.i.d., or even finite range dependent, environments), see [15,17,18]. Further, some of the examples discussed in this article, and in particular those of Subsect. "Non-Ballistic Results", see Remark 18, do not seem to be consistent with the formal renormalization analysis.

An attempt to put the analysis on a rigorous foundation was made in [22]. Among other things, they introduced the following isotropy condition:

Definition 20 The law P on the environment is isotropic if, for any rotation matrix $\mathcal{O}$ acting on $\mathbb{R}^d$ that fixes $\mathbb{Z}^d$, the laws of $(\omega(0, \mathcal{O}e))_{e:|e|=1}$ and $(\omega(0, e))_{e:|e|=1}$ coincide.

In particular, if P is isotropic then $Ed_0 = 0$. The main result of [22] is the following:

Theorem 21 (Bricmont–Kupiainen) *Assume $d \geq 3$. There exists an $\epsilon_0 = \epsilon_0(d)$ such that if P is i.i.d. and*

isotropic, and an ϵ perturbation of simple random walk with $\epsilon < \epsilon_0$, then for some deterministic $\sigma^2 > 0$ and for P almost every ω, the sequence $X_n/\sigma\sqrt{n}$ converges in distribution, under P^0_ω, to a standard Gaussian random variable.

The approach of [22] is to introduce a (diffusive) rescaling in time and space, and propagate an estimate on both the large scale behavior of the RWRE, as well as about the existence of local traps that have the potential to destroy, at the next level, the diffusivity properties. The restriction to $d \geq 3$ is useful because the underlying simple random walk for $d \geq 3$ is transient, and hence Green function computations can be performed.

Several attempts have recently been made to provide an rescaling argument (alternative to [22]) that is more transparent. The first approach [103], which is closest to Theorem 21, has been in the context of diffusions in random environments. In the remainder of this section, we describe another approach, due to [13], that yields a result concerning the exit measure of (isotropic) RWRE from large balls.

Let $V_L = \{x \in \mathbb{Z}^d : |x| \leq L\}$ be the ball of radius L in $\mathbb{Z}^d$ (where we recall that $|\cdot|$ is the euclidean norm), and let $\partial V_L = \{y \in \mathbb{Z}^d : d(y, V_L) = 1\}$ denote the boundary of V_L. Let $\tau_L = \min\{n : X_n \notin V_L\}$ denote the exit time of the RWRE from V_L, and for $x \in V_L, z \in \partial V_L$, let $\Pi_L(x,z) = P^x_\omega(X_{\tau_L} = z)$ denote the exit measure of the RWRE from V_L, and let $\pi_L(x,z)$ denote the corresponding quantity for simple random walk. Finally, let $\Pi^s_{L,l}(x,z) = \Pi_L \star \pi_{\eta l}$, where $\star$ denotes convolution and η is a random variable with smooth density supported on $(1,2)$. $\Pi^s_{L,l}$ is a smoothed version of Π_L, where the smoothing is at scale l.

One expects that for an isotropic environment that is a small perturbation of simple random walk, the exit measure Π_L approaches that of simple random walk, except for small nonvanishing correction that are due to localized perturbations near the boundary, and that as soon as some additional smoothing is applied, convergence occurs. Under the assumptions of Theorem 21, this is indeed the case. In what follows, for probability measures μ, ν we write $\|\mu - \nu\|$ for the variational distance between μ and ν.

Theorem 22 *Assume $d \geq 3$. There exists a $\delta_0 = \delta_0(d) > 0$ with the following property: for each $\delta < \delta_0$ there exists an $\epsilon_0 = \epsilon_0(d, \delta)$ such that if $\epsilon < \epsilon_0$ and P is an i.i.d. and isotropic law which is an ϵ perturbation of simple random walk, then*

$$\limsup_{L\to\infty} \|\Pi_L(0,\cdot) - \pi_L(0,\cdot)\| \leq \delta\,. \tag{7}$$

Further,

$$\limsup_{L\to\infty} \|\Pi^s_{L,l}(0,\cdot) - \pi_L \star \pi_{\eta l}(0,\cdot)\| \leq c_l \to_{l\to\infty} 0\,. \tag{8}$$

Diffusions in Random Environments

The model of RWRE possesses a natural analogue in the setup of diffusion processes.

One Dimensional Generators

For dimension $d = 1$, the study of analogues of the RWRE model goes back to [23] and [86]. Formally, one looks at solutions to the stochastic differential equation

$$\mathrm{d}X_t = -\frac{1}{2}V'(X_t)\mathrm{d}t + \mathrm{d}\beta_t\,, \quad X_0 = 0\,, \tag{9}$$

where β is a standard Brownian motion and V, the potential, is itself an (independent of β) Brownian motion with constant drift. Of course, (9) does not make sense as written, but one can express the solution to (9) for smooth V in a way that makes sense also when V is replaced by Brownian motion, by saying that conditioned on the environment V, X_t is a diffusion with generator

$$\frac{1}{2}\mathrm{e}^{V(x)}\frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{-V(x)}\frac{\mathrm{d}}{\mathrm{d}x}\right)\,. \tag{10}$$

The diffusion in (9) inherits many of the asymptotic properties of the RWRE model. Additional tools, borrowed from stochastic calculus, are often needed to obtain sharp statements. We refer to [89] for details and additional references.

Multi Dimensional Diffusions: Finite Range Dependence

Like the RWRE in dimension $d = 1$, the model (9) leads to a reversible diffusion. A direct generalization of (9) via the expression (10) for the generator, see for example [68,69], preserves the reversibility of the process, and thus for our purpose does not serve as a true analogue of the RWRE model. Instead, we consider diffusions satisfying the equation in $\mathbb{R}^d$:

$$\mathrm{d}X_t = b(X_t,\omega)\mathrm{d}t + \sigma(X_t,\omega)\mathrm{d}W_t\,, \quad X_0 = 0\,, \tag{11}$$

with generator

$$L = \frac{1}{2}\sum_{i,j=1}^d a_{ij}(x,\omega)\,\partial^2_{ij} + \sum_{i=1}^d b_i(x,\omega)\,\partial_i\,, \tag{12}$$

where $a = \sigma\sigma^T$ is a d-by-d matrix and the coefficients a, b are assumed to satisfy the following:

Assumption 23

(a) *The functions $a(\cdot,\omega)$ and $b(\cdot,\omega)$ are uniformly (in ω) bounded by K, with Lipschitz norm bounded by K, and a is uniformly elliptic, i. e. $a(x,\omega) - \kappa I$ is positive definite for some $\kappa > 0$ independent of x or ω.*

(b) *The random field $\big(a(x,\omega), b(x,\omega)\big)_{x\in\mathbb{R}^d}$ is stationary with respect to shifts in $\mathbb{R}^d$.*

(c) *The collection of random variables $\big(a(x,\cdot), b(x,\cdot)\big)_{x\in A}$ and $\big(a(y,\cdot), b(y,\cdot)\big)_{y\in B}$ are independent when $d(A,B) > R$.*

Part (a) of Assumption 23 ensures that (11) possesses a unique strong solution. Part (c) of Assumption 23 is a "finite range dependence" condition. We continue to write P_ω for the quenched law of the trajectories of the diffusion.

Many of the results described in Subsect. "Ergodic Properties and a 0–1 Law" and Subsect. "Ballistic Behavior and Sznitman's Conditions" have been proved also in the context of diffusions, when Assumption 23 holds. We refer to [45,84,85,88] for details.

Isotropic Diffusions in the Perturbative Regime

The analogue of the isotropy condition Assumption 23 (b) in the diffusion context is the following:

Assumption 24 (Isotropy) *For any rotation matrix $\mathcal{O}$ preserving the union of coordinate axes of $\mathbb{R}^d$,*

$$\big(a(\mathcal{O}x,\omega), b(\mathcal{O}x,\omega)\big)_{x\in\mathbb{R}^d} \quad \text{has same law under } P \text{ as} \quad \big(\mathcal{O}a(x,\omega)\mathcal{O}^T, \mathcal{O}b(x,\omega)\big)_{x\in\mathbb{R}^d}.$$

The analogue of Theorem 21 is the following theorem, due to [103]. Its proof again uses multi-scale arguments, and is based on controlling the (scaled) Hölder norm of the operator associated with the transition probability of the diffusion.

Theorem 25 *Let Assumptions 23 and 24 hold. Then, there exists a constant $\epsilon_0 = \epsilon_0(d,K,R)$ such that if $|a(x,\omega) - I| \le \epsilon_0$ and $|b(x,\omega)| \le \epsilon_0$, for all $x \in \mathbb{R}^d, \omega \in \Omega$, then for some deterministic $\sigma^2 > 0$, for a.e. ω, the sequence of random variables $X_t/\sigma\sqrt{t}$ converges in distribution to a standard Gaussian random variable.*

(A full quenched invariance principle also holds under the assumptions of Theorem 25.)

Topics Left Out and Future Directions

As pointed out in the text, many open problems remain in the study of RWRE's and motion in random media, with the most important one being the general validity of the law of large numbers, and the existence of diffusive limits for dimensions $d \ge 2$. We briefly mention in the rest of this section several topics that are related to this review but that we have not covered in details.

Random Conductance Model

We have concentrated in this review on RWRE's in i.i.d. environments, which give rise in the multi-dimensional case to non-reversible Markov processes. Although mentioned in several places, we did not discuss in details the reversible case, where homogenization techniques using the environment viewed from the point of view of the particle are very efficient (note that the reversible case is a very particular case of an environment which is not i.i.d. but rather dependent with finite range dependence). The prototype for such reversible models is the "random conductance model", where each edge (x,y) of $\mathbb{Z}^d$ is associated a (random, i.i.d.) conductance $C_{x,y}$, and the transition probability between x and y is $C_{x,y}/(\sum_{z:\,|z-x|=1} C_{x,z})$. Annealed CLT's for the random conductance model are provided in [57,66]. See also [2] for a related model with symmetric transitions. The quenched CLT is obtained in [16] and [90].

One of the motivations to consider the random conductance model is the analysis of random walk on supercritical percolation clusters. The annealed CLT is covered by [66]. Several recent papers discuss the quenched case, first in dimension $d \ge 4$ [90], and then in all dimensions, see [5,70]. In another direction, when one discusses *biased* walks on a percolation cluster, new phenomena occur, for example the lack of monotonicity of the speed of the walk in the strength of the bias, which is again a manifestation of the trapping phenomenon. We refer to [6] and [100] for details.

Brownian Motion in a Field of Random Obstacles

Another closely related (reversible) model is the model of Brownian motion in a field of obstacles in $\mathbb{R}^d$. Here, one defines a potential $V(x,\omega) = \sum_i W(x - x_i)$ where the collection $\{x_i\}$ is a (random) configuration of points in $\mathbb{R}^d$ (usually, taken according to a Poisson law) and W is a fixed nonnegative shape function. Of interest are the properties of Brownian motion $(X_t)_{t\in[0,T]}$, perturbed by the change of measure

$$\Lambda_T = \frac{1}{Z_T(\omega)} \exp\left(-\int_0^T V(X_s,\omega)\mathrm{d}s\right).$$

It is common to distinguish between "soft traps", with W bounded and typically of compact support, and "hard trap", where $W = \infty \mathbf{1}_C$ where C is a given compact set. One is interested in understanding various path properties, as T gets large, or in understanding the quenched partition function $Z_T(\omega)$ and its annealed counterpart EZ_T. Due to reversibility, the problem is closely related to the study of the bottom λ_ω of the spectrum of $-\Delta/2 + V$, and the difficulty is in understanding the structure of those traps that influence λ_ω. A good overview of the model and the techniques developed to analyze it, including the "method of enlargement of obstacles", can be found in [95].

Time Dependent RWRE

An interesting variant of the RWRE model has been proposed in [8]. In this model, the random environment is dynamic, i. e. changes with time, and so we write $\omega(x, x+e, n)$ where we wrote before $\omega(x, x+e)$. In the simplest version, the collection of random vectors $(\omega(x, x+\cdot, n))_{x\in\mathbb{Z}^d, n\in\mathbb{N}}$ is i.i.d. Annealed, the RWRE is then a simple random walk in an averaged environment, but the true interest lies in obtaining quenched statements. Those were obtained in [8,10] by a perturbative approach. An alternative, simpler proof is given by [94]. Another approach to the quenched CLT, that covers other cases of random walk "with a forbidden direction", is developed in [81], based on a general pointwise CLT for additive functionals of Markov chains due to [34].

An interpolation between the RWRE model and the i.i.d. dynamical environment model is when the collection $(\omega(x, x+\cdot, n))_{x\in\mathbb{Z}^d, n\in\mathbb{N}}$ is i.i.d. in x but Markovian in n. This case has been analyzed by perturbative methods in [9], and by regeneration techniques in [3]. In both cases, an annealed CLT holds in any dimension, but the quenched CLT was obtained only in high dimension. Recently, a dynamical approach was developed in [36], that proves the quenched CLT in all dimensions, subject to fast enough mixing of the Markov chain. It is still open to determine whether in the Markovian setup, there are uniformly elliptic, exponentially mixing examples where the quenched CLT fails.

RWRE on Trees and Other Graphs

We have already mentioned the interest in considering random walks on random subgraphs of $\mathbb{Z}^d$, and in particular percolation clusters. Of course, one may consider instead random walk (or biased random walk) on other random graphs. A particularly important class of models treats random walks on random trees, and in particular Galton–Watson trees. We refer to [65] for an excellent overview of the properties and ergodic theory of such random walks, and to [74] for recent results concerning the CLT. See also [50] for slowdown estimates for the analog of the RWRE on the binary tree. We emphasize that these models are all reversible.

Non Nearest Neighbor RWRE

Many of the techniques described in this survey have a natural generalization to non nearest neighbor walks. In particular, the results in [80,106] are already stated in terms of compactly supported transition probabilities, and the development of regeneration times can easily be extended, following the techniques in [26,27], to the non-nearest neighbor, finite range setup. However, to the best of my knowledge, no systematic study of RWRE for non-nearest neighbor RWRE's in dimension $d \geq 2$ has appeared in the literature.

The situation is different in dimension $d = 1$, where the RWRE is not reversible anymore. It was early realized, see [53,63], that ergodic theorems involve the study of certain Lyapounov exponents associated with the product of random matrices. For some recent results, we refer to [11] and [21].

Bibliography

Primary Literature

1. Alili S (1999) Asymptotic behaviour for random walks in random environments. J Appl Probab 36:334–349
2. Anshelevich VV, Khanin KM, Sinai YG (1982) Symmetric random walks in random environments. Commun Math Phys 85:449–470
3. Bandyopadhyay A, Zeitouni O (2006) Random Walk in Dynamic Markovian Random Environment. ALEA 1:205–224
4. Berger N (2006) On the limiting velocity of high-dimensional random walk in random environment. Arxiv: math.PR/0601656
5. Berger N, Biskup M (2007) Quenched invariance principles for simple random walk on percolation clusters. Probab Theory Relat Fields 137:83–120
6. Berger N, Gantert N, Peres Y (2003) The speed of biased random walk on percolation clusters. Probab Theory Relat Fields 126:221–242
7. Berger N, Zeitouni O (2008) A quenched invariance principle for certain ballistic random walks in i.i.d. environments. In: Siduravicius V, Vares ME (eds) In and out of equilibrium, Progress in probability, vol. pp 137–160
8. Boldrighini C, Minlos RA, Pellegrinotti A (1997) Almost-sure central limit theorem for a Markov model of random walk in dynamical random environment. Probab Theory Relat Fields 109:245–273
9. Boldrighini C, Minlos RA, Pellegrinotti A (2000) Random walk in a fluctuating random environment with Markov evolution.

In: On Dobrushin's way. From probability theory to statistical physics. Amer Math Soc Transl Ser 198(2):13–35; Amer Math Soc, Providence
10. Boldrighini C, Minlos RA, Pellegrinotti A (2004) Random walks in quenched i.i.d. space-time random environment are always a.s. diffusive. Probab Theory Relat Fields 129:133–156
11. Bolthausen E, Goldsheid I (2000) Recurrence and transience of random walks in random environments on a strip. Commun Math Phys 214:429–447
12. Bolthausen E, Sznitman AS (2002) On the static and dynamic points of view for certain random walks in random environment. Method Appl Analysis 9:345–375
13. Bolthausen E, Zeitouni O (2007) Multiscale analysis of exit distributions for random walks in random environments. Prob Theor Fields 138:581–645
14. Bolthausen E, Sznitman AS, Zeitouni O (2003) Cut points and diffusive random walks in random environments. Ann Inst H Poincare 39:527–555
15. Bouchaud JP, Georges A, Le Doussal P (1987) Anomalous diffusion in random media: trapping, correlations and central limit theorems. J Phys 48:1855–1860
16. Boivin D, Depauw J (2003) Spectral homogenization of reversible random walks on $\mathbb{Z}^d$ in a random environment. Stoch Proc App 104:29–56
17. Bramson M (1991) Random walk in random environment: a counterexample without potential. J Stat Phys 62:863–875
18. Bramson M, Durrett R (1988) Random walk in random environment: a counterexample? Commun Math Phys 119:199–211
19. Bramson M, Zeitouni O, Zerner MPW (2006) Shortest spanning trees and a counterexample for random walks in random environments. Ann Probab 34:821–856
20. Brémont J (2004) Behavior of random walks on $\mathbb{Z}$ in Gibbsian medium. C R Math Acad Sci Paris 338:895–898
21. Brémont J (2004) Random walks in random medium on $\mathbb{Z}$ and Lyapunov spectrum. Ann Inst H Poincar Probab Stat 40:309–336
22. Bricmont J, Kupiainen A (1991) Random walks in asymmetric random environments. Com mun Math Phys 142:345–420
23. Brox T (1986) A one-dimensional diffusion process in a Wiener medium. Ann Probab 14:1206–1218
24. Cheliotis D (2005) Diffusions in random environments and the renewal theorem. Ann Probab 33:1760–1781
25. Comets F, Gantert N, Zeitouni O (2000) Quenched, annealed and functional large deviations for one dimensional random walk in random environment. Probab Theory Relat Fields 118:65–114
26. Comets F, Zeitouni O (2004) A law of large numbers for random walks in random mixing environments. Ann Probab 32:880–914
27. Comets F, Zeitouni O (2005) Gaussian fluctuations for random walks in random mixing environments. Isr J Math 148:87–114
28. Dembo A, Gantert N, Peres Y, Shi Z (2007) Valleys and the maximum local time for random walk in random environment. Probab Theory Relat Fields 137:443–473
29. Dembo A, Zeitouni O (1998) Large deviations techniques and applications. 2nd edn. Springer, New York
30. Dembo A, Guionnet A, Zeitouni O (2001) Aging properties of Sinai's random walk in random environment. Arxiv:math.PR/0105215
31. Dembo A, Gantert N, Peres Y, Zeitouni O (2002) Large deviations for random walks on Galton–Watson trees: averaging and uncertainty. Probab Theory Relat Fields 122:241–288
32. Dembo A, Gantert N, Zeitouni O (2004) Large deviations for random walk in random environment with holding times. Ann Probab 32:996–1029
33. Dembo A, Peres Y, Zeitouni O (1996) Tail estimates for one-dimensional random walk in random environment. Commun Math Physics 181:667–684
34. Derriennic Y, Lin M (2003) The central limit theorem for Markov chains started at a point. Probab Theory Relat Fields 125:73–76
35. Derrida B, Luck JM (1983) Diffusion on a random lattice: weak-disorder expansion in arbitrary dimension. Phys Rev B 28:7183–7190
36. Dolgopyat D, Keller G, Liverani C (2007) Random walk in markovian environment. ArXiv:math/0702100v1 [math.PR] (preprint)
37. Donsker MD, Varadhan SRS (1983) Asymptotic evaluation of certain Markov process expectations for large time, IV. Commun Pure Appl Math 36:183–212
38. Doyle PG, Snell JL (1984) Random walks and electric networks, Carus Mathematical Monographs, 22. Mathematical Association of America, Washington
39. Enriquez N, Sabot C, Zindy O (2007) Limit laws for transient random walks in random environments on $\mathbb{Z}$. ArXiv:math/0703660v1 [math.PR] (preprint)
40. Erdös P, Taylor SJ (1960) Some intersection properties of random walks paths. Acta Math Acad Sci Hungar 11:231–248
41. Fisher DS (1984) Random walks in random environments. Phys Rev A 30:60–964
42. Gantert N (2002) Subexponential tail asymptotics for a random walk with randomly placed one-way nodes. Ann Inst H Poincaré – Probab Statist 38:1–16
43. Gantert N, Shi Z (2002) Many visits to a single site by a transient random walk in random environment. Stoch Process Appl 99:159–176
44. Gantert N, Zeitouni O (1998) Quenched sub-exponential tail estimates for one-dimensional random walk in random environment. Commun Math Physics 194:177–190
45. Goergen L (2006) Limit velocity and zero-one laws for diffusions in random environment. Ann Appl Probab 16:1086–1123
46. Goldsheid I (2007) Simple transient random walks in one-dimensional random environment: the central limit theorem. Probab Theory Related Fields 139:41–64
47. Golosov AO (1985) On limiting distributions for a random walk in a critical one dimensional random environment. Commun Moscow Math Soc 199:199–200
48. Greven A, den Hollander F (1994) Large deviations for a random walk in random environment. Ann Probab 22:1381–1428
49. Hu Y, Shi Z (2000) The problem of the most visited site in random environment. Probab Theory Relat Fields 116:273–302
50. Hu Y, Shi Z (2007) A subdiffusive behaviour of recurrent random walk in random environment on a regular tree. Probab Theory Related Fields 138:521–549
51. Kalikow SA (1981) Generalized random walks in random environment. Ann Probab 9:753–768
52. Kesten H (1986) The limit distribution of Sinai's random walk in random environment. Physica A 138:299–309

53. Key ES (1984) Recurrence and transience criteria for random walk in a random environment. Ann Probab 12:529–560
54. Kosygina E, Rezakhanlou F, Varadhani RS (2006) Stochastic homogenization of Hamilton-Jacobi-Bellman Equations. Comm Pure Appl Math 59:1489–1521
55. Kesten H, Kozlov MV, Spitzer F (1975) A limit law for random walk in a random environment. Comput Math 30:145–168
56. Kozlov SM (1985) The method of averaging and walks in inhomogeneous environments. Russian Math Surv 40:73–145
57. Kunnemann R (1983) The diffusion limit of reversible jump processes in $\mathbb{Z}^d$ with ergodic random bond conductivities. Commun Math Phys 90:27–68
58. Kuo HJ, Trudinger NS (1990) Linear elliptic difference inequalities with random coefficients. Math Comput 55:37–53
59. Lawler GF (1982) Weak convergence of a random walk in a random environment. Commun Math Phys 87:81–87
60. Lawler GF (1989) Low-density expansion for a two-state random walk in a random environment. J Math Phys 30:145–157
61. Lawler GF (1991) Intersections of Random walks. Birkhauser, Basel
62. Le Doussal P, Monthus C, Fisher D (1999) Random walkers in one-dimensional random environment: exact renormalization group analysis. Phys Rev E 59:4795–4840
63. Ledrappier F (1984) Quelques propriétés des exposants charactéristiques, Lecture Notes in Mathematics, vol 1097. Springer, New York
64. Liggett TM (1985) An improved subadditive ergodic theorem. Ann Probab 13:1279–1285
65. Lyons R, Peres Y Probability on trees and networks. http://mypage.iu.edu/~rdlyons/prbtree/prbtree.html
66. De Masi A, Ferrari PA, Goldstein S, Wick WD (1989) An invariance principle for reversible Markov processes. Applications to random motions in random environments. J Stat Phys 55:787–855
67. Mayer-Wolf E, Roitershtein A, Zeitouni O (2004) Limit theorems for one-dimensional transient random walks in Markov environments. Ann Inst H Poincaré Probab Stat 40:635–659
68. Mathieu P (1994) Zero white noise limit through Dirichlet forms, with applications to diffusions in a random medium. Probab Theory Relat Fields 99:549–580
69. Mathieu P (1995) Limit theorems for diffusions with a random potential. Stoch Proc App 60:103–111
70. Mathieu P, Piatnitski A (2007) Quenched invariance principles for random walks on percolation clusters. Proc R Soc Lond Ser A Math Phys Eng Sci 463:2287–2307
71. Molchanov SA (1994) Lectures on random media. Lecture Notes in Mathematics, vol 1581. Springer, New York
72. Papanicolaou GC, Varadhan SRS (1982) Diffusions with random coefficients. In: Kallianpur G, Krishnaiah PR, Ghosh JK (eds) Statistics and probability: essays in honor of C. R. Rao. North-Holland, Amsterdam, pp 547–552
73. Pemantle R (1988) Phase transition in reinforced random walk and RWRE on trees. Ann Probab 16:1229–1241
74. Peres Y, Zeitouni O (2008) A Quenched CLT for biased random walks on Golton–Watson trees, Prob Theory Rel Fieldsx 140:595–629
75. Peterson J (2008) Ph.D. thesis. University of Minnesota
76. Peterson J, Zeitouni O (2007) Quenched limits for transient, zero speed one-dimensional random walk in random environment. ArXiv:0704.1778v1 [math.PR] (preprint)
77. Pisztora A, Povel T (1999) Large deviation principle for random walk in a quenched random environment in the low speed regime. Ann Probab 27:1389–1413
78. Pisztora A, Povel T, Zeitouni O (1999) Precise large deviation estimates for a one-dimensional random walk in a random environment. Probab Theory Relat Fields 113:191–219
79. Rassoul-Agha F (2003) The point of view of the particle on the law of large numbers for random walks in a mixing random environment. Ann Probab 31:1441–1463
80. Rassoul-Agha F (2004) Large deviations for random walks in a mixing random environment and other (non-Markov) random walks. Commun Pure Appl Math 57:1178–1196
81. Rassoul-Agha F, Seppäläinen T (2005) An almost sure invariance principle for random walks in a space-time random environment. Probab Theory Relat Fields 133:299–314
82. Rassoul-Agha F, Seppäläinen T (2007) Almost sure functional central limit theorem for non-nestling random walk in random environment. ArXiv:0704.1022v1 [math.PR] (preprint)
83. Sabot C (2004) Random walks in random environment at low disorder. Ann Probab 32:2996–3023
84. Schmitz T (2006) Diffusions in random environment and ballistic behavior. Ann Inst H Poincaré Prob Statist 42:683–714
85. Schmitz T (2006) Examples of condition (T) for diffusions in random environment. Electron J Probab 11:540–562
86. Schumacher S (1985) Diffusions with random coefficients. Contemp Math 41:351–356
87. Shen L (2002) Asymptotic properties of certain anisotropic walks in random media. Ann Appl Probab 12:477–510
88. Shen L (2003) On ballistic diffusions in random environment. Ann Inst H Poincare Probab Statist 39:839–876
89. Shi Z (2001) Sinai's walk via stochastic calculus. In: Comets F, Pardoux E (eds) Milieux Aléatoires. Panoramas et Synthèses 12. Soc Math Fr, pp 53–74
90. Sidoravicius V, Sznitman AS (2004) Quenched invariance principles for walks on clusters of percolation or among random conductances. Probab Theory Relat Fields 129:219–244
91. Sinai YG (1982) The limiting behavior of a one-dimensional random walk in random environment. Theor Prob Appl 27:256–268
92. Solomon F (1975) Random walks in random environments. Ann Probab 3:1–31
93. Stroock D, Varadhan SRS (1979) Multidimensional diffusion processes. Springer, New York
94. Stannat W (2004) A remark on the CLT for a random walk in a random environment. Probab Theory Relat Fields 130:377–387
95. Sznitman AS (1998) Brownian motion, obstacles and random media. Springer, New York
96. Sznitman AS (1999) On a class of transient random walks in random environment. Ann Probab 29:724–765
97. Sznitman AS (2000) Slowdown estimates and central limit theorem for random walks in random environment. JEMS 2:93–143
98. Sznitman AS (2002) An effective criterion for ballistic behavior of random walks in random environment. Probab Theory Relat Fields 122:509–544
99. Sznitman AS (2003) On new examples of ballistic random walks in random environment. Ann Probab 31:285–322
100. Sznitman AS (2003) On the anisotropic walk on the supercritical percolation cluster. Commun Math Phys 240.123–148

101. Sznitman AS (2004) Topics in random walks in random environment. School and Conference on Probability Theory. ICTP Lect Notes XVII Abdus Salam, Int Cent Theor Phys Trieste, pp 203–266
102. Sznitman AS, Zerner M (1999) A law of large numbers for random walks in random environment. Ann Probab 27:1851–1869
103. Sznitman AS, Zeitouni O (2006) An invariance principle for isotropic diffusions in random environments. Invent Math 164:455–567
104. Temkin DE (1972) One dimensional random walk in a two-component chain. Soviet Math Dokl 13:1172–1176
105. Varadhan SRS (1966) Asymptotic probabilities and differential equations. Commun Pure Appl Math 9:261–286
106. Varadhan SRS (2003) Large deviations for random walks in a random environment. Commun Pure Appl Math 56:1222–1245
107. Zeitouni O (2004) Random walks in random environment. XXXI Summer School in Probability, St. Flour, 2001. Lect Notes Math, vol 1837. Springer, New York, pp 193–312
108. Zeitouni O (2006) Random walks and diffusions in random environments. J Phys A: Math Gen 39:R433–R464
109. Zerner MPW (1998) Lyapounov exponents and quenched large deviations for multidimensional random walk in random environment. Ann Probab 26:1446–1476
110. Zerner MPW (2002) A non-ballistic law of large numbers for random walks in i.i.d. random environment. Electron Commun Probab 7(19):191–197
111. Zerner MPW, Merkl F (2001) A zero-one law for planar random walks in random environment. Ann Probab 29:1716–1732

Books and Reviews

Bolthausen E, Sznitman AS (2002) Ten lectures on random media DMV Seminar, 32. Birkhäuser, Basel
Hughes BD (1996) Random walks and random environments. Oxford University Press, Oxford
Révész P (2005) Random walk in random and non-random environments, 2nd edn. World Scientific, Hackensack
Revuz D, Yor M (1999) Continuous martingales and Brownian motion, 3rd edn. Springer, New York
Spitzer F (1976) Principles of random walk, 2nd edn. Springer, New York
Zeitouni O (2002) Random Walks in Random Environments. Proceedings of ICM 2002. Documenta Mathematica III:117–127

Rational, Goal-Oriented Agents

Rosaria Conte
LABSS-ISTC, CNR, Rome, Italy

Article Outline

Glossary

Agent A system that uses its power to bring about a given state in a physical, social, or mental environment.

Autonomy It is a property of a subset of agents, which decide to act on the grounds of their internal criteria (internal representations).

Computing agent An entity that executes the instructions contained in an algorithm running on a computer.

Goal-directed agents are intelligent agents that have an internal representation of the goals they achieve.

Goal-oriented agents are entities designed to achieve a certain state of the world wanted by either the agent itself, which in such a case is also a goal-directed system, or the user/designer.

Intelligent agents are goal-oriented agents using their knowledge to solve problems, including taking decisions and planning actions.

Internal representation Knowledge stored into internal agent structures or distributed over sets of interconnected internal units.

Knowledge Information about the environment, which is stored in agents' memory (archives).

Rational agents are intelligent agents that use their (limited) knowledge to maximize the difference between benefits and costs.

Definition of the Subject

In this article, goal-oriented agents, i. e. agents designed to achieve a given goal, will be treated as a highly general category within agent theory and application, and argued to subsume both rational agents, which maximize the distance between benefits and costs, and goal-directed agents, which have a representation of the goal they achieve. These sub-types of agents exemplify two different but non-exclusive approaches to agent theory, endogenous and exogenous. The endogenous approach is aimed at modeling agents in terms of their internal mechanisms of regulation. The exogenous approach is used to describe them from an external point of view, in terms of the effects they achieve in the world. In this article, these two approaches will be combined with another important classification, weak and strong agents. The resulting specific models and applications will be described. Finally, goal-directed and rational agents will be compared and future directions of research in the agent field will be discussed.

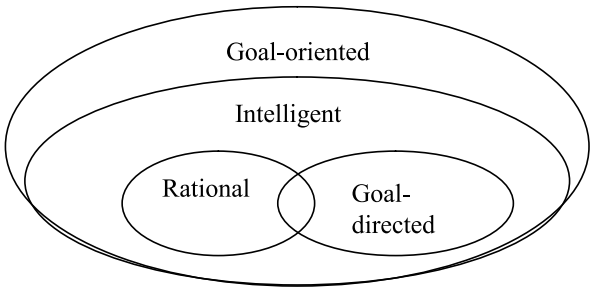
The field of agent theory and technology is growing fast. Suffice it to consider that the community of scientists interested in multi agent systems, which counted at most a few dozen people in the early nineties, is now one of the largest in the area of computational and technological sciences (The 2008 AAMAS (Autonomous Agents and Multi Agent Systems) conference received around 700 submissions.). Furthermore, a growing number of disciplines within the social and behavioral sciences, not to speak of the science of complexity, are turning to agent-based computational modeling in the aim to promote explicit, controllable models of social and behavioral phenomena, and to realize computational laboratories for visualizing such phenomena and conducting repeatable experiments about them.

Despite, or perhaps because of such a fast development, the field of agent theory and application did not receive an adequate scientific re-elaboration, nor commonly shared views and definitions. Agent is still a controversial notion, specified – as somebody observed in [81] – by a growing number of adjectives, for example, autonomous, goal-directed, rational, etc., rather than by explicit properties. Developed by different disciplines – physics, biology, AI, psychology, computer science, economy, decision theory, organization science, etc. – with their own conventions and vocabulary, the notion of agent would necessitate an explicit confrontation, which is underestimated by the different communities. The result is an unpleasant conceptual mix and consequent loss of scientific credibility of the field.

This article is therefore intended to offer a contribution of clarification concerning some properties of agency, in order to compare at least a subset of models and theories currently insufficiently interacting, and help envisage future trends.

Introduction

In the present contribution, rational agents are seen as a subset of intelligent agents (see Fig. 1), and these in turn as a subset of goal-oriented ones. Agents are said to be goal-oriented when they exhibit a finalistic behavior [3,9], i. e. when their behavior brings about a state of the world that achieves somebody's, not necessarily the user's, goals. To be noticed, goal-oriented systems need not be goal-*directed* either. The latter are guided by a representation of the goal they are oriented to [76]: a software agent designed to bring about a given effect is goal-oriented. It becomes goal-directed if, for any reason, the effect needs to be represented as an internal state of the agent.



Rational, Goal-Oriented Agents, Figure 1
Set relationships among some types of goal-oriented agents

The distinction between goal-oriented and goal-directed action is crucial, since it allows two main components of computational agents to be distinguished, the design-goals – i. e. the goals of the user/designer that need not be represented into the agent although incorporated in the way it operates and is constructed – and the internal goals, which are represented in the agent that achieves them. Of course, goal-oriented modeling works as a filter for the designer, which enables her to recognize important properties of her domain of interest. It also tells her which internal mechanisms the system ought to possess in order to achieve its design-goal, and whether and which among these goals ought to be internalized.

Within the broad category of goal-oriented agents, we identify intelligent agents, as those entities that use *their own knowledge* to bring about certain effects in the world. Intelligent agents include goal-directed agents, achieving their own goals, and rational agents, maximizing their own utility.

Rational and goal-directed agents are here taken as paradigmatic examples of two complementary approaches in the agent field, endogenous and exogenous: the latter describes agents' behavior as perceived by an external observer; the former models it in terms of the agent's internal mechanisms. Either approach will be characterized separately.

In the final section, rational and goal-oriented agents will be compared, their differences will be outlined and the respective advantages for both theory and application will be pointed out.

Agents

Agents are systems acting on some environment, including their minds. In this case, their action consists of taking decisions, manipulating information, etc. More generally, they act on the physical or social environment producing transitions from one state to another.

They are modeled in more or less formal terms. In general, agent models are abstract, formal, often computational. The latter describe agents as entities carrying out the operations described in an algorithm, accepting inputs from the environment, and performing actions that depend upon the current inputs and produce outputs that modify the environment.

By an abstract model of the agent, it is meant an ideal-type entity endowed with defining properties as specified in the underlying theory. One such entity, which may never be found in nature, facilitates general explanation of real-world phenomena involving natural agents, and provides inspiring models for concrete applications to be designed.

By a formal description, it is meant an abstract description that is also explicit, non-equivocal and sound. It is by no means necessary that such a description be translated into a standard logic language, although this is sometimes the case. Strictly speaking, a formal model of the agent is not necessarily transformed into a computational model either, although this is often the case.

In this contribution, agent *computational* models will mainly be addressed, although the discussion of rational agents will involve mathematical notions, and that of BDI agents – a special type of goal-directed systems – will bring into play *logic* definitions of mental states. In all of these cases, however, formal technicalities will not be provided, and the models will only be reported on in informal terms.

Criteria for Classifying Agent Models

As shall be argued throughout the article, goal-directed and rational agents represent two complementary ways, endogenous and exogenous, to approach intelligent agents, which in turn largely depends on whether agents are instruments or objects of scientific inquiry.

Taken as scientific instruments, agents are modeled and implemented on a computer in order to generate phenomena [33] that are themselves the primary objects of scientific investigation. This is the case with macro-social phenomena *grown* in artificial societies, where agents are specified only to the extent that is required to bring about the (macroscopic) effects of interest. This usually consists of an interesting dynamics at the aggregate level, such as emergent macro-social effects like segregation, public goods provision, opinion dynamics, etc.; or stylized social facts, such as game theoretic scenarios, in order to explore social or collective problems and dilemmas and account for prosocial action, such as cooperation [6], altruism [34], fairness and justice [10], etc.

As targets of scientific interest, agents – in particular autonomous agents [41,96,100] – draw the scientist's attention on the mechanisms of regulation that activate and rule the activity of any system of action. In this sense, agents are studied and described in their internal mechanisms of regulation. Ideally, the language, the concepts, and the formal instruments applied to them, describe agents and their behaviors precisely in the same mechanisms that operate when they behave. Such a description of course corresponds to the real processes as suggested by the current scientific standards.

Which consequences do these different ways of conceptualizing agents bear? The distinction discussed above leads to the emergence of two different approaches to agent modeling, *endogenous* and *exogenous*, which have not as yet been systematically analyzed.

Endogenous vs. Exogenous Approach Generally speaking, endogenous variables are meant as properties originating into the system under study. Exogenous variables instead are properties originating from outside the system, which are not under its control nor are manipulated by the system itself, but can be manipulated by the observer/experimenter.

In computational models, exogenous variables are parameters set to defined values by the computer scientist. It is interesting to observe that an extended parameters' space poses a serious problem of computability to the analyzer (see [99], for a recent review of alternatives and prospects), who must arbitrarily choose which parameters to analyze extensively. Therefore, attention is currently paid to the so-called parameters' "endogeneization" [19].

By endogenous models, it will be here meant to refer to models describing agents and their behaviors in terms of a special type of endogenous variable, i. e. their generative mechanisms. These are supposed to operate *within* the agents when they exhibit any given behavior, including responses to external stimuli. Exogenous models, instead, describe agent's behavior from the outside in terms of its effects or functions (e. g., utility or fitness function). An exogenous model is indifferent to how behavior is actually ruled; it is limited to apply a formal description that accounts for the behavior under study as this is perceived by the external observer.

Notice that the exogenous and endogenous approaches are not necessarily concurrent: they tackle agency from two different perspectives. For example, on the grounds of an exogenous model, one can assume that a certain opinion and the corresponding behavior spread over a population as an effect of the physical space (see [68]). It is interesting to notice that so far, computa-

tional models have adopted exclusively an exogenous approach to the study of opinion dynamics [30,51,71,73,90], modeling the external dynamics of opinions, i. e. their dynamics in society, independent of the endogenous mechanisms allowing agents to accept [84], update [58], revise [12], transmit [72], retreat and finally drop their opinions. However, at least in principle, an endogenous model of opinion dynamics is necessary to unfold the social dynamics: to understand how physical or social distance impacts on opinion dynamics, we ought to understand how opinions *gain ground*, agent by agent, mind by mind. In other words, we ought to model the process by means of which agents influence one other and accept others' opinions.

One might say that these two approaches tackle different aspects of the same phenomena, or different steps of the same processes, the exogenous model being focused on the so called distal causes, the endogenous model addressing the proximal. At a more careful sight, however, this does not seem to be the case. Indeed, it should be noticed that proximal causes are sometimes exogenously modeled. This is the case with game-theoretic models that describe decision-making (proximal cause of behavior) by means of a utility function taking into account wealth distribution and risk sensitivity [42,67]. This can hardly be conceived of as an internal mechanism ruling agents' decision-making. While taking what frequently proves to be an intelligent, adaptive decision, agents haven't the faintest idea that the process they have gone through might well be described by means of a mathematical function. Far from identifying the mechanisms generating the agent's choice from within, this function describes behavior from the outside.

Game-theoretic models have often been charged with the accusation of proceeding from implausible assumptions. This accusation is warranted from the point of view of an endogenous model, not from an exogenous standpoint. That agents maximize their expected utility is probably a truism, at least if utility is not meant to coincide with self-interest [39]. If utility is meant to coincide with own desires or goals [5], the rational assumption is perhaps too obvious, even tautological, but not implausible. What makes it implausible is the description in which it is conveyed: agents do *not* calculate mathematically their expected utility, although many times they *act as if they were doing so*. The utility function is a valid exogenous model, of course, but is an implausible endogenous explanation.

The problem is, what is a good endogenous explanation? What are the internal mechanisms that allow agents to maximize their utility? This contribution does not, nor is intended to provide an answer to this question. However, a good step forward in the direction to "endogenize" factors of regulation, is to turn to current scientific theories that aim explicitly at elaborating an endogenous model.

Analogously, we should refrain from the common fallacy that distal explanations can easily do without endogenous models. Things are not so simple. If one aims at generating the behavioral effect starting from a far removed cause, one needs to construct the whole chain between distal causes and wanted effects. Without an endogenous model, one can get nothing better than a partial and non-generative explanation. This is not equal to saying that this explanation has no right of its own. On the contrary, and somewhat in contrast with latest generativists [33], the present writer's claim is that exogenous models are non-generative statements that may have explanatory value [22].

As Nwana [81] argued in his overview of the field of computational agents, one of the symptoms of the ontological deficiency characterizing this field is the number of adjectives often accompanying the word agent. Rather than adequate definitions of these adjectives, which are easily accessible both in Nwana's or in Wooldridge and Jennings' papers [101], the present contribution attempts at discussing major criteria for categorizing agents and combining them with the epistemological approaches discussed above.

Weak vs. Strong Agents Wooldridge and Jennings [101] distinguished between a weak and a strong form of computing agents. Weak agents are characterized by a number of properties.

Autonomy Computing agents operate without the direct intervention of humans or others, and have some kind of control over their actions and internal state. Agents can be placed on a continuous dimension of autonomy, as several intermediate cases occur between full autonomy and full slavery. Indeed, limited autonomy [23] is the paradigmatic case of humans. Sources of limited autonomy can be found both in resource scarcity and in agents' sharing a common environment. Due to resource scarcity agents may not possess all of the means needed for achieving their goals (limited self-sufficiency). In a common environment, agents interfere with one another positively and negatively. Both factors make them liable to all sorts of mutual influences.

Social ability Agents interact with other agents (and possibly humans) via some kind of agent-communication language [44].

Reactivity Agents perceive and respond to their environment, including other agents.

Rational, Goal-Oriented Agents, Table 1
Set relationships among types of goal-oriented agents

		Goal-Oriented	
		Endogenous approach	*Exogenous approach*
	Sub-Agents	Reflex	Spin glasses Cellular automata
INTELLIGENT	**Weak-Agents**	Learning Evolutionary Neural nets Goal-directed	
	Strong-Agents	Bounded rational	
		Deliberative (BDI, Cognitive)	Rational

Pro-activeness Agents act in response to their environment, but also take the initiative [101].

In addition to these properties, strong agents are computer systems characterized by mentalistic notions, such as knowledge, belief, intention, and obligation [93] as well as emotions [7,8].

What happens if we combine the dimension of weak and strong agency, with the endogenous and exogenous approaches?

In Tab. 1, goal-oriented agents are classified. In the first row, sub-agents appear, i. e. those lacking one or another of the minimal properties for weak agency identified above. In particular, none of the systems classified as sub-agents are proactive. Below, intelligent agents are shown: they include weak and strong agents, endogenously and exogenously modeled.

Overview: Weak Agents

Weak agents are essentially autonomous, interconnected and communicating systems that effectuate state transitions. Unlike simpler and stronger models, no exogenous examples of weak agents are found. They are all described in terms of their internal mechanisms. Verisimilarly, simpler models lend themselves to be described with mathematical formalisms, which are also applied to model some complex capacities, like rational choice (see rational agents). Instead, weak agents are not modeled to *describe* the effects they produce in the world, but to find out algorithms generating them. The interest of these systems is intrinsically computational, even technological, and they gave a strong impulse to endogenous modeling.

Endogenous Weak endogenous agents are described in their internal mechanisms, essentially rules or subsymbolic representations. However simplified, these enable the agents to behave autonomously, react to external stimuli, and show also some kind of proactive behavior.

Evolutionary In their early time, agents were modeled as rigid machines, which might assume a fairly limited set of states on the grounds of a number of fixed rules. This was the case with cellular automata, which at the origin were fixed to given locations on a toroidal grid and could assume either an active or a nonactive state depending on the states of neighbors. Action was taken according to a fixed set of rules [26].

Several factors contributed to promote flexibility, the principal of which probably was the growing interest in evolutionary phenomena and the capacity of agents to modify and adapt to a changing environment. Fitness function was designed as a function measuring how good a given property or trait is in a pool of traits (for example a gene in a population): the agents performing above a given value would survive and reproduce whereas those performing below would extinguish.

Let us look more closely to computational models based on a fitness function. In these models, agents' behavior is the expression of a vector in which given traits are represented. As the metaphor is clearly the evolutionary mechanism, mutations in the vector may occur with a probability that is established offline. Mutations that contribute to the reproductive success of the agent will be transmitted to the future generations after a recombination with the partner's traits.

Genetic algorithms (GAs) are a particular class of algorithms inspired by evolutionary biology, and implementing evolutionary mechanisms and processes, such as inheritance, mutation, selection, and crossover (also called recombination). Starting from a population of randomly generated individuals, at every generation the fitness of every individual in the population is evaluated, the current population genotype is modified (recombined and possibly randomly mutated) and the new population is then used in the next iteration of the algorithm. The fitness function is problem dependent and in some problems, hard to define.

A GA proceeds through iterated executions of mutation, crossover, and selection operators. Whereas selection is an important genetic operator, the importance of crossover versus mutation is more debatable (references in Fogel [36] support the importance of mutation-based search). To be noticed, for specific problems, other optimization algorithms may do better than GAs with equal

speed, for example simulated annealing. (Simulated annealing (SA) is an algorithm for locating a good approximation to the global optimum of a given function in a large search space. It is often used when the search space is discrete (e. g., all tours that visit a given set of cities). In favorable cases, simulated annealing may be more effective than exhaustive enumeration of the search space [59].)

As with other computational models, one main problem is the parameters' space search (mutation probability, recombination probability, etc.) To which value should one set the mutation rate, when no reference is made to evolutionary phenomena known in nature? A very small mutation rate may lead to genetic drift. On the other hand, a high mutation rate may lead to loss of good solutions.

GAs are an invaluable means for optimizing fitness. A fascinating explanation, the so-called building block hypothesis (BBH), has been proposed to explain this effect [46]. Although heavily criticized (see [102]) and poorly supported by experimental evidence, BBH provides an intuitive reason for the mechanism of adaptation, which is said to result from iterated recombination of "building blocks", or low order elements: at any recombination (or generation), past solutions that proved fitter than their own building blocks are recombined to increase the optimality of new solutions. By this means, genetic algorithm achieves step-wisely optimal performance through the iterated recombination of building blocks.

If GAs are meant to optimize the genotype of any given population, genetic programming (GP) is used to optimize ability to perform computational tasks. After some pioneer computer scientists, and especially John Holland in the early 1970s, John R. Koza ([62,63,64,65,66]) has pioneered the application of genetic programming in various contexts. Results obtained by the application of GP include innovations in hardware as well as computer programs.

Learning Learning algorithms used in agent based models include classifier systems, reinforcement learning, Q-learning, and neural nets.

Classifier Systems Classifier systems (CSs) are strongly related to genetic algorithms. Introduced by John Holland [52], CSs consist of sets of rules optimized thanks to a genetic algorithm that operates on reinforcement rather than on a fitness function. There are two main types of CSs. In one case, the genetic algorithm recombines different sets of rules; in the other, the genetic algorithm operates within one set only.

Reinforcement Learning Inspired to behavioral science, reinforcement has been applied to machine learning as a reward mechanism (see [54]). Reinforcement based algorithms have been applied in economics and game theory as a boundedly rational interpretation of how agents endowed with the same learning mechanisms may converge on one equilibrium. Given a set of world-states, a set of actions A, and a set of "rewards", at each time the agent chooses an action and receives a certain reward. Reinforcement will cause agent to maximize the quantity of reward obtained.

Recently, reinforcement learning has been used in cognitive models of human problem solving (e. g., [43,47]) and error-processing [53].

Neural Nets Borrowed from neuroscience, where they are identified as groups of neurons performing a specific physiological function, neural nets currently refer to artificial neural networks [50]. These are made up of interconnecting artificial neurons, and used either to study biological circuits – a simplified view of artificial neural networks may be used to simulate properties of neural networks – or to solve artificial intelligence problems, such as speech recognition, in order to construct software agents or autonomous robots.

Neural networks have been viewed as simplified models of the brain, but the relation between this model and the brain architecture is far from obvious. After all, computers entail sequential processing and are based on explicit instructions, whereas neural networks model biological systems as performing parallel processing on implicit instructions.

An artificial neural network is an adaptive system. Its structure gets modified based on information spreading through the network.

Neural nets are based on sub-symbolic representations. These are physical sets of interconnected units, with variable weighs on connections. The resulting networks are said to take inspiration from the organization and structure of neuronal synapses. As they *stand for* nothing, they hardly fit the notion of "representation" as such. Nonetheless, they effectively play a number of important functionalities of representations, especially pattern-recognition. Given a certain input, a given set of internal units may be gradually "taught" to "recognize" it by contrast with other inputs, thanks to a suitable modification of the weighs of given connections. Subsymbolic representations are

- Quantifiable, whereas symbolic representations vary, and can be compared on, a qualitative base.
- Gracefully degradable, whereas symbolic representations are none-or-all. Hence, unlike these latter, they al-

low for partial and reparable damages of the long-term memory.
- Compatible with non-decided upon and non-controlled activity.

Applications of artificial neural networks include pattern and sequence recognition [87] (radar systems, face identification, object recognition, etc.), sequence recognition (gesture, speech, text recognition), medical diagnosis [39], data mining, etc.

Goal-Directed Agents Goal-directed agents are systems ruled by, and described in terms of, their goals.

In a psychological-behavioral sense, a goal is a motivational factor defined on the grounds of a number of observable features, among which vigorous attainment, persistence in the face of obstacles, resumption after disruption ([69], etc.).

Another notion of goal derives from the theory of systems and cybernetic circuits. Control theory introduced the notion of feedback [15] to control states or outputs of a dynamical system [40]. Its name comes from the information processed in the system: inputs (e. g. a given state of the world) have an effect on the outputs (for example, elicit a given action) that is measured with sensors, and the result of which (the worldstate brought about by the action) is used as input to the process.

On the grounds of such a notion of goal, a theory of behavior as planning activity [78] has been constructed, which has played a foundational role in the development of cognitive science.

Building on control theory, cognitive scientists define a goal as a wanted state of the world that, if discrepant from the currently perceived worldstate, activates the agent and rules its action. A goal is therefore both a trigger and a regulatory state: actions are activated and ruled by their goals [23].

Further analysis allowed several categories of goals to be identified.

End- and sub-goals. A goal can either be the final end of activity – f.i., survival is probably the supreme finality of natural systems – or a condition for its execution, which may not be verified in real matters. In such a case, in terms of AI planning, the initial end-goal gives rise to further sub-goals, i. e. to verify the necessary conditions for action execution.

Innate vs.learned. Whereas end-goals are usually innate (but examples of subgoals that become ends in themselves exist), subgoals are formed and dropped after successful execution.

Side-goals. Goals may guide the action they trigger, or they may rule actions triggered by others goals. The latter are side-goals. A typical example is utility-maximization, which rules a number of activities, especially economic, triggered by other goals.

Achievement and maintenance goals. Goals may be already verified in real matters, i. e. coincide with currently perceived worldstates. In this case, they are usually dropped – as is the case with purchasing a new car, or finding a job. However, in many cases, these fortunate circumstances may suggest a new goal, i. e. to keep one's job or car. In the BDI language (see below), the goal to maintain a current worldstate is a maintenance goal, whereas the goal to realize a given worldstate is an achievement goal [20].

Realizable vs. unrealizable goals. In the logic-based treatment of goals, these must be consistent and realizable. Unrealizable goals are treated as a subset of desires that are never chosen for action.

The first application of goal-directed agents goes back to the AI planning systems [89,97], which soon gave impulse to a Copernican revolution in planning and plan execution, i. e. Distributed Artificial Intelligence (DAI) systems [11]. In the early nineties, the first studies and scientific events [31] on Multi Agent Systems (MAS) took their moves from the necessity to increase effective autonomy of distributed systems and their interactive capacity. Both of these requirements, in turn, necessitated developments of mental capacities, especially in the direction of social intelligence [21], what soon led to the appearance of stronger models of agency.

Overview: Strong Agents

Below, we will review the main approaches to account for complex tasks, like deliberative choice. As was premised above, both the endogenous and the exogenous approaches concurred to this aim. Let us examine them in turn.

Exogenous Here, rational agents are shown to represent the paradigmatic case of exogenous models of strong agents, whereas BDI and cognitive agents examplify strong agents endogenously modeled.

Rational Agents One way of conceptualizing rational action describes it as *reason-based* action. Hence rational agents are sometimes identified with intelligent agents at large [88,100].

In this contribution we go back to the classic notion of rationality as utility-maximization, and define rational

agents as a subset of intelligent agents that act *so as* to maximize their utility [2,80]. In this more specific sense, a rational agent takes actions expected to maximize its chances of success [4]. Its action is said to bring about this effect *no matter how*. In other words, a rational agent model is indifferent to the mechanisms that effectively operate within the agent to ensure the hypothesized effect [1]. At most, what is attributed to a rational agent is (limited) knowledge of its environment, consisting of (a memory of) past experience, current information about the environment (possibly including other agents and their expectations [70]), as well as the estimated benefits and the chances of success of own actions [88].

The classic conceptualization of rational agents derives from the abstract model of *homo oeconomicus* defined by John Stuart Mill [77] as an entity which aims at possessing wealth and is able to evaluate and adopt the means for achieving it. On these assumptions, rational choice took on the specific meaning of *self-interested* action. Thus meant, the term "rational" implies a utility function: independent of the benefits themselves, what is said to be rational is the choice that maximizes the distance between the benefits obtained and the costs sustained to obtain them.

Rational action theory provides a formal model of practical behavior, mainly applied to social and economic domains. We might consider it as the still dominant paradigm in economy, but its hegemony extends to all of the social sciences and to the computational field of agent theory and technology. Essentially, it assumes individuals to choose the best action according to stable utility functions. Rather than describing reality, rational choice theory aims at supporting reasoning (prescriptive use) and modeling social behavior in highly stylized scenarios.

Requirements of Rational Action Let us examine the view of the agent implied by the rationality assumptions. Choosing an action rationally requires individual preferences, a set of options or alternatives for actions, and a set of expectations about their outcomes. Two important assumptions concern preferences:

Completeness All actions are ranked in an order of preference.

Transitivity If any given action a_1 is preferred to action a_2, and the latter is preferred over a third option a_3, then a_1 is preferred to a_3.

Consequently, an individual can always order the available options on the grounds of its preferences, which will always be consistent. This gives rise to a utility function, i. e. an ordinal number assigned to alternative actions – such as: $u(a_i) > u(a_j)$ – with preferences being defined as relations between these assignments. Unrealistic assumptions are often made, such as full or perfect information and the ability and time to weigh every choice against every other choice. Relaxations of these assumptions are included in theories of bounded rationality, which will be examined later.

Although empirical support to rational choice theory is far from satisfactory [48], this paradigm presents two main advantages for the social sciences.

First, precisely because it proceeds from stylized facts and abstract phenomena, it proves a powerful instrument for the scientific study of social action.

Secondly, and consequently, rational action theory in principle allows future actions to be predicted. Compared to stochastic modeling, the predictive power of social scientific explanation is expected to increase. However, it should be observed that the tautological and a posteriori explanation of behavior deriving from the principle of utility maximization strongly reduces the predictive potential of the theory. As preferences and goals are subjective, any choice results in the maximization of utility, whatever its outcome is and can be said to be rational (tautology), once it has been made (*a posteriori*).

Third, the abstract and strong nature of the theory allows testable hypotheses to be formulated and data to be collected. The experimental potential of social scientific explanation is therefore bound to increase.

However, rationality theory has been criticized on empirical grounds. In traditional societies, people's social action, for example the way they exchange goods, follows rules that could not be expected on the grounds of rationality assumptions and which, thanks to the famous work by Mauss [75], seemed to point to a regime of "gift", rather than market, economy. In addition, according to experimental scientists (cf. the wide literature cited in [82]), people are much more cooperative than expected by the rational choice theory. Finally, the rationality assumptions are commonly said to necessitate too much understanding of macroeconomics and economic forecasting for people to take rational decision. We will get back on this argument later in the next section.

Theoretical critiques are no less strong. Among others, it has been observed that the role of intrinsic motivations [29], for example that of altruism, is under-estimated to the advantage of extrinsic, incentive-based, motivations, an objection that is not entirely warranted in consideration of the indetermined nature of preferences: rational agents can be subjectively inclined to altruism. In such a case, even one's extreme sacrifice to the benefit of others' is perfectly "rational".

Rather, the inadequate account of preferences dynamics poses a twofold problem to the social scientist. On one

hand, totally indeterminate preferences reduce the predictive potential of the rationality paradigm. On the other, if intrinsic motivations are acknowledged, little is said about how they can be acquired. As a consequence, the origin of altruistic motivations, like that of any other taste and inclination, and the effects of social influence, training, education, and the like on the utility function are overlooked.

Another consequence of rational agents' poor internal dynamics is the insufficient account of inner conflicts like that between motivations and goals (for example, passing the exam and playing around) or that between types of motivations or finally between individual goals and societal values, which may strongly affect and impair individuals' decision-making.

Substantially, these critiques converge on arguing against a purely exogenous approach, which ignores the acquisition, modification and integration of internal states. To say that agents' choices proceed from their preferences is neither particularly interesting nor informative, but to make the model more realistic does not help much and may render it tautological. The missed point here is a theory of the internal dynamics of preferences. However, this requires an endogenous approach to rationality.

Endogenous Strong models of the agent aiming at modeling internal states and mechanisms include a variant of the rational choice theory, i. e. bounded rationality, as well as theories and architectures of cognitive agents, including but not reduced to BDI agents.

Bounded Rationality Bounded rationality is a variant of rational choice theory that took its move from the obvious consideration that perfect knowledge is never available, and that agents, even investors, deal with uncertainty and risk in business. Far from regretting the special attention paid to economic behavior and the *homo oeconomicus* model, the proposers of bounded rationality questioned the utility of full rationality as a paradigm for the study of economic behavior itself. As Fox and Tversky [37] observed, the very notion of uncertainty was first questioned by theoretical economists. In [61], Knight distinguished between measurable uncertainty or risk, and unmeasurable uncertainty, suggesting that entrepreneurs deal with unmeasurable uncertainty, rather than risk. At the same time, Keynes [59] distinguished between probability of occurrence of a given event and the weight of the evidence supporting it. Nonetheless, game theorists showed little interest in the issue of vagueness of probability [91].

When one considers it carefully, bounded rationality is situated at the borderline between endogenous and exogenous approaches. Prospect Theory [56], one of the main theoretical contributions of bounded rationality, was proposed as a psychologically realistic alternative to rationality theory, in particular to expected utility theory. The authors meant to account for empirical evidence showing that when people evaluate potential losses and gains under risk, e. g. in financial decisions, the probability of occurrence acts differently on the values of alternative options depending on whether agents are gaining utility or avoiding losses. Rather than constructing a model that shows how this happens in the minds of the agents, the authors expressed this evidence in an asymmetric utility function. This shows a bigger impact of losses than of gains (loss aversion) and expresses that people tend to overreact to small probability events, but underreact to medium and large probabilities. The theory, meant as an endogenous development of full rationality, is expressed into an exogenous formalism.

Nonetheless, bounded rationality has done a great job in revising the assumptions of full rational theory. What is more, Herbert Simon [94] did actually argue for an endogenous model, when he claimed that in order to account for rational action, it is necessary to study the mental processes and mechanisms allowing agents to formulate and solve complex problems and process (receive, store, retrieve, transmit) information (see also [98], p. 553, quoting Simon). One important aspect of the mental endowment is represented by heuristics [57,83], i. e. powerful informal methods to solve problems in complex situations. In psychology, heuristics are simple, efficient rules, rules of thumb, which either evolve or are learned and support people in taking decisions, coming to judgments, and solving complex problems under incomplete information. When agents cannot compute the expected utility of every alternative action, they resort to heuristics, which are applied in different contexts and might have been acquired over the course of life, through repeated decision-making processes. These rules work well under most circumstances, but in certain cases lead to systematic cognitive biases, i. e. very basic statistical and memory errors that are common to all human beings.

To be noticed, most models of bounded rationality did not strictly follow Simon's ideas. Some authors preferred to model people's decisions as a sub-optimal version of rationality [32] or to study how people cope with their inability to optimize. For others [45], alternatives to full rationality in decision making frequently lead to better decisions.

Deliberative Agents Let us now turn our attention to fully endogenous models of complex agents, such as BDI architectures and more generally cognitive agents. Both

are goal-directed systems, not incompatible with utility maximization but actually implementing it. However, goal-directed agents do not necessarily imply neither explicit computation of utility maximization, nor heuristics achieving it.

Cognitive systems. These are goal-directed agents endowed with mentalistic notions, such as beliefs, intentions, obligations and the capacity to manipulate mental symbols and accomplish a number of operations on them, including inference, reasoning, problem-solving, planning, etc. Symbolic representations *stand for* the states of the world they refer to in such a manner that these world states can be:

- Compared: a given representation of a worldstate ws1 can be compared with another to check whether they are equal or not. This is fundamental in goal-directed action.
- Manipulated, e. g. asserted or denied, evaluated and possibly selected along common criteria, modified and revised and therefore used as instruments for reasoning, action planning, etc.
- Embedded into one another, thereby giving rise to meta-beliefs. Among other reasons, meta-beliefs are useful for the representation of others' beliefs (the so called "theory of mind"), which is of vital importance in social life [16,85]. Of late, theory of mind has been used to refer to a specific cognitive capacity: the ability to attribute mental states – beliefs, intents, desires, pretending, knowledge, etc. – to oneself and others and to understand that others have beliefs, desires and intentions that are different from one's own [27,49].

BDI agents. A BDI agent is an architecture for a special type of cognitive agent, more specifically for intention-driven agents, endowed with particular mental attitudes, viz: Beliefs, Desires and Intentions (BDI).

The philosophical bases of BDI agents is Bratman's theory of practical reasoning, but has been formally described by Anand Rao and Michael Georgeff's [86], which combine temporal logic with a modal logic enriched with beliefs, desires and intentions.

Beliefs Beliefs represent the agent's representations about the world (including itself and other agents), which are neither necessarily true nor complete. Inference rules generating new beliefs allow to enrich the agent belief base.

Desires Desires represent the motivational state of the agent. Goals are a subset of desires, i. e. consistent and achievable ones.

Intentions Intentions are executable goals, i. e. goals that the agent has chosen for action, to which the agent has to some extent committed (in implemented systems, this means the agent has begun executing a plan).

Aimed to represent the future as an epistemic tree, allowing agents to reason upon alternative courses of action, BDI architectures lend themselves to model the interplay among different mentalistic notions. In particular, an extension of BDI, the BOID architecture, concerns the representation of the interaction among beliefs, obligations, intentions, and desires, and has been used to support reasoning and deliberative choice on legal norms [14].

Mental Representations and Their Dynamics An endogenous approach to the computational model of intelligent agents revolves around the main question as to how agents represent their goals and the information needed to achieve them [28,74,95]. This issue, called knowledge representation, arises at the intersection between cognitive science and Artificial Intelligence: AI scientists have borrowed methods of knowledge representation from cognitive scientists, since it was of primary importance for them to design programs that could store knowledge and re-apply it to analogous problems.

Knowledge Representation techniques – such as frames, scripts, etc. – stem from theories of human information processing. The main goal is to represent knowledge in such a way as to facilitate reasoning, planning, problem solving and the like.

The main questions posed by AI researchers concerning knowledge representation focuses on the format of representations, the nature of knowledge, the preferability of general purpose vs. domain-specific representation schemata, the degree of expressiveness and detail of such schemata, whether it should be declarative or procedural etc.

In AI, many methods of knowledge representations have been tried since the early seventies, e. g. theorem proving and expert systems, with variable degree of success. The major knowledge representation systems go back to the eighties when for example large databases of language information were built, rendering knowledge representations more feasible. Several programming languages were developed, such as Prolog, KL-ONE and XML, facilitating information retrieval and data mining.

The necessity to store and manipulate knowledge in a formal way so that it may be used by mechanisms to accomplish a given task has drawn the attention of knowl-

edge representation scientists. Expert systems, machine translation, and information retrieval are good examples of applications.

Probably the most extensively used structure for knowledge representation from the 1960s, is the knowledge *frame* [35,79], with its own name and set of attributes, or slots that contain values. Frames have been used for expert systems in object-oriented programming, with inheritance of features described by an IS-A link, a type of relation that has posed a number of problems of semantic interpretation [13]. However, in an endogenous perspective, it should probably be observed that frame structures are well-suited for the representation of stereotypes and other socio-cognitive patterns.

A *script* [92] is a frame that describes events, especially behavioral events. The usual example is going to a restaurant, with steps including waiting to be seated, receiving a menu, ordering, etc.

Goals Typologies of goals are not limited to those examined in Sect. "Goal-Directed Agents", but include other sub-categories, which are based on the interplay among goals and other mental notions, such as beliefs and obligations.

Active vs. inactive goals. Maintenance goals are not always active. A goal is active when it rules the system's behavior. When a goal is represented in the agent's mind but does not rule its behavior is inactive. A maintenance goal may be activated when it is likely to be thwarted. An inactive goal may be a maintenance or side-goal, or an achievement goal that becomes (momentarily) inactive because incompatible with a more important one (emergency or conflict).

Executed vs. waiting. An active goal may be chosen for action, or may be interrupted during its execution and kept waiting in a queue, because the conditions for its execution turn out to be false, and must be realized for the goal to be achieved.

Chosen. An active goal may be chosen for action and transformed into an intention when (a) it is not realized, but it is (b) realizable and (c) more important than other active goals, if any. These clauses are conjunctive: in case the goal is realized, or outcompeted, it becomes inactive. In case it is unrealizable, it is dropped (or retrocedes to the status of a simple desire).

Individual vs. social. Goals are individual when mentioning no-one else except the hosting agent. Otherwise, they are weak or strong social goals. Weak social goals mention others as models to imitate or as *sources* of influence – for example their future actions are perceived as either obstacles or favoring events. Strong social goals mention others as *targets* of influence (simple request or persuasion or manipulation or coercion).

Prosocial vs. aggressive. A goal is prosocial when it aims at realizing a worldstate *as long as* and *why* this is wanted by another agent. The process by means of which an agent comes to have another's goal as its own is called goal-adoption. This may be either an end in itself (benevolence) or a sub-goal for an individual goal (e. g., exchange).

An aggressive goal is a goal that aims at thwarting another's goal, as long as and why this is wanted by the latter. Again, an aggressive goal may be an end-goal (hate), or a sub-goal, and in the latter case it may be instrumental to a selfish or a pro-social goal (prevent crime), even to the benefit of the victim (surgery, education, etc.).

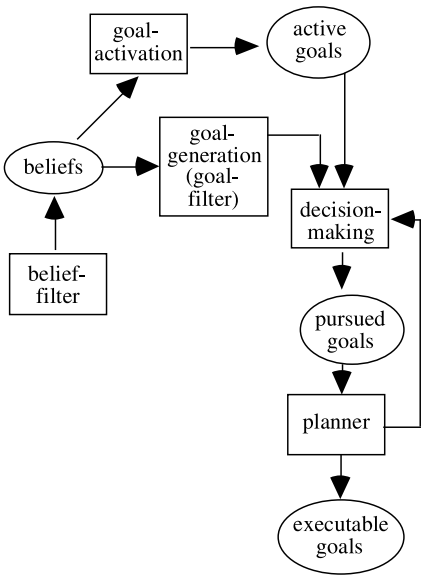
Shared vs. collective. Finally, a goal may be shared by a set of agents. It is collective [25] when it mentions a set of agents in which any one is necessary, but none is sufficient to achieve the goal shared by them all (an orchestra or a football team are typical examples of collective goals).

Cognitive Dynamics In a rather generic sense, an autonomous agent is a self-interested agent. In a more specific sense, an autonomous agent is one, which has internal criteria to select among inputs. Inputs might generate beliefs and goals. An autonomous agent is therefore characterized by a "double filter architecture", allowing both beliefs and goals to be selected (cf. [17]). These two filters are sequential, but at the same time they allow for an integrated processing of mental representations.

Filtering beliefs. Thanks to this filter, agents have control over the beliefs they form. This filter is rather complex and implies that a number of tests be executed over a candidate belief against several distinct criteria. These are pragmatic or epistemic criteria. Epistemic criteria include

- Credibility, with agents controlling, among other properties, the coherence of candidate beliefs with previous beliefs, the reliability of the source of the candidate belief: agents accept information from other agents (Gricean principles) provided they have no reasons to doubt their sincerity or competence.
- Non-negotiability, or the Pascal law. To believe or not is a "decision". However, it cannot be made in view of one's pragmatic utility, but only in view of one's epistemic utility. In social interaction, we cannot use threats ("Argumentum ad baculum") or promise to make people believe something. The difference between persuading to do and persuading to believe is crucial. Since beliefs control goals, this represents a further protection of agents' autonomy.

Pragmatic criteria concern the reasons for believing something. Typologies of beliefs generally concern the format of their representation (declarative, procedural); the degrees of certainty; the levels of nesting (one can believe something, without believing that one believes . . .). These typologies are known too well. Perhaps it is less obvious that beliefs may have a different "status" in the mind according to the motives for acceptance. The language provides a rich vocabulary: superstitious belief, creed, faith, credence, doctrine, postulate, axiom, principle, conception, idea, view, opinion, and many others. These beliefs vary on several, often quantitative, dimensions, such as certitude (subjective truth value), retractability (how likely it will be modified), connectivity (how much a belief is connected with other beliefs). One interesting quantitative dimension is the "force" of beliefs (cf. the role of this dimension in the Social Impact Theory, cf. [68]): beliefs vary on how strongly they are held. This is related to certitude, but also to the motives of acceptance, which may lead the agent to ignore the belief's truth-value; for example,



Rational, Goal-Oriented Agents, Figure 2
The "double filter" architecture

- Self-protection and self-enhancement, agents may be led to accept one among several competing beliefs because of the belief's positive effect on their self-esteem or self-concept.
- Commitment to a given (set of) belief(s) provides one important reason for accepting further, consistent beliefs, despite or independent of incompatible evidence: agents that accept beliefs out of commitment don't check their truth value.
- Hypothetical and counter-factual reasoning, beliefs may be (transitorily) held as means to reason and carry out operations (demonstrations, proofs, experiments). A god example is the priest accepting the atheist's point of view to dismantle it.
- Communication: A psychotherapist may "accept" the delusions of her patient in order to communicate with him and give a clinical sense to his fantasies; here the goal is not to carry out a counterfactual argumentation, but to understand the meaning of delusions.
- Empathy: agents may want to share the views of their close connections.
- Risk-taking or gambling: agents may participate in lotteries, accepting one alternative and investing (money) in it; in such a case, agents will hold an uncertain belief but behave as if it were certain.
- Prudence, agents may accept uncertain information (for example, rumours, gossip, even calumnies) and behave as if they were certain: unlike the preceding situation, a risk-minimization strategy applies.

Filtering goals. There are at least two fundamental tests which are performed on goals: self-interested goal-generation (an agent is autonomous if whatever new goal it comes to have, there is at least another goal of that agent, for which, in the agent's beliefs, the former is a means), and belief-driven goal-processing (any modification of an autonomous agent's goals can only be allowed by a modification of its beliefs) (cf. Fig. 2).

Both these filters bear interesting social and cultural consequences. First, agents' minds are modified thanks to a process of belief-formation or belief-revision. Secondly, belief-formation and -revision are decision-based and selective processes. With Cavalli-Sforza and Feldman [18], we will speak of beliefs' "acceptance". A cultural process is a process interspersed with decisions taken by the agents involved. But a decision-based process is not necessarily explicit and reflected upon: mental filters do not necessarily operate consciously: agents may not be able to report on them. Thirdly, agents will never accept beliefs under threat, or in order to obtain a benefit in return (non-negotiability). Fourth, agents may accept beliefs for different reasons, and these will affect the probability that such beliefs are held, their strength, and their transmission. The social mechanisms of influence and transmission are strongly intertwined with criteria and motives of acceptance. This leads us to the question of limited autonomy.

Limited autonomy. The model of agent outlined above appears rather abstract and unrealistic. In the real life,

agents appear liable to external influence, prone to accept and transmit prejudices, victim of superstition, prey to false doctrines and creeds. Indeed, autonomy is limited both at the belief and at the goal level: agents are liable to being influenced by external (including social) inputs.

Both at the level of goals and beliefs, autonomy is limited in a very elementary sense: agents are designed to take into account external inputs, if only to discard them later. If an input is received, the filter processing is activated (cf. [17]). At the goal level, agents cannot avoid accept very elementary requests: if somebody asks a passer-by the time, this will not keep on ignoring the request. At most, the passer-by can pretend that she did not perceive it. But if this perception cannot be concealed, an answer whatsoever will be given, if only to say that one has no idea what the time is (minimal level of adoption). Of course, this type of influence is rather superficial and ephemeral. But it paves the way for other more relevant types of influence. Obviously, agents' autonomy is limited because they are not always self-sufficient. They may need the help of other agents to achieve their goals (social dependence), and this causes agents to adopt others' goals and to accept their requests. However, one adoption of others' goals will always be a means for the achievement of one's goals for example through social exchange or cooperation. In turn, these social actions favor the transmission of beliefs, including action plans, techniques, procedures, rules, conventions, social beliefs. Finally, agents' autonomy is limited by norms, which are aimed to regulate agents' behaviors. But agents may accept or reject norms, comply with or violate them, always according to their internal criteria for acceptance.

At the level of beliefs, agents' liability varies depending upon the type of beliefs. For example, social agents are strongly permeable to social evaluations, rumors, gossip, even calumnies (cf. [24]). Rumors and gossip are accepted for prudence, and, as we will see, this favors their spreading. Indeed, these are important phenomena of memetic transmission, which spread social labels, stigmas, and prejudices, but also reputation, social hierarchies and other institutions.

Concluding Remarks and Future Directions

Which are the similarities and different advantages of these different paradigms in agent modeling, especially among rational and goal-directed agents? As initially stated, these should not be viewed as mutually exclusive. Hence, the question arises as to when, for which scientific objectives, and how they could be integrated. But first, let us see what they have in common and where they differ.

Rational and goal-directed agents have three major common properties: autonomy, (limited) knowledge, and strategic action. These attributes characterize the fundamental endowment of intelligent agents, allowing them to act on their own criteria, reasons, and current beliefs.

To be noticed, however, these properties are differently characterized within the two approaches.

Autonomy. Whereas rational agents are self-interested systems – with self-interest being closely related to the notion of adaptation and fitness in a biological sense – goal-directed agents are systems pursuing their own goals, whether these are means for own fitness or not. A suicide is a goal-directed agent, which autonomously decides to take out its life. As this example shows, goal-directed agency is an abstract and general category that can easily be applied to natural systems, which not always pursue their self-interest.

Limited knowledge. Both approaches model agents' beliefs as limited and uncertain. However, the rational approach does not account for the internal dynamics of beliefs, their acquisition, revision, and eventual dropping. Even adaptive views of agents fail to account for the internal dynamics of their motivations, and for the role these have in the management, update, and decay of their beliefs. However, the double filter architecture allows cognitive dynamics of goals and beliefs to be accounted for.

Strategic action. As to this aspect, i. e. the social application of rational action, one should acknowledge a superiority of the rational agents, paradigm, as this is inherently devoted to the solution of social and collective dilemmas. However, in the rationality paradigm, strategic action is meant exclusively in the weak social sense meant above, i. e. as the ability to take into account what other agents know, can do and will do. The strong meaning instead is essentially overlooked, since rational agents are not modeled as goal-directed: they cannot be said to pursue the goal to *modify* others' beliefs, goals, and actions. This is one of the main reasons why it is difficult to integrate social influence, training, education, and the other phenomena of social change in the rationality paradigm.

Despite the aforesaid shortcomings, the rationality paradigm is undeniably influential not only in the social, but also in the computational field. In comparison, the impact of goal-directed agents is far from satisfactory.

There are multiple reasons for such a state of affairs.

First, a persistent cultural hegemony of Homo Oeconomicus should not be ignored. However indeterminate subjective preferences are in principle declared to be, investment and economic business are tacitly assumed to be the master domains of application of rationality.

Secondly, rationality paradigm has all the advantages of an exogenous, in particular a mathematical, approach: (a) quantitative data, (b) theorem-proving, and (c) the reduction of computational complexity, as the rational agent architecture is rather simplified.

Thirdly, and most importantly, the rationality paradigm enjoys the powerful simplicity of abstraction, the invaluable scientific appeal of stylized facts. No other paradigm, so far, has been able to design such simple, highly abstract scenarios, characterized by the same heuristic power, as those represented by the *games* worked out by rational choice theory for the study of strategic action.

Making a difficult but fascinating exercise, let us endeavor to envisage the future of these paradigms. Whilst the rationality paradigm can reasonably be expect to maintain its hegemony in the short term, what can we expect to happen in the middle and long run? The growing success of the neuroscientific approach (think of *mirror neurons*) leads us but to foresee a dominance of the neural nets approach not only within the behavioral sciences (e. g., psychology), but also – and thanks to embodied agents and robotics – within the computational sciences. How and to what extent can this model answer the questions currently answered by rational agents, on one hand, and by goal-directed agents, on the other?

Largely, the answer depends on whether or not the predictable advances of the neuroscientific approach will empower agent technologies, for example, in the direction of social and collective intelligence, acquisition of norms and social values, formation and recognition of institutions etc. However, the answer might also depend on the potential cross-over between the neuroscientific approach and the complexity one. To what extent can embodied agents and neural nets be combined with sociophysical models? The future status of the rationality paradigm depends on the probability of this cross-fertilization. More generally, the future of the agent paradigms depends on the probability of integration between endogenous and exogenous paradigms.

Bibliography

1. Allingham M (2002) Choice theory: a very short introduction. Oxford University Press, Oxford
2. Anand P (1993) Foundations of rational choice under risk. Oxford University Press, Oxford
3. Aristotle (2006) Metaphysics book theta. Oxford University Press, Oxford [translated with an introduction and commentary by Stephen Makin]
4. Arrow KJ (1987) Economic theory and the hypothesis of rationality. In: The new palgrave: a dictionary of economics, vol 2. Macmillan, Houndmills, pp 69–75
5. Aumann RJ (2006) War and peace. Nobel Lect PNAS 103(46):17075–17078. November 14, 2006
6. Axelrod R (1997) The complexity of cooperation. Princeton University Press, Princeton
7. Bates J (1994) The role of emotion in believable agents. Commun ACM 37(7):122–125
8. Bates J, Loyall BA, Reilly SW (1992) An architecture for action, emotion, and social behaviour. Technical Report CMU-CS-92-144. School of Computer Science, Carnegie-Mellon University, Pittsburgh
9. Bigelow J, Rosenblueth A, Wiener N (1943) Behavior, purpose and teleology. Philos Sci 10:18–24
10. Binmore K (1994) Playing fair: game theory and the social contract, vol 1. MIT Press, Cambridge. Vol. 2: Just Playing: Game theory and the Social Contract
11. Bond AH, Gasser L (eds) (1988) Readings in distributed artificial intelligence. Morgan Kaufmann Publishers, San Matteo
12. Boutilier C (1995) Generalized update: belief change in dynamic settings. In: Proc of the 14th international joint conference on artificial intelligence (IJCAI'95). Morgan, Kaufmann, Montreal, pp 1550–1556
13. Brachman RJ (1983) What IS-A is and isn't. An analysis of taxonomic links in semantic networks. IEEE Comput 16(10):30–36
14. Broersen J, Dastani M, Hulstijn J, Huang Z, van der Torre L (2001) The BOID architecture conflicts between beliefs, obligations, intentions and desires. In: Mueller JP, Andre E, Sen S, Frasson C (eds) Proc of the 5th international conference on autonomous agents, Montreal, 28 May–01 June 2001, pp 9–17
15. Bush V (1929) Operational circuit analysis. Wiley, New York
16. Carruthers P, Smith PK (eds) Theories of theories of mind. Cambridge University Press, Cambridge
17. Castelfranchi C (1997) Representation and integration of multiple knowledge sources: issues and questions. In: Cantoni V, Di Gesù V, Setti A, Tegolo D (eds) Human and machine perception: information fusion. Plenum, New York, pp 236–254
18. Cavalli-Sforza LL, Feldman MW (1981) Cultural transmission and evolution: a quantitative approach. Princeton University Press, Princeton
19. Chavalarias D (2007) Endogenous distribution in multi-agents models: the example of endogeneization of ends and time constants, M2M 2007. In: 3rd international Model-to-Model workshop, Marseille, France, 15–16 March 2007
20. Cohen PR, Levesque HJ (1990) Rational interaction as the basis for communication. In: Cohen PR, Morgan J, Pollack ME (eds) Intentions in communication. MIT Press, Cambridge, pp 221–256
21. Conte R (1999) Artificial social intelligence: a necessity for agent systems' developments. Knowl Eng Rev 14:109–118
22. Conte R (2007) From simulation to theory (and backward). In: Squazzoni F (ed) Proc of the epistemological perspectives of simulation, 2nd edn. Springer (in press)
23. Conte R, Castelfranchi C (1995) Cognitive social action. UCL Press, London
24. Conte R, Paolucci M (2002) Reputation in artificial societies: social beliefs for social order. Kluwer, New York
25. Conte R, Turrini P (2006) Argyll-feet giants: a cognitive analysis of collective autonomy. Cogn Syst Res 7:209–219
26. Conway JH (1976) On numbers and games. Academic Press, London

27. Courtin C, Melot A-M (2005) Metacognitive development of deaf children: lessons from the appearance-reality and false belief tasks. J Deaf Stud Deaf Educ 5:266–276
28. Davis R, Shrobe H, Szolovits P (1993) What is a knowledge representation? AI Mag 14(1):17–33
29. Deci EL (1975) Intrinsic motivation. Plenum, New York
30. Deffuant G, Amblard F, Weisbuch G, Faure, T (2002) How can extremism prevail? A study based on the relative agreement interaction model. J Artif Soc Soc Simul 5(4). http://jasss.soc.surrey.ac.uk/5/4/1.html
31. Demazeau Y, Mueller JP (eds) (1990) Decentralized AI. Elsevier, North Holland
32. Elster J (1983) Sour grapes: studies in the subversion of rationality. Cambridge University Press, Cambridge
33. Epstein JM (2007) Generative social science. studies in agent-based computational modelling. Princeton University Press, Princeton
34. Fehr E, Fischbacher U (2003) The nature of human altruism. Nature 425:785–791
35. Fikes RE, Kehler T (1985) The role of frame-based representation in knowledge representation and reasoning. Commun ACM 28(9):904–920
36. Fogel DB (ed) (1998) Evolutionary computation: the fossil record. IEEE Press, New York
37. Fox CR, Tversky A (1995) Ambiguity aversion and comparative ignorance. Q J Econ 110(3):585–603
38. Frank MJ (2005) Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and non-medicated parkinsonism. J Cogn Neurosci 17:51–72
39. Frank RH (1987) If homo economicus could choose his own utility function, would he want one with a conscience? Amer Econ Rev 77(4):593–604
40. Franklin GF, Powell JD, Emami-Naeini A (1994) Feedback control of dynamic systems. Addison-Wesley, Reading
41. Franklin S, Graesser A (1997) Is it an agent, or just a program? A taxonomy for autonomous agents. In: Intelligent agents, vol III. Springer, Berlin, pp 21–35
42. Friedman M, Savage L J (1948) The utility analysis of choices involving risk. J Polit Econ 56(4):279–304
43. Fu W-T, Anderson JR (2006) From recurrent choice to skill learning: a reinforcement-learning model. J Exp Psychol Gen 135(2):184–206
44. Genesereth MR, Ketchpel SP (1994) Software agents. Commun ACM 37(7):48–53
45. Gigerenzer G, Selten R (2002) Bounded rationality. MIT Press, Cambridge. reprint edition
46. Goldberg DE (1989) Genetic algorithms in search, optimization and machine learning. Kluwer, Boston
47. Gray WD, Sims CR, Fu W-T, Schoelles MJ (2006) The soft constraints hypothesis: a rational analysis approach to resource allocation for interactive behavior. Psychol Rev 113(3):461–482
48. Green TP, Shapiro I (1994) Pathologies of rational choice theory: A critique of application in political science. Yale University Press, London
49. Hare B, Call J, Tomasello M (2001) Do chimpanzees know what conspecifics know and do not know? Anim Behav 61:139–151
50. Haykin S (1999) Neural networks: a comprehensive foundation. Prentice Hall, Englewood Cliffs
51. Hegselmann R, Krause U (2002) Opinion dynamics and bounded confidence: models, analysis and simulation. J Artif Soc Soc Simul 5(3). http://jasss.soc.surrey.ac.uk/5/3/2.html
52. Holland J (1989) Using classifier systems to study adaptive nonlinear networks. In: Stein DL (ed) Lectures in the sciences of complexity. Addison Wesley, New York
53. Holroyd CB, Coles MGH (2002) The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol Rev 109:679–709
54. Kaelbling LP, Littman ML, Moore AW (1996) Reinforcement learning: a survey. J Artif Intell Res 4:237–285
55. Kahneman D (2003) Maps of bounded rationality: psychology for behavioral economics. Amer Econ Rev 93(5):1449–1475
56. Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. Econometrica XLVII:263–291
57. Kahneman D, Tversky A, Slovic P (eds) (1982) Judgement under uncertainty: heuristics and biases. Cambridge University Press, Cambridge
58. Katsuno H, Mendelzon AO (1991) On the difference between updating a knowledge base and revising it. In: Proc of the 2nd international conference on the principles of knowledge representation and reasoning (KR'91). Morgan Kaufmann, San Matteo, pp 387–394
59. Keynes JM (1921) Treatise on probability. MacMillan, London
60. Kirkpatrick S, Gelatt CD, Vecchi MP (1983) Optimization by simulated annealing. Science 220(4598):671–680. http://www.cs.virginia.edu/cs432/documents/sa-1983.pdf; http://citeseer.ist.psu.edu/kirkpatrick83optimization.html
61. Knight FH (1921) Risk, uncertainty, and profit. In: Hart, Schaffner and Marx prize essays, vol 31. Houghton Mifflin, Boston, New York
62. Koza JR (1990) Genetic programming: a paradigm for genetically breeding populations of computer programs to solve problems. Stanford University Computer Science Department technical report STAN-CS-90-1314
63. Koza JR (1992) Genetic programming: on the programming of computers by means of natural selection. MIT Press, Cambridge
64. Koza JR (1994) Genetic programming II: automatic discovery of reusable programs. MIT Press, Cambridge
65. Koza JR, Bennett FH, Andre D, Keane MA (1999) Genetic programming III: darwinian invention and problem solving. Morgan Kaufmann, San Mateo
66. Koza JR, Keane MA, Streeter MJ, Mydlowec W, Yu J, Lanza G (2003) Genetic programming IV: routine human-competitive machine intelligence. Kluwer, Norwell
67. Kuznar LA, Frederick WG (2003) Environmental constraints and sigmoid utility: implications for value, risk sensitivity, and social status. Ecol Econ 46(2):293–306. September 2003
68. Latané B, Nowak A, Liu JH (1995) Distance matters: physical space and social impact. Pers Soc Psychol Bull 21:795–805
69. Lewin K (1921) Das Problem der Willensmessung und das Grundgesetz der Assoziation, vol I. Psychol Forsch 1:191–302
70. Lewis DK (1969) Convention: a philosophical study. Harvard University Press, Cambridge
71. Lorenz J (2007) Continuous opinion dynamics of multidimensional allocation problems under bounded confidence: More dimensions lead to better chances for consensus. Eur J Econ Soc Syst EJESS 19:213–227. Beliefs, Norms, and Markets

72. Lynch A (1999) Thought contagion: how belief spreads through society by. Basic Books, New York
73. Malarz K (2006) Truth seekers in opinion dynamics models. Int J Mod Phys C 17(10):1521–1524
74. Markman AB (1998) Knowledge representation. Lawrence Erlbaum Associates, Hillsdale
75. Mauss M (2006) The gift. The form and reason for exchange in archaic societies. Routledge, London
76. McFarland D (1995) Opportunity versus goals in robots, animals and people. In: Roitblat HL, Meyer J-A (eds) Comparative approaches to cognitive science. MIT Press, Cambridge, pp 415–433
77. Mill JS (1874) On the definition of political economy; and on the method of investigation proper to it. In: London and Westminster review. Essays on some unsettled questions of political economy, 2nd edn. Longmans, Green, Reader & Dyer, London. 1874, essay 5, paragraphs 38 and 48
78. Miller GA, Galanter E, Pribram KH (1960) Plans and the structures of behaviour. Holt, New York
79. Minsky M (1975) A framework for representing knowledge. In: Winston PH (ed) The psychology of computer vision. McGraw-Hill, New York
80. Nozick R (1993) The nature of rationality. Princeton University Press, Princeton
81. Nwana HS (1996) Software agents: an overview. Knowl Eng Rev 11(3):1–40. September 1996
82. Ostrom E (2000) Collective action and the evolution of social norms. J Econ Pers 14(3):137–158
83. Pearl J (1983) Heuristics: intelligent search strategies for computer problem solving. Addison-Wesley, New York. p vii
84. Perry J (1980) Belief and acceptance. Midwest Stud Philos V:533–542
85. Premack DG, Woodruff G (1978) Does the chimpanzee have a theory of mind? Behav Brain Sci 1:515–526
86. Rao AS, Georgeff MP (1995) BDI agents from theory to practice. In: Proc of the 1st international conference on multi agent systems ICMAS. AAAI Press, MIT Press, San Francisco
87. Ripley BD (1996) Pattern recognition and neural networks. Cambridge University Press, Cambridge
88. Russell SJ, Norvig P (2003) Artificial intelligence: a modern approach, 2nd edn. Prentice Hall, Upper Saddle River
89. Sacerdoti ED (1977) A structure for plans and behavior. Elsevier, New York
90. Salzarulo L (2006) A continuous opinion dynamics model based on the principle of meta-contrast. J Artif Soc Soc Simul 9(1). http://jasss.soc.surrey.ac.uk/9/1/13.html
91. Savage L J (1954) The foundations of statistics. Wiley, New York
92. Schank RC, Abelson RP (1977) Scripts, plans, goals, and understanding: an inquiry into human knowledge structures. Lawrence Erlbaum, Hillsdale
93. Shoham Y (1993) Agent-oriented programming. Artif Intell 60(1):51–92
94. Simon H (1957) A behavioral model of rational choice. In: Models of man, social and rational: mathematical essays on rational human behavior in a social setting. Wiley, New York
95. Sowa JF (2000) Knowledge representation: logical, philosophical, and computational foundations. Brooks/Cole, New York
96. Sun R (2002) Duality of the mind. Erlbaum, New York
97. Wilensky R (1983) Planning and understanding: a computational approach to human reasoning. Advanced book program. Addison-Wesley, Reading
98. Williamson O (1981) The economies of organization: the transaction cost approach. Amer J Sociol 87:548–577
99. Windrum P, Fagiolo G, Moneta A (2007) Empirical validation of agent based models: alternatives and prospects. JASSS J Artif Soc Soc Simul 10(2):8. http://jasss.soc.surrey.ac.uk/10/2/8.html
100. Wooldridge MJ (2000) Reasoning about rational agents. MIT Press, Cambridge
101. Wooldridge M, Jennings NR (1995) Intelligent agents: theory and practice. Knowl Eng Rev 10(2):115–152
102. Wright AH, Vose MD, Rowe JE (2003) Implicit parallelism. In: Cantu-Paz E et al (eds) Proc of GECCO (the genetic and evolutionary computation conference). Lecture notes in computer science. Springer, Berlin, pp 1003–1014

Reaction-Diffusion Computing

Andrew Adamatzky
University of the West of England, Bristol, UK

Article Outline

Glossary

Belousov–Zhabotinsky (BZ) reaction is a chemical reaction where the organic substrate is oxidized by bromate ions in the presence of acid and a one electron transfer redox catalyst. The reaction produces oscillations in well-stirred reactors and traveling waves in thin layers.

Cellular automaton is an array of locally connected finite automata, which update their discrete states in discrete time depending on the states of their neighbors; all automata of the array update their states in parallel.

Collision-based computer is a uniform homogeneous medium which employs mobile compact patterns (particles, wave fragments) which travel in space and

perform computation (e. g. implement logical gates) when they collide with each other. Truth values of logical variables are given by either the absence or presence of a localization or by various types of localizations.

Excitable medium is spatially distributed assembly of coupled excitable systems; spatial distribution and coupling allow for propagation of excitation waves.

Excitable system is a system with a single steady quiescent state that is stable to small perturbations, but responds with an excursion from its quiescent state (excitation event) if the perturbation is above a critical threshold level. After excitation the system enters a refractory period during which time it is insensitive to further excitation before returning to its steady state.

Glider as related to cellular automata, is a compact (neither infinitely expanding nor collapsing) pattern of non-quiescent states that travels along the cellular-automaton lattice.

Image processing is a transformation of an input image to an output image with desirable properties, using manipulation of images to enhance or extract information.

Logical gate is an elementary building block of a digital, or logical, circuit, which represents (mostly) binary logical operations, e. g. AND, OR, XOR, with two input terminals and one output terminal. In Boolean logic terminals are in one of two binary conditions (e. g. low voltage and high voltage) corresponding to TRUE and FALSE values of logical variables.

Logically universal processor is a system which can realize a functionally complete set of logical operations in its development, e. g. conjunction and negation.

Oregonator is a system of three (or two in a modified version) coupled differential equations aimed to simulate oscillatory phenomena in the **Belousov–Zhabotinsky reaction**.

Shortest-path problem is the problem of finding a path between two sites (e. g. vertices of the graph, locations in space) such that the length (sum of the weights of the graph edges or travel distances) of the path is minimized.

Skeleton of a planar contour is a set of centers of bi-tangent circles lying inside the contour.

Subexcitable medium is a medium whose steady state lies between the excitable and the unexcitable domains. In excitable media waves initiated by perturbations of a sufficient size propagate throughout the media. In an unexcitable medium no perturbation is large enough to trigger a wave. In a subexcitable medium wave fragments with open ends are formed.

Voronoi diagram of a planar set **P** of planar points is a partition of the plane into regions, each for any element of **P**, such that a region corresponding to a unique point p contains all those points of the plane that are closer to p than to any other node of **P**.

Wave fragment is an excitation wave formed in a subexcitable medium; this is a segment with free ends, which either expand or contract, depending on their size and the medium's excitability.

Definition of the Subject

A reaction-diffusion computer is a spatially extended chemical system, which processes information using interacting growing patterns, excitable and diffusive waves. In reaction-diffusion processors, both the data and the results of the computation are encoded as concentration profiles of the reagents. The computation is performed via the spreading and interaction of wave fronts.

A reaction-diffusion computer is a thin layer of a reagent mixture which reacts to changes of one reagent's concentration—data configuration—in a predictable way to form a stationary pattern corresponding to the concentration of the reagent—result configuration. A computation in the chemical processor is implemented via the spreading and interaction of diffusive or phase waves [1,7].

The reaction-diffusion computers are parallel because myriads of their micro-volumes update their states simultaneously, and molecules diffuse and react in parallel. Liquid-phase chemical media are wet-analogs of massive-parallel (millions of elementary processors in a small chemical reactor) and locally connected (every micro-volume of the medium changes its state depending on states of its closest neighbors) processors. They have parallel input and outputs, e. g. optical input—control of initial excitation dynamics by illumination masks, output is parallel because the concentration profile representing the results of computation is visualized by indicators. The reaction-diffusion computers show fault-tolerance and are capable of automatic reconfiguration, namely if we remove some quantity of the computing substrate, the topology is restored almost immediately.

Introduction

Reaction-diffusion computers are based on three principles of physics-inspired computing. First, physical action measures amount of information: we exploit active processes in non-linear systems and interpret dynamics of the systems as computation. Second, physical information travels only a finite distance: this means that computation is local and we can assume the non-linear medium is

a spatial arrangement of elementary processing units connected locally, i. e. each unit interacts with closest neighbors. Third, nature is governed by waves and spreading patterns: computation is therefore spatial.

Surface tension, propagating waves, electricity and chemical reactions have been principal "engines" of nature-inspired computers for over two centuries [1]. However, only reaction-diffusion computers utilize all these phenomena at once, to solve large-scale NP-complete problems in a parallel and stable manner. Experimental studies and designs of reaction-diffusion computers could be traced back to the pioneer discovery of Kuhnert [28]. In 1986 he demonstrated that some very basic image transformations can be implemented in a light-sensitive Belousov–Zhabotinsky system [28]. The ideas by Kuhnert, Krinsky and Agladze [29] on image and planar shape transformations in two-dimensional excitable chemical medium were further developed and modified by Rambidi with colleagues, see e. g. [42,43]. At that time, mid and late 1990s, a range of chemical logical gates was experimentally built in the Showalter [52] and Yoshikawa laboratories [30]. Computation of shorter, one of classical optimization problems, has also been implemented in these laboratories using the Belousov–Zhabotinsky medium [7,10,51].

Reaction-diffusion computers give us the best examples of unconventional computers, they feature, following Jonathan Mills' classification of conventional vs. unconventional [34]: wet-ware, non-silicon computing substrate; parallel processing; computation occurring everywhere in substrate space; computation is based on analogies; spatial increase in precision; holistic and spatial programming; visual structure; implicit error correcting. A theory of reaction-diffusion computing was established and a range of practical applications outlined in [1]. Recent discoveries were published in a collective monograph [7].

Herein we will provide an account of achievements in reaction-computing obtained in the latter 1990s and early 2000s. Our present article in no way serves as a substitute for these books but rather an introduction to the field and case study of several characteristic examples. We give a case-study introduction to the novel paradigm of wave-based computing in chemical systems. We show how selected problems and tasks of computational geometry, robotics and logics can be solved by encoding data in configuration of chemical medium's disturbances and programming wave dynamics and interaction. We describe and analyze working reaction-diffusion algorithms for image processing, computational geometry, logical and arithmetical circuits, memory devices, path planning and robot navigation, and control of massive parallel actuators.

Computational Geometry

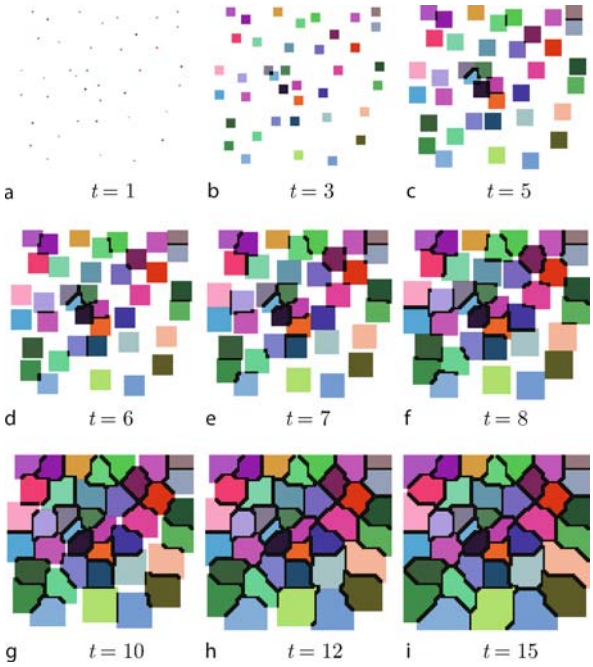
The Voronoi diagram, or plane subdivision, is our most favorable NP-complete problem, for demonstrating "mechanics" and computational complexity of reaction-diffusion computers.

Let $\mathbf{P}$ be a nonempty finite set of planar points. A planar Voronoi diagram of the set $\mathbf{P}$ is a partition of the plane onto such regions, that for any element of $\mathbf{P}$, a region corresponding to a unique point p contains all those points of the plane which are closer to p than to any other node of $\mathbf{P}$. A unique region $\mathrm{vor}(p) = \{z \in \mathbf{R}^2 : d(p,z) < d(p,m) \quad \forall\ m \in \mathbf{R}^2,\ m \neq z\}$ assigned to point p is called a Voronoi cell of the point p. The boundary of the Voronoi cell of a point p is built of segments of bisectors separating pairs of geographically closest points of the given planar set $\mathbf{P}$. A union of all boundaries of the Voronoi cells determines the *planar Voronoi diagram*: $\mathrm{VD}(\mathbf{P}) = \cup_{p \in \mathbf{P}} \partial \mathrm{vor}(p)$. A variety of Voronoi diagrams and algorithms of their construction can be found in [27].

The basic concept of constructing Voronoi diagrams with reaction-diffusion systems is based on a very simple intuitive technique for detecting the bisector points separating two given points of the set $\mathbf{P}$. If we drop reagents at the two data points the diffusive waves, or phase waves if the computing substrate is active, spread outwards from the drops with the same speed. The waves travel the same distance from the sites of origination before they meet one another. The points, where the waves meet, are the bisector points. This idea of a Voronoi diagram computation was originally implemented in cellular automata models and in experimental parallel chemical processors, see the extensive bibliography in [1,7].

Assuming that the computational space is homogeneous and locally connected, and every site (micro-volume of the chemical medium or cell of the automaton array) is coupled to its closest neighbors by the same diffusive links, we can easily draw a parallel between distance and time, and thus put our wave-based approach into action. In a cellular-automaton representation of the physical reality cell neighborhood u determines that all processes in the cellular automata model are constrained to the discrete metric L_∞. So, when studying automata models we should think rather about the discrete Voronoi diagram than its Euclidean representation. Chemical laboratory prototypes of reaction-diffusion computers do approximate a continuous Voronoi diagram as we will see further.

A discrete Voronoi diagram can be defined on lattices or arrays of cells where, e. g. a two-dimensional lattice $\mathbf{Z}^2$. The distance $d(\cdot,\cdot)$ is calculated not in Euclidean but in one of the discrete metrics, e. g. L_1 and L_∞. A dis-



Reaction-Diffusion Computing, Figure 1
Computation of a Voronoi diagram in a cellular-automaton model of a chemical processor with *O*(*n*) reagents. Precipitate is shown in *black*

crete bisector of nodes x and y of $\mathbf{Z}^2$ is determined as $B(x,y) = \{z \in \mathbf{Z}^2 : d(x,z) = d(y,z)\}$. However, following such definition we sometimes generate bisectors that fill a quarter of the lattices or produce no bisector at all [1]. If we want the constructed diagrams to be closer to the real-world then we could re-define the discrete bisector as follows $B(x,y) = \{z \in \mathbf{Z}^2 : |d(x,z) - d(y,z)| \leq 1\}$. The redefined bisector will comprise edges of Voronoi diagrams constructed in discrete, cellular-automaton models of reaction-diffusion and excitable media.

Now we will discuss several versions of reaction-diffusion wave-based constructions of Voronoi diagrams, from a naïve model, where the number of reagents grow proportionally to the number of data points, to a minimalist implementation with just one reagent and one substrate [1].

Let us start with the $O(n)$-reagent model. In a naïve version of reaction-diffusion computation of a Voronoi diagram one needs two reagents and a precipitate to mark a bisector separating two points. Therefore $n+2$ reagents, including precipitate and substrate, are required to approximate a Voronoi diagram of n points. We place n unique reagents on n points of the given data set $\mathbf{P}$, waves of these reagents spread around the space and interact with each other where they meet. When at least two different reagents meet at the same or adjacent sites of the space, they react and form a precipitate—sites that contain the precipitate represent edges of the Voronoi cell, and therefore constitute the Voronoi diagram. In "chemical reaction" equations the idea looks as follows, α and β are different reagents and # is a precipitate: $\alpha + \beta \rightarrow \#$. This can be converted to a cellular-automaton interpretation as follows:

$$x^{t+1} = \begin{cases} \rho\,, & \text{if} \quad x^t = \bullet \quad \text{and} \quad \Psi(x)^t \subset \{\rho, \bullet\} \\ \#\,, & \text{if} \quad x^t \neq \# \quad \text{and} \quad |\Psi(x)^t/\#| > 1 \\ x^t\,, & \text{otherwise} \end{cases}$$

where $\bullet$ is a resting state (cell in this state does not contain any reagents), $\rho \in \mathbf{R}$ is a reagent from the set $\mathbf{R}$ of n reagents, and $\Psi(x)^t = \{y^t : y \in u(x)\}$ characterizes what reagents are present in the local neighborhood $u(x)$ of the cell x at time step t.

The first transition of the above rule symbolizes diffusion. A resting cell takes the state ρ if only this reagent is present in the cell's neighborhood. If there are two different reagents in the cells neighborhood then the cell takes the precipitate state #. Diffusing reagents halt because formation of precipitate reduces the number of "vacant" resting cells. The precipitate does not diffuse. Cell in state # remain in this state indefinitely. An example of a cellular-automaton simulation of $O(n)$-reagent chemical processor is shown in Fig. 3.

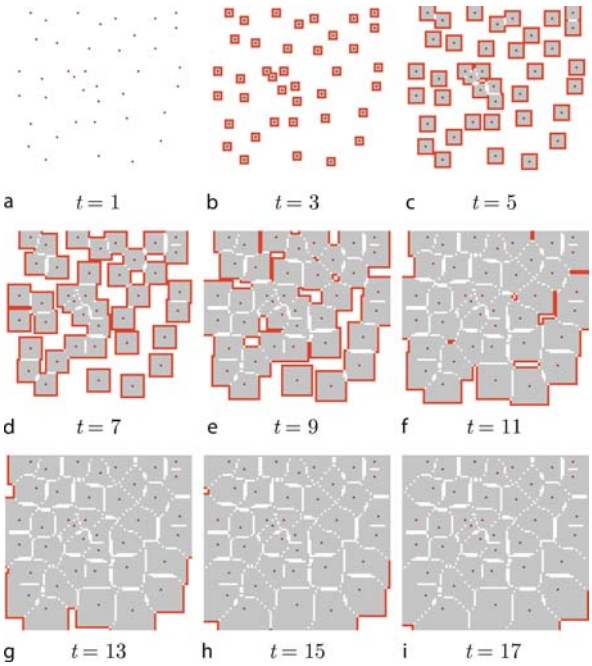
The $O(n)$-reagent model is demonstrative, however, computationally inefficient. Clearly we can reduce the number of reagents to four—using map coloring theorems—but pre-processing time will be unfeasibly high. The number of participating reagents can be sufficiently reduced to $O(1)$ when the topology of the spreading waves is taken into account [1].

Now we go from one extreme to another and consider a model with just one reagent, and a substrate. The reagent α diffuses from sites corresponding to points of a planar data set $\mathbf{P}$. When two diffusing wave fronts meet a super-threshold concentration of reagents they do not spread further. A cellular-automaton model represents this as follows.

Every cell has two possible states: resting or substrate state $\bullet$ and reagent state α. If the cell is in state α it remains in this state indefinitely. If the cell is in state $\bullet$ and between one and four of its neighbors are in state α, then the cell takes the state α. Otherwise, the cell remains in the state $\bullet$—this reflects the "super-threshold inhibition" idea. A cell state transition rule is as follows:

$$x^{t+1} = \begin{cases} \alpha\,, & \text{if} \quad x^t = \bullet \quad \text{and} \quad 1 \leq \sigma(x)^t \leq 4 \\ x^t\,, & \text{otherwise} \end{cases}$$

where $\sigma(x)^t = |y \in u(x) : y^t = \alpha|$.



Reaction-Diffusion Computing, Figure 2
An example of Voronoi diagram computing in an automaton model of reaction-diffusion medium with one reagent and one substrate. Reactive parts of wave fronts are shown in *black*. The precipitate is *gray* and the edges of the Voronoi diagram are *white*

Increasing the number of reagents to two (one reagent and one precipitate) would make life easy. A reagent β diffuses on a substrate, from the initial points (drop of reagent) of **P**, and forms a precipitate in the reaction $m\beta \to \alpha$, where $1 \le m \le 4$.

Every cell takes three states: $\bullet$ (resting cell, no reagents), α (e. g. colored precipitate) and β (reagent). The cell updates its states by the rule:

$$x^{t+1} = \begin{cases} \beta\,, & \text{if} \quad x^t = \bullet \quad \text{and} \quad 1 \le \sigma(x)^t \le 4 \\ \alpha\,, & \text{if} \quad x^t = \beta \quad \text{and} \quad 1 \le \sigma(x)^t \le 4 \\ x^t\,, & \text{otherwise} \end{cases}$$

where $\sigma(x)^t = |y \in u(x)\colon y^t = \beta|$.

An example of a Voronoi diagram computed in an automaton model of a reaction-diffusion medium with one reagent and one substrate is shown in Fig. 2.

By increasing the number of cell-state and enlarging cell neighborhood in the cellular automaton model we can produce more realistic—almost perfectly matching outcomes of chemical laboratory experiments—Voronoi diagrams (Fig. 3).

Let us consider the following model. Cells of the automaton take their state from interval $[\rho, \alpha]$, where ρ is a minimum refractory value, and α is a maximum excitation value; $\rho = -2$ and $\alpha = 5$ in our experiments. Cell x's state transitions are strongly determined by normalized local excitation $\sigma_x^t = \sum_{y \in u_x} y^t / \sqrt{(|u_x|)}$. Every cell x updates its state at time $t + 1$, depending on its state x^t and state u_x^t of its neighborhood u_x—in experiments we used 15×15 cell neighborhood—as follows:

$$x^{t+1} = \begin{cases} \alpha\,, & \text{if} \quad x^t = 0 \quad \text{and} \quad \sigma_x^t \ge \alpha \\ 0, & \text{if} \quad x^t = 0 \quad \text{and} \quad \sigma_x^t < \alpha \\ x^t + 1\,, & \text{if} \quad x^t < 0 \\ x^t - 1\,, & \text{if} \quad x^t > 1 \\ \rho, & \text{if} \quad x^t = 1\,. \end{cases}$$

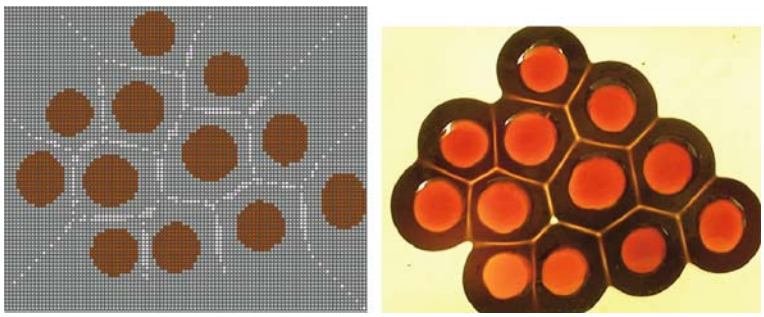
This rule represents spreading of "excitation", or simply phase wave-fronts, in computational space, interaction and annihilation of the wave-fronts. To allow the reaction-diffusion computer to "memorize" sites of wave collision we add a precipitate state p_x^t. Concentration p_x^t of precipitate at site x at moment t is calculated as $p_x^{t+1} \sim |\{y \in u_x\colon y^t = \alpha\}|$.

As shown in Fig. 4 the model represents cellular-automaton Voronoi diagrams in "unlike phase" with experimental chemical representation of the diagram. Sites of higher concentration of precipitate in cellular-automaton configurations correspond to sites with lowest precipitate concentration in experimental processors.

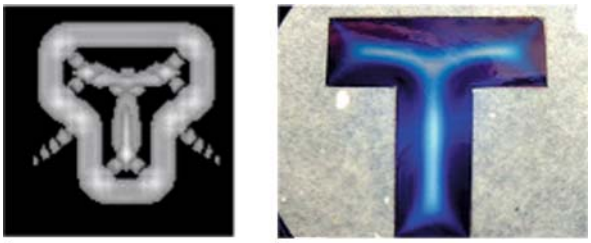
Logical Universality

Certain families of thin-layer reaction-diffusion chemical media can implement sensible transformation of initial (data) spatial distribution of chemical species concentrations to a final (result) concentration profile [1,50]. In these reaction-diffusion computers a computation is realized via spreading and interaction of diffusive or phase waves. Specialized, intended to solve a particular problem, experimental chemical processors implement basic operations of image processing [7,29,42,43], computation of optimal paths [7,10,51] and control of mobile robots [7].

A device is called computationally universal if it implements a functionally complete system of logical gates, e. g. a tuple of negation and conjunction, in its space-time dynamics. A number of computationally universal reaction-diffusion devices were implemented, the findings include logical gates [49,52] and diodes [18,30,35] in the Belousov–Zhabotinsky (BZ) medium, and the XOR gate in the palladium processor [4].



Reaction-Diffusion Computing, Figure 3
Planar Voronoi diagram computed in cellular automaton (*left*) and palladium reaction-diffusion chemical processor (*right*) [7]



Reaction-Diffusion Computing, Figure 4
Skeleton (internal Voronoi diagram) of planar T-shape constructed in multi-state cellular-automaton model (a) and chemical laboratory Prussian blue reaction-diffusion processor (b) [11]

The most known so far experimental prototypes of reaction-diffusion processors exploit interaction of wave fronts in a geometrically constrained chemical medium, i. e. the computation is based on a stationary architecture of a medium's inhomogeneities. Constrained by stationary wires and gates reaction-diffusion chemical universal processors pose a little computational novelty and no dynamical reconfiguration ability because they simply imitate architectures of conventional silicon computing devices. To appreciate in full massive-parallelism of thin-layer chemical media and to free the chemical processors from limitations of fixed computing architectures we adopt an unconventional paradigm of architecture-less, or collision-based, computing. An architecture-based, or stationary, computation implies that a logical circuit is embedded into the system in such a manner that all elements of the circuit are represented by the system's stationary states. The architecture is static. If there is any kind of "artificial" or "natural" compartmentalization the medium is classified as an architecture-based computing device. Personal computers, living neural networks, cells, and networks of chemical reactors are typical examples of architecture-based computers.

A collision-based, or dynamical, computation employs mobile compact finite patterns, mobile self-localized excitations or simply localizations, in active non-linear medium. Essentials of collision-based computing are as follows. Information values (e. g. truth values of logical variables) are given by either absence or presence of the localizations or other parameters of the localizations. The localizations travel in space and do computation when they collide with each other. There are no predetermined stationary wires, a trajectory of the traveling pattern is a momentary wire. Almost any part of the medium space can be used as a wire. Localizations can collide anywhere within a space sample, there are no fixed positions at which specific operations occur, nor location specified gates with fixed operations. The localizations undergo transformations, form bound states, annihilate or fuse when they interact with other mobile patterns. Information values of localizations are transformed as a result of collision and thus a computation is implemented [2].

The paradigm of collision-based computing originates from the technique of proving computational universality of the Game of Life [14], conservative logic and billiard-ball model [21] and their cellular-automaton implementations [32].

Solitons, defects in tubulin microtubules, excitons in Scheibe aggregates and breather in polymer chains are the most frequently considered candidates for a role of information carrier in nature-inspired collision-based computers, see overview in [1]. It is experimentally difficult to reproduce all these artifacts in natural systems, therefore the existence of mobile localizations in an experiment-friendly chemical media would open new horizons for fabrication of collision-based computers.

The basis for material implementation of collision-based universality of reaction-diffusion chemical media was discovered by Sendiña-Nadal et al. [48]. They experimentally proved the existence of localized excita-

tions—traveling wave fragments which behave like quasi-particles—in a photosensitive sub-excitable Belousov–Zhabotinsky medium.

We show how logical circuits can be fabricated in a sub-excitable BZ medium via collisions between traveling wave fragments. While implementation collision-based logical operations is relatively straightforward [7], more attention should be paid to control of signal propagation in the homogeneous medium. It has been demonstrated that applying light of varying intensity we can control the excitation dynamic in a Belousov–Zhabotinsky medium [13,23,39], wave velocity [46], and pattern formation [55]. Of particular interest are experimental evidences of light-induced back propagating waves, wave-front splitting and phase shifting [63]; we can also manipulate a medium's excitability by varying the intensity of the medium's illumination [16]. On the basis of these facts we show how to control signal-wave fragments by varying geometric configuration of excitatory and inhibitory segments of impurity-reflectors.

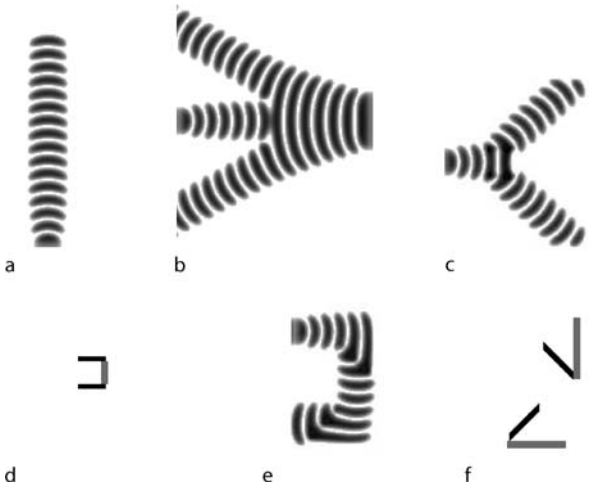
We built our model on a two-variable Oregonator equation [19,53] adapted to a light-sensitive BZ reaction with applied illumination [13]:

$$\frac{\partial u}{\partial t} = \frac{1}{\epsilon}\left(u - u^2 - (fv + \phi)\frac{u - q}{u + q}\right) + D_u \nabla^2 u$$

$$\frac{\partial v}{\partial t} = u - v$$

where variables u and v represent local concentrations of bromous acid $HBrO_2$ and the oxidized form of the catalyst ruthenium Ru(III), ϵ sets up a ratio of time scale of variables u and v, q is a scaling parameter depending on reaction rates, f is a stoichiometric coefficient, ϕ is a light-induced bromide production rate proportional to intensity of illumination (an excitability parameter—moderate intensity of light will facilitate the excitation process, a higher intensity will produce excessive quantities of bromide which suppresses the reaction). We assumed that the catalyst is immobilized in a thin-layer of gel, therefore there is no diffusion term for v. To integrate the system we used the Euler method with a five-node Laplasian operator, time step $\Delta\ t = 10^{-3}$ and grid point spacing $\Delta\ x = 0.15$, with the following parameters: $\phi = \phi_0 + A/2$, $A = 0.0011109$, $\phi_0 = 0.0766$, $\epsilon = 0.03$, $f = 1.4$, $q = 0.002$.

The chosen parameters correspond to a region of "higher excitability of the sub-excitability regime" outlined in [48] (see also how to adjust f and q in [41]) that supports propagation of sustained wave fragments (Fig. 5a). These wave fragments are used as quanta of information in our design of collision-based logical circuits. The waves



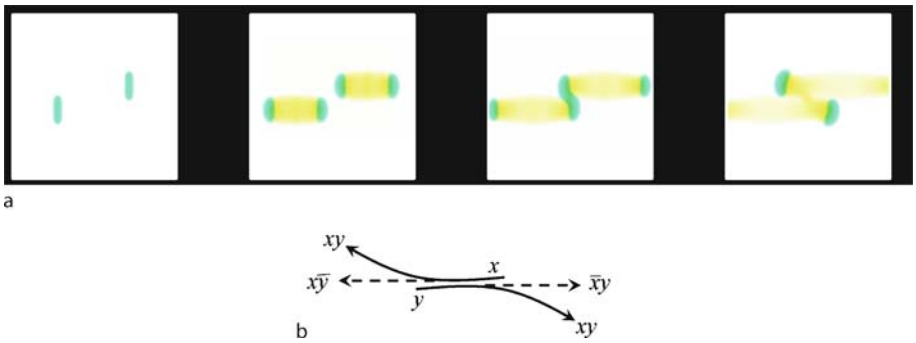
Reaction-Diffusion Computing, Figure 5

Basic operations with signals. Overlay of images taken every 0.5 time units. Exciting domains of impurities are shown in black, inhibiting domains of impurities are shown in gray. a Wave fragment traveling north. b Signal branching without impurities: a wave fragment traveling east splits into two wave fragments (traveling south-east and north-east) when it collides with a smaller wave fragment traveling west. c Signal branching with impurity: wave fragment traveling west is split by impurity (d) into two waves traveling north-west and south-west. e Signal routing (U-turn) with impurities: wave fragment traveling east is routed north and then west by two impurities (f). An impurity-reflector consists of inhibitory (*gray*) and excitatory (*black*) chains of grid sites

were initiated by locally disturbing initial concentrations of species, e. g. ten grid sites in a chain are given the value $u = 1.0$ each, this generated two or more localized wave fragments, similarly to counter-propagating waves induced by temporary illumination in experiments [63]. The traveling wave fragments keep their shape for around $4 \cdot 10^3$–10^4 steps of simulation (4–10 time units), then decrease in size and vanish. The wave's life-time is sufficient however to implement logical gates; this also allows us not to worry about "garbage collection" in the computational medium.

We model signals by traveling wave fragments [13,48]: a relatively stable propagating wave fragment (Fig. 5a) represents a TRUE value of a logical variable corresponding to the wave's trajectory (momentarily wire).

To demonstrate that a physical system is logically universal it is enough to implement negation and conjunction or disjunction in space-time dynamics of the system. To realize a fully functional logical circuit we must also know how to operate input and output signals in the system's dynamics, namely to implement signal branching and routing; delay can be realized via appropriate routing.



Reaction-Diffusion Computing, Figure 6
Implementation of conservative gate in Belousov–Zhabotinsky system. a Elastic collision of two wave-fragments, one traveling West another East. The fragments change the direction of their motion to North-West and South-East, respectively, as a result of the collision. b Scheme of the gate. In a logical variables are represented as $x = 1$ and $y = 1$

We can branch a signal using two techniques. Firstly, we can collide a smaller auxiliary wave to a wave fragment representing the signal, the signal-wave will split then into two signals (these daughter waves shrink slightly down to stable size and then travel with constant shape further $4 \cdot 10^3$ time steps of the simulation) and the auxiliary wave will annihilate (Fig. 5b).

Secondly, we can temporarily and locally apply illumination impurities on a signal's way to change properties of the medium and thus cause the signal to split (Fig. 5c,d). We must mention, it was already demonstrated in [63], that a wave front influenced by strong illumination (inhibitory segments of the impurity) splits and its ends do not form spirals, as in typical situations of excitable media.

A control impurity, or reflector, consists of a few segments of sites where the illumination level is slightly above or below the overall illumination level of the medium. Combining excitatory and inhibitory segments we can precisely control the wave's trajectory, e. g. realize a U-turn of a signal (Fig. 5e,f).

A typical billiard-ball model interaction gate [21,32] has two inputs—x and y, and four outputs—$x\bar{y}$ (ball x moves undisturbed in the absence of ball y), $\bar{x}y$ (ball y moves undisturbed in the absence of ball x), and twice xy (balls x and y change their trajectories when they collide with each other). Such a conservative interaction gate can be implemented via elastic collision of the wave-fragment, see Fig. 6.

The elastic collision is not particularly common in laboratory prototypes of chemical systems, more often interacting waves either fuse or one of the waves annihilates as a result of the collision with another wave. This leads to a non-conservative version of the interaction gate with two inputs and three outputs, i. e. just one xy output instead of two. Such a collision gate is shown in Fig. 7.

Reaction-Diffusion Computing, Figure 7
Two wave fragments undergo angle collision and implement interaction gate $\langle x, y\rangle \rightarrow \langle x\bar{y}, xy, \bar{x}y\rangle$. a In this example $x = 1$ and $y = 1$, both wave fragments are present initially. Overlay of images taken every 0.5 time units. b Scheme of the gate. In *upper left* and *bottom left* corners of a we see domains of wave generation, two echo wave fragments are also generated, they travel outwards from the gate area and thus do not interfere with computation

The rich dynamic of a sub-excitable Belousov–Zhabotinsky medium allows us also to implement complicated logical operations just in a single interaction event, for details see [7].

Memory

Memory in chemical computers can be represented in several ways as follows. In precipitating systems any site with precipitate is a memory element. However, they are not rewritable. In "classical" excitable chemical systems, like Belousov–Zhabotinsky dynamics one can construct memory as a configuration of sources of spiral or target waves. We used this technique to program movement of a wheeled robot controlled by an on-board chemical reactor with the Belousov–Zhabotinsky system [7]. The method has the same drawback as precipitating memory—as soon as reaction space is divided by spiral or target waves, it is quite

difficult if not impossible to sensibly move the source of the waves. This is only possible with external inhibition or complete reset of the medium.

In geometrically constrained excitable chemical medium, as demonstrated in [36], we can employ old-time techniques of storing information in induction coils and other types of electrical circuits, i. e. dynamical memory. A ring with an input channel is prepared from the reaction substrate. The ring is broken by a small gap and the input is also separated from the ring with a gap of similar width [36], the gaps play the role of one-way gates to prevent excitation from spreading backwards. The waves enter the ring via the input channel and travel along the ring "indefinitely" (as long as the substrate lasts) [36]. The approach aims to split the reaction-diffusion system into many compartments, and thus does not fit our paradigm of computing in a uniform medium.

In our search for real-life chemical systems exhibiting both mobile and stationary localizations we discovered a cellular-automaton model [58] of an abstract activator-inhibitor reaction-diffusion system, which ideally fits the framework of the collision-based computing paradigm and reaction-diffusion computing. The phenomenology of the automaton was discussed in detail in our previous work [58], therefore in the present paper we draw together the computational properties of the reaction-diffusion cellular hexagonal automaton. The automaton imitates spatio-temporal dynamics of the following reaction equations:

$$\begin{array}{llll} A + 6S \to A & A + I \to I & & A + 3I \to I \\ & & A + 2I \to S & 2A \to I \\ & & 3A \to A & \beta A \to I \\ & & & I \to S\,. \end{array}$$

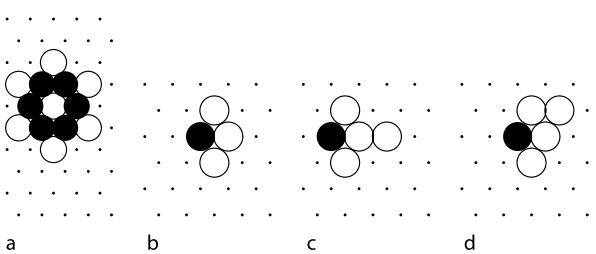
Each cell of the automaton takes three states – substrate S, activator A and inhibitor I. Adopting formalism from [9], we represent the cell-state transition rule as a matrix $\mathbf{M} = (m_{ij})$, where $0 \le i \le j \le 7, 0 \le i + j \le 7$, and $m_{ij} \in \{I, A, S\}$. The output state of each neighborhood is given by the row-index i, the number of neighbors in cell-state I, and column-index j (the number of neighbors in cell-state A). We do not have to count the number of neighbors in cell-state S, because it is given by $7 - (i + j)$. A cell with a neighborhood represented by indexes i and j will update to cell-state M_{ij} which can be read off the matrix. In terms of the cell-state transition function this can be presented as follows: $x^{t+1} = M_{\sigma_2(x)^t \sigma_1(x)^t}$, where $\sigma_i(x)^t$ is a sum of cell x's neighbors in state i, $i = 1, 2$, at time step t. The exact matrix structure, which corresponds to matrix M_3 in [58], is as follows:

$$M = \left\{ \begin{array}{cccccccc} S & A & I & A & I & I & I & I \\ S & I & I & A & I & I & I & \\ S & S & I & A & I & I & & \\ S & I & I & A & I & & & \\ S & S & I & A & & & & \\ S & S & I & & & & & \\ S & S & & & & & & \\ S & & & & & & & \end{array} \right\}.$$

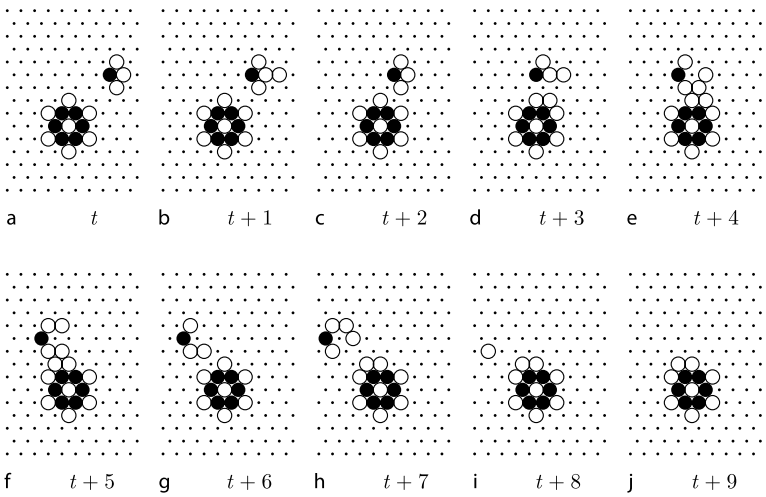
The cell-state transition rule reflects the nonlinearity of activator-inhibitor interactions for sub-threshold concentrations of the activator. Namely, for a small concentration of the inhibitor and for threshold concentrations, the activator is suppressed by the inhibitor, while for critical concentrations of the inhibitor both inhibitor and activator dissociate producing the substrate. In exact words, $M_{01} = A$ symbolizes the diffusion of activator A, $M_{11} = I$ represents the suppression of activator A by the inhibitor I, $M_{z2} = I$ $(z = 0, \ldots, 5)$ can be interpreted as self-inhibition of the activator in particular concentrations. $M_{z3} = A$ $(z = 0, \ldots, 4)$ means a sustained excitation under particular concentrations of the activator. $M_{z0} = S$ $(z = 1, \ldots, 7)$ means that the inhibitor is dissociated in absence of the activator, and that the activator does not diffuse in sub-threshold concentrations. And, finally, $M_{zp} = I$, $p \ge 4$ is an upper-threshold self-inhibition.

Amongst non-trivial localizations, see full "catalog" in [5], found in the medium we selected eaters gliders G_4 and G_{34}, mobile localizations with activator head and inhibitor tail, and eaters E_6, stationary localizations transforming gliders colliding with them, as components of the memory unit.

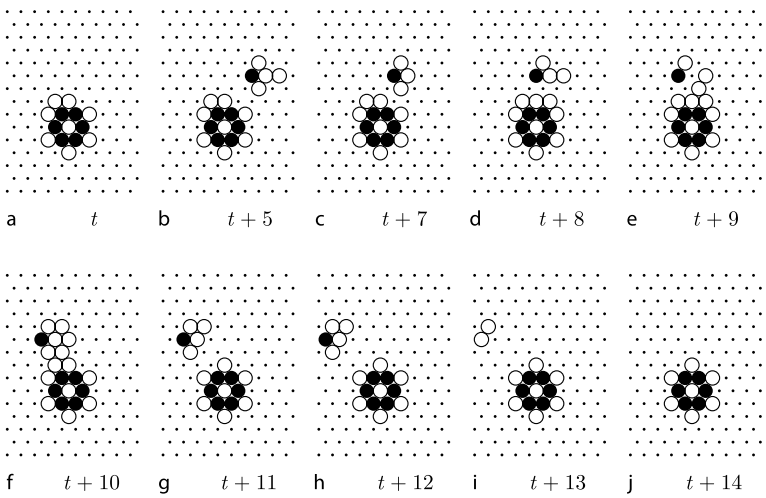
The eater E_6 can play the role of a 6-bit flip-flop memory device. The substrate-sites (bit-down) between inhibitor-sites (Fig. 8) can be switched to an inhibitor-



Reaction-Diffusion Computing, Figure 8
Localizations in reaction-diffusion hexagonal cellular-automaton. Cells with inhibitor *I* are empty circles, cells with activator *A* are black disks. a Stationary localization eater E_6, b–c two forms of glider G_{34}, d glider G_4 [5]



Reaction-Diffusion Computing, Figure 9
Write bit [5]



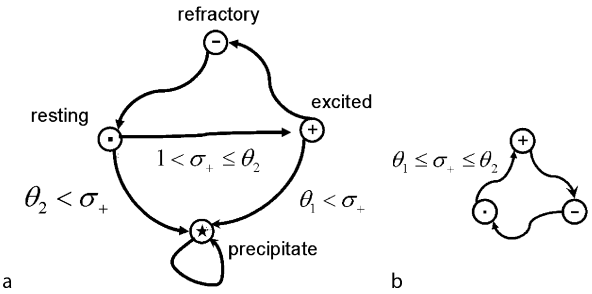
Reaction-Diffusion Computing, Figure 10
Read and erase bit [5]

state (bit-up) by a colliding glider. An example of writing one bit of information in E_6 is shown in Fig. 9. Initially E_6 stores no information. We aim to write one bit in the substrate-site between the northern and north-western inhibitor-sites (Fig. 9a). We generate a glider G_{34} (Fig. 9b,c) traveling West. G_{34} collides with (or brushes past) the North edge of E_6 resulting in G_{34} being transformed to a different type of glider, G_4 (Fig. 9g,h). There is now a record of the collision—evidence that writing was successful. The structure of E_6 now has one site (between the northern and north-western inhibitor-sites) changed to an inhibitor-state (Fig. 9j) —a bit was saved [5].

To read a bit from the E_6 memory device with one bit-up (Fig. 10a), we collide (or brush past) with glider G_{34} (Fig. 10b). Following the collision, the glider G_{34} is transformed into a different type of basic glider, G_{34} (Fig. 10g), and the bit is erased (Fig. 10j).

Programmability

In the chemical laboratory the term programmabilty means controllability. How real chemical systems can be controlled? The majority of the literature, related to theoretical and experimental studies concerning the controllability of the reaction-diffusion medium, deals with application of an electric field. For example, in a thin-layer Belousov–Zhabotinsky reactor stimulated by an electric field the following phenomena are observed: the velocity of ex-



Reaction-Diffusion Computing, Figure 11
Cell state transition diagrams: a model of precipitating reaction-diffusion medium, b model of excitable system

citation waves is increased by a negative and decreased by a positive electric field, a wave is split into two waves that move in opposite directions if a very high electric field is applied across the evolving medium, crescent waves are formed not commonly observed in the field absent evolution of the BZ reaction, stabilization and destabilization of wave fronts, see [7].

Other control parameters may include temperature (to e. g. program transitions between periodic and chaotic oscillations), substrate's structure (controlling formation, annihilation and propagation of waves), and illumination (inputting data and routing signals in light-sensitive chemical systems).

Let us demonstrate a concept of control-based programmability in models of reaction-diffusion processors. Firstly, we show how to adjust reaction rates in chemical medium to make it perform a computation of a Voronoi diagram over a set of given points. Secondly, we show how to switch an excitable system between specialized-processor and universal-processor modes, see [7] for additional examples and details.

Let a cell x of a two-dimensional lattice take four states: resting $\circ$, excited $(+)$, refractory $(-)$ and precipitate $\bigstar$, and update their states in discrete time t depending on the number $\sigma^{t(x)}$ of excited neighbors in its eight-cell neighborhood as follows (Fig. 11a).

A resting cell x becomes excited if $0 < \sigma^t(x) \le \theta_2$ and precipitates if $\theta_2 < \sigma^t(x)$.

An excited cell "precipitates" if $\theta_1 < \sigma^t(x)$ or otherwise becomes refractory.

A refractory cell recovers to the resting state unconditionally, and the precipitate cell does not change its state.

Initially we perturb the medium, excite it in several sites, thus inputting data. Waves of excitation are generated, they grow, collide with each other and annihilate as a result of the collision. They may form a stationary inactive concentration profile of a precipitate, which represents the result of the computation. Thus, we can only be concerned with reactions of precipitation: $+ \to^{k_1} \star$ and $\circ \boxplus + \to^{k_2} \star$, where k_1 and k_2 are inversely proportional to θ_1 and θ_2, respectively. Varying θ_1 and θ_2 from 1 to 8, and thus changing precipitation rates from the maximum possible to the minimum, we obtain various kinds of precipitate patterns, as shown in Fig. 12.

Precipitate patterns developed for relatively high ranges of reaction rates ($3 \le \theta_1, \theta_2 \le 4$) represent discrete Voronoi diagrams (a given "planar" set, represented by sites of initial excitation, is visible in pattern $\theta_1 = \theta_2 = 3$ as white dots inside the Voronoi cells) derived from the set of initially excited sites, see Fig. 13a and b. This example demonstrates that by externally controlling precipitation rates we can force the reaction-diffusion medium to compute a Voronoi diagram.

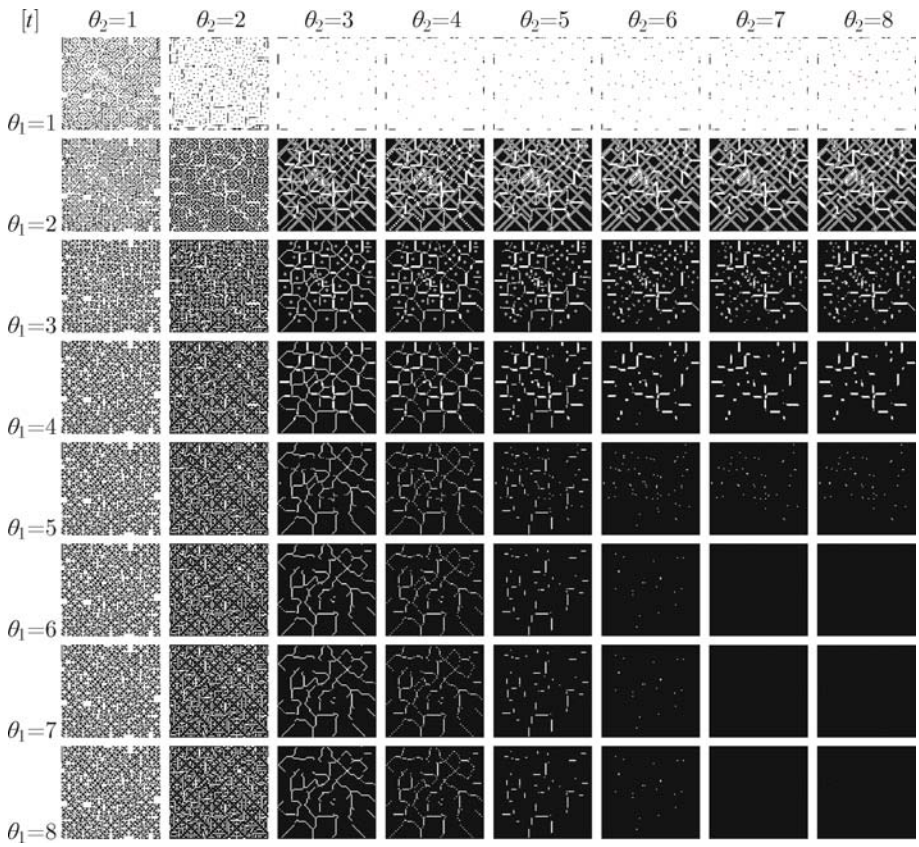
When dealing with excitable media excitability is the key parameter for tuning spatio-temporal dynamics. In [1] we demonstrated that by varying excitability we can force the medium to exhibit almost all possible types of excitation dynamics.

Let each cell of the 2D automaton take three states: resting $(\cdot)$, exciting $(+)$ and refractory $(-)$, and update its state depending on number σ_+ of excited neighbors in its 8-cell neighborhood (Fig. 11a). A cell goes from excited to refractory and from refractory to resting states unconditionally, and a resting cell excites if $\sigma_+ \in [\theta_1, \theta_2]$, $1 \le \theta_1 \le \theta_2 \le 8$. By changing θ_1 and θ_2 we can move the medium dynamics in a domain of "conventional" excitation waves, useful for image processing and robot navigation [7] (Fig. 14a), as well as to make it exhibit mobile localized excitations (Fig. 14b), quasi-particles, discrete analogs of dissipative solitons, employed in collision-based computing [1].

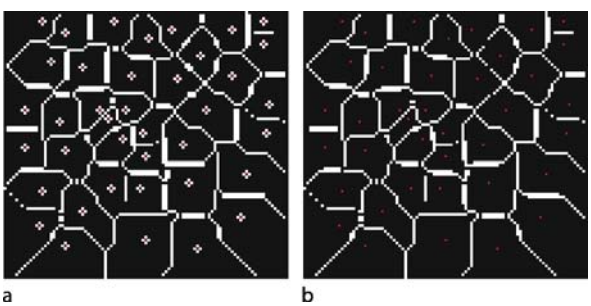
Robot Navigation and Massive Manipulation

As we have seen in previous sections, reaction-diffusion chemical systems can solve complex problems and implement logical circuits. Embedded controllers for non-traditional robotics architectures would be yet another potentially huge field of application of reaction-diffusion computers. The physico-chemical artifacts are well-known to be capable of sensible motion. Most famous are Belousov–Zhabotinsky vesicles [26], self-propulsive chemo-sensitive drops [25,37] and ciliar arrays. Their motion is directional but somewhere lacks sophisticated control mechanisms.

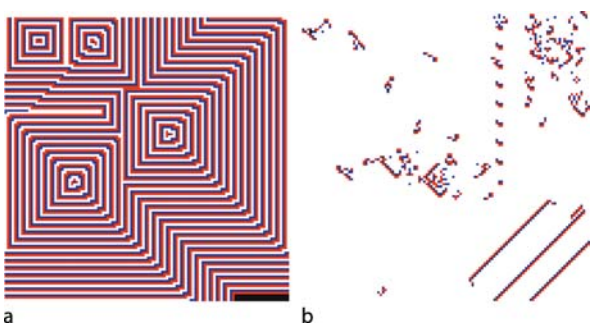
At the present stage of reaction-diffusion computing research it seems to be difficult to provide effective solutions for experimental prototyping of combined sensing, decision-making and actuating. However, as a proof-of-



Reaction-Diffusion Computing, Figure 12
Final configurations of reaction-diffusion medium for $1 \preceq \theta_1 \preceq \theta_2 \preceq 2$. Resting sites are *black*, precipitate is *white* [3]



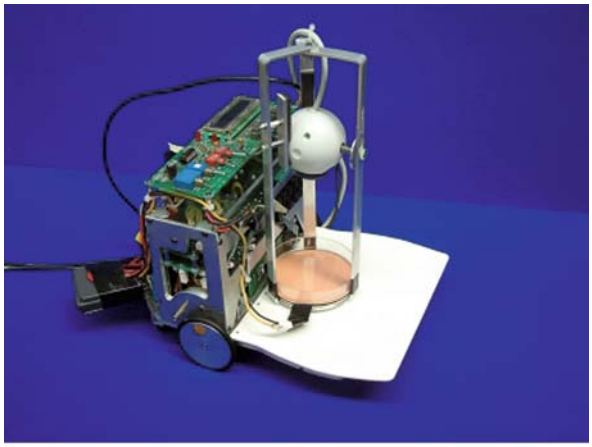
Reaction-Diffusion Computing, Figure 13
Exemplary configurations of reaction-diffusion medium for a $\theta_1 = 3$ and $\theta_2 = 3$, b $\theta_1 = 4$ and $\theta_2 = 3$. Resting sites are *black*, precipitate is *white* [7]



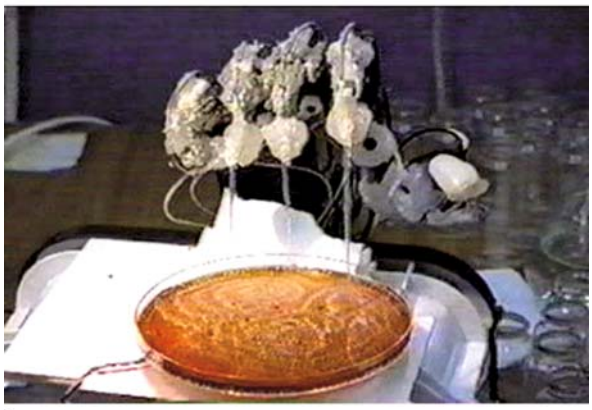
Reaction-Diffusion Computing, Figure 14
Snapshots of space-time excitation dynamics for excitability $\sigma_+ \in [1, 8]$ (a) and $\sigma_+ \in [2, 2]$ (b)

concept we can always consider hybrid "wetware + hardware" systems. For example, to fabricate a chemical controller for a robot (Fig. 15a), we can place a reactor with Belousov–Zhabotinsky solution on-board of a wheeled robot, and allow the robot to observe excitation wave dynamics in the reactor. When the medium is stimulated at one point, target waves are formed. The robot becomes aware of the direction toward source of stimulation from the topology of the wave-fronts [7].

A set of remarkable experiments were undertaken by Hiroshi Yokoi and Ben De Lacy Costello. They built an interface (Fig. 15b) between a robotic hand and Belousov–



a



b

Reaction-Diffusion Computing, Figure 15
Robots controlled by BZ chemical medium: a mobile robot, b robotic hand (courtesy of Hiroshi Yokoi)

Zhabotinsky chemical reactor [62]. Excitation waves propagating in the reactor were sensed by photo-diodes, which triggered finger motion. Bending fingers touched the chemical medium with their glass nails filled with colloid silver, which triggered circular waves in the medium [7]. Starting from any initial configuration, the chemical-robotic system does always reach a coherent activity mode, where fingers move in regular, somewhat melodic patterns, and few generators of target waves govern dynamics of excitation in the reactor [62].

The chemical processors for navigating the wheeled robot and for controlling, and actively interacting with, a robotic hand are well discussed in our recent monograph [7], therefore we will not go into detail in herein. Instead we will concentrate on rather novel findings on coupling of a reaction-diffusion system with a massive parallel array of virtual actuators.

How a reaction-diffusion medium can manipulate objects? To find out we couple a simulated abstract parallel manipulator with an experimental Belousov–Zhabotinsky (BZ) chemical medium, so the excitation dynamics in the chemical system are reflected in changing the OFF-ON mode of elementary actuating units. In this case, we convert experimental snapshots of the spatially distributed chemical system to a force vector field and then simulate the motion of manipulated objects in the force field, thus achieving reaction-diffusion medium controlled actuation. To build an interface between the recordings of space-time snapshots of the excitation dynamics in the BZ medium and simulated physical objects we calculate force fields generated by mobile excitation patterns and then simulate the behavior of an object in this force field.

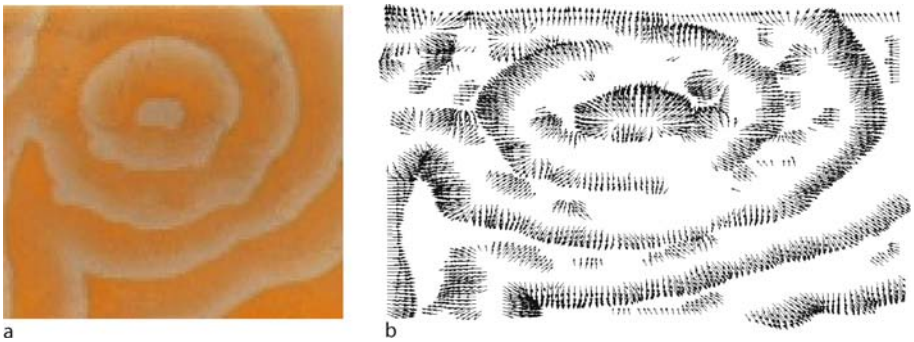
The chemical medium to perform actuation is prepared following the typical recipe[1], see [8,20], based on a ferroin catalyzed BZ reaction. A silica gel plate is cut and soaked in a ferroin solution. The gel sheet is placed in a Petri dish and BZ solution added. The dynamics of the chemical system are recorded at 30-second intervals using a digital camera.

The cross-section profile of the BZ wave-front recorded on a digital snapshot shows a steep rise of red color values in the pixels at the wave-front's head and a gradual descent in the pixels along the wave-front's tail. Assuming that excitation waves push the object local force vectors generated at each site—pixel of the digitized image—of the medium should be oriented along local gradients of the red color values. From the digitized snapshot of the BZ medium we extract an array of red components from the snapshot's pixels and then calculate the projection of a virtual vector force at the pixel. Force fields generated by the excitation patterns in a BZ system Fig. 16 result in tangential forces being applied to a manipulated object, thus causing translational and rotational motions of the object [8].

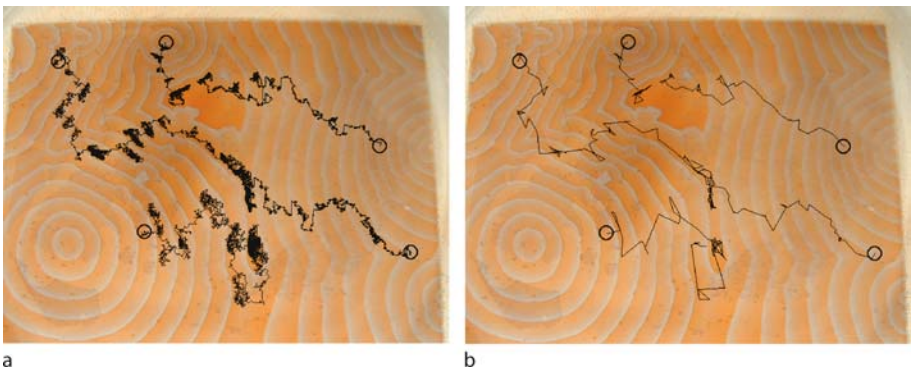
Non-linear medium controlled actuators can be used for sorting and manipulating both small objects, comparable in size to the elementary actuating unit, and larger objects, with lengths of tens or hundreds of actuating units. Therefore, we demonstrate here two types of experiments with BZ-based manipulation: of pixel-sized objects and of planar convex shapes.

Pixel-objects, due to their small size, are subjected to random forces, caused by impurities of the physical medium and imprecision of the actuating units. In this case, no averaging of forces is allowed and the pixel-ob-

[1]Chemical laboratory experiments are undertaken by Dr. Ben De Lacy Costello (UWE, Bristol, UK).



Reaction-Diffusion Computing, Figure 16
Force vector field (b) calculated from BZ medium's image (a) [8]



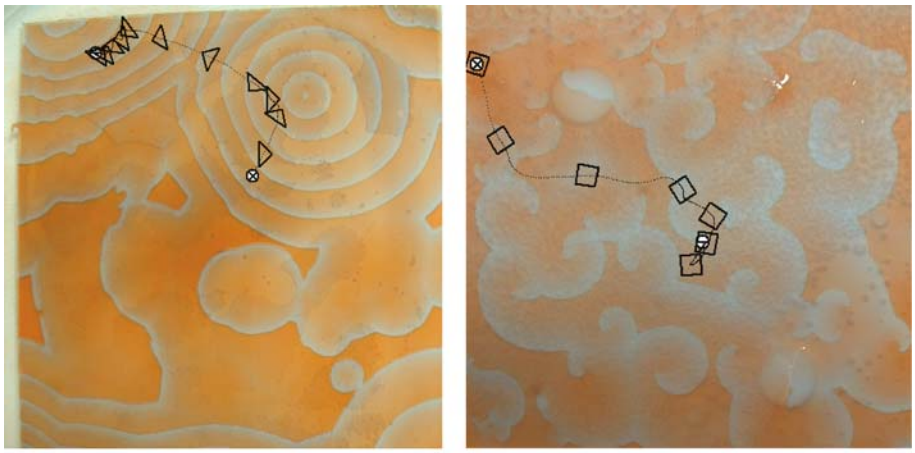
Reaction-Diffusion Computing, Figure 17
Examples of manipulating five pixel-objects using the BZ medium: a trajectories of pixel-objects, b jump-trajectories of pixel-objects recorded every 100th time step. Initial positions of the pixel-objects are shown by *circles* [8]

jects themselves sensitively react to a single force vector. Therefore, we adopt the following model of manipulating a pixel-object: if all force vectors at the 8-pixel neighborhood of the current site of the pixel-object are nil then the pixel-object jumps to a randomly chosen neighboring pixel of its neighborhood, otherwise the pixel-object is translated by the maximum force vector in its neighborhood.

When placed on the simulated manipulating surface, pixel-objects move at random in the domains of the resting medium, however by randomly drifting each pixel-object does eventually encounter a domain of co-aligned vectors (representing the excitation wave front in the BZ medium) and is translated along the vectors. An example of several pixel-objects transported on a "frozen" snapshot of the chemical medium is shown in Fig. 17. Trajectories of pixel-objects (Fig. 17a) show distinctive intermittent modes of random motion separated by modes of directed "jumps" guided by traveling wave fronts. Smoothed trajectories of pixel-objects (Fig. 17b) demonstrate that despite a very strong chaotic component in manipulation, pixel-objects are transported to the sites of the medium where two or more excitation wave-fronts meet.

The overall speed of pixel-object transportation depends on the frequency of wave generations by sources of target waves. As a rule, the higher the frequency the faster the objects are transported. This is because in parts of the medium spanned by low frequency target waves there are lengthy domains of resting system, where no force vectors are formed. Therefore, pixel-sized objects can wander randomly for a long time till climbing the next wave front [8].

To calculate the contribution of each force we partitioned the object into fragments, using a square grid, in which each cell of the grid corresponds to one pixel of the image. We assume that the magnitude of the force applied to each fragment above a given pixel is proportional to the area of the fragment, and is co-directional with a force vector. A momentum of inertia of the whole object with respect to the axis normal to the object and passing through the object's center of mass is calculated from the position



Reaction-Diffusion Computing, Figure 18
Manipulating planar object in BZ medium. a *Right-angled triangle* moved by fronts of target waves. b *Square object* moved by fronts of fragmented waves in sub-excitable BZ medium. Trajectories of center of mass of the square are shown by the *dotted line*. Exact orientation of the objects is displayed every 20 steps. Initial position of the object is shown by ⊖ and the final position by ⊗ [8]

of the center of mass and the mass of every fragment. Since the object's shape and size are constant it is enough to calculate the moment of inertia only at the beginning of simulation. We are also taking into account principal rotational momentum created by forces and angular acceleration of the object around its center of mass. Therefore, object motion in our case can be sufficiently described by the coordinates of its center of mass and its rotation at every moment of time [8].

Spatially extended objects follow the general pattern of motion observed for the pixel-sized objects. However, due to integration of many force vectors the motion of planar objects is smoother and less sensitive to the orientation of any particular force-vector.

Outcome of manipulation depends on the size of the object, with increasing size of the object—due to larger numbers of local vector-forces acting on the object—the objects become more controllable by the excitation wave-fronts (Fig. 18).

Future Directions

The field of reaction-diffusion computing started 20 years ago [28] as a sub-field of physics and chemistry dealing with image processing operations in uniform thin-layer excitable chemical media. The basic idea was to apply input data as a two-dimensional profile of heterogeneous illumination, then allow excitation waves to spread and interact with each, and then optically record the result of the computation. The first reaction-diffusion computers were already massively parallel, with parallel optical input and outputs. Later computer engineers entered the field, and started to exploit traditional techniques—wires were implemented by channels where wave-pulses travel, and specifically shaped junctions acted as logical valves. In this manner, the most "famous" chemical computing devices were implemented, including Boolean gates, coincidence detectors, memory units and more. The idea of reaction-diffusion computation was if not ruined then forced into a cul-de-sac of non-classical computation. The breakthrough happened when paradigms and solutions from the field of dynamical, collision-based computing and conservative logic, were mapped onto realms of spatially extended chemical systems. The computers became uniform and homogeneous.

Over several examples we demonstrated that reaction-diffusion chemical systems are capable of solving combinatorial problems with natural parallelism. In spatially distributed chemical processors, the data and the results of the computation are encoded as concentration profiles of the chemical species. The computation per se is performed via the spreading and interaction of wave fronts.

The reaction-diffusion computers are parallel because the chemical medium's micro-volumes update their states simultaneously, and molecules diffuse and react in parallel. During the last decades a wide range of experimental prototypes of reaction-diffusion computing devices have been fabricated and applied to solve various problems of computer science, including: image processing, pattern recognition, path planning, robot navigation, computational ge-

ometry, logical gates in spatially distributed chemical media, arithmetical and memory units.

These important—but scattered across many scientific fields— results convince us that reaction-diffusion systems can do a lot. Are they capable enough to be intelligent? Yes, reaction-diffusion systems are smart—showing a state of readiness to respond, able to cope with difficult situations, capable of determining something by mathematical and logical methods—and endowed with the capacity to reason. Reaction-diffusion computers allow for massive parallel input of data. Equivalently, reaction-diffusion robots would need no dedicated sensors, each micro-volume of the medium, each site of the matrix gel, is sensitive to changes in one or another physical characteristic of the environment. Electric field, temperature and illumination are "sensed" by reaction-diffusion devices, and these are the three principle parameters in controlling and programming reaction-diffusion robots.

Hard computational problems of geometry, image processing and optimization on graphs are resource-efficiently solved in reaction-diffusion media due to intrinsic natural parallelism of the problems [1]. Herein we demonstrated the efficiency of reaction-diffusion computers with the example of constructing Voronoi diagrams. The Voronoi diagram is a subdivision of plane by data planar set. Each point of the data set is represented by a drop of a reagent. The reagent diffuses and produces a color precipitate when reacting with the substrate. When two or more diffusive fronts of the "data" chemical species meet, no precipitate is produced (due to concentration-dependent inhibition). Thus, uncolored domains of the computing medium represent bisectors of the Voronoi diagram. The precipitating chemical processor can also compute a skeleton. The skeleton of a planar shape is computed in a similar manner. A contour of the shape is applied to the computing substrate as a disturbance in reagent concentrations. The contour concentration profile induces diffusive waves. A reagent diffusing from the data-contour reacts with the substrate and the precipitate is formed. The precipitate is not produced at the sites of collision of diffusive waves. The uncolored domains correspond to the skeleton of the data shape. To compute a collision-free shortest path in a space with obstacles, we can couple two reaction-diffusion media. Obstacles are represented by local disturbances of concentration profiles in one of the media. The disturbances induce circular waves traveling in the medium and approximating a scalar distance-to-obstacle field. This field is mapped onto the second medium, which calculates a tree of "many-sources-one-destination" shortest paths by spreading wave-fronts [7].

Just few words of warning—when thinking about chemical algorithms some of you may realize that diffusive and phase waves are pretty slow in physical time. The sluggishness of computation is the only point that may attract criticism to reaction-diffusion chemical computers. There is however a cure—to speed up we are implementing the chemical medium in silicon, micro-processor LSI analogs of reaction-diffusion computers [12]. Further miniaturization of the reaction-diffusion computers can be reached when the system is implemented as a two-dimensional array of single-electron nonlinear oscillators diffusively coupled to each other [38]. Yet another point of developing reaction-diffusion computers—is to design embedded controllers for soft-bodied robots, where usage of conventional silicon materials seem to be inappropriate.

Acknowledgments

Chemical laboratory prototypes of reaction-diffusion computers, discussed in the article, were implemented by Ben De Lacy Costello. I am grateful to Andy Wuensche (hexagonal cellular automata), Hiroshi Yokoi (robotic hand controlled by Belousov–Zhabotinsky reaction), Chris Melhuish (control of robot navigation), Sergey Skachek (massive parallel manipulation), Tetsuya Asai (LSI prototypes of reaction-diffusion computers) and Genaro Martinez (binary-state cellular automata) for their cooperation. Some pictures, where indicated, where adopted from our co-authored publications. My sincere thanks go to Soichiro Tsuda and Tomohiro Shirakawa who ignited my interest in experimenting with plasmodium and were my patient advisers.

Bibliography

Primary Literature

1. Adamatzky A (2001) Computing in nonlinear media and automata collectives. Institute of Physics Publishing, London
2. Adamatzky A (ed) (2003) Collision based computing. Springer, London
3. Adamatzky A (2005) Programming reaction-diffusion computers. In: Unconventional programming paradigms. Springer, New York
4. Adamatzky A, De Lacy Costello BPJ (2002) Experimental logical gates in a reaction-diffusion medium: the XOR gate and beyond. Phys Rev E 66:046112
5. Adamatzky A, Wuensche A (2006) Computing in 'spiral rule' reaction-diffusion hexagonal cellular automaton. Complex Syst 16(4):277–298
6. Adamatzky A, De Lacy Costello BPJ, Melhuish C, Ratcliffe N (2003) Experimental reaction-diffusion chemical processors for robot path planning, J Intell Robt Syst 37:233–249
7. Adamatzky A, De Lacy Costello BPJ, Asai T (2005) Reaction-diffusion computers. Elsevier, Amsterdam

8. Adamatzky A, De Lacy Costello BPJ, Skachek S, Melhuish C (2005) Manipulating objects with chemical waves: open loop case of experimental Belousov–Zhabotinsky medium. Phys Lett A 350(3–4):201–209
9. Adamatzky A, Wuensche A, De Lacy Costello BPJ (2006) Glider-based computation in reaction-diffusion hexagonal cellular automata. Chaos, Solitons Fractals 27:287–295
10. Agladze K, Magome N, Aliev R, Yamaguchi T, Yoshikawa K (1997) Finding the optimal path with the aid of chemical wave. Phys D 106:247–254
11. Asai T, De Lacy Costello BPJ, Adamatzky A (2005) Silicon implementation of a chemical reaction-diffusion processor for computation of Voronoi diagram. Int J Bifurc Chaos 15(1)
12. Asai T, Kanazawa Y, Hirose T, Amemiya Y (2005) Analog reaction-diffusion chip imitating Belousov–Zhabotinsky reaction with hardware oregonator model. Int J Unconv Comput 1:123–147
13. Beato V, Engel H (2003) Pulse propagation in a model for the photosensitive Belousov–Zhabotinsky reaction with external noise. In: Schimansky-Geier L, Abbott D, Neiman A, Van den Broeck C (eds) Noise in complex systems and stochastic dynamics Proc. SPIE, vol 5114. pp 353–362
14. Berlekamp ER, Conway JH, Guy RL (1982) Winning ways for your mathematical plays, vol 2. Academic Press
15. Bode M, Liehr AW, Schenk CP, Purwins H-G (2000) Interaction of dissipative solitons: particle-like behavior of localized structures in a three-component reaction-diffusion system. Phys D 161:45–66
16. Brandtstädter H, Braune M, Schebesch I, Engel H (2000) Experimental study of the dynamics of spiral pairs in light-sensitive Belousov–Zhabotinskii media using an open-gel reactor. Chem Phys Lett 323:145–154
17. Courant R, Robbins H (1941) What is mathematics? Oxford University Press
18. Dupont C, Agladze K, Krinsky V (1998) Excitable medium with left-right symmetry breaking. Phys A 249:47–52
19. Field RJ, Noyes RM (1974) Oscillations in chemical systems: IV. Limit cycle behavior in a model of a real chemical reaction. J Chem Phys 60:1877–1884
20. Field R, Winfree AT (1979) Traveling waves of chemical activity in the Zaikin–Zhabotinsky–Winfree reagent. J Chem Educ 56:754
21. Fredkin F, Toffoli T (1982) Conservative logic. Int J Theor Phys 21:219–253
22. Gerhardt M, Schuster H, Tyson JJ (1990) A cellular excitable media. Phys D 46:392–415
23. Grill S, Zykov VS, Müller SC (1996) Spiral wave dynamics under pulsatory modulation of excitability. J Phys Chem 100:19082–19088
24. Hartman H, Tamayo P (1990) Reversible cellular automata and chemical turbulence. Phys D 45:293–306
25. Kitahata H, Yoshikawa K (2005) Chemo-mechanical energy transduction through interfacial instability. Phys D 205:283–291
26. Kitahata H, Aihara R, Magome N, Yoshikawa K (2002) Cinvective and periodic motion driven by a chemical wave. J Chem Phys 116:5666
27. Klein R (1990) Concrete and abstract voronoi diagrams. Springer, Berlin
28. Kuhnert L (1986) A new photochemical memory device in a light sensitive active medium. Nature 319:393
29. Kuhnert L, Agladze KL, Krinsky VI (1989) Image processing using light-sensitive chemical waves. Nature 337:244–247
30. Kusumi T, Yamaguchi T, Aliev R, Amemiya T, Ohmori T, Hashimoto H, Yoshikawa K (1997) Numerical study on time delay for chemical wave transmission via an inactive gap. Chem Phys Lett 271:355–360
31. Maeda S, Hashimoto S, Yoshida R (2004) Design of chemo-mechanical actuator using self-oscillating gels. In: Proc. IEEE International Conference on Robotics and Biomimetics (ROBIO 2004). p 313
32. Margolus N (1984) Physics-like models of computation. Phys D 0:81–95
33. Markus M, Hess B (1990) Isotropic cellular automata for modeling excitable media. Nature 347:56–58
34. Mills J (2005) The new computer science and its unifying principle: complementarity and unconventional computing. In: The Grand Challenge in Nonclassical Computation, Int. Workshop, Position Papers. York, 18–19 April 2005
35. Motoike IN, Yoshikawa K (2003) Information operations with multiple pulses on an excitable field. Chaos, Solitons Fractals 17:455–461
36. Motoike IN, Yoshikawa K, Iguchi Y, Nakata S (2001) Real-time memory on an excitable field. Phys Rev E 63:036220
37. Nagai K, Sumino Y, Kitahata H, Yoshikawa K (2005) Mode selection in the spontaneous motion of an alcogol droplets. Phys Rev E 71:065301
38. Oya T, Asai T, Fukui T, Amemiya Y (2005) Reaction-diffusion systems consisting of single-electron oscillators. Int J Unconvent Comput 1:179–196
39. Petrov V, Ouyang Q, Swinney HL (1997) Resonant pattern formation in a chemical system. Nature 388:655–657
40. Pour-El MB (1974) Abstract computability and its relation to the general purpose analog computer (some connections between logic, differential equations and analog computers). Trans Am Math Soc 199:1–28
41. Qian H, Murray JD (2001) A simple method of parameter space determination for diffusion-driven instability with three species. Appl Math Lett 14:405–411
42. Rambidi NG (1998) Neural network devices based on reaction-diffusion media: an approach to artificial retina. Supramol Sci 5:765–767
43. Rambidi NG, Shamayaev KR, Peshkov GY (2002) Image processing using light-sensitive chemical waves. Phys Lett A 298:375–382
44. Ramos JJ (2003) Oscillatory dynamics of inviscid planar liquid sheets. Appl Math Comput 143:109–144
45. Saltenis V (1999) Simulation of wet film evolution and the Euclidean Steiner problem. Informatica 10:457–466
46. Schebesch I, Engel H (1998) Wave propagation in heterogeneous excitable media. Phys Rev E 57:3905–3910
47. Schenk CP, Or-Guil M, Bode M, Purwins H-G (1997) Interacting pulses in three-component reaction-diffusion systems on two-dimensional domains. Phys Rev Lett 78:3781–3784
48. Sediña-Nadal I, Mihaliuk E, Wang J, Pérez-Muñuzuri V, Showalter K (2001) Wave propagation in subexcitable media with periodically modulated excitability. Phys Rev Lett 86:1646–1649
49. Sielewiesiuk J, Gorecki J (2001) Logical functions of a cross junction of excitable chemical media. J Phys Chem A 105:8189–8195

50. Sienko T, Adamatzky A, Rambidi N, Conrad M (eds) (2003) Molecular computing. MIT Press
51. Steinbock O, Toth A, Showalter K (1995) Navigating complex labyrinths: optimal paths from chemical waves. Science 267:868–871
52. Tóth A, Showalter K (1995) Logic gates in excitable media. J Chem Phys 103:2058–2066
53. Tyson JJ, Fife PC (1980) Target patterns in a realistic model of the Belousov–Zhabotinskii reaction. J Chem Phys 73:2224–2237
54. Vergis A, Steiglitz K, Dickinson B (1986) The complexity of analog computation. Math Comput Simul 28:91–113
55. Wang J (2001) Light-induced pattern formation in the excitable Belousov–Zhabotinsky medium. Chem Phys Lett 339:357–361
56. Weaire D, Hutzler S, Cox S, Kern N, Alonso MD, Drenckhan W (2003) The fluid dynamics of foams. J Phys Condens Matter 15:S65–S73
57. Wuensche A (2005) Glider dynamics in 3-value hexagonal cellular automata: the beehive rule. Int J Unconvent Comput 1:375–398
58. Wuensche A, Adamatzky A (2006) On spiral glider-guns in hexagonal cellular automata: Activator-inhibitor paradigm. Int J Mod Phys 17:1009–1026
59. Yaguma S, Odagiri K, Takatsuka K (2004) Coupled-cellular-automata study on stochastic and pattern-formation dynamics under spatiotemporal fluctuation of temperature. Phys D 197:34–62
60. Yang X (2004) Pattern formation in enzyme inhibition and cooperativity with parallel cellular automata. Parallel Comput 30:741–751
61. Yang X (2006) Computational modeling of nonlinear calcium waves. Appl Math Model 30:200–208
62. Yokoi H, Adamatzky A, De Lacy Costello B, Melhuish C (2004) Excitable chemical medium controlled for a robotic hand: closed loop experiments. Int J Bifurc Chaos 14(9)3347–3354
63. Yoneyama M (1996) Optical modification of wave dynamics in a surface layer of the Mn-catalyzed Belousov–Zhabotinsky reaction. Chem Phys Lett 254:191–196
64. Young D (1984) A local activator–inhibitor model of vertebrate skin patterns. Math Biosci 72:51

Books and Reviews

Adamatzky A (2001) Computing in nonlinear media and automata collectives. Institute of Physics Publ, London
Adamatzky A (ed) (2003) Collision-Based computing. Springer, London
Adamatzky A, Teuscher C (eds) (2006) From utopian to genuine unconventional computers. Luniver Press, London
Adamatzky A, De Lacy Costello B, Asai T (2005) Reaction-Diffusion computers. Elsevier, Amsterdam
Chua L (1998) CNN: a paradigm for complexity. World Scientific
Gray P, Scott SK (2002) Chemical oscillations and instabilities: nonlinear chemical kinetics. Oxford University Press, Oxford
Scott SK (1994) Oscillations, waves and chaos in chemical kinetics. Oxford University Press, Oxford
Teuscher C, Adamatzky A (eds) (2005) Proceedings of the 2005 workshop on unconventional computing: from cellular automata to wetware. Luniver Press, London
Toffoli T, Margolus N (1987) Cellular automata machines. MIT Press

Reaction Kinetics in Fractals

Ezequiel V. Albano
Facultad de Ciencias Exactas, UNLP, CONICET, Instituto de Investigaciones Fisicoquímicas Teóricas y Aplicadas (INIFTA) CCT La Plata, La Plata, Argentina

Article Outline

Glossary

Fractals Fractal geometry is a mathematical tool well suited to treating complex systems that exhibit scale invariance or, equivalently, the absence of any characteristic length scale. Scale invariance implies that objects are self-similar: if we take a part of the object and magnify it by the same magnification factor in all directions, the obtained object cannot be distinguished from the original one. Self-similar objects are often characterized by non-integer dimensions, a fact that led B. Mandelbrot, early in the 1980s, to coin the name "fractal dimension". Also, all objects described by fractal dimensions are generically called *fractals*. Within the context of the present work, fractals are the underlying media where the *reaction kinetics* among atoms, molecules, or particles in general, is studied.

Reaction kinetics The description of the time evolution of the concentration of reacting particles (ρ_i, where $i = 1, 2, \ldots, N$ identifies the type of particle) is achieved, far from a stationary regime, by formulating a kinetic rate equation $\dot{\rho}_i(t) = F[\rho_i(t)]$, where F is a function. This description, known in physical chemistry as the *law of mass action*, states that the rate of a chemical reaction is proportional to the concentration of reacting species and was formulated by Waage and Guldberg in 1864. Often, especially when dealing with reactions occurring in homogeneous media, F involves integer powers (also known as the reaction orders) of the concentrations, leading to *classical reaction kinetics*. However, as in most cases treated in this arti-

cle, if a reaction takes place in a fractal, one may also have kinetic rate equations involving non-integer powers of the concentration that lead to *fractal reaction kinetics*. Furthermore, it is usual to find that the slowest step involved in a kinetic reaction determines its rate, leading to a *process-limited reaction*, where e. g. the process could be diffusion or mass transport, adsorption, reaction, etc.

Heterogeneously catalyzed reactions A reaction limited by at least one rate-limiting step could be prohibitively slow for practical purposes when, e. g., it occurs in a homogeneous media. The role of a good solid-state catalyst in contact with the reactants – in the gas or fluid phase – is to obtain an acceptable output rate of the products. Reactions occurring in this way are known as heterogeneously catalyzed. This type of reaction involves at least the following steps: (i) the first one comprises trapping, sticking and adsorption of the reactants on the catalytic surface. Particularly important, from the catalytic point of view, is that molecules that are stable in the homogeneous phase, e. g. H_2, N_2, O_2, etc., frequently undergo dissociation on the catalyst surface. This process is essential in order to speed up the reaction rate. (ii) After adsorption, species may diffuse or remain immobile (chemisorbed) on the surface. The actual reaction step occurs between neighboring adsorbed species of different kinds. The result of the reaction is the formation of products that can either be intermediates of the reaction or its final output. (iii) The final step is the desorption of the products, which is essential not only for the practical purpose of collecting and storing the desired output, but also in order to regenerate the catalytically active surface sites.

Definition of the Subject

Processes involving reaction among atoms, molecules, and particles in general, are ubiquitous both in nature and in laboratories. After the introduction of the concept of fractals by B. Mandelbrot [89] in the 1980s, it has been realized that a wide variety of reactions take place in fractal media, leading to the observation of anomalous behavior, i. e. the so called "fractal reaction kinetics" [33,67,81]. The study of this subject has received considerable theoretical attention [38,43,76,77,103] and a huge numerical effort has been made to improve its understanding [13,17,18,20,34,80,92]. Furthermore, among the practical examples of fractal reaction kinetics one can quote exciton annihilation in composite materials, chemical reactions in membranes pores, charge recombination in colloids and clouds, coagulation, polymerization and growth of dendrites, [15,79,83,94] etc. Another scenery for the study of *reaction kinetics in fractals* is in the field of heterogeneous catalysis. In fact, it is well known that most catalysts are composed of small, catalytically active particles, supported by highly porous (fractal) substrates [21,22,23,30,75,95,106]. So, it is not surprising that this field of research has become particularly active due not only to its practical and technological relevance, but also to the occurrence of quite interesting and challenging phenomena such as irreversible phase transitions, oscillatory behavior, chaos and stochastic resonance, propagation and interference of chemical waves, interface coarsening, metastability and hysteretic effects, etc. [9,10,52,68,70,71,84,85,90,109].

Introduction

The description of the kinetic behavior of reaction processes occurring in homogeneous media can be found in most textbooks on chemical physics. However, care must be taken considering *reaction kinetics in fractals* because conventional textbook equations are based on mean-field approaches that neglect not only the fractal structure of the underlying media, but also fluctuations in the concentration of the reactants, many-particle effects, etc. that may lead to the observation of anomalous behavior [81]. In fact, soon after the recognition of the relevance of the concept of fractals for the description of the structure and properties of physical objects by Mandelbrot [89], more than three decades ago, it was realized that random transport in low-dimensional and disordered media may be anomalous [33,67,105], as are diffusion-limited reactions occurring in those media [13,15,17,18,20,34,43,79,80,81,83,92,94]. Therefore, the classical reaction kinetic approach becomes unsatisfactory in a wide variety of situations, such as when the reactants are spatially constrained by walls, phase boundaries or force fields. Due to these studies, early in the 1980s, it was realized that even elementary reactions (e. g. $A + A \rightarrow$ inert, $A + B \rightarrow$ products, etc.) can no longer be described by classical rate equations with integer exponents (the so-called "reaction order") but instead, fractal orders were identified leading to *fractal reaction kinetics* [89]. The emerging new theory also predicts the occurrence of self-ordering and self-unmixing of reactants [38,76,77,103], as well as time-dependent rate "constants", i. e. rate coefficients with temporal "memories" [89].

On the other hand, after the exhaustive study and characterization of the fractal nature of a wide variety of substrates used as support in most catalysts, due to Avnir et

al. [21,22,23,95], the study of heterogeneously catalyzed reactions in fractal media has also attracted growing attention [1,2,5,7,41,72,86,87]. These studies are focused on the understanding of irreversible phase transitions occurring between an active state with reaction and an inactive state where the reaction ceases irreversibly. This latter state is due to the poisoning of the catalyst, by the reactants and/or their subproducts, and is known as the absorbing state: a system trapped in the absorbing state can never escape from it [9,10,68,85,90].

In order to cover those topics, this article is organized as follows: in Sect. "Fractals and Some of Their Relevant Properties", the basic properties of fractals are presented and discussed. Section "Random Walks" is devoted to the description of the behavior of random walks in fractal media, while in Sect. "Diffusion-limited Reactions" archetypical cases of diffusion limited reactions among random walks are considered. Irreversible phase transitions occurring in heterogeneously catalyzed reactions are addressed in Sect. "Irreversible Phase Transitions in Heterogeneously Catalyzed Reactions". Finally, in Sect. "Future Directions" promising directions for future works are briefly outlined.

Fractals and Some of Their Relevant Properties

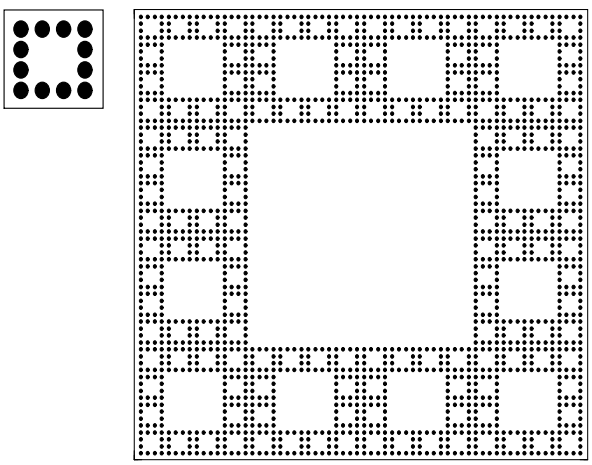
Classical Euclidean geometry is useful for describing the properties of regular objects such as circles, spheres, cones, etc. However, disordered objects such as clouds, dielectric breakdown patterns, coastlines, mountains, landscapes, etc., have not been so far satisfactorily described by using the Euclidean geometry. B. Mandelbrot overcome this shortcoming by introducing the fractal geometry as a suitable tool for the treatment of disordered or fractal media [89]. Along this article, fractals will be used as media where reactions of interest take place. A relevant property of fractal media is self-similarity or symmetry under dilatation [27,33,40,61,67]. Here, it is worth noting the main difference between regular Euclidean space and fractal geometry: whereas the former has translation symmetry, this type of symmetry is violated in the latter, which exhibits a new symmetry known as scale invariance or invariance under dilatation. Since these concepts may appear to be obscure, it is convenient to develop an intuitive understanding and outline some basic definitions. Fractals can either be deterministic or non-deterministic (often called random fractals). However, it is worth mentioning that many non-deterministic fractals may be obtained as a result of complex dynamic (non-stochastic) processes. From the large variety of known deterministic fractals, e. g. the Koch curve, the Julia set, Sierpinski gaskets, carpets, and sponges, etc., let us focus our attention on the construction of a Sierpinski Carpet (SC). In order to build up a generic SC(a,c), a square in $d = 2$ dimensions is segmented into a^d subsquares and c of them are then removed. This square, of side a, is known as the generating cell. The segmentation process is then iterated on the remaining squares a number k of segmentation steps. Figure 1 shows the SC(4,4) obtained after $k = 4$ segmentation steps. These kinds of fractals have a lower cutoff length given by the side of the generating cell, but, in principle, they lack an upper cutoff length for $k \to \infty$. However, in practice one performs a finite number of segmentation steps.

Let us now remind the reader about the concept of dimension in regular systems. In this case, the dimension d characterizes the dependence of the mass $M(L)$ as a function of the linear size L of the system. If we now consider a smaller part of the system of size $bL(a < 1)$, then $M(aL)$ will be decreased by a factor of a^d, so that

$$\tilde{M}(a, L) = \mu(a)M(aL) = a^d M(L) . \tag{1}$$

The solution of Eq. (1) gives $M(L) = AL^d$, where A is a constant. So, for a wire one has $d = 1$, while for a thin plate $d = 2$ is obtained, etc. Intuitively, the SC(4,4) shown in Fig. 1 seems to be "denser" than a wire but also "sparser" than a thin plate. So, Eq. (1) has to be generalized for fractals, leading to

$$M(aL) = a^{d_f} M(L) , \tag{2}$$



Reaction Kinetics in Fractals, Figure 1
Sierpinski carpet obtained after $k = 4$ iterations on the generating cell shown in the *top-left corner*, i. e. the SC(4,4). The lattice side is $L = 64$ and the fractal dimension of this object is $D_f = log(12)/log(4) \simeq 1.7925$

which yields the following solution

$$M(L) = BL^{d_f}\,, \tag{3}$$

where B is a constant, and d_f is a non-integer dimension known, after Mandelbrot, as the fractal dimension. By applying Eq. (2) to the SC(4,4) shown in Fig. 1, one has $M(1/4\,L) = 1/12\,M(L)$, so that $d_f = \log(12)/\log(4) \simeq 1.7925$. In general, for an SC one has $M(1/aL) = 1/(a^2 - c)\,M(L)$, which yields $d_f = \log(a^2 - c)/\log(a)$.

The factor a in Eqs. (1) and (3) could be an arbitrary real number, leading to *continuous* scale invariance. However, many fractals exhibit *discrete* scale invariance (DSI) [99], which is a weak kind of scale invariance such that a is no longer an arbitrary real number, but it can only take specific discrete values of the form $a_n = (a_1)^n$, where a_1 is a fundamental scaling ratio. Then, for the case of DSI, the solution of Eq. (1) yields

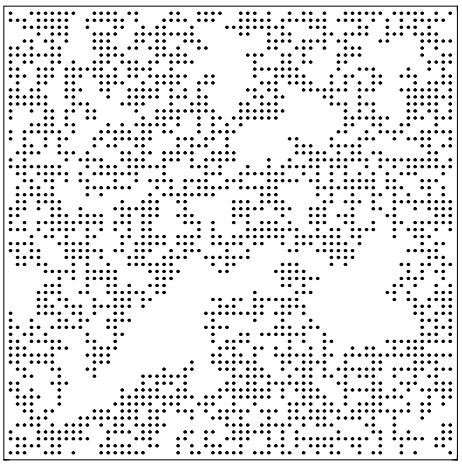
$$M(L) = L^{d_f} F\left(\frac{\log(L)}{\log(a_1)}\right)\,, \tag{4}$$

where F is a periodic function of period one. The measurement of soft oscillations in spatial domain [99] is a signature of spatial DSI. For instance, the $SC(4,4)$ shown in Fig. 1 exhibits DSI with $a_1 = 4$.

Deterministic fractals can be finitely ramified or infinitely ramified objects. A fractal is classified as finitely ramified if any bounded set of the fractal can be separated from the whole structure just by removing a finite number of bonds between the individual constituting entities. This is not possible for the case of the SC(4,4) shown in Fig. 1, so that it is an infinitely ramified fractal. However, other fractals such as the Sierpinski Gasket and the Koch curve are finitely ramified. One advantage of finitely ramified clusters is that various physical properties, such as the conductivity and the vibrational excitations, can be calculated exactly helping to understand the anomalous behavior of some observables. However, it is worth mentioning that classical critical phenomena, e. g. the Ising model, are not observed in finitely ramified fractals [58,59,60]. This is not the case of irreversible critical phenomena (e. g. the contact process [68]), criticality occurring at a trivial value of the control parameter, e. g. the coarsening without surface tension observed in the voter model at $T = 0$ [48], which is still present in finitely ramified fractals [102], etc.

For a careful description of the building rules and relevant properties of various deterministic fractals see e. g. [40].

In the large variety of non-deterministic fractals, there are many examples of random fractals such as diffusion-limited aggregates [27,40,107], cluster-cluster aggregates [101], the incipient percolation cluster [101], etc. So,



Reaction Kinetics in Fractals, Figure 2
Percolating cluster obtained at the critical threshold. The lattice side is $L = 64$ and the cluster contains 2100 particles. The fractal dimension of the infinite percolation cluster is $D_f \simeq 1.89$

in order to acquaint the reader with this subject, let us briefly describe the percolation model in $d = 2$ dimensions. Considering the square lattice, one has that each site can be occupied randomly with probability p or left empty with probability $(1 - p)$. This model mimics, for example, the deposition of conducting metallic particles in an isolating substrate, so that p is the surface density of deposited particles. At lower densities, one has isolated clusters of particles and the films do not conduct electrical current between the edges of the sample. However, by properly increasing the concentration, the onset of electrical conductivity is abruptly observed at a certain critical value p_c, known as the percolation threshold [101]. So, at criticality one has a spanning cluster that connects opposite edges of the sample, which is known as the Incipient Percolation Cluster (IPC), see Fig. 2. Equations (2) and (3) also hold for random fractals, but the mass M has to be averaged over different realizations of the fractal in order to achieve reliable statistics. The IPC is composed of several fractal substructures such as the backbone, blobs, red bonds, dangling ends, etc., so that one needs to define different fractal dimensions in order to achieve a better description.

Several methods have been developed in order to measure the fractal dimension of non-deterministic fractals, the "sandbox" and the "box counting" methods being among the most extensively used, for further details see e. g., [40].

Random Walks

The properties of random walks are useful for the understanding of a great variety of phenomena in virtually all

sciences, e. g. in Physics, Chemistry, Astronomy, Biology, Ecology, and even in Economics [105].

Throughout this article, we will focus our attention on simple random walks on lattices, either regular or fractal, as a model for diffusion. The simple discrete random walk is a stochastic process, such that the walk advances one step in unit time. The walker steps from its present position to another site of the lattice according to a specified random rule. Since the rule is independent of the history of the walk, the process is Markovian. Along this article, we will consider cases where the step is performed, with the same probability, to one of the nearest neighbor sites of the lattice. While that choice is always possible in regular lattices, it is no longer the case for random walks in fractals, because all nearest neighbor sites may not belong to the substrate. In this situation one often considers steps performed with equal probability to any of the nearest neighbor sites belonging to the substrate.

Let us now consider the displacement of the random walk. After n steps (since one has discrete time scale, $n = t$ holds) the net displacement ($R(t)$) is given by [33,67,105]

$$R(t) = \sum_{j=1}^{n} \mathbf{u}_j \,, \tag{5}$$

where $\mathbf{u}_j$ is a unitary vector pointing to a nearest-neighbor site, so that it represents the jth step of the walk. Due to the fact that $\langle \mathbf{u}_j \rangle = 0$, the displacement averaged over a large number of realizations of the walk vanishes, i. e. $\langle R(t) \rangle = 0$. So, a more interesting and useful kinetic observable of random walks is the rms displacement from the origin (R^2), given by

$$\langle R^2(t) \rangle = \left\langle \left(\sum_{j=1}^{n} \mathbf{u}_j \right)^2 \right\rangle = t + 2 \sum_{j>i}^{n} \langle \mathbf{u}_j . \mathbf{u}_i \rangle = t \,, \tag{6}$$

since $\langle \mathbf{u}_j . \mathbf{u}_j \rangle = 1$ and $\langle \mathbf{u}_j . \mathbf{u}_i \rangle = 0$ for $i \neq j$ because the steps are independent. So, the classical result for Euclidean space is obtained.

A distinctive feature of transport in fractal media is that the linear time dependence of the rms displacement of the walk given by Eq. (6) has to be replaced by

$$R^2 \propto t^{2/d_w} \,, \tag{7}$$

where d_w is the anomalous diffusion exponent. In most studied models the fractal dimension d_w exceeds 2, due to the fact that disorder tends, on average, to slow down the diffusion of the walk moving on those media.

Another interesting observable of a walk is the average number of distinct sites visited by a single random walk after N steps (S_N), which is also known as the exploration space of the walk [17,19,67,105] and references therein. Assuming that $N \propto t$, one has [96]

$$S_N \propto t^{d_s/2} \,, \quad t \to \infty \,, \tag{8}$$

where d_s is the spectral dimension and Eq. (8) holds for $d_s < 2$ [14,96], which corresponds to low-dimensional and fractal media, leading to the so-called "anomalous diffusion" behavior [19,33,67]. Furthermore, $d = 2$ is the upper critical dimension, such as for a d-dimensional regular space one has $d_s = d > 2$ and Eq. (8) becomes $S_N \propto t$, leading to classical diffusion (for further details see also Eq. (17) below).

Diffusion, or random walks, are also related to the density of states for harmonic excitations of the media [78] ($h(\epsilon)$) through the probability that the random walk returns to the origin. So, one has

$$h(\epsilon) \propto \epsilon^{\frac{d_f}{d_w}-1} \equiv \epsilon^{\frac{d_s}{2}-1} \,, \tag{9}$$

where the energy and vibrational density of states are related through $h(\epsilon)d\epsilon = g(\omega)d\omega$. The relationship (9) is similar to the well-known density of states in Euclidean space, except that d has to be replaced by

$$d_s = 2\frac{d_f}{d_w} \,, \tag{10}$$

which is also known as the fracton dimensionality, since the vibration modes are called fractons instead of phonons.

Very recently we have found evidence of discrete scale invariance in the *time* domain by measuring the relaxation of the magnetization in the Ising model on Sierpinski Carpets [25,26]. Subsequently, we have conjectured that physical processes characterized by an observable $O(t)$, occurring in fractal media with DSI, and that develop a monotonically increasing time-dependent characteristic length $\xi(t)$, may also exhibit *time* DSI. In fact, by assuming $\xi \propto t^{1/z}$, where z is a dynamic exponent, it can be shown that $O(t)$ has to obey time DSI according to [53]

$$O(t) = Ct^{\alpha/z} F\left(\phi + \frac{\log(t)}{\log(a_1^z)} \right) \,, \tag{11}$$

where C and ϕ are constants. So, the conjecture given by Eq. (11) implies the existence of a logarithmic periodic modulation of time observables characterized by a time-scaling ratio τ given by

$$\tau = a_1^z \,, \tag{12}$$

see also Eq. (4). So, in the case of random walks on fractal media with DSI, Eqs. (7) and (8) have to be generalized according to Eq. (11) with $d_w = z$.

Diffusion-limited Reactions

Reaction kinetics is influenced by the characteristic time of the processes involved. The rate of a heterogeneous reaction is often determined by the adsorption of the reactants from a fluid phase on the catalyst surface. For reactions occurring in a single phase the transport of the reactants and the time of reaction influence the overall reaction rate. Here, we focus on diffusion-limited reaction processes for which the transport time, given by the typical time required for the reactants to meet, is much longer than the reaction time. Since the diffusion of particles in fractals is often anomalous, it is expected that this behavior will affect diffusion-limited reactions. Furthermore, one also needs to account for density fluctuations of the reactants occurring at all length scales. For these reasons, as well as for the occurrence of many particle effects, the study of diffusion-limited reactions is difficult and often eludes straightforward mean-field approaches.

In this section, simple and archetypical reactions occurring in fractal media will be discussed. For a more general overview of diffusion-limited reactions in homogeneous and disordered media, see e. g. [33,67].

One-Species Reactions

So far, the most studied case is the one-species annihilation process, where species A diffuse and annihilate upon encounter [33,67], according to the reaction scheme

$$A + A \to 0 \,, \tag{13}$$

such that the reaction is instantaneous. This example also includes the case in which the product 0 is some inert particle that does not influence the overall kinetics. A closely related reaction is the one-species coalescence process given by [33,67]

$$A + A \to A \,. \tag{14}$$

In the mean-field limit, which holds for the reaction-limited case, the rate equation for both reactions (13) and (14) is given by

$$\frac{\mathrm{d}\rho_A(t)}{\mathrm{d}t} = -K\rho_A^X(t) \,, \tag{15}$$

where $\rho_A(t)$ is the concentration of A-particles at time t and K is the reaction constant. Eq. (15), where $X = 2$ is the reaction order, is the typical second-order rate equation that can be solved yielding

$$\rho_A(t) = \frac{1}{kt + \frac{1}{\rho_A(0)}} \,, \tag{16}$$

so that in the long-time limit one has $\rho_A(t) \approx t^{-1}$. On the other hand, for the diffusion-controlled regime, it can be proved rigorously that the concentration decays, for both processes (13) and (14) according to [33,67]

$$\rho_A(t) \propto \frac{1}{t^{d/2}} \,, \quad d < d_\mathrm{c} = 2 \,, \tag{17}$$

where $d_\mathrm{c} = 2$ is the upper critical dimension, such that for $d > d_\mathrm{c}$ the density decays according to the mean-field prediction, given by Eq. (16). So, for $d < 2$ the kinetics of the diffusion-controlled process is anomalous and just at d_c one has logarithmic corrections to Eq. (17).

So, let us focus our attention on heterogeneous chemical reactions occurring in media having a fractal structure with $d_\mathrm{f} < d_\mathrm{c} = 2$. Typical examples for these structures are catalysts, porous glass [12], diffusion-limited aggregates [107], percolation clusters [79], zeolites, etc. Since, as already discussed, diffusion in such media is anomalous [14,96], so are expected to be diffusion-limited chemical reactions [13,15,17,18,45,80,92,103]. Then, the simple textbook elementary reaction given by Eq. (15) has to be modified. A straightforward way to understand the nature of the modifications can be envisioned just by considering that S_N (see Eq. (8)) is the exploration space of the random walk, so that the rate constant K can be written in terms of the visitation efficiency γ [80] as follows

$$K \propto \gamma \equiv \frac{\mathrm{d}S_N}{\mathrm{d}t} \propto t^{d_\mathrm{s}/2-1} \,. \tag{18}$$

Of course, for standard diffusion one has $d_\mathrm{s} = 2$, K is actually a constant and the classical reaction order $X = 2$ is recovered. However, it should be noticed that for anomalous diffusion one has $d_\mathrm{s} < 2$, so that K is no longer a constant but depends on time. In fact, by inserting Eq. (8) into Eq. (18) it follows that

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} \propto -t^{d_\mathrm{s}/2-1}\rho^2 \,, \quad t \to \infty \,, \tag{19}$$

and the integrated rate equation that is obtained by using Eq. (19) reads

$$\rho^{-1} - \rho_o^{-1} = kt^{d_\mathrm{s}/2} \,, \tag{20}$$

where $\rho_o \equiv \rho(t = 0)$ is the initial density of random walks and, of course, Eqs. (19) and (20) are expected to hold in the low density limit since they are derived from the behavior of a single walker. Now, by comparing Eqs. (8) and (20), it follows that both the asymptotic (low-density) regime of the density of species in the diffusion-limited reaction and the exploration space of the random walk are

governed by the same exponent, given by the spectral dimension of the media where the physical process actually takes place. Furthermore, by replacing Eq. (20) into Eq. (19), it follows that

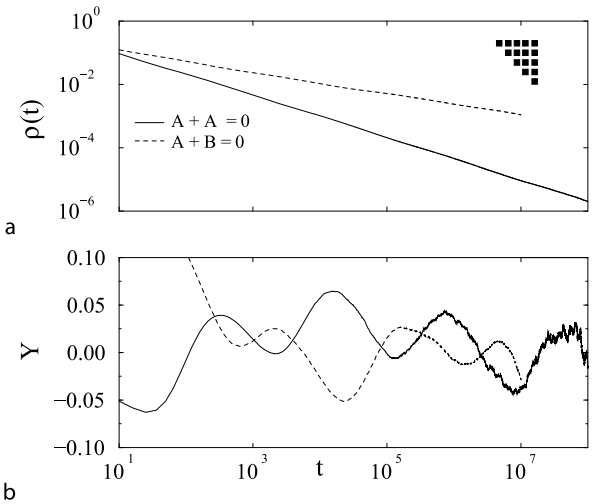
$$\frac{d\rho}{dt} = -k\rho^{1+2/d_s}, \quad t \to \infty, \tag{21}$$

and the reaction order for the elementary reaction given by Eq. (15) can be written as [15,17,45,81,92]

$$X = 1 + \frac{2}{d_s}, \tag{22}$$

which yields the textbook result ($X = 2$) for classical diffusion, but one has $X > 2$ for anomalous diffusion that is expected to occur in low-dimensional and fractal media for $d_s < 2$, e. g. for $d = 1$ one has $d_s = 1$ and the reaction order is $X = 3$. This relationship for the reaction order, based on the conjecture on the time dependence of the effective rate constant given by Eq. (18), was later derived rigorously by Clément et al. [16] based on calculations of the pair-correlation function and the macroscopic reaction law for Euclidean dimensions $d \leq 3$ and self-similar fractal structures with spectral dimensions $1 \leq d_s \leq 2$.

Figure 3a shows the time dependence of the decay of the density of random walks as a result of the reaction given by Eq. (20). The results correspond to the Sierpinski Gasket SG(5,10). A power-law decay with exponent $d_s/2 = 0.661 \pm 0.003$ is observed in agreement with Eq. (20). However, a careful inspection of the curve allows us to distinguish soft log-periodic oscillations modulating the power-law behavior. These oscillations become even more evident in Fig. 3b, which was obtained from the original data already shown in Fig. 3a after subtraction and subsequent normalization by the fitted power law. The presence of log-periodic oscillations is the signature of the occurrence of *time* DSI (see Eq. (11)) due to the coupling of the reaction kinetics and the *spatial* DSI of the substrate (see Eq. (4)), as discussed in Sect. "Fractals and Some of Their Relevant Properties" and "Random Walks".



Reaction Kinetics in Fractals, Figure 3
a Log-log plots of the decay of the density of reacting species as obtained by means of Monte Carlo simulations in the SG(5,10) whose generating cell is shown in the *upper-right corner*. b Linear-log plot of the raw data shown in a but after proper subtraction and subsequent normalization by the corresponding power-law behavior fitted in a. In both panels, the *dashed* and *full lines* correspond to the reactions $A + B \to 0$ and $A + A \to 0$, respectively. More details in the text

Two-Species Annihilation

The study of the two-species diffusion-limited reaction of the type

$$A + B \to 0 \tag{23}$$

has received considerable attention due to the fact that, most likely, it is the simplest case where fluctuations in the local concentration of the reactants play a prevailing role in the emerging kinetic behavior [33,38,43,67,76,103]. According to Eq. (23) the concentration difference of both species is conserved, so that $\rho_A(t) - \rho_B(t) = \rho_A(0) - \rho_B(0) \equiv$ const. If the initial concentrations are different, the mean-field approach predicts an exponential decrease of the minority species [33,67]. However, for $\rho_A(0) = \rho_B(0)$ one expects an algebraic decay of the form $\rho \propto 1/t$.

Let us now briefly review the effect caused by spatial fluctuations in the concentrations [33,38,43,67,76,103]. For the sake of simplicity, let us assume that both species have the same diffusion constant (D) and that the initial concentrations are equal. For a random initial distribution of particles, a region of the space of linear size l would initially have $N_A = \rho_A(0)l^d \pm \sqrt{\rho_A(0)l^d}$ particles of type A, and equivalently for B-particles. The second term accounts for the local fluctuation, at a scale of the order of l, in the number of particles. By assuming classical diffusion, one has that for a time t, such as $Dt \simeq l^{1/2}$, all the particles within the considered region would have time enough to react with each other. In the absence of fluctuations, the concentrations of both species will vanish. However, at the end of the reaction process one still has of the order of $\sqrt{\rho(0)l^d}$ particles of the majority species. Then, the concentration of either A- or B-particles at time t is actually

given by

$$\rho_{A,B} \propto \frac{\sqrt{\left(\rho_{A,B}(0) l^d\right)}}{l^d} \propto \frac{\sqrt{\left(\rho_{A,B}(0)\right)}}{D t^{d/4}} , \qquad (24)$$

and the decay exponent is twice as small than in the case of the one-species processes (Eq. (17)). According to the above scaling reasoning, one would expect the formation of alternating domains of A- and B-particles. This segregation of the reactants, known as the Toussaint–Wilczek effect [103], causes the slowdown of the kinetics since the reaction actually takes place along the boundary between domains. A careful analysis reveals that for $d \leq 4$, the domains are unstable, so that $d_c = 4$ is the upper critical dimension above which the mean-field approach holds. So, Eq. (24) holds for $d \leq 4$, and segregation has qualitatively been observed, e. g. in $d = 2$, by means of simulations [103] (see also Fig. 3). In order to gain a quantitative insight into the segregation of the reactants, it is useful to evaluate the excess density of either A- or B-particles within spatial domains of size l^d, defined as

$$\Delta\rho_{A,B}(l) = \frac{|N_A - N_B|}{l^d} , \qquad (25)$$

where $N_A(N_B)$ is the number of A- (B-) particles within the domain.

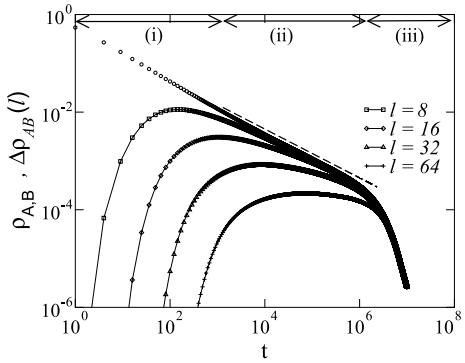
Figure 4 shows log-log plots of the time dependence of the decay of the density of particles, as well as that of the excess density, as measured for different values of l. Results are obtained by means of Monte Carlo simulations in $d = 2$ dimensions. Three different regimes can be observed, namely: (i) For the short-time behavior ($t \leq 10^3$ in Fig. 4) one observes deviations from Eq. (24), which are in agreement with the fact that at early times the segregation of the reactants, as evidenced by the measurements of the excess density, is almost negligible. (ii) During an intermediate time regime ($10^3 < t \leq 10^6$ in Fig. 4) the Toussaint–Wilczek effect becomes dominant – the excess densities are maxima in all measured length scales – and the density decays according to Eq. (24) with exponent $d/4 = 1/2$. (iii) Finally, for long times ($t > 10^6$ in Fig. 4) one observes the fast decay of both the density and the excess density due to the finite size of the sample used.

Let us now analyze the expected behavior for the case of reactions occurring in fractals. Here one has that the characteristic scaling time is of the order of $t \propto l^{1/d_w}$, and according to Eq. (24) the residual concentration can be written as

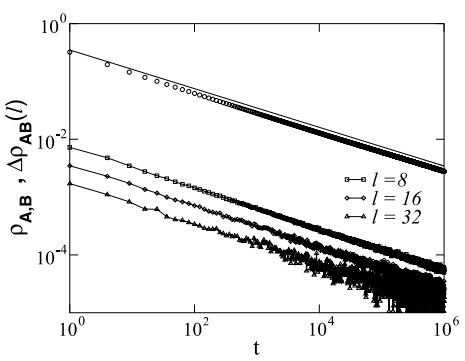
$$\rho_{A,B}(t) \propto \frac{l^{d_f/2}}{l^{d_f}} \propto \frac{t^{d_f/2d_w}}{t^{d_f/d_w}} \propto \frac{1}{t^{d_s/4}} , \qquad (26)$$

where Eq. (10) has been used. Therefore, for fractals one has to replace d by the spectral dimension d_s of the fractal in Eq. (24).

Figure 5 shows log-log plots of the time dependence of the decay of the density of particles and the excess density, as in the case of Fig. 4, but obtained by means of Monte Carlo simulations on random-fractal substrates given by percolation clusters at the critical threshold (see



Reaction Kinetics in Fractals, Figure 4
Log-log plots of the density of species (o) and the excess density measured for different length scales as indicated. Results obtained for the $A + B \rightarrow 0$ reaction (Eq. (23)) in $d = 2$ dimensions by using a lattice of side $L = 4096$. Results are averaged over 100 different realizations. The *double-arrow lines* at the *top* indicate the three different time regimes discussed in the text. The length scales used to measure the excess density are identified by the corresponding symbols. The *dashed line*, with slope $d/4 = 1/2$ shows the expected behavior according to Eq. (24)



Reaction Kinetics in Fractals, Figure 5
Log-log plots of the density of species (o) and the excess density measured for different length scales as indicated. Results obtained for the $A + B \rightarrow 0$ reaction (Eq. (23)) in incipient percolation clusters in $d = 2$ dimensions by using a lattice of side $L = 4096$. Results are averaged over 100 different realizations. The length scales used to measure the excess density are identified by the corresponding symbols. The *full line* (slightly shifted *up* for the sake of clarity), with slope $d_s/4 = 0.3344$, is the best fit of the data according to Eq. (26)

Fig. 2). In contrast to the $d = 2$ dimensional case discussed within the context of Fig. 4, the disordered nature of the fractal causes the occurrence of density fluctuations at all the spatial scales even for the initial configuration. So, the Toussaint–Wilczek effect starts to play a dominant role from the very beginning of the reaction. Consequently, the density decays according to Eq. (26) along the whole time interval measured. The best fit of the data yields $d_s/4 = 0.3344$, in excellent agreement with reported values [20]. Of course, finite-size effects, which are not observed in Fig. 5, could be expected to occur at later times. It is also worth mentioning that Fig. 5 confirms that the excess density also decays with the same power-law behavior as the density, in agreement with the arguments outlined in order to obtain Eqs. (24) and (26).

Figure 3a also shows the time dependence of the decay of the density of particles for the $A + B \to 0$ reaction given by Eq. (26). These results were obtained in the SG(5,10). The best fit of the data gives $d_s/4 = 0.331 \pm 0.003$, in excellent agreement with the value obtained for the annihilation reaction (Eq. (20)) given by $d_s/2 = 0.663 \pm 0.003$ (see also Fig. 3a). The soft log-periodic oscillation of the density that can be observed in Fig. 3a is enhanced in Fig. 3b. Again, this behavior is the fingerprint of the occurrence of *time* DSI, as discussed in Sect. "Fractals and Some of Their Relevant Properties" and "Random Walks" (see Eqs. (4) and (11), respectively).

Summing up, for the simplest example of a bimolecular irreversible decay of the form $A + B \to 0$ the obtained universal kinetic law depends on the dimensionality, the initial concentration and the particle-conservation law that holds for the system.

So far, the transient $A + B \to 0$ reaction is fundamentally different from the $A + A \to 0$ reaction due to the Toussaint–Wilczek effect. This effect represents a quite delicate balance that may become elusive to be properly identified, as discussed for the case of $d = 2$ dimensions (Fig. 4). Therefore, it is interesting to study its persistence under tiny perturbations, e. g. the steady feeding of reactants. Early simulations by Anacker et al. [93] have revealed that, in a cubic lattice, the Toussaint–Wilczek effect is destroyed by the addition of a steady source of walkers. Furthermore, in contrast to the transient result where the reaction order is $X = 3$, the steady-state $A + B \to 0$ reaction gives $X = 2.00 \pm 0.02$, for steady-state densities as low as 0.001. Since in $d = 2$ and for the transient regime, the occurrence of the effect is confined to a narrow density window (see Fig. 4) careful measurements will be needed in order to test if segregation still remains relevant in $d = 2$ under steady-state conditions. Segregation of the reactants under steady-state conditions has also been observed to occur in percolation clusters [93,100]. Furthermore, dramatic segregation effects are observed in the Sierpinski Gasket under steady-state conditions [93,100]. In this case the value $X = 2.00 \pm 0.2$ has been reported for the reaction order. All these results point out that steady-state segregation cannot simply be due to the Toussaint–Wilczek effect and the understanding of this behavior remains elusive, even after more than 20 years.

Irreversible Phase Transitions in Heterogeneously Catalyzed Reactions

Classical phase transitions, such as those observed in magnets, fluids, alloys, etc., which take place under equilibrium conditions, are reversible [28,35,111]. For example, one can change the phase of a magnet, from the ferromagnetic to the paramagnetic one, just by properly tuning the temperature. In contrast, irreversible phase transitions (IPTs) [9,10,68,85,90] occurring in reaction systems take place between an active state, characterized by a steady reaction among the reactants and the output of the products, and an inactive – also called poisoned or absorbing – state, where the reaction stops irreversibly. The inactive state is absorbing in the sense that after undergoing trapping, the system can never escape from it. Consequently, it is impossible to change from the absorbing state to the active one, just by tuning the control parameter, and the transition is irreversible.

IPTs are typically studied by using lattice-gas reaction models of heterogeneously catalyzed reactions. Within this context IPTs observed upon the catalytic oxidation of CO, H_2 and NO have extensively been studied [42,85,88,108]. Further studies of IPTs comprise the generic monomer-monomer ($A + B \to 0$) reaction [90], as well as a wide variety of reaction processes such as epidemic propagation [84], forest-fire models [8,97], prey-predator systems [51], etc., for reviews see e. g. [68]. Since a wide variety of catalysts are composed by a solid-state fractal support containing catalytically active particles [21,22,23,95], most of the above-mentioned reactions have also been studied in fractal media [1,2,5,7,41,72,86,87]. A brief overview of the state of the art in the characterization of IPTs in fractal substrates will be presented below.

The ZGB Model for the Catalytic Oxidation of Carbon Monoxide

It is well known that the catalytic oxidation of carbon monoxide, namely $2CO + O_2 \to 2CO_2$, which is likely the most studied reaction system, proceeds according to the Langmuir–Hinshelwood mechanism [50,65], i.e with

both reactants adsorbed on the catalyst surface

$$CO(g) + S \rightarrow CO(a)\,, \tag{27}$$

$$O_2(g) + 2S \rightarrow 2O(a)\,, \tag{28}$$

$$CO(a) + O(a) \rightarrow CO_2(g)\,, \tag{29}$$

where S is an empty site on the surface, while (a) and (g) refer to the adsorbed and gas phase, respectively. The reaction takes place with the catalyst, e. g. a transition metal surface such as Pt, in contact with a reservoir of CO and O_2 whose partial pressures are P_{CO} and P_{O_2}, respectively.

Equation (27) describes the irreversible molecular adsorption of CO on a single site of the catalyst surface. It is known that under suitable temperature and pressure reaction conditions, CO molecules diffuse on the surface. Furthermore, there is a small probability of CO desorption that increases as the temperature is raised [71]. Equation (28) corresponds to the irreversible adsorption of O_2 molecules that involves the dissociation of such species and the resulting O atoms occupy two sites of the catalytic surface. Under reaction conditions, both the diffusion and the desorption of oxygen are negligible. Due to the high stability of the O_2 molecule, the whole reaction does not occur in the homogeneous phase due to the lack of O_2 dissociation. So, Eq. (28) dramatically shows the role of the catalyst that makes the rate-limiting step of the reaction feasible. Finally, Eq. (29) describes the formation of the product (CO_2) that desorbs from the catalyst surface. This final step is essential for the regeneration of the catalytically active surface.

For the practical implementation of the Ziff–Gulari–Barshad (ZGB) model [108], the catalyst surface is replaced by a lattice, so a lattice-gas reaction model is actually considered. The catalyst is assumed to be in contact with an infinitely large reservoir containing the reactants in the gas phase. Adsorption events are treated stochastically neglecting energetic interactions. Furthermore, diffusion and desorption of CO are also neglected. Then, on the two-dimensional square lattice the Monte Carlo simulation algorithm for the ZGB model is as follows:

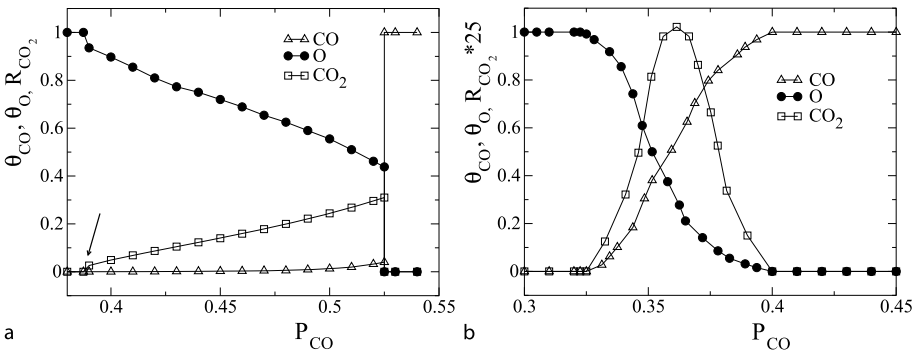
(i) CO or O_2 molecules are selected at random with relative probabilities P_{CO} and P_O, respectively. These probabilities are the relative impingement rates of both species, which are proportional to their partial pressures in the gas phase in contact with the catalyst. Due to the normalization, $P_{CO} + P_O = 1$, the model has a single parameter, i. e. P_{CO}. If the selected species is CO, one surface site is selected at random, and if that site is vacant, CO is adsorbed on it according to Eq. (27). Otherwise, if that site is occupied, the trial ends and a new molecule is selected. If the selected species is O_2, a pair of nearest-neighbor sites is selected at random and the molecule is adsorbed on them only if they are both vacant, as required by Eq. (28).

(ii) After each adsorption event, the nearest-neighbors of the added molecule are examined in order to account for the reaction given by Eq. (29). If more than one $[CO(a), O(a)]$ pair is identified, a single one is selected at random and removed from the surface.

Assuming irreversible adsorption-reaction steps, as in the case of Eqs. (28)–(29), it may be expected that in the limit of large P_{CO} and small P_{O_2} (small P_{CO} and large P_{O_2}) values, the catalyst surface would become saturated by CO- (O_2-) species and the reaction would stop. In fact, the catalyst surface fully covered by a single type of species, where further adsorption of the other species is no longer possible, corresponds to an inactive state of the system. This state is known as 'poisoned', in the sense that adsorbed species on the catalyst surface are the poison that causes the reaction to stop. Physicists used to call such a state (or configuration) 'absorbing' because a system becomes trapped by it forever, with no possibility of escape [49]. These concepts are clearly illustrated in Fig. 6a and b, which show plots of the rate of CO_2 production (R_{CO_2}) and the surface coverage with CO and O_2 (θ_{CO} and θ_O, respectively), versus the partial pressure of CO (P_{CO}), as obtained by simulating the ZGB model [108]. Figure 6a corresponds to simulations performed in homogeneous, two-dimensional lattices. In this case, for $P_{CO} \leq P_{1CO} \simeq 0.38975$ the surface becomes irreversibly poisoned by O species with $\theta_{CO} = 0$, $\theta_O = 1$ and $R_{CO_2} = 0$. In contrast, for $P_{CO} \geq P_{2CO} \simeq 0.525$ the catalyst is irreversibly poisoned by CO molecules with $\theta_{CO} = 1$, $\theta_O = 0$ and $R_{CO_2} = 0$. However, as shown in Fig. 6a, between these absorbing states there is a reaction window, namely for $P_{1CO} < P_{CO} < P_{2CO}$, such that a steady state with sustained production of CO_2 is observed.

As follows from Fig. 6a, when approaching the oxygen absorbing state from the reactive phase, all quantities of interest change smoothly until they adopt the values corresponding to the absorbing state. This behavior typically corresponds to a second-order IPT [9,10,68,85,90]. Remarkably, the behavior of the system is quite different upon approaching the CO absorbing state from the reactive regime (see Fig. 6a). In this case, all quantities of interest exhibit a marked discontinuity close to $P_{2CO} \simeq 0.525$. This is a typical first-order IPT and P_{2CO} is the coexistence point.

Experimental results for the catalytic oxidation of carbon monoxide on Pt(210) and Pt(111) [3,36] are in qualitative agreement with simulation results of the ZGB



Reaction Kinetics in Fractals, Figure 6

Phase diagrams of the ZGB model showing the dependence of the surface coverage with CO (θ_{CO}, △) and Oxygen (θ_O, •), and the rate of CO_2 production (R_{CO_2}, ⋄) on the partial pressure of CO (P_{CO}) in the gas phase. (a) Results obtained in $d = 2$ homogeneous lattices. The irreversible phase transition (IPT) occurring at $P_{1CO} \simeq 0.38975$ (second-order) is shown by an *arrow*, while at $P_{2CO} \simeq 0.525$ one has a sharp (first-order) IPT. (b) Results obtained by performing simulations in incipient percolation clusters of fractal dimension $d_f \simeq 1.89$ in $d = 2$ dimensions. Second-order IPTs occurring at $P_{1CO} \simeq 0.325$ and $P_{2CO} \simeq 0.400$ can clearly be observed. Notice that the rate of CO_2 production has been amplified by a factor 25 for the sake of clarity

model. A remarkable agreement is the (almost) linear increase in the reaction rate observed when the CO pressure is raised and the abrupt drop of the reactivity when a certain 'critical' pressure is reached. However, two essential differences are worth discussing: (i) the oxygen-poisoned phase exhibited by the ZGB model within the CO low-pressure regime is not observed experimentally. (ii) The CO-rich phase exhibiting low reactivity found experimentally resembles the CO-poisoned state predicted by the ZGB model. However, in the experiments the non-vanishing CO-desorption probability prevents the system from entering into a truly absorbing state and the abrupt, 'first-order-like' transition, observed in the experiments is actually reversible. Of course, these and other disagreements are not surprising since the lattice gas reaction model, with a single parameter, is a simplified approach to the actual catalytic reaction that is far more complex.

In order to observe the influence caused by the fractal nature of the substrate on the phase diagram of the ZGB model, Fig. 6b shows results obtained by means of simulations performed in incipient percolating clusters at the critical threshold in two dimensions (see also Fig. 2). In this case, for $P_{CO} \leq P_{1CO} \simeq 0.325$ the surface becomes irreversibly poisoned by O species, while for $P_{CO} \geq P_{2CO} \simeq 0.400$ the catalyst is irreversibly poisoned by CO molecules [2]. Also, a reaction window between these absorbing states is found. So, by comparing Fig. 6a and b at least three main differences can be identified [2]: (i) The IPT observed at a low CO pressure in homogeneous lattices becomes largely shifted towards lower pressures in the case of the IPC. This effect is due to the constraint imposed by the disordered cluster on the adsorption of oxygen molecules that need two adjacent sites for deposition. (ii) The sharp first-order IPT characteristic of the homogeneous lattice becomes of second-order for the case of the fractal. (iii) The reaction window, which is of the order of $\Delta P_{CO} = P_{2CO} - P_{1CO} \simeq 0.135$ in the homogeneous lattice, becomes narrowed up to $\Delta P_{CO} \simeq 0.085$ for the case of the fractal. These results are in agreement with the findings of Casties et al. [41], who also studied the influence of CO diffusion on the location of the critical points. In particular, it is found that P_{2CO} increases when the rate of CO-diffusion also increases.

On the other hand, simulations of the ZGB model performed in disordered clusters below the critical threshold, i. e. in the so-called 'lattice animals' in percolation [101], show that poisoning can be caused either by CO- or O-species [41,91] and that due to finite-size effects the finite-width reaction window vanishes for lattices of side $L \simeq 20$ [41,91]. These observations are consistent with the fact that the characteristic lengths of the fluctuations in the coverages are larger than the sizes of the animals, playing a dominant role in the catalytic behavior of small aggregates, such as those used in actual catalysts [21,75,106].

The common feature of Monte Carlo simulations of the ZGB model performed in Sierpinski Carpets is that the first-order IPT characteristic of the homogeneous lattice is no longer observed, so that transition becomes of second-order [5,41]. In fact, even for the $SC(5,1)$ with fractal dimension $d_f = \log(24)/\log(25) \simeq 1.975$ one observes a second-order IPT at high CO pressures [5]. Simulations performed in the $SC(3,1)$ with a fractal dimension $d_f \simeq 1.89$, which is very close to that of the IPC ran-

dom fractal, yield $P_{1CO} \simeq 0.397$ and $P_{2CO} \simeq 0.52$ [41], i. e. two critical points that are very close to those measured in $d = 2$ homogeneous media (see e. g. Fig. 6a) but far from the results obtained for the IPC (see e. g. Fig. 6b). These findings point out that the location of the critical points depends on topological properties of the fractals that are not fully accounted for a single fractal dimension. Also, the finite-width reaction window, which lies between both second-order IPTs, becomes narrowed when decreasing d_f [5,41]. It has even been suggested that for Sierpinski Carpets the reaction window may become of zero width close to $d_f \approx 1.6$ [5]. This statement is in agreement with the fact that Casties et al. [41] have reported that the reaction window of the ZGB model in an SC with $d_f = 1.59$ is very narrow, i. e. $\Delta P_{CO} \simeq 0.02$ It is also worth mentioning that, in $d = 1$, the ZGB model lacks a finite-width reaction window.

For second-order IPTs, as in the case of their reversible counterparts, it is possible to define an order parameter, which for the former is given by the concentration of minority species (θ_{CO}, in the case of the second-order IPT of the catalytic oxidation of CO). Furthermore, it is known that θ_{CO} vanishes according to a power law upon approaching the critical point [64], so that

$$\theta_{CO} \propto (P_{CO} - P_{1CO})^{\beta} , \tag{30}$$

where β is the order parameter critical exponent and P_{1CO} is the critical point. Careful studies performed in the $d = 2$ homogeneous lattice [63,73,104] have confirmed that this IPT belongs to the universality class of directed percolation [62,82] in $d + 1$ dimensions.

Generic Lattice-Gas Reaction Models of Catalyzed Reactions

The simplest reaction system studied proceeding according to the Langmuir–Hinshelwood mechanism [50,65] is the monomer-monomer (MM) model. In the MM model A- and B-species adsorb, react and desorb according to the following scheme

Adsorption:

$$A(g) + S \to A(a) , \tag{31}$$

$$B(g) + S \to B(a) , \tag{32}$$

Reaction:

$$A(a) + B(a) \to AB(g) + 2S , \tag{33}$$

Desorption:

$$A(a) \to A(g) + S , \tag{34}$$

$$B(a) \to B(g) + S , \tag{35}$$

where S is an empty site on the surface, while (a) and (g) refer to the adsorbed and gas phase, respectively. Adsorption of A- and B-species takes place with probability p_A and p_B, respectively. So, by taking $p_A + p_B = 1$ one has a single parameter for the adsorption process, given by $p_{ad} \equiv p_A = 1 - p_B$. Also the rates of reaction and desorption of A- and B-species are k_R, P_{dA} and P_{dB}, respectively.

By properly selecting the parameters one can study different regimes of the MM model. Let us first consider the *adsorption-limited* regime that corresponds to $k_R \equiv \infty$ and $P_{dA} = P_{dB} \equiv 0$. Here, for $p_{ad} \neq 1/2$ one has that any sample becomes irreversibly poisoned by the majority species and the final states are always absorbing. However, for $p_{ad} = 1/2$ a slow coarsening process with the formation of solid domains of A- and B-species on the surface of the catalyst is observed. By mapping the MM model into a kinetic Ising model, it can be proved that any finite sample of N_s sites will ultimately become irreversibly poisoned due to fluctuations in the coverages after a time T_P of the order of $T_P \propto N_s \ln(N_s)$ [32]. This exact result has early been predicted by means of numerical simulations by ben-Avraham et al. [4].

On the other hand, by considering desorption of a single species, say $P_{dA} \equiv 0$ and $P_{dB} > 0$, one has that the MM model exhibits a second-order IPT between an active state and a single poisoned state with A-species [6]. The critical points are found to be $P_{dB} = 0.5099 \pm 0.0003$, $P_{dB} = 0.6150 \pm 0.0003$ and $P_{dB} = 0.5562 \pm 0.0003$ for the case of $d = 2$, $d = 1$ homogeneous substrates and the IPC in $d = 2$ [54], respectively.

Interesting results are also obtained by taking $P_d \equiv P_{dA} = P_{dB}$, as proposed by Fichthorn et al. [46]. In fact, by means of numerical simulations, they have shown that for that choice of parameters the MM with desorption exhibits a noise-induced transition from monostability to bistability. This result was subsequently also obtained by means of a mean-field approach [44]. Later on, the occurrence of the transition was proved rigorously [32]. On the other hand, for a fractal lattice and in the limit of $P_d \to 0$ it has been conjectured that any finite sample of N_s sites would become saturated by either A- or B-species when

$$\chi \equiv P_d^{d_s/2} N_s \simeq 1 \tag{36}$$

where d_s is the spectral dimension of the underlying fractal and χ is a reduced parameter [55]. Furthermore, for an intermediate rate of desorption, a simple scaling behavior, which depends on the spectral dimension, is expected. In fact, for $\chi \gg 1$, a segregation regime with a reaction rate

per site $R(P_d)$ given by

$$R(P_d) \simeq P_d^{1-d_s/2} \tag{37}$$

has been predicted [55]. Also, for $\chi \ll 1$, one expects the saturation of any finite cluster due to a fluctuation-dominated regime, and the rate of reaction is expected to depend on the cluster size according to

$$R(P_d) = N_s P_d \,. \tag{38}$$

All these analytical results, namely equations (36), (37) and (38), have been validated by means of numerical simulations by using IPCs in $d = 2$ and $d = 3$ Euclidean dimensions [55] as fractal substrates.

Another interesting generic system is the dimer-dimer (DD) lattice-gas reaction model, which is described by the following Langmuir–Hinshelwood scheme

Adsorption:

$$A_2(g) + 2S \to 2A(a) \,, \tag{39}$$

$$B_2(g) + 2S \to 2B(a) \,, \tag{40}$$

Reaction:

$$A(a) + B(a) \to AB(g) + 2S \,, \tag{41}$$

where S is an empty site on the surface, while (a) and (g) refer to the adsorbed and gas phase, respectively. In this case we are interested in the reaction-controlled regime, so that the adsorption of either A_2- or B_2-species (see e. g. Eqs. (39) and (40)) takes place with the same probability and proceeds instantaneously on every pair of empty sites left behind by the reaction (Eq. (41)) [37]. It is worth mentioning that the DD model can be *exactly* mapped onto the so-called "voter" model [37]. The voter model is a spin system with two possible orientations, namely spin-up and spin-down. In the voter model, a randomly selected spin adopts the orientation of a randomly selected neighbor. Of course, if a spin is surrounded by equally oriented spins, it does not change its orientation. So, the voter dynamics is zero-temperature in nature. In fact, the voter model defines a broad universality class that describes coarsening without surface tension [48]. The voter model is also a simple model for opinion formation [57]: individuals with different opinions, labeled by A and B, respectively, change their opinions according to the voter dynamics. A remarkable feature of the DD model, or equivalently the voter model, is that it is one of the very few models that can be solved exactly in any integer dimension d.

The order parameter of the voter model is the density of interfaces with different states of opinion (C_{AB}), which corresponds to the density of A-B interfaces in the DD model. Starting form a fully disordered configuration that maximizes the interfaces, it has been found that the the density of interface evolves according to [37]

$$C_{AB}(t) = \begin{cases} t^{-\alpha} \,, & d < 2 \\ (\ln(t))^{-1} \,, & d = 2 \\ 1 \,, & d > 2 \end{cases} \tag{42}$$

where $\alpha = 1 - d/2$. So, in the long-time regime on has that $C_{AB} \to 0$ when $d < 2$, and the coarsening of domains is observed. However, for $d > 2$ single-species domains do not appear. The borderline is the two-dimensional case where coarsening proceeds logarithmically, so that $d_c = 2$ is the upper critical dimension of the voter model.

Monte Carlo simulations of the voter model performed in both Sierpinski Gaskets and Carpets with $d_f < 2$ have confirmed the coarsening of the domains. An interesting feature observed in the plots of $C_{AB}(t)$ versus t, as obtained for various fractals, is the presence of soft log-periodic modulations of the power-law decay predicted by Eq. (42) [53,102]. As already discussed in Sect. "Fractals and Some of Their Relevant Properties" and "Random Walks" (see Eqs. (4) and (11), respectively), these oscillations are the signature of *time* discrete scale invariance caused by the coupling of the voter dynamics with the *spatial* discrete scale invariance of the underlying fractal.

Epidemic Simulations for the Characterization of IPTs

The study of second-order IPTs by using standard simulation methods is hindered by the fact that, due to large fluctuations occurring close to critical points, any finite system will ultimately become irreversibly trapped by the absorbing state. So, the measurements are actually performed within metastable states facing two competing constraints: on the one hand, the measurement time has to be long enough in order to allow the system to develop the corresponding correlations and, on the other hand, such time must be short enough to prevent poisoning of the sample. In view of these shortcomings, experience indicates that the best approach to second-order IPTs is to complement the standard approach, consisting in performing finite-size scaling studies of stationary quantities, with epidemic simulations. In fact, the application of the Epidemic Method (EM) to the study of IPTs has become a useful tool for the evaluation of critical points, dynamic critical exponents and eventually for the identification of universality classes [62,82,104]. The idea behind the EM is to initialize the simulation using a configuration very close to the absorbing state. Such a configuration can be achieved by removing some species from the center of an otherwise

fully poisoned sample. Then a small patch of empty sites is left. In the case of the ZGB model, this can be done by filling the whole lattice with O, except for a small patch. Patches consisting of 3–6 neighboring empty sites are frequently employed, but it is known that the asymptotic results are independent of the size of the initial patch. Such patch is the kernel of the subsequent epidemic.

After the generation of the starting configuration, the time evolution of the system is followed, and during this dynamic process the following quantities are recorded: (i) the average number of empty sites ($N(t)$), (ii) the survival probability $P(t)$, which is the probability that the epidemic is still alive at time t, and (iii) the average mean square distance, $R^2(t)$, over which the empty sites have spread. Of course, each single epidemic stops if the sample is trapped in the poisoned state with $N(t) = 0$ and, since these events may happen after very short times (depending on the patch size), results have to be averaged over many different epidemics. It should be noticed that $N(t)$ ($R^2(t)$) is averaged over all (surviving) epidemics.

If the epidemic is performed just at criticality, a power-law behavior (scaling invariance) can be assumed and the following *Ansätze* are expected to hold,

$$N(t) \propto t^{\eta} , \tag{43}$$

$$P(t) \propto t^{-\delta} , \tag{44}$$

and

$$R^2(t) \propto t^{z_{\mathrm{epi}}} , \tag{45}$$

where η, δ and z_{epi} are dynamic critical exponents. Thus, at the critical point log-log plots of $N(t)$, $P(t)$ and $R^2(t)$ will asymptotically show a straight line behavior, while off-critical points will exhibit curvature. This behavior allows the determination of the critical point, and from the slopes of the plots the critical exponents can also be evaluated quite accurately [62].

The validity of Eqs. (43), (44) and (45) for second-order IPTs taking place in substrates of integer dimensions ($1 \leq d \leq 3$) is very well established. Furthermore, the observation of a power-law behavior for second-order IPTs is in agreement with the ideas developed in the study of equilibrium (reversible) phase transitions: scale invariance reflects the existence of a diverging correlation length at criticality. For the ZGB model in $d = 2$ dimensions the best reported values of the exponents are [63]: $\eta = 0.2295 \pm 0.001, \delta = 0.4505 \pm 0.001$ and $z_{\mathrm{epi}} = 1.1325 \pm 0.001$. These values confirm that the IPT belongs to the universality class of directed percolation in $d + 1$ dimensions, where the values reported by Grassberger [82] are $\eta = 0.229 \pm 0.003, \delta = 0.451 \pm 0.003$ and $z_{\mathrm{epi}} = 1.133 \pm 0.002$.

Early epidemic studies of the ZGB model performed in both the IPC and the SC(3,1), with almost the same fractal dimension of $d_{\mathrm{f}} \simeq 1.89$, confirmed that the IPTs are of second-order [7]. The values of the critical exponents obtained for both types of fractals are almost the same, within error bars, and are given by $\eta \approx 0.23$ and $\delta \simeq 0.41$. These values are also in agreement with epidemic calculations of the contact process [84] performed by Jensen [72], who has reported $\eta = 0.235 \pm 0.010$ and $\delta \simeq 0.40 \pm 0.01$. Subsequently, Gao and Yang [110] have studied the influence of the lacunarity of the SC on the epidemic behavior of the ZGB model. The epidemic exponents are very different for lattices of the same fractal dimension but different lacunarity, suggesting that a single fractal dimension is not enough to fully specify the critical points and exponents of the critical process. This finding is in qualitative agreement with the same conclusion obtained by means of stationary measurements and discussed in the previous section.

It is worth mentioning that the better statistics and the longer measurement time achieved by Gao et al. [110], as compared e. g. to references [7,72], allow us to observe a subtle log-periodic modulation of the power-law behavior of some dynamic observables. This effect is particularly clear for the case of the time dependence of $N(t)$ and it is the signature of *time* discrete-scale invariance (DSI). In fact, since the underlying fractals used by Gao et al. [110] exhibit *spatial* DSI with a fundamental scaling ratio $a_1 = 4$ (see also the discussion on Eq. (4)), we expect that the dynamics of any critical process occurring in those fractals would become coupled to the spatial DSI [53] through the growing correlation length $\xi \propto t^{1/z}$, where z is a dynamic exponent. This coupling would ultimately lead to *time* DSI characterized by a soft log-periodic modulation of the power laws, according to Eq. (11). Then the characteristic time-scaling ratio τ is given by Eq. (12) with $z = 2/z_{\mathrm{epi}}$. Of course, the accurate evaluation of τ by fitting the reported epidemic measurements is no longer possible because it will require a careful fit of the data. However, based on the fact that $z_{\mathrm{epi}} \simeq 1$ [110], we predict $\tau \simeq 4^2 \simeq 16$.

The Constant Coverage Ensemble Applied to the Study of First-Order IPTs

It is well known that first-order transitions are characterized by hysteresis effects, long-lived metastabilities, etc., which make difficult their study by means of conventional simulation methods. In order to avoid these shortcomings,

Brosilow and Ziff [39] have proposed the constant coverage (CC) ensemble for the study of first-order IPTs. The method has early been applied to the ZGB model [39], and subsequently the same authors [11] have also performed CC simulations of the model proposed by Yaldran and Khan [88] for the NO + CO catalyzed reaction.

For the case of the ZGB model, the CC method can be implemented by means of the following procedure. First, a stationary configuration of the system is achieved using the standard algorithm, as described in Sect. "The ZGB Model for the Catalytic Oxidation of Carbon Monoxide". For this purpose, one selects a value of the parameter close to the coexistence point, e. g. $P_{CO} = 0.51$. After achieving the stationary state the simulation is actually switched to the CC method. So, in order to maintain the CO coverage as constant as possible around a prefixed value θ_{CO}^{CC}, only oxygen (CO) adsorption attempts take place whenever $\theta_{CO} > \theta_{CO}^{CC}$, ($\theta_{CO} < \theta_{CO}^{CC}$). Let $\mathcal{N}_{CO}$ and $\mathcal{N}_{OX}$ be the number of carbon monoxide and oxygen attempts respectively. Then, the value of the "pressure" of CO in the CC ensemble (P_{CO}^{CC}) is determined just as the ratio $P_{CO}^{CC} = \mathcal{N}_{CO}/(\mathcal{N}_{CO} + \mathcal{N}_{OX})$. Subsequently, the coverage is increased by an small amount, say $\Delta\theta_{CO}^{CC}$. A transient period τ_P is then disregarded for the proper relaxation of the system to the new θ_{CO}^{CC} value, and finally averages of P_{CO}^{CC} are taken over a certain measurement time τ_M. The CC ensemble allows to determine spinodal points and investigate hysteresis effects at first-order IPTs [11,29,39], which are expected to be relevant as compared with their counterpart in equilibrium (reversible) conditions [9,10,85].

Future Directions

In spite of the large effort involved in the study of the kinetics of reactions occurring in low-dimensional and fractal media, the subject still poses both theoretical and experimental challenges whose understanding would be of a wide interdisciplinary interest.

It is largely recognized that the development and characterization of new materials and nanostructures is essential in order to boost the technology of the present century. Within this context, one has that reaction kinetics in nanosystems and confined materials exhibits a quite interesting physical behavior that is very different from that of the bulk. The average number of reactants in a nanosystem is typically very small and eventually it may be of the order of the fluctuations. So, mean-field-like treatments are far from being useful and new analytical tools, often guided by numerical simulations, need to be developed. Relevant experiments in this field are the study of the relaxation dynamics of photoexited charge carriers in semiconducting nanotubes and nanoparticles [56]. Also, the presence of catalytically active micro- and nanoparticles could lead to a considerable enhancement of the reaction rate of heterogeneously catalyzed reactions, the exhaustive understanding of this phenomena being an open challenge [30,75,106].

During the 1990s physicists and mathematicians started the study of the properties of complex networks opening a new interdisciplinary branch of science that also involves biology, ecology, economics, social and computer sciences, and others. Early studies in the science of complex networks were aimed to discover general laws governing the creation and growth of such structures. Subsequently, growing interest has been addressed to the understanding of processes taking place in networks, such as e. g. the description of the behavior of random walkers, the characterization of qualitatively new critical phenomena, the study of the kinetics of reaction-diffusion processes, the discovery of the emergency of complex social behavior, etc. Due to the unique properties of the networks, which involve, among others, the combination of the compactness, their inherent inhomogeneity, their complex architectures, etc., reaction-diffusion processes in those media are essentially different from both classical and anomalous results already discussed. In fact, annihilation rates are abnormally high, no segregation is observed in the archetypical $A + B \to 0$ reaction, and depletion zones are absent in the annihilation $A + A \to 0$ reaction [69]. So, we expect that this will be an active field of interdisciplinary research in the study of reaction kinetics in complex media.

The recently formulated conjecture that the coupling between the spatial discrete scale invariance of fractal media and the dynamics and kinetics of physical processes occurring in those media may lead to the observation of time discrete scale invariance [53], opens new theoretical perspectives for the study of diffusion-limited reactions. This kind of study is further stimulated by recent reports suggesting that spatial discrete scale invariance is no longer restricted to deterministic fractals, but it may also emerge as a consequence of different growing mechanisms in some disordered fractals, such as e. g. diffusion-limited aggregates, animals in percolation, etc. [66,74]. In order to achieve some progress in this subject, of course, extensive simulations and careful measurements would be necessary because the phenomenon is quite subtle and the standard averaging methods may destroy it as if was simple noise [66].

The propagation of fronts has attracted considerable attention since its understanding is relevant in many areas of research and technology. Among others, some examples

of front propagation include: material growth and interfaces, forest-fire and epidemic propagation, particle-diffusion fronts, chemical pattern formation, biological invasion problems, displacement of an unstable phase by a stable one close to first-order phase transitions, etc. Within this context, the properties of reaction-diffusion fronts in simple reactions where the reactants are initially segregated, such as $A + B \rightarrow O$, have extensively been studied [31,47]. However, most of the studies addressed reaction-diffusion fronts occurring in homogeneous media while less attention has been devoted to the case of, increasingly important inhomogeneous media [98]. Inhomogeneities can have different origin: i. e. physical, chemical, geometrical, etc. It is expected that the study of front propagation generated by diffusion-limited reactions in fractal media would also be a promising field of future research, where anomalous diffusion may lead to the occurrence of very interesting phenomena. In a related context, theoretical studies of the diffusion of *A*- and *B*-species in the bulk of a confined (homogeneous) media, leading to reaction at the surface with desorption of the product, predict the development of an annihilation catastrophe [98]. This phenomenon is due to the self-organized explosive growth of the surface concentration of species leading to an abrupt desorption peak. The influence of anomalous diffusion in fractal media leading to surface reaction will certainly be addressed in the near future, hopefully deserving the observation of interesting physical phenomena.

Bibliography

Primary Literature

1. Albano EV (1990) Finite-size effects in kinetic phase transitions of a model reaction on a fractal surface: Scaling approach and Monte Carlo investigation. Phys Rev B 42:R10818–R10821
2. Albano EV (1990) Monte Carlo simulation of the oxidation of carbon monoxide on fractal surfaces. Surf Sci 235:351–359
3. Albano EV (1991) On the self-poisoning of small particles upon island formation of the reactants in a model for a heterogenoeusly catalyzed reaction. J Chem Phys 94:1499–1504
4. Albano EV (1992) Critical exponents for the irreversible surface reaction $A + B \rightarrow AB$ with *B*-desorption on homogeneous and fractal media. Phys Rev Lett 69:656–659
5. Albano EV (1992) Irreversible phase transitions in the dimer-monomer surface reaction process on fractal media. Phys Lett A 168:55–58
6. Albano EV (1992) Study of the critical behaviour of an irreversible phase transition in the $A + B \rightarrow AB$ reaction with *B*-desorption on a fractal surface- Phys Rev A 46:5020–5025
7. Albano EV (1994) Critical behaviour of the irreversible phase transitions of a dimer-monomer process on fractal media. J Phys A (Math Gen) 27:431–436
8. Albano EV (1995) Spreading analysis and finite-size scaling study of the critical behavior of a forest fire model with immune trees. Physica A 216:213–226
9. Albano EV (1996) The Monte Carlo simulation method: A powerful tool for the study of reaction processes. Heter Chem Rev 3:389–418
10. Albano EV (2000) Borówko M (ed) Computational methods in surface and colloid Science. Marcel Dekker Inc, New York, pp 387–435
11. Albano EV (2001) Monte Carlo simulations of teh short-time dynamics of a first-oreder irreversible phase transition. Phys Lett A 288:73–78
12. Albano EV, Mártin HO (1988) Study of recombination reactions of particles adsorbed on fractal and multifractal substrata. Appl Phys A 47:399–407
13. Albano EV, Mártin HO (1988) Temperature-programmed reactions with anomalous diffusion. J Phys Chem 92:3594–3597
14. Alexander S, Orbach R (1982) Density of states on fractals: Fractons. J Physique Lett 43:L625–L631
15. Anacker LW, Kopelman R (1984) Fractal chemical kinetic: Simulations and experiments. J Chem Phys 81:6402–6403
16. Anacker LW, Kopelman R (1987) Steady-state chemical kinetics on fractals: Segregation of reactants. Phys Rev Lett 58:289–291
17. Anacker LW, Kopelman R, Newhouse JS (1984) Fractal chemical kinetics: Reacting random walkers. J Stat Phys 36:591–602
18. Anacker LW, Parson RP, Kopelman R (1985) Diffusion-controlled reaction kinetics on fractal and Euclidian lattices: Transient and steady-state annihilation. J Phys Chem 89:4758–4761
19. Argyrakis P, Kopelman R (1986) Fractal behaviour of correlated random walk on percolation clusters. J Chem Phys 84:1047–1048
20. Argyrakis P, Kopelman R (1992) Diffusion-controlled binary reactions in low dimensions: Refined simulations. Phys Rev A 45:5814–5819
21. Avnir D (1991) Fractal geometry: a new approach to heterogeneous catalysis. Chem Ind 12:912–916
22. Avnir D (1997) Applications of fractal geometry methods in heterogeneous catalysis. In: Ertl G, Knozinger H, Weitkamp J (eds) Handbook of heterogeneous catalysis, vol 2. Wiley-VCH, Weinheim, pp 598–614
23. Avnir D, Farin D, Pfeifer P (1983) Chemistry in noninteger dimensions between two and three. II Fractal surfaces of adsorbents. J Chem Phys 79:3566–3571
24. Bab MA, Fabricius G, Albano EV (2005) Critical behavior of an Ising system on the Sierpinski carpet: A short-time dynamics study. Phys Rev E 71:036139
25. Bab MA, Fabricius G, Albano EV (2006) Discrete scale invariance effects in the nonequilibrium critical behavior of the Ising magnet on a fractal substrate. Phys Rev E 74:041123
26. Bab MA, Fabricius G, Albano EV (2008) On the occurrence of oscillatory modulations in the power-law behavior of dynamic and kinetic processes in fractal media. EPL 81(1)
27. Barabási AL, Stanley HE (1995) Fractal concepts in surface growth. Cambridge University Press, Cambridge
28. Barber MN (1983) Domb C, Lewobitz JL (eds) Phase transitions and Critical Phenomena, vol II. Academic Press, London, p 146

29. Barzykin AV, Tachiya M (2007) Stochastic models of charge carrier dynamics in semiconducting nanosystems. J Phys C Condens Matter 19:065113
30. Bell AT (2003) The impact of nanoscience on heterogeneous catalysis. Science 14:1688–1691
31. Bena I, Droz M, Martens K, Rácz Z (2007) Reaction-diffusion fronts with inhomogeneous initial conditions. J Phys C Condens Matter 19:065103
32. ben-Avraham D, Considine DB, Meakin P, Redner S, Takayasu H (1990) Saturation transition in a monomer-monomer model of heterogeneous catalysis. J Phys A (Math Gen) 23:4297–4312
33. ben-Avraham D, Havlin S (2000) Diffussion and reactions in fractals and disordered systems. Cambridge University Press, Cambridge
34. Berry H (2002) Monte Carlo simulations of enzyme reactions in two dimensions: Fractal kinetics and spatial segregation. Biophys J 83:1891–1901
35. Binder K, Heermann D (2002) Monte Carlo simulations in statistical physics. An introduction, 4th edn. Springer, Berlin
36. Block JH, Ehsasi M, Gorodetskii V (1993) Dynamic Studies of surface reactions with microscopic techniques. Prog Surf Sci 42:143–168
37. Bordogna CM, Albano EV (2007) Statistical methods applied to the study of opinion formation models: a brief overview and results of a numerical study of a model based on the social impact theory. J Phys C Condens Matter 19:065144
38. Bramson M, Lebowitz JL (1988) Asymptotic behavior of densities in difussion-dominated annihilation reactions. Phys Rev Lett 61:2397–2400
39. Brosilow BJ, Ziff RM (1992) Comment on: NO-CO reaction on square and hexagonal surfaces: A Monte Carlo simulation. J Catal 136:275–278
40. Bunde A, Havlin S (1995) A brief introduction to fractal geometry. In: Bunde A, Havlin S (eds) Fractals in science. Springer, Berlin, pp 1–25
41. Casties A, Mai J, von Niessen W (1993) A Monte Carlo study of the CO oxidation on probabilistic fractal. J Chem Phys 99:3082–3091
42. Clar S, Drossel B, Schwabl F (1996) Forest fires and other examples of self-organized criticality. J Phys C Condens Matter 8:6803–6824
43. Clément E, Kopelman R, Sanders LM (1991) The diffusion-limited reaction $A + B \rightarrow 0$ on a fractal lattice. J Stat Phys 65:919–924
44. Clément E, Leroux-Hugon P, Argyrakis P (1994) Catalysis on a fractal lattice: A model for poisoning. Phys Rev E 49:4857–4864
45. Clément E, Sanders LM, Kopelman R (1989) Steady-state diffusion-controlled $A + A \rightarrow 0$ reaction in Euclidean and fractal dimensions: Rate laws and particle self-ordering. Phys Rev A 39:6472–77
46. Considine DB, Redner S, Takayasu H (1989) Comment on: Noise-induced bistability in a Monte Carlo surface-reaction model. Phys Rev Lett 63:2857
47. Cross MC, Hohenberg PC (1993) Pattern formation outside of equilibrium. Rev Mod Phys 65:851–1112
48. Dornic I, Chaté H, Chave J, Hinrichsen H (2001) Critical coarsening without surface tension: The universality class of the Voter model. Phys Rev Lett 87:045701
49. Ehsasi M, Matloch M, Franck O, Block JH, Christmann K, Rys FS, Hirschwald W (1989) Steady and nonsteady rates of reaction in a heterogeneously catalyzed reaction: Oxidation of CO on platinum, experiments and simulations. J Chem Phys 91:4949–4960
50. Engel T (1978) A molecular beam investigation of He, CO, and O_2 scattering from Pd(111). J Chem Phys 69:373–385
51. Engel T, Ertl G (1978) A molecular beam investigation of the catalytic oxidation of CO on Pd(111). J Chem Phys 69: 1267–1281
52. Ertl G (1990) Oscillatory catalytic reactions at single-crystal surfaces. Adv Catal 37:213–277
53. Even U, Rademann K, Jortner J, Manor N, Reisfeld R (1984) Electronic energy transfer on fractals. Phys Rev Lett 42:2164–2167
54. Fichthorn K, Gulari E, Ziff RM (1989) Noise-induced bistability in a Monte Carlo surface-reaction model. Phys Rev Lett 63:1527–1530
55. Frachebourg L, Krapivsky PL (1996) Exact results for kinetics of catalytic reactions. Phys Rev E 53:R3009–R3012
56. Gallos LK, Argyrakis P (2007) Influence of complex network substrate on reaction-diffusion processes. J Phys C Condens Matter 19:065123
57. Gao Z, Yang ZR (1999) Dynamic behavior of the Ziff–Gulari–Barshad model on regular fractal lattices: The influence of lacunarity. Phys Rev E 59:2795–2800
58. Gefen Y, Aharony A, Mandelbrot BB (1983) Phase transitions on fractals. I Quasi-linear lattices. J Phys A (Math Gen) 16:1267–1278
59. Gefen Y, Aharony A, Mandelbrot BB (1984) Phase transitions on fractals. III Infinitely ramified lattices. J Phys A (Math Gen) 17:1277–1289
60. Gefen Y, Mandelbrot BB, Aharony AY (1980) Critical phenomena on fractal lattices. Phys Rev Lett 45:855–858
61. Gouyet JF (1996) Physics and fractal structures. Springer, Paris
62. Grassberger P (1989) Directed percolation in 2+1 dimensions. J Phys A (Math Gen) 22:3673–3679
63. Grassberger P, de la Torre A (1979) Reggeon field theory (Schlögl's first model) on a lattice: Monte Carlo calculations of critical behaviour. Ann Phys (New York) 122:373–396
64. Grinstein G, Lai Z-W, Browne DA (1989) Critical phenomena in a nonequilibrium model of heterogeneous catalysis. Phys Rev A 40:4820–4823
65. Grinstein G, Muñoz MA (1997) Garrido PL, Marro J (eds) Fourth Granada lectures in computational physics. Springer, Berlin, p 223
66. Gálfi L, Rácz Z (1988) Properties of the reaction front in an A+B=C type reaction-diffusion process. Phys Rev A 38:3151–3154
67. Havlin S, ben-Avraham D (1987) Diffusion in disordered media. Adv Phys 36:695–798
68. Hinrichsen H (2000) Non-equilibrium critical phenomena and phase transitions into absorbing states. Adv Phys 49:815–958
69. Huang Y, Ouillon G, Saleur H, Sornette D (1997) Spontaneous generation of discrete scale invariance in growth models. Phys Rev E 55:6433–6447
70. Imbhil R, Ertl G (1995) Oscillatory kinetics in heterogeneous catalysis. Chem Rev 95:697–733
71. Imbihl R (1993) Oscillatory reactions on single crystal surfaces. Prog Surf Sci 44:185–343
72. Jensen I (1991) Non-equilibrium critical behaviour on fractal lattices. J Phys A (Math Gen) 24:L1111–L1117

73. Jensen I, Fogedby H, Dickman R (1990) Critical exponents for an irreversible surface reaction model. Phys Rev A 41:R3411–R3414
74. Johansen A, Sornette D (1998) Evidence of discrete scale invariance in DLA and time-to-failure by canonical averaging. Int J Mod Phys C 9:433–447
75. Johnson BFG (2003) Nanoparticles in catalysis. Topics catal 24:147–159
76. Kang K, Redner S (1984) Scaling approach for the kinetics of recombination processes. Phys Rev Lett 52:955–958
77. Kang K, Redner S (1985) Fluctuation-dominated kinetics in diffusion-controlled reactions. Phys Rev A 32:435–447
78. Kittel C (1986) Introduction to solid state physics, 8th edn. Wiley, New York
79. Kopelman R (1976) In: Fong FK (ed) Radiationless processes in molecules and condensed phases, vol 15. Springer, Berlin, p 297
80. Kopelman R (1986) Rate processes on fractals: Theory, simulations, and experiments. J Stat Phys 42:185–200
81. Kopelman R (1988) Fractal reaction kinetics. Science 241:1620–1626
82. Krapivsky PL (1992) Kinetics of monomer-monomer surface catalytic reactions. Phys Rev A 45:1067–1072
83. Lee Koo Y-E, Kopelman R (1991) Space- and time-resolved diffusion-limited binary reaction kinetics in capillaries: experimental observation of segregation, anomalous exponents, and depletion zone. J Stat Phys 65:893–918
84. Liggett TM (1985) Interacting particle systems. Springer, New York
85. Loscar E, Albano EV (2003) Critical behaviour of irreversible reaction systems. Rep Prog Phys 66:1343–1382
86. Mai J, Casties A, von Niessen W (1992) A model for the catalytic oxidation of CO on fractal lattices. Chem Phys Lett 196:358–362
87. Mai J, Casties A, von Niessen W (1993) A Monte Carlo simulation of the catalytic oxidation of CO on DLA clusters. Chem Phys Lett 211:197–202
88. Maltz A, Albano EV (1992) Kinetic phase transitions in dimer-dimer surface reaction models studied by means of mean-field and Monte Carlo methods. Surf Sci 277:414–428
89. Mandelbrot BB (1983) The fractal geometry of nature. Freeman, San Francisco
90. Marro J, Dickman R (1999) Nonequilibrium phase transitions and critical phenomena. Cambridge Univ Press, Cambridge
91. Meakin P, Scalapino DJ (1987) Simple models for heterogeneous catalysis: Phase transition-like behavior in nonequilibrium systems. J Chem Phys 87:731–741
92. Newhouse JS, Kopelman R (1985) Fractal chemical kinetics: Binary steady-state reaction on a percolating cluster. Phys Rev B 31:1677–1678
93. Newhouse JS, Kopelman R (1988) Steady-state chemical kinetics on surface clusters and islands: Segregation of reactants. J Phys Chem 92:1538–1541
94. Parus SJ, Kopelman R (1989) Self-ordering in diffusion-controlled reactions: Exciton fusion experiments and simulations on naphthalene powder, percolation clusters, and impregnated porous silica. Phys Rev B 39:889–892
95. Pfeifer P, Avnir D (1983) Chemistry in noninteger dimensions between two and three. I Fractal theory of heterogeneous surfaces. J Chem. Phys 79:3558–3565
96. Rammal R, Toulouse G (1983) Random walks on fractal structures and percolation clusters. J Physique Lett 44:L13–L22
97. Rozenfeld AF, Albano EV (2001) Critical and oscillatory behavior of a system of smart preys and predators. Phys Rev E 63:061907
98. Shipilevsky BM (2007) Formation of a finite-time singularity in diffusion-controlled annihilation dynamics. J Phys C Condens Matter 19:065106
99. Sornette D (1998) Discrete scale invariance and complex dimensions. Phys Rep 297:239–270
100. Stanley HE (1971) Introduction to phase transitions and critical phenomena. Oxford University Press, New York
101. Stauffer D, Aharoni A (1992) Introduction to the percolation theory, 2nd edn. Taylor and Francis, London
102. Suchecki K, Hołyst JA (2006) Voter model on Sierpinski fractals. Physica A 362:338–344
103. Toussaint D, Wilczek F (1983) Particle-antiparticle annihilation in diffusive motion. J Chem Phys 78:2642–2647
104. Voigt CA, Ziff RM (1997) Epidemic analysis of the second-order transition in the Ziff–Gulari–Barshad surface-reaction model. Phys Rev E 56:R6241–R6244
105. Weiss GH (1995) A primer of random walkology. In: Bunde A, Havlin S (eds) Fractals in science. Springer, Berlin, p 119
106. Wieckowski A, Savinova ER, Vayenas CG (2003) Catalysis and electrocatalysis at nanoparticle surfaces. Marcel Dekker, New York, pp 1–959
107. Witten TA, Sander LM (1983) Diffusion-limited aggregation. Phys Rev B 27:5686–5697
108. Yaldram K, Khan MA (1991) NO-CO reaction on square and hexagonal surfaces: A Monte Carlo simulation. J Catal 131:369–377
109. Zhdanov VP, Kasemo B (1994) Kinetic phase transitions in simple reactions on solid surfaces. Surf Sci Rep 20:111–189
110. Ziff RM, Brosilow BJ (1992) Investigation of the first-order phase transition in the $A - B_2$ reaction model using a constant-coverage kinetic ensemble. Phys Rev A 46:4630–4633
111. Ziff RM, Gulari E, Barshad Y (1986) Kinetic phase transitions in an irreversible surface-reaction model. Phys Rev Lett 56:2553–2556

Books and Reviews

Binder K (1997) Applications of Monte Carlo methods to statistical physiscs. Rep Prog Phys 60:487–560

Bunde A, Havlin S (1991) Fractals and disordered media. Springer, New York

Cardy J (1997) Goddard P, Yeomas JM (eds) Scaling and renormalization in statistical physics. Cambridge University Press, Cambridge

Christmann K (1991) Introduction to surface physical chemistry. Steinkopff, Darmstadt, p 1

Egelhoff WF Jr (1982) King DA, Woodruff DP (eds) Fundamental studies of heterogeneous catalysis, vol 4. Elsevier, Amsterdam

Lindenberg K, Oshanin G Tachiya M (2007) Chemical kinetics beyond the texbook: Fluctuations, Many-particle effects and anomalous dynamics. Special issue of J Phys C Condens Matter 19

Walgraef D (1997) Spatio-temporal pattern formation. Springer, New York

Record Statistics and Dynamics

Paolo Sibani[1], Henrik, Jeldtoft Jensen[2]
[1] Institut for Fysik og Kemi, SDU, Odense, Denmark
[2] Department of Mathematics, Imperial College London, London, UK

Article Outline

Glossary

Aging The slow time evolution of a class of complex systems, which are brought into a far-from-equilibrium state by the sudden change of an external parameter, e. g. the temperature.

Complex dynamics The collective or emergent time dependent properties of interacting multi-component and multi agent systems.

Marginal stability A metastable state is marginally stable if it can be destroyed by slight perturbations.

Metastability The ability of a non-equilibrium system to remain in, or close to, the same state for a certain characteristic time.

Record In a time ordered series of random numbers, a record is an entry larger than all preceding entries.

Scale invariant process A scale invariant process looks the same under re-scaling of time and/or space variables.

Stationary process A stationary process is time homogeneous, i. e. it is invariant under time translations.

Definition of the Subject

The term *record statistics* covers the statistical properties of records within an ordered series of numerical data obtained from observations or measurements. A record within such series is simply a value larger (or smaller) than all preceding values. The mathematical properties of records strongly depend on the properties of the series from which they are extracted. These properties have been investigated for many different cases, the simplest cases perhaps being series of independent random numbers drawn from the same (arbitrary) distribution, and series produced by a diffusion process with independent random increments.

The term *record dynamics* covers the rather new idea that records may, in special situations, have measurable dynamical consequences. The approach applies to the aging dynamics of glasses and other systems with multiple metastable states. The basic idea is that record sizes fluctuations of e. g. the energy are able to push the system past some sort of 'edge of stability', inducing irreversible configurational changes, whose statistics then closely follows the statistics of record fluctuations.

Introduction

Floods, droughts and earth-quakes are spectacular, potentially disastrous, and routinely monitored extremal events. Yet, even with no large scale effects involved, record performances of all kinds attract great interest, as witnessed by the famous *Guinness Book of Records*. Aside from their popular appeal, extremes and their properties are an important topic of mathematical statistics [16,28].

Formally, a value larger (or smaller) than all its predecessors within a time ordered series of numerical data constitutes a record. The number of records in a series encodes information on how the data are produced. E. g. measuring every second the velocity of a steadily accelerating object produces a time series where each entry is a record. Conversely, with data generated by a random process, e. g. the coordinates of a drunkard taking a step in any direction with equal probability, considerably fewer records occur. An intermediate case, which is indicative of a climatic warming trend, are temperature records more frequent than expected for fluctuations around a constant [35]. More generally, record-sized events in spatio-temporal series of data, such as seismic shocks, are indicative of correlations and causal relationships [14].

In some contexts, record sized events have important implications: e. g. in population dynamics, mutants surpassing the current standard for highest reproductive success contribute disproportionately to the genetic pool of the next generation, a bias which automatically raises the bar for future improvements. This naturally leads to the expectation that record events be of significance to biological evolution, as perhaps in particular suggested by the catch-phrase 'survival of the fit*test*'. Darwinism's impact on biology – and human culture at large – can hardly

be exaggerated. In recent years computer optimization methods have appeared, e. g. Genetic Algorithms [17] and Extremal Optimization [5], which are inspired by evolution dynamics and which, implicitly or explicitly, utilize records.

In a computer program, or in a social context, the function, entity or organization in charge of memorizing the existing record and detecting its eventual obliteration is clearly distinct from the process being monitored. By contrast, record keeping in biological evolution is fully integrated in the dynamics, and we will see in examples below that record dynamics can emerge as a collective property of a collection of co-evolving organisms.

In cases where records act as seeds for an amplifying dynamical process, the magnitude of the seed bears no direct relation to the magnitude of the final outcome of the amplification. E. g. a snow avalanche can be triggered by sound vibrations above a certain threshold. However, its impact is more related to the amount of snow on the ground and to the slope and shape of the valley than to the intensity of the sound.

Two important and related characteristic of record-induced dynamics are: *i)* its temporal non-homogeneity, meaning that record fluctuations and any induced effects occur at a decreasing rate and *ii)* the presence of an embedded memory mechanism. Not coincidentally, the same properties are present in a large class of systems known as *glassy*. In the sequel, we focus on the significance of record sized fluctuations for *glassy dynamics*, expanding on the observation that record fluctuations induce transitions between metastable states [43], leaving a so-called intermittent signal as a fingerprint [44].

Complex Dynamics

Systems of many interacting components and with a coupling to an external reservoir of e. g. energy and/or particles, are usually described using Master- or Fokker–Planck equations [57]. These equations combine external and internal deterministic forces with a randomizing influence of the reservoir, and determine the probability flow between a set of allowed configurations or states. A simple yet illustrative example is the Brownian motion of a grain of pollen immersed in a fluid. The collisions with the surrounding molecules exert random forces on the molecules. Furthermore, an external force may be applied to the pollen by means of e. g. an electric field. If the motion of the grain is constrained by the walls of a container, the probability distribution of its position reaches on a certain time scale – called the relaxation time – a stationary form which only depends on the geometry of the container. In this case, all memory of the initial condition is lost. Master equations or other dynamical equations often lead to a stationary state. The latter can be unique, or it can be one out of many, each state draining a basin of attraction in configuration space. An interesting situation arises when the stationarity is only approximate, i. e. when, given enough time, it is possible to escape from a basin of attraction and enter a different one. Each attraction basin then becomes a metastable region of configuration space, which is characterized by a certain *escape time*: the time typically needed for a trajectory to exit the basin.

Gases, liquids and crystalline solids are mostly found in a stationary state where relevant physical observables, e. g. the local density, fluctuate reversibly around fixed values. A stationary fluctuation signal is invariant under time translations, its average is a constant and its autocorrelation function only depends on the difference between its two time arguments. In some cases, e. g. for equilibrium systems fluctuating near a critical temperature, correlations decay algebraically and may lack an associated time-scale. Interestingly, a similar situation arises in certain slowly driven systems, epitomized by the famous 'sand pile' model [4,22], which self-organizes into marginally stable configurations lacking characteristic space and time scales. Friction dominated granular systems, e. g. rice piles, are experimental realizations of this behavior, which is known as Self Organized Critical, or SOC.

In SOC and some other driven dissipative systems the initial state is statistically equivalent to the states subsequently visited. In contrast, complex materials such as glasses, polymers, colloids and disordered magnetic materials, undergo a slow but systematic physical aging following a rapid quench of e. g. the temperature or the particle density. Typically, these materials get trapped in metastable basins, where physical variables appear to fluctuate reversibly on observation time scales shorter than the age. During this pseudo-equilibrium regime, fluctuations have Gaussian distributions with zero averages [44]. On longer time scales, intermittent events add an asymmetric, and typically exponential, tail to the Probability Density Function (PDF) of the fluctuations. The tail is caused by the non-equilibrium drift which, through a sequence of metastable basins of subtly different nature, slowly erases the memory of the initial state and changes physical properties. e. g. the linear response to a small perturbation.

The Distribution of Records in Stationary Time Series

Record-sized entries within a stationary series of independent and identically distributed random numbers form

a sub-series whose interesting statistical properties follow from simple arguments [16,42,45]. E. g., as the magnitude required to qualify as the next record, increases with each new entry, records appear at a decreasing rate. For sufficiently long time series, the number of records found between the first and the tth draw can be shown to have a Poisson distribution, with average $\ln t$. Equivalently, if the kth record appears at times t_k, the ratios $\ln(t_k/t_{k-1})$ are independent random numbers with an exponential distribution.

Consider a sequence of independent random numbers drawn from the same distribution at times $1, 2, 3, \ldots, t$. We exclude distributions supported on a finite set, because these eventually produce a record which cannot be beaten[1]. The first number drawn is by definition a record. Subsequent trials lead to a record if their outcome is larger than the previous record. We seek the probability $P_n(t)$ of finding precisely n records in t successive trials, where $1 \leq n \leq t$. In the derivation we need the auxiliary function $P_{(1,m_1,\ldots m_{k-1})}(t)$, which is the joint probability that k records happen at times $1 < m_1, \cdots < m_{k-1}$, with $m_{k-1} \leq t$. $P_1(t)$ is simply the probability that the first outcome be largest among t. As each outcome has, by symmetry, the same probability of being the largest, it follows that $P_1(t) = 1/t$. In order to obtain two records at times 1 and m, the largest of the first $m-1$ random numbers must be drawn at the very first trial. This happens with probability $1/(m-1)$. Secondly, the mth outcome must be the largest among t. This happens with probability $1/t$, independently of the position of the largest outcome in the first $m-1$ trials. Accordingly,

$$P_{(1,m)}(t) = \frac{1}{(m-1)t} \,. \tag{1}$$

Summing the above over all possible values of m we then obtain

$$P_2(t) = \sum_{m=2}^{t} \frac{1}{(m-1)t} \approx \ln(t)/t \,, \tag{2}$$

and, in the more general case of n events, we similarly obtain

$$P_{(1,m_1,\ldots,m_{n-1})}(t) = \frac{1}{\prod_{i=1}^{n-1}(m_i-1)t} \,. \tag{3}$$

We now take $q_i = m_i - 1$ and sum over all possible values of the q_i's, leading to

$$P_n(t) = \sum_{q_1=1}^{t-n+1} \frac{1}{q_1} \cdots \sum_{q_{n-1}=q_{n-2}+1}^{t-1} \frac{1}{q_{n-1}} \frac{1}{t} \,. \tag{4}$$

[1] Apart from this constraint, the form of the distribution is immaterial.

An approximate closed form expression can now be obtained by replacing the sums by integrals, which is reasonable for $t \gg n \gg 1$. The integrals can then be evaluated, finally yielding

$$P_n(t) = \frac{(\ln t)^{n-1}}{(n-1)!} \frac{1}{t} \,. \tag{5}$$

As anticipated, Eq. 5 is a Poisson distribution, with $\ln t$ replacing the usual time argument t. Clearly, the number of records between extractions t_w and $t > t_w$ is the difference of two Poisson process, and hence itself a Poisson process with average $\ln(t) - \ln(t_w) = \ln(t/t_w)$. The discrete series of random numbers can be replaced by stationary random white noise, in which case t becomes a continuous time variable, and the restriction on the asymptotic validity of Eq. 5 for $t \gg n$ can be lifted.

Let $\overline{n(t)}$ and $\sigma_n^2(t)$ be the average and variance of the number of events in time t. As an immediate consequence of Eq. 5 we note that

$$\overline{n(t)} = \sigma_n^2(t) = \ln t \,. \tag{6}$$

Furthermore, the 'current', i. e. the average number of events per unit of time decays as

$$\frac{\mathrm{d}\overline{n}}{\mathrm{d}t} = \frac{1}{t} \,. \tag{7}$$

A third consequence of Eq. 5 is the following: Let $t_1 = 1 < t_2 < \cdots < t_k < \cdots$ be the times at which the record breaking events occur, and let $\tau_1 = \ln t_1 = 0 < \tau_2 \cdots < \tau_k = \ln t_k < \cdots$ be the corresponding natural logarithms. The stochastic variables $\Delta_k = \tau_{k+1} - \tau_k = \ln(t_{k+1}/t_k)$ are independent and identically distributed. Their common distribution is an exponential with unit average. By writing: $\tau_k = \Delta_{k-1} + \Delta_{k-2} + \cdots \Delta_1$ we have that $(\tau_{k+1} - k)/\sqrt{k}$ approaches a standard Gaussian distribution for large k. Hence, the waiting time t_k for the kth event is approximately log-normal, and the average of its logarithm grows linearly in k. By Jensen's inequality [37] we also find

$$\ln(\overline{t_k}) \geq \overline{\ln(t_k)} = k \,. \tag{8}$$

We see that the average waiting time from the kth to the $(k+1)$th event grows at least exponentially in k.

Consider finally the case in which a system is made up of α independent parts. The total number of records is the sum of contributions from each part, and hence remains Poisson distributed. For the most general case where records between times t_w and $t > t_w$ are of interest, the Poisson process has average

$$\langle n \rangle (t_w, t) = \alpha \ln(t/t_w) \,. \tag{9}$$

The cumulative distribution of the 'logarithmic waiting times' $\delta_k = \ln(t_k/t_{k-1})$ is correspondingly

$$P(\delta_k < x) = 1 - \exp(-\alpha x); \quad x > 0; \quad k = 2, 3, \ldots \quad (10)$$

In glassy dynamics, noise fluctuations are drawn from a (quasi-) equilibrium distribution characterizing the current metastable region. This distribution heavily penalizes large excursions from the metastable configuration, making record events rare and uncorrelated.

Importantly, spatially extended systems with short range interactions may break down into a set of weakly interacting subsystems, or domains, which are localized in space and which have independent fluctuations and independent metastable states. E. g. spin glasses are characterized by a thermal correlation length, corresponding to the linear size of the equilibrated domains. The latter grows slowly and remains at all times rather small compared to the linear size of the sample. Each independently thermalized domain evolves in parallel with other domains, from which it is shielded by a backbone of mainly inactive degrees of freedom. In accord with observations discussed below, fluctuation spectra of physical quantities have a Gaussian component with zero average, and a non-Gaussian tail. The former describes the sum of independent equilibrium-like fluctuations, and the latter describes the rare non-equilibrium large events, or *quakes*, which carry the drift. In this situation, the parameter α of Eq. 9 can be identified with the number of domains. Its value enters the width of the Gaussian, which grows as the square root of α, as well as the statistical weight of the tail, which grows linearly with α.

Aging and Metastability

The term glassy dynamics usually refers to the extremely slow relaxation observed in complex systems with many interacting components, where reaching a steady, time independent, state is typically far beyond experimentally accessible time scales. E. g., when melted alloys are cooled down rapidly, they do not enter a crystalline ordered state. Instead the atoms retain the amorphous arrangement characteristic of the liquid phase while the mobility of the molecules decreases by many orders of magnitude. This colossal change in the characteristic dynamical time scales hinders a structural glass at low temperature to reach thermodynamic equilibrium. Nevertheless, the properties of glasses over short observation time scales may appear to be time independent as in thermal equilibrium.

The low temperature out-of-equilibrium behavior of structural glasses and of a host of other complex materials is intriguingly similar. E. g. spin glasses [29] a class of disordered magnetic alloys, polymers [53], colloids and gels [10] and type II superconductors [31,32] all remain out of equilibrium, with their macroscopic physical properties changing algebraically or logarithmically as a function of the time elapsed since the initial quench. Small temperature and field variations applied during the aging process uncover many fascinating 'memory' effects see e. g. [19,23,59], which indicate the presence of a hierarchy of time scales for relaxation in configuration space, a hierarchy which is also explicitly uncovered in numerical studies [49,50].

The topography of complex energy landscapes has been thoroughly investigated, see e. g. [2,30,38,51,52], and various heuristic approaches have been proposed in order to link topography and dynamics. For glassy dynamics, heuristic ideas such as hierarchical mesoscopic models [20,58], and the highly popular trap model [6] have been considered. While many aspects of aging in glassy systems are known and partly understood, important open questions remain. E. g. how are configuration space properties and real space morphology to be combined in a seamless fashion? How should one weigh thermodynamical versus dynamical properties, and, more specifically what is the rôle of equilibrium statistical properties, a rôle initially strongly advocated from different camps [7,15,33], and hotly debated ever since [24,25,34]. In retrospect, knowing that systems with similar aging behavior have completely different equilibrium properties, e. g. vibrated granular matter versus thermalizing polymers, the connection between aging and equilibrium properties appears rather tenuous.

A radically different approach was initiated by Coppersmith and Littlewood who studied memory effects in Charge Density Waves [12] using a simple model with multiple attractors. The focus is there on the mechanism for attractor *selection*, and the suggestion is that the least stable attractors, are those typically selected following a quench, in spite of having little statistical weight. One development of the idea [45] points to fluctuations records, which are likewise statistically insignificant, as key events in aging dynamics.

The focus of both experiments and theory has recently shifted from macroscopic averages to the fluctuations around these averages (Buisson et al. [8]), which are observable in meso- and nanoscaled systems. As mentioned, the fluctuation spectra have a Gaussian part, pertaining to pseudo-equilibrium fluctuations around the 'current' metastable value of the macroscopic average of the quantity of interest, and an asymmetric exponential tail, describing rare and irreversible jumps from one metastable

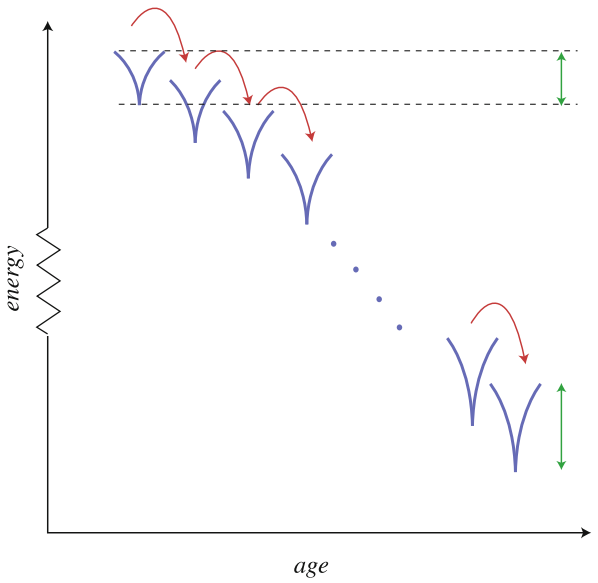
configuration to the next. Such 'quakes' carry the full macroscopic evolution of the system. Their decreasing rate is indicative of an entrenchment into gradually more stable configurations.

Marginal Stability

Broadly speaking, marginal stability relates to how metastable attractor basins are *selected* during the off-equilibrium evolution of metastable complex systems. Diverse applications include driven systems, e. g. charge density waves [12,45] and the sand-pile model [4] and its conceptual ancestor [3,41,56], thermal aging of e. g. spin-glasses [44] and type II superconductors [32]. The relevance to biological evolution is suggested by a number of model studies, some where a fitness function is explicitly defined and drives the dynamics [46,48] and others based on co-evolutionary interactions without fitness function [1].

Record size fluctuations are irrelevant in stationary processes, where the possibility of triggering permanent changes simply does not exist. They are likewise irrelevant in processes characterized by one or few time scales, e. g. thermal equilibration between two metastable basins: A record has a rank, but lacks an inherent magnitude which could match the scale of the process. For record fluctuations to trigger irreversible changes, many inequivalent meta-stable basins must be accessible to the dynamics. We envisage that each metastable basin is associated to a *finite* characteristic escape time τ, the typical time it takes to leave the basin. For thermal dynamics, the relevant time scale is the Arrhenius time corresponding to a free energy barrier $b = Tk_B \ln(\tau)$, where T is the temperature and k_B is the Boltzmann constant.

Assuming that attractor basins have a broad range of stability, with the more fickle basins out-numbering the more stable ones, a quench from a random initial condition typically leads to very fickle basins. At any stage of the subsequent evolution, all moves leading from the less to the more fickle can be reversed on the time scale at which they first happen. Irreversible moves into more stable attractors, the quakes, are important for the aging dynamics, and typically entail a tiny, or marginal, increase of stability. The marginal increase implies that a random fluctuation slightly larger than the previous largest fluctuation, i. e. a record fluctuation, can trigger the next quake. De-facto irreversibility of the quakes can be strengthened by a thermodynamic mechanism, i. e. the basin change may entail a large decrease in energy, and/or a large increase of entropy.



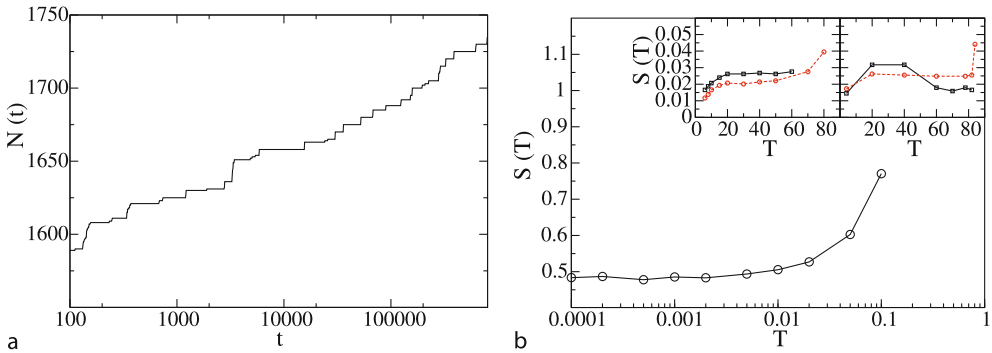
Record Statistics and Dynamics, Figure 1
The cartoon, taken from [43], illustrates the link between marginal stability and record dynamics: the wedges represent metastable region of a complex system, with any reference to their internal hierarchical structure omitted for simplicity. The vertical axis is the energy. Jumps from one metastable region to the next (quakes) are indicated by unidirectional arrows. The thermal stability of each metastable region is represented by the vertical distance from the bottom to the boundary of the wedge, i. e. the energy barrier for the corresponding region. Importantly, the barrier of the current attractor only changes minutely through each quake

The record dynamics scenario sketched above implies that, as the systems age, the metastable basins successively explored develop a nested hierarchical substructure, each layer composed of basins of greater and greater stability. If all metastable regions maintain a similar structure, the noise distribution from which the records are drawn does not itself change during the aging process.

The temporal statistics of record fluctuations and quakes has universal properties which are disconnected from the physical changes associated to the quakes. These effects may e. g. depend on the physical quantity being measured and on the noise strength. Using the *subordination* principle [32,39,47] briefly discussed in the next Section, the changes can be described mathematically in a general, albeit approximate, fashion.

Record Statistics and Intermittency in Glassy Dynamics

The left panel of Fig. 2 (taken from [32]) illustrates important qualitative features of intermittency in complex systems. The figure depicts the time evolution of the number



Record Statistics and Dynamics, Figure 2

The figure is taken from [32]. Left panel: The time variation of the total number of vortices $N(t)$ in the system for a single realization of the pinning potential and the thermal noise in a 8 × 8 × 8 lattice for $T = 0.1$. The monotonous step function character of the time series, and the fact that the steps are approximately the same duration on a *logarithmic* time scale indicates the decelerating nature of the dynamics. Right panel: Numerical results for the creep rate versus T for the time interval between $t = 1000$ and $t = 10000$. In the low temperature region the creep rate is constant within our numerical precision for about two orders of magnitude – we observe a nonzero creep rate in the $T \to 0$ limit. *Insets:* experimental results for the creep rate versus T. The *right inset* shows data from Keller et al. [26,27] for melt processed YBCO crystals with the magnetic field applied along the *c*-axis (squares) and *ab* plane (circles). The *left inset* shows data from Civale et al. [11] for un-irradiated (squares) and 3 MeV proton-irradiated (circles) YBCO flux grown crystals with a 1 *T* magnetic Field applied parallel to the *c*-axis

of vortices in the Random Occupancy Model for magnetic flux creep in type II superconductors [31].

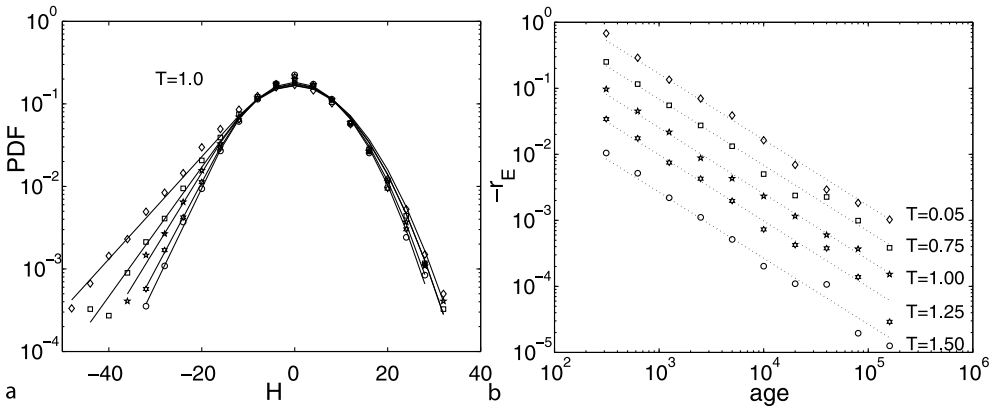
The number $n(t)$ of vortices inside the system changes in a step-wise fashion. Each horizontal plateau corresponds to the time spent in a metastable configuration. The plateaus appear to have the same average duration on a *logarithmic* time scale, showing that the dynamics entrenches itself into gradually more stable basins. Each vertical jump corresponds to a quake, leading from one basin to the next. The size of the quake is the height of the corresponding jump.

In a record-dynamics scenario the rate of quakes is independent of the statistical properties of the noise. Specifically for thermally activated dynamics, it is independent of the temperature. The rate of change of a physical observable may have a temperature dependence, which enters through the quake size distribution, i. e. the statistical distribution of the vertical jumps. The right panel of Fig. 2, also taken from [32], shows the striking temperature independence of the creep rate in the ROM model, and the insert shows its experimental counterpart.

Quenching a glassy system from a high temperature invariably leads to a configuration with an energy far above the equilibrium value at the low temperature. It is then of interest to understand how the excess energy leaves the system during the aging process. The release in a spin-glass [44] occurs through sporadic, intermittent quakes, and the average rate of energy flow (nearly) decreases with the inverse of the system age. The p-spin model [54], an Ising spin model with plaquette interactions, has this type of behavior as well. The left panel of Fig. 3 (taken from [39]) shows the Probability Density Function (PDF) of the amount of heat given off by the system over short intervals of length δt, for several values of the system age t_w. The Gaussian part of the PDF, which is nearly independent of t_w, describe reversible energy fluctuations of zero average. The exponential tail describes the out-flow of heat. As t_w increases, all other parameters being constant, the tail becomes less prominent. The right panel shows, for a number of low temperatures, that the rate of energy flow decays as the reciprocal of the age.

Experimental probes of glassy dynamics often involve applying a small perturbing field at a certain age t_w and measuring the linear response, $R(t, t_w)$, for $t > t_w$. The glassy response depends on both t_w and t, while in a stationary situation the only dependence is on the difference $t - t_w$. There is a large and growing body of experimental and numerical results concerning linear response functions in aging systems and their many facets[23,36,47,55,58].

Attempts to describe the off-equilibrium response as analogous to equilibrium response lead to the idea of effective temperature [13]. Being defined in the limits $t_w \to \infty$ and $t \to \infty$, with t/t_w finite, the latter quantity is, in practice, difficult to measure [21]. At a more basic level, one may ask whether a glassy system is mainly responding to the applied field, or rather mainly responding to the initial quench. In the latter case, the field has the auxiliary rôle of biasing the course of the quakes, but no influence on their temporal statistics. Technically, the re-



Record Statistics and Dynamics, Figure 3

The figure is taken from [39]. Left panel: The PDF of the heat exchanged between system and thermal bath over a time $\delta t = 5$. Negative values correspond to an energy outflow. The data are based on 200 independent runs, taken at temperature $T = 1$ in the intervals $[t_w, t_w + 500]$ for $t_w = 1000$ (diamonds), $t_w = 2000$ (squares) $t_w = 4000$ (polygons) $t_w = 8000$ (hexagons) and $t_w = 16000$ (circles). Right panel: The average rate of flow of the energy is plotted versus the age for the five temperatures shown. The full line has the form $y = ct_w^{-1}$, with the proportionality constant c estimated as the mean of $t_w r_E$. Since data sets for different temperatures are almost overlapping, they are vertically shifted in the plot for the sake of typographical clarity

sponse is then *subordinated* to the quakes, which are themselves subordinated to record-sized fluctuations. Following this approach [39,47], physical observables are treated as (stochastic) functions of the number of quakes k occurring during the observation interval. Generic eigenvalue expansions in the variable k are available, and the time and age dependence are extracted by averaging them over all possible values of k according to the Poisson distribution with the average given in Eq. 9. The subordination approach is not limited to linear response functions, e. g. it can be applied to a calculation of the configurational autocorrelation function [40] or indeed any other quantity which can be argued to be (mainly) a function of k [32].

The potential relevance of records for the evolution of ecosystems has been investigated using the Tangled Nature model of evolutionary ecology [9,18]. The model deals with a population of individuals undergoing mutation-prone reproduction in a type space. All individuals are equally likely to be removed or killed. Their reproduction, however, depends on the interactions among subsets of individuals: The reproduction probability of an individual of type, say, A will change through interactions with a type B individual which comes into being through a mutation of another existing type. Depending on whether the interaction is cooperative or antagonistic, the reproduction rate of A will increase, respectively decrease.

Metastable ecosystems spontaneously arise as sets of types for which the reproduction and death probability of the individuals are balanced. Such metastable configurations last for periods of varying duration, and are eventually replaced by new metastable configurations. The model does not assume a fitness function, and no optimization procedure is explicitly invoked at the level of the microscopic dynamics. Interestingly, the times spent in consecutive metastable configurations have a statistics in qualitative agreement with the predictions of record dynamics [1]. This illustrates how record dynamics can be identified by a relatively straightforward analysis of time series, and how it can emerge at the level of collective evolution even when the precise nature of the observable undergoing records is not known.

Future Directions

Metastable systems are ubiquitous in nature, and their interaction with an external noisy environment affects their dynamics in significant ways. When the system evolves irreversibly through metastable configurations of marginally increasing stability, fluctuation records determine the course of the dynamics. The temporal statistics of records is an in-road to an approximate analytical description of the time evolution of complex systems, ranging from physical material to biological systems. It is an open challenge to test and develop these idea on the large variety of model and observational data where irreversibility and marginal stability appear in different guises.

Acknowledgments

The authors are indebted to P. Anderson, J. Dall, L. Oliveira and S. Boettcher. P. Sibani did part of this work while visiting the Cherry L. Emerson Center for Scientific Computation at Emory University.

Bibliography

Primary Literature

1. Anderson P, Jensen HJ, Oliveira LP, Sibani P (2004) Evolution in complex systems. Complexity 10:49–56
2. Angelani L, Di Leonardo R, Parisi G, Ruocco G (2003) Topological description of the aging dynamics in simple glasses. Phys Rev Lett 87:055502
3. Bak P (1997) How Nature Works. The science of self-organized criticality. Oxford University Press, Oxford
4. Bak P, Tang C, Wiesenfeld K (1987) Self-organized criticality: An explanation of 1/f noise. Phys Rev Lett 59:381–384
5. Boettcher S, Percus A (2001) Optimization with extremal dynamics. Phys Rev Lett 86:5211–5214
6. Bouchaud JP (1992) Weak ergodicity breaking and aging in disordered systems. J Phys I France 2:1705–1713
7. Bray AJ, Moore MA (1987) Chaotic nature of the spin-glass phase. Phys Rev Lett 58:57–60
8. Buisson L, Bellon L, Ciliberto S (2003) Intermittency in aging. J Phys Condens Matter 15:S1163
9. Christensen K, di Collobiano SA, Hall M, Jensen HJ (2002) Tangled Nature: A Model of Evolutionary Ecology. J Theor Biol 216:73–84
10. Cipelletti L, Manley S, Ball RC, Weitz DA. Universal aging features in the restructuring of fractal colloidal gels. Phys Rev Lett 84:2275–2278
11. Civale L, Marwick AD, McElfresh MW, Worthington TK, Malozemoff AP, Holtzberg FH, Thompson JR, Kirk MA (1990) Defect independence of the irreversibility line in proton-irradiated y-ba-cu-o crystals. Phys Rev Lett 65(9):1164–1167
12. Coppersmith SN, Littlewood PB (1987) Pulse-duration memory effect and deformable charge-density waves. Phys Rev B 36:311–317
13. Cugliandolo LF, Kurchan J, Peliti L (1997) Energy flow, partial equilibration, and effective temperature in systems with slow dynamics. Phys Rev E 55:3898–3914
14. Davidsen J, Grassberger P, Paczuski M (2007) Networks of recurrent events, a theory of records, and an application to finding causal signatures in seismicity. cond-mat/0701190
15. Fisher DS, Huse DA (1986) Ordered phase of short-range Ising spin-glasses. Phys Rev Lett 56:1601–1604
16. Glick N (1978) Breaking records and breaking boards. Am Math Mon 85:2–26
17. Goldberg DE (1989) Genetic algorithms in search, optimization and machine learning. Addison-Wesley, Reading
18. Hall M, Christensen K, di Collabiano SA, Jensen HJ (2002) Time dependent extinction rate and species abundance in the Tangled Nature model of biological evolution. Phys Rev E 66:011904
19. Hammann J, Bouchaud J-P, Dupuis V, Vincent E (2001) Separation of time and length scales in spin glasses: Temperature as a microscope. Phys Rev B 65:024439
20. Hoffmann KH, Sibani P (1988) Diffusion in hierarchies. Phys Rev A 38:4261–4270
21. Hérisson D, Ocio M (2002) Fluctuation-dissipation ratio of a spin glass in the aging regime. Phys Rev Lett 88:257202
22. Jensen HJ (1998) Self-Organized Complex Systems. Cambridge University Press, Cambridge
23. Jonason K, Vincent E, Hammann J, Bouchaud JP, Nordblad P (1998) Memory and Chaos Effects in Spin Glasses. Phys Rev Lett 81:3243–3246
24. Jönsson PE, Takayama H, Katori AH, Ito A (2005) Dynamical breakdown of the ising spin-glass order under a magnetic field. Phys Rev B 71:180412(R)
25. Katzgraber HG, Young AP (2005) Probing the Almeida-Thouless line away from the mean-field model. Phys Rev B 72:184416
26. Keller C, Kupfer H, Meier-Hirmer R, Wiech U, Selvamanickam V, Salama K (1990) Irreversible behaviour of oriented grained *yba2cu3ox*. part 2: relaxation phenomena. Cryogenics 30:410–416
27. Keller C, Kupfer H, Meier-Hirmer R, Wiech U, Selvamanickam V, Salama K (1990) Irreversible behaviour of oriented grained YBa2Cu3Ox. part 1: transport and shielding currents. Cryogenics 30:401–409
28. Krug J (2007) Records in a changing world. cond-mat/0702136
29. Lundgren L, Nordblad P, Sandlund L (1986) Memory behaviour of the spin glass relaxation. Europhys Lett 1:529–534
30. Mossa S, Ruocco G, Sciortino F, Tartaglia PP (2002) Quenches and crunches: does the system explore in ageing the same part of the configuration space explored in equilibrium? Phil Mag B 82:695–705
31. Nicodemi M, Jensen HJ (2001) Aging and memory phenomena in magnetic and transport properties of vortex matter: a brief review. J Phys A 34:8425
32. Oliveira LP, Jensen HJ, Nicodemi M, Sibani P (2005) Record dynamics and the observed temperature plateau in the magnetic creep rate of type ii superconductors. Phys Rev B 71: 104526
33. Parisi G (1979) Infinite number of order parameters for spin-glasses. Phys Rev Lett 43:1754–1756
34. Parisi G, Marinari E, Ruiz-Lorenzo JJ (1998) On the phase structure of the 3D Edwards-Anderson spin-glass. Phys Rev B 58:14852
35. Redner S, Petersen MR (2007) Role of global warming on the statistics of record-breaking temperatures. Phys Rev E 74:061114
36. Rodriguez GF, Kenning GG, Orbach R (2003) Full Aging in Spin Glasses. Phys Rev Lett 91:037203
37. Rudin W (1966) Real and complex analysis. McGraw Hill, New York
38. Schön JC, Putz H, Jansen M (1996) Studying the energy hypersurface of continuous systems–the threshold algorithm. J Phys: Condens Matter 8:143–156
39. Sibani P (2006) Aging and intermittency in a p-spin model. Phys Rev E 74:031115
40. Sibani P (2006) Mesoscopic fluctuations and intermittency in aging dynamics. Europhys Lett 73:69–75
41. Sibani P, Andersen CM (2001) Aging and self-organized criticality in driven dissipative systems. Phys Rev E 64:021103
42. Sibani P, Brandt M, Alstrøm P (1998) Evolution and extinction dynamics in rugged fitness landscapes. Int J Modern Phys B 12:361–391
43. Sibani P, Dall J (2003) Log-Poisson statistics and pure aging in glassy systems. Europhys Lett 64:8–14
44. Sibani P, Jensen HJ (2005) Intermittency, aging and extremal fluctuations. Europhys Lett 69:563–569
45. Sibani P, Littlewood PB (1993) Slow Dynamics from Noise Adaptation. Phys Rev Lett 71:1482–1485
46. Sibani P, Pedersen A (1999) Evolution dynamics in terraced NK landscapes. Europhys Lett 48:346–352

47. Sibani P, Rodriguez GF, Kenning GG (2006) Intermittent quakes and record dynamics in the thermoremanent magnetization of a spin-glass. Phys Rev B 74:224407
48. Sibani P, Schmidt M, Alstrøm P (1995) Fitness optimization and decay of the extinction rate through biological evolution. Phys Rev Lett 75:2055–2058
49. Sibani P, Schriver P (1994) Phase-structure and low-temperature dynamics of short range Ising spin glasses. Phys Rev B 49:6667–6671
50. Sibani P, Schön C, Salamon P, Andersson J-O (1993) Emergent hierarchical structures in complex system dynamics. Europhys Lett 22:479–485
51. Stillinger FH, Weber TA (1983) Dynamics of structural transitions in liquids. Phys Rev A 28:2408–2416
52. Stillinger FH, Weber TA (1984) Packing Structures and Transitions in Liquids and Solids. Science 225:983–989
53. Struik LCE (1978) Physical aging in amorphous polymers and other materials. Elsevier, New York
54. Swift MR, Bokil H, Travasso RDM, Bray AJ (2000) Glassy behavior in a ferromagnetic p-spin model. Phys Rev B 62:11494–11498
55. Takayama H, Hukushima K (2002) Numerical Study on Aging Dynamics in 3D Ising Spin-Glass Model. III. Cumulative Memory and 'Chaos' Effects in the Temperature Shift Protocol. J Phys Soc JPN 71:3003–3010
56. Tang C, Wiesenfeld K, Bak P, Coppersmith S, Littlewood P (1987) Phase Organization. Phys Rev Lett 58:1161–1164
57. Van Kampen NG (1992) Stochastic Processes in Physics and Chemistry. North Holland, Amsterdam
58. Vincent E (1991) Slow dynamics in spin glasses and other complex systems. In: Ryan DH (ed) Recent progress in random magnets. McGill University, Montreal, pp 209–246
59. Vincent E, Hammann J, Ocio M, Bouchaud J-P, Cugliandolo LF (1996) Slow dynamics and aging in spin-glasses. SPEC-SACLAY-96/048

Books and Reviews

Tinkham M (1998) Introduction to Superconductivity. Krieger, Melbourne
Drossel B (2001) Biological evolution and statistical physics. Adv Phys 50:209–295

Regional Climate Models: Linking Global Climate Change to Local Impacts

DANIELA JACOB
Max-Planck-Institute for Meteorology, Hamburg, Germany

Article Outline

Glossary

Climate models They are mathematical representations of the Earth system, in which physical and biogeochemical processes are described numerically. Climate models can be of a global scale or focus on a sub-region (regional climate model).

Downscaling Dynamical and statistical techniques to interpret global climatic changes in specific regions.

IPCC emission scenario Description of possible developments of the socio-economic system expressed in terms of emissions into the atmosphere.

Projection Simulation of possible climatic changes in the future, dependent on emission scenarios, land-use changes and natural variability in the climate system.

Validation Comparison of observed data against model result for quality assessment of the model.

Definition of the Subject

A variety of observations demonstrates that during the last decades the climate has changed. As reported by the *Intergovernmental Panel on Climate Change* (IPCC, 2001, 2007), a mean increase of temperature by 0.09 K per decade was observed globally from 1951 to 1989. Up to now, 2007, this trend has continued. Europe experienced an extraordinary heat wave in summer 2003, with daily mean temperatures being about 10° warmer locally than the long term mean. The increase of temperature varies depending on the region and season.

The temperature change seems to be accompanied by changes in several meteorological and hydrological quantities, like number and duration of heat waves, frost periods, storminess or monthly mean precipitation. In Germany, for example, winter precipitation has increased in parts by more than 30% within the last four decades. In addition, very intense precipitation was observed in summer 2002 in parts of the Elbe drainage basin, which faced a severe flooding.

It can be expected that extreme weather situations will occur more often in a warming world. Therefore, a growing demand from decision-makers and the general public for detailed information on possible future climate development is evident, worldwide. The quantification of risks

associated with changing climates is a prerequisite for the formulation and implementation of realistic adaptation and mitigation strategies [1].

Global climate models (GCM) have been developed to study the Earth's climate system in the past and future. Unfortunately, even today, global climate models provide information only at a relatively coarse spatial scale, which is often not suitable for regional climate change assessments. To fill this gap, two different principles to transfer the information from a global model to the region of interest have been developed accordingly: statistical downscaling and dynamical downscaling. Statistical downscaling techniques connect the climate change signal provided by the GCM with observations from measurement stations in the region to achieve higher resolved climate change signals.

Dynamical downscaling uses high resolution three-dimensional regional climate models (RCM), which are nested into GCMs. RCMs are similar to numerical weather forecasting models, which are taken into account non-linear processes in the climate system. The results of both downscaling methods depend on both the quality of the global and regional models. In the following, the focus will be on dynamical downscaling, in order to be able to also detect more easily new extremes, which have not been observed so far, and to take into account possible feedback mechanisms, which might appear under climate change conditions, and which influence the extent of regional climatic changes. Regional feedback mechanisms are, for example, snow-albedo/temperature feedbacks or soil moisture-temperature feedbacks. If snow melts the reflectivity of the surface changes from bright (white snow) to dark (vegetation or soil). This enhances the absorption of incoming radiation and leads to warming of the surface, which in turn accelerates the snow melt nearby. Evaporation from soils and vegetation increases with temperature and decreases soil moisture. Dryer soils evaporate less, so that cooling due to evaporation is decreasing, which in turn increases temperatures regionally.

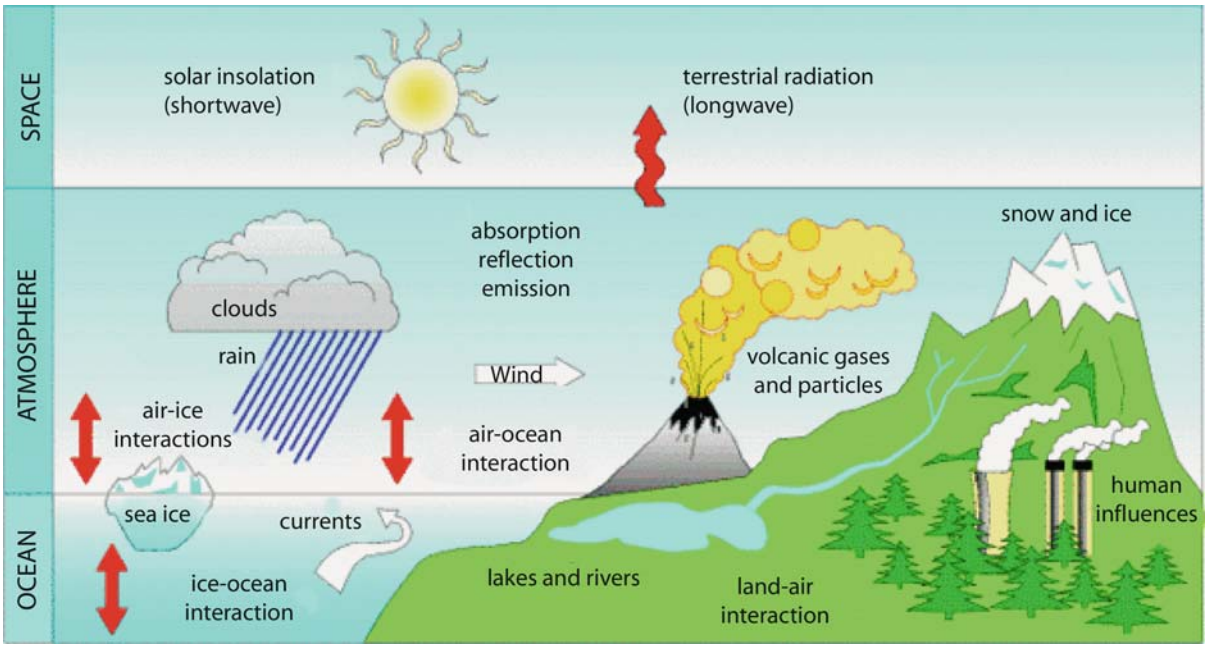
Introduction

The climate on Earth varies from region to region, and is characterized by sequences of weather events. These events are determined by the atmospheric flow, established through a wide range of interacting scales. The interacting scales cover large-scale features of the order of thousands to hundreds of km, mostly determined by the distribution of the continents and oceans, solar radiation and the composition of the atmosphere, regional features of the order of a few hundred km to a few km, forced through complex topography and vegetation distribution, and small scale features, like convection. The description of the Earth's climate needs to consider all scales, respectively. Therefore global climate models (GCMs) have been developed. They are mathematical representations of the Earth system (Fig. 1), in which physical and biogeochemical processes are described numerically to simulate the climate system as realistically as possible. Today GCMs develop into Earth system models (ESM), which are not only coupled atmosphere-ocean general circulation models (AOGCM), but also take into account some biogeochemical feedbacks, like the carbon cycle or dynamical vegetation. They are the most advanced numerical tools for climate modeling and describe changes due to large scale forcing.

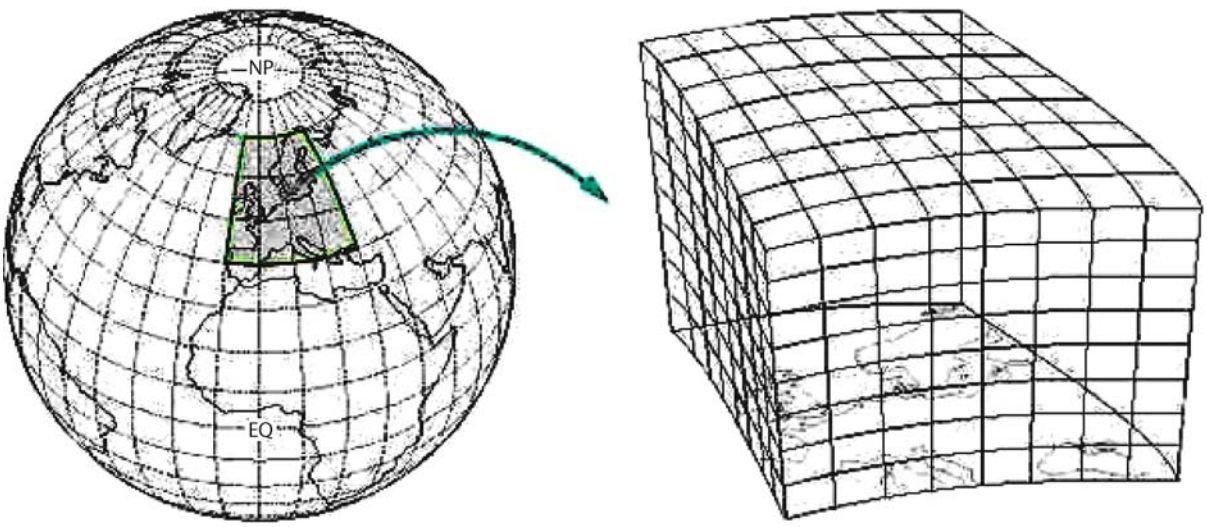
Increasing greenhouse gas (GHG) concentration, changing aerosol composition and load as well as land surface changes are influencing the climate of the Earth, globally and regionally. Therefore the demands for fine scale regional climate information were growing and in response to this the development of regional climate models started about 20 years ago. It is obvious that the simulations of regional climate changes requires the simulations of processes from global to local scales, so that very high resolution AOGCMs with grid sizes of about 10 km could be seen as the best solution. However, until today the horizontal resolution is still relatively coarse (100 to 250 km grid size) due to limitations in computer resources. Hence AOGCMs cannot provide regional details.

To overcome the deficiency two different approaches are used: statistical downscaling and dynamical downscaling. Both translate the information from the global model to the region of interest. Statistical downscaling techniques connect the climate change signal provided by the GCM with observations from measurement stations in the region to achieve higher resolved climate change signals. Dynamical downscaling uses high resolution three-dimensional regional climate models (RCM), which are nested into GCMs (Fig. 2). RCMs are similar to numerical weather prediction models, in which non-linear processes in the climate system are taken into account. The results of both methods depend on both the quality of the global and regional models. In the following, the focus will be on dynamical downscaling, to be able to detect more easily new extremes, which have not been observed so far, and to take into account possible feedback mechanisms, which might appear under climate change conditions, like the snow-albedo/temperature feedback or the soil moisture-temperature feedback.

Regional climate models are limited area models. They only cover the area of interest, which can be resolved to



Regional Climate Models: Linking Global Climate Change to Local Impacts, Figure 1
The physical climate system



Regional Climate Models: Linking Global Climate Change to Local Impacts, Figure 2
Nesting technique

a much higher degree than GCMs. Therefore RCMs describe the effects of regional and small scale processes within the simulation domain and are connected to the global flow using the nesting technique. This technique was developed to get higher resolution climate information on regional scales, and it is very similar to the nesting procedure in NWP. For initialization and at the lateral boundaries the GCM, in which the RCM is nested, provides information about the state of the atmosphere and the surface conditions. Usually atmospheric fields like wind, pressure, temperature and humidity are provided as well as sea surface temperatures [14,17]. Soil temperature and soil moisture are initialized once, but calculated within the RCMs during the simulation.

The development of regional climate models started in the USA. Filippo Giorgi at NCAR was the first one running the MM4 model in a so-called climate mode [10], which means simulations longer than a few days, as it was common for numerical weather prediction (NWP). For many years RCMs were applied to simulations covering one month. The extension of NWP models to month-long simulations required changes in the formulation of physical processes which were taken into account within the model. In NWP models processes acting on time scales longer than weeks are not important and so not included.

The development of longer-term climate simulations happened very fast and simultaneously in several modeling centers of the world. In the early 1990s, the first multi-year simulations were carried out by Giorgi et al. [11,12], whereas Jones et al. [21,22] and McGregor et al. [27] succeeded in ten-year simulations. Nowadays regional climate simulations stretch from several decades, first achieved by Machenhauer et al. [25], up to more than a century in transient climate change mode [20]. Currently RCMs are widely used for regional climate studies for almost all regions of the world, with horizontal grid spacing ranging from more than 100 km to 10 km. A more detailed overview can be found in Giorgi [9] and for example in [7] focusing on regional climate modeling in the Arctic.

The basic features of RCMs will be explained in Sect. "Basic Features and Model Characteristics" and examples of applications will be presented in Sect. "Validation", climate scenarios in Sect. "IPCC-Scenarios", and examples of applications in Sects. "Regional Climate Change" and "Regional Extremes". A discussion of future perspectives and concluding remarks follow in Sect. "Future Directions".

Basic Features and Model Characteristics

Until today, most RCMs are three-dimensional hydrostatic circulation models, solving the discretized primitive equations of the atmospheric motion. Summaries of model characteristics can be found in many publications, e. g. Jacob et al. [18,19]. As an example for the development and characteristics of many RCMs, the standard set-up of REMO, the regional climate model developed and used at the Max-Planck-Institute for Meteorology is described in more detail below.

The development of REMO started in 1994 utilizing the existing NWP model (EM) of the German Weather Service DWD [26]. Additionally, the physical parametrization package of the general circulation model ECHAM4 [33] has been implemented. During the last decade it could be shown in several applications that the combination of the EM dynamical core plus the ECHAM4 physical parametrization scheme is able to realistically reproduce regional climatic features and therefore became the standard setup.

The atmospheric prognostic variables of REMO are the horizontal wind components, surface pressure, temperature and specific humidity, as well as cloud liquid water. The temporal integration is accomplished by a leap-frog scheme with semi-implicit correction and time filtering after Asselin [2]. REMO is a grid box model, with grid box centers defined on a rotated latitude–longitude coordinate system. For horizontal discretization the model uses a spherical Arakawa-C grid in which all variables except the wind components are defined in the center of the respective grid box. In the vertical, a hybrid vertical coordinate system is applied [35]. Details about the physical parameterizations can be found in Jacob 2001 [16], but will not be explained here in more detail, since they vary slightly from RCM to RCM (see for example [18,19]).

The resolution of the horizontal grids in RCMs varies from about 100 km to 10 km and has increased constantly. For many years 1/2° grid size could be seen as a standard horizontal resolution, which was used in many experiments, even for model inter-comparison studies (e. g. [4,18,25,32]).

REMO uses horizontal grids with 1/12°, 1/6° or 1/2°, corresponding to horizontal resolutions of about 10 km, 18 km and 55 km. In the vertical 20 to 40 levels are applied.

Applying the nesting technique for regional climate models requires large scale atmospheric flow fields to *drive* the RCMs at their lateral boundaries. These fields can be derived from different sources depending on the application. Regional climate simulations require climate change information from AOGCMs, whereas the simulations of the last decades are driven by global analyzes of observations. The analyzes consist of observations, which have been interpolated in space and time using global models. They can be interpreted as the best available representation of the observed atmospheric flow conditions; however, systematic biases cannot be excluded due to the utilization of numerical models for interpolation. In regional climate modeling the use of driving data from analyzes or re-analyzes products are referred to as simulations with perfect boundary conditions (PBC). These experiments have the clear advantage to be directly comparable to observations for the actual time periods and they build the basis for model validation experiments.

In all cases, the relaxation scheme according to Davies [5] is applied in REMO, meaning that the prognostic variables are adjusted towards the large-scale forcing in a lateral sponge zone of 8 grid boxes. Within this zone the

influence of the lateral boundary conditions decreases exponentially towards the inner model domain.

At the lower boundary, RCMs are determined through the interaction with the land surface and, over sea, by the sea surface temperature (SST) and sea ice distribution. The SST can either be interpolated from the large-scale forcing or from observational datasets, or it can be calculated online by a regional ocean model coupled to the RCM, e. g. [23]. The same is true for the sea ice extent, which can as a further option also be diagnosed from the SST. The land surface with its ongoing changes plays a major role in the climate system. Therefore, in all RCMs the exchange between surface and atmosphere is realized by the implementation of a land surface scheme [30]. Generally, one surface grid box can either be covered by water, sea ice or land or can include fractions of land and water areas, all characterized by their own roughness length and albedo. The land fraction of the surface can be covered by bare soil or by vegetation of different type. Depending on the complexity of the land surface scheme the exchange between the atmosphere and the underlying surface is realized through turbulent surface fluxes and the surface radiation flux, which are calculated separately for each fraction and weighted averages of the fluxes are used within the lowest atmospheric model level. Physical properties of the soil and vegetation control the exchange of heat, moisture and momentum over land. In REMO these properties include for example the surface roughness length, the soil field capacity, the water holding capacity of the vegetation, the background albedo, the fractional vegetation cover and the leaf area index (LAI). Some of these parameters strongly depend on the physiological state of the vegetation and are variable between the growing and the dormancy season [31].

There are two options to use the nesting technique. Within the one-way mode a GCM drives a RCM at the lateral boundaries, but no information is given back to the GCM. This method is the standard one used until today in regional climate modeling. It is relatively easy to implement and allows the use of RCMs without running a GCM. The RCM adds information on scales smaller than the driving GCM (e. g. topographical forcing), but is strongly dependent on the superimposed large scale flow. Hence RCMs cannot correct large scale flows originating from GCMs, which might have large errors. However, Giorgi et al. [13], showed that some modulation of the large scale flow is possible within the RCM simulation, stimulated by regional scale forcing.

In the two-way mode, both models GCM and RCM run simultaneously and the RCM feeds back information to the GCM every GCM time step. This method has recently been established for regional climate modeling studies [24]; it has the clear advantage that the atmospheric flow generated within the RCM domain can modulate the large scale flow in areas with strong energetic input from the surface to the atmosphere (e. g. the maritime continent).

If RCM experiments are carried out with high horizontal resolution, like 10 or 20 km, it can be required to use the so-called double nesting technique to avoid mismatch in scales along the lateral boundaries due to the coarse resolution of the driving GCMs. Double nesting means that first a RCM simulation will be carried out with a relatively coarse resolution to generate lateral boundary conditions for further nesting. For REMO, sometimes a sequence of nests is calculated [20].

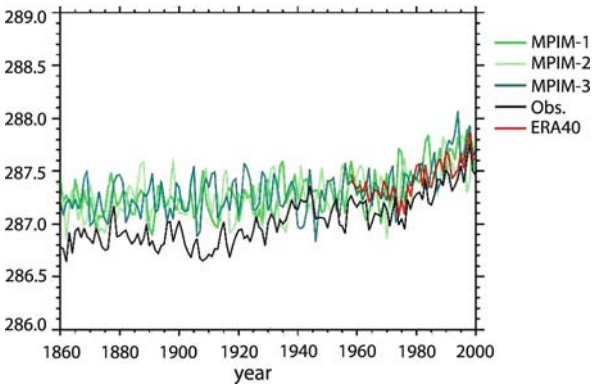
Finally there are two independent options to run a regional climate model with PBC: the forecast mode and the climate mode. In climate mode the RCM simulation is initialized once from analyzes, and then it is continuously calculated forward in time, driven by regularly up-dated lateral boundaries. In forecast mode, a sequence of short runs (e. g. 30 hours), each initialized every 30 hours from analyzes, is carried out. The forecast mode has the advantage to force the RCM flow to be very close to the observed one, but it has the disadvantage to suppress mesoscale flow features. These mesoscale processes can be excited within the RCM domain by land–sea contrasts or topography and are too small to be taken into account in the GCM.

Validation

The quality of the RCM simulations depends strongly on the performance of the driving model due to the one-way nesting procedure. Therefore, it is extremely important to validate the driving large scale fields before applying RCMs. The model quality, however, can only be judged in comparison with independent observations. Therefore, time periods of the past are simulated and the model results are compared against measurements before the models are used for climate change studies. These comparisons are also part of model development and testing.

As an example, Fig. 3 shows time series of observed and simulated global mean near surface temperatures for the period 1860 to 2000. The simulated results from the global coupled climate model ECHAM5/MPI-OM (Max-Planck-Institute for Meteorology) are in good agreement with ERA40 data, but about 0.5° warmer than the reconstructed observations. The observed increase during the last decades is clearly visible.

As for GCMs, the model quality of RCMs needs to be analyzed. Therefore RCMs are nested into re-analyzed



Regional Climate Models: Linking Global Climate Change to Local Impacts, Figure 3
Time series of the global mean near surface air temperature (K): observed (*black*), from reanalyzes (*red*), and from the MPIM global climate model simulations (*green*)

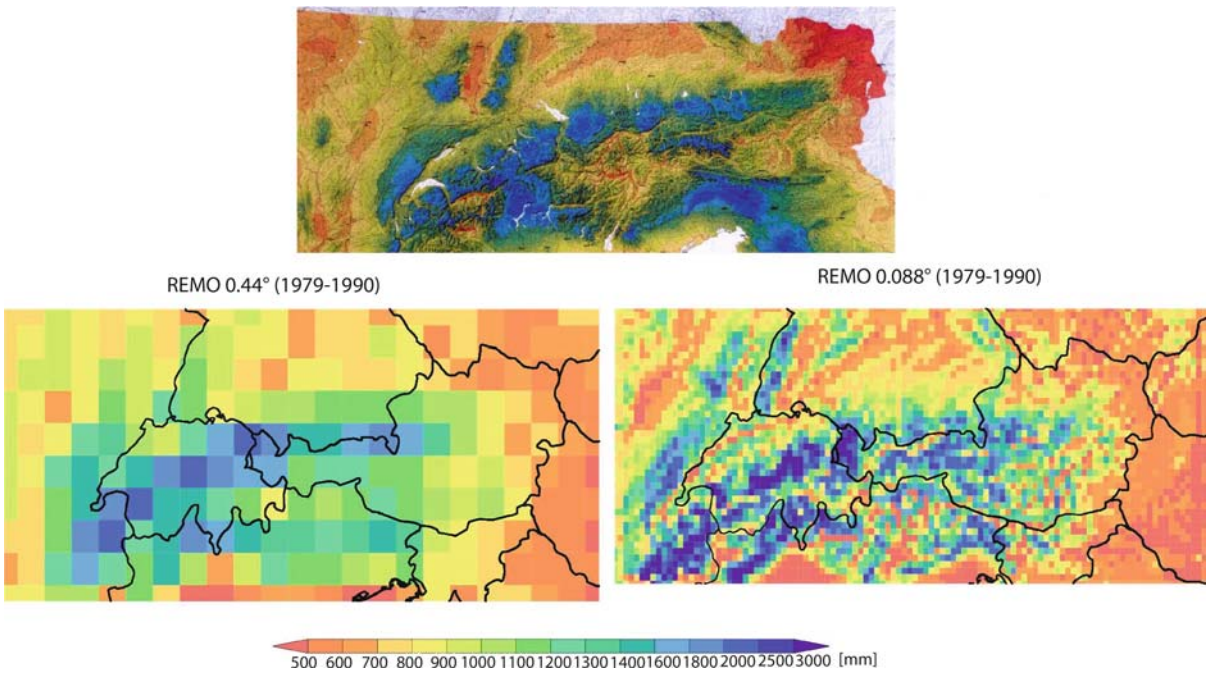
data, which can be seen as close to reality as possible (see above). The results of the RCM simulations of the last decades are compared against independent observations, means as well as extremes are considered. As an example, simulated precipitation climatologies calculated with REMO with two different horizontal grid sizes are compared against observations [8].

The total amount of precipitation and the horizontal pattern are much better resolved using the very high horizontal resolution of about 10 km (Fig. 4). Regional maxima, like the one in the Black Forest, and minima, like in the central valleys of the Alps, are detectable. However, the resolution is still too coarse for climate change studies in individual alpine valleys. The 50 km grid is much finer than standard GCM grids (about 150 to 250 km), but is still insufficient for studying regional details if the regions are too small.

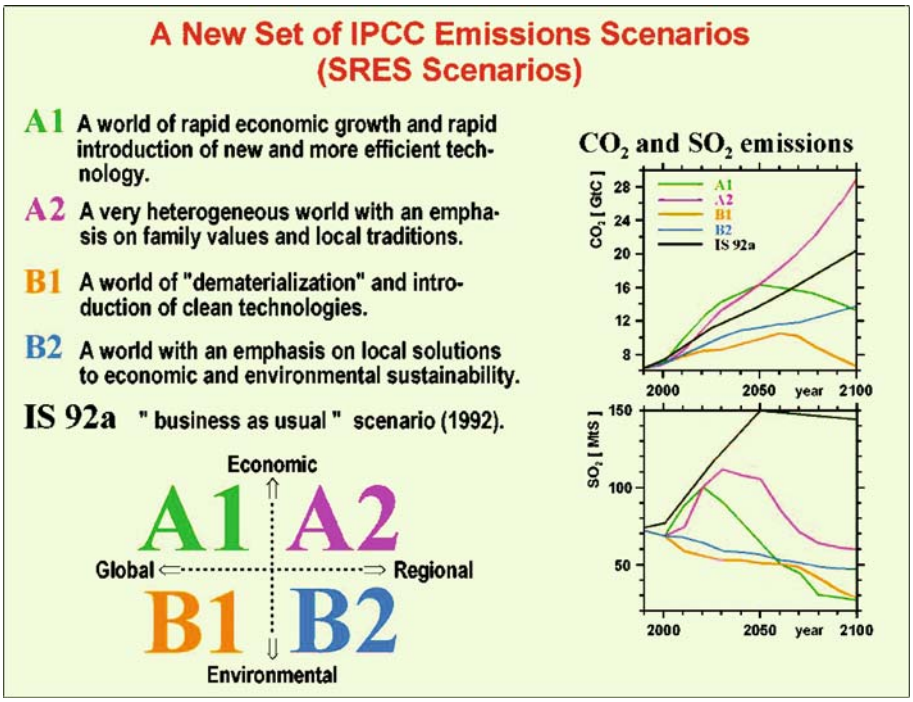
IPCC-Scenarios

The investigation of possible future climate changes requires information about possible changes in the *drivers* of climate change. So-called *drivers* are for example, amount and distribution of aerosols and green house gases (GHG) in the atmosphere, which depend directly on natural and man-made emissions. The IPCC emissions scenarios (Fig. 5) follow so-called story lines, describing possible developments of the socioeconomic system [29].

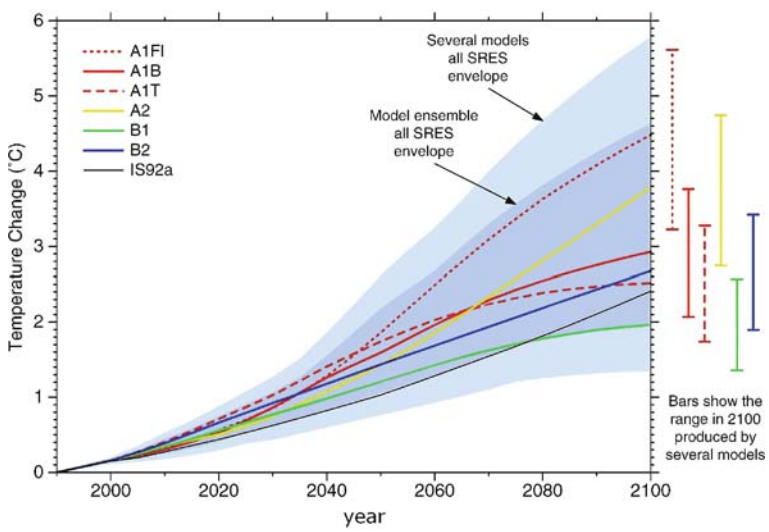
The emissions are directly used within GCMs and RCMs and they initiate changes in global and regional climates through numerous non-linear feedback mechanisms. As an example, Fig. 6 shows possible developments



Regional Climate Models: Linking Global Climate Change to Local Impacts, Figure 4
Annual total precipitation (mm), observed (1971–1990, *upper panel*) and simulated with about 50 km grid size (*left*) and 10 km grid size (*right*)



Regional Climate Models: Linking Global Climate Change to Local Impacts, Figure 5
SRES Scenarios, which shows the four major storylines together with the associated developments of CO_2 and SO_2 emissions from 2000 until 2100



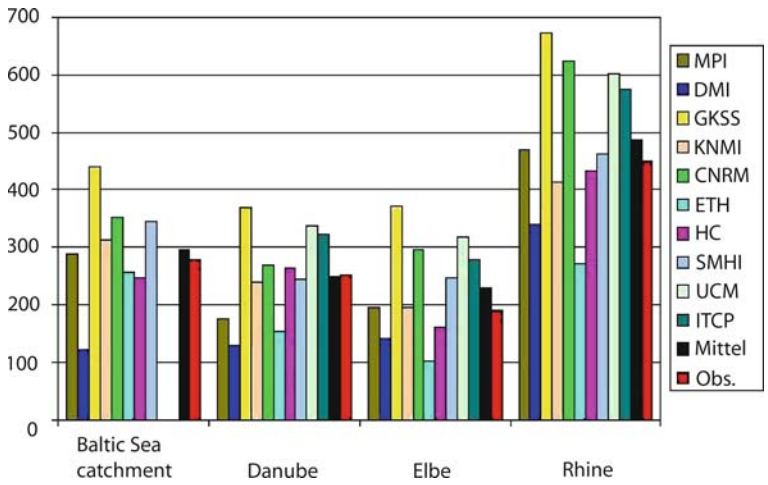
Regional Climate Models: Linking Global Climate Change to Local Impacts, Figure 6
Changes in global mean near surface air temperature as calculated by several GCMs under seven emissions scenarios until 2100

of global mean near surface temperatures calculated by several models for different scenarios.

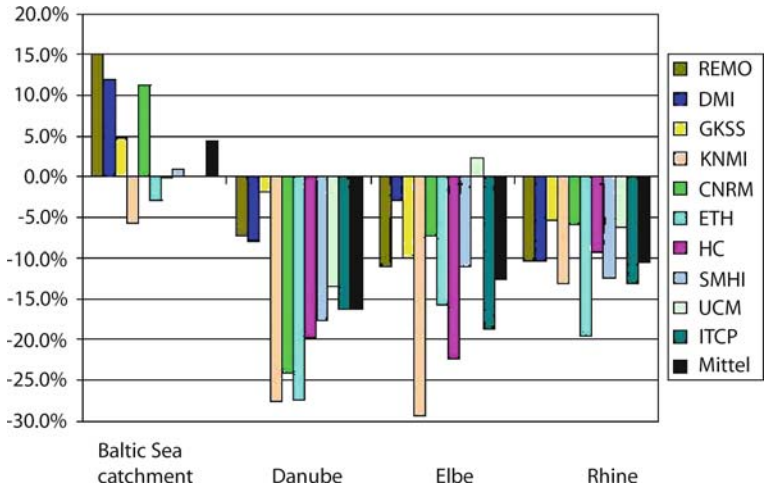
The global mean changes in near surface temperature until 2050 is about 1.5°C, whereas until the end of the century a wide spread appears from 1.5°C to 5.5°C.

Regional Climate Change

In order to achieve information about the probability, e. g. for the intensification of the hydrological cycle over Europe, several models from different European climate re-



Regional Climate Models: Linking Global Climate Change to Local Impacts, Figure 7
Simulated and observed river run-off (precipitation P – evaporation E) for 1961 to 1990 [15] in the Baltic Sea, Danube, Elbe and Rhine catchments



Regional Climate Models: Linking Global Climate Change to Local Impacts, Figure 8
Simulated and observed change in river run-off (precipitation P – evaporation E) for the period 2071 to 2100 compared to 1961 to 1990 [15]. Baltic Sea, Danube, Elbe and Rhine catchments

search institutes are used, as it was done in the EU project PRUDENCE [4].

Following the climate change scenario A2 projecting a strong future increase of greenhouse gases until the year 2100 (IPCC, 2001) and a subsequent global mean temperature increase of about 3.5°, numerous simulations were conducted within PRUDENCE [19]. An analysis of their results for different river catchments [15] shows significant differences between the projected changes over northern and central Europe for the time period 2070–2100 compared to the current climate (1961–1990, Fig. 7).

For the Baltic Sea catchment, a precipitation increase of about +10% for the annual mean is projected, with the largest increase of up to +40% in winter, while a slight reduction of precipitation is calculated for the late summer. Evapotranspiration will increase during the entire year with a maximum increase in winter. These rises in precipitation and evapotranspiration would lead to an increase of river discharge into the Baltic Sea of more than 20% in winter and early spring. Here, the seasonal distribution of discharge is largely influenced by the onset of spring snowmelt.

For the catchments of Rhine, Elbe and Danube, a different change in the water balance components is projected. While the annual mean precipitation will remain almost unchanged, it will increase in late winter (January–

March) and decrease significantly in summer. The evapotranspiration will rise during the entire year, except for the summer, with a maximum increase in winter. These changes lead to a large reduction of 10 to 20% in the annual mean discharge (Fig. 8). Especially for the Danube, the projected summer drying has a strong impact on the discharge that is reduced up to 20% throughout the year except for the late winter (February/March) when the increased winter precipitation causes a discharge increase of about 10%. These projected changes in the mean discharge will have significant impacts on water availability and usability in the affected regions.

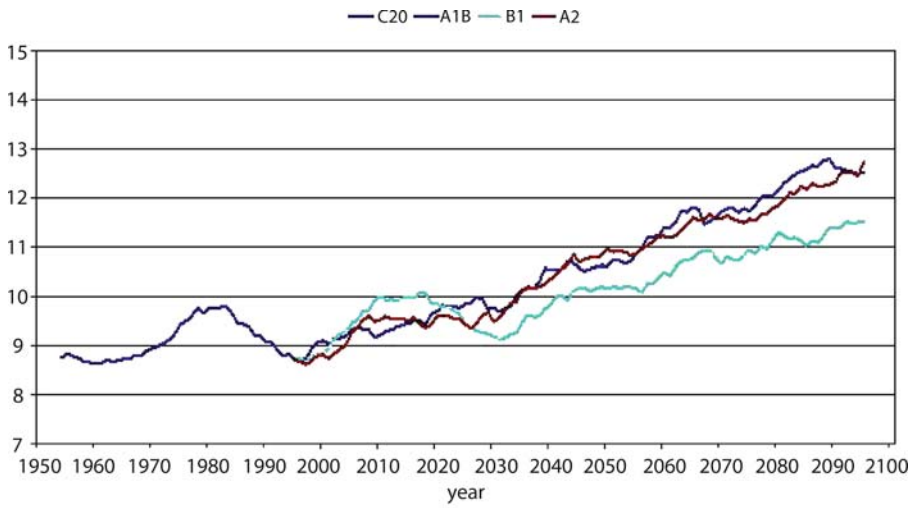
Under climate change conditions not only the absolute amounts of precipitation may change but also the precipitation intensities, i. e. the amount of precipitation within a certain time period. The simulation of precipitation intensities or extreme precipitation events requires however a considerably higher resolution than the A2 results presented above so that for example the influence of the topography of the Alps on the formation of precipitation over the Rhine catchment could be adequately calculated. High resolution RCM results show that the global warming until 2050 will lead to an increase of high precipitation events over the Alpine part of the Rhine catchment, especially in summer. This climate change signal becomes clearly visible in the Pre-Alps, but a similar trend is also seen in the high resolution simulations over large parts of Europe.

A major breakthrough was possible with the regional climate change simulations on 10 km grid scale. Within a co-operation with the national environmental agency (UBA), REMO was used for a control simulation from 1950 to 2000 and three transient run for the IPCC SRES scenarios A2, A1B and B1. The simulation domain covers Germany, Austria and Switzerland [20]. As an example the most important results for Germany at the end of this century are summarized as follows:

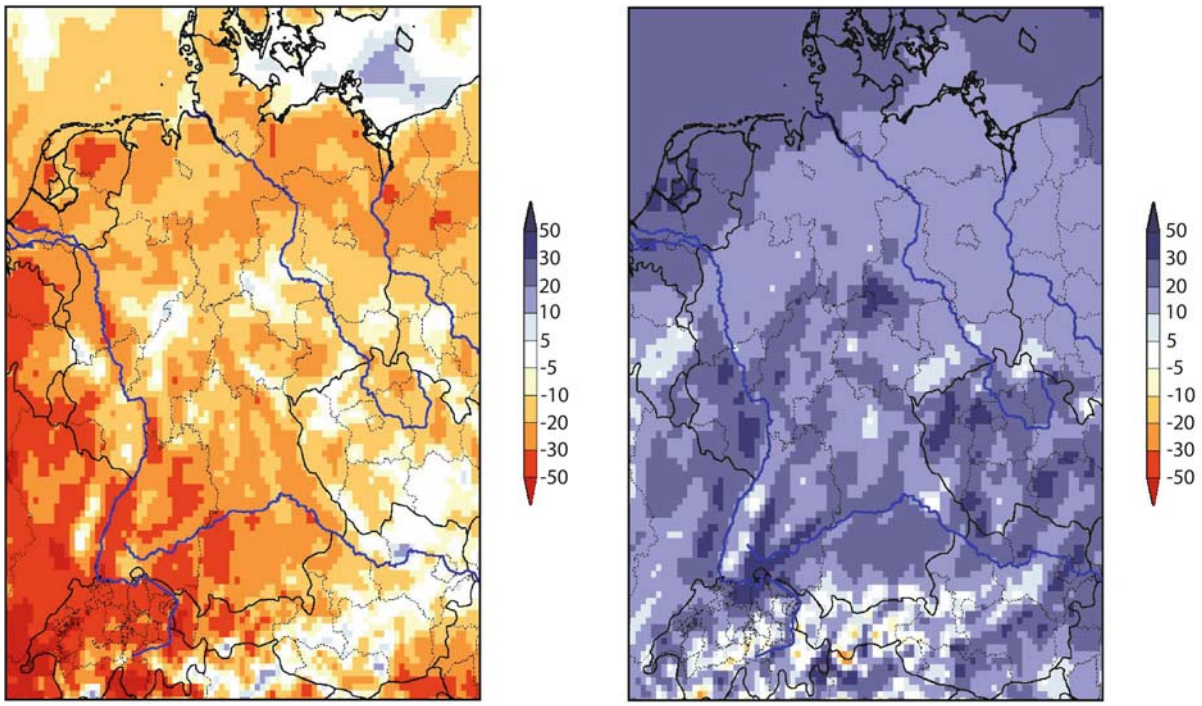
The simulated annual mean near surface temperature is increasing up to 3.5°C depending on the emission scenario (Fig. 9). The regional pattern of temperature changes shows that the south and southeast warm more than all other areas in the simulation domain. The warming is associated with a decrease of precipitation amount in wide areas of Germany during summer and an increase of precipitation in south and southwest regions during the winter (Fig. 10). The winter precipitation is mostly rain and less precipitation falls as snow.

Regional Extremes

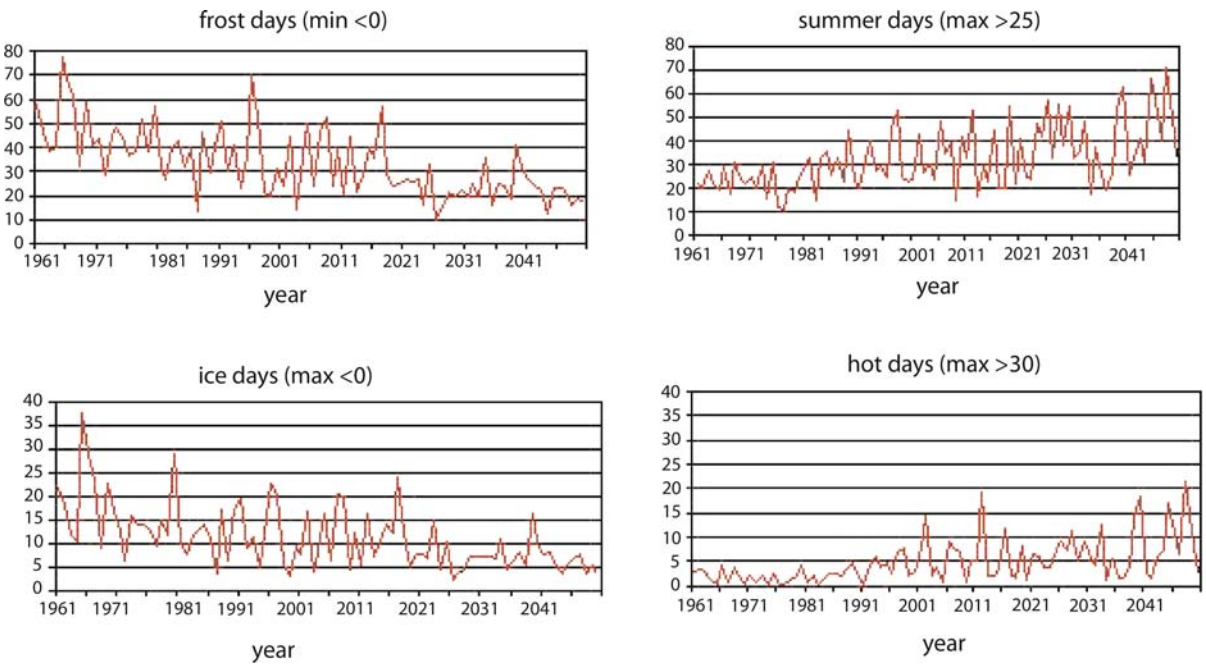
The calculated rapid and strong changes of climate parameters can have severe impacts on humans and the environment. As an example, REMO results for the Rhine basin are presented for a B2 scenario until 2050. Between 1960 and 2050 the near surface temperature will rise by about 3°C and the number of summer days and hot days will increase (Fig. 11). In addition, the number of periods with summer days, this is the period of consecutive days with a daily maximum temperature above 25°C (not shown), will be higher in the future decades. Winter temperature also increases, leading to a decrease in frost and ice days.



Regional Climate Models: Linking Global Climate Change to Local Impacts, Figure 9
Changes in annual mean near surface temperature (°C) from 1950 to 2100 for three different IPCC SRES scenarios



Regional Climate Models: Linking Global Climate Change to Local Impacts, Figure 10
Climate change signals for summer (*left*) and winter (*right*) precipitation (%) in scenario A1B for 2071 to 2100 compared to 1961 to 1990



Regional Climate Models: Linking Global Climate Change to Local Impacts, Figure 11
REMO B2 scenario for the Rhine catchment: frost days (*upper left*), ice days (*lower left*), summer days (*upper right*) and hot days (*lower right*)

The investigation of probability distribution functions for temperature and precipitation using the 10 km horizontal resolutions simulations for Germany shows possible monthly mean temperature of more than 30°C for July and >10°C for January appearing in the Rhine valley under the assumption of A1B scenario until the end of this century. In addition, possible increases in monthly mean precipitations are projected for A1B until 2100 in the area of Leipzig (Elbe drainage basin) for January as well as for July (Bülow, Ph D thesis, in preparation).

The projected changes in extremes are coherent with findings of Schär et al. [34] who studied the occurrence of summer heat waves in Europe today and in the future. He showed that the summer 2003 heat wave was extraordinary but can appear much more often in the future.

Another robust finding is the increase of heavy precipitation events in summer, which goes together with a decrease in monthly mean summer precipitation in central Europe [3]. Such short-term, strong convective summer precipitation events have the potential of causing damages, e. g. for agriculture, but also in cities, when sewage systems might be flooded.

Future Directions

Regional climate modeling has made major progress during the last decades and it could be shown that the added value lies mainly in the provision of details in space and time, which is demanded for impact assessments as well as by the public. RCMs are ready for operational use as powerful tools to simulate regional climatic features and their changes in all regions of the world, for time periods from today until the end of this century.

Now, further research on regional climate changes should focus on the reliability in projected regional and local climate change pattern. This can be done by ensemble calculations of GCMs-RCMs, which are also needed for the establishment of probability distribution functions to analyze extreme situations. Here special focus will be on the use of RCMs for regional climate prediction for the next 10 to 30 years, for which it is unclear if weather and climate extremes may change, where it might happen and if regionas will face extreme situations in the future, in which this never happened before. The dissemination of regional climate "change information" can efficiently be done with the establishment of climate information systems in individual regions of the World.

The use of RCMs as intelligent interpolators of observed data in data sparse regions need to be proven, but RCMs have a great potential for this. In addition, the added value of RCMs compared to GCMs needs to be proven by applying the two-way nesting technique. In mountainous regions like the Alps an impact on the large scale flow can be expected.

The extension of regional climate models (RCM) to regional system models (RSM) is a major challenge for the next years. The coupling to regional ocean models, land use and hydrological models has recently started. Along with this the carbon and nitrogen cycles will be implemented on a regional scale. This allows a much better simulation of additional regional feedbacks as mentioned in Sect. "Definition of the Subject", which might have the potential to modulate the regional climate signal projected by GCMs for the region of interest.

Finally, it must be stated that along with improvements of RCMs, the development of GCMs has to be continued. Their performance in individual regions can be quite poor and an improvement is urgently needed. Here RCMs can deliver important information about regional climatic details and help advancing global climate change simulations.

Acknowledgments

I like to thank the REMO-grooup at the Max Planck Institute for Meteorology for their enthusiastic support.

Bibliography

1. Arnell NM (1996) Global warming, river flows and water resources. Wiley, Chichester
2. Asselin R (1972) Frequency filter for time integrations. Mon Weather Rev 100:487–490
3. Christensen JH, Christensen OB (2004) Climate modelling: severe summer flooding in Europe. Nature 421:805–806
4. Christensen JH, Christensen OB (2007) A summary of the PRUDENCE model projections of changes in European climate by the end of this century. Clim Chan 81(1):7–30
5. Davies HC (1976) A latery boundary formulation for multi-level prediction models. Quart J R Meteorol Soc 102:405–418
6. Dethloff K, Abegg C, Rinke A, Hebestad I, Romanov V (2001) Sensitivity of Arctic climate simulations to different boundary layer parameterizations in a regional climate model. Tellus 53(A):1–26
7. Dethloff K, Rinke A, Lynch A, Dorn W, Saha S, Handorf D (2008) Arctic climate change – The ACSYS decade and beyond. Chapter 8: Arctic regional climate models (in press)
8. Frei C, Christensen JH, Deque M, Jacob D, Jones RG, Vidale PL (2003) Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps. J Geophys Res 108(D3):4124. doi: 10.1029/2002JD002287
9. Giorgi F (2006) Regional climate modeling: Status and perspectives. J Phys IV France 139(2006):101–118. doi: 10.1051/jp4:2006139008
10. Giorgi F, Bates GT (1989) The climatological skill of a regional model over complex terrain. Mon Weather Rev 117:2325–2347

11. Giorgi F, Bates GT, Niemann SJ (1993) The multi-year surface climatology of a regional atmospheric model over the western United States. J Clim 6:75–95
12. Giorgi F, Brodeur CS, Bates GT (1994) Regional climate change scenarios over the United States produced with a nested regional climate model. J Clim 7:375–399
13. Giorgi F, Mearns LO, Shields C, McDaniel L (1998) Regional nested model simulations of present day and 2XCO2 climate over the Central Plains of the US. Clim Chan 40:457–493
14. Giorgi F, Mearns LO (1999) Introduction to special section: Regional climate modelling revisited. J Geophys Res 104:6335–6352
15. Hagemann S, Jacob D (2007) Gradient in the climate change signal of European discharge predicted by a multi-model ensemble. Clim Chan 81(1):309–327
16. Jacob D (2001) A note to the simulation of the annual and inter-annual variability of the water budget over the Baltic Sea drainage basin. Meteorol Atmos Phys 77:61–73
17. Jacob D, Podzun R (1997) Sensitivity studies with the regional climate model REMO. Meteorol Atmos Phys 63:119–129
18. Jacob D, Van den Hurk BJJM, Andrae U, Elgered G, Fortelius C, Graham LP, Jackson SD, Karstens U, Köpken C, Lindau R, Podzun R, Rockel B, Rubel F, Sass BH, Smith RNB, Yang X (2001) A comprehensive model inter-comparison study investigating the water budget during the BALTEX-PIDCAP period. Meteorol Atmos Phys 77:19–43
19. Jacob D, Bärring L, Christensen OB, Christensen JH, Hagemann S, Hirschi M, Kjellström E, Lenderink G, Rockel B, Schär C, Seneviratne SI, Somot S, van Ulden A, van den Hurk B (2007) An inter-comparison of regional climate models for Europe: Design of the experiments and model performance. Clim Chan 81(1):31–52
20. Jacob D, Göttel H, Kotlarski S, Lorenz P, Sieck K (2008) Klimaauswirkungen und Anpassung in Deutschland – Phase 1: Erstellung regionaler Klimaszenarien für Deutschland. Abschlussbericht zum UFOPLAN-Vorhaben 204 41 138, Berichtszeitraum: 1. Oktober 2004 bis 30. September 2007. Max-Planck-Institut für Meteorologie (MPI-M), Hamburg
21. Jones RG, Murphy JM, Noguer M (1995) Simulations of climate change over Europe using a nested regional climate model. I: Assessment of control climate, including sensitivity to location of lateral boundaries. Quart J R Meteorol Soc 121:1413–1449
22. Jones RG, Murphy JM, Noguer M, Keen AB (1997) Simulation of climate change over Europe using a nested regional climate model. II: Comparison of driving and regional model responses to a doubling of carbon dioxide. Quart J R Meteorol Soc 123:265–292
23. Lehmann A, Lorenz P, Jacob D (2004) Modelling the exceptional Baltic Sea flow events in 2002–2003. Geophys Res Lett 31:L21308. doi: 0.1029/2004GL020830
24. Lorenz P, Jacob D (2005) Influence of regional scale information on the global circulation: A two-way nesting climate simulation. Geophys Res Lett 32:L18706. doi: 0.1029/2005GL023351
25. Machenhauer B, Windelband M, Botzet M, Christensen JH, Déqué M, Jones RG, Ruti PM, Visconti G (1998) Validation and analysis of regional present-day climate and climate change simulations over Europe. MPI Report No. 275. MPI, Hamburg
26. Majewski D (1991) The Europa-Modell of the Deutscher Wetterdienst. In: ECMWF seminar on numerical methods in atmospheric models, vol 2. ECMWF, Reading
27. McGregor JL, Katzfey JJ, Nguyen KC (1995) Seasonally varying nested climate simulations over the Australian region. 3rd Int Conf Model Glob Clim Chan Var. Hamburg, Germany, 4–8 Sept 1995
28. McGregor JL, Katzfey JJ, Nguyen KC (1999) Recent regional climate modelling experiments at CISRO. In: Ritchie H (ed) Research activities in atmospheric and oceanic modelling. CAS/JSC Working Group on Numerical Experimentation Report 28. WMO/TD – no. 942. WMO, Geneva, pp 7.37–7.38
29. Nakicenovic N, Alcamo J, Davis G, de Vries B, Fenhann J, Gaffin S, Gregory K, Grübler A, Jung TY, Kram T, La Rovere EL, Michaelis L, Mori S, Morita T, Pepper W, Pitcher H, Price L, Raihi K, Roehrl A, Rogner HH, Sankovski A, Schlesinger M, Shukla P, Smith S, Swart R, van Rooijen S, Victor N, Dadi Z (2000) IPCC special report on emissions scenarios. Cambridge University Press, Cambridge
30. Pitman A (2003) Review: The evolution of, and revolution in, land surface schemes designed for climate models. Int J Climatol 23:479–510
31. Rechid D, Jacob D (2006) Influence of seasonally varying vegetation on the simulated climate in Europe. Meteorol Z 15:99–116
32. Rinke A, Marbaix P, Dethloff K (2004) Internal variability in Arctic regional climate simulations: Case study fort he SHEBA year. Clim Res 27:197–209
33. Roeckner E, Arpe K, Bengtsson L, Christoph M, Claussen M, Dümenil L, Esch M, Giorgetta M, Schlese U, Schulzweida U (1996) The atmospheric general simulation model ECHMA-4: Model description and simulation of present-day climate. Report 218. Max Planck Institute for Meteorology, Hamburg
34. Schär C, Vidale PL, Lüthi D, Frei C, Häberli C, Liniger MA, Appenzeller C (2004) The role of increasing temperature variability in European summer heatwaves. Nature 427:332–336
35. Simmons AJ, Burridge DM (1981) An energy and angular- momentum conserving vertical finite-difference scheme and hybrid vertical coordinate. Mon Weather Rev 109:758–766

Relaxation Oscillations

JOHAN GRASMAN
Wageningen University and Research Centre, Wageningen, The Netherlands

Article Outline

Glossary

Dynamical system At any time t the state of a n-dimensional dynamical system is determined by the state variables forming a vector $x(t)$ in R^n. The change in time of this vector is given by the vector differential equation $\mathrm{d}x/\mathrm{d}t = f(x)$. In the state space R^n solutions of this equation are the trajectories or orbits of the system. If the right-hand side of the equation also depends directly on t, $\mathrm{d}x/\mathrm{d}t = f(t, x)$, then the system is forced and called nonautonomous.

Oscillator An oscillator is a time periodic solution of the system forming a closed orbit in state space. If trajectories, close to a periodic solution, tend to this periodic solution for $t \to \pm\infty$, then the oscillator is called a limit cycle. If a nonautonomous system is periodically forced, $f(t, x) = f(t + T, x)$, then the system may get entrained: Periodic solutions with the same period T or a period kT with some integer larger than 1 (subharmonic solution) may occur.

Singular perturbations If a dynamical system contains a small parameter ε and we let $\varepsilon \to 0$ then the system may become degenerate meaning that it cannot satisfy generic initial- or periodicity conditions. In fact such a system has two time scales: The fast variables vector x and the slow variables vector y: $\varepsilon \mathrm{d}x/\mathrm{d}t = f(x, y)$, $\mathrm{d}y/\mathrm{d}t = g(x, y)$. For $\varepsilon = 0$ initial values outside the manifold $M^{(0)}$: $f(x_0, y_0) = 0$ cannot be satisfied. For small positive ε we see that in the initial phase x changes rapidly until a manifold $M^{(\varepsilon)}$ near $M^{(0)}$ is reached at a part where the points (x_0, y_0) are stable equilibria of $\varepsilon \mathrm{d}x/\mathrm{d}t = f(x, y_0)$. Next the system will enter a quasi-stationary state and changes slowly within $M^{(\varepsilon)}$. With singular perturbations separate approximate solutions are constructed for the two phases. They have the form of power series expansions in ε. Integration constants are determined by a matching procedure.

Relaxation oscillation A relaxation oscillation is a limit cycle of a singularly perturbed dynamical system. Within a cycle at least once the system leaves and returns to the manifold $M^{(\varepsilon)}$.

Canard If within a cycle the relaxation oscillation comes near a part of $M^{(0)}$ with unstable equilibrium points of $\varepsilon \mathrm{d}x/\mathrm{d}t = f(x, y_0)$ then we have a so-called canard type of oscillation. In addition to the parameter ε a second parameter a can be identified that passes a Hopf bifurcation point, where a stable equilibrium changes into a stable relaxation oscillation. Then just before the regular relaxation oscillation arises a canard appears. For very small ε a canard is not easily detected.

Definition of the Subject

A relaxation oscillation is a type of periodic behavior that occurs in physical, chemical and biological processes. To describe it mathematically, a system of coupled nonlinear differential equations is formulated. Such a system is studied with qualitative and quantitative methods of mathematical analysis. The well-known linear pendulum (harmonic oscillator) is not the appropriate system to model real life oscillations. Characteristic for a relaxation oscillator is its nonlinearity and the presence of phases in the cycle with different time scales: A phase of slow change is followed by a short phase of rapid change in which the system jumps to the next stage of slow variation. These oscillations belong to the class of nonlinear systems that give rise to a self-sustained oscillation meaning that the system goes alternately through phases in which energy dissipates and is taken up again.

Van der Pol [66] studied such a type of oscillation in a triode circuit. For small values of a system parameter he found an almost sinusoidal oscillation, while for larger values the system exhibited the type of slow-fast dynamics described above. In the last case the period of the oscillation is almost proportional to that parameter. The name relaxation oscillation, introduced by Van der Pol, refers to this characteristic time constant of the system. In a next publication Van der Pol [67] points out that not only in electronics but in far more fields of science relaxation oscillations may be present. In this respect the physiology of nerve excitation and, more specifically, the heart beat was given special attention by him. In new generations of electronic and electromagnetic systems, such as transistor circuits, Josephson junctions and laser systems, relaxation oscillations are prominently present. Applications are also found in chemistry such as the Zhabotinskii reaction [56], and in geophysics with e. g. the irregular pattern of earthquakes at folds [70]. It is noted that the stick-slip model [71] in the form of a Brownian relaxation oscillator is often brought up in these earthquake studies. In addition to the physiological applications, relaxation oscillations are found in the biology of interacting populations such as in epidemiology [39] and in prey-predator systems, see [12,53,54]. Also in the humanities comparable phenomena are met such as business cycles in economics, see e. g. [31,69].

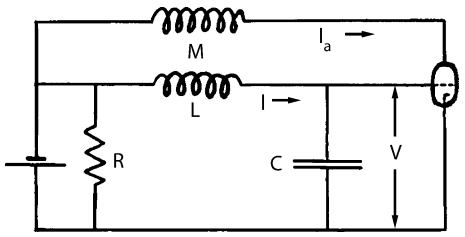
One way to handle mathematically relaxation oscillations is to exploit the presence of a small number. Dynamical systems with a small parameter ε multiplying the time derivative of some of the state variables degenerate when this parameter is set to zero, as it is not possible anymore to satisfy all initial- or periodicity conditions. In the

theory of singular perturbations the limit process $\varepsilon \to 0$ is followed and the nonuniform convergence of the solution to some limit function is taken in consideration. At the point in time, where this limit solution is discontinuous, a new time scale is introduced by a stretching transformation. Then again the limit $\varepsilon \to 0$ is taken leading to a new (locally valid) limit solution. Integration constants in both limit solutions are found from a matching procedure [15]. In a second approach, based on nonstandard analysis [55], it is not necessary to consider at each step of the computational process this limit procedure. The set of real numbers is extended with infinitesimal numbers being numbers that are nonzero and have an absolute value that is smaller than any real number. Then a solution of a differential equation with an infinitesimally small parameter lies infinitely close to the limit solution in an appropriate function space, see [58,79]. The nonstandard approach of a type of relaxation oscillations known as canards or "French ducks" [6] has given the method an important place in the literature. In a third type of approach the attention fully goes into the analysis of the vector field related to the dynamical system. The onset of a phase of fast change in the period is marked by the passage of a special point in state space (the fold point). By a blow up of the vector field at this point [14] the periodic trajectory can be rigorously described over a time interval containing this point and therefore a full description over the entire period is at hand. It has been worked out for relaxation oscillations [38] and is connected to the dynamical systems theory known as geometrical singular perturbations [19].

Introduction

Periodic processes control our daily life. External influences such as the dynamics of sun, earth and moon play an important role in e. g. the seasons, the tides and our day–night rhythm. However, internally we also have our circadian clock as well as other autonomous periodic processes such as the regulation of our metabolic functions. Periodic electric activities in our brain and heart play a key role in the functioning of these organs. Through the work of Hodgkin and Huxley [34] we have a good understanding of the transport of electric pulses in nerve cells. Looking back we notice that the result of Balthasar van der Pol on periodicity in electric circuits is one of the scientific achievements forming the basis of this breakthrough.

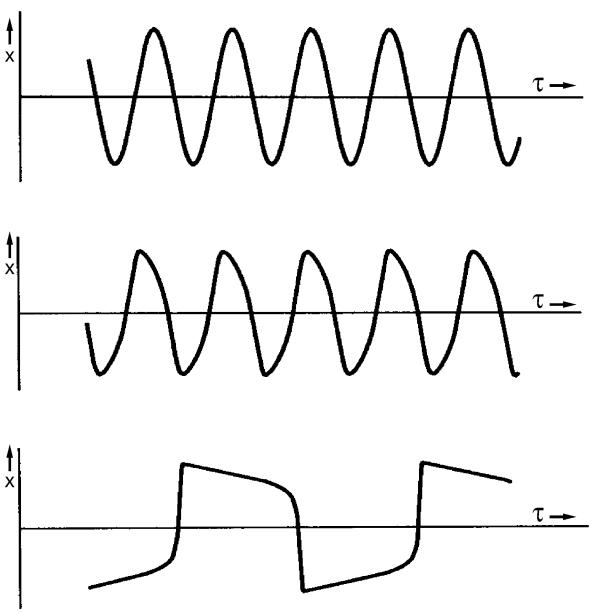
In Fig. 1 Van der Pol's triode circuit is depicted. It produces a self-sustained oscillation due to the nonlinearity in the triode characteristic given by $I_a = V - 1/3V^3$. Replacing V by a scaled potential x and scaling also the time variable we obtain from circuit theory the following differential equation



Relaxation Oscillations, Figure 1
Triode circuit giving rise to a self-sustained oscillation. *L* is a self-inductance, *M* a mutual inductance and *R* a resistance. *I* and *V* are, respectively, the current and the grid voltage

$$\frac{d^2x}{d\tau^2} + \mu(x^2 - 1)\frac{dx}{d\tau} + x = 0, \quad \mu = M/\sqrt{LC} - R\sqrt{C/L}. \tag{1}$$

For μ small a nearly sinusoidal oscillation is found with amplitude close to 2 and a period close to 2π. For $\mu \gg 1$ an almost discontinuous solution appears for which the period is nearly proportional with this parameter, see Fig. 2. The best way to study this last type of oscillation is to rewrite the second-order differential Eq. (1) as a system of two coupled first-order differential equations and to introduce the small parameter $\varepsilon = 1/\mu^2$. Furthermore, a new time scale is introduced $t = \tau/\mu$. It leads to the so-



Relaxation Oscillations, Figure 2
Periodic solution of Eq. (1) for a $\mu = 0.1$, b $\mu = 1$, c $\mu = 10$

called Lienard-type of representation of the system [41]:

$$\varepsilon \frac{\mathrm{d}x}{\mathrm{d}t} = y - F(x), \quad F(x) = 1/3\, x^3 - x, \quad 0 < \varepsilon \ll 1, \quad (2)$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = -x + a\,, \quad a = 0\,. \quad (3)$$

Discontinuous Limit Solution

In Fig. 3a it is seen how the periodic solution of (2) behaves in the limit $\varepsilon \to 0$. It has the form of a discontinuous solution consisting of two time intervals in which the system follows alternately in state space two branches (DA and BC) of the graph of $y = F(x)$ with $|x| > 1$, see Fig. 3b. These are the stable branches; they are rapidly approached because of a fast changing variable x if $y \neq F(x)$. In between the two branches there is the unstable branch AC at which two trajectories move away from the unstable equilibrium at the origin. At the point where a stable branch becomes unstable the system jumps to the opposite stable branch. Clearly the amplitude of the discontinuous oscillation has the value 2 in the x variable. The period follows from

$$T_0 = 2\int_{-2/3}^{2/3} \frac{\mathrm{d}t}{\mathrm{d}y}\mathrm{d}y = 2\int_{-2}^{-1} \frac{1}{-x} F'(x)\mathrm{d}x = 3 - 2\ln 2\,. \quad (4)$$

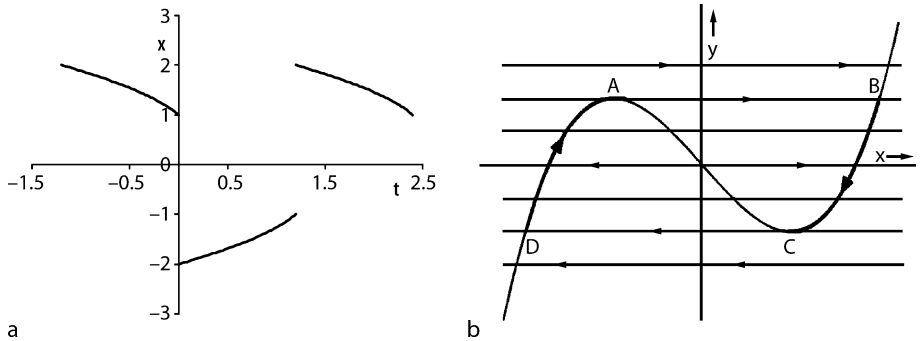
This discontinuous solution can be seen as a zero-order asymptotic approximation of the solution with respect to the small parameter ε. In Sect. "Asymptotic Solution of the Van der Pol Oscillator" we will deal with higher-order approximations using the theory of singular perturbations.

Canards

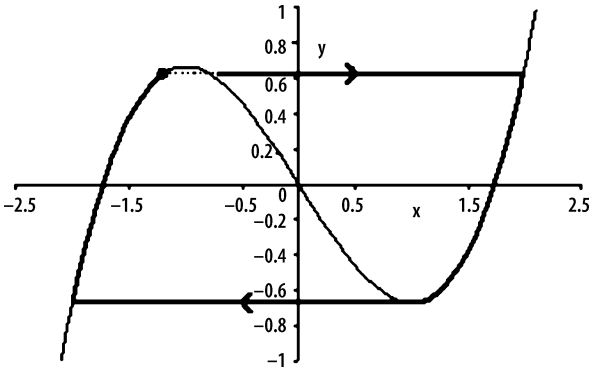
If in formula (2) the parameter a is varied an intriguing type of Hopf bifurcation arises. When passing the values $a = \pm 1$ for decreasing absolute value of a, a stable equilibrium turns unstable and a relaxation oscillation arises. The scenario of a common Hopf bifurcation is that a stable equilibrium changes into an unstable one at the bifurcation point and that a periodic solution branches off with an amplitude that grows in the beginning quadratically as a function of the distance of the parameter to the just passed bifurcation point. However, the above discontinuous periodic solution suggests that directly after the bifurcation point a fully developed relaxation oscillation arises. Using concepts of nonstandard analysis [55] French mathematicians [6] explain how the curious emergence of a relaxation type of oscillation at a Hopf bifurcation can be understood. Also by an intricate asymptotic analysis the phenomenon can be described with singular perturbation theory [16]. In Sect. "Canards" we sketch the result. Finally, as part of a dynamical systems approach (Sect. "Dynamical Systems Approach") a geometrical singular perturbation analysis can be carried out as well [38].

Bonhoeffer–Van der Pol Equation or FitzHugh–Nagumo Equation

When we take in (3) $a = -1 - \delta, 0 < \delta \ll 1$ with δ independent of ε, a stable equilibrium $(\underline{x}, \underline{y})$ arises with $\underline{x} = -1 - \delta$. A small perturbation of the x-component of the system, causing a deviation that stays below 2δ, will damp out quickly. A positive perturbation just above this threshold (dotted line) will trigger a cycle of the system as depicted in Fig. 4. FitzHugh [21] constructed a similar variant of the Van der Pol Equation (2) providing a mathematical model of a nerve excitation as described by Bonhoeffer [3]. Firing of a neuron occurs when an electric stimulation of a dendrite is above some threshold value triggering an electric pulse in the neuron itself which is passed through the axon to other neurons. Also the pres-



Relaxation Oscillations, Figure 3
The discontinuous approximation of the solution of the Van der Pol equation for $\varepsilon \to 0$. The parts AB and DC of the orbit are taken infinitely fast. a The function $x(t)$ for $\varepsilon \to 0$. b The closed orbit in the x, y-plane



Relaxation Oscillations, Figure 4
A perturbation (*dotted line*) of the equilibrium (●) above threshold triggers a cycle. During the cycle perturbations do not have a large effect (refractory period)

ence of a so-called refractory period can be understood from the above system. The refractory period is the time directly after an above threshold stimulation. During this time the neuron is insensitive to perturbations. Later on more refined models evolved [57]. The introduction of spatial structure and diffusion made it possible to consider traveling pulses [1,34].

Orbital Stability, Entrainment, Chaos, Quenching, and Noise

In applications we typically meet nonlinear oscillations that are orbitally highly stable but rather easily speeded up or slowed down in their cycle. In a biological context organisms are provided in this way with a mechanism that helps them to adapt to external circumstances. Also in physics we meet such behavior in systems with strong energy exchange with the environment (lasers), as opposed to conservative systems without dissipation such as in celestial systems.

In (2) entrainment is found if we let the parameter a depend periodically on time with a period say T. If the period $T^{(0)}$ of the autonomous system with $a = 0$ in (2) is sufficiently close to the forcing period T or if the amplitude of the forcing is sufficiently large, then the system will take over this period T. If $T^{(0)}$ is near a value nT, $n = 2, 3, \ldots$ a subharmonic solution with period nT may arise. It also may occur that two stable subharmonics with different n values co-exist. Then the starting value determines which one is chosen. Van der Pol and Van der Mark [68] and Littlewood [43,44] already concluded on respectively, experimental and theoretical grounds that in such a case also other "strange" (chaotic) solutions may be present in the Van der Pol-type oscillator with periodic forcing, see [4,40,42]. For autonomous systems chaotic dynamics can be found in systems consisting of at least three components [23]. A system with one fast and two slow variables already exhibits a wealth of interesting dynamical features, including chaos. Quenching of an oscillation can be achieved by extending the system with a set of differential equations including a feedback to the original system such that the amplitude of the oscillation is reduced by choosing appropriate values for the parameters. For an application to relaxation oscillations see [72,74].

Relaxation oscillations are met in a wide range of applications. Since in practice the action of random forces mostly cannot be excluded, stochastic oscillation forms an essential part of the theory of periodic phenomena in the description of natural processes as well as in the engineering sciences, see [27,33]. An additional reason to take stochasticity in account comes from the fact that the phase velocity of the relaxation oscillator is easily influenced.

Higher-Order Systems and Coupling of Oscillators

Making a generalization we consider the periodic solution of a system of differential equations of the form

$$\varepsilon \frac{dx}{dt} = f(x, y; \varepsilon), \quad \frac{dy}{dt} = g(x, y; \varepsilon), \quad 0 < \varepsilon \ll 1, \tag{5}$$

where x and y are respectively k- and l-dimensional vector functions of time. The vector functions f and g remain bounded for $\varepsilon \to 0$. Then the (slow) dynamics is governed by $y'(t) = g(x, y; 0)$ with constraint $0 = f(x, y; 0)$ and for the fast dynamics we have $\varepsilon x'(t) = f(x, y; 0)$ with y constant. An extension of the Van der Pol-type dynamics to such a higher-dimensional system is not difficult to conceive. However, formulating precisely the conditions for the system (5) to have a periodic solution and to work out a full proof for the existence of such a solution is a rather complex task [46]. As we mentioned before, chaotic solutions may occur too.

Another way to arrive at higher-order systems is to couple relaxation oscillators [76]. For relaxation oscillators of type (2) with different autonomous periods we take

$$\begin{aligned} \varepsilon \frac{dx_i}{dt} &= y_i - F(x_i)\,, \\ \frac{dy_i}{dt} &= -c_i x_i + \sum_{j=0_{j\neq i}}^{n} \delta p_{ij} x_j\,, \quad 0 < \varepsilon \ll \delta \ll 1\,. \end{aligned} \tag{6}$$

For a system of $n = 100$ oscillators with a high degree of coupling, e. g. $p_{ij} \neq 0$ for all $j \neq i$, and widely different frequencies, e. g. $c_i = 0.5 + i/n$, we will observe (after a spin up) a spectrum of frequencies with distinct peaks.

These peaks are related by a simple ratio of their frequencies such as 1:2 or 3:2 and so on, see [17,25]. Considering spatially distributed oscillators with nearest neighbor coupling we meet interesting phenomena such as traveling phase waves. The corresponding fronts may have a preferred direction given by the gradient in the autonomous periods (inherently faster oscillators are likely ahead in phase). Locally, oscillators tend to have an equal actual period (phase velocity) although their intrinsic periods are different. This phenomenon, called plateau behavior, has been investigated by Ermentrout and Kopell [18]. Wave fronts may also move in spirals or travel randomly breaking down when meeting each other as we see in the Zhabotinskii reaction [20] or during fibrillation of the heart [62].

Asymptotic Analysis with Respect to the Small Parameter

By letting $\varepsilon \to 0$ the Van der Pol oscillator (2) reduces to the system:

$$(x_0^2 - 1)\frac{dx_0}{dt} + x_0 = 0 \tag{7}$$

having a discontinuous solution

$$t = \ln(x_0) - \tfrac{1}{2}(x_0^2 - 1) \quad \text{for} \quad -\tfrac{1}{2}T_0 < t < 0\,, \tag{8}$$

$$t = \ln(-x_0) - \tfrac{1}{2}(x_0^2 - 1) + \tfrac{1}{2}T_0 \quad \text{for} \quad 0 < t < \tfrac{1}{2}T_0 \tag{9}$$

with T_0 given by (4). This approximation, see Fig. 3a, is improved in Sect. "Asymptotic Solution of the Van der Pol Oscillator" by the construction of higher-order terms with respect to the small parameter ε. Contrary to regular perturbation problems, where a power series expansion with respect to ε holds over the entire period, we have to consider here separate power series expansions for the stable branches. Moreover, we have to exclude a small neighborhood of the point $t = 0$ where the discontinuity occurs. Near $(t, x) = (0, 1)$ a local approximation is made based on fractional powers of ε. It is followed by an internal layer solution approximating the fast change in x from the value 1 to the starting value -2 at the next stable branch. Unknown (integration) constants are determined by matching the local solutions at $t = 0$ to the regular solutions at the two stable branches, see Sect. "Asymptotic Solution of the Van der Pol Oscillator". Matching of local asymptotic solutions typically applies to differential equations with a small parameter multiplying the highest derivative such as in (1) with τ and μ replaced by t and ε, see [15,50]. In the literature on relaxation oscillations the procedure has been carried out in different ways depending on whether one chooses to solve the problem in the x, t-plane [8], the $x, \dot{x}$-plane [13], or the Lienard plane. For a survey of this literature see [25]. For higher-order systems the construction of matched local asymptotic solutions that involve higher-order terms with respect to ε turns out to be rather complicated, see [46].

A different type of approximation, based on power series expansion with respect to the parameter $\mu = 1/\sqrt{\varepsilon}$, was made possible by the use of computerized formula manipulation packages. In this way elements of a periodic solution such as period and amplitude are approximated also for large values of the parameter μ, see [2,10].

Topological Methods, Mappings and the Dynamical Systems Approach

The existence of a periodic solution for (2) has been proved with the Poincaré–Bendixson theorem. This method, based on topological arguments, only applies to two-dimensional systems. If in the x, y-plane we can construct a domain that does not contain equilibrium points and all trajectories crossing the boundary are entering the domain, then the domain contains a limit cycle. If the domain is a narrow annulus enclosing the limit cycle the method also produces an approximation of the solution together with its period and amplitude [52].

Another important tool in the dynamical systems approach is the Poincaré map or transition map. For autonomous n-dimensional systems we may consider starting values at a bounded transversal $(n-1)$-dimensional surface and analyze how the trajectories intersect some other surface in the state space or the same surface (returnmap). In periodically forced systems we may consider the mapping of the full state space into itself after a time interval equal to the forcing period. Such a type of mapping may also apply to the phase of an oscillator or to the phases of a system of coupled oscillators. A periodic solution corresponds with a fixed point of the return map or, in the case of a forced system, with a fixed point of the map of the state space into itself after one or more periods of the forcing term. Chaotic solutions may occur due to the presence of a so-called horse-shoe map [61]. Levi [40] used a map of this type to describe chaotic solutions of a piecewise linear Van der Pol equation with a sinusoidal forcing term.

Presently, relaxation oscillations are also analyzed with geometrical singular perturbation theory [64]. Geometrical singular perturbations deal with the structure of the slow manifold $f(x, y; 0) = 0$ containing trajectories of (5) with $\varepsilon = 0$ that vary slowly in time [19]. A point of the slow manifold is an equilibrium of the fast system in the

time scale $\xi = t/\varepsilon$: $\mathrm{d}x/\mathrm{d}\xi = f(x, y^{(0)}; 0)$. It is assumed that the eigenvalues of this system linearized at the equilibrium have real parts that are bounded away from zero (hyperbolicity condition). For relaxation oscillations this condition is not satisfied [73]. More details are given in Sect. "Dynamical Systems Approach".

Asymptotic Solution of the Van der Pol Oscillator

From the different methods to approximate asymptotically the solution of (2) we choose the approach of Carrier and Lewis [8], who consider the solution in the t, x-plane, see Fig. 5. Writing (2) again as a second-order differential equation

$$\varepsilon \frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + (x^2 - 1)\frac{\mathrm{d}x}{\mathrm{d}t} + x = 0\,, \tag{10}$$

we can expand the solution valid for the stable branch as

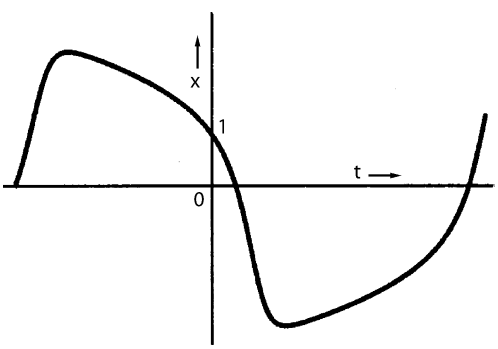
$$x(t;\varepsilon) = x_0(t) + \varepsilon x_1(t) + \varepsilon^2 x_2(t) + \cdots\,. \tag{11}$$

Substitution of (11) in (10) gives, after equating terms with equal powers in ε, a recurrent system of differential equations for the coefficients $x_i(t)$ of (11) with $x_0(t)$ given by (7) when taking the branch for $t < 0$. We see from (8) that for $t \uparrow 0$ this first term behaves as $x_0 \approx 1 + \sqrt{-t}$, so the second-order derivative cannot be neglected near the origin. To analyze the local behavior of the solution in more detail we apply a stretching transformation

$$t = \xi\varepsilon^{2/3}\,, \quad x = 1 + \varepsilon^{1/3} v(\xi)\,. \tag{12}$$

Substitution in (10) yields for the leading terms the equation

$$\frac{\mathrm{d}^2 v_0}{\mathrm{d}\xi^2} + 2v_0\frac{\mathrm{d}v_0}{\mathrm{d}\xi} + 1 = 0 \quad \text{or} \quad \frac{\mathrm{d}v_0}{\mathrm{d}\xi} + v_0^2 + \xi = B\,. \tag{13}$$



Relaxation Oscillations, Figure 5
The solution in the t, x-plane with the fast change just after $t = 0$ starting with a specific local behavior near the point (0,1)

In order to match x_0 for $\xi \to -\infty$ the solution must satisfy $v_0(\xi) \approx \sqrt{-\xi}$ when taking this limit. The solution

$$v_0(\xi) = -\frac{Ai'(-\xi)}{Ai(-\xi)} \tag{14}$$

with $Ai(z)$ the Airy function complies with this matching condition. As $-\xi$ approaches $\alpha = -2.33811\ldots$, being the first zero of the Airy function, the solution will behave as $v_0(\xi) \approx 1/(\xi - \alpha)$. Consequently, the first two terms of (10) will be leading in the left hand side of the equation and make the solution enter the phase of fast change for which we choose the appropriate time variable η using the transformation

$$t = \alpha\varepsilon^{2/3} + \eta\varepsilon\,, \tag{15}$$

so

$$\frac{\mathrm{d}^2 w_0}{\mathrm{d}\eta^2} + (w_0^2 - 1)\frac{\mathrm{d}w_0}{\mathrm{d}\eta} = 0 \quad \text{or} \quad \frac{\mathrm{d}w_0}{\mathrm{d}\eta} = w_0 - \frac{1}{3}w_0^3 + D\,. \tag{16}$$

Matching with v_0 for $\eta \to -\infty$ yields $D = -2/3 - \alpha\varepsilon^{2/3}$ and integrating Eq. (16) we obtain

$$-\eta = \frac{1}{1 - w_0} + \frac{1}{3}\ln\left(\frac{w_0 + 2 + 1/3\alpha\varepsilon^{2/3}}{1 - w_0}\right)\,. \tag{17}$$

For $\eta \to \infty$ the solution behaves as $w_0 = -2 - 1/3\alpha\varepsilon^{2/3} + O(\exp(-3\eta))$. This must match the regular asymptotic solution at the stable branch for $t > 0$ for which the zero-order approximation satisfies

$$t = \ln(-x_0) - \tfrac{1}{2}(x_0^2 - 1) + E \tag{18}$$

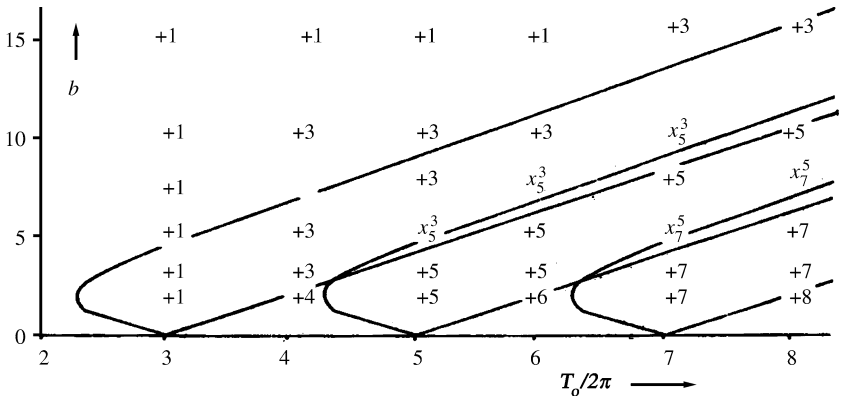
giving $E = 3/2 - \ln(2) + 3/2\alpha\varepsilon^{2/3}$. At $x_0 = -1$ the variable t is at the value $\frac{1}{2}T_0$, so

$$T_0 \approx 3 - 2\ln(2) + 3\alpha\varepsilon^{2/3}\,. \tag{19}$$

The amplitude A follows from the minimal value of w_0 which is approached in (17) for $\eta \to \infty$:

$$A \approx 2 + 1/3\alpha\varepsilon^{2/3}\,. \tag{20}$$

For a higher-order approximation of both the period and the amplitude we refer to Grasman [25]. In Grasman et al. [30] the Lyapunov exponents are approximated in a similar way.



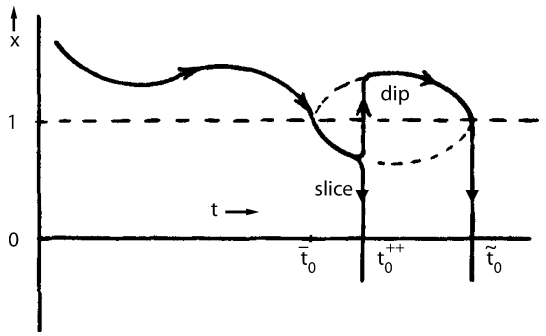
Relaxation Oscillations, Figure 6
Numerical solutions of period $2\pi m$ with m depending on μ and b

Sinusoidal Forcing

We next consider the Van der Pol oscillator with a sinusoidal forcing term

$$\frac{d^2x}{d\tau^2} + \mu(x^2-1)\frac{dx}{d\tau} + x = (b + c\mu)\cos(k\tau)\,, \quad \mu \gg 1\,. \tag{21}$$

A first mathematical investigation of this problem was made by Cartwright and Littlewood [9] followed by Littlewood [43,44]. Subharmonic solutions were constructed. Solutions with different periods may coexist and also solutions that behave chaotically may turn up. Following the method of Carrier and Lewis for this nonautonomous Eq. (21) with $c = 0$ and $k = 1$ we go through the same sequence of asymptotic solutions: The regular approximation of the type (11), the local solution when the trajectory crosses the lines $x = \pm 1$, giving rise to an expression with Airy functions, and a fast changing part directly after with a local approximation similar to (17), see [29]. We may construct various periodic solutions. The most prominent ones have period $2\pi m$ with m an odd number. The value of m depends on μ. In Fig. 6 we depicted the domains in the parameter plane where these solutions are found. Their boundaries follow from the construction of the asymptotic solution. The result is compared with the numerical solution of the system. Notice the overlap of these domains and the coexistence of different numerical solutions. Also chaotic solutions can be approximated asymptotically. They come with the phenomenon of dips and slices which also occurs in the general case for $c \neq 0$ [28]. The main difference with the previous case is that instead of Airy functions parabolic cylinder functions turn up. In Fig. 7 we see how dips and slices arise. They occur when the trajectory stays close to the regular solution



Relaxation Oscillations, Figure 7
A dip and a slice for slightly different initial values chosen such that upon arrival at the line $x = 1$ a point is passed where stable ($x > 1$) and unstable ($x < 1$) regular solutions meet

below the line $x = 1$, while switching from a stable branch ($x > 1$) to an unstable branch ($x < 1$) of this regular solution. It can be seen as a canard type of phenomenon as noted by Smolyan and Wechselberger [63].

Coupled Van der Pol-Type Oscillators

The discontinuous periodic solution of (2) for $\varepsilon \to 0$ is given by (7) and has the form of the vector $\{x_0(t), y_0(t)\}$. We next consider (2) with the parameter a replaced by a small amplitude piecewise continuous periodic function $\delta h(t)$ with period T:

$$\varepsilon\frac{dx}{dt} = y - F(x)\,, \quad \frac{dy}{dt} = -cx + \delta h(t)\,, \quad c = 1\,, \quad 0 < \varepsilon \ll \delta \ll 1\,. \tag{22}$$

For $\varepsilon \to 0$ and δ sufficiently small the solution will take the same orbit as for $\delta = 0$; only the velocity on the limit cycle

is influenced. Consequently, a solution of Eq. (22) can be approximated by

$$\{x_0(\phi(t)), y_0(\phi(t))\}\,. \tag{23}$$

Substitution in the Eq. (7) yields

$$\frac{\mathrm{d}y_0}{\mathrm{d}\phi}\frac{\mathrm{d}\phi}{\mathrm{d}t} = -x_0(\phi(t)) + \delta h(t) \quad \text{or} \quad \frac{\mathrm{d}\phi}{\mathrm{d}t} = 1 - \frac{\delta h(t)}{x_0(\phi(t))}\,. \tag{24}$$

Starting with a phase $\phi(0) = \alpha_0$ and integrating over one period T gives a phase shift

$$\delta\psi(\alpha_0) = \int_0^T \frac{\delta h(t)}{x_0(\phi(t))}\mathrm{d}t\,. \tag{25}$$

Thus, the dynamics of the oscillator is described in a stroboscopic way by the iteration map

$$\alpha_{k+1} = \alpha_k + T + \delta\psi(\alpha_k)(\mathrm{mod})T_0\,. \tag{26}$$

Besides a periodic solution of period T different types of subharmonic solutions can be found as well as chaotic solutions depending on the function $\psi(\alpha)$. An entrained solution corresponds with a stationary point α_s of the map (26). Such a point exists if for some $m = 1, 2, \ldots$

$$\{\delta\psi(\alpha)\}_{\min} < mT_0 - T < \{\delta\psi(\alpha)\}_{\max}\,. \tag{27}$$

Mapping of the phase in the above way is a well-known approach of modeling entrainment by the use of circle maps [51]. It is noted that here we go back further and relate this map to an oscillator given by a differential equation.

If we consider a system of identical oscillators $\{x_0(\phi_i(t)), y_0(\phi_i(t))\}$, $i = 1, 2, \ldots, n$ with a forcing of oscillator i by the others of the form $\delta\Sigma h_{ij}(x_j, y_j)$, we are in the position to analyze different forms of mutual entrainment such as wave phenomena in systems of spatially distributed oscillators with nearest neighbor coupling as described in Sect. "Introduction". A system of spatially distributed oscillators with nearest neighbor coupling can be seen as a discrete representation of a nonlinear diffusion problem with oscillatory behavior. A time delay in the coupling can be handled as well, see [7,26].

It is also allowed that the orbits and the autonomous period of the oscillators are different. As an example we take 25 piecewise linear oscillators with in (22)

$$F(x) = x + 2 \quad \text{for} \quad x < -1\,,$$
$$F(x) = x - 2 \quad \text{for} \quad x > 1\,,$$
$$F(x) = x \quad \text{for} \quad |x| \le 1$$

and $\delta = 0.0025$. For this example an analytic expression for $\psi(\alpha)$ can be derived. For oscillator i the parameter c is in (6) replaced by $c_i(\delta) = 1 - q_i\delta$ with q_i being a random number generated by the normal distribution with expected value zero and standard deviation, so that the autonomous period is $T_0^{(i)} = T_0(1 + \delta q_i)$. As the forcing function of oscillator i we choose $h = \Sigma h_{ij}(x_j) = \Sigma x_j$ with $j \neq i$. The evolution of the dynamics with the oscillators starting in a uniformly distributed random phase has the form of an iteration map of the type (26) and is depicted in Fig. 8. It is observed that by the type of coupling the oscillators are slowed down considerably and that in the almost fully entrained state ($t = 50T_0$) the inherently faster oscillators are slightly ahead in phase.

Canards

We return to the Van der Pol equation in the Lienard-form (2–3) and apply a translation of the state variables x and y as well as the parameter a such that $(x, y; a) = (1, -2/3; 1)$ moves to the origin $(0, 0; 0)$:

$$\varepsilon\frac{\mathrm{d}x}{\mathrm{d}t} = y - F(x), \quad F(x) = 1/3x^3 + x^2, \qquad 0 < \varepsilon \ll 1, \tag{28}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = -x + a\,. \tag{29}$$

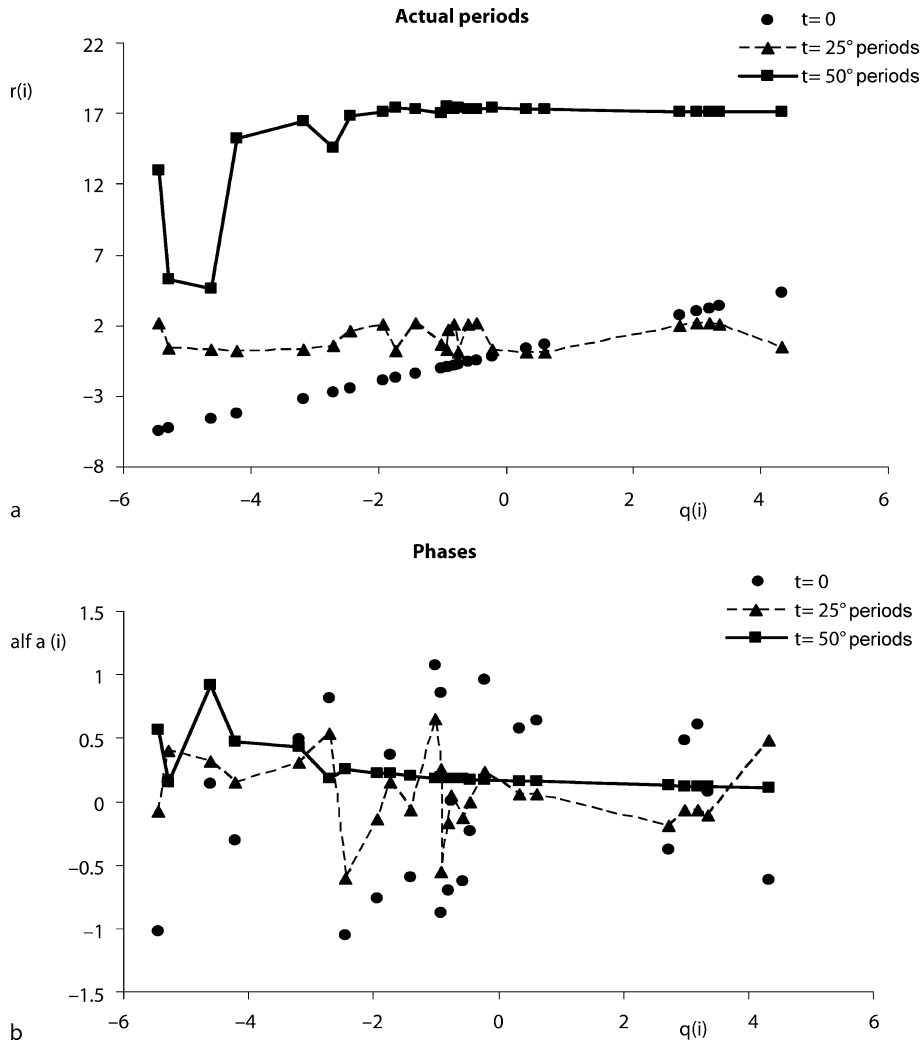
For $-2 < a < 0$ with a fixed and $\varepsilon \to 0$ a relaxation oscillation is found as we analyzed in Sect. "Asymptotic Solution of the Van der Pol Oscillator" using the asymptotic method of Carrier and Lewis [8]. The computation was done for $a = -1$, but any value at this interval would have given the same result. For $a < -2$ or $a > 0$ a stable equilibrium turns up. In a small interval near either the point -2 or the point 0 the solution exhibits a drastic change. Eckhaus [16] analyzes that more explicitly near these values critical points exist, where in a small interval the solution rapidly changes from an equilibrium into a full grown relaxation oscillation. Near $a = 0$ we have the critical point

$$\alpha_c(\varepsilon)\} = -1/8\varepsilon + \cdots\,, \tag{30}$$

where such a change takes place. The following stretching transformation reveals the continuous change of the limit solution as a function of the parameter a:

$$a = \alpha_c(\varepsilon) + \sigma\varepsilon^{3/2}\exp(-k^2/\varepsilon)\,. \tag{31}$$

In Fig. 9 we sketch two periodic limit solutions with σ having different signs. First notice that for $\sigma = o(1)$ a solution



Relaxation Oscillations, Figure 8

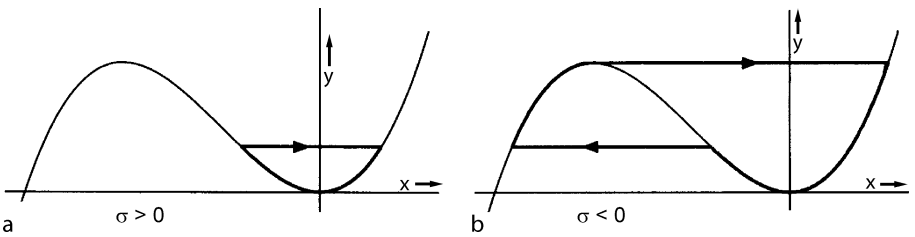
A system of coupled piecewise linear Van der Pol relaxation oscillators with autonomous period $T_0^{(i)} = T_0(1 + \delta q_i)$, $i = 1, \ldots, 25$. The phase of each oscillator runs from $-\frac{1}{2}T_0 = -\ln(3)$ to $-\frac{1}{2}T_0$. The actual phase and period at times $t = nT_0$ are depicted. The actual period $T_i(n)$ is given by its deviation from T_0: $r_i = (T_i(n) - T_0)/\delta$. This expression compares with q_i (for $n = 0$, the *points* are situated at the diagonal)

that is on the stable branch ($x > 0$) will keep following the unstable branch until the point where it meets the other stable branch ($x = -2$) and will jump then back to the stable branch where it started. For $\sigma > 0$ we have a periodic solution that in the limit for $\sigma \to \infty$ transforms into an equilibrium solution (Fig. 9a). For $\sigma < 0$ we have a periodic solution having the shape of a duck that in the limit for $\sigma \to -\infty$ takes the form of a fully developed relaxation oscillation (Fig. 9b).

The phenomenon, that is revealed here, is called a "canard explosion". It is basically due to the fact that a trajectory is very close to a stable branch and that at the moment the branch becomes unstable it keeps on going close to the unstable branch because the deviation from this branch does not grow rapidly, see also [24,60]. A similar phenomenon is met in "slow passage through bifurcation" [45].

Dynamical Systems Approach

We start from the system (5) with vector functions $x(t)$ and $y(t)$ and consider the fast dynamics by making a transformation to the fast time scale $\xi = t/\varepsilon$ and keeping $y = y^{(0)}$



Relaxation Oscillations, Figure 9
Two stages in the canard explosion with parameter values given by (31)

fixed:

$$\frac{dx}{d\xi} = f(x, y^{(0)}; 0)\,. \tag{32}$$

In the dynamical systems approach trajectories of (5) with $\varepsilon = 0$ exist, called fibers having $y = y^{(0)}$ for all t and with x given by (32). Let them move to the point $(x^{(0)}, y^{(0)})$ at the smooth manifold $M^{(0)}$ given by $f(x, y; 0) = 0$. Thus, the point $(x^{(0)}, y^{(0)})$ at $M^{(0)}$ is a stable equilibrium of the fast system (32). The equilibria are the base points of the fibers. Then next at the manifold $M^{(0)}$ the slow dynamics is governed by

$$\frac{dy}{dt} = g(x, y; 0) \quad \text{with constraint} \quad f(x, y; 0) = 0\,. \tag{33}$$

In the geometrical singular perturbation theory of Fenichel [19] a condition for the equilibrium $(x^{(0)}, y^{(0)})$ of (32) is formulated. Fenichel only considers parts of $M^{(0)}$ for which at all points $(x^{(0)}, y^{(0)})$ the Jacobian $\partial f(x, y; 0)/\partial x$ has eigenvalues with a real part that is bounded away from zero. This is called the hyperbolicity condition. It is noted that eigenvalues with positive real parts are allowed. Then the corresponding eigenvectors span a subspace $V^{(u)}$ for which in the system (32) linearized at $(x^{(0)}, y^{(0)})$ all trajectories move away from the equilibrium. For (32) itself a manifold $W^{(u)}$ is defined with $V^{(u)}$ tangent to this manifold at $(x^{(0)}, y^{(0)})$. The manifold $W^{(u)}$ consists of the set of unstable fibers. Similarly the manifold $W^{(s)}$ is defined containing the set of stable fibers. It is proved that, for ε positive and sufficiently small, near $M^{(0)}$ a slow manifold $M^{(\varepsilon)}$ exists satisfying (5). The fast dynamics is governed by the fibers keeping in mind that we move from one fiber to its neighboring one as described by the slow change of the fiber base point at $M^{(\varepsilon)}$. When approximating asymptotically the solution, a transformation is preferred that differs slightly from the matched asymptotic expansions approach. We illustrate it with the following simple system with scalar functions $x(t)$ and $y(t)$:

$$\varepsilon\frac{dx}{dt} = y - x\,, \quad \frac{dy}{dt} = 1\,. \tag{34}$$

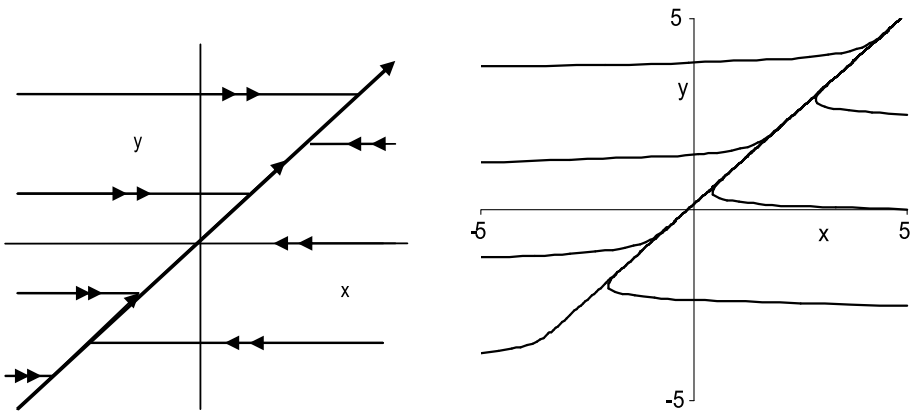
Then for $\varepsilon = 0$ we have fast fibers with $y =$ constant moving towards the slow manifold $M^{(0)}$: $y = x$, see Fig. 10. For small positive ε the slow manifold is defined by $M^{(\varepsilon)}$: $y = x + \varepsilon$. The motion at this manifold (and the change of the base point of the actual fiber) is governed by $dy/dt = 1$. For the fast motion parallel to the fibers we make the transformation $x = y - \varepsilon + v$, so that the new variable v represents the distance to the base point of the fiber. In the fast time scale we then obtain for the motion parallel to the fibers $dv/d\xi = -v$. For more complex systems the expressions for $M^{(\varepsilon)}$ and for the solutions of the slow and fast systems take the form of power series expansions with respect to ε. For more details we refer to [36,37].

For relaxation oscillations the hyperbolicity condition is not satisfied. This is easily verified from the Van der Pol Eq. (2) where $M^{(0)}$: $y = 1/3x^3 - x$, because at the fold points $x = \pm 1$ the eigenvalue of the Eq. (2) linearized at these x values equals zero. To restore hyperbolicity a blow up technique [14] is used. It applies to the system (2) extended with the differential equation $d\varepsilon/dt = 0$ for which the fold point $(x, y, \varepsilon) = (1, -2/3, 0)$ as well as $(-1, 2/3, 0)$ and their direct neighborhoods are put under a magnifying glass. We follow the exposition by Krupa and Szmolyan [38] and consider the neighborhood of a fold point situated at the origin for the following system in the fast time scale

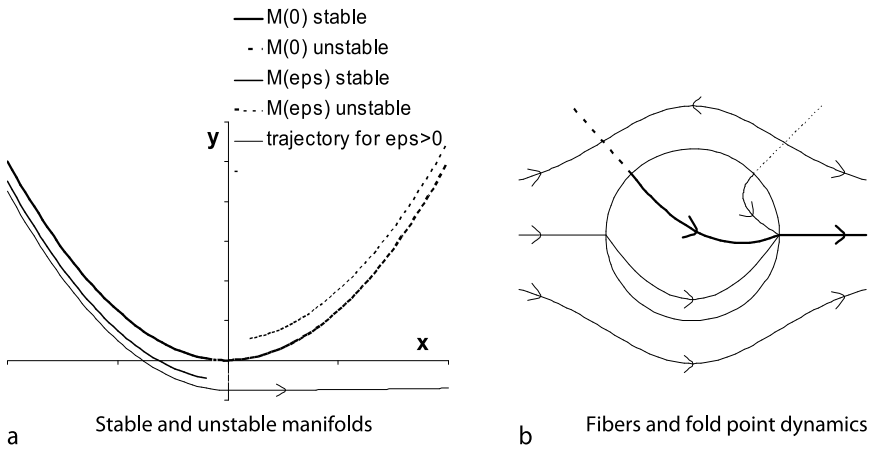
$$\frac{dx}{d\xi} = -y + x^2 + a(x, y, \varepsilon)\,, \quad \frac{dy}{d\xi} = -\varepsilon + b(x, y, \varepsilon)\,, \quad \frac{d\varepsilon}{d\xi} = 0\,, \tag{35}$$

where the terms $a(x, y, \varepsilon)$ and $b(x, y, \varepsilon)$ can be neglected if we are sufficiently close to the origin, see Fig. 11a for the dynamics near the fold point. The blow up transformation $(x, y, \varepsilon) \to (\bar{x}, \bar{y}, \bar{\varepsilon}, \bar{r})$ takes the form

$$x = \bar{r}\bar{x}\,, \quad y = \bar{r}^2\bar{y}\,, \quad \varepsilon = \bar{r}^3\bar{\varepsilon}\,. \tag{36}$$



Relaxation Oscillations, Figure 10
The dynamics of (34) for $\varepsilon = 0$ (*left*) with the fast flow represented by the lines y is constant (fibers) and the slow dynamics by the line $y = x$ being the manifold $M^{(0)}$. For small positive ε (*right*) the trajectories cross the fibers (given for $\varepsilon = 0$) until they arrive at an exponentially small neighborhood of the slow manifold $M^{(\varepsilon)}$



Relaxation Oscillations, Figure 11
Dynamics near and in a fold point, see the text for a detailed description

It is such that in R^3 a ball $B = S^2 \times [0, \rho]$ with ρ sufficiently small is mapped onto R^3. By this transformation the new coordinates $\bar{x}$, $\bar{y}$ and $\bar{\varepsilon}$ indicate a point at a spherical surface S^2. In Fig. 11b a circle is depicted that corresponds with the fold point ($\bar{r} = 0$) in a blow up. Inside the circle, where $\bar{\varepsilon}$ is positive and $\bar{r} = 0$, we are in S^2 above the equator, while at the circle we are at the equator where $\bar{\varepsilon} = 0$. Outside the circle we have $\bar{\varepsilon} = \varepsilon = 0$ but now with $\bar{r} > 0$. There transformed fibers of the limit system near the fold point are drawn. Note that the slow manifold consists of equilibria because of the fast time scale. The bold lines indicate the path of the orbit as it passes the fold point. Away from the equator the dynamics near the fold point is described by taking the fixed value $\bar{\varepsilon} = 1$. We introduce a new time scale $\bar{\xi} = \bar{r}\xi$ and let $\bar{r} \to 0$. In that way we arrive at the system

$$\frac{\mathrm{d}\bar{x}}{\mathrm{d}\bar{\xi}} = -\bar{y} + \bar{x}^2 , \quad \frac{\mathrm{d}\bar{y}}{\mathrm{d}\bar{\xi}} = -1 , \tag{37}$$

having the same solution as (13). For more details we refer to Krupa and Szmolyan [38], who proceed with a similar analysis of a canard explosion. Furthermore, the method has been extended to relaxation oscillations in R^3 [64], including canards, see [5,11,63]. In [75] the method is applied to a prey-predator system.

Future Directions

Over more than 75 years studies on relaxation oscillations have appeared in the literature. It is noticed that each

time a new development takes place in physics or mathematics, relaxation oscillations turn out to be an interesting subject to be investigated with such a novel theory. In electronics we find the use of nonlinear devices starting with triodes in electric circuits [66]. In mathematics we see that relaxation oscillation is a rewarding subject to which many new theories can be applied. In chronological order we mention: Rigorous theories for estimating solutions of differential equations [9], singular perturbation theory and matched asymptotic expansions [8], nonstandard analysis [6], catastrophe theory [69], chaos theory [40], and geometrical singular perturbations [38]. The reason for this lies in the fact that a relaxation oscillation is a highly nonlinear phenomenon with internal mechanisms that are close to real life functions like adaptation by entrainment and threshold behavior in case of triggering. Possible new topics in the field of nonlinear oscillation can, in particular, be found in biology. Literature in which models of biological processes are formulated may function as a source of inspiration. In particular we mention Winfree [78] and Murray [47,48]. To be more explicit about new opportunities we bring up the following cases.

The study of the dynamics of oscillations by circle maps or phase equations turned out to be very useful in particular for the description of entrainment of biological oscillators. In Sect. "Coupled Van der Pol-Type Oscillators" we made the connection between a differential equation model for state variables and the phase equations, see (22)–(26), see also [35]. However, the result only holds for the discontinuous solution as we see from (25), where the denominator of the integrand jumps over the value zero. How in the limit for $\varepsilon \to 0$ we arrive at this result has not yet been worked out. We furthermore mention that spatial distribution of oscillators and the way they are coupled may give rise to interesting entrainment phenomena. In the first place there is the nearest neighbor coupling. Taking an appropriate limit in case of diffusive coupling we may model the dynamics by a nonlinear diffusion equation. Although studies in this direction exist, see e. g. [49], still a lot of work can be done. Furthermore, observations of rhythms in neural networks with long distance coupling show special forms of entrainment, see [65]. Mathematical analysis of such a network of coupled relaxation oscillators may help to understand more from these phenomena. Not only the distance can be varied also the type of coupling; we may consider excitatory and inhibitory forcing referring to the effect they have on an oscillator of the Bonhoeffer–Van der Pol-type, see [77]. Besides easy entrainment of coupled relaxation oscillators there is a second form of adaptation which is comparable with that of neural computing. The coupling between two oscillators in a network may increase if they are entrained for a longer period of time and decrease if they are not. In this way the network is structured by the input and if the input changes the network will follow. The neural theory of cell assemblies [32] is based on such a mechanism known as plasticity of the network, see [22,59]. Effects due to a combination of long range coupling and plasticity may occur as well [79].

Studies on systems with complex interactions between many oscillators as we sketched above may take different directions. On one hand large-scale computer simulations may offer a wealth of results, but models that are more conceptual and are investigated (partly) analytically may give more insight into the fundamental properties of such systems. However, it is not easy to see at forehand that a study of the second type will lead to a meaningful result and is for that reason a more risky enterprise.

Acknowledgment

The author is grateful to Ed Veling (Technical University Delft, The Netherlands) for his valuable advice.

Bibliography

Primary Literature

1. Agüera y Arcas B, Fairhall AL, Bialek W (2003) Computation on a single neuron: Hodglin and Huxley revisited. Neural Comput 15:1715–1749
2. Andersen CM, Geer J (1982) Power series expansions for the frequency and period of the limit cycle of the Van der Pol equation. SIAM J Appl Math 42:678–693
3. Bonhoeffer KF (1948) Activation of passive iron as a model for the excitation of nerve. J Gen Physiol 32:69–91
4. Braaksma B, Grasman J (1993) Critical dynamics of the Bonhoeffer–Van der Pol equation and its chaotic response to periodic stimulation. Physica D 68:265–280
5. Brons M, Krupa M, Wechselberger M (2006) Mixed mode oscillations due to the generalized canard phenomenon. Fields Inst Comm 49:39–63
6. Callot JL, Diener F, Diener M (1978) Le problème de la 'chasse au canard'. CR Acad Sc Paris Sér A 286:1059–1061
7. Campbell SR, Wang DL (1998) Relaxation oscillators with time delay coupling. Physica D 111:151–178
8. Carrier GF, Lewis JA (1953) The relaxation oscillations of the Van der Pol oscillator. Adv Appl Mech 3:12–16
9. Cartwright M, Littlewood J (1947) On nonlinear differential equations of the second order II. Ann Math 48:472–494
10. Dadfar MB, Geer JF (1990) Resonances and power series solutions of the forced Van der Pol oscillator. SIAM J Appl Math 50:1496–1506
11. De Maesschalck P, Dumortier F (2005) Canard solutions at generic turning points. In: Dumortier F, Broer HW, Mawhin J, Vanderbauwhede A, Van Duyn Lunel S (eds) Proc Equadiff 2003. World Scientific, Singapore, pp 900–905
12. Deng B (2004) Food chain chaos with canard explosion. Chaos 14:1083–1092

13. Dorodnicyn AA (1947) Asymptotic solution of Van der Pol's Equation. Akad Nauk SSSR Prikl Mat Mech 11:313–328, Russian (1962, Am Math Soc Transl Series 1 4:1–23)
14. Dumortier F (1993) Techniques in the theory of local bifurcations: Blow up, normal forms, nilpotent bifurcations, singular perturbations. In: Schlomiuk D (ed) Structures in bifurcations and periodic orbits of vector fields. Kluwer, Dordrecht, pp 19–73
15. Eckhaus W (1973) Matched asymptotic expansions and singular perturbations. North-Holland Mathematics Studies 6. North Holland, Amsterdam
16. Eckhaus W (1983) Relaxation oscillations including a standard chase on French ducks. In: Verhulst F (ed) Asymptotic analysis II. Lecture Notes in Math, vol 985. Springer, Berlin, pp 449–494
17. Ermentrout GB (1985) Synchronization in a pool of mutually coupled oscillators with random frequencies. J Math Biol 22:1–10
18. Ermentrout GB, Kopell N (1984) Frequency plateaus in a chain of weakly coupled oscillators I. SIAM J Math Anal 15:215–237
19. Fenichel N (1997) Geometrical singular perturbation theory for ordinary differential equations. J Differ Equ 31:53–98
20. Fife PC (1985) Understanding the patterns in the BZ reagent. J Stat Phys 39:687–703
21. FitzHugh R (1955) Mathematical models of threshold phenomena in the nerve membrane. Bull Math Biophys 17:257–278
22. Gerstner W, Kistler WM (2002) Spiking neuron models: Single neurons, populations, plasticity. Cambridge University Press, Cambridge
23. Ginoux JM, Rossetto B (2006) Differential geometry and mechanics applications to chaotic dynamical systems. Int J Bifurc Chaos 16:887–910
24. Gorelov GN, Sobolev VA (1992) Duck-trajectories in a thermal explosion problem. Appl Math Lett 5:3–6
25. Grasman J (1987) Asymptotic methods for relaxation oscillations and applications. Springer, New York
26. Grasman J, Jansen MJW (1979) Mutually synchronized relaxation oscillators as prototypes of oscillating systems in biology. J Math Biol 7:171–197
27. Grasman J, Roerdink JBTM (1989) Stochastic and chaotic relaxation oscillations. J Stat Phys 54:949–970
28. Grasman J, Nijmeijer H, Veling EJM (1984) Singular perturbations and a mapping on an interval for the forced Van der Pol relaxation oscillator. Physica D 13:195–210
29. Grasman J, Veling EJM, Willems GM (1976) Relaxation oscillations governed by a Van der Pol equation with periodic forcing term. SIAM J Appl Math 31:667–676
30. Grasman J, Verhulst F, Shih S-H (2005) The Lyapunov exponents of the Van der Pol oscillator. Math Meth Appl Sci 28:1131–1139
31. Hamburger L (1934) Note on economic cycles and relaxation oscillations. Econometrica 2:112
32. Hebb DO (1949) The organization of behavior: A neuropsychological theory. Wiley, New York
33. Hilborn RC, Erwin RJ (2005) Fokker–Planck analysis of stochastic coherence in models of an excitable neuron with noise in both fast and slow dynamics. Phys Rev E 72:031112
34. Hodgkin A, Huxley A (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. J Physiol 117:500–544
35. Izhikevich EM (2000) Phase equations for relaxation oscillators. SIAM J Appl Math 60:1789–1804
36. Jones CKRT (1994) Geometric singular perturbation theory In: Arnold L (ed) Dynamical systems, montecatini terme, lecture notes in mathematics, vol 1609. Springer, Berlin, pp 44–118
37. Kaper T (1999) An introduction to geometric methods and dynamical systems theory for singular perturbation problems. In: O'Malley RE Jr, Cronin J (eds) Analyzing multiscale phenomena using singular perturbation methods, proceedings of symposia in applied mathematics, vol 56. American Mathematical Society, Providence, pp 85–132
38. Krupa M, Szmolyan P (2001) Extending geometric singular perturbation theory to nonhyperbolic points-fold and canard points in two dimensions. SIAM J Math Anal 33:286–314
39. Lenbury Y, Oucharoen R, Tumrasvin N (2000) Higher-dimensional separation principle for the analysis of relaxation oscillations in nonlinear systems, application to a model of HIV infection. IMA J Math Appl Med Biol 17:243–161
40. Levi M (1981) Qualitative analysis of periodically forced relaxation oscillations. Memoirs of the American Mathematical Society, no 244. American Mathematical Society, Providence
41. Lienard A (1928) Etude des oscillations entretenues. Revue Générale de l'Electricité 23:901–946
42. Lin KK (2006) Entrainment and chaos in a pulse driven Hodgkin–Huxley oscillator. SIAM J Appl Dyn Syst 5:179–204
43. Littlewood JE (1957) On non-linear differential equations of the second order III $y'' - k(1-y^2)y' + y = b\mu k \cos \mu + \alpha$ for large k and its generalisations. Acta Math 97:267–308
44. Littlewood JE (1957) On non-linear differential equations of the second order IV. Acta Math 98:1–110
45. Marée GJM (1996) Slow passage through a pitchfork bifurcation. SIAM J Appl Math 56:889–918
46. Mishchenko EF, Rosov NK (1980) Differential equations with small parameters and relaxation oscillations. Plenum, New York
47. Murray JD (2002) Mathematical biology I: An introduction. Springer, New York
48. Murray JD (2003) Mathematical biology II: Spatial models and biomedical applications. Springer, New York
49. Nishiura Y, Mimura M (1989) Layer oscillations in reaction-diffusion systems. SIAM J Appl Math 49:481–514
50. O'Malley RE Jr (1991) Singular perturbation methods for ordinary differential equations. Springer, New York
51. Osipov CV, Kurths J (2001) Regular and chaotic phase synchronization of coupled circle maps. Phys Rev E 65:016216
52. Ponzo PJ, Wax N (1965) On certain relaxation oscillations: Asymptotic solutions. SIAM J Appl Math 13:740–766
53. Rinaldi S, Muratori S (1992) Slow-fast limit cycles in predator-prey models. Ecol Model 61:287–308
54. Rinaldi S, Muratori S, Kuznetsov YA (1993) Multiple attractors, catastrophes and chaos in seasonally perturbed predator-prey communities. Bull Math Biol 55:15–35
55. Robinson A (1966) Non-standard analysis. North-Holland, Amsterdam
56. Rössler OE, Wegmann K (1978) Chaos in the Zhabotinskii reaction. Nature 271:89–90
57. Rossoreanu C, Georgescu A, Giurgteanu N (2000) The FitzHugh–Nagumo model – bifurcation and dynamics. Kluwer, Dordrecht
58. Sari T (1996) Nonstandard perturbation theory of differential equations, at symposium on nonstandard analysis and its applications. ICMS, Edinburgh

59. Seliger P, Young SC, Tsimring LS (2002) Plasticity and learning in a network of coupled phase oscillators. Phys Rev E 65:041906
60. Shchepakina E, Sobolev V (2001) Integral manifolds, canards and black swans. Nonlinear Anal Ser A: Theory Meth 44:897–908
61. Smale S (1967) Differentiable dynamical systems. Bull Am Math Soc 73:747–817
62. Strumiłło P, Strzelecki M (2006) Application of coupled neural oscillators for image texture segmentation and modeling of biological rhythms. Int J Appl Math Comput Sci 16:513–523
63. Szmolyan P, Wechselberger M (2001) Canards in R^3. J Differential Equ 177:419–453
64. Szmolyan P, Wechselberger M (2004) Relaxation oscillations in R^3. J Differential Equ 200:69–104
65. Traub RD, Whittington MA, Stanford IM, Jefferys JGR (1991) A mechanism for generation of long-range synchronous fast oscillators in the cortex. Nature 383:621–624
66. Van der Pol B (1926) On relaxation oscillations. Phil Mag 2:978–992
67. Van der Pol B (1940) Biological rhythms considered as relaxation oscillations. Acta Med Scand Suppl 108:76–87
68. Van der Pol B, Van der Mark J (1927) Frequency demultiplication. Nature 120:363–364
69. Varian HR (1979) Catastrophe theory and the business cycle. Econ Inq 17:14–28
70. Vasconcelos GL (1996) First-order phase transition in a model for earthquakes. Phys Rev Lett 76:4865–4868
71. Vatta F (1979) On the stick-slip phenomenon. Mech Res Commun 6:203–208
72. Verhulst F (2005) Quenching of self-excited vibrations. J Engin Math 53:349–358
73. Verhulst F (2007) Periodic solutions and slow manifolds. Int J Bifurc Chaos 17:2533–2540
74. Verhulst F, Abadi (2005) Autoparametric resonance of relaxation oscillations. Z Angew Math Mech 85:122–131
75. Verhulst F, Bakri T (2007) The dynamics of slow manifolds. J Indon Math Soc 13:1–10
76. Wang DL (1999) Relaxation oscillators and networks. In: Webster JG (ed) Wiley encyclopedia of electrical and electronics engineering, vol 18. Wiley, Malden, pp 396–405
77. Wang DL, Terman D (1995) Locally excitatory globally inhibitory oscillator networks. IEEE Trans Neural Netw 6:283–286
78. Winfree AT (2000) The geometry of biological time, 2nd edn. Springer, New York
79. Womelsdorf T, Schoffelen JM, Oosterveld R, Singer W, Desimore R, Engel AK, Fries P (2007) Modulation of neural interactions through neural synchronization. Science 316:1609–1612
80. Zvonkin AK, Shubin MA (1984) Nonstandard analysis and singular perturbations of ordinary differential equations. Russ Math Surveys 39:69–131

Books and Reviews

Beuter A, Glass L, Mackey MC, Titcombe MS (eds) (2003) Nonlinear dynamics in physiology and medicine. Springer, New York
Guckenheimer J, Holmes P (1983) Nonlinear oscillations, dynamical systems, and bifurcations of vector fields. Springer, New York
Verhulst F (2005) Methods and applications of singular perturbations: Boundary layers and multiple timescale dynamics. Springer, New York

Repeated Games with Complete Information

Olivier Gossner[1], Tristan Tomala[2]
[1] PSE, UMR CNRS-EHESS-ENPC-ENS 8545, Northwestern University, Paris, France
[2] Economics and Finance Department, HEC Paris, Paris, France

Article Outline

Glossary

Repeated game A model of repeated interaction between agents.

Behavioral strategy A decision rule that prescribes a randomized choice of actions for every possible history.

Nash equilibrium A strategy profile from which no unilateral deviation is profitable.

Sequential equilibrium A strategy profile and Bayesian beliefs on past histories such that after every history, every agent is acting optimally given his beliefs.

Monitoring structure A description of player's observation of each other's strategic choices. It specifies, for every profile of actions, the probability distribution over the profiles of individual signals received by the agents.

Definition of the Subject

Repeated interactions arise in several domains such as Economics, Computer Science, and Biology.

The theory of repeated games models situations in which a group of agents engage in a strategic interaction over and over. The data of the strategic interaction is fixed over time and is known by all the players. This is in contrast with stochastic games, for which the data of the strategic interaction is controlled by player's choices, and repeated games with incomplete information, where the stage game is not common knowledge among players (the reader is referred to the corresponding chapters of this Encyclopedia). Early studies of repeated games include Luce and Raiffa [48] and Aumann [4]. In the context of production games, Friedman [25] shows that, while the compet-

itive outcome is the only one compatible with individual profit maximization under a static interaction, collusion is sustainable at an equilibrium when the interaction is repeated.

Generally, repeated games provide a framework in which individual utility maximization by selfish agents is compatible with welfare maximization (common good), while this is known to fail for many classes of static interactions.

Introduction

The discussion of an example shows the importance of repeated games and introduces the questions studied.

Consider the following game referred to as the Prisoner's Dilemma:

	C	D
C	4, 4	5, 5
D	5, 0	1, 1

The Prisoner's Dilemma

Player 1 chooses the row, player 2 chooses the column, and the pair of numbers in the corresponding cell are the payoffs to players 1 and 2 respectively.

In a one-shot interaction, the only outcome consistent with game theory predictions is (D, D). In fact, each player is better off playing D whatever the other player does.

On the other hand, if players engage in a repeated Prisoner's Dilemma, if they value sufficiently future payoffs compared to present ones, and if past actions are observable, then (C, C) is a sustainable outcome. Indeed, if each player plays C as long as the other one has always done so in the past, and plays D otherwise, both player have an incentive to always play C, since the short term gain that can be obtained by playing D is more than offset by the future losses entailed by the opponent playing D at all future stages.

Hence, a game theoretical analysis predicts significantly different outcomes from a repeated game than from static interaction. In particular, in the Prisoner's Dilemma, the cooperative outcome (C, C) can be sustained in the repeated game, while only the non-cooperative outcome (D, D) can be sustained in one-shot interactions.

In general, what are the equilibrium payoffs of a repeated game and how can they be computed from the data of the static game? Is there a significant difference between games repeated a finite number of times and infinitely repeated ones? What is the role played by the degree of impatience of players? Do the conclusions obtained for the Prisoner's Dilemma game and for other games rely crucially on the assumption that each player perfectly observes other player's past choices, or would imperfect observation be sufficient? The theory of repeated games aims at answering these questions, and many more.

Games with Observable Actions

This section focuses on repeated games with perfect monitoring in which, after every period of the repeated game, all strategic choices of all the players are publicly revealed.

Data of the Game, Strategies, Payoffs

Data of the Stage Game There is a finite set I of players. A stage game is repeated over and over. Each player i's action set in this stage game is denoted A_i, and $S_i = \Delta(A_i)$ is the set of player i's mixed actions (for any finite set X, $\Delta(X)$ denotes the set of probabilities over X). Every degenerate lottery in S_i (which puts probability 1 to one particular action in A_i) is associated to the corresponding element in A_i. A choice of action for every player i determines an outcome $a \in \prod_i A_i$. The payoff function of the stage game is $g\colon A \to \mathbb{R}^I$. Payoffs are naturally associated to profiles of mixed actions $s \in S = \prod_i S_i$ using the expectation: $g(s) = \mathbf{E}_s g(a)$.

Repeated Game After every repetition of the stage game, the action profile previously chosen by the players is publicly revealed. After the t first repetitions of the game, a player's information consists of the publicly known history at stage t, which is an element of $H_t = A^t$ ($H_0 = \{\emptyset\}$ by convention). A strategy in the repeated game specifies the choice of a mixed action at every stage, as a function of the past observed history. More specifically, a *behavioral* strategy for player i is of the form $\sigma_i\colon \cup_t H_t \to S_i$. When all the strategy choices belong to A_i ($\sigma_i\colon \cup_t H_t \to A_i$), σ_i is called a *pure* strategy.

Other Strategy Specifications A behavioral strategy allows the player to randomize his action depending on past history. If, at the start of the repeated game, the player was to randomize over the set of behavioral strategies, the result would be equivalent to a particular behavioral strategy choice. This result is a consequence of Kuhn's theorem [5,41]. Furthermore, behavioral strategies are also equivalent to randomizations over the set of pure strategies.

Induced Plays Every choice of pure strategies $\sigma = (\sigma_i)_i$ by all the players induces a play $h = (a_1, a_2, \dots) \in A^\infty$ in the repeated game, defined inductively by $a_1 = (\sigma_{i,0}(\emptyset))$ and $a_t = (\sigma_{i,t-1}(a_1, \dots, a_{t-1}))$. A profile of behavioral strategies σ defines a probability distribution P_σ over plays.

Preferences To complete the definition of the repeated game, it remains to define player's preferences over plays. The literature commonly distinguishes infinitely repeated games with or without discounting, and finitely repeated games.

In infinitely repeated games with no discounting, the players care about their long-run stream of stage payoffs. In particular, the payoff in the repeated game associated to a play $h = (a_1, a_2, \dots) \in A^\infty$ coincides with the limit of the Cesaro means of stage payoffs when this limit exists. When this limit does not exist, the most common evaluation of the stream of payoffs is defined through a Banach limit of the Cesaro means (a Banach limit is a linear form on the set of bounded sequences that lies always between the liminf and the limsup).

In infinitely repeated games with discounting, a discount factor $0 < \delta < 1$ characterizes the player's degree of impatience. A payoff of 1 at stage $t+1$ is equivalent to a payoff of δ at stage t. Player i's payoff in the repeated game for the play $h = (a_1, a_2, \dots) \in A^\infty$ is the normalized sum of discounted payoffs: $(1-\delta)\sum_{t\geq 1} \delta^{t-1} g_i(a_t)$.

In finitely repeated games, the game ends after some stage T. Payoffs induced by the play after stage T are irrelevant (and a strategy needs not specify choices after stage T). The payoff for a player is the average of the stage payoffs during stages 1 up to T: $\frac{1}{T}\sum_{t=1}^{T} g_i(a_t)$.

Equilibrium Notions What plays can be expected to be observed in repeated interactions of players who observe each other's choices? Non-cooperative Game Theory focuses mainly on the idea of stable convention, i. e. of strategy profiles from which no player has incentives to deviate, knowing the strategies adopted by the other players.

A strategy profile forms a Nash Equilibrium (Nash [55]) when no player can improve his payoff by choosing an alternative strategy, as long as other players follow the prescribed strategies.

In some cases, the observation of past play may not be consistent with the prescribed strategies. When, for every possible history, each player's strategy maximizes the continuation stream of payoffs, assuming that other players abide with their prescribed strategies at all future stages, the strategy profile forms a subgame perfect equilibrium (Selten [66]).

Perfect equilibrium is a more robust and often considered a more satisfactory solution concept than Nash equilibrium. The construction of perfect equilibria is in general also more demanding than the construction of Nash equilibria.

The main objective of the theory of repeated games is to characterize the set of payoff vectors that can be sustained by some Nash or perfect equilibrium of the repeated game.

Necessary Conditions on Equilibrium Payoffs

Some properties are common to all equilibrium payoffs. First, under the common assumption that all players evaluate the payoff associated to a play in the same way, the resulting payoff vector in the repeated game is a convex combination of stage payoffs. That is, the payoff vector in the repeated game is an element of the convex closure of $g(A)$, called the set of *feasible payoffs* and denoted F.

A notable exception is the work of Lehrer and Pauzner [47] who study repeated games where players have heterogeneous time preferences. The payoff vector resulting from a play does not necessarily belong to F if players have different evaluations of payoff streams. For instance, in a repetition of the Prisoner's Dilemma, if player 1 cares only about the payoff in stage 1 and player 2 cares only about the payoff in stage 2, it is possible for both players to obtain a payoff of 4 in the repeated game.

Now consider a strategy profile σ, and let τ_i be a strategy of player i that plays after every history $(a_1, \dots, a_t)$ a best response to the profile of mixed actions chosen by the other players in the next stage. At any stage of the repeated game, the expected payoff for player i using τ_i is no less than

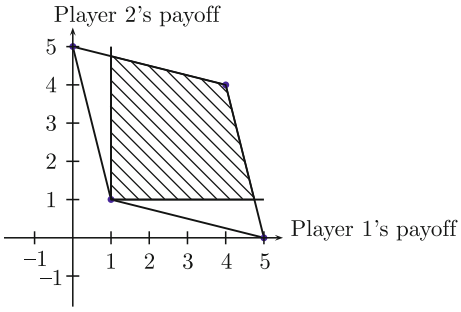
$$v_i = \min_{s_{-i} \in S_{-i}} \max_{a_i \in A_i} g_i(s_{-i}, a_i) \tag{1}$$

where $s_{-i} = (s_j)_{j \neq i}$ (we use similar notations throughout the paper: for a family of sets $(E_i)_{i\in I}$, e_{-i} denotes an element of $E_{-i} = \Pi_{j\neq i} E_j$ and a profile $e \in \Pi_j E_j$ is denoted $e = (e_i, e_{-i})$ when the ith component is stressed).

The payoff v_i is referred to as player i's min max payoff. A payoff vector that provides each player i with at least [resp. strictly more than] v_i is called *individually rational* [resp. strictly individually rational], and IR [resp. IR^*] denotes the set of such payoff vectors. Since for any strategy profile, there exists a strategy of player i that yields a payoff no less than v_i, all equilibrium payoffs have to be individually rational.

Also note that players $j \neq i$ collectively have a strategy profile in the repeated game that forces player i's payoff down to v_i: they play repeatedly a mixed strategy profile that achieves the minimum in the definition of v_i. Such a strategy profile in the one-shot game is referred to as punishing strategy, or min max strategy against player i.

For the Prisoner's Dilemma game, F is the convex hull of $(1,1)$, $(5,0)$, $(0,5)$ and $(4,4)$. Both player's min max levels are equal to 1. Figure 1 illustrates the set of feasible and individually rational payoff vectors (hatched area):



Repeated Games with Complete Information, Figure 1
***F* and *IR* for the Prisoner's Dilemma**

The set of feasible and individually rational payoffs can be directly computed from the stage game data.

Infinitely Patient Players

The following result has been part of the folklore of Game Theory at least since the mid 1960s. Its authorship is obscure (see the introduction of Aumann [7]). For this reason, it is commonly referred to as the "Folk Theorem". By extension, characterization of sets of equilibrium payoffs in repeated games are also referred to as "Folk Theorems".

Theorem 1 *The set of equilibrium payoffs of the repeated game with no discounting coincides with the set of feasible and individually rational payoffs.*

Aumann and Shapley [8,9] and Rubinstein [62,64] show that restricting attention to perfect equilibria does not narrow down the set of equilibrium payoffs. They prove that:

Theorem 2 *The set of perfect equilibrium payoffs of the repeated game with no discounting coincides with the set of feasible and individually rational payoffs.*

We outline a proof of Theorem 2. It is established that any equilibrium payoff is in $F \cap IR$. We need only to prove that every element of $F \cap IR$ is a subgame perfect equilibrium payoff. Let $x \in F \cap IR$, and let $h = a_1, \ldots, a_t, \ldots$ be a play inducing x. Consider the strategies that play a_t in stage t; if player i does not respect this prescription at stage t_0, the other players punish player i for t_0 stages by repeatedly playing the min max strategy profile against player i. After the punishment phase is over, players revert to the play of h, hence playing $a_{2t_0+1}, \ldots$.

Now we explain why these strategies form a subgame perfect equilibrium. Consider a strategy of player i starting after any history. The induced play by this strategy for player i and by other player's prescribed strategies is, up to a subset of stages of null density, defined by the sequence h with interweaved periods of punishment for player i. Hence the induced long-run payoff for player i is a convex combination of his punishment payoff and of the payoff induced by h. The result follows since the payoff for player i induced by h is no worse than the punishment payoff.

Impatient Players

The strategies constructed in the proof of the Folk Theorem for repeated games with infinitely patient players (Theorem 1) do not necessarily constitute a subgame perfect equilibrium if players are impatient. Indeed, during a punishment phase, the punishing players may be receiving low stage payoffs, and these stage payoffs matter in the evaluation of their stream of payoffs. When constructing subgame perfect equilibria of discounted games, one must make sure that after a deviation of player i, players $j \neq i$ have incentives to implement player i's punishment.

Nash Reversion Friedman [25] shows that every feasible payoff that Pareto dominates a Nash equilibrium payoff of the static game is a subgame perfect equilibrium payoff of the repeated game provided that players are patient enough. In Friedman's proof, punishments take the simple form of reversion to the repeated play of the static Nash equilibrium forever. In the Prisoner's Dilemma, (D, D) is the only static Nash equilibrium payoff, and thus $(4, 4)$ is a subgame perfect Nash equilibrium payoff of the repeated game if players are patient enough. Note however that in some games, the set of payoffs that Pareto dominate some equilibrium payoff may be empty. Also, Friedman's result constitutes a partial Folk Theorem only in that it does not characterize the full set of equilibrium payoffs.

The Recursive Structure Repeated games with discounting possess a structure similar to dynamic programming problems. At any stage in time, players choose actions that maximize the sum of the current payoff and the payoff at the subsequent stages. When strategies form a subgame perfect equilibrium, the payoff vector at subsequent stages must be an equilibrium payoff, and players must have incentives to follow the prescribed strategies at the current stage. This implies that subgame perfect equilibrium payoffs have a recursive structure, first studied by Abreu [1]. Subsection "A Recursive Structure" presents the recursive structure in more details for the more general model of games with public monitoring.

The Folk Theorem for Discounted Games Relying on Abreu's recursive results, Fudenberg and Maskin [20]

prove the following Folk Theorem for subgame perfect equilibria with discounting:

Theorem 3 *If the number of players is 2 or if the set feasible payoff vectors has non-empty interior, then any payoff vector that is feasible and strictly individually rational is a subgame perfect equilibrium of the discounted repeated game, provided that players are sufficiently patient.*

Forges, Mertens and Neyman [24] provide an example for which a payoff which is individually rational but not strictly individually rational is not an equilibrium payoff of the discounted game.

Abreu, Dutta and Smith [2] show that the non-empty interior condition of the theorem can be replaced by a weaker condition of "non equivalent utilities": no pair of players have the same preferences over outcomes. Wen [73] and Fudenberg, Levine and Takahashi [22] show that a Folk Theorem still holds when the condition of non equivalent utilities fails if one replaces the minmax level defining individually rational payoffs by some "effective minmax" payoffs.

An alternative representation of impatience to discounted payoffs in infinitely repeated games is the overtaking criterion, introduced by Rubinstein [63]: the play $(a_1, a_2, \dots)$ is strictly preferred by player i to the play $(a'_1, a'_2, \dots)$ if the inferior limit of the difference of the corresponding streams of payoffs is positive, i. e. if $\liminf_T \sum_{t=1}^{T} g_i(a_t) - g_i(a'_t) > 0$. Rubinstein [63] proves a Folk Theorem with the overtaking criterion.

Finitely Repeated Games

Strikingly, equilibrium payoffs in finitely repeated games and in infinitely repeated games can be drastically different. This effect is best exemplified in repetitions of the Prisoner's Dilemma.

The Prisoner's Dilemma Recall that in an infinitely repeated Prisoner's Dilemma, cooperation at all stages is achieved at a subgame perfect equilibrium if players are patient enough. By contrast, at every Nash equilibrium of any finite repetition of the Prisoner's Dilemma, both players play D at every stage with probability 1.

Now we present a short proof of this result. Consider any Nash equilibrium of the Prisoner's Dilemma repeated T times. Let $a_1, \dots, a_T$ be a sequence of action profiles played with positive probability at the Nash equilibrium. Since each player can play D at the last stage of the repetition, and D is a dominating action, $a_T = (D, D)$. We now prove by induction on τ that for any such τ, $a_{T-\tau}, \dots, a_T = (D, D), \dots, (D, D)$. Assume the induction hypothesis valid for $\tau - 1$. Consider a strategy for player i that follows the equilibrium strategy up to stage $T - \tau - 1$, then plays D from stage $T - \tau$ on. This strategy obtains the same payoff as the equilibrium strategy at stages $1, \dots, T - \tau - 1$, and as least as much as the equilibrium strategy at stages $T - \tau + 1, \dots, T - \tau$. Hence, this strategy cannot obtain more than the equilibrium strategy at stage $T - \tau$, and therefore the equilibrium strategy plays D at stage $T - \tau$ with probability 1 as well.

Sorin [68] proves the more general result:

Theorem 4 *Assume that in every Nash equilibrium of G, all players are receiving their individually rational levels. Then, at every Nash equilibrium of any finitely repeated version of G, all players are receiving their individually rational levels.*

The proof of Theorem 4 relies on a backwards induction type of argument, but it is striking that the result applies for *all* Nash equilibria and not only for *subgame perfect* Nash equilibria. This result shows that, unless some additional assumptions are made on the one-shot game, a Folk Theorem cannot obtain for finitely repeated games.

Games with Unique Nash Payoff Using a proof by backwards induction, Benoît and Krishna [10] obtain the following result.

Theorem 5 *Assume that G admits x as unique Nash equilibrium payoff. Then every finite repetition of G admits x as unique subgame perfect equilibrium payoff.*

Theorems 4 and 5 rely on the assumption that the last stage of repetition, T, is common knowledge between players. Neyman [56] shows that a Folk Theorem obtains for the finitely repeated Prisoner's Dilemma (and for other games) if there is lack of common knowledge on the last stage of repetition.

Folk Theorems for Finitely Repeated Games A Folk Theorem can be obtained when there are two Nash equilibrium payoffs for each player. The following result is due to Benoît and Krishna [10] and Gossner [27].

Theorem 6 *Assume that each player has two distinct Nash equilibrium payoffs in G and that the set of feasible payoffs has non-empty interior. Then, the set of subgame perfect equilibrium payoffs of the T times repetition of G converges to the set of feasible and individually rational payoffs as T goes to infinity.*

Hence, with at least two equilibrium payoffs per player, the sets of equilibrium payoffs of finitely repeated games and infinitely repeated games are asymptotically the same.

The condition that each player has two distinct Nash equilibrium payoffs in the stage game can be weakened, see Smith [67]. Assume for simplicity that *one* player has two distinct Nash payoffs. By playing one of the two Nash equilibria in the last stages of the repeated game, it is possible to provide incentives for this player to play actions that are not part of Nash equilibria of the one-shot game in previous stages. If this construction leads to perfect equilibria in which a player $j \neq i$ has distinct payoffs, we can now provide incentives for both players i and j. If successive iterations of this procedure yield distinct subgame perfect equilibrium payoffs for all players, a Folk Theorem applies.

Games with Non-observable Actions

For infinitely repeated games with perfect monitoring, a complete and simple characterization of the set of equilibrium payoffs is obtained: feasible and individually rational payoff vectors. In particular, cooperation can be sustained at equilibrium. How equilibrium payoffs of the repeated game depend on the quality of player's monitoring of each other's actions is the subject of a very active area of research.

Repeated games with imperfect monitoring, in which players observe imperfectly other player's action choices, were first motivated by economic applications. In Stigler [69], two firms are repeatedly engaged in price competition over market shares. Each firm observes its own sales, but not the price set by the rival. While it is in the best interest for both firms to set a collusive price, each firm has incentives to secretly undercut the rival's price. Upon observing plunging sales, should a firm deduce that the rival firm is undercutting prices, and retaliate by setting lower prices, or should lower sales be interpreted as a result of an exogenous shock on market demand? Whether collusive behavior is sustainable or not at equilibrium is one of the motivating questions in the theory of repeated games with imperfect monitoring.

It is interesting to compare repeated games with imperfect monitoring with their perfect monitoring counterparts.

The structure of equilibria used to prove the Folk Theorem with perfect monitoring and no discounting is rather simple: if a player deviates from the prescribed strategies, the deviation is detected and the deviating player is identified, all other players can then punish the deviator. With imperfect monitoring, not all deviations are detectable, and when a deviation is detected, deviators are not necessarily identifiable. The notions of detection and identification allow fairly general Folk Theorems for undiscounted games. We present these results in Subsect. "Detection and Identification".

With discounting, repeated games with perfect monitoring possess a recursive structure that facilitates their study. Recursive methods can also be successfully applied to discounted games with public monitoring. We review the major results of this branch of the literature in Subsect. "Public Equilibria".

Almost-perfect monitoring is the natural framework to study the effect of small departures from the perfect or public monitoring assumptions. We review this literature in Subsect. "Almost Perfect Monitoring".

Little is known about general discounted games with imperfect private monitoring. We present the main known results in Subsect. "General Stochastic Signals".

With perfect monitoring, the worst equilibrium payoff for a player is given by the min max of the one-shot game, where punishing (minimizing) players choose an independent profile of mixed strategies. With imperfect monitoring, correlation past signals for the punishing players may lead to more efficient punishments. We present results on punishment levels in Subsect. "Punishment Levels".

Model

In this section we define repeated games with imperfect monitoring, and describe several classes of monitoring structures of particular interest.

Data of the Game Recall that the one-shot strategic interaction is described by a finite set I of players, a finite action set A_i for each player i, and a payoff function $g\colon A \to \mathbb{R}^I$. Player's observation of each other's actions is described by a *monitoring structure* given by a finite set of signals Y_i for each player i and by a transition probability $Q\colon A \to \Delta(Y)$ (with $A = \prod_{i\in I} A_i$ and $Y = \prod_{i\in I} Y_i$). When the action profile chosen is $a = (a_i)_{i\in I}$, a profile of signals $y = (y_i)_{i\in I}$ is drawn with probability $Q(y|a)$ and y_i is observed by player i.

Perfect Monitoring Perfect monitoring is the particular case in which each player observes the action profile chosen: for each player i, $Y_i = A$ and $Q((y_i)_{i\in I}|a) = \mathbb{1}_{\{\forall i,\, y_i = a\}}$.

Almost Perfect Monitoring The monitoring structure is ε-perfect (see Mailath and Morris [52]) when each player can identify the other player's action with a probability of error less than ε. This is the case if there exist functions $f_i\colon A_i \times Y_i \to A_{-i}$ for all i such that, for all $a \in A$:

$$Q(\forall i, f_i(a_i, y_i) = a_{-i}|a) \geq 1 - \varepsilon\,.$$

Almost-perfect monitoring refers to ε-perfect monitoring for small values of ε.

Canonical Structure The monitoring structure is *canonical* when each player's observation corresponds to an action profile of the opponents, that is, when $Y_i = A_{-i}$.

Public and Almost Public Signals Signals are *public* when all the players observe the same signal, i. e., $Q(\forall i, j, y_i = y_j | a) = 1$, for every action profile a. For instance, in Green and Porter [33], firms compete over quantities, and the public signal is the realization of the price. Firms can then make inferences on rival's quantities based on their own quantity and market price.

The case in which $Q(\forall i, j, y_i = y_j | a)$ is close to 1 for every a is referred to as *almost public monitoring* (see Mailath and Morris [52]).

Private signals refer to the case when these signals are not public.

Deterministic Signals Signals are *deterministic* when the signal profile is uniquely determined by the action profile. When a is played, the signal profile y is given by $y = f(a)$, where f is called the *signaling function*.

Observable Payoffs Payoffs are *observable* when each player i can deduce his payoff from his action and his signal. This is the case if there exists a mapping $\varphi : A_i \times Y_i \to \mathbb{R}$ such that for every action profile a, $Q(\forall i, g_i(a) = \varphi(a_i, y_i) | a) = 1$.

The Repeated Game The game is played repeatedly and after every stage t, the profile of signals y_t received by the players is drawn according to the distribution $Q(y_t | a_t)$, where a_t is the profile of action chosen at stage t. A player's information consists of his past actions and signals. We let $H_{i,t} = (A_i \times Y_i)^t$ be the set of *player i's histories* of length t. A strategy for player i now consists of a mapping $\sigma_i : \cup_{t \geq 0} H_{i,t} \to S_i$. The set of complete histories of the game after t stages is $H_t = (A \times Y)^t$, it describes chosen actions and received signals for all the players at all past stages. A strategy profile $\sigma = (\sigma_i)_{i \in I}$ induces a probability distribution P_σ on the set of plays $H_\infty = (A \times Y)^\infty$.

Equilibrium Notions

Nash Equilibria Player's preferences over game plays are defined according to the same criteria as for perfect monitoring. We focus on infinitely repeated games, both discounted and undiscounted. A choice of players' preferences defines a set of Nash equilibrium payoffs in the repeated game.

Sequential Equilibria The most commonly used refinement of Nash equilibrium for repeated games with imperfect monitoring is the *sequential equilibrium* concept (Kreps and Wilson, [42]), which we recall here.

A *belief assessment* is a sequence $\mu = (\mu_{i,t})_{t \geq 1, i \in I}$ with $\mu_{i,t} : H_{i,t} \to \Delta(H_t)$, i. e., given the private history h_i of player i, $\mu_{i,t}(h_i)$ is the probability distribution representing the belief that player i holds on the full history.

A sequential equilibrium of the repeated game is a pair (σ, μ) where σ is a strategy profile and μ is a belief assessment such that: 1) for each player i and every history h_i, σ_i is a best reply in the continuation game, given the strategies of the other players and the belief that player i holds regarding the past; 2) the beliefs must be consistent in the sense that (σ, μ) is the limit of a sequence (σ^n, μ^n) where for every n, σ^n is a completely mixed strategy (it assigns positive probability to every action after every history) and μ^n is the unique belief derived from Bayes' law under P_{σ^n}.

Sequential equilibria are defined both on the discounted game and the undiscounted versions of the repeated game.

For undiscounted games, the set of Nash equilibrium payoffs and sequential equilibrium payoffs coincide. The two notions also coincide for discounted games when the monitoring has full support (i. e. under every action profile, all signal profiles have positive probability). The results presented in this survey all hold for sequential equilibria, both for discounted and undiscounted games.

Extensions of the Repeated Game When players receive correlated inputs or may communicate between stages of the repeated game, the relevant concepts are correlated and communication equilibria.

Correlated Equilibria A correlated equilibrium (Aumann [6]) of the repeated game is an equilibrium of an extended game in which: at a preliminary stage, a mediator chooses a profile of correlated random inputs and informs each player of his own input; then the repeated game is played. A characterization of the set of correlated equilibrium payoffs for two-player games is obtained by Lehrer [45].

Correlation arises endogenously in repeated games with imperfect monitoring, as the signals received by the players can serve as correlated inputs that influence player's continuation strategies. This phenomenon is called *internal correlation*, and was studied by Lehrer, [44], Gossner and Tomala, [29,30].

Communication Equilibria An (extensive form) communication equilibrium (Myerson [54], Forges [23]) of

a repeated game is an equilibrium of an extension of the repeated game in which after every stage, players send messages to a mediator, and the mediator sends back private outputs to the players. Characterizations of the set of communication equilibrium payoffs are obtained under weak conditions on the monitoring structure, see e. g. Kandori and Matsushima [38], Compte [15], and Renault and Tomala [61].

Detection and Identification

Equivalent Actions A player's deviation is detectable when it induces a different distribution of signals for other players. When two actions induce the same distribution of actions for other players, they are called equivalent (Lehrer [43,44,45,46]):

Definition 1 Two actions a_i and b_i of player i are equivalent, and we note $a_i \sim b_i$, if they induce the same distribution of other players' signals:

$$Q(y_{-i}|a_i, a_{-i}) = Q(y_{-i}|b_i, a_{-i}) , \quad \forall a_{-i} .$$

Example 1 Consider the two-player repeated Prisoner's Dilemma where player 2 receives no information about the actions of player 1 (e. g. Y_2 is a singleton). The two actions of player 1 are thus equivalent. The actions of player 2 are independent of the actions of player 1: player 1 has no impact on the behavior of player 2. Player 2 has no power to threat player 1 and in any equilibrium, player 1 defects at every stage. Player 2 also defects at every stage: since player 1 always defects, he also loses his threatening power. The only equilibrium payoff in this repeated game is thus $(1, 1)$.

Example 1 suggests that between two equivalent actions, a player chooses at equilibrium the one that yields the highest stage payoff. This is indeed the case when the information received by a player does not depend on his own action. Lehrer [43] studies particular monitoring structures satisfying this requirement. Recall from Lehrer [43] the definition of *semi-standard* monitoring structures: each action set A_i is endowed with a partition $\bar{A}_i$, when player i plays a_i, the corresponding partition cell $\bar{a}_i$ is publicly announced. In the semi-standard case, two actions are equivalent if and only if they belong to the same cell: $a_i \sim b_i \iff \bar{a}_i = \bar{b}_i$ and the information received by a player on other player's action does not depend on his own action.

If player i deviates from a_i to b_i, the deviation is undetected if and only if $a_i \sim b_i$. Otherwise it is detected by all other players. A profile of mixed actions is called *immune to undetectable deviations* if no player can profit by a unilateral deviation that maintains the same distribution of other players' signals. The following result, due to Lehrer [43], characterizes equilibrium payoffs for undiscounted games with semi-standard signals:

Theorem 7 *In an undiscounted repeated game with semi-standard signals, the equilibrium payoffs are the individually rational payoffs that belong to the convex hull of payoffs generated by mixed action profiles that are immune to undetectable deviations.*

More Informative Actions When the information of player i depends on his own action, some deviations may be detected in the course of the repeated game even though they are undetectable in the stage game.

Example 2 Consider the following modification of the Prisoner's Dilemma. The action set of player 1 is $A_1 = \{C_1, D_1\} \times \{C_2, D_2\}$ and the action set of player 2 is $\{C_2, D_2\}$. An action for player 1 is thus a pair $a_1 = (\tilde{a}_1, \tilde{a}_2)$. When the action profile $(\tilde{a}_1, \tilde{a}_2, a_2)$ is played, the payoff to player i is $g_i(\tilde{a}_1, a_2)$. We can interpret the component $\tilde{a}_1$ as a *real* action (it impacts payoffs) and the component $\tilde{a}_2$ as a *message* sent to player 2 (it does not impact payoffs). The monitoring structure is as follows:

- player 2 only observes the message component $\tilde{a}_2$ of the action of player 1 and,
- player 1 perfectly observes the action of player 2 if he chooses the cooperative real action ($\tilde{a}_1 = C_1$), and gets no information on player 2's action if he defects ($\tilde{a}_1 = D_1$).

Note that the actions (C_1, C_2) and (D_1, C_2) of player 1 are equivalent, and so are the actions (C_1, D_2) and (D_1, D_2). However, it is possible to construct an equilibrium that implements the cooperative payoff along the following lines:

i) Using his message component, player 1 reports at every stage $t > 1$ the previous action of player 2. Player 1 is punished in case of a non matching report.
ii) Player 2 randomizes between both actions, so that player 1 needs to play the cooperative action in order to report player 2's action accurately. The weight on the defective action of player 2 goes to 0 as t goes to infinity to ensure efficiency.

Player 2 has incentives to play C_2 most of the time, since player 1 can statistically detect if player 2 uses the action D_2 more frequently than prescribed. Player 1 also has incentives to play the real action C_1, as this is the only way to

observe player 2's action, which needs to be reported later on.

The key point in the example above is that the two real actions C_1 and D_1 of player 1 are equivalent but D_1 is *less informative* than C_1 for player 1. For general monitoring structures an action a_i is *more informative* than an action b_i if: *whenever player i plays a_ii, he can reconstitute the signal he would have observed, had he played b_i*. The precise definition of the *more informative* relation relies on Blackwell's ordering of stochastic experiments [11]:

Definition 2 The action a_i of player i is more informative than the action b_i if there exists a transition probability $f\colon Y_i \to \Delta(Y_i)$ such that for every a_{-i} and every profile of signals y,

$$\sum_{y_i} f(y_i'|y_i)Q(y_i, y_{-i}|a_i, a_{-i}) = Q(y_i', y_{-i}|b_i, a_{-i}) .$$

We denote $a_i \succeq b_i$ if $a_i \sim b_i$ and a_i is more informative than b_i.

Assume that prescribed strategies require player i to play b_i at stage t, and let $a_i \succeq b_i$. Consider the following deviation from player i: play a_i at stage t, and reconstruct a signal at stage t that could have arisen from the play of b_i. In all subsequent stages, play as if no deviation took place at stage t, and as if the reconstructed signal had been observed at stage t. Not only such a deviation would be undetectable at stage t, since $a_i \sim b_i$, but it would also be undetectable at all subsequent stages, as it would induce the same probability distribution over plays as under the prescribed strategy. This argument shows that, if an equilibrium strategy specifies that player i plays a_i, there is no $b_i \succeq a_i$ that yields a higher expected stage payoff than a_i.

Definition 3 A distribution of actions profiles $p \in \Delta(A)$ is *immune to undetectable deviations* if for each player i and pair of actions a_i, b_i such that $b_i \succeq a_i$:

$$\sum_{a_{-i}} p(a_i, a_{-i})g_i(a_i, a_{-i}) \geq \sum_{a_{-i}} p(a_i, a_{-i})g_i(b_i, a_{-i})$$

If p is immune to undetectable deviations, and if player i is supposed to play a_i, any alternative action b_i that yields a greater expected payoff can not be such that $b_i \succeq a_i$.

The following proposition gives a necessary condition on equilibrium payoffs that holds both in the discounted and in the undiscounted cases.

Proposition 1 *Every equilibrium payoff of the repeated game is induced by a distribution that is immune to undetectable deviations.*

The condition of Proposition 1 is tight for some specific classes of games, all of them assuming two players and no discounting.

Following Lehrer [45], signals are *non-trivial* if, for each player i, there exist an action a_i for player i and two actions a_j, b_j for i's opponent such that the signal for player i is different under (a_i, a_j) and (a_i, b_j). Lehrer [45] proves:

Theorem 8 *The set of correlated equilibrium payoffs of the undiscounted game with deterministic and non-trivial signals is the set of individually rational payoffs induced by distributions that are immune to undetectable deviations.*

Lehrer [46] assumes that payoffs are observable, and obtains the following result:

Theorem 9 *In a two-player repeated game with no discounting, non-trivial signals and observable payoffs, the set of equilibrium payoffs is the set of individually rational payoffs induced by distributions that are immune to undetectable deviations.*

Finally, Lehrer [44] shows that in some cases, one may dispense with the correlation device of Theorem 8, as all necessary correlation can be generated endogenously through the signals of the repeated game:

Proposition 2 *In two-player games with non-trivial signals such that either the action profile is publicly announced or a blank signal is publicly announced, the set of equilibrium payoffs coincides with the set of correlated equilibrium payoffs.*

Identification of Deviators A deviation is identifiable when every player can infer the identity of the deviating player from his observations. For instance, in a game with public signals, if separate deviations from players i and j induce the same distribution of public signals, these deviations from i or j are not identifiable. In order to be able to punish the deviating player, it is sometimes necessary to know his identity. Detectability and identifiability are two separate issues, as shown by the following example.

Example 3 Consider the following 3-player game where player 1 chooses the row, player 2 chooses the column and player 3 chooses the matrix.

	W		M		E	
	L	R	L	R	L	R
T	1, 1, 1	4, 4, 0	0, 3, 0	0, 3, 0	3, 0, 0	3, 0, 0
B	4, 4, 0	4, 4, 0	0, 3, 0	0, 3, 0	3, 0, 0	3, 0, 0

Consider the monitoring structure in which actions are not observable and the payoff vector is publicly announced.

The payoff $(1, 1, 1)$ is feasible and individually rational. The associated action profile (T, L, W) is immune to undetectable deviations since any individual deviation from (T, L, W) changes the payoff.

However, $(1, 1, 1)$ is not an equilibrium payoff. The reason is that, player 3, who has the power to punish either player 1 or player 2, cannot punish both players simultaneously: punishing player 1 rewards player 2 and vice-versa. More precisely, whatever weights player 3 puts on the action M and E, the sum of player 1 and player 2's payoffs is greater than 3. Any equilibrium payoff vector $v = (v_1, v_2, v_3)$ must thus satisfy $v_1 + v_2 \geq 3$. In fact, it is possible to prove that the set of equilibrium payoffs of this repeated game is the set of feasible and individually rational payoffs that satisfy this constraint.

Approachability When the deviating player cannot be identified, it may be necessary to punish a group of suspects altogether. The notion of a payoff that is enforceable under group punishments is captured by the definition of approachable payoffs:

Definition 4 A payoff vector v is *approachable* if there exists a strategy profile σ such that, for every player i and unilateral deviation τ_i of player i, the average payoff of player i under (τ_i, σ_{-i}) is asymptotically less than or equal to v_i.

Blackwell's [12] approachability theorem and its generalization by Kohlberg [40] provide simple geometric characterizations of approachable payoffs. It is straightforward that approachability is a necessary condition on equilibrium payoffs:

Proposition 3 *Every equilibrium payoff of the repeated game is approachable.*

Renault and Tomala [61] show that the conditions of Proposition 1 and 3 are tight for communication equilibria:

Theorem 10 *For every game with imperfect monitoring, the set of communication equilibrium payoffs of the repeated game with no discounting is the set of approachable payoffs induced by distributions which are immune to undetectable deviations.*

Tomala [70] shows that pure strategy equilibrium payoffs of undiscounted repeated games with public signals are also characterized through identifiability and approachability conditions (the approachability definition then uses pure strategies). Tomala [71] provides a similar characterization in mixed strategies for a restricted class of public signals.

Identification Through Endogenous Communication A deviation may be identified in the repeated game even though it cannot be identified in the stage game. In a *network game*, players are located at nodes of a graph, and each player monitors his neighbors' actions. Each player can use his actions as messages that are broadcasted to all the neighbors in the graph. The graph is called *2-connected* if no single node deletion disconnects the graph. Renault and Tomala [60] show that when the graph is 2-connected, there exists a communication protocol among the players that ensures that the identity of any deviating player becomes common knowledge among all players in finite time. In this case, identification takes place through communication over the graph.

Public Equilibria

In a seminal paper, Green and Porter [33] introduce a model in which firms are engaged in a production game and publicly observe market prices, which depend both on quantities produced and on non-observable exogenous market shocks. Can collusion be sustained at equilibrium even if prices convey imperfect information on quantities produced? This motivates the study of *public equilibria* for which sharp characterizations of equilibrium payoffs are obtained.

Signals are public when all sets of signals are identical, i. e. $Y_i = Y_{\text{pub}}$ for each i and $Q(\forall i, j, y_i = y_j | a) = 1$ for every a. A *public history* of length t is a record of t public signals, i. e. an element of $H_{\text{pub},t} = (Y_{\text{pub}})^t$. A strategy σ_i for player i is a *public strategy* if it depends on the public history only: if $h_i = (a_{i,1}, y_1, \ldots, a_{i,t}, y_t)$ and $h'_i = (a'_{i,1}, y'_1, \ldots, a'_{i,t}, y'_t)$ are two histories for player i such that $y_1 = y'_1, \ldots, y_t = y'_t$, then $\sigma_i(h_i) = \sigma_i(h'_i)$.

Definition 5 A *perfect public equilibrium* is a profile of public strategies such that after every public history, each player's continuation strategy is a best reply to the opponents' continuation strategy profile.

The repetition of a Nash equilibrium of the stage game is a perfect public equilibrium, so that perfect public equilibria exist. Every perfect public equilibrium is a sequential equilibrium: any consistent belief assigns probability one to the realized public history and thus correctly forecasts future opponents' choices.

A Recursive Structure A perfect public equilibrium (PPE henceforth) is a profile of public strategies that forms

an equilibrium of the repeated game and such that, after every public history, the continuation strategy profile is also a PPE. The set of PPEs and of induced payoffs therefore possesses a recursive structure, as shown by Abreu, Pearce and Stachetti [3]. The argument is based on a dynamic programming principle. To state the main result, we first introduce some definitions.

Given a mapping $f\colon Y_{\text{pub}} \to \mathbb{R}^I$, $G(\delta, f)$ represents the one-shot game where each player i choose actions in A_i and where payoffs are given by:

$$(1-\delta)g_i(a) + \delta \sum_{y \in Y_{\text{pub}}} Q(y|a) f_i(y)\,.$$

In $G(\delta, f)$, the stage game is played, and players receive $f(y)$ as an additional payoff if y is the realized public signal. The weights $1-\delta$ and δ are the relative weights of present payoffs versus all future payoffs in the repeated game.

Definition 6 A payoff vector $v \in \mathbb{R}^I$ is decomposable with respect to the set $W \subset \mathbb{R}^I$ if there exists a mapping $f\colon Y_{\text{pub}} \to W$ such that v is a Nash equilibrium payoff of $G(\delta, f)$. $F_\delta(W)$ denotes the set of payoff vectors which are decomposable with respect to W.

Let $E(\delta)$ be the set of perfect public equilibrium payoffs of the repeated game discounted at the rate δ. The following result is due to Abreu et al. [3]:

Theorem 11 *$E(\delta)$ is the largest bounded set W such that $W \subseteq F_\delta(W)$.*

Fudenberg and Levine [19] derive an asymptotic characterization of the set of PPE payoffs when the discount factor goes to 1 as follows. Given a vector $\lambda \in \mathbb{R}^I$, define the *score* in the direction λ as:

$$k(\lambda) = \sup \langle \lambda, v \rangle$$

where the supremum is taken over the set of payoff vectors v that are Nash equilibrium payoffs of $G(\delta, f)$, where f is any mapping such that,

$$\langle \lambda, v \rangle \geq \langle \lambda, f(y) \rangle\,, \quad \forall y \in Y_{\text{pub}}\,.$$

Scores are independent of the discount factor. The following theorem is due to Fudenberg and Levine [19]:

Theorem 12 *Let C be the set of vectors v such that for every $\lambda \in \mathbb{R}^I$, $\langle \lambda, v \rangle \leq k(\lambda)$. If the interior of C is non-empty, $E(\delta)$ converges to C (for the Hausdorff topology) as δ goes to 1.*

Fudenberg, Levine and Takahashi [22] relax the non-empty interior assumption. They provide an algorithm for computing the affine hull of $\lim_{\delta \to 1} E(\delta)$ and provide a corresponding characterization of the set C with continuation payoffs belonging to this affine hull.

Folk Theorems for Public Equilibria The recursive structure of Theorem 11 and the asymptotic characterization of PPE payoffs given by Theorem 12 are essential tools for finding sufficient conditions under which every feasible and individually rational payoff is an equilibrium payoff, i. e. conditions under which a Folk Theorem holds.

The two conditions under which a Folk Theorem in PPEs holds are a 1) a condition of detectability of deviations and 2) a condition of identifiability of deviating players.

Definition 7 A profile of mixed actions $s = (s_i, s_{-i})$ has individual full rank if for each player i, the probability vectors (in the vector space $\mathbb{R}^{Y_{\text{pub}}}$)

$$\{Q(\cdot|a_i, s_{-i})\colon a_i \in A_i\}$$

are linearly independent.

If s has individual full rank, no player can change the distribution of his actions without affecting the distribution of public signals. Individual full rank is thus a condition on detectability of deviations.

Definition 8 A profile of mixed actions s has pairwise full rank if for every pair of players $i \neq j$, the family of probability vectors

$$\{Q(\cdot|a_i, s_{-i})\colon a_i \in A_i\} \cup \{Q(\cdot|a_j, s_{-j})\colon a_j \in A_j\}$$

has rank $|A_i| + |A_j| - 1$.

Under the condition of pairwise full rank, deviations from two distinct players induce distinct distributions of public signals. Pairwise full rank is therefore a condition of identifiability of deviating players.

Fudenberg et al. [21] prove the following theorem:

Theorem 13 *Assume the set of feasible and individually rational payoff vectors F has non-empty interior. If every pure action profile has individual full rank and if there exists a mixed action profile with pairwise full rank, then every convex and compact subset of the interior of F is a subset of $E(\delta)$ for δ large enough.*

In particular, under the conditions of the theorem, every feasible and individually rational payoff vector is arbitrarily close to a PPE payoff for large enough discount factors. Variations of this result can be found in [21] and [19].

Extensions

The Public Part of a Signal The definition of perfect public equilibria extends to the case in which each player's signals consists of two components: a public component and

a private component. The public components of all players' signals are the same with probability one. A public strategy is then a strategy that depends only on the public components of past signals, and all the analysis carries through.

Public Communication In the public communication extension of the repeated game, players make public announcements between any two stages of the repeated game. The profile of public announcements then forms a public signal, and recursive methods can be successfully applied. The fact that public communication is a powerful instrument to overcome the difficulties arising from private signals was first observed by Matsushima [49,50]. Ben Porath and Kahneman [14], Kandori and Matsushima [38], and Compte [15] prove Folk Theorems in games with private signals and public communication. Kandori [37] shows that in games with public monitoring, public communication allows to relax the conditions for the Folk Theorem of Fudenberg et al. [21].

Private Strategies in Games with Public Monitoring PPE payoffs do not cover the full set of sequential equilibrium payoffs, even when signals are public, as some equilibria may rely on players using *private strategies*, i. e. strategies that depend on past chosen actions and past private signals. See [53] and [39] for examples. In a *minority game*, there is an odd number of players, each player chooses between actions *A* and *B*. Players choosing the least chosen (minority) action get a payoffs of 1, other players get 0. The public signal is the minority action. Renault et al. [58,59] show that, for minority games, a Folk Theorem holds in private strategies but not in public strategies. Only few results are known concerning the set of sequential equilibrium payoffs in privates strategies of games with public monitoring. A monotonicity property is obtained by Kandori [36] who shows that the set of payoffs associated to sequential equilibria in pure strategies is increasing with respect to the quality of the public signal.

Almost Public Monitoring Some PPEs are robust to small perturbations of public signals. Considering strategies with finite memory, Mailath and Morris [52] identify a class of public strategies which are sequential equilibria of the repeated game with imperfect private monitoring, provided that the monitoring structure is close enough to a public one. They derive a Folk Theorem for games with almost public and almost perfect monitoring. Hörner and Olszewski [35] strengthen this result and prove a Folk Theorem for games with almost public monitoring. Under detectability and identifiability conditions, they prove that feasible and individually rational payoffs can be achieved by sequential equilibria with finite memory.

Almost Perfect Monitoring

Monitoring is almost perfect when each player can identify the action profile of his opponents with near certainty. Almost perfect monitoring is the natural framework to study the robustness of the Folk Theorem to small departures from the assumption that actions are perfectly observed.

The first results were obtained for the Prisoner's Dilemma. Sekiguchi [65] shows that the cooperative outcome can be approximated at equilibrium when players are sufficiently patient and monitoring is almost perfect. Under the same assumptions, Bhaskar and Obara [13], Piccione [57] and Ely and Valimaki [16] show that a Folk Theorem obtains.

Piccione [57] and Ely and Valimaki [16] study a particular class of equilibria called *belief free*. Strategies form a *belief free* equilibrium if, whatever player *i*'s belief on the opponent's private history, the action prescribed by *i*'s strategy is a best response to the opponent's continuation strategy.

Ely, Hörner and Olszewski [17] extend the belief free approach to general games. However, they show that, in general, belief free strategies are not enough to reconstruct a Folk Theorem, even when monitoring is almost perfect.

For general games and with any number of players, Hörner and Olszewski [34] prove a Folk Theorem with almost perfect monitoring. The strategies that implement the equilibrium payoffs are defined on successive blocks of a fixed length, and are block-belief-free in the sense that, at the beginning of every block, every player is indifferent between several continuation strategies, independently on his belief as to which continuation strategies are used by the opponents. This result closes the almost perfect monitoring case by showing that equilibrium payoffs in the Folk Theorem are robust to a small amount of imperfect monitoring.

General Stochastic Signals

Besides the case of public (or almost public) monitoring, little is known about equilibrium payoffs of repeated games with discounting and imperfect signals.

The Prisoner's Dilemma game is particularly important for economic applications. In particular, it captures the essential features of collusion with the possibility of secret price cutting, as in Stigler [69].

When signals are imperfect, but independent conditionally on the pair of actions chosen (a condition called conditional independence), Matsushima [51] shows that

the efficient outcome of the repeated Prisoner's Dilemma game is an equilibrium outcome if players are sufficiently patient. In the equilibrium construction, every player's action is constant in every block. The conditional independence assumption is crucial in that it implies that, during every block, a player has no feedback as to what signals the other player has received. The conditional independence assumption is non-generic: it holds for a set of monitoring structures of empty interior.

Fong, Gossner, Hörner, and Sannikov [18] prove that efficiency can be obtained at equilibrium without conditional independence. Their main assumption is that there exists a sufficiently informative signal, but this signal needs not be almost perfectly informative. Their result holds for a family of monitoring structures of non empty interior. It is the first result that establishes cooperation in the Prisoner's Dilemma with impatient players for truly imperfect, private and correlated signals.

Punishment Levels

Individual rationality is a key concept for Folk Theorems and equilibrium payoff characterizations. Given a repeated game, define the individually rational (IR) level of player i as the lowest payoff down to which this player may be punished in the repeated game.

Definition 9 The individual rational level of player i is:

$$\lim_{\delta\to 1}\min_{\sigma_{-i}}\max_{\sigma_i}\mathbf{E}_{\sigma_i,\sigma_{-i}}\left[\sum_t(1-\delta)\delta^{t-1}g_{i,t}\right]$$

where the min runs over profiles of behavior strategies for player $-i$, and the max over behavior strategies of player i.

That is, the individually rational level is the limit (as the discount factor goes to one) of the min max value of the discounted game (other approaches, through undiscounted games or limits of finitely repeated games, yield equivalent definitions, see [29]).

Comparison of the IR Level with the min max With perfect monitoring, the IR level of player i is player i's min max in the one-shot game, as defined by equation (1). With imperfect monitoring, the IR level for player i is never larger than v_i since player i's opponents can force down player i to v_{-i} by repeatedly playing the min max strategy against player i.

With two players, it is a consequence of von-Neumann's min max theorem [72] that v_i is the IR level for player i.

For any number of players, Gossner and Hörner [28] show that v_i is equal to the min max in the one-shot game whenever there exists a garbling of player i's signal such that, conditionally on i's garbled signal, the signals of i's opponents are independent. Furthermore, the condition in [28] is also a necessary condition in normal form games extended by correlation devices (as in Aumann [6]). A continuity result in the IR level also applies for monitoring structure close to those that satisfy the conditional independence condition.

The following example shows that, in general, the IR level can be lower than v_i:

Example 4 Consider the following three-player game. Player 1 chooses the row, player 2 the column and player 3 the matrix. Players 1 and 2 perfectly observe the action profile while player 3 observes player 2's action only. As we deal with the IR level of player 3, we specify the payoff for this player only.

	W		E	
	L	R	L	R
T	0	0	−1	0
B	0	−1	0	0

A simple computation shows that $v_3 = -\frac{1}{4}$ and that the min max strategies of players 1 and 2 are uniform. Consider the following strategies of players 1 and 2 in the repeated game: randomize uniformly at odd stages, play (T, L) or (B, R) depending on player 1's previous action at even stages. Against these strategies, player 3 cannot obtain better than $-\frac{1}{4}$ at odd stages and $-\frac{1}{2}$ at even stages, resulting in an average payoff of $-\frac{3}{8}$.

Entropy Characterizations The exact computation of the IR level in games with imperfect monitoring requires to analyze the optimal trade-off for punishing players between the production of correlated and private signals and the use of these signals for effective punishment. Gossner and Vieille, [31] and Gossner and Tomala [29] develop tools based on information theory to analyze this trade-off. At any stage, the amount of correlation generated (or spent) by the punishing players is measured using the entropy function. Gossner and Tomala [30] derive a characterization of the IR level for some classes of monitoring structures. Gossner, Laraki, and Tomala [32] provide methods explicit computations of the IR level. In particular, for the above example, the IR level computed and is about $-.401$. Explicit computations of IR levels for other games are derived by Goldberg [26].

Acknowledgments

The authors are grateful to Johannes Hörner for insightful comments.

Bibliography

Primary Literature

1. Abreu D (1988) On the theory of infinitely repeated games with discounting. Econometrica 56:383–396
2. Abreu D, Dutta P, Smith L (1994) The folk theorem for repeated games: a NEU condition. Econometrica 62:939–948
3. Abreu D, Pearce D, Stacchetti E (1990) Toward a theory of discounted repeated games with imperfect monitoring. Econometrica 58:1041–1063
4. Aumann RJ (1960) Acceptable points in games of perfect information. Pac J Math 10:381–417
5. Aumann RJ (1964) Mixed and behavior strategies in infinite extensive games. In: Dresder M, Shapley LS, Tucker AW (eds) Advances in Game Theory. Princeton University Press, New Jersey, pp 627–650
6. Aumann RJ (1974) Subjectivity and correlation in randomized strategies. J Math Econ 1:67–95
7. Aumann RJ (1981) Survey of repeated games. In: Aumann RJ (ed) Essays in game theory and mathematical economics in honor of Oskar Morgenstern. Wissenschaftsverlag, Bibliographisches Institut, Mannheim, pp 11–42
8. Aumann RJ, Shapley LS (1976) Long-term competition – a game theoretic analysis. Re-edited in 1994. See [9]
9. Aumann RJ, Shapley LS (1994) Long-term competition – a game theoretic analysis. In: Megiddo N (ed) Essays on game theory. Springer, New York, pp 1–15
10. Benoît JP, Krishna V (1985) Finitely repeated games. Econometrica 53(4):905–922
11. Blackwell D (1951) Comparison of experiments. In: Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, Berkeley, pp 93–102
12. Blackwell D (1956) An analog of the minimax theorem for vector payoffs. Pac J Math 6:1–8
13. Bhaskar V, Obara I (2002) Belief-based equilibria in the repeated prisoners' dilemma with private monitoring. J Econ Theory 102:40–70
14. Ben EP, Kahneman M (1996) Communication in repeated games with private monitoring. J Econ Theory 70(2):281–297
15. Compte O (1998) Communication in repeated games with imperfect private monitoring. Econometrica 66:597–626
16. Ely JC, Välimäki J (2002) A robust folk theorem for the prisoner's dilemma. J Econ Theory 102:84–106
17. Ely JC, Hörner J, Olszewski W (2005) Belief-free equilibria in repeated games. Econometrica 73:377–415
18. Fong K, Gossner O, Hörner J, Sannikov Y (2007) Efficiency in a repeated prisoner's dilemma with imperfect private monitoring. mimeo
19. Fudenberg D, Levine DK (1994) Efficiency and observability with long-run and short-run players. J Econ Theory 62:103–135
20. Fudenberg D, Maskin E (1986) The folk theorem in repeated games with discounting or with incomplete information. Econometrica 54:533–554
21. Fudenberg D, Levine DK, Maskin E (1994) The folk theorem with imperfect public information. Econometrica 62(5):997–1039
22. Fudenberg D, Levine DK, Takahashi S (2007) Perfect public equilibrium when players are patient. Games Econ Behav 61:27–49
23. Forges F (1986). An approach to communication equilibria. Econometrica 54:1375–1385
24. Forges F, Mertens J-F, Neyman A (1986) A counterexample to the folk theorem with discounting. Econ Lett 20:7–7
25. Friedman J (1971) A noncooperative equilibrium for supergames. Rev Econ Stud 38:1–12
26. Goldberg Y (2007) Secret correlation in repeated games with imperfect monitoring: The need for nonstationary strategies. Math Oper Res 32:425–435
27. Gossner O (1995) The folk theorem for finitely repeated games with mixed strategies. Int J Game Theory 24:95–107
28. Gossner O, Hörner J (2006) When is the individually rational payoff in a repeated game equal to the minmax payoff? DP 1440, CMS-EMS
29. Gossner O, Tomala T (2006) Empirical distributions of beliefs under imperfect observation. Math Oper Res 31(1):13–30
30. Gossner O, Tomala T (2007) Secret correlation in repeated games with signals. Math Oper Res 32:413–424
31. Gossner O, Vieille N (2002) How to play with a biased coin? Games Econ Behav 41:206–226
32. Gossner O, Laraki R, Tomala T (2009) Informationally optimal correlation. Math Programm B 116:147–112
33. Green EJ, Porter RH (1984) Noncooperative collusion under imperfect price information. Econometrica 52:87–100
34. Hörner J, Olszewski W (2006) The folk theorem with private almost-perfect monitoring. Econometrica 74(6):1499–1544
35. Hörner J, Olszewski W (2007) How robust is the folk theorem with imperfect public monitoring? mimeo
36. Kandori M (1992) The use of information in repeated games with imperfect monitoring. Rev Econ Stud 59:581–593
37. Kandori M (2003) Randomization, communication, and efficiency in repeated games with imperfect public monitoring. Econometrica 71:345–353
38. Kandori M, Matsushima H (1998) Private observation, communication and collusion. Rev Econ Stud 66:627–652
39. Kandori M, Obara I (2006) Efficiency in repeated games revisited: The role of private strategies. Econometrica 74:499–519
40. Kohlberg E (1975) Optimal strategies in repeated games with incomplete information. Int J Game Theory 4:7–24
41. Kuhn HW (1953) Extensive games and the problem of information. In: Kuhn HW, Tucker AW (eds) Contributions to the Theory of Games, vol II. Annals of Mathematical Studies, vol 28. Princeton University Press, New Jersey, pp 193–216
42. Kreps DM, Wilson RB (1982) Sequential equilibria. Econometrica 50:863–894
43. Lehrer E (1990) Nash equilibria of *n*-player repeated games with semi-standard information. Int J Game Theory 19:191–217
44. Lehrer E (1991) Internal correlation in repeated games. Int J Game Theory 19:431–456
45. Lehrer E (1992) Correlated equilibria in two-player repeated games with nonobservable actions. Math Oper Res 17:175–199
46. Lehrer E (1992) Two players repeated games with non observable actions and observable payoffs. Math Oper Res 17:200–224

47. Lehrer E, Pauzner A (1999) Repeated games with differential time preferences. Econometrica 67:393–412
48. Luce RD, Raiffa H (1957) Games and Decisions: Introduction and Critical Survey. Wiley, New York
49. Matsushima H (1991) On the theory of repeated games with private information: Part i: Anti-folk theorem without communication. Econ Lett 35:253–256
50. Matsushima H (1991) On the theory of repeated games with private information: Part ii: Revelation through communication. Econ Lett 35:257–261
51. Matsushima H (2004) Repeated games with private monitoring: Two players. Econometrica 72:823–852
52. Mailath G, Morris S (2002) Repeated games with almost-public monitoring. J Econ Theory 102:189–229
53. Mailath GJ, Matthews SA, Sekiguchi T (2002) Private strategies in finitely repeated games with imperfect public monitoring. B E J Theor Econ 2
54. Myerson RB (1982) Optimal coordination mechanisms in generalized principal-agent problems. J Math Econ 10:67–81
55. Nash JF (1951) Noncooperative games. Ann Math 54:289–295
56. Neyman A (1999) Cooperation in repeated games when the number of stages is not commonly known. Econometrica 67:45–64
57. Piccione M (2002) The repeated prisoner's dilemma with imperfect private monitoring. J Econ Theory 102:70–84
58. Renault J, Scarlatti S, Scarsini M (2005) A folk theorem for minority games. Games Econ Behav 53:208–230
59. Renault J, Scarlatti S, Scarsini M (2008) Discounted and finitely repeated minority games with public signals. Math Soc Sci 56:44–74
60. Renault J, Tomala T (1998) Repeated proximity games. Int J Game Theory 27:539–559
61. Renault J, Tomala T (2004) Communication equilibrium payoffs of repeated games with imperfect monitoring. Games Econ Behav 49:313–344
62. Rubinstein A (1977) Equilibrium in supergames. Center for Research in Mathematical Economics and Game Theory. Res Memo 25
63. Rubinstein A (1979) Equilibrium in supergames with the overtaking criterion. J Econ Theory 21:1–9
64. Rubinstein A (1994) Equilibrium in supergames. In: Megiddo N (ed) Essays on game theory. Springer, New York, pp 17–28
65. Sekiguchi T (1997) Efficiency in repeated prisoner's dilemma with private monitoring. J Econ Theory 76:345–361
66. Selten R (1965) Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit. Z Gesamte Staatswiss 12:201–324
67. Smith L (1995) Necessary and sufficient conditions for the perfect finite horizon folk theorem. Econometrica 63:425–430
68. Sorin S (1986) On repeated games with complete information. Math Oper Res 11:147–160
69. Stigler G (1964) A theory of oligopoly. J Political Econ 72:44–61
70. Tomala T (1998) Pure equilibria of repeated games with public observation. Int J Game Theory 27:93–109
71. Tomala T (1999) Nash equilibria of repeated games with observable payoff vector. Games Econ Behav 28:310–324
72. von Neumann J (1928) Zur Theorie der Gesellschaftsspiele. Math Ann 100:295–320
73. Wen Q (1994) The "folk theorem" for repeated games with complete information. Econometrica 62:949–954

Books and Reviews

Aumann RJ (1981) Survey of repeated games. In: Aumann RJ (ed) Essays in game theory and mathematical economics in honor of Oskar Morgenstern. Wissenschaftsverlag, Bibliographisches Institut, Mannheim, pp 11–42
Mailath GJ, Samuelson L (2006) Repeated Games and Reputations: Long-Run Relationships. Oxford University Press, Oxford
Mertens J-F (1986) Repeated games. In: Proceedings of the international congress of Mathematicians. Berkeley, California, pp 1528–1577
Mertens J-F, Sorin S, Zamir S (1994) Repeated games. CORE discussion paper. pp 9420–9422, Univeritsé Catholique de Levain, Louvain-la-neuve

Repeated Games with Incomplete Information

JÉRÔME RENAULT
Ceremade, Université Paris Dauphine, Paris, France

Article Outline

Glossary

Repeated game with incomplete information A situation where several players repeat the same stage game, the players having different knowledge of the stage game which is repeated.

Strategy of a player A rule, or program, describing the action taken by the player in any possible situation which may happen, depending on the information available to this player in that situation.

Strategy profile A vector containing a strategy for each player.

Lack of information on one side Particular case where all the players but one perfectly know the stage game which is repeated.

Zero-sum games 2-player games where the players have opposite payoffs.

Value Solution (or price) of a zero-sum game, in the sense of the fair amount that player 1 should give to player 2 to be entitled to play the game.

Equilibrium Strategy profile where each player's strategy is in best reply against the strategy of the other players.

Completely revealing strategy Strategy of a player which eventually reveals to the other players everything known by this player on the selected state.

Non revealing strategy Strategy of a player which reveals nothing on the selected state.

The simplex of probabilities over a finite set For a finite set S, we denote by $\Delta(S)$ the set of probabilities over S, and we identify $\Delta(S)$ to $\{p = (p_s)_{s\in S} \in \mathbb{R}^S, \forall s \in S\ p_s \geq 0 \text{ and } \sum_{s\in S} p_s = 1\}$. Given s in S, the Dirac measure on s will be denoted by δ_s. For $p = (p_s)_{s\in S}$ and $q = (q_s)_{s\in S}$ in $\mathbb{R}^S$, we will use, unless otherwise specified, $\|p - q\| = \sum_{s\in S} |p_s - q_s|$.

Definition of the Subject

Introduction

In a repeated game with incomplete information, there is a basic interaction called stage game which is repeated over and over by several participants called players. The point is that the players do not perfectly know the stage game which is repeated, but rather have different knowledge about it. As illustrative examples, one may think of the following situations: an oligopolistic competition where firms don't know the production costs of their opponents, a financial market where traders bargain over units of an asset which terminal value is imperfectly known, a cryptographic model where some participants want to transmit some information (e. g., a credit card number) without being understood by other participants, a conflict when a particular side may be able to understand the communications inside the opponent side (or might have a particular type of weapons), ...

Natural questions arising in this context are as follows. What is the optimal behavior of a player with a perfect knowledge of the stage game? Can we determine which part of the information such a player should use? Can we price the value of possessing a particular information? How should one player behave while having only a partial information?

Foundations of games with incomplete information have been studied in [28] and [56]. Repeated games with incomplete information have been introduced in the sixties by Aumann and Maschler [6], and we present here the basic and fundamental results of the domain. Let us start with a few well known elementary examples [6,91].

Basic Examples In each example, there are two players, and the game is zero-sum, i. e. player 2's payoff always is the opposite of player 1's payoff. There are two states a and b, and the possible stage games are given by two real matrices G^a and G^b with identical size. Initially a true state of nature $k \in \{a, b\}$ is selected with even probability between a and b, and k is announced *to player 1 only*. Then the matrix game G^k is repeated over and over: at every stage, simultaneously player 1 chooses a row i, whereas player 2 chooses a column j, the stage payoff for player 1 is then $G^k(i, j)$ but only i and j are publicly announced before proceeding to the next stage. Players are patient and want to maximize their long-run average expected payoffs.

Example 1

$$G^a = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad G^b = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}.$$

This example is very simple. In order to maximize his payoff, player 1 just has to play, at any stage, the *Top* row if the state is a and the *Bottom* row if the state is b. This corresponds to playing optimally in each possible matrix game.

Example 2

$$G^a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad G^b = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

A naive strategy for player 1 would be to play at stage 1: *Top* if the state is a, and *Bottom* if the state is b. Such a strategy is called completely revealing, or CR, because it allows player 2 to deduce the selected state from the observation of the actions played by player 1. This strategy of player 1 would be optimal here if a single stage was to be played, but it is a very weak strategy on the long run and does not guarantee more than zero at each stage $t \geq 2$ (because player 2 can play *Left* or *Right* depending on player 1's first action).

On the opposite, player 1 may not use his information and play a non revealing, or NR, strategy, i. e. a strategy which is independent of the selected state. He can consider the average matrix

$$\frac{1}{2}G^a + \frac{1}{2}G^b = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix},$$

and play independently at each stage an optimal mixed action in this matrix, i. e. here the unique mixed action $\frac{1}{2}Top + \frac{1}{2}Bottom$. It will turn out that this is here the optimal behavior for player 1, and the value of the repeated game is the value of the average matrix, i. e. 1/4.

Example 3

$$G^a = \begin{pmatrix} 4 & 0 & 2 \\ 4 & 0 & -2 \end{pmatrix} \quad \text{and} \quad G^b = \begin{pmatrix} 0 & 4 & -2 \\ 0 & 4 & 2 \end{pmatrix}.$$

Playing a CR strategy for player 1 does not guarantee more than zero in the long-run, because player 2 will eventually be able to play *Middle* if the state is a, and *Left* if the state is b. But a NR strategy will not do better, because the average matrix

$$\frac{1}{2}G^a + \frac{1}{2}G^b \text{ is } \begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \end{pmatrix},$$

hence has value 0.

We will see later that an optimal strategy for player 1 in this game is to play as follows. Initially, player 1 chooses an element s in $\{T, B\}$ as follows: if $k = a$, then $s = T$ with probability 3/4, and thus $s = B$ with probability 1/4; and if $k = b$, then $s = T$ with probability 1/4, and $s = B$ with probability 3/4. Then at each stage player 1 plays row s, independently of the actions taken by player 2. The conditional probabilities satisfy: $P(k = a|s = T) = 3/4$, and $P(k = a|s = B) = 1/4$. At the end of stage 1, player 2 will have learned, from the action played by his opponent, something about the selected state: his belief on the state will move from $\frac{1}{2}a + \frac{1}{2}b$ to $\frac{3}{4}a + \frac{1}{4}b$ or to $\frac{1}{4}a + \frac{3}{4}b$. But player 2 still does not know perfectly the selected state. Such a strategy of player 1 is called *partially revealing*.

General Definition

Formally, a repeated game with incomplete information is given by the following data. There is a set of players N, and a set of states K. Each player i in N has a set of actions A^i and a set of signals U^i, and we denote by $A = \prod_{i\in N} A^i$ the set of action profiles and by $U = \prod_{i\in N} U^i$ the set of signal profiles. Every player i has a payoff function $g^i : K\times A \longrightarrow \mathbb{R}$. There is a signaling function $q: K \times A \longrightarrow \Delta(U)$, and an initial probability $\pi \in \Delta(K \times U)$. In what follows, we will always assume the sets of players, states, actions and signals to be non empty and finite.

A repeated game with incomplete information can thus be denoted by $\Gamma = (N, K, (A^i)_{i\in N}, (U^i)_{i\in N}, (g^i)_{i\in N}, q, \pi)$. The progress of the game is the following.

- Initially, an element $(k, (u_0^i)_i)$ is selected according to π: k is the realized state of nature and will remain fixed, and each player i learns u_0^i (and nothing more than u_0^i).
- At each integer stage $t \geq 1$, simultaneously every player i chooses an action a_t^i in A^i, and we denote by $a_t = (a_t^i)_i$ the action profile played at stage t. The stage payoff of a player i is then given by $g^i(k, a_t)$. A signal profile $(u_i^t)_i$ is selected according to $q(k, a_t)$, and each player i learns u_t^i (and nothing more than u_i^t) before proceeding to the next stage.

Remark

1. We will always assume that during the play, each player remembers the past actions he has chosen, as well as the past signals he has received. And players will be allowed to select their action independently at random.
2. The players do not necessarily know their stage payoff after each stage (as an illustration, imagine the players bargaining over units of an asset which terminal value will only be known "at the end" of the game). This is without loss of generality, because it is possible to add hypotheses on q so that each player will be able to deduce his stage payoff from his realized stage signal.
3. Repeated games with complete information are a particular case, corresponding to the situation where each initial signal u_0^i reveals the selected state. Such games are studied in the chapter "Repeated games with complete information".
4. Games where the state variable k evolve from stage to stage, according to the actions played, are called stochastic games. These games are not covered here, but in a specific chapter entitled "Stochastic games".
5. The most standard case of signaling function is when each player exactly learns, at the end of each stage t, the whole action profile a_t. Such games are usually called games with "perfect monitoring", "full monitoring", "perfect observation" or with "observable actions".

Strategies, Payoffs, Value and Equilibria

Strategies

A (behavior) strategy for player i is a rule, or program, describing the action taken by this player in any possible case which may happen. These actions may be chosen at random, so a strategy for player i is an element $\sigma^i = (\sigma_t^i)_{t\geq 1}$, where for each t, σ_t^i is a mapping from $U^i \times (U^i \times A^i)^{t-1}$ to $\Delta(A^i)$ giving the lottery played by player i at stage t as a function of the past signals and actions of player i. The set of strategies for player i is denoted by Σ^i.

A history of length t in Γ is a sequence $(k, u_0, a_1, u_1, \ldots, a_t, u_t)$, and the set of such histories is the finite set $K \times U \times (A \times U)^t$. An infinite history is called a play, the set of plays is denoted by $\Omega = K \times U \times (A \times U)^\infty$ and is endowed with the product σ-algebra. A strategy profile $\sigma = (\sigma^i)_i$ naturally induces, together with the initial probability π, a probability distribution over the set of histories of length t. This probability uniquely extends to a probability over plays, and is denoted by $\mathbb{P}_{\pi,\sigma}$.

Payoffs

Given a time horizon T, the average expected payoff of player i, up to stage T, if the strategy profile σ is played

is denoted by:

$$\gamma_T^i(\sigma) = \mathbb{E}_{\mathbb{P}_{\pi,\sigma}}\left(\frac{1}{T}\sum_{t=1}^{T} g^i(k,a_t)\right).$$

The T-stage game is the game Γ_T where simultaneously, each player i chooses a strategy σ^i in Σ^i, then receives the payoff $\gamma_T^i((\sigma^j)_{j\in N})$.

Given a discount factor λ in $(0,1]$, the λ-discounted payoff of player i is denoted by:

$$\gamma_\lambda^i(\sigma) = \mathbb{E}_{\mathbb{P}_{\pi,\sigma}}\left(\lambda\sum_{t=1}^{\infty}(1-\lambda)^{t-1} g^i(k,a_t)\right).$$

The λ-discounted game is the game Γ_λ where simultaneously, each player i chooses a strategy σ^i in Σ^i, then receives the payoff $\gamma_\lambda^i((\sigma^j)_{j\in N})$.

Remark A strategy for player i is called *pure* if it always plays in a deterministic way. A *mixed strategy* for player i is defined as a probability distribution over the set of pure strategies (endowed with the product σ-algebra). Kuhn's theorem (see [3,38] or [85] for a modern presentation) states that mixed strategies or behavior strategies are equivalent, in the following sense: for each behavior strategy σ^i, there exists a mixed strategy τ^i of the same player such that $\mathbb{P}_{\pi,\sigma^i,\sigma^{-i}} = \mathbb{P}_{\pi,\tau^i,\sigma^{-i}}$ for any strategy profile σ^{-i} of the other players, and *vice-versa* if we exchange the words "behavior" and "mixed". Unless otherwise specified, the word strategy will refer here to a behavior strategy, but we will also sometimes equivalently use mixed strategies, or even mixtures of behavior strategies.

Value of Zero-Sum Games

By definition the game is zero-sum if there are two players, say player 1 and player 2, with opposite payoffs. The T-stage game Γ_T can then be seen as a matrix game, hence by the minmax theorem it has a value $v_T = \sup_{\sigma^1}\inf_{\sigma^2}\gamma_T^1(\sigma^1,\sigma^2) = \inf_{\sigma^2}\sup_{\sigma^1}\gamma_T^1(\sigma^1,\sigma^2)$. Similarly, one can use Sion's theorem [74] to show that the λ-discounted game has a value $v_\lambda = \sup_{\sigma^1}\inf_{\sigma^2}\gamma_\lambda^1(\sigma^1,\sigma^2) = \inf_{\sigma^2}\sup_{\sigma^1}\gamma_\lambda^1(\sigma^1,\sigma^2)$.

To study long term strategic aspects, it is also important to consider the following notion of uniform value. Players are asked to play well uniformly in the time horizon, i. e. simultaneously in all game Γ_T with T sufficiently large (or similarly uniformly in the discount factor, i. e. simultaneously in all game Γ_λ with λ sufficiently low).

Definition 1 Player 1 can guarantee the real number v in the repeated game Γ if: $\forall\varepsilon>0, \exists\sigma^1\in\Sigma^1, \exists T_0, \forall T\geq T_0, \forall\sigma^2\in\Sigma^2, \gamma_T^1(\sigma^1,\sigma^2)\geq v-\varepsilon$. Similarly, Player 2 can guarantee v in Γ if $\forall\varepsilon>0, \exists\sigma^2\in\Sigma^2, \exists T_0, \forall T\geq T_0, \forall\sigma^1\in\Sigma^1, \gamma_T^1(\sigma^1,\sigma^2)\leq v+\varepsilon$. If both player 1 and player 2 can guarantee v, then v is called the uniform value of the repeated game. A strategy σ^1 of player 1 satisfying $\exists T_0, \forall T\geq T_0, \forall\sigma^2\in\Sigma^2, \gamma_T^1(\sigma^1,\sigma^2)\geq v$ is then called an optimal strategy of player 1 (optimal strategies of player 2 are defined similarly).

The uniform value, whenever it exists, is necessarily unique. Its existence is a strong property, which implies that both v_T, as T goes to infinity, and v_λ, as λ goes to zero, converge to the uniform value.

Equilibria of General-Sum Games

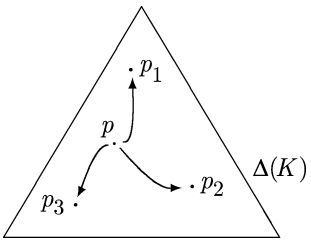
In the general case, the T-stage game Γ_T can be seen as the mixed extension of a finite game, and consequently possesses a Nash equilibrium. Similarly, the discounted game Γ_λ always has, by the Nash Glicksberg theorem, a Nash equilibrium. Concerning uniform notions, couples of optimal strategies are generalized as follows.

Definition 2 A strategy profile $\sigma=(\sigma^i)_{i\in N}$ is a uniform Nash equilibrium of Γ if: 1) $\forall\varepsilon>0, \sigma$ is an ε-Nash equilibrium in every finitely repeated game sufficiently long, that is: $\exists T_0, \forall T\geq T_0, \forall i\in N, \forall\tau^i\in\Sigma^i, \gamma_T^i(\tau^i,\sigma^{-i})\leq\gamma_T^i(\sigma)+\varepsilon$, and 2) the sequence of payoffs $\left((\gamma_T^i(\sigma))_{i\in N}\right)_T$ converges to a limit payoff $(\gamma^i(\sigma))_{i\in N}$ in $\mathbb{R}^N$.

Remark The initial probability π will play a great role in the following analyzes, so we will often write $\gamma_T^{i,\pi}(\sigma)$ for $\gamma_T^i(\sigma)$, $v_T(\pi)$ for the value v_T, etc.

The Standard Model of Aumann and Maschler

This famous model has been introduced in the sixties by Aumann and Maschler (see the reedition [6]). It deals with zero-sum games with lack of information on one side and observable actions, as in the basic examples previously presented. There is a finite set of states K, an initial probability $p=(p^k)_{k\in K}$ on K, and a family of matrix games G^k with identical size $I\times J$. Initially, a state k in K is selected according to p, and announced to player 1 (called the informed player) only. Then the matrix game G^k is repeated over and over: at every stage, simultaneously player 1 chooses a row i in I, whereas player 2 chooses a column j in J, the stage payoff for player 1 is then $G^k(i,j)$ but only i and j are publicly announced before proceeding to the next stage. Denote by M the constant $\max_{k,i,j}|G^k(i,j)|$.



Repeated Games with Incomplete Information, Figure 1
Splitting

Basic Tools: Splitting, Martingale, Concavification, and the Recursive Formula

The following aspects are simple but fundamental. The initial probability $p = (p^k)_{k\in K}$ represents the initial belief, or a priori, of player 2 on the selected state of nature. Assume that player 1 chooses his first action (or more generally a message or signal s from a finite set S) according to a probability distribution depending on the state, i. e. according to a transition probability $x = (x^k)_{k\in K} \in \Delta(S)^K$. For each signal s, the probability that s is chosen is denoted $\lambda(x,s) = \sum_k p^k x^k(s)$, and given s such that $\lambda(x,s) > 0$ the conditional probability on K, or a posteriori of player 2, is $\hat{p}(x,s) = \left((p^k x^k(s))/(\lambda(x,s))\right)_{k\in K}$. We clearly have:

$$p = \sum_{s\in S} \lambda(x,s)\hat{p}(x,s)\,. \tag{1}$$

So the a priori p lies in the convex hull of the a posteriori. The following lemma expresses a reciprocal: player 1 is able to induce any family of a posteriori containing p in its convex hull.

Lemma 1 (Splitting) *Assume that p is written as a convex combination $p = \sum_{s\in S} \lambda_s p_s$ with positive coefficients. There exists a transition probability $x \in \Delta(S)^K$ such that $\forall s \in S, \lambda_s = \lambda(x,s)$ and $p_s = \hat{p}(x,s)$.*

Proof Just put $x^k(s) = \frac{\lambda_s p_s^k}{p^k}$ if $p^k > 0$. □

Equation (1) not only tells that the a posteriori contains p in their convex hull, but also that the expectation of the a posteriori is the a priori. We are here in a repeated context, and for every strategy profile σ one can define the process $(p_t(\sigma))_{t\geq 0}$ of the a posteriori of player 2. We have $p_0 = p$, and $p_t(\sigma)$ is the random variable of player 2's belief on the state after the first t stages. More precisely, we define for any $t \geq 0$, $h_t = (i_1, j_1, \ldots, i_t, j_t) \in (I\times J)^t$ and k in K:

$$p_t^k(\sigma, h_t) = \mathbb{P}_{p,\sigma}(k|h_t) = \frac{p^k \mathbb{P}_{\delta^k,\sigma}(h_t)}{\mathbb{P}_{p,\sigma}(h_t)}\,.$$

$p_t(\sigma, h_t) = (p_t^k(\sigma, h_t))_{k\in K} \in \Delta(K)$ (arbitrarily defined if $\mathbb{P}_{p,\sigma}(h_t) = 0$) is the conditional probability on the state of nature given that σ is played and h_t has occurred in the first t stages. It is easy to see that as soon as $\mathbb{P}_{p,\sigma}(h_t) > 0$, $p_t(\sigma, h_t)$ does not depend on player 2's strategy σ^2, nor on player 2's last action j_t. It is fundamental to notice that:

Lemma 2 (Martingale of a posteriori) *$(p_t(\sigma))_{t\geq 0}$ is a $\mathbb{P}_{p,\sigma}$-martingale with values in $\Delta(K)$.*

This is indeed a general property of Bayesian learning of a fixed unknown parameter: the expectation of what I will know tomorrow is what I know today. This martingale is controlled by the informed player, and the splitting lemma shows that this player can essentially induce any martingale issued from the a priori p. Notice that, to be able to compute the realizations of the martingale, player 2 needs to know the strategy σ^1 used by player 1.

The splitting lemma also easily gives the following concavification result. Let f be a continuous mapping from $\Delta(K)$ to $\mathbb{R}$. The smallest concave function above f is denoted by $\mathrm{cav} f$, and we have:

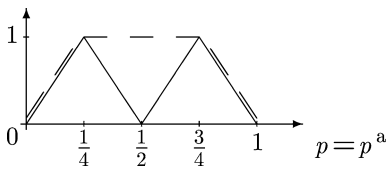
$$\mathrm{cav} f(p) = \max\left\{ \sum_{s\in S} \lambda_s f(p_s),\ S \text{ finite},\right.$$
$$\left. \forall s \lambda_s \geq 0,\ p_s \in \Delta(K),\ \sum_{s\in S} \lambda_s = 1,\ \sum_{s\in S} \lambda_s p_s = p \right\}\,.$$

Lemma 3 (Concavification) *If for any initial probability p, the informed player can guarantee $f(p)$ in the game $\Gamma(p)$, then for any p this player can also guarantee $\mathrm{cav} f(p)$ in $\Gamma(p)$.*

Non Revealing Games

As soon as player 1 uses a strategy which depends on the selected state, the martingale of a posteriori will move and player 2 will have learned something on the state. This is the dilemma of the informed player: he can not use the information on the state without revealing information. Imagine now that player 1 decides to reveal no information on the selected state, and plays independently of it. Since payoffs are defined via expectations, it is as if the players were repeating the average matrix game $G(p) = \sum_{k\in K} p^k G^k$. Its value is:

$$u(p) = \max_{x\in\Delta(I)} \min_{y\in\Delta(J)} \sum_{i,j} x(i)y(j)G(p)(i,j)$$
$$= \min_{y\in\Delta(J)} \max_{x\in\Delta(I)} \sum_{i,j} x(i)y(j)G(p)(i,j)\,.$$



Repeated Games with Incomplete Information, Figure 2
u* and cav*u

u is a Lispchitz function, with constant M, from $\Delta(K)$ to $\mathbb{R}$. Clearly, player 1 can guarantee $u(p)$ in the game $\Gamma(p)$ by playing i. i. d. at each stage an optimal strategy in $G(p)$. By the concavification lemma, we obtain:

Proposition 1 *Player 1 can guarantee* cav$u(p)$ *in the game* $\Gamma(p)$.

Let us come back to the examples. In example 1, we have

$$u(p) = \mathrm{Val}\begin{pmatrix} -(1-p) & 0 \\ 0 & -p \end{pmatrix} = -p(1-p)\,,$$

where $p \in [0,1]$ stands here for the probability of state a. This is a convex function of p, and cav$u(p) = 0$ for all p. In example 2, $u(p) = p(1-p)$ for all p, hence u is already concave and cav$u = u$. Regarding example 3, the following picture show the functions u (regular line), and cav$u$ (dashed line).

Let us consider again the partially revealing strategy previously described. With probability 1/2, the a posteriori will be $3/4a + 1/4b$, and player 1 will play *Top* which is optimal in $3/4G^a + 1/4G^b = \begin{pmatrix} 3 & 1 & 1 \\ 3 & 1 & -1 \end{pmatrix}$. Similarly with probability 1/2, the a posteriori will be $1/4a + 3/4b$ and player 1 will play an optimal strategy in $1/4G^a + 3/4G^b$. Consequently, this strategy guarantees $1/2u(3/4) + 1/2u(1/4) = \mathrm{cav}u(1/2) = 1$ to player 1.

Player 2 can Guarantee the Limit Value

In the infinitely repeated game with initial probability p, player 2 can play as follows: T being fixed, he can play an optimal strategy in the T-stage game $\Gamma_T(p)$, then forget everything and play again an optimal strategy in the T-stage game $\Gamma_T(p)$, etc.… By doing so, he guarantees $v_T(p)$ in the game $\Gamma(p)$. So he can guarantee $\inf_T v_T(p)$ in this game, and this implies that $\limsup_T v_T(p) \le \inf_T v_T(p)$. As a consequence, we obtain:

Proposition 2 *The sequence* $(v_T(p))_T$ *converges to* $\inf_T v_T(p)$*, and this limit can be guaranteed by player 2 in the game* $\Gamma(p)$.

Uniform Value: cav *u* Theorem

We will see here that $\lim_T v_T(p)$ is nothing but cav$u(p)$, and since this quantity can be guaranteed by both players this is the uniform value of the game $\Gamma(p)$. The idea of the proof is the following. The martingale $(p_t(\sigma))_{t\ge 0}$ is bounded, hence will converge almost surely, and we have a bound on its L^1 variation (see Lemma 4 below). This means that after a certain stage the martingale will essentially remain constant, so approximately player 1 will play in a non revealing way, so will not be able to have a stage payoff greater than $u(q)$, where q if a "limit a posteriori". Since the expectation of the a posteriori is the a priori p, player 1 can not guarantee more than $\max\{\sum_{s\in S}\lambda_s u(p_s), S \text{ finite}, \forall s \in S \lambda_s \ge 0, p_s \in \Delta(K), \sum_{s\in S}\lambda_s = 1, \sum_{s\in S}\lambda_s p_s = p\}$, that is more than cav$u(p)$. Let us now proceed to the formal proof.

Fix a strategy σ^1 of player 1, and define the strategy σ^2 of player 2 as follows: play at each stage an optimal strategy in the matrix game $G(p_t)$, where p_t is the current a posteriori in $\Delta(K)$. Assume that $\sigma = (\sigma^1, \sigma^2)$ is played in the repeated game $\Gamma(p)$. To simplify notations, we write $\mathbb{P}$ for $\mathbb{P}_{p,\sigma}$, $p_t(h_t)$ for $p_t(\sigma, h_t)$, etc. We use everywhere norms $\|.\|_1$. To avoid confusion between variables and random variables in the following computations, we will use tildes to denote random variables, e. g. $\tilde{k}$ will denote the random variable of the selected state.

Lemma 4

$$\forall T \ge 1,\quad \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}(\|p_{t+1}-p_t\|) \le \frac{\sum_{k\in K}\sqrt{p^k(1-p^k)}}{\sqrt{T}}.$$

Proof This is a property of martingales with values in $\Delta(K)$ and expectation p. We have for each state k and $t \ge 0$: $\mathbb{E}\left((p^k_{t+1}-p^k_t)^2\right) = \mathbb{E}(\mathbb{E}((p^k_{t+1}-p^k_t)^2|\mathcal{H}_t))$, where $\mathcal{H}_t$ is the σ-algebra on plays generated by the first t action profiles. So $\mathbb{E}\left((p^k_{t+1}-p^k_t)^2\right) = \mathbb{E}(\mathbb{E}((p^k_{t+1})^2 + (p^k_t)^2 - 2p^k_{t+1}p^k_t|\mathcal{H}_t)) = \mathbb{E}((p^k_{t+1})^2) - \mathbb{E}((p^k_t)^2)$. So $\mathbb{E}\left(\sum_{t=0}^{T-1}(p^k_{t+1}-p^k_t)^2\right) = \mathbb{E}\left((p^k_T)^2\right) - (p^k)^2 \le p^k(1-p^k)$. By Cauchy- Schwartz inequality, we also have for each k,

$$\mathbb{E}\left(\frac{1}{T}\sum_{t=0}^{T-1}\left|p^k_{t+1}-p^k_t\right|\right) \le \sqrt{\frac{1}{T}\mathbb{E}\left(\sum_{t=0}^{T-1}(p^k_{t+1}-p^k_t)^2\right)}$$

and the result follows. □

For h_t in $(I\times J)^t$, $\sigma^1_{t+1}(k,h_t)$ is the mixed action in $\Delta(I)$ played by player 1 at stage $t+1$ if the state is k and h_t has previously occurred, and we write $\bar{\sigma}^1_{t+1}(h_t)$ for the law of the action of player 1 of stage $t+1$ after h_t: $\bar{\sigma}^1_{t+1}(h_t) =$

$\sum_{k\in K} p_t^k(h_t)\sigma_{t+1}^1(k,h_t) \in \Delta(I)$. $\bar{\sigma}_{t+1}(h_t)$ can be seen as the average action played by player 1 after h_t, and will be used as a non revealing approximation for $(\sigma_{t+1}^1(k,h_t))_k$. The next lemma precisely links the variation of the martingale $(p_t(\sigma))_{t\geq 0}$, i. e. the information revealed by player 1, and the dependence of player 1's action on the selected state, i. e. the information used by player 1.

Lemma 5

$$\forall t \geq 0, \forall h_t \in (I\times J)^t, \mathbb{E}\left(\|p_{t+1}-p_t\|\,|h_t\right) = \mathbb{E}\left(\left\|\sigma_{t+1}^{\tilde{k}}(h_t)-\bar{\sigma}_{t+1}(h_t)\right\|\,|h_t\right).$$

Proof Fix $t\geq 0$ and h_t in $(I\times J)^t$ s.t. $\mathbb{P}_{p,\sigma}(h_t)>0$. For (i_{t+1},j_{t+1}) in $I\times J$, one has:

$$\begin{aligned} p_{t+1}^k(h_t,i_{t+1},j_{t+1}) &= \mathbb{P}(\tilde{k}=k|h_t,i_{t+1}) \\ &= \frac{\mathbb{P}(\tilde{k}=k|h_t)\mathbb{P}(i_{t+1}|k,h_t)}{\mathbb{P}(i_{t+1}|h_t)} \\ &= \frac{p_t^k(h_t)\sigma_{t+1}^1(k,h_t)(i_{t+1})}{\bar{\sigma}_{t+1}^1(h_t)(i_{t+1})}. \end{aligned}$$

Consequently,

$$\begin{aligned} &\mathbb{E}\left(\|p_{t+1}-p_t\||h_t\right) \\ &= \sum_{i_{t+1}\in I}\bar{\sigma}_{t+1}^1(h_t)(i_{t+1})\sum_{k\in K}|p_{t+1}^k(h_t,i_{t+1})-p_t^k(h_t)| \\ &= \sum_{i_{t+1}\in I}\sum_{k\in K}|p_t^k(h_t)\sigma_{t+1}^1(k,h_t)(i_{t+1}) \\ &\quad -\bar{\sigma}_{t+1}^1(h_t)(i_{t+1})p_t^k(h_t)| \\ &= \sum_{k\in K}p_t^k(h_t)\|\sigma_{t+1}^1(k,h_t)-\bar{\sigma}_{t+1}^1(h_t)\| \\ &= \mathbb{E}\left(\|\sigma_{t+1}^1(\tilde{k},h_t)-\bar{\sigma}_{t+1}^1(h_t)\||h_t\right). \end{aligned}$$

□

We can now control payoffs. For $t\geq 0$ and h_t in $(I\times J)^t$:

$$\begin{aligned} &\mathbb{E}\left(G^{\tilde{k}}(\tilde{i}_{t+1},\tilde{j}_{t+1})|h_t\right) \\ &= \sum_{k\in K}p_t^k(h_t)G^k(\sigma_{t+1}^1(k,h_t),\sigma_{t+1}^2(h_t)) \\ &\leq \sum_{k\in K}p_t^k(h_t)G^k(\bar{\sigma}_{t+1}^1(h_t),\sigma_{t+1}^2(h_t)) \\ &\quad + M\sum_{k\in K}p_t^k(h_t)\|\sigma_{t+1}^1(k,h_t)-\bar{\sigma}_{t+1}^1(h_t)\| \\ &\leq u(p_t(h_t)) + M\sum_{k\in K}p_t^k(h_t)\|\sigma_{t+1}^1(k,h_t)-\bar{\sigma}_{t+1}^1(h_t)\|, \end{aligned}$$

where $u(p_t(h_t))$ comes from the definition of σ^2. By Lemma 5, we get:

$$\begin{aligned} &\mathbb{E}\left(G^{\tilde{k}}(\tilde{i}_{t+1},\tilde{j}_{t+1})|h_t\right) \\ &\qquad \leq u(p_t(h_t)) + M\mathbb{E}\left(\|p_{t+1}-p_t\||h_t\right). \end{aligned}$$

Applying Jensen's inequality yields:

$$\mathbb{E}\left(G^{\tilde{k}}(\tilde{i}_{t+1},\tilde{j}_{t+1})\right) \leq \mathrm{cav}u(p) + M\mathbb{E}\left(\|p_{t+1}-p_t\|\right).$$

We now apply Lemma 4 and obtain:

$$\begin{aligned} \gamma_T^{1,p}(\sigma^1,\sigma^2) &= \mathbb{E}\left(\frac{1}{T}\sum_{t=0}^{T-1}G^{\tilde{k}}(\tilde{i}_{t+1},\tilde{j}_{t+1})\right) \\ &\leq \mathrm{cav}u(p) + \frac{M}{\sqrt{T}}\sum_{k\in K}\sqrt{p^k(1-p^k)}. \end{aligned}$$

This is true for any strategy σ^1 of player 1. Considering the case of an optimal strategy for player 1 in the T-stage game $\Gamma_T(p)$, we have shown:

Proposition 3 *For p in $\Delta(K)$ and $T\geq 1$,*

$$v_T(p) \leq \mathrm{cav}u(p) + \frac{M\sum_{k\in K}\sqrt{p^k(1-p^k)}}{\sqrt{T}}.$$

It remains to conclude about the existence of the uniform value. We have seen that player 1 can guarantee $\mathrm{cav}u(p)$, that player 2 can guarantee $\lim_T v_T(p)$, and we obtain from proposition 3 that $\lim_T v_T(p) \leq \mathrm{cav}u(p)$. This is enough to deduce Aumann and Maschler's celebrated "cav$u$" theorem.

Theorem 1 (Aumann and Maschler [6]) *The game $\Gamma(p)$ has a uniform value which is* $\mathrm{cav}u(p)$.

T-Stage Values and the Recursive Formula

As the T-stage game is a zero-sum game with incomplete information where player 1 is informed, we can write:

$$\begin{aligned} v_T(p) &= \inf_{\sigma^2\in\Sigma^2}\sup_{\sigma^1\in\Sigma^1}\gamma_T^{1,p}(\sigma), \\ &= \inf_{\sigma^2\in\Sigma^2}\sup_{\sigma^1\in\Sigma^1}\sum_{k\in K}p^k\gamma_T^{1,\delta_k}(\sigma), \\ &= \inf_{\sigma^2\in\Sigma^2}\sum_{k\in K}p^k\left(\sup_{\sigma^1\in\Sigma^1}\gamma_T^{1,\delta_k}(\sigma)\right). \end{aligned}$$

This shows that v_T is the infimum of a family of affine functions of p, hence is a concave function of p. This concavity represents the advantage of player 1 to

possess the information on the selected state. Clearly, we have $v_T(p) \geq u(p)$, hence we get the inequalities: $\forall T \geq 1, \text{cav}u(p) \leq v_T(p) \leq \text{cav}u(p) + (M\sum_{k\in K}\sqrt{p^k(1-p^k)})/\sqrt{T}$.

It is also easy to prove that the T-stage value functions satisfy the following recursive formula:

$$v_{T+1}(p) = \frac{1}{T+1}\max_{x\in\Delta(I)^K}\min_{y\in\Delta(J)} \cdot\left(G(p,x,y) + T\sum_{i\in I}x(p)(i)v_T(\hat{p}(x,i))\right)$$
$$= \frac{1}{T+1}\min_{y\in\Delta(J)}\max_{x\in\Delta(I)^K} \cdot\left(G(p,x,y) + T\sum_{i\in I}x(p)(i)v_T(\hat{p}(x,i))\right),$$

where $x = (x^k(i))_{i\in I,k\in K}$, with x^k the mixed action used at stage 1 by player 1 if the state is k, $G(p,x,y) = \sum_{k,i,j}p^k G^k(x^k(i),y(j))$ is the expected payoff of stage 1, $x(p)(i) = \sum_{k\in K}p^k x^k(i)$ is the probability that action i is played at stage 1, and $\hat{p}(x,i)$ is the conditional probability on K given i.

The recursive formula simply is a generalization of the dynamic programming principle. The following property interprets easily: the advantage of the informed player can only decrease as the number of stages increases (for a proof, one can show that $v_{T+1} \leq v_T$ by induction on T, using the concavity of v_T).

Lemma 6 *The T-stage value $v_T(p)$ is non increasing in T.*

Optimal Strategies

In order to determine the optimal behavior of the players, it is important to be able to compute optimal strategies.

The recursive formula enables to compute, by induction on T, an optimal strategy for player 1 in the T-stage game $\Gamma_T(p)$. But it can not be used to compute an optimal strategy for player 2 in the finitely repeated games, because such a strategy should not depend on player 1's strategy, and consequently on the a posteriori of player 2. Constructing such an optimal strategy for player 2 can be done via the recursive formula of a *dual* game, see Subsect. "Zero-Sum Games".

These strategies of player 2 may afterwards be used to construct an optimal strategy in the infinitely repeated game (see Subsect. "Player 2 can Guarantee the Limit Value"): define consecutive blocks of stages $B^1,\ldots,B^T,\ldots$ of respective cardinalities $1,\ldots,T$, and play independently at each block B^T an optimal strategy for player 2 in the T-stage game $\Gamma_T(p)$. This strategy guarantees $\limsup v_T(p) = \text{cav}u(p)$ for the uninformed player, hence is optimal for player 2, in $\Gamma(p)$. In the next section we will also see how to directly construct an *explicit* optimal strategy for player 2 in $\Gamma(p)$, taking care simultaneously of all possible states k.

It is much more simpler to construct an optimal strategy for player 1 in the infinitely repeated game: since player 1 has to guarantee $\text{cav}u(p)$, this can be done using the concavification lemma, see proposition 1.

Vector Payoffs and Approachability

The following model has been introduced by D. Blackwell [8] and is, strictly speaking, not part of the general definition given in Sect. "Definition of the Subject". We still have a family of $I\times J$ matrices $(G^k)_{k\in K}$, where K is a finite set of parameters. At each stage t, simultaneously player 1 chooses $i_t \in I$ and player 2 chooses $j_t \in J$, and the stage "payoff" is the full vector $G(i_t,j_t) = (G^k(i_t,j_t))_{k\in K}$ in $\mathbb{R}^K$. Notice that there is no initial probability or true state of nature here, and both players have a symmetric role. We assume here that after each stage both players observe exactly the stage vector payoff (but one can check that assuming that the action profiles are observed wouldn't change the results). A natural question is then to determine the sets C in $\mathbb{R}^K$ such that player 1 (for example) can force the average long term payoff to belong to C? Such sets will be called *approachable* by player 1.

In Sect. "Vector Payoffs and Approachability", we use Euclidean distances and norms. Denote by $F = \{(G^k(i,j))_{k\in K}, i\in I, j\in J\}$ the finite set of possible stage payoffs, and by M a constant such that $\|u\| \leq M$ for each u in F. A strategy for player 1, resp. player 2, is an element $\sigma = (\sigma_t)_{t\geq 1}$, where σ_t maps F^{t-1} into $\Delta(I)$, resp. $\Delta(J)$. Strategy spaces for player 1 and 2 are respectively denoted by Σ and $\mathcal{T}$. A strategy profile (σ,τ) naturally induces a unique probability on $(I\times J\times F)^\infty$ denoted by $\mathbb{P}_{\sigma,\tau}$. Let C be a "target" set, that will always be assumed, without loss of generality, a closed subset of $\mathbb{R}^K$. We denote by g_t the random variable, with value in F, of the payoff of stage t, and we use $\bar{g}_t = \frac{1}{t}\sum_{t'=1}^t g_{t'} \in \text{conv}(F)$, and finally $d_t = d(\bar{g}_t, C)$ for the distance from $\bar{g}_t$ to C.

Definition 3 C is approachable by player 1 if: $\forall\varepsilon > 0, \exists\sigma\in\Sigma, \exists T, \forall\tau\in\mathcal{T}, \forall t\geq T, \mathbb{E}_{\sigma,\tau}(d_t) \leq \varepsilon$. C is excludable by player 1 if there exist $\delta > 0$ such that $\{z\in\mathbb{R}^K, d(z,C)\geq\delta\}$ is approachable by player 1.

Approachability and excludability for player 2 are defined similarly. C is approachable by player 1 if for each

$\varepsilon > 0$, this player can force that for t large we have $\mathbb{E}_{\sigma,\tau}(d_t) \leq \varepsilon$, so the average payoff will be ε-close to C with high probability. A set cannot be approachable by a player as well as excludable by the other player. In the usual case where K is a singleton, we are in dimension 1 and the Minmax theorem implies that for each t, the interval $[t, +\infty[$ is either approachable by player 1, or excludable by player 2, depending on the comparison between t and the value $\max_{x\in\Delta(I)} \min_{y\in\Delta(J)} G(x,y) = \min_{y\in\Delta(J)} \max_{x\in\Delta(I)} G(x,y)$.

Necessary and Sufficient Conditions for Approachability

Given a mixed action x in $\Delta(I)$, we write xG for the set of possible vector payoffs when player 1 uses x, i. e. $xG = \{G(x,y), y \in \Delta(J)\} = \text{conv}\{\sum_{i\in I} x_i G(i,j), j \in J\}$. Similarly, we write $Gy = \{G(x,y), x \in \Delta(I)\}$ for y in $\Delta(J)$.

Definition 4 The set C is a B(lackwell)-set for player 1 if for every $z \notin C$, there exists $z' \in C$ and $x \in \Delta(I)$ such that: (i) $\|z' - z\| = d(z, C)$, and (ii) the hyperplane containing z' and orthogonal to $[z, z']$ separates z from xG.

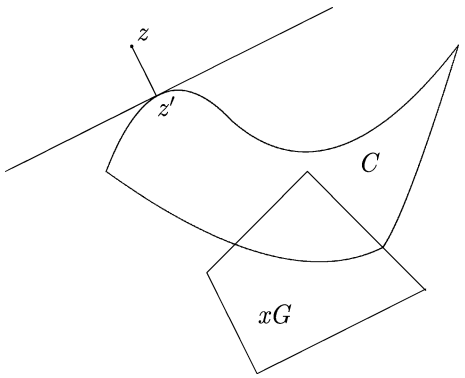
For example, any set xG, with x in $\Delta(I)$, is a B-set for player 1. Given a B-set for player 1, we now construct a strategy σ adapted to C as follows. At each positive stage $t+1$, player 1 considers the current average payoff $\bar{g}_t$. If $\bar{g}_t \in C$, or if $t = 0$, σ plays arbitrarily at stage $t+1$. Otherwise, σ plays at stage $t+1$ a mixed action x satisfying the previous definition for $z = \bar{g}_t$.

Theorem 2 *If C is a B-set for player 1, a strategy σ adapted to C satisfies:*

$$\forall \tau \in \mathcal{T}, \forall t \geq 1,\ \mathbb{E}_{\sigma,\tau}(d_t) \leq \frac{2M}{\sqrt{t}} \quad \text{and} \quad d_t \longrightarrow_{t\to\infty} 0 \mathbb{P}_{\sigma,\tau}\ a.s.$$

As an illustration, in dimension 1 and for $C = \{0\}$, this theorem implies that a bounded sequence $(x_t)_t$ of reals, such that the product $x_{T+1}\left(\frac{1}{T}\sum_{t=1}^{T} x_T\right)$ is non-positive for each T, Cesaro converges to zero.

Proof Assume that player 1 plays σ adapted to C, whereas player 2 plays some strategy τ. Fix $t \geq 1$, and assume that $\bar{g}_t \notin C$. Consider $z' \in C$ and $x \in \Delta(I)$ satisfying (i) and (ii) of definition 4 for $z = \bar{g}_t$. We have:



Repeated Games with Incomplete Information, Figure 3
The Blackwell property

$$\begin{aligned} d_{t+1}^2 &= d(\bar{g}_{t+1}, C)^2 \leq \|\bar{g}_{t+1} - z'\|^2 \\ &= \|\frac{1}{t+1}\sum_{l=1}^{t+1} g_l - z'\|^2 \\ &= \|\frac{1}{t+1}(g_{t+1} - z') + \frac{t}{t+1}(\bar{g}_t - z')\|^2 \\ &= \left(\frac{1}{t+1}\right)^2 \|g_{t+1} - z'\|^2 + \left(\frac{t}{t+1}\right)^2 {d_t}^2 \\ &\quad + \frac{2t}{(t+1)^2}\langle g_{t+1} - z', \bar{g}_t - z'\rangle . \end{aligned}$$

By hypothesis, the expectation, given the first t action profiles $h_t \in (I \times J)^t$, of the above scalar product is non-positive, so $\mathbb{E}\left(d_{t+1}^2 | h_t\right) \leq (t/(t+1))^2 {d_t}^2 + (1/(t+1))^2 \mathbb{E}\left(\|g_{t+1} - z'\|^2 | h_t\right)$. Since $\mathbb{E}\left(\|g_{t+1} - z'\|^2 | h_t\right) \leq \mathbb{E}\left(\|g_{t+1} - \bar{g}_t\|^2 | h_t\right) \leq (2M)^2$, we have:

$$\mathbb{E}\left(d_{t+1}^2 | h_t\right) \leq (\frac{t}{t+1})^2 {d_t}^2 + (\frac{1}{t+1})^2 4M^2 . \qquad (2)$$

Taking the expectation, we get, whether $\bar{g}_t \notin C$ or not: $\forall t \geq 1, \mathbb{E}\left(d_{t+1}^2\right) \leq (t/(t+1))^2 \mathbb{E}({d_t}^2) + (1/(t+1))^2 4M^2$. By induction, we obtain that for each $t \geq 1, \mathbb{E}(d_t^2) \leq (4M^2)/t$, and $\mathbb{E}(d_t) \leq (2M)/\sqrt{t}$.

Put now, as in Sorin 2002 [85], $e_t = d_t^2 + \sum_{t'>t} \frac{4M^2}{t'^2}$. Inequality (2) gives $\mathbb{E}(e_{t+1}|h_t) \leq e_t$, so (e_t) is a non-negative supermartingale which expectation goes to zero. By a standard probability result (see, e. g., Meyer 1966 [58]), we obtain $e_t \longrightarrow_{t\to\infty} 0 \mathbb{P}_{\sigma,\tau}$ a.s., and finally $d_t \longrightarrow_{t\to\infty} 0 \mathbb{P}_{\sigma,\tau}$ a.s. □

This theorem implies that any B-set for player 1 is approachable by this player. The converse is true for convex sets.

Theorem 3 *Let C be a closed convex subset of $\mathbb{R}^K$.*

$$\begin{aligned} & (i)\ C \text{ is a B-set for player 1}\ , \\ \Leftrightarrow\quad & (ii)\ \forall y \in \Delta(J), Gy \cap C \neq \emptyset\ , \\ \Leftrightarrow\quad & (iii)\ C \text{ is approachable by player 1}\ , \\ \Leftrightarrow\quad & (iv)\ \forall q \in \mathbb{R}^K, \\ & \max_{x\in\Delta(I)} \min_{y\in\Delta(J)} \sum_{k\in K} q^k G^k(x,y) \geq \inf_{c\in C} \langle q, c\rangle\ . \end{aligned}$$

Proof The implication $(i) \Longrightarrow (iii)$ comes from Theorem 2. Proof of $(iii) \Longrightarrow (ii)$: assume there exists $y \in \Delta(J)$ such that $Gy \cap C = \emptyset$. Since Gy is approachable by player 2, then C is excludable by player 2 and thus C is not approachable by player 1. Proof of $(ii) \Longrightarrow (i)$: Assume that $Gy \cap C \neq \emptyset \forall y \in \Delta(J)$. Consider $z \notin C$ and define z' as its projection onto C. Define the matrix game where payoffs are projected towards the direction $z' - z$, i. e. the matrix game $\sum_{k\in K}(z'^k - z^k)G^k$. By assumption, one has: $\forall y \in \Delta(J), \exists x \in \Delta(I)$ such that $G(x,y) \in C$, hence such that:

$$\langle z' - z, G(x,y)\rangle \geq \min_{c\in C} \langle z' - z, c\rangle = \langle z' - z, z'\rangle\ .$$

So $\min_{y\in\Delta(J)} \max_{x\in\Delta(I)} \langle z' - z, G(x,y)\rangle \geq \langle z' - z, z'\rangle$. By the minmax theorem, there exists x in $\Delta(I)$ such that $\forall y \in \Delta(J), \langle z' - z, G(x,y)\rangle \geq \langle z' - z, z'\rangle$, that is $\langle z' - z, z' - G(x,y)\rangle \leq 0$.

(iv) means that any half-space containing C is approachable by player 1. $(iii) \Longrightarrow (iv)$ is thus clear. $(iv) \Longrightarrow (i)$ is similar to $(ii) \Longrightarrow (i)$. □

Up to minor formulation differences, Theorems 2 and 3 are due to Blackwell [8]. More recently, X. Spinat [86] has proved the following characterization.

Theorem 4 *A closed set is approachable for player 1 if and only if it contains a B-set for player 1.*

As a consequence, it shows that adding the condition $d_t \longrightarrow_{t\to\infty} 0 \mathbb{P}_{\sigma,\tau}$ a.s. in the definition of approachability does not modify the notion.

Approachability for Player 1 Versus Excludability for Player 2

As a corollary of Theorem 3, we obtain that: *A closed convex set in $\mathbb{R}^K$ is either approachable by player 1, or excludable by player 2.*

One can show that when K is a singleton, then any set is either approachable by player 1, or excludable by player 2. A simple example of a set which is neither approachable for player 1 nor excludable by player 2 is given in dimension 2 by:

$$G = \begin{pmatrix} (0,0) & (0,0) \\ (1,0) & (1,1) \end{pmatrix}, \quad \text{and}$$
$$C = \{(1/2, v), 0 \leq v \leq 1/4\} \bigcup \{(1, v), 1/4 \leq v \leq 1\}$$

(see [85]).

Weak Approachability

On can weaken the definition of approachability by giving up time uniformity.

Definition 5 C is weakly approachable by player 1 if: $\forall \varepsilon > 0, \exists T, \forall t \geq T, \exists \sigma \in \Sigma, \forall \tau \in \mathcal{T}, \mathbb{E}_{\sigma,\tau}(d_t) \leq \varepsilon$. C is weakly excludable by player 1 if there exists $\delta > 0$ such that $\{z \in \mathbb{R}^K, d(z, C) \geq \delta\}$ is weakly approachable by player 1.

N. Vieille [87] has proved, via the consideration of certain differential games:

Theorem 5 *A subset of $\mathbb{R}^K$ is either weakly approachable by player 1 or weakly excludable by player 2.*

Back to the Standard Model

Let us come back to Aumann and Maschler's model with a finite family of matrices $(G^k)_{k\in K}$, and an initial probability p on $\Delta(K)$. By Theorem 1, the repeated game $\Gamma(p)$ has a uniform value which is cav$u(p)$, and Blackwell approachability will allow for the construction of an *explicit* optimal strategy for the uninformed player. Considering a hyperplane which is tangent to cav$u$ at p, we can find a vector l in $\mathbb{R}^K$ such that

$$\begin{aligned} & \langle l, p\rangle = \text{cav}u(p) \\ & \quad \text{and} \quad \forall q \in \Delta(K), \langle l, q\rangle \geq \text{cav}u(q) \geq u(q)\ . \end{aligned}$$

Define now the orthant $C = \{z \in \mathbb{R}^K, z^k \leq l^k \forall k \in K\}$. Recall that player 2 does not know the selected state, and an optimal strategy for him can not depend on player 1' strategy, and consequently on a martingale of a posteriori. He will play in a way such that player 1's long term payoff is, *simultaneously for each k in K*, not greater than l^k if the state is k.

Fix $q = (q^k)_k$ in $\mathbb{R}^K$. If there exists k with $q^k > 0$, we clearly have $\inf_{c\in C}\langle q, c\rangle = -\infty \leq \max_{y\in\Delta(J)} \min_{x\in\Delta(I)} \sum_{k\in K} q^k G^k(x,y)$. Assume now that $q^k \leq 0$ for

each k, with $q \neq 0$. Write $s = \sum_k (-q^k)$.

$$\begin{aligned}\inf_{c\in C}\langle q, c\rangle &= \sum_{k\in K} q^k l^k \\ &= -s\left\langle l, \frac{-q}{s}\right\rangle \\ &\leq -su\left(\frac{-q}{s}\right) \\ &\leq -s \max_{x\in\Delta(I)} \min_{y\in\Delta(J)} \sum_{k\in K} \frac{-q^k}{s} G^k(x,y) \\ &= \max_{y\in\Delta(J)} \min_{x\in\Delta(I)} \sum_{k\in K} q^k G^k(x,y)\end{aligned}$$

This is condition (iv) of Theorem 3, adapted to player 2. So C is a B-set for player 2, and a strategy τ adapted to C satisfies by Theorem 2: $\forall \sigma \in \Sigma, \forall k \in K$,

$$\begin{aligned}&\mathbb{E}_{\sigma,\tau}\left(\frac{1}{T}\sum_{t=1}^{T} G^k(\tilde{i}_t, \tilde{j}_t) - l^k\right) \\ &\quad\leq \mathbb{E}_{\sigma,\tau}\left(d\left(\frac{1}{T}\sum_{t=1}^{T} G^k(\tilde{i}_t, \tilde{j}_t), C\right)\right) \leq \frac{2M}{\sqrt{T}},\end{aligned}$$

(where M is here an upper bound for the Euclidean norms of the vectors $(G^k(i,j))_{k\in K}$, with $i \in I$ and $j \in J$.) And this holds as well for any strategy σ of player 1 in the repeated game with incomplete information. So for any such strategy σ,

$$\begin{aligned}\gamma_T^{1,p}(\sigma,\tau) &= \sum_{k\in K} p^k \left(\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_{\sigma,\tau}(G^k(\tilde{i}_t, \tilde{j}_t))\right) \\ &\leq \langle p, l\rangle + \frac{2M}{\sqrt{T}} = \mathrm{cav}u(p) + \frac{2M}{\sqrt{T}}.\end{aligned}$$

As shown by Kohlberg [35], the approachability strategy τ is thus an optimal strategy for player 2 in the repeated game $\Gamma(p)$.

Zero-Sum Games with Lack of Information on Both Sides

The following model has also been introduced by Aumann and Maschler [6]. We are still in the context of zero-sum repeated games with observable actions, but it is no longer assumed that one of the players is fully informed. The set of states is here a product $K \times L$ of finite sets, and we have a family of matrices $(G^{k,l})_{(k,l)\in K\times L}$ with size $I \times J$, as well as initial probabilities p on K, and q on L. In the game $\Gamma(p,q)$, a state of nature (k,l) is first selected according to the product probability $p \otimes q$, then k, resp. l, is announced to player 1, resp. player 2 only. Then the matrix game $G^{k,l}$ is repeated over and over: at every stage, simultaneously player 1 chooses a row i in I, whereas player 2 chooses a column j in J, the stage payoff for player 1 is $G^{k,l}(i,j)$ but only i and j are publicly announced before proceeding to the next stage.

The average payoff for player 1 in the T-stage game is written: $\gamma_T^{1,p,q}(\sigma^1,\sigma^2) = \mathbb{E}^{p,q}_{\sigma^1,\sigma^2}\left(\frac{1}{T}\sum_{t=1}^{T} G^{\tilde{k},\tilde{l}}(\tilde{i}_t,\tilde{j}_t)\right)$, and the T-stage value is written $v_T(p,q)$. Similarly, the λ-discounted value of the game will be written $v_\lambda(p,q)$

The non revealing game now corresponds to the case where player 1 plays independently of k and player 2 plays independently of l. Its value is denoted by:

$$u(p,q) = \max_{x\in\Delta(I)} \min_{y\in\Delta(J)} \sum_{k,l} p^k q^l G^{k,l}(x,y). \tag{3}$$

Given a continuous function $f\colon \Delta(K)\times\Delta(L) \longrightarrow \mathbb{R}$, we denote by $\mathrm{cav}_{\mathrm{I}} f$ the concavification of f with respect to the first variable: for each (p,q) in $\Delta(K)\times\Delta(L)$, $\mathrm{cav}_{\mathrm{I}} f(p,q)$ is the value at p of the smallest concave function from $\Delta(K)$ to $\mathbb{R}$ which is above $f(.,q)$. Similarly, we denote by $\mathrm{vex}_{\mathrm{II}} f$ the convexification of f with respect to the second variable. It can be shown that $\mathrm{cav}_{\mathrm{I}} f$ and $\mathrm{vex}_{\mathrm{II}} f$ are continuous, and we can compose $\mathrm{cav}_{\mathrm{I}}\mathrm{vex}_{\mathrm{II}} f$ and $\mathrm{vex}_{\mathrm{II}}\mathrm{cav}_{\mathrm{I}} f$. These functions are both concave in the first variable and convex in the second variable, and they satisfy $\mathrm{cav}_{\mathrm{I}}\mathrm{vex}_{\mathrm{II}} f(p,q) \leq \mathrm{vex}_{\mathrm{II}}\mathrm{cav}_{\mathrm{I}} f(p,q)$.

Maxmin and Minmax of the Repeated Game

Theorem 1 generalizes as follows.

Theorem 6 ([6]) *In the repeated game $\Gamma(p,q)$, the greatest quantity which can be guaranteed by player 1 is $cav_I vex_{II} u(p,q)$, and the smallest quantity which can be guaranteed by player 2 is $vex_{II} cav_I u(p,q)$.*

Aumann, Maschler and Stearns also showed that $\mathrm{cav}_{\mathrm{I}}\mathrm{vex}_{\mathrm{II}} u(p,q)$ can be *defended* by player 2, uniformly in time, i. e. that $\forall \varepsilon > 0, \forall \sigma^1, \exists T_0, \exists \sigma^2, \forall T \geq T_0, \gamma_T^{p,q}(\sigma^1,\sigma^2) \leq \mathrm{cav}_{\mathrm{I}}\mathrm{vex}_{\mathrm{II}} u(p,q)v + \varepsilon$. Similarly, $\mathrm{vex}_{\mathrm{II}}\mathrm{cav}_{\mathrm{I}} u(p,q)$ can be *defended* by player 1.

The proof uses the martingales of a posteriori of each player, and a useful notion is that of the informational content of a strategy: for a strategy σ^1 of the first player, it is defined as: $I(\sigma^1) = \sup_{\sigma^2} \mathbb{E}^{p,q}_{\sigma^1,\sigma^2}\left(\sum_{k\in K}\sum_{t=0}^{\infty} (p^k_{t+1}(\sigma^1) - p^k_t(\sigma^1))^2\right)$, where $p_t(\sigma^1)$ is the a posteriori on K of player 2 after stage t given that player 1 uses σ^1. By linearity of the expectation, the supremum can be re-

stricted to strategies of player 2 which are both pure and independent of l.

Theorem 6 implies that $\mathrm{cav_I vex_{II}} u(p,q) = \sup_{\sigma^1 \in \Sigma^1} \liminf_T \left(\inf_{\sigma^2 \in \Sigma^2} \gamma_T^{1,p,q}(\sigma^1, \sigma^2) \right)$, and $\mathrm{cav_I vex_{II}} u(p,q)$ is called the *maxmin* of the repeated game $\Gamma(p,q)$. Similarly, $\mathrm{vex_{II} cav_I} u(p,q) = \inf_{\sigma^2 \in \Sigma^2} \limsup_T \left(\sup_{\sigma^1 \in \Sigma^1} \gamma_T^1(\sigma^1, \sigma^2) \right)$ is called the *minmax* of $\Gamma(p,q)$. As a corollary, we obtain that the repeated game $\Gamma(p,q)$ has a uniform value if and only if: $\mathrm{cav_I vex_{II}} u(p,q) = \mathrm{vex_{II}} \mathrm{cav_I} u(p,q)$. This is not always the case, and there exist counter-examples to the existence of the uniform value.

Example 4 $K = \{a, a'\}$, and $L = \{b, b'\}$, with p and q uniform.

$$G^{a,b} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 \end{pmatrix} \quad G^{a,b'} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$G^{a',b} = \begin{pmatrix} -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad G^{a',b'} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

Mertens and Zamir [52] have shown that here, $\mathrm{cav_I vex_{II}} u(p,q) = -\frac{1}{4} < 0 = \mathrm{vex_{II} cav_I} u(p,q)$.

Limit Values

It is easy to see that for each T and λ, the value functions v_T and v_λ are concave in the first variable, and convex in the second variable. They are all Lipschitz functions, with the same constant $M = \max_{i,j,k,l} |G^{k,l}(i,j)|$, and here also, recursive formula can be given. In the following result, v_T and v_λ are viewed as elements of the set C of continuous mappings from $\Delta(K) \times \Delta(L)$ to $\mathbb{R}$.

Theorem 7 (Mertens and Zamir [52]) *$(v_T)_T$, as T goes to infinity, and $(v_\lambda)_\lambda$, as λ goes to zero, both uniformly converge to the unique solution f of the following system:*

$$\begin{cases} f = vex_{II} \max\{u, f\} \\ f = cav_I \min\{u, f\} \end{cases}$$

Besides, the above system can also be fruitfully studied without reference to repeated games (see [39,40,55,81]). For a proof of Theorem 7, one can also see Zamir [91] or Sorin [85]. Mertens and Zamir notably consider responses of a player, to a given strategy of his opponent, which are of the following type: play non revealing up to a particular stopping time, and then start using the information by playing optimally in the remaining subgame.

Remark Let U be the set of all non revealing value functions, i. e. of functions from $\Delta(K) \times \Delta(L)$ to $\mathbb{R}$ satisfying Eq. (3) for some family of matrices $(G^{k,l})_{k,l}$. One can easily show that any mapping in C is a uniform limit of elements in U.

Correlated Initial Information

A more general model can be written, where it is no longer assumed that the initial information of the players are independent. The set of states is now denoted by R (instead of $K \times L$), initially a state r in R is chosen according to a known probability $p = (p^r)_{r \in R}$, and each player receives a deterministic signal depending on r. Equivalently, each player i has a partition R^i of R and observes the element of his partition which contains the selected state.

After the first stage, player 1 will play an action $x = (x^r)_{r \in R}$ which is measurable with respect to R^1, i. e. $(r \longrightarrow x^r)$ is constant on each atom of R^1. After having observed player 1's action at the first stage, the conditional probability on R necessarily belongs to the set:

$$\Pi^{\mathrm{I}}(p) = \left\{ (\alpha^r p^r)_{r \in R}, \forall r \alpha^r \geq 0, \sum_r \alpha^r p^r = 1 \right. \\ \left. \text{and } (\alpha^r)_r \text{ is } R^1\text{-measurable} \right\}.$$

$\Pi^{\mathrm{I}}(p)$ contains p, and is a convex compact subset of $\Delta(R)$. A mapping f from $\Delta(R)$ to $\mathbb{R}$ is now said to be I-concave if for each p in $\Delta(R)$, the restriction of f to $\Pi^{\mathrm{I}}(p)$ is concave. And given $g\colon \Delta(R) \longrightarrow \mathbb{R}$ which is bounded from above, we define the concavification $\mathrm{cav_I} g$ as the smallest function above g which is I-concave. Similarly one can define the set $\Pi^{\mathrm{II}}(p)$ and the notions of II-convexity and II-convexification. With these generalized definitions, the results of Theorem 6 and 7 perfectly extend [52].

Non Zero-Sum Games with Lack of Information on One Side

We now consider the generalization of the standard model of Sect. "The Standard Model of Aumann and Maschler" to the non-zero sum case. Hence two players infinitely repeat the same bimatrix game, with player 1 only knowing the bimatrix. Formally, we have a finite set of states K, an initial probability p on K, and families of $I \times J$-payoff matrices $(A^k)_{k \in K}$ and $(B^k)_{k \in K}$. Initially, a state k in K is selected according to p, and announced to player 1 only. Then the bimatrix game (A^k, B^k) is repeated over and over: at every stage, simultaneously player 1 chooses a row i in I, whereas player 2 chooses a column j in J, the stage payoff for player 1 is then $A^k(i,j)$, the stage payoff for player 2 is $B^k(i,j)$, but only i and j are publicly announced before proceeding to the next stage. Without loss of generality, we assume that $p^k > 0$ for each k, and that each player has at least two actions.

Given a strategy pair (σ^1, σ^2), it is here convenient to denote the expected payoffs up to stage T by:

$$\alpha_T^p(\sigma^1,\sigma^2) = \mathbb{E}_{p,\sigma^1,\sigma^2}\left(\frac{1}{T}\sum_{t=1}^{T} A^{\tilde{k}}(\tilde{i}_t,\tilde{j}_t)\right) = \sum_{k\in K} p^k \alpha_T^k(\sigma^1,\sigma^2)\,.$$

$$\beta_T^p(\sigma^1,\sigma^2) = \mathbb{E}_{p,\sigma^1,\sigma^2}\left(\frac{1}{T}\sum_{t=1}^{T} B^{\tilde{k}}(\tilde{i}_t,\tilde{j}_t)\right) = \sum_{k\in K} p^k \beta_T^k(\sigma^1,\sigma^2)\,.$$

Given a probability q on K, we write $A(q) = \sum_k q^k A^k$, $B(q) = \sum_k q^k B^k$, $u(q) = \max_{x\in\Delta(I)} \min_{y\in\Delta(J)} A(q)(x,y)$ and $v(q) = \max_{y\in\Delta(J)} \min_{x\in\Delta(I)} B(q)(x,y)$. If $\gamma = (\gamma(i,j))_{(i,j)\in I\times J} \in \Delta(I\times J)$, we put $A(q)(\gamma) = \sum_{(i,j)\in I\times J} \gamma(i,j)A(q)(i,j)$ and similarly $B(q)(\gamma) = \sum_{(i,j)\in I\times J} \gamma(i,j)\, B(q)(i,j)$.

Existence of Equilibria

The question of existence of an equilibrium has remained unsolved for long. Sorin [79] proved the existence of an equilibrium for two states of nature, and the general case has been solved by Simon et al. [76].

Exactly as in the zero-sum case, a strategy pair σ induces a sequence of a posteriori $(p_t(\sigma))_{t\geq 0}$ which is a $\mathbb{P}_{p,\sigma}$- martingale with values in $\Delta(K)$. We will concentrate on the cases where this martingale moves only once.

Definition 6 A joint plan is a triple (S,λ,γ), where:

- S is a finite non empty set (of messages),
- $\lambda = (\lambda^k)_{k\in K}$ (signaling strategy) with for each $k, \lambda^k \in \Delta(S)$ and for each s, $\lambda_s =_{\text{def}} \sum_{k\in K} p^k \lambda_s^k > 0$,
- $\gamma = (\gamma_s)_{s\in S}$ (contract) with for each s, $\gamma_s \in \Delta(I\times J)$.

The idea is due to Aumann, Maschler and Stearns. Player 1 observes k, then chooses $s\in S$ according to λ^k and announces s to player 2. Then the players play pure actions corresponding to the frequencies $\gamma_s(i,j)$, for i in I and j in J. Given a joint plan (S,λ,γ), we define:

- $\forall s\in S$, $p_s = (p_s^k)_{k\in K} \in \Delta(K)$, with $p_s^k = \frac{p^k\lambda_s^k}{\lambda_s}$ for each k. p_s is the a posteriori on K given s.
- $\varphi = (\varphi^k)_{k\in K} \in \mathbb{R}^K$, with for each k, $\varphi^k = \max_{s\in S} A^k(\gamma_s)$.
- $\forall s\in S$, $\psi_s = B(p_s)(\gamma_s)$ and $\psi = \sum_{k\in K} p^k \sum_{s\in S} \lambda_s^k B^k(\gamma_s) = \sum_{s\in S} \lambda_s \psi_s$.

Definition 7 A joint plan (S,λ,γ) is an equilibrium joint plan if:

(i) $\forall s\in S, \psi_s \geq \operatorname{vex}v(p_s)$,
(ii) $\forall k\in K, \forall s\in S$ s.t. $p_s^k > 0$, $A^k(\gamma_s) = \varphi^k$, and
(iii) $\forall q\in\Delta(K)$, $\langle\varphi, q\rangle \geq u(q)$.

Condition (ii) can be seen as an incentive condition for player 1 to choose s according to λ^k. Given an equilibrium joint plan (S,λ,γ), one define a strategy pair $(\sigma^{1*},\sigma^{2*})$ adapted to it. For each message s, first fix a sequence $(i_t^s, j_t^s)_{t\geq 1}$ of elements in $I\times J$ such that for each (i,j), the empirical frequencies converge to the corresponding probability: $\frac{1}{T}|\{t, 1\leq t\leq T, (i_t^s, j_t^s) = (i,j)\}| \longrightarrow_{T\to\infty} \gamma_s(i,j)$. We also fix an injective mapping f from S to I^l, where l is large enough, corresponding to a code between the players to announce an element in S. σ^{1*} is precisely defined as follows. Player 1 observes the selected state k, then chooses s according to λ^k, and announces s to player 2 by playing $f(s)$ at the first l stages. Finally, σ^{1*} plays i_t^s at each stage $t > l$ as long as player 2 plays j_t^s. If at some stage $t > l$ player 2 does not play j_t^s then player 1 punishes his opponent by playing an optimal strategy in the zero-sum game with initial probability p_s and payoffs for player 1 given by $(-B^k)_{k\in K}$. We now define σ^{2*}. Player 2 arbitrarily plays at the beginning of the game, then compute at the end of stage l the message s sent by player 1. Next he plays at each stage $t > l$ the action j_t^s as long as player 1 plays i_t^s. If at some stage $t > l$, player 1 does not play i_t^s, or if the first l actions of player 1 correspond to no message, then player 2 plays a punishing strategy $\bar{\sigma}^2$ such that: $\forall\varepsilon > 0, \exists T_0, \forall T\geq T_0, \forall\sigma^1\in\Sigma^1, \forall k\in K, \alpha_T^k(\sigma^1,\bar{\sigma}^2) \leq \varphi^k + \varepsilon$. Such a strategy $\bar{\sigma}^2$ exists because of condition (iii): it is an approachability strategy for player 2 of the orthant $\{x\in\mathbb{R}^K, \forall k\in K\, x^k \leq \varphi^k\}$ (see Subsect. "Back to the Standard Model").

Lemma 7 ([79]) *A strategy pair adapted to an equilibrium joint plan is a uniform equilibrium of the repeated game.*

Proof The payoffs induced by $(\sigma^{1*},\sigma^{2*})$ can be easily computed: $\forall k, \alpha_T^k(\sigma^{1*},\sigma^{2*}) \longrightarrow_{T\to\infty} \sum_{s\in S}\lambda_s^k A^k(\gamma_s) = \varphi^k$ because of (*ii*), and $\beta_T^p(\sigma^{1*},\sigma^{2*}) \longrightarrow_{T\to\infty} \sum_{k\in K} p^k \sum_{s\in S}\lambda_s^k B^k(\gamma_s) = \psi$. Assume that player 2 plays σ^{2*}. The existence of $\bar{\sigma}^2$ implies that no detectable deviation of player 1 is profitable, so if the state is k, player 1 will gain no more than $\max_{s'\in S} A^k(\gamma_{s'})$. But this is just φ^k. The proof can be made uniform in σ^1 and we obtain: $\forall\varepsilon > 0\ \exists T_0\ \forall T\geq T_0, \forall k\in K, \forall\sigma^1\in\Sigma^1, \alpha_T^k(\sigma^1,\sigma^{2*}) \leq \varphi^k + \varepsilon$. Finally assume that player 1 plays σ^{1*}. Condition (*i*) implies that if player 2 uses σ^{2*},

the payoff of this player will be at least $\text{vex}v(p_s)$ if the message is s. Since $\text{vex}\, v(p_s)(= -\text{cav}(-v(p_s)))$ is the value, from the point of view of player 2 with payoffs $(B^k)_k$, of the zero-sum game with initial probability p_s, player 2 fears the punition by player 1, and $\forall \varepsilon > 0, \exists T_0, \forall T \geq T_0, \forall \sigma^2 \in \Sigma^2, \beta_T^p(\sigma^{1*}, \sigma^2) \leq \sum_{s\in S} \lambda_s \psi_s + \varepsilon = \psi + \varepsilon$. □

To prove the existence of equilibria, we then look for equilibrium joint plans. The first idea is to consider, for each probability r on K, the set of payoff vectors φ compatible with r being an a posteriori. This leads to the consideration of the following correspondence (for each r, $\Phi(r)$ is a subset of $\mathbb{R}^K$):

$$\Phi : \Delta(K) \rightrightarrows \mathbb{R}^K$$
$$r \mapsto \{(A^k(\gamma))_{k\in K}, \quad \text{where } \gamma \in \Delta(I \times J) \text{ satisfies } B(r)(\gamma) \geq \text{vex}\, v(r)\}$$

It is easy to see that the graph of Φ, i. e. the set $\{(r, \varphi) \in \Delta(K) \times \mathbb{R}^K, \varphi \in \Phi(r)\}$, is compact, that Φ has non empty convex values, and satisfies: $\forall r \in \Delta(K), \forall q \in \Delta(K), \exists \varphi \in \Phi(r), \langle \varphi, q\rangle \geq u(q)$.

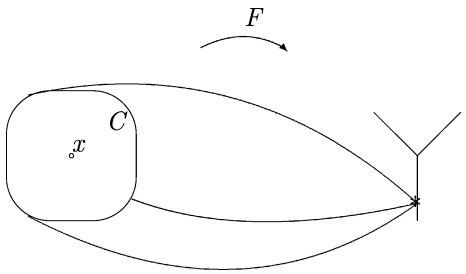
Assume now that one can find a finite family $(p_s)_{s\in S}$ of probabilities on K, as well as vectors φ and, for each s, φ_s in $\mathbb{R}^K$ such that: 1) $p \in \text{conv}\{p_s, s \in S\}$, 2) $\langle \varphi, q\rangle \geq u(q) \forall q \in \Delta(K)$, 3) $\forall s \in S, \varphi_s \in \Phi(p_s)$, and 4) $\forall s \in S, \forall k \in K, \varphi_s^k \leq \varphi^k$ with equality if $p_s^k > 0$. It is then easy to construct an equilibrium joint plan. Thus we get interested in proving the following result.

Proposition 4 *Let p be in $\Delta(K)$, $u : \Delta(K) \longrightarrow \mathbb{R}$ be a continuous mapping, and $\Phi : \Delta(K) \rightrightarrows \mathbb{R}^K$ be a correspondence with compact graph and non empty convex values such that: $\forall r \in \Delta(K), \forall q \in \Delta(K), \exists \varphi \in \Phi(r), \langle \varphi, q\rangle \geq u(q)$. Then there exists a finite family $(p_s)_{s\in S}$ of elements of $\Delta(K)$, as well as vectors φ and, for each s, φ_s in $\mathbb{R}^K$ such that:*

- $p \in \text{conv}\{p_s, s \in S\}$,
- $\langle \varphi, q\rangle \geq u(q) \forall q \in \Delta(K)$,
- $\forall s \in S, \varphi_s \in \Phi(p_s)$,
- $\forall s \in S, \forall k \in K, \varphi_s^k \leq \varphi^k$ *with equality if* $p_s^k > 0$.

The proof of proposition 4 relies, as explained in [62] or [75], on a fixed point theorem of Borsuk–Ulam type proved by Simon, Spież and Toruńczyk [76] via tools from algebraic geometry. A simplified version of this fixed point theorem can be written as follows:

Theorem 8 ([76]) *Let C be a compact subset of an n-dimensional Euclidean space, $x \in C$ and Y be a finite union of affine subspaces of dimension $n-1$ of an Euclidean space. Let F be a correspondence from C to Y with compact graph and non empty convex values. Then there exists $L \subset \partial C$ and $y \in Y$ such that: $\forall l \in L, y \in F(l)$ and $x \in \text{conv}(L)$.*



Repeated Games with Incomplete Information, Figure 4
A Borsuk–Ulam type theorem by Simon, Spież and Toruńczyk

Notice that for $n = 1$ (corresponding to 2 states of nature), the image by F of the connected component of C containing x necessarily is a singleton, hence the result is clear. In the general case, one finally obtains:

Theorem 9 ([76]) *There exists an equilibrium joint plan. Thus there exists a uniform equilibrium in the repeated game $\Gamma(p)$.*

Characterization of Equilibrium Payoffs

Characterizing equilibrium payoffs, as the Folk theorem does for repeated games with complete information, has been a challenging problem. We denote here by p_0 the initial probability in the interior of $\Delta(K)$. We are interested in the set of equilibrium payoffs, in the convenient following sense:

Definition 8 A vector (a, b) in $\mathbb{R}^K \times \mathbb{R}$ is called an equilibrium payoff of the repeated game $\Gamma(p_0)$ if there exists a strategy pair $(\sigma^{1*}, \sigma^{2*})$ satisfying:

(i) $\forall \varepsilon > 0\ \exists T_0\ \forall T \geq T_0, \forall k \in K, \forall \sigma^1 \in \Sigma^1, \alpha_T^k(\sigma^1, \sigma^{2*}) \leq \alpha_T^k(\sigma^{1*}, \sigma^{2*}) + \varepsilon$, and $\forall \varepsilon > 0\ \exists T_0 \forall T \geq T_0, \forall \sigma^2 \in \Sigma^2, \beta_T^{p_0}(\sigma^{1*}, \sigma^2) \leq \beta_T^p(\sigma^{1*}, \sigma^{2*}) + \varepsilon$, and

(ii) $(\alpha_T^k(\sigma^{1*}, \sigma^{2*}))_{k,T}$ and $(\beta_T^{p_0}(\sigma^{1*}, \sigma^{2*}))_T$ respectively converge to a and b.

Since p lies in the interior of $\Delta(K)$, the first line of (i) is equivalent to: $\forall \varepsilon > 0\ \exists T_0\ \forall T \geq T_0, \forall \sigma^1 \in \Sigma^1, \alpha_T^p(\sigma^1, \sigma^{2*}) \leq \alpha_T^p(\sigma^{1*}, \sigma^{2*}) + \varepsilon$. The strategy pair $(\sigma^{1*}, \sigma^{2*})$ is thus a uniform equilibrium of the repeated game, with the additional requirement that expected average payoffs of player 1 converge in each state k. In some sense, player 1

is viewed here as $|K|$ different types or players, and we require the existence of the limit payoff of each type. We will only consider such uniform equilibria in the sequel.

Notice that the above definition implies: $\forall k \in K$, $\forall \varepsilon > 0, \exists T_0, \forall T \geq T_0, \forall \sigma^1 \in \Sigma^1, \alpha_T^k(\sigma^1, \sigma^{2*}) \leq a^k + \varepsilon$. So the orthant $\{x \in \mathbb{R}^K, x^k \leq a^k \forall k \in K\}$ is approachable by player 2, and by Theorem 3 and Subsect. "Back to the Standard Model" one can obtain that:

$$\langle a, q \rangle \geq u(q) \quad \forall q \in \Delta(K) \tag{4}$$

Condition (4) is called the individual rationality condition for player 1, and does not depend on the initial probability in the interior of $\Delta(K)$. Regarding player 2, we have: $\forall \varepsilon > 0 \; \exists T_0 \; \forall T \geq T_0, \forall \sigma^2 \in \Sigma^2, \beta_T^{p_0}(\sigma^{1*}, \sigma^2) \leq \beta + \varepsilon$, so by Theorem 1:

$$\beta \geq \operatorname{vex} v(p_0)\,. \tag{5}$$

Condition (5) is the individual rationality condition for player 2: at equilibrium, this player should have at least the value of the game where player 1's plays in order to minimize player 2's payoffs.

Imagine now that σ^{1*} is a non revealing strategy for player 1, and that the players play actions with empirical frequencies corresponding to a given probability distribution $\pi = (\pi_{i,j})_{(i,j)\in I\times J} \in \Delta(I \times J)$. We will have: $\forall k \in K$, $a^k = \sum_{i,j} \pi_{i,j} A^k(i,j)$ and $\beta = \sum_k p_0^k \sum_{i,j} \pi_{i,j} B^k(i,j)$, and if the individual rationality conditions are satisfied, no detectable deviation of a player can be profitable. This leads to the definition of the following set, where M is the constant $\max\{|A^k(i,j)|, |B^k(i,j)|, (i,j) \in I \times J\}$, and $\mathbb{R}_M = [-M, M]$.

Definition 9 Let G be the set of triples $(a, \beta, p) \in \mathbb{R}_M^K \times \mathbb{R}_M \times \Delta(K)$ satisfying:

1. $\forall q \in \Delta(K), \langle a, q \rangle \geq u(q)$,
2. $\beta \geq \operatorname{vex} v(p)$,
3. $\exists \pi \in \Delta(I \times J)$ s.t. $\beta = \sum_k p^k \sum_{i,j} \pi_{i,j} B^k(i,j)$ and $\forall k \in K, a^k \geq \sum_{i,j} \pi_{i,j} A^k(i,j)$ with equality if $p^k > 0$.

We need to considerate every possible initial probability because the main state variable of the model is, here also, the belief, or a posteriori, of player 2 on the state of nature. $\{(a,\beta), (a,\beta,p_0) \in G\}$ is the set of payoffs of non revealing equilibria of $\Gamma(p_0)$. The importance of the following definition will appear with Theorem 10 below (which unfortunately has not led to a proof of existence of equilibrium payoffs).

Definition 10 G^* is defined as the set of elements $g = (a, \beta, p) \in \mathbb{R}_M^K \times \mathbb{R}_M \times \Delta(K)$ such that there exist a probability space $(\Omega, \mathcal{A}, Q)$, an increasing sequence $(\mathcal{F}_n)_{n\geq 1}$ of finite sub-σ-algebras of $\mathcal{A}$, and a sequence of random variables $(g_n)_{n\geq 1} = (a_n, \beta_n, p_n)_{n\geq 1}$ defined on $(\Omega, \mathcal{A})$ with values in $\mathbb{R}_M^K \times \mathbb{R}_M \times \Delta(K)$ satisfying: (i) $g_1 = g$ a.s., (ii) $(g_n)_{n\geq 1}$ is a martingale adapted to $(\mathcal{F}_n)_{n\geq 1}$, (iii) $\forall n \geq 1, a_{n+1} = a_n$ a.s. or $p_{n+1} = p_n$ a.s., and (iv) $(g_n)_n$ converges a.s. to a random variable g_∞ with values in G.

Let us forget for a while the component of player 2's payoff. A process $(g_n)_n$ satisfying (ii) and (iii) may be called a bi-martingale, it is a martingale such that at every stage, one of the two components remains a.s. constant. So the set G^* can be seen as the set of starting points of converging bi-martingales with limit points in G.

Theorem 10 (Hart [29]) *Let (a, β) be in $\mathbb{R}^K \times \mathbb{R}$.*

$$(a, \beta) \text{ is an equilibrium payoff of } \Gamma(p_0) \iff (a, \beta, p_0) \in G^*\,.$$

Theorem 10 is too elaborate to be proved here, but let us give a few ideas about the proof. First consider the implication $\Longrightarrow$, and fix an equilibrium $\sigma^* = (\sigma^{1*}, \sigma^{2*})$ of $\Gamma(p_0)$ with payoff (a, β). The sequence of a posteriori $(p_t(\sigma^*))_{t\geq 0}$ is a $\mathbb{P}_{p_0,\sigma^*}$- martingale. Modify now slightly the time structure so that at each stage, player 1 plays first, and then player 2 plays without knowing the action chosen by player 1. At each half-stage where player 2 plays, his a posteriori remains constant. At each half-stage where player 1 plays, the "expectation of player 1's future payoff" (which can be properly defined) remains constant. Hence, the heuristic apparition of the bimartingale. And since bounded martingale converge, for large stages everything will be fixed and the players will approximately play a non revealing equilibrium at a "limit a posteriori", so the convergence will be towards elements of G.

Consider now the converse implication $\Longleftarrow$. Let (a, β) be such that $(a, \beta, p_0) \in G^*$, and assume for simplification that the associated bi-martingale (a_n, β_n, p_n) converges in a fixed number N of stages: $\forall n \geq N, (a_n, \beta_n, p_n) = (a_N, \beta_N, p_N) \in G$. One can construct an equilibrium $(\sigma^{1*}, \sigma^{2*})$ of $\Gamma(p_0)$ with payoff (a, β) along the following lines. For each index n, (a_n, β_n) will be an equilibrium payoff of the repeated game with initial probability p_n. Eventually, player 1 will play independently of the state, the a posteriori of player 2 will be p_N, and the players will end up playing a non revealing equilibrium of the repeated game $\Gamma(p_N)$ with payoff (a_N, β_N). What should be played before? Since we are in an undiscounted setup, any finite number of stages can be used for communication without

influencing payoffs. Let $n < N$ be such that $a_{n+1} = a_n$. To move from (a_n, β_n, p_n) to $(a_n, \beta_{n+1}, p_{n+1})$, player 1 can simply use the splitting lemma (Lemma 1) in order to signal part of the state to player 2. Let now $n < N$ be such that $p_{n+1} = p_n$, so that we want to move from (a_n, β_n, p_n) to $(a_{n+1}, \beta_{n+1}, p_n)$. Player 1 will play independently of the state, and both players will act so as to convexify their future payoffs. This convexification is done through procedures called "jointly controlled lotteries" and introduced in the sixties by Aumann and Maschler [6], with the following simple and brilliant idea. Imagine that the players have to decide with even probability whether to play the equilibrium $E1$ with payoff (a^1, β^1) or to play the equilibrium $E2$ with payoff (a^2, β^2). The players may not be indifferent between $E1$ and $E2$, e. g. player 1 may prefer $E1$ whereas player 2 prefers $E2$. They will proceed as follows, with i and i', respectively j and j', denoting two distinct actions of player 1, resp. player 2. Simultaneously and independently, player 1 will select i or i' with probability 1/2, whereas player 2 will behave similarly with j and j'.

$$\begin{array}{cc} & \begin{array}{cc} j & j' \end{array} \\ \begin{array}{c} i \\ i' \end{array} & \begin{pmatrix} \times & \\ & \times \end{pmatrix} \end{array} .$$

Then the equilibrium $E1$ will be played if the diagonal has been reached, i. e. if (i, j) or (i', j') has been played, and otherwise the equilibrium $E2$ will be played. This procedure is robust to unilateral deviations: none of the players can deviate and prevent $E1$ and $E2$ to be chosen with probability 1/2. In general, jointly controlled lotteries are procedures allowing to select an alternative among a finite set according to a given probability (think of binary expansions if necessary), in a way which is robust to deviations by a single player. S. Hart has precisely shown how to combine steps of signaling and jointly controlled lotteries to construct an equilibrium of $\Gamma_\infty(p_0)$ with payoff (a, β).

Biconvexity and Bimartingales

The previous analysis has lead to the introduction and study of biconvexity phenomena. The reference here is [4]. Let X and Y be compact convex subsets of Euclidean spaces, and let $(\Omega, \mathcal{F}, \mathcal{P})$ be an atomless probability space.

Definition 11 A subset B of $X \times Y$ is biconvex if for every x in X and y in Y, the sections $B_{x.} = \{y' \in Y, (x, y') \in B\}$ and $B_{.y} = \{x' \in X, (x', y) \in B\}$ are convex. If B is biconvex, a mapping $f : B \longrightarrow \mathbb{R}$ is called biconvex if for each $(x, y) \in X \times Y$, $f(., y)$ and $f(x, .)$ are convex.

As in the usual convexity case, we have that if f is biconvex, then for each α in $\mathbb{R}$, the set $\{(x, y) \in B, f(x, y) \leq \alpha\}$ is biconvex.

Definition 12 A sequence of random variables $Z_n = (X_n, Y_n)_{n\geq 1}$ with values in $X \times Y$ is called a bimartingale if:

(1) there exists an increasing sequence $(\mathcal{F}_n)_{n\geq 1}$ of finite sub-σ-algebra of $\mathcal{F}$ such that $(Z_n)_n$ is a $(\mathcal{F}_n)_{n\geq 1}$-martingale.
(2) $\forall n \geq 1, X_n = X_{n+1}$ a.s. or $Y_n = Y_{n+1}$ a.s.
(3) Z_1 is a.s. constant.

Notice that $(Z_n)_{n\geq 1}$ being a bounded martingale, it converges almost surely to a limit Z_∞.

Definition 13 Let A be a measurable subset of $X \times Y$.

$$A^* = \{z \in X \times Y, \text{ there exists a bimartingale } (Z_n)_{n\geq 1} \text{ converging to a limit } Z_\infty \text{ such that } Z_\infty \in A \text{ a.s. and } Z_1 = z \text{ a.s.}\} .$$

One can show that any atomless probability space $(\Omega, \mathcal{F}, \mathcal{P})$, or any product of convex compact spaces $X \times Y$ containing A, induce the same set A^*. One can also substitute condition (2) by: $\forall n \geq 1, (X_n = X_{n+1}$ or $Y_n = Y_{n+1})$ a.s. Notice that without condition (2), the set A^* would just be the convex hull of A.

We always have $A \subset A^* \subset \text{conv}(A)$, and these inclusions can be strict. For example, if $X = Y = [0, 1]$ and $A = \{(0, 0), (1, 0), (0, 1)\}$, it is possible to show that $A^* = \{(x, y) \in [0, 1] \times [0, 1], x = 0 \text{ or } y = 0\}$. A^* always is biconvex and thus contains $\text{biconv}(A)$, which is defined as the smallest biconvex set which contains A. The inclusion $\text{biconv}(A) \subset A^*$ can also be strict, as shown by the following example:

Example 5 Put $X = Y = [0, 1]$, $v_1 = (1/3, 0)$, $v_2 = (0, 2/3)$, $v_3 = (2/3, 1)$, $v_4 = (1, 1/3)$, $w_1 = (1/3, 1/3)$, $w_2 = (1/3, 2/3)$, $w_3 = (2/3, 2/3)$ et $w_4 = (2/3, 1/3)$, and $A = \{v_1, v_2, v_3, v_4\}$.

A is biconvex, so $A = \text{biconv}(A)$. Consider now the following Markov process $(Z_n)_{n\geq 1}$, with $Z_1 = w_1$. If $Z_n \in A$, then $Z_{n+1} = Z_n$. If $Z_n = w_i$ for some i, then $Z_{n+1} = w_{i+1(\text{mod } 4)}$ with probability 1/2, and $Z_{n+1} = v_i$ with probability 1/2. $(Z_n)_n$ is a bimartingale converging a.s. to a point in A, hence $w_1 \in A^* \backslash \text{biconv}(A)$.

We now present a geometric characterization of the set A^*, and assume here that A is closed. For each biconvex subset B of $X \times Y$ containing A, we denote by $nsc(B)$ the set of elements of B which can not be separated from A by

a continuous bounded biconvex function on A. More precisely, $nsc(B) = \{z \in B, \forall f\colon B \longrightarrow \mathbb{R}$ bounded biconvex, and continuous on A, $f(z) \leq \sup\{f(z'), z' \in A\}\}$.

Theorem 11 ([4]) *A^* is the largest biconvex set B containing A such that $nsc(B) = B$.*

Let us now come back to repeated games and to the notations of Subsect. "Approachability for Player 1 Versus Excludability for Player 2". To be precise, we need to add the component of player 2's payoff, and consequently to slightly modify the definitions. G is closed in $\mathbb{R}^K_M \times \mathbb{R}_M \times \Delta(K)$. For $B \subset \mathbb{R}^K_M \times \mathbb{R}_M \times \Delta(K)$, B is biconvex if for each a in $\mathbb{R}^K_M$ and for each p in $\Delta(K)$, the sections $\{(\beta, p'), (a, \beta, p') \in B\}$ and $\{(a', \beta), (a', \beta, p) \in B\}$ are convex. A real function f defined on a biconvex set B is said to be biconvex if $\forall a$, $\forall p$, $f(a, ., .)$ and $f(., ., p)$ are convex.

Theorem 12 ([4]) *G^* is the largest biconvex set B containing G such that: $\forall z \in B, \forall f\colon B \longrightarrow \mathbb{R}$ bounded biconvex, and continuous on A, $f(z) \leq \sup\{f(z'), z' \in G\}$.*

Non-observable Actions

We now consider the case where, as in the general definition of Sect. "Definition of the Subject", there is a signaling function $q\colon K \times A \longrightarrow \Delta(U)$ giving the distributions of the signals received by the players as a function of the state of nature and the action profile just played. The particular case where $q(k, a)$ does not depend on k is called *state independent signaling*. The previous models correspond to the particular case of perfect observation, where the signals received by the players exactly reveal the action profile played.

Theorem 1 has been generalized [6] to the general case of signaling function. We keep the notations of Sect. "The Standard Model of Aumann and Maschler". Given a mixed action $x \in \Delta(I)$, an action j in J and a state k, we denote by $Q(k, x, j)$ the marginal distribution on U^2 of the law $\sum_{i \in I} x(i)q(k, i, j)$, i. e. $Q(k, x, j)$ is the law of the signal received by player 2 if the state is k, player 1 uses x and player 2 plays j. The set of non revealing strategies of player 1 is then defined as:

$$\begin{aligned} NR(p) = \{x = (x^k)_{k \in K} \in \Delta(I)^K,\ & \\ \forall k \in K, \forall k' \in K \text{ s.t. } & p^k p^{k'} > 0, \\ \forall j \in J, Q(k, x^k, j) = Q(k', x^{k'}, j)\} & . \end{aligned}$$

If the initial probability is p and player 1 plays a strategy x in $NR(p)$ (i. e. plays x^k if the state is k), the a posteriori of player 2 will remain a.s. constant: player 2 can deduce no information on the selected state k. The value of the non revealing game becomes:

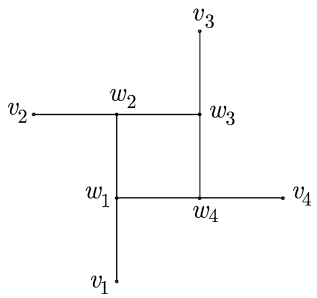
$$\begin{aligned} u(p) &= \max_{x \in NR(p)} \min_{y \in \Delta(J)} \sum_{k \in K} p^k G^k(x^k, y) \\ &= \min_{y \in \Delta(J)} \max_{x \in NR(p)} \sum_{k \in K} p^k G^k(x^k, y)\ , \end{aligned}$$

where $G^k(x^k, y) = \sum_{i,j} x^k(i)y(j)G^k(i, j)$, and the convention $u(p) = -\infty$ if $NR(p) = \emptyset$. Theorem 1 perfectly extends here: The repeated game with initial probability p has a uniform value given by $\mathrm{cav}u(p)$.

The explicit construction of an optimal strategy of player 2 (see Subsect. "Back to the Standard Model") has also been generalized to the general signaling case (see [35], and part B, p. 234 [57] for random signals).

Regarding zero-sum games with lack of information on both sides, the results of Sect. "Zero-Sum Games with Lack of Information on Both Sides" have been generalized to the case of state independent signaling (see [50,52] and [55]). Attention has been paid to the speed of convergence of the value function $(v_T)_T$, and bounds are identical for both models of lack of information on one side and on both sides, if we assume state independent signaling: this speed is of order $1/T^{1/2}$ for games with perfect observation, and of order $1/T^{1/3}$ for games with signals (these orders are optimal, both for lack of information on one side and lack of information on both sides, see [89,90]). For state dependent signaling and lack of information on one side, it was shown by Mertens [51] that the convergence occurs with worst case error $\sim (\ln n/n)^{1/3}$.

A particular class of zero-sum repeated games with state dependent signaling has been studied (games with no signals, see [54,88] and [82]). In these games, the state k is first selected according to a known probability and is not announced to the players; then after each stage both players receive the same signal which is either "nothing" or "the state is k". It is not possible to obtain here a standard recursive formula with state space $\Delta(K)$, or even



Repeated Games with Incomplete Information, Figure 5
The "four frogs" example of Aumann and Hart: $A^* \neq$ biconv(A)

$\Delta(K) \times \Delta(K)$, because even when the strategies are given, during the play none of the players can compute the a posteriori of the other player. It was shown that the maxmin and the minmax may differ, although $\lim_T v_T$ always exists.

In non zero-sum repeated games with lack of information on one side, the existence of "joint plan" equilibria have been generalized to the case of state independent signaling [62], and more generally to the case where "player 1 can send non revealing signals to player 2" [77]. The existence of a uniform equilibrium in the general signaling case is still an open question see [78].

Miscellaneous

Zero-Sum Games

In games with lack of information on one side, it is important that player 1 knows not only the selected state k, but also the a priori p. [84] provides an example of a game with lack of information on "one and a half" side with no uniform value. More precisely, in this example nature first chooses p in $\{p_1, p_2\}$ according to a known probability, and announces p to player 2 only; then k is selected according to p, and announced to player 1 only; finally the matrix game G^k is played.

For games with lack of information on one side, the value function v_T is a concave piecewise linear function of the initial probability p (see [61] for more generality). On the contrary, the discounted value v_λ can be quite a complex function of p: in example 2 of Sect. "Definition of the Subject", Mayberry [49] has proved that for $2/3 < \lambda < 1$, v_λ is, at each rational value of p, non differentiable.

Convergence of the value functions $(v_T)_T$ and $(v_\lambda)_\lambda$ have been widely studied. We have already discussed the speed of convergence in Sect. "Non-observable Actions", but much more can be said.

Example 6 Standard model of lack of information on one side and observable actions.

$$K = \{a, b\}\,,$$
$$G^a = \begin{pmatrix} 3 & -1 \\ -3 & 1 \end{pmatrix} \quad \text{and} \quad G^b = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}.$$

One can show [53] that for each $p \in [0, 1]$, viewed as the initial probability of state a, the sequence $\sqrt{T} v_T(p)$ converges to $\varphi(p)$, where $\varphi(p) = 1/\sqrt{2\pi}\, e^{-x_p^2/2}$, and x_p satisfies $1/\sqrt{2\pi} \int_{-\infty}^{x_p} e^{-x^2/2} dx = p$. So the limit of $\sqrt{T} v_T(p)$ is the standard normal density function evaluated at its p-quantile.

The apparition of the normal distribution is by no way an isolated phenomenon, but rather an important property of some repeated games ([12,13,14,15,18], ...).

B. de Meyer introduced the notion of "dual game" (see the previous references and also [17,19,41,69]). Let us now illustrate this on the standard model of Sect. "The Standard Model of Aumann and Maschler".

Let z be a parameter in $\mathbb{R}^K$. In the dual game $\Gamma_T^*(z)$, player 1 first secretly chooses the state k. Then at each stage $t \le T$, the players choose as usual actions i_t and j_t which are announced before proceeding to the next stage. With time horizon T, player 1's payoff finally is $\frac{1}{T}\sum_{t=1}^T G^k(i_t, j_t) - z^k$. This player is thus now able to fix the state equal to k, but has to pay z^k for it. It can be shown that the T-stage dual game $\Gamma_T^*(z)$ has a value $w_T(z)$. w_T is convex, and is linked to the value of the primal game by the conjugate formula:

$$w_T(z) = \max_{p \in \Delta(K)} (v_T(p) - \langle p, z \rangle), \quad \text{and}$$
$$v_T(p) = \inf_{z \in \mathbb{R}^K} (w_T(z) + \langle p, z \rangle)\,.$$

And $(w_T)_T$ satisfies the dual recursive formula:

$$w_{T+1}(z) = \min_{y \in \Delta(J)} \max_{i \in I} \frac{T}{T+1} w_T \cdot \left(\frac{T+1}{T} z - \frac{1}{T} \sum_{j \in J} y_j \left(G^k(i,j) \right)_k \right).$$

There are also strong relations between the optimal strategies of the players in the primal and dual games, and this gives a way to compute recursively optimal strategies of the uninformed player in the finite game (see also [32] on this topic).

Repeated games with incomplete information, as well as stochastic games, can also be studied in a functional analysis setup called the operator approach. This general approach is based on the study of the recursive formula [40,70,85].

In [65], the standard model, as well as the proof of Theorem 1, is generalized to the case where the state is not fixed at the beginning of the game, but evolves according to a Markov chain uniquely observed by player 1 (see also [59] for non observable actions, [48] and [33] for the difficulty of computing the value, [67] for the generalization to a state process controlled and observed by player 1, and [71] for several kinds of stochastic games with lack of information on one side). It is known since [80] that

the uniform value may not exist in general for stochastic games with lack of information on one side (where the stochastic game to be played is first randomly selected and announced to player 1 only).

Blackwell's approachability theorem has been extended to infinite dimensional spaces by Lehrer [43]. Approachability theory has strong links with the existence of no-regret strategies (first studied in [31], see also [10,26,30,44,72] and the recent book [9]), convergence of simple procedures to the set of correlated equilibria [31], and calibration [25,42]. The links between merging, reputation phenomena and repeated games with incomplete information have been studied in [83], where several existing results are unified. Finally, no-regret and approachability have also been studied when the players have bounded computational capacities (finite automata, bounded recall strategies) [45,46].

Let us mention also that de Meyer and Moussa Saley studied the modelization via Brownian motions in financial models [18]. They introduced a marked game based on a repeated game with lack of information on one side, and showed the endogenous apparition of a Brownian motion (see [16] for incomplete information on both sides).

Non Zero-Sum Games

In the setup of Sect. "Non Zero-Sum Games with Lack of Information on One Side", it is interesting to study the number of communication stages which is needed to construct the different equilibria. This number is linked with the convergence of the associated bimartingales (see [4,6,21,24]). Let us mention also that F. Forges [23] gave a similar characterization of equilibrium payoffs, for a larger notion of equilibria called communication equilibria (see also [22] for correlated equilibria). Amitai [2] studied the set of equilibrium payoffs in case of lack of information on both sides. Aumann and Hart [5] characterized the equilibrium payoffs in two player games with lack of information on one side when long, payoff-irrelevant, preplay communication is allowed (see [1] for incomplete information on both sides).

The particular case where each player knows his own payoffs is particularly worthwhile studying (*known own payoffs*). In the two-player case with lack of information on one side, this amounts to say that player 2's payoffs do not depend on the selected state. In this case, Shalev [73] showed that any equilibrium payoff can be obtained as the payoff of an equilibrium which is completely revealing. This result generalizes to the non zero-sum case of lack of information of both sides (see the unpublished manuscript [37]), but unfortunately uniform equilibria may fail to exist even though both players known their own payoffs.

Another model deals with the symmetric case, where the players have an incomplete, but identical, knowledge of the selected state. After each stage they receive the same signal, which may depend on the state. A. Neyman and S. Sorin have proved the existence of equilibrium payoffs in the case of two players (see [60], the zero-sum case being solved in [36] and [20]).

Few papers study the case of more than 2 players. The existence of uniform equilibrium has been studied for 3 players and lack of information on one side [63], and in the case of two states of nature it appears that a completely revealing equilibria, or a joint plan equilibria by one of the informed players, always exists. Concerning n-player repeated games with incomplete information and signals, several papers study how the initial information can be strategically transmitted, independently of the payoffs ([64,66], and [68] for an application to a cryptographic model). As an application, the existence of completely revealing equilibria is obtained in particular cases.

Repeated games with incomplete information have been used to study perturbations of repeated games with complete information (see [27] and [11] for Folk theorem-like results, [7] for enforcing cooperation in games with a Pareto-dominant outcome, and [34] for a perturbation with known own payoffs). The case where the players have different discount factors has also been investigated [11,47].

Future Directions

Several open problems are well formulated and deserve attention. Does a uniform equilibrium always exist in two-player repeated games with lack of information on one side and general signaling, or in n-player repeated games with lack of information on one side? More conceptually, one should look for classes of n-player repeated games with incomplete information which allow for the existence of equilibria, and/or for a tractable description of equilibrium payoffs (or at least of some of these payoffs). Regarding applications, there is certainly a lot of room in the vast fields of financial markets, cryptology, and sequential decision problems.

Acknowledgments

I thank Françoise Forges, Sergiu Hart, Dinah Rosenberg, Robert Simon and Eilon Solan for their comments on a preliminary version of this chapter.

Bibliography

Primary Literature

1. Amitai M (1996) Cheap-Talk with Incomplete Information on Both Sides. Ph D Thesis, The Hebrew University of Jerusalem, http://ratio.huji.ac.il/dp/dp90.pdf
2. Amitai M (1996) Repeated games with incomplete information on Both Sides. Ph D Thesis, The Hebrew University of Jerusalem, http://ratio.huji.ac.il/dp/dp105.pdf
3. Aumann RJ (1964) Mixed and behaviour strategies in infinite extensive games. In: Dresher, Shapley, Tucker (eds) Advances in game theory. Annals of Mathematics Study vol 52, Princeton University Press, Princeton, pp 627–650
4. Aumann RJ, Hart S (1986) Bi-convexity and bi-martingales. Israel J Math 54:159–180
5. Aumann RJ, Hart S (2003) Long cheap talk. Econometrica 71:1619–1660
6. Aumann RJ, Maschler M (1995) Repeated games with incomplete information, with the collaboration of Stearns RE. MIT Press, Cambridge (contains a reedition of chapters of Reports to the US Arms Control and Disarmament Agency ST-80, 116 and 143, Mathematica, 1966–1967–1968)
7. Aumann RJ, Sorin S (1989) Cooperation and bounded recall. Games Econ Behav 1:5–39
8. Blackwell D (1956) An analog of the minmax theorem for vector payoffs. Pac J Math 65:1–8
9. Cesa-Bianchi N, Lugosi G (2006) Prediction, Learning and Games. Cambridge University Press, Cambridge
10. Cesa-Bianchi N, Lugosi G, Stoltz G (2006) Regret minimization under partial monitoring. Math Oper Res 31:562–580
11. Cripps MW, Thomas JP (2003) Some asymptotic results in discounted repeated games of one-sided incomplete information. Math Oper Res 28:433–462
12. de Meyer B (1996) Repeated games and partial differential equations. Math Oper Res 21:209–236
13. de Meyer B (1996) Repeated games, duality and the central limit theorem. Math Oper Res 21:237–251
14. de Meyer B (1998) The maximal variation of a bounded martingale and the central limit theorem. Annales de l'Institut Henri Poincaré, Probabilités et statistiques 34:49–59
15. de Meyer B (1999) From repeated games to Brownian games. Annales de l'Institut Henri Poincaré, Probabilités et statistiques 35:1–48
16. de Meyer B, Marino A (2004) Repeated market games with lack of information on both sides. DP 2004.66, MSE Université Paris I
17. de Meyer B, Marino A (2005) Duality and optimal strategies in the finitely repeated zero-sum games with incomplete information on both sides. DP 2005.27, MSE Université Paris I
18. de Meyer B, Moussa H (2003) Saley On the strategic origin of Brownian motion in finance. Int J Game Theory 31:285–319
19. de Meyer B, Rosenberg D (1999) "*Cavu*" and the dual game. Math Oper Res 24:619–626
20. Forges F (1982) Infinitely repeated games of incomplete information: symmetric case with random signals. Int J Game Theory 11:203–213
21. Forges F (1984) A note on Nash equilibria in repeated games with incomplete information. Int J Game Theory 13:179–187
22. Forges F (1985) Correlated equilibria in a class of repeated games with incomplete information. Int J Game Theory 14:129–149
23. Forges F (1988) Communication equilibria in repeated games with incomplete information. Math Oper Res 13:191–231
24. Forges F (1990) Equilibria with communication in a job market example. Q J Econ 105:375–398
25. Foster D (1999) A Proof of Calibration via Blackwell's Approachability Theorem. Games Eco Behav 29:73–78
26. Foster D, Vohra R (1999) Regret in the On-Line Decision Problem. Games Econ Behav 29:7–35
27. Fudenberg D, Maskin E (1986) The folk theorem in repeated games with discounting or with incomplete information. Econometrica 54:533–554
28. Harsanyi J (1967–68) Games with incomplete information played by 'Bayesian' players, parts I-III. Manag Sci 8:159–182, 320–334, 486–502
29. Hart S (1985) Nonzero-sum two-person repeated games with incomplete information. Math Oper Res 10:117–153
30. Hart S (2005) Adaptative Heuristics. Econometrica 73:1401–1430
31. Hart S, Mas-Colell A (2000) A simple adaptative procedure leading to Correlated Equilibrium. Econometrica 68:1127–1150
32. Heuer M (1992) Optimal strategies for the uninformed player. Int J Game Theory 20:33–51
33. Hörner J, Rosenberg D, Solan E, Vieille N (2006) On Markov Games with Incomplete Information on One Side. Discussion paper 1412, Center for Mathematical Studies in Economics and Management Science, Northwestern University, Evanston
34. Israeli E (1999) Sowing Doubt Optimally in Two-Person Repeated Games. Games Econ Behav 28:203–216
35. Kohlberg E (1975) Optimal strategies in repeated games with incomplete information. Int J Game Theory 4:7–24
36. Kohlberg E, Zamir S (1974) Repeated games of Incomplete information: The Symmetric Case. Ann Stat 2:40–41
37. Koren G (1992) Two-person repeated games where players know their own payoffs. master thesis, Tel-Aviv University, Tel-Aviv http://www.ma.huji.ac.il/hart/papers/koren.pdf
38. Kuhn HW (1953) Extensive games and the problem of information. In: Kuhn, Tucker (eds) Contributions to the Theory of Games, vol II. Annals of Mathematical Studies 28, Princeton University Press, Princeton, pp 193–216
39. Laraki R (2001) Variational inequalities, system of functional equations and incomplete information repeated games. SIAM J Control Optim 40:516–524
40. Laraki R (2001) The splitting game and applications. Int J Game Theory 30:359–376
41. Laraki R (2002) Repeated games with lack of information on one side: the dual differential approach. Math Oper Res 27:419–440
42. Lehrer E (2001) Any inspection is manipulable. Econometrica 69:1333–1347
43. Lehrer E (2003) Approachability in infinite dimensional spaces. Int J Game Theory 31:253–268
44. Lehrer E (2003) A wide range no-regret theorem. Games Econ Behav 42:101–115
45. Lehrer E, Solan E (2003) No regret with bounded computational capacity. DP 1373, Center for Mathematical Studies in Economics and Management Science, Northwestern University, Evanston
46. Lehrer E, Solan E (2006) Excludability and bounded computational capacity. Math Oper Res 31:637–648

47. Lehrer E, Yariv L (1999) Repeated games with lack of information on one side: the case of different discount factors. Math Oper Res 24:204–218
48. Marino A (2005) The value of a particular Markov chain game. Chaps 5 and 6, Ph D thesis, Université Paris I http://alexandre.marino.free.fr/theseMarino.pdf
49. Mayberry J-P (1967) Discounted repeated games with incomplete information. Report of the US Arms control and disarmament agency, ST116, Chap V. Mathematica, Princeton, pp 435–461
50. Mertens J-F (1972) The value of two-person zero-sum repeated games: the extensive case. Int J Game Theory 1:217–227
51. Mertens J-F (1998) The speed of convergence in repeated games with incomplete information on one side. Int J Game Theory 27:343–357
52. Mertens J-F, Zamir S (1971) The value of two-person zero-sum repeated games with lack of information on both sides. Int J Game Theory 1:39–64
53. Mertens J-F, Zamir S (1976) The normal distribution and repeated games. Int J Game Theory 5:187–197
54. Mertens J-F, Zamir S (1976) On a repeated game without a recursive structure. Int J Game Theory 5:173–182
55. Mertens J-F, Zamir S (1977) A duality theorem on a pair of simultaneous functional equations. J Math Anal Appl 60:550–558
56. Mertens J-F, Zamir S (1985) Formulation of Bayesian analysis for games with incomplete information. Int J Game Theory 14:1–29
57. Mertens J-F, Sorin S, Zamir S (1994) Repeated games. CORE discussion paper. Université Catholique de Louvain, Belgium, pp 9420–9422
58. Meyer PA (1966) Probability and Potentials. Blaisdell, New York (in French: Probabilités et potentiel, Hermann, 1966)
59. Neyman A (2008) Existence of Optimal Strategies in Markov Games with Incomplete Information. Int J Game Theory 37:581–596
60. Neyman A, Sorin S (1998) Equilibria in Repeated Games with Incomplete Information: The General Symmetric Case. Int J Game Theory 27:201–210
61. Ponssard JP, Sorin S (1980) The LP formulation of finite zero-sum games with incomplete information. Int J Game Theory 9:99–105
62. Renault J (2000) 2-player repeated games with lack of information on one side and state independent signalling. Math Oper Res 4:552–572
63. Renault J (2001) 3-player repeated games with lack of information on one side. Int J Game Theory 30:221–246
64. Renault J (2001) Learning sets in state dependent signalling game forms: a characterization. Math Oper Res 26:832–850
65. Renault J (2006) The value of Markov chain games with lack of information on one side. Math Oper Res 31:490–512
66. Renault J, Tomala T (2004) Learning the state of nature in repeated games with incomplete information and signals. Games Econ Behav 47:124–156
67. Renault J (2007) The value of Repeated Games with an informed controller. Technical report, Université Paris–Dauphine, Ceremade
68. Renault J, Tomala T (2008) Probabilistic reliability and privacy of communication using multicast in general neighbor networks. J Cryptol 21:250–279
69. Rosenberg D (1998) Duality and Markovian strategies. Int J Game Theory 27:577–597
70. Rosenberg D, Sorin S (2001) An operator approach to zero-sum repeated games. Israel J Math 121:221–246
71. Rosenberg D, Solan E, Vieille N (2004) Stochastic games with a single controller and incomplete information. SIAM J Control Optim 43:86–110
72. Rustichini A (1999) Minimizing Regret: The General Case. Games Econ Behav 29:224–243
73. Shalev J (1994) Nonzero-Sum Two-Person Repeated Games with Incomplete Information and Known-Own Payoffs. Games Econ Behav 7:246–259
74. Sion M (1958) On General Minimax Theorems. Pac J Math 8:171–176
75. Simon RS (2002) Separation of joint plan equilibrium payoffs from the min-max functions. Games Econ Beha 1:79–102
76. Simon RS, Spież S, Toruńczyk H (1995) The existence of equilibria in certain games, separation for families of convex functions and a theorem of Borsuk–Ulam type. Israel J Math 92:1–21
77. Simon RS, Spież S, Toruńczyk H (2002) Equilibrium existence and topology in some repeated games with incomplete information. Trans AMS 354:5005–5026
78. Simon RS, Spież S, Toruńczyk H (2008) Equilibria in a class of games and topological results implying their existence. RACSAM Rev R Acad Cien Serie A Mat 102:161–179
79. Sorin S (1983) Some results on the existence of Nash equilibria for non- zero sum games with incomplete information. Int J Game Theory 12:193–205
80. Sorin S (1984) Big match with lack of information on one side, Part I. Int J Game Theory 13:201–255
81. Sorin S (1984) On a pair of simultaneous functional equations. J Math Anal Appl 98:296–303
82. Sorin S (1989) On recursive games without a recursive structure: existence of lim v_n. Int J Game Theory 18:45–55
83. Sorin S (1997) Merging, reputation, and repeated games with incomplete information. Games Econ Behav 29:274–308
84. Sorin S, Zamir S (1985) A 2-person game with lack of information on 1 and 1/2 sides. Math Oper Res 10:17–23
85. Sorin S (2002) A first course on zero-sum repeated games. Mathématiques et Applications, vol 37. Springer
86. Spinat X (2002) A necessary and sufficient condition for approachability. Math Oper Res 27:31–44
87. Vieille N (1992) Weak approachability. Math Oper Res 17:781–791
88. Waternaux C (1983) Solution for a class of repeated games without recursive structure. Int J Game Theory 12:129–160
89. Zamir S (1971) On the relation between finitely and infinitely repeated games with incomplete information. Int J Game Theory 1:179–198
90. Zamir S (1973) On repeated games with general information function. Int J Game Theory 21:215–229
91. Zamir S (1992) Repeated Games of Incomplete Information: zero-sum. In: Aumann RJ, Hart S (eds) Handbook of Game Theory, I. Elsevier Science Publishers, Amsterdam, pp 109–154

Books and Reviews

Forges F (1992) Repeated Games of Incomplete Information: Non-zero sum. In: Aumann RJ, Hart S (eds) Handbook of Game Theory, vol I. Elsevier Science Publishers, Amsterdam, pp 155–177

Laraki R, Renault J, Tomala T (2006) Théorie des Jeux, Introduction à la théorie des jeux répétés. Editions de l'Ecole Polytechnique, Journées X-UPS, in French (Chapter 3 deals with repeated games with incomplete information). Palaiseau

Mertens J-F (1987) Repeated games. In: Proceedings of the International Congress of Mathematicians 1986. American Mathematical Society, Berkeley, pp 1528–1577

Reputation Effects

George J. Mailath
Department of Economics, University of Pennsylvania, Philadelphia, USA

Article Outline

Glossary

Action type A type of player who is committed to playing a particular action, also called a *commitment type* or *behavioral type*.

Complete information Characteristics of all players are common knowledge.

Flow payoff Stage game payoff.

Imperfect monitoring Past actions of all players are not public information.

Incomplete information Characteristics of some player are not common knowledge.

Long-lived player Player subject to intertemporal incentives, typically has the same horizon as length of the game.

Myopic optimum An action maximizing stage game payoffs.

Nash equilibrium A strategy profile from which no player has a profitable unilateral deviation (i. e., it is self-enforcing).

Nash reversion In a repeated game, permanent play of a stage game Nash equilibrium.

Normalized discounted value The discounted sum of an infinite sequence $\{a_t\}_{t\geq 0}$, calculated as $(1-\delta)\sum_{t\geq 0}\delta^t a_t$, where $\delta \in (0,1)$ is the discount value.

Perfect monitoring Past actions of all players are public information.

Repeated game The finite or infinite repetition of a stage game.

Reputation bound The lower bound on equilibrium payoffs of a player that the other player(s) believe may be a simple action type (typically the Stackelberg type).

Short-lived player Player not subject to intertemporal incentives, having a one-period horizon and so is myopically optimizing.

Simple action type An action who plays the same (pure or mixed) stage-game action in every period, regardless of history.

Stage game A game played in one period.

Stackelberg action In a stage game, the action a player would commit to, if that player had the chance to do so, i. e., the optimal commitment action.

Stackelberg type A simple action type that plays the Stackelberg action.

Subgame In a repeated game with perfect monitoring, the game following any history.

Subgame perfect equilibrium A strategy profile that induces a Nash equilibrium on every subgame of the original game.

Type The characteristic of a player that is not common knowledge.

Definition of the Subject

Repeated games have many equilibria, including the repetition of stage game Nash equilibria. At the same time, particularly when monitoring is imperfect, certain plausible outcomes are not consistent with equilibrium. *Reputation effects* is the term used for the impact upon the set of equilibria (typically of a repeated game) of perturbing the game by introducing incomplete information of a particular kind. Specifically, the characteristics of a player are not public information, and the other players believe it is possible that the distinguished player is a type that necessarily plays some action (typically the Stackelberg action). Reputation effects fall into two classes: "Plausible" phenomena that are not equilibria of the original repeated game are equilibrium phenomena in the presence of incomplete information, and "implausible" equilibria of the original game are not equilibria of the incomplete information game. As such, reputation effects provide an important qualification to the general indeterminacy of equilibria.

Introduction

Repeating play of a stage game often allows for equilibrium behavior inconsistent with equilibrium of that stage game. If the stage game has multiple Nash equilibrium

	C	D
C	$2, 2$	$-1, 3$
D	$3, -1$	$0, 0$

Reputation Effects, Figure 1
The prisoners' dilemma. The cooperative action is labeled *C*, while defect is labeled *D*

payoffs, a large finite number of repetitions provide sufficient intertemporal incentives for behavior inconsistent with stage-game Nash equilibria to arise in some subgame perfect equilibria. However, many classic games do not have multiple Nash equilibria. For example, mutual defection *DD* is the unique Nash equilibrium of the prisoners' dilemma, illustrated in Fig. 1.

A standard argument shows that the finitely repeated prisoner's dilemma has a unique subgame perfect equilibrium, and in this equilibrium, *DD* is played in every period: In any subgame perfect equilibrium, in the last period, *DD* must be played independently of history, since the stage game has a unique Nash equilibrium. Then, since play in the last period is independent of history, there are no intertemporal incentives in the penultimate period, and so *DD* must again be played independently of history. Proceeding recursively, *DD* must be played in every period independently of history. (In fact, the finitely repeated prisoners' dilemma has a unique Nash equilibrium outcome, given by *DD* in every period.)

This contrasts with intuition, which suggests that if the prisoners' dilemma were repeated a sufficiently large (though finite) number of times, the two players would find a way to play cooperatively (*C*) at least in the initial stages. In response, Kreps, Milgrom, Roberts and Wilson [15] argued that intuition can be rescued in the finitely repeated prisoners' dilemma by introducing incomplete information. In particular, suppose each player assigns some probability to their opponent being a behavioral type who mechanistically plays tit-for-tat (i. e., plays *C* in the first period or if the opponent had played *C* in the previous period, and plays *D* if the opponent had played *D* in the previous period) rather than being a rational player. No matter how small the probability, if the number of repetitions is large enough, the rational players will play *C* in early periods, and the fraction of periods in which *CC* is played is close to one.

This is the first example of a *reputation effect*: a small degree of incomplete information (of the right kind) both rescues the intuitive *CC* for many periods as an equilibrium outcome, and eliminates the unintuitive always *DD* as one. In the same issue of the *Journal of Economic Theory* containing Kreps, Milgrom, Roberts and Wilson [15], Kreps and Wilson[14] and Milgrom and Roberts [18] explored reputation effects in the finite chain store of Selten [22], showing that intuition is again rescued, this time by introducing the possibility that the chain store is a "tough" type who always fights entry.

Reputation effects describe the impact upon the set of equilibria of the introduction of small amounts of incomplete information of a particular form into repeated games (and other dynamic games). Reputation effects fall into two classes: "Plausible" phenomena that are not equilibria of the complete information game are equilibrium phenomena in the presence of incomplete information, and "implausible" equilibria of the complete information game are not equilibria of the incomplete information game.

Reputation effects are distinct from the equilibrium phenomenon in complete information repeated games that are sometimes described as capturing *reputations*. In this latter use, an equilibrium of the complete information repeated game is selected, involving actions along the equilibrium path that are not Nash equilibria of the stage game. As usual, incentives to choose these actions are created by attaching less favorable continuation paths to deviations. Players who choose the equilibrium actions are then interpreted as maintaining a reputation for doing so, with a punishment-triggering deviation interpreted as causing the loss of one's reputation. For example, players who cooperate in the *infinitely* repeated prisoners' dilemma are interpreted as having (or maintaining) a cooperative reputation, with any defection destroying that reputation. In this usage, the link between past behavior and expectations of future behavior is an equilibrium phenomenon, holding in some equilibria, but not in others. The notion of reputation is used to interpret an equilibrium strategy profile, but otherwise adds nothing to the formal analysis.

In contrast, the approach underlying reputation effects begins with the assumption that a player is uncertain about key aspects of her opponent. For example, player 2 may not know player 1's payoffs, or may be uncertain about what constraints player 1 faces on his ability to choose various actions. This incomplete information is a device that introduces an intrinsic connection between past behavior and expectations of future behavior. Since incomplete information about players' characteristics can have dramatic effects on the *set* of equilibrium payoffs, reputations in this approach do not describe certain equilibria, but rather place constraints on the set of possible equilibria.

An Example

While reputation effects were first studied in a symmetric example with two long-lived players, they arise in their

	h	ℓ
H	2, 3	0, 2
L	3, 0	1, 1

Reputation Effects, Figure 2
The product-choice game

purest form in infinitely repeated games with one long-lived player playing against a sequence of short-lived players. The chain store game of Selten [22] is a finitely repeated game in which a chain store (the long-lived player) faces a finite sequence of potential entrants in its different markets. Since each entrant only cares about its own decision, it is short-lived.

Consider the "product-choice" game of Fig. 2. The row player (player 1), who is long-lived, is a firm choosing between high (H) and low (L) effort, while the column player (player 2), who is short-lived, is a customer choosing between a high (h) or low (ℓ) priced product. (Mailath and Samuelson [17] illustrate various aspects of repeated games and reputation effects using this example.) Player 2 prefers the high-priced product if the firm has exerted high effort, but prefers the low-priced product if the firm has not. The firm prefers that customers purchase the high-priced product and is willing to commit to high effort to induce that choice by the customer. In a simultaneous move game, however, the firm cannot observably choose effort before the customer chooses the product. Since high effort is costly, the firm prefers low effort, no matter the choice of the customer.

The stage game has a unique Nash equilibrium, in which the firm exerts low effort and the customer purchases the low-priced product. Suppose the game is played infinitely often, with *perfect monitoring* (i. e., the history of play is public information). The firm is *long-lived* and discounts flow profits by the discount factor $\delta \in (0, 1)$, and is *patient* if δ is close to 1. The role of the customer is taken by a succession of *short-lived* players, each of whom plays the game only once (and so myopically optimizes). It is standard to abuse language by treating the collection of short-lived players as a single myopically optimizing player.

When the firm is sufficiently patient, there is an equilibrium outcome in the repeated game in which the firm always exerts high effort and customers always purchase the high-priced product. The firm is deterred from taking the immediate myopically optimal action of low effort by the prospect of future customers then purchasing the low-priced product. Purchasing the high-priced product is a best response for the customer to high effort, so that no incentive issues arise concerning the customer's behavior. In this equilibrium, the long-lived player's payoff is 2 (the firm's payoffs are calculated as the normalized discounted sum, i. e., as the discounted sum of flow payoffs normalized by $(1 - \delta)$, so that payoffs in the infinite horizon game are comparable to flow payoffs). However, there are many other equilibria, including one in which low effort is exerted and low price purchased in every period, leading to a payoff of 1 for the long-lived player. Indeed, for $\delta \geq 1/2$, the set of pure-strategy subgame-perfect-equilibrium player 1 payoffs is given by the entire interval $[1, 2]$.

Reputation effects effectively rule out any payoff less than 2 as an equilibrium payoff for player 1. Suppose customers are not entirely certain of the characteristics of the firm. More specifically, suppose they attach high probability to the firm's being "normal," that is, having the payoffs given above, but they also entertain some (possibly very small) probability that they face a firm who fortuitously has a technology or some other characteristic that ensures high effort. Refer to the latter as the "H-action" type of firm. Since such a type necessarily plays H in every period, it is a type described by behavior (not payoffs), and such a type is often called a *behavioral* or *commitment* type.

This is now a game of *incomplete information*, with the customers uncertain of the firm's type. Since the customers assign high probability to the firm being "normal," the game is in some sense close to the game of complete information. None the less, reputation effects are present: For a sufficiently patient firm, in any Nash equilibrium of the repeated game, the firm's payoff cannot be significantly less than 2. This result holds no matter how unlikely customers think the H-action type to be, though increasing patience is required from the normal firm as the action type becomes less likely.

The intuition behind this result is most easily seen by considering pure strategy Nash equilibria of the incomplete information game where the customers believe the firm is either the normal or the H-action type. In that case, there is no pure strategy Nash equilibrium with a payoff less than 2δ (which is clearly close to 2 for δ close to 1). In the pure strategy Nash equilibrium, either the firm always plays H, (in which case, the customers always play h and the firm's payoff is 2), or there is a first period (say t) in which the firm plays L, revealing to future customers that he is the normal type (since the action type plays H in every period). In such an equilibrium, customers play h before t (since both types of firm are choosing H). After observing H in period t, customers conclude the firm is the H-action type. Consequently, as long as H is always chosen thereafter, customers subsequently play h (since they continue to believe the firm is the H-action type, and

so necessarily plays H). An easy lower bound on the normal firm's equilibrium payoff is then obtained by observing that the normal firm's payoff must be at least the payoff from mimicking the action type in every period. The payoff from such behavior is at least as large as

$$\underbrace{(1-\delta)\sum_{\tau=0}^{t-1}\delta^{\tau}2}_{\substack{\text{payoff in } \tau < t \text{ from pooling}\\ \text{with } H\text{-action type}}} + \underbrace{(1-\delta)\delta^{t}\times 0}_{\substack{\text{payoff in } t \text{ from playing } H \text{ when}\\ L \text{ may be myopically optimal}}} + \underbrace{(1-\delta)\sum_{\tau=t+1}^{\infty}\delta^{\tau}2}_{\substack{\text{payoff in } \tau > t \text{ from playing like}\\ \text{and being treated as the } H\text{-action type}}}$$

$$= (1-\delta^{t})2 + \delta^{t+1}2$$
$$= 2 - 2\delta^{t}(1-\delta)$$
$$\geq 2 - 2(1-\delta) = 2\delta\,.$$

The outcome in which the stage game Nash equilibrium $L\ell$ is played in every period is thus eliminated.

Since reputation effects are motivated by the hypothesis that the short-lived players are uncertain about some aspect of the long-lived player's characteristics, it is important that the results are not sensitive to the precise nature of that uncertainty. In particular, the lower bound on payoffs should not require that the short-lived players *only* assign positive probability to the normal and the H-action type (as in the game just analyzed). And it does not: The customers in the example may assign positive probability to the firm being an action type that plays H on even periods, and L on odd periods, as well as to an action type that plays H in every period before some period t' (that can depend on history), and then always plays L. Yet, as long as the customers assign positive probability to the H-action type, for a sufficiently patient firm, in any Nash equilibrium of the repeated game, the firm's payoff cannot be significantly less than 2.

Reputation effects are more powerful in the presence of imperfect monitoring. Suppose that the firm's choice of H or L is not observed by the customers. Instead, the customers observe a public signal $y \in \{y, \bar{y}\}$ at the end of each period, where the signal $\bar{y}$ is realized with probability $p \in (0,1)$ if the firm chose H, and with the smaller probability $q \in (0,p)$ if the firm chose L. Interpret $\bar{y}$ as a good meal: while customers do not observe effort, they do observe a noisy signal (the quality of the meal) of that effort, with high effort leading to a good meal with higher probability. In the game with complete information, the largest equilibrium payoff to the firm is now given by

$$\bar{v}_1 \equiv 2 - \frac{1-p}{p-q}\,, \tag{1}$$

reflecting the imperfect monitoring of the firm's actions (the firm is said to be subject to binding moral hazard, see Sect. 7.6 in [17]). Since deviations from H cannot be detected for sure, there are no equilibria with the deterministic outcome path of Hh in every period. In some periods after some histories, $L\ell$ must be played in order to provide the appropriate intertemporal incentives to the firm.

As under perfect monitoring, as long as customers assign positive probability to the H-action type in the incomplete information game with imperfect monitoring, for a sufficiently patient firm, in any Nash equilibrium of the repeated game, the firm's payoff cannot be significantly less than 2 (in particular, this lower bound exceeds $\bar{v}_1$). Thus, in this case, reputation effects provide an intuitive lower bound on equilibrium payoffs that both rules out "bad" equilibrium payoffs, as well as rescues outcomes in which Hh occurs in most periods.

Proving that a reputation bound holds in the imperfect monitoring case is considerably more involved than in the perfect monitoring case. In perfect-monitoring games, it is only necessary to analyze the evolution of the customers' beliefs when always observing H, the action of the H-action type. In contrast, imperfect monitoring requires consideration of belief evolution on all histories that arise with positive probability.

None the less, the intuition is the same: Consider a putative equilibrium in which the normal firm receives a payoff less than $2-\varepsilon$. Then the normal and action types must be making different choices over the course of the repeated game, since an equilibrium in which they behave identically would induce customers to choose h and would yield a payoff of 2. As in the perfect monitoring case, the normal firm has the option of mimicking the behavior of the H-action type. Suppose the normal firm does so. Since the customers expect the normal type of firm to behave differently from the H-action type, they will more often see signals indicative of the H-action type (rather than the normal type), and so must eventually become convinced that the firm is the H-action type. Hence, in response to this deviation, the customers will eventually play their best response to H of h. While "eventually" may take a while, that time is independent of the equilibrium (indeed of the discount factor), depending only on the imperfection in the monitoring and the prior probability assigned to the H-action type. Then, if the firm is sufficiently patient, the payoff from mimick-

ing the H-action type is arbitrarily close to 2, contradicting the existence of an equilibrium in which the firm's payoff fell short of $2 - \varepsilon$.

At the same time, because monitoring is imperfect, as discussed in Sect. "Temporary Reputation Effects", the reputation effects are necessarily transient. Under general conditions in imperfect-monitoring games, the incomplete information that is at the core of reputation effects is a short-run phenomenon. Player 2 must eventually come to learn player 1's type and continuation play must converge to an equilibrium of the complete information game.

Reputation effects arise for very general specifications of the incomplete information as long as the customers assign strictly positive probability to the H-action type. It is critical, however, that the customers do assign strictly positive probability to the H-action type. For example, in the product-choice game, the set of Nash equilibria of the repeated game is not significantly impacted by the possibility that the firm is either normal or the L-action type only. While reputation effects per se do not arise from the L-action type, it is still of interest to investigate the impact of such uncertainty on behavior using stronger equilibrium notions, such as Markov perfection (see Mailath and Samuelson [16]).

A Canonical Model

The Stage Game

The stage game is a two-player simultaneous-move finite game of public monitoring. Player i has action set A_i, $i = 1, 2$. Pure actions for player i are denoted by $a_i \in A_i$, and mixed actions are denoted by $\alpha_i \in \Delta(A_i)$, where $\Delta(A_i)$ is the set of probability distributions over A_i. Player 2's actions are public, while player 1's are potentially private. The public signal of player 1's action, denoted by y is drawn from a finite set Y, with the probability that y is realized under the pure action profile $a \in A \equiv A_1 \times A_2$ denoted by $\rho(y \mid a)$. Player 1's ex post payoff from the action profile a and signal realization y is $r_1(y, a)$, and so the ex ante (or expected) flow payoff is $u_1(a) \equiv \sum_y r_1(y, a)\rho(y \mid a)$. Player 2's ex post payoff from the action profile a and signal realization y is $r_2(y, a_2)$, and so the ex ante (or expected) flow payoff is $u_2(a) \equiv \sum_y r_2(y, a_2)\rho(y \mid a)$. Since player 2's ex post payoff is independent of player 1's actions, player 1's actions only affect player 2's payoffs through the impact on the distribution of the signals and so on ex ante payoffs. While the ex post payoffs r_i play no explicit role in the analysis, they justify the informational assumptions to be made. In particular, the model requires that histories of signals and past actions are the only information players receive, and so it is important that stage game payoffs u_i are not informative about the action choice (and this is the critical feature delivered by the assumptions that ex ante payoffs are not observable and that payer 2's ex post payoffs do *not* depend on a_1).

Perfect monitoring is the special case where $Y = A_1$ and $\rho(y \mid a) = 1$ if $y = a_1$, and 0 otherwise.

The results in this section hold under significantly weaker monitoring assumptions. In particular, it is not necessary that the actions of player 2 be public. If these are also imperfectly monitored, then the ex post payoff for player 1 is independent of player 2 actions. Since player 2 is short-lived, when player 2's actions are not public, it is then natural to also assume that the period t player 2 does not know earlier player 2's actions.

The Complete Information Repeated Game

The stage game is infinitely repeated. Player 1 is long-lived, with payoffs given by the normalized discounted value $(1 - \delta) \sum_{t=0}^{\infty} \delta^t u_1^t$, where $\delta \in (0, 1)$ is the discount factor and u_1^t is player 1's period t flow payoff. Player 1 is *patient* if δ is close to 1. As in our example, the role of player 2 is taken by a succession of *short-lived* players, each of whom plays the game only once (and so myopically optimizes).

Player 1's set of private histories is $\mathcal{H}_1 \equiv \cup_{t=0}^{\infty}(Y \times A)^t$ and the set of public histories (which coincides with the set of player 2's histories) is $\mathcal{H} \equiv \cup_{t=0}^{\infty}(Y \times A_2)^t$. If the game has perfect monitoring, histories $h = (y^0, a^0; y^1, a^1; \ldots; y^{t-1}, a^{t-1})$ in which $y^\tau \neq a_1^\tau$ for some $\tau \leq t - 1$ arise with zero probability, independently of behavior, and so can be ignored. A strategy σ_1 for player 1 specifies a probability distribution over 1's pure action set for each possible private history, i. e., $\sigma_1 : \mathcal{H}_1 \to \Delta(A_1)$. A strategy σ_2 for player 2 specifies a probability distribution over 2's pure action set for each possible public history, i. e., $\sigma_2 : \mathcal{H} \to \Delta(A_2)$.

Definition 1 The strategy profile (σ_1^*, σ_2^*) is a *Nash equilibrium* if

1. there does not exist a strategy σ_1 yielding a strictly higher payoff for player 1 when player 2 plays σ_2^*, and
2. in all periods t, after any history $h^t \in \mathcal{H}$ arising with positive probability under (σ_1^*, σ_2^*), $\sigma_2^*(h^t)$ maximizes $E[u_2(\sigma_1^*(h_1^t), a_1) \mid h^t]$, where the expectation is taken over the period t-private histories that player 1 may have observed.

The Incomplete Information Repeated Game

In the incomplete information game, the type of player 1 is unknown to player 2. A possible type of player 1 is denoted by $\xi \in \Xi$, where Ξ is a finite or countable set (see Fudenberg and Levine [12] for the uncountable case). Player 2's prior belief about 1's type is given by the distribution μ, with support Ξ. The set of types is partitioned into a set of *payoff types* Ξ_1, and a set of *action types* $\Xi_2 \equiv \Xi \setminus \Xi_1$. Payoff types maximize the average discounted value of payoffs, which depend on their type and which may be nonstationary,

$$u_1 : A_1 \times A_2 \times \Xi_1 \times \mathbb{N}_0 \to \mathbb{R} .$$

Type $\xi_0 \in \Xi_1$ is the *normal type* of player 1, who happens to have a stationary payoff function, given by the stage game in the benchmark game of complete information,

$$u_1(a, \xi_0, t) = u_1(a) \quad \forall a \in A, \forall t \in \mathbb{N}_0 .$$

It is standard to think of the prior probability $\mu(\xi_0)$ as being relatively large, so the games of incomplete information are a seemingly small departure from the underlying game of complete information, though there is no requirement that this be the case.

Action types (also called *commitment* or *behavioral types*) do not have payoffs, and simply play a specified repeated game strategy. For any repeated-game strategy from the complete information game, $\hat{\sigma}_1 : \mathcal{H}_1 \to \Delta(A_1)$, denote by $\xi(\hat{\sigma}_1)$ the action type committed to the strategy $\hat{\sigma}_1$. In general, a commitment type of player 1 can be committed to any strategy in the repeated game. If the strategy in question plays the same (pure or mixed) stage-game action in every period, regardless of history, that type is called a *simple action type*. For example, the H-action type in the product-choice game is a simple action type. The (simple action) type that plays the pure action a_1 in every period is denoted by $\xi(a_1)$ and similarly the simple action type committed to $\alpha_1 \in \Delta(A_1)$ is denoted by $\xi(\alpha_1)$. As will be seen soon, allowing for mixed action types is an important generalization from simple pure types.

A strategy for player 1, also denoted by $\sigma_1 : \mathcal{H}_1 \times \Xi \to \Delta(A_1)$, specifies for each type $\xi \in \Xi$ a repeated game strategy such that for all $\xi(\hat{\sigma}_1) \in \Xi_2$, the strategy $\hat{\sigma}_1$ is specified. A strategy σ_2 for player 2 is as in the complete information game, i. e., $\sigma_2 : \mathcal{H} \to \Delta(A_2)$.

Definition 2 The strategy profile (σ_1^*, σ_2^*) is a *Nash equilibrium* of the incomplete information game if

1. for all $\xi \in \Xi_1$, there does not exist a repeated game strategy σ_1 yielding a strictly higher payoff for payoff type ξ of player 1 when player 2 plays σ_2^*, and
2. in all periods t, after any history $h^t \in \mathcal{H}$ arising with positive probability under (σ_1^*, σ_2^*) and $\mu, \sigma_2^*(h^t)$ maximizes $E[u_2(\sigma_1^*(h_1^t, \xi), a_1) \mid h^t]$, where the expectation is taken over both the period t-private histories that player 1 may have observed and player 1's type.

Example 1 Consider the product-choice game (Fig. 2) under perfect monitoring. The firm is willing to commit to H to induce h from customers. This incentive to commit is best illustrated by considering a sequential version of the product-choice game: The firm first publicly commits to an effort, and then the customer chooses between h and ℓ, knowing the firm's choice. In this sequential game, the firm chooses H in the unique subgame perfect equilibrium. Since Stackelberg [23] was the first investigation of such leader-follower interactions, it is traditional to call H the Stackelberg action, and the H-action type of player 1 the Stackelberg type, with associated Stackelberg payoff 2. Suppose $\Xi = \{\xi_0, \xi(H), \xi(L)\}$. For $\delta \geq 1/2$, the grim trigger strategy profile of always playing Hh, with deviations punished by Nash reversion, is a subgame perfect equilibrium of the complete information game. Consider the following adaptation of this profile in the incomplete information game:

$$\sigma_1(h^t, \xi) = \begin{cases} H, & \text{if } \xi = \xi(H), \\ & \quad \text{or } \xi = \xi_0 \text{ and } a^\tau = Hh \\ & \qquad\qquad \text{for all } \tau < t, \\ L, & \text{otherwise,} \end{cases}$$

and

$$\sigma_2(h^t) = \begin{cases} h, & \text{if } a^\tau = Hh \text{ for all } \tau < t, \\ \ell, & \text{otherwise .} \end{cases}$$

In other words, player 2 and the normal type of player 1 follow the strategies from the Nash-reversion equilibrium in the complete information game, and the action types $\xi(H)$ and $\xi(L)$ play their actions.

This is a Nash equilibrium for $\delta \geq 1/2$ and $\mu(\xi(L)) < 1/2$. The restriction on $\mu(\xi(L))$ ensures that player 2 finds h optimal in period 0. Should player 2 ever observe L, then Bayes' rule causes her to place probability 1 on type $\xi(L)$ (if L is observed in the first period) or the normal type (if L is first played in a subsequent period), making her participation in Nash reversion optimal. The restriction on δ ensures that Nash reversion provides sufficient incentive to make H optimal for the normal player 1. After observing $a_1^0 = H$ in period 0, player 2 assigns zero probability to $\xi = \xi(L)$. However, the posterior probability that 2 assigns to the Stackelberg type does not converge to 1. In period 0, the prior probability is $\mu(\xi(H))$.

After one observation of H, the posterior increases to $\mu(\xi^*)/[\mu(\xi^*)+\mu(\xi_0)]$, after which it is constant. By stipulating that an observation of H in a history in which L has previously been observed causes player 2 to place probability one on the normal type of player 1, a specification of player 2's beliefs that is consistent with sequentiality is obtained.

As seen in the introduction, for δ close to 1, $\sigma_1(h^t, \xi_0) = L$ for all h^t is not part of any Nash equilibrium.

The Reputation Bound

Which type would the normal type most like to be treated as? Player 1's *pure-action Stackelberg payoff* is defined as

$$v_1^* = \sup_{a_1 \in A_1} \min_{\alpha_2 \in B(a_1)} u_1(a_1, \alpha_2), \quad (2)$$

where $B(a_1) = \arg\max_{a_2} u_2(a_1, a_2)$ is the set of player 2 myopic best replies to a_1. If the supremum is achieved by some action a_1^*, that action is an associated Stackelberg action,

$$a_1^* \in \arg\max_{a_1 \in A_1} \min_{\alpha_2 \in B(a_1)} u_1(a_1, \alpha_2).$$

This is a pure action to which player 1 would commit, if player 1 had the chance to do so (and hence the name "Stackelberg" action, see the discussion in Example 1), given that such a commitment induces a best response from player 2. If there is more than one such action for player 1, the action can be chosen arbitrarily.

However, player 1 would typically prefer to commit to a mixed action. In the product-choice game, for example, a commitment by player 1 to mixing between H and L, with slightly larger probability on H, still induces player 2 to choose h and gives player 1 a larger payoff than a commitment to H. Define the mixed-action Stackelberg payoff as

$$v_1^{**} \equiv \sup_{\alpha_1 \in \Delta(A_1)} \min_{\alpha_2 \in B(\alpha_1)} u_1(\alpha_1, \alpha_2), \quad (3)$$

where $B(\alpha_1) = \arg\max_{a_2} u_2(\alpha_1, a_2)$ is the set of player 2's best responses to α_1. In the product-choice game, $v_1^* = 2$, while $v_1^{**} = 5/2$. Typically, the supremum is not achieved by any mixed action, and so there is no mixed-action Stackelberg type. However, there are mixed action types that, if player 2 is convinced she is facing such a type, will yield payoffs arbitrarily close to the mixed-action Stackelberg payoff.

As with imperfect monitoring, simple mixed action types under perfect monitoring raise issues of monitoring, since a deviation by the normal type from the distribution α_1 of a mixed action type $\xi(\alpha_1)$, to some action in the support cannot be detected. However, when monitoring of the pure actions is perfect, it is possible to statistically detect deviations, and this will be enough to imply the appropriate reputation lower bound.

When monitoring is imperfect, the public signals are statistically informative about the actions of the long-lived player under the next assumption (Lemma 1).

Assumption 1 *For all $a_2 \in A_2$, the collection of probability distributions $\{\rho(y \mid (a_1, a_2)\colon a_1 \in A_1\}$ is linearly independent.*

This assumption is trivially satisfied in the perfect monitoring case. Reputation effects still exist when this assumption fails, but the bounds are more complicated to calculate (see [12] or Sect. 15.4.1 in [17]).

Fixing an action for player 2, a_2, the mixed action α_1 implies the signal distribution $\sum_{a_1} \rho(y \mid (a_1, a_2))\alpha_1(a_1)$.

Lemma 1 *Suppose ρ satisfies Assumption 1. Then, if for some a_2,*

$$\sum_{a_1} \rho(y \mid (a_1, a_2))\alpha_1(a_1) = \sum_{a_1} \rho(y \mid (a_1, a_2))\alpha_1'(a_1), \quad \forall y, \quad (4)$$

then $\alpha_1 = \alpha_1'$.

Proof Suppose (4) holds for some a_2. Let R denote the $|Y| \times |A_1|$ matrix whose y-a_1 element is given by $\rho(y \mid (a_1, a_2))$ (so that the a_1-column is the probability distribution on Y implied by the action profile $a_1 a_2$). Then, (4) can be written as $R\alpha_1 = R\alpha_1'$, or more simply as $R(\alpha_1 - \alpha_1') = 0$. By Assumption 1, R has full column rank, and so $x = 0$ is the only vector $x \in \mathbb{R}^{|A_1|}$ solving $Rx = 0$. □

Consequently, if player 2 believes that the long-lived player's behavior implies a distribution over the signals close to the distribution implied by some particular action α_1', then player 2 must believe that the long-lived player's action is also close to α_1'. Since A_2 is finite, this then implies that when player 2 is best responding to some belief about the long-lived player's behavior implying a distribution over signals sufficiently close to the distribution implied by α_1', then player 2 is in fact best responding to α_1'.

We are now in a position to state the main reputation bound result. Let $\underline{v}_1(\xi_0, \mu, \delta)$ be the infimum over the set of the normal player 1's payoffs in any (pure or mixed) Nash equilibrium in the incomplete information repeated game, given the distribution μ over types and the discount factor δ.

Proposition 1 (Fudenberg and Levine [11,12]) *Suppose ρ satisfies Assumption 1 and let $\hat{\xi}$ denote the simple action type that always plays $\hat{\alpha}_1 \in \Delta(A_1)$. Suppose $\mu(\xi_0)$, $\mu(\hat{\xi}) > 0$. For every $\eta > 0$, there is a value K such that for all δ,*

$$\underline{v}_1(\xi_0, \mu, \delta) \geq (1-\eta)\delta^K \min_{\alpha_2 \in B(\hat{\alpha}_1)} u_1(\hat{\alpha}_1, \alpha_2) + (1-(1-\eta)\delta^K) \min_{a \in A} u_1(a) . \quad (5)$$

This immediately yields the pure action Stackelberg reputation bound. Fix $\varepsilon > 0$. Taking $\hat{\alpha}_1$ in the proposition as the degenerate mixture that plays the Stackelberg action a_1^* with probability 1, Eq. (5) becomes

$$\begin{aligned}\underline{v}_1(\xi_0, \mu, \delta) &\geq (1-\eta)\delta^K v_1^* + (1-(1-\eta)\delta^K) \min_{a \in A} u_1(a) \\ &\geq v_1^* - (1-(1-\eta)\delta^K) 2M ,\end{aligned}$$

where $M \equiv \max_a |u_1(a)|$. This last expression is at least as large as $v_1^* - \varepsilon$ when $\eta < \varepsilon/(2M)$ and δ is sufficiently close to 1.

The mixed action Stackelberg reputation bound is also covered:

Corollary 1 *Suppose ρ satisfies Assumption 1 and μ assigns positive probability to some sequence of simple types $\{\xi(\alpha_1^k)\}_{k=1}^{\infty}$ with each α_1^k in $\Delta(A_1)$ satisfying*

$$v_1^{**} = \lim_{k \to \infty} \min_{\alpha_2 \in B(\alpha_1^k)} u_1(\alpha_1^k, \alpha_2) .$$

For all $\varepsilon' > 0$, there exists $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$,

$$\underline{v}_1(\xi_0, \mu, \delta) \geq v_1^{**} - \varepsilon' .$$

The remainder of this subsection outlines a proof of Proposition 1. Fix a strategy profile (σ_1, σ_2) (which may be Nash, but at this point of the discussion, need not be). The beliefs μ then induce a probability distribution $\mathbf{P}$ on the set of outcomes, which is the set of possible infinite histories (denoted by h^∞) and realized types, $(Y \times A)^\infty \times \Xi \equiv \Omega$. The probability measure $\mathbf{P}$ describes how the short-lived players believe the game will evolve, given their prior beliefs μ about the types of the long-lived player. Let $\hat{\mathbf{P}}$ denote the probability distribution on the set of outcomes induced by (σ_1, σ_2) and the action type $\hat{\xi}$. The probability measure $\hat{\mathbf{P}}$ describes how the short-lived players believe the game will evolve if the long-lived player's type is $\hat{\xi}$. Finally, let $\tilde{\mathbf{P}}$ denote the probability distribution on the set of outcomes induced by (σ_1, σ_2) conditioning on the long-lived player's type not being the action type $\hat{\xi}$. Then, $\mathbf{P} \equiv \hat{\mu}\hat{\mathbf{P}} + (1-\hat{\mu})\tilde{\mathbf{P}}$, where $\hat{\mu} \equiv \mu(\hat{\xi})$.

The discussion after Lemma 1 implies that the optimal behavior of the short-lived player in period t is determined by that player's beliefs over the signal realizations in that period. These beliefs can be viewed as a *one-step ahead prediction* of the signal y that will be realized conditional on the history h^t, $\mathbf{P}(y \mid h^t)$. Let $\hat{\mu}^t(h^t) = \mathbf{P}(\hat{\xi} \mid h^t)$ denote the posterior probability after observing h^t that the short-lived player assigns to the long-lived player having type $\hat{\xi}$. Note also that if the long-lived player is the action type $\hat{\xi}$, then the true probability of the signal y is $\hat{\mathbf{P}}(y \mid h^t) = \rho\big(y \mid (H, \sigma_2(h^t))\big)$. Then,

$$\mathbf{P}(y \mid h^t) = \hat{\mu}^t(h^t)\hat{\mathbf{P}}(y \mid h^t) + (1-\hat{\mu}^t(h^t))\tilde{\mathbf{P}}(y \mid h^t) .$$

The key step in the proof of Proposition 1 is a statistical result on merging. The following lemma essentially says that the short-lived players cannot be surprised too many times. Note first that an infinite public history h^∞ can be thought of as a sequence of ever longer finite public histories h^t. Consider the collection of infinite public histories with the property that player 2 often sees histories h^t that lead to very different one-step ahead predictions about the signals under $\tilde{\mathbf{P}}$ and under $\hat{\mathbf{P}}$ and have a "low" posterior that the long-lived player is $\hat{\xi}$. The lemma asserts that if the long-lived player is in fact the action type $\hat{\xi}$, this collection of infinite public histories has low probability. Seeing the signals more likely under $\hat{\xi}$ leads the short-lived players to increase the posterior probability on $\hat{\xi}$. The posterior probability fails to converge to 1 under $\hat{\mathbf{P}}$ only if the play of the types different from $\hat{\xi}$ leads, on average, to a signal distribution similar to that implied by $\hat{\xi}$. For the purely statistical statement and its proof, see Section 15.4.2 in [17].

Lemma 2 *For all $\eta, \psi > 0$ and $\mu^\dagger \in (0, 1]$, there exists a positive integer K such that for all $\mu(\hat{\xi}) \in [\mu^\dagger, 1)$, for every strategy $\sigma_1 : \mathcal{H}_1 \times \Xi \to \Delta(A_1)$ and $\sigma_2 : \mathcal{H} \to \Delta(A_2)$,*

$$\hat{\mathbf{P}}\Big(h^\infty : \big|\big\{t \geq 1 : (1-\hat{\mu}^t(h^t)) \max_y \big|\tilde{\mathbf{P}}(y \mid h^t) - \hat{\mathbf{P}}(y \mid h^t)\big| \geq \psi\big\}\big| \geq K\Big) \leq \eta . \quad (6)$$

Note that the bound K holds for all strategy profiles (σ_1, σ_2) and all prior probabilities $\mu(\hat{\xi}) \in [\mu^\dagger, 1)$. This allows us to bound equilibrium payoffs.

Proof of Proposition 1 Fix $\eta > 0$. From Lemma 1, by choosing ψ sufficiently small in Lemma 2, with $\hat{\mathbf{P}}$-probability at least $1 - \eta$, there are at most K periods in which the short-lived players are not best responding to $\hat{\alpha}_1$.

Since a deviation by the long-lived player to the simple strategy of always playing $\hat{\alpha}_1$ induces the same distribution on public histories as $\hat{\mathbf{P}}$, the long-lived player's expected payoff from such a deviation is bounded below by the right side of (5). □

Temporary Reputation Effects

Under perfect monitoring, there are often pooling equilibria in which the normal and some action type of player 1 behave identically on the equilibrium path (as in Example 1). Deviations on the part of the normal player 1 are deterred by the prospect of the resulting punishment. Under imperfect monitoring, such pooling equilibria do not exist. The normal and action types may play identically for a long period of time, but the normal type always eventually has an incentive to cheat at least a little on the commitment strategy, contradicting player 2's belief that player 1 will exhibit commitment behavior. Player 2 must then eventually learn player 1's type.

In addition to Assumption 1, disappearing reputation effects require full support monitoring.

Assumption 2 *For all $a \in A$, $y \in Y$, $\rho(y \mid a) > 0$.*

This assumption implies that Bayes' rule determines the beliefs of player 2 about the type of player 1 after *all* histories.

Suppose there are only two types of player 1, the normal type ξ_0 and a simple action type $\hat{\xi}$, where $\hat{\xi} = \xi(\hat{\alpha}_1)$ for some $\hat{\alpha}_1 \in \Delta(A_1)$. The analysis is extended to many commitment types in Section 6.1 in Cripps et al. [8]. It is convenient to denote a strategy for player 1 as a pair of functions $\tilde{\sigma}_1$ and $\hat{\sigma}_1$ (so $\hat{\sigma}_1(h_1^t) = \hat{\alpha}_1$ for all $h_1^t \in \mathcal{H}_1$), the former for the normal type and the latter for the action type.

Recall that $\mathbf{P} \in \Delta(\Omega)$ is the unconditional probability measure induced by the prior μ, and the strategy profile $(\hat{\sigma}_1, \tilde{\sigma}_1, \sigma_2)$, while $\hat{\mathbf{P}}$ is the measure induced by conditioning on $\hat{\xi}$. Since $\{\xi_0\} = \Xi \setminus \{\hat{\xi}\}$, $\tilde{\mathbf{P}}$ is the measure induced by conditioning on ξ_0. That is, $\hat{\mathbf{P}}$ is induced by the strategy profile $\hat{\sigma} = (\hat{\sigma}_1, \sigma_2)$ and $\tilde{\mathbf{P}}$ by $\tilde{\sigma} = (\tilde{\sigma}_1, \sigma_2)$, describing how play evolves when player 1 is the commitment and normal type, respectively.

The action of the commitment type satisfies the following assumption.

Assumption 3 *Player 2 has a unique stage-game best response to $\hat{\alpha}_1$ (denoted by $\hat{a}_2$) and $\hat{\alpha} \equiv (\hat{\alpha}_1, \hat{a}_2)$ is not a stage-game Nash equilibrium.*

Let $\hat{\sigma}_2$ denote the strategy of playing the unique best response $\hat{a}_2$ to $\hat{\alpha}_1$ in each period independently of history. Since $\hat{\alpha}$ is not a stage-game Nash equilibrium, $(\hat{\sigma}_1, \hat{\sigma}_2)$ is not a Nash equilibrium of the complete information infinite horizon game.

Proposition 2 (Cripps, Mailath and Samuelson [8]) *Suppose the monitoring distribution ρ satisfies Assumptions 1 and 2, and the commitment action $\hat{\alpha}_1$ satisfies Assumption 3. In any Nash equilibrium of the game with incomplete information, the posterior probability assigned by player 2 to the commitment type, $\hat{\mu}^t$, converges to zero under $\tilde{\mathbf{P}}$, i. e.,*

$$\hat{\mu}^t(h^t) \to 0, \qquad \tilde{\mathbf{P}}\text{-a.s.}$$

The intuition is straightforward: Suppose there is a Nash equilibrium of the incomplete information game in which both the normal and the action type receive positive probability in the limit (on a positive probability set of histories). On this set of histories, player 2 cannot distinguish between signals generated by the two types (otherwise player 2 could ascertain which type she is facing), and hence must believe that the normal and action types are playing the same strategies on average. But then player 2 must play a best response to this strategy, and hence to the action type. Since the action type's behavior is not a best response for the normal type (to this player 2 behavior), player 1 must eventually find it optimal to *not* play the action-type strategy, contradicting player 2's beliefs.

Assumption 3 requires a unique best response to $\hat{\alpha}_1$. For example, in the product-choice game, every action for player 2 is a best response to player 1's mixture α_1' that assigns equal probability to H and L. This indifference can be exploited to construct an equilibrium in which (the normal) player 1 plays α_1' after every history (Section 7.6.2 in [17]). This will still be an equilibrium in the game of incomplete information in which the commitment type plays α_1', with the identical play of the normal and commitment types ensuring that player 2 never learns player 1's type. In contrast, player 2 has a unique best response to any other mixture on the part of player 1. Therefore, if the commitment type is committed to any mixed action other than α_1', player 2 will eventually learn player 1's type.

As in Proposition 1, a key step in the proof of Proposition 2 is a purely statistical result on updating. Either player 2's expectation (given her history) of the strategy played by the normal type ($\tilde{E}[\tilde{\sigma}_1^t \mid h^t]$, where $\tilde{E}$ denotes expectation with respect to $\tilde{\mathbf{P}}$) is in the limit identical to the strategy played by the action type ($\hat{\alpha}_1$), or player 2's posterior probability that player 1 is the action type ($\hat{\mu}^t(h^t)$) converges to zero (given that player 1 is indeed

normal). This is a merging argument and closely related to Lemma 2. If the distributions generating player 2's signals are different for the normal and action type, then these signals provide information that player 2 will use in updating her posterior beliefs about the type she faces. This (converging, since beliefs are a martingale) belief can converge to an interior probability only if the distributions generating the signals are asymptotically uninformative, which requires that they be asymptotically identical.

Lemma 3 *Suppose the monitoring distribution ρ satisfies Assumptions 1 and 2. Then in any Nash equilibrium,*

$$\lim_{t\to\infty} \hat{\mu}^t \max_{a_1} \left|\hat{\alpha}_1(a_1) - \tilde{E}[\,\tilde{\sigma}_1^t(a_1) \mid h^t\,]\right| = 0, \quad \tilde{\mathbf{P}}\text{-}a.s. \tag{7}$$

Given Proposition 2, it should be expected that continuation play converges to an equilibrium of the complete information game, and this is indeed the case. See Theorem 2 [8] for the formal statement.

Proposition 2 leaves open the possibility that for any period T, there may be equilibria in which uncertainty about player 1's type survives beyond T, even though such uncertainty asymptotically disappears in any equilibrium. This possibility cannot arise. The existence of a sequence of Nash equilibria with uncertainty about player 1's type persisting beyond period $T \to \infty$ would imply the (contradictory) existence of a limiting Nash equilibrium in which uncertainty about player 1's type persists.

Proposition 3 (Cripps, Mailath and Samuelson [9]) *Suppose the monitoring distribution ρ satisfies Assumptions 1 and 2, and the commitment action $\hat{\alpha}_1$ satisfies Assumption 3. For all $\varepsilon > 0$, there exists T such that for any Nash equilibrium of the game with incomplete information,*

$$\tilde{\mathbf{P}}(\hat{\mu}^t < \varepsilon,\ \forall t > T) > 1 - \varepsilon\,.$$

Example 2 Recall that in the product-choice game, the unique player 2 best response to H is to play h, and Hh is not a stage-game Nash equilibrium. Proposition 1 ensures that the normal player 1's expected value in the repeated game of incomplete information with the H-action type is arbitrarily close to 2, when player 1 is very patient. In particular, if the normal player 1 plays H in every period, then player 2 will at least eventually play her best response of h. If the normal player 1 persisted in mimicking the action type by playing H in each period, this behavior would persist indefinitely. It is the feasibility of such a strategy that lies at the heart of the reputation bounds on expected payoffs. However, this strategy is not optimal. Instead, player 1 does even better by attaching some probability to L, occasionally reaping the rewards of his reputation by earning a stage-game payoff even larger than 2. The result of such equilibrium behavior, however, is that player 2 must eventually learn player 1's type. The *continuation* payoff is then bounded below 2 (recall (1)).

Reputation effects arise when player 2 is uncertain about player 1's type, and there may well be a long period of time during which player 2 is sufficiently uncertain of player 1's type (relative to the discount factor), and in which play does not resemble an equilibrium of the complete information game. Eventually, however, such behavior must give way to a regime in which player 2 is (correctly) convinced of player 1's type.

For any prior probability $\hat{\mu}$ that the long-lived player is the commitment type and for any $\varepsilon > 0$, there is a discount factor δ sufficiently large that player 1's expected payoff is close to the commitment-type payoff. This holds no matter how small $\hat{\mu}$. However, for any fixed δ and in any equilibrium, there is a time at which the posterior probability attached to the commitment type has dropped below the corresponding critical value of $\hat{\mu}$, becoming too small (relative to δ) for reputation effects to operate.

A reasonable response to the results on disappearing reputation effects is that a model of *long-run* reputations should incorporate some mechanism by which the uncertainty about types is continually replenished. For example, Holmström [13], Cole, Dow and English [6], Mailath and Samuelson [16], and Phelan [19] assume that the type of the long-lived player is governed by a stochastic process rather than being determined once and for all at the beginning of the game. In such a situation, reputation effects can indeed have long-run implications.

Reputation as a State

The posterior probability that short-lived players assign to player 1 being $\hat{\xi}$ is sometimes interpreted as player 1's reputation, particularly if $\hat{\xi}$ is the Stackelberg type. When Ξ contains only the normal type and $\hat{\xi}$, the posterior belief $\hat{\mu}^t$ is a state variable of the game, and attention is sometimes restricted to Markov strategies (i. e., strategies that only depend on histories through their impact on the posterior beliefs of the short-lived players). An informative example is Benabou and Laroque [2], who study the Markov perfect equilibria of a game in which the uninformed players respond continuously to their beliefs. They show that the informed player eventually reveals his type in any Markov perfect equilibrium. On the other hand, Markov equilibria

need not exist in finitely repeated reputation games (Section 17.3 in [17]).

The literature on reputation effects has typically not restricted attention to Markov strategies, since the results do not require the restriction.

Two Long-Lived Players

The introduction of nontrivial intertemporal incentives for the uninformed player significantly reduces reputation effects. For example, when only simple Stackelberg types are considered, the Stackelberg payoff may not bound equilibrium payoffs. The situation is further complicated by the possibility of non-simple commitment types (i. e., types that follow nonstationary strategies).

Consider applying the logic from Sect. "The Reputation Bound" to obtain the Stackelberg reputation bound when both players are long-lived and player 1's characteristics are unknown, under perfect monitoring. The first step is to demonstrate that, if the normal player 1 persistently plays the Stackelberg action and there exists a type committed to that action, then player 2 must eventually attach high probability to the event that the Stackelberg action is played in the future. This argument, a simple version of Lemma 2, depends only upon the properties of Bayesian belief revision, independently of whether the person holding the beliefs is a long-lived or short-lived player.

When player 2 is short-lived, the next step is to note that if she expects the Stackelberg action, then she will play a best response to this action. If player 2 is instead a long-lived player, she may have an incentive to play something other than a best response to the Stackelberg type.

The key step when working with two long-lived players is thus to establish conditions under which, as player 2 becomes increasingly convinced that the Stackelberg action will appear, player 2 must eventually play a best response to that action. One might begin such an argument by observing that, as long as player 2 discounts, any losses from not playing a current best response must be recouped within a finite length of time. But if player 2 is "very" convinced that the Stackelberg action will be played not only now but for sufficiently many periods to come, there will be no opportunity to accumulate subsequent gains, and hence player 2 might just as well play a stage-game best response.

Once it is shown that player 2 is best responding to the Stackelberg action, the remainder of the argument proceeds as in the case of a short-lived player 2. The normal player 1 must eventually receive very nearly the Stackelberg payoff in each period of the repeated game. By making player 1 sufficiently patient (*relative to player 2*, so that discount factors differ), this consideration dominates player 1's payoffs, putting a lower bound on the latter. Hence, the obvious handling of discount factors is to fix player 2's discount factor δ_2, and to consider the limit as player 1 becomes patient, i. e., δ_1 approaching one.

This intuition misses the following possibility. Player 2 may be choosing something other than a best response to the Stackelberg action out of fear that a current best response may trigger a disastrous future punishment. This punishment would not appear if player 2 faced the Stackelberg type, but player 2 can be made confident only that she faces the Stackelberg *action*, not the Stackelberg type. The fact that the punishment lies off the equilibrium path makes it difficult to assuage player 2's fear of such punishments. Short-lived players in the same situation are similarly uncertain about the future ramifications of best responding, but being short-lived, this uncertainty does not affect their behavior.

Consequently, reputation effects are typically weak with two long-lived players under perfect monitoring: Celentani, Fudenberg, Levine and Pesendorfer [3] and Cripps and Thomas [7], describe examples with only the normal and the Stackelberg types of player 1, in which the future play of the normal player 1 is used to punish player 2 for choosing a best response to the Stackelberg action when she is not supposed to, and player 1's payoff is significantly below the Stackelberg payoff. Moreover, the robustness of reputation effects to additional types beyond the Stackelberg type, a crucial feature of settings with one long-lived player, does not hold with two long-lived players. Schmidt [21] showed that the possibility of a "punishment" type can prevent player 2 best responding to the Stackelberg action, while Evans and Thomas [10] showed that the Stackelberg bound is valid if in addition to the Stackelberg type, there is an action type who punishes player 2 for *not* behaving appropriately (see Sections 16.1 and 16.5 in [17]).

Imperfect monitoring (of *both* players' actions), on the other hand, rescues reputation effects. With a sufficiently rich set of commitment types, player 1 can be assured of at least his Stackelberg payoff. Indeed, player 1 can often be assured of an even higher payoff, in the presence of commitment types who play nonstationary strategies [3]. At the same time, these reputation effects are temporary (Theorem 2 in [9]).

Finally, there is a literature on reputation effects in bargaining games (see [1,4,5,20]), where the issues described above are further complicated by the need to deal with the bargaining model itself.

Future Directions

The detailed structure of equilibria of the incomplete information game is not well understood, even for the canonical game of Sect. "A Canonical Model". A more complete description of the structure of equilibria is needed.

While much of the discussion was phrased in terms of the Stackelberg type, Proposition 1 provides a reputation bound for *any* action type. While in some settings, it is natural that the uninformed players assign strictly positive probability to the Stackelberg type, it is not natural in other settings. A model endogenizing the nature of action types would be an important addition to the reputation literature.

Finally, while the results on reputation effects with two long-lived players are discouraging, there is still the possibility that some modification of the model will rescue reputation effects in this important setting.

Acknowledgments

I thank Eduardo Faingold, KyungMin Kim, Antonio Penta, and Larry Samuelson for helpful comments.

Bibliography

1. Abreu D, Gul F (2000) Bargaining and Reputation. Econometrica 68(1):85–117
2. Benabou R, Laroque G (1992) Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility. Q J Econ 107(3):921–958
3. Celentani M, Fudenberg D, Levine DK, Pesendorfer W (1996) Maintaining a Reputation Against a Long-Lived Opponent. Econometrica 64(3):691–704
4. Chatterjee K, Samuelson L (1987) Bargaining with Two-Sided Incomplete Information: An Infinite Horizon Model with Alternating Offers. Rev Econ Stud 54(2):175–192
5. Chatterjee K, Samuelson L (1988) Bargaining with Two-Sided Incomplete Information: The Unrestricted Offers Case. Oper Res 36(4):605–638
6. Cole HL, Dow J, English WB (1995) Default, Settlement, and Signalling: Lending Resumption in a Reputational Model of Sovereign Debt. Int Econ Rev 36(2):365–385
7. Cripps MW, Thomas JP (1997) Reputation and Perfection in Repeated Common Interest Games. Games Econ Behav 18(2):141–158
8. Cripps MW, Mailath GJ, Samuelson L (2004) Imperfect Monitoring and Impermanent Reputations. Econometrica 72(2):407–432
9. Cripps MW, Mailath GJ, Samuelson L (2007) Disappearing Private Reputations in Long-Run Relationships. J Econ Theory 134(1):287–316
10. Evans R, Thomas JP (1997) Reputation and Experimentation in Repeated Games with Two Long-Run Players. Econometrica 65(5):1153–1173
11. Fudenberg D, Levine DK (1989) Reputation and Equilibrium Selection in Games with a Patient Player. Econometrica 57(4):759–778
12. Fudenberg D, Levine DK (1992) Maintaining a Reputation When Strategies Are Imperfectly Observed. Rev Econ Stud 59(3):561–579
13. Holmström B (1982) Managerial Incentive Problems: A Dynamic Perspective. In: Essays in Economics and Management in Honour of Lars Wahlbeck, pp 209–230. Swedish School of Economics and Business Administration, Helsinki. Published in: Rev Econ Stud 66(1):169–182
14. Kreps D, Wilson R (1982) Reputation and Imperfect Information. J Econ Theory 27:253–279
15. Kreps D, Milgrom PR, Roberts DJ, Wilson R (1982) Rational Cooperation in the Finitely Repeated Prisoner's Dilemma. J Econ Theory 27:245–252
16. Mailath GJ, Samuelson L (2001) Who Wants a Good Reputation? Rev Econ Stud 68(2):415–441
17. Mailath GJ, Samuelson L (2006) Repeated Games and Reputations: Long-Run Relationships. Oxford University Press, New York
18. Milgrom PR, Roberts DJ (1982) Limit pricing and entry under incomplete information: An equilibrium analysis. Econometrica 50:443–459
19. Phelan C (2006) Public Trust and Government Betrayal. J Econ Theory 130(1):27–43
20. Schmidt KM (1993) Commitment Through Incomplete Information in a Simple Repeated Bargaining Game. J Econ Theory 60(1):114–139
21. Schmidt KM (1993) Reputation and Equilibrium Characterization in Repeated Games of Conflicting Interests. Econometrica 61(2):325–351
22. Selten R (1978) Chain-store paradox. Theory Decis 9:127–59
23. Stackelberg HV (1934) Marktform und Gleichgewicht. Springer, Vienna

Resonances in Electronic Transport Through Quantum Wires and Rings

VASSILIOS VARGIAMIDIS
Department of Physics, Aristotle University, Thessaloniki, Greece

Article Outline

Glossary

Scattering Scattering is a general physical process whereby some forms of radiation, such as light, sound or moving particles, for example, are forced to deviate from a straight trajectory by one or more localized non-uniformities in the medium through which it passes. These non-uniformities are, sometimes, known as scatterers or scattering centers. In quantum transport, a scattering center may be provided by an impurity potential.

Transmission probability The transmission probability, denoted by T, for a quantum mechanical particle to pass through a scattering potential is the ratio between the flux of particles that emerges from the potential and the flux that arrives at it. Equivalently, T is the fraction of incident particles that succeed in passing through the scattering potential.

Transmission resonance Transmission resonance in electronic transport of quantum wires is an abrupt variation of the transmission probability that occurs over a very small interval the incident electron's energy.

Mode In a low-dimensional system where the electrons are confined in one or more directions, the energy eigenstates in the confinement direction(s) represent the modes of the system.

Evanescent mode When the electrons are free to propagate in one direction but are confined in the other directions, a mode whose energy is greater than the Fermi energy cannot propagate. This is called an evanescent mode.

Conductance quantization To measure the conductance of a sample (such as, a constriction) we divide the current through the sample by the electric potential difference between the reservoirs which are connected to the sample. If N_c is the number of occupied subbands, the expression for the conductivity pertaining to the two-terminal measurement is $G = 2e^2 N_c/h$. According to this description, the conductance is quantized such that it increases in steps by an amount equal to the quantum of conductance $2e^2/h$ whenever a new subband opens. This can be achieved either by widening of the constriction (and thus by lowering the subband energies) or by increasing the density of electrons (and thus by raising the Fermi energy).

Definition of the Subject

Since the 1980s advances in the growth techniques and new electronic materials developed therefrom have provided almost defect-free electronic devices, which have dimensions in one or more directions on the quantum scale. New quantum regimes governing such systems of lower dimensionality have led to novel electronic properties with potential applications. Quantum wells, wires, and dots, which have been implemented in the terminology of condensed-matter physics, indicate not only different dimensionality but also exhibit dramatically different electronic properties. Particularly, the electronic transport properties in lower dimensionality have several important features, which have stimulated a great deal of theoretical and experimental research. In electronic transport studies, if the size of the sample (or device) is smaller than the phase breaking length, the transport is phase coherent. In this case, electrons have a well-defined phase throughout the device, even though they may experience elastic scattering. Numerous publications on this type of transport have appeared, thus contributing to a field called mesoscopic physics, a term that indicates a new length scale for physical events between macroscopic and microscopic.

Phase coherent electronic transport through mesoscopic systems has been shown to exhibit resonant behavior as a function of the incident electron energy. This type of resonant behavior is of great interest both as a basis for the creation of new resonant nanoelectronic devices and for revealing a fundamental aspect of quantum mechanics. In various model calculations on electronic transport of microstructures, transmission resonances have been found to exhibit two types of resonance behavior: i) symmetric resonances (Breit–Wigner type), and ii) asymmetric resonances (Fano type). In particular, one often encounters points of vanishing transmission or reflection as a resonance is crossed. The interference effects that give rise to these resonances in the above-mentioned nanostructures are reviewed here, with emphasis given to the more recent interesting effects. This article is intended for those who are familiar with the field but the somewhat detailed derivations and explanations make it also accessible to beginners or researchers from similar disciplines.

Introduction

Since the early 1990s, resonances of the Fano type have been treated theoretically in various condensed matter systems including transport through quantum wires with attractive impurities or quantum dots [4,6,22,23,30,31,35,36,37,38,39,40]. In these systems, the coupling between a bound state of the impurity and the continuum leads to a quasibound (resonant) state. The interference between the direct (nonresonant) transmission pathway and the transmission via the quasibound state gives rise to an asymmetric Fano resonance. The Fano effect is a ubiqui-

tous phenomenon and has been observed in a large variety of experiments including atomic photoionization [13], neutron scattering [1], Raman scattering [8], optical absorption [11], and transport through mesoscopic systems with embedded quantum dots [5,16,19,24,26].

The purpose of this article is to review the resonance phenomena in the electronic transport of non-interacting electrons through: i) infinite rectilinear quantum wires with impurities, and ii) one-dimensional rings with impurities connected to current leads. In Sect. "Electronic Transport Through Quantum Wires", I will review the coupled-channel Feshbach theory, which is particularly suitable for describing the resonant behavior in ballistic electronic transport of quantum wires, and apply it to the case of an impurity with finite-range and smooth profile along the propagation direction. In Sect. "Resonances in the Transmission Probability of a Quantum Wire", it will be shown that varying the strength and size of the impurity causes the transformation of an asymmetric Fano resonance into a symmetric antiresonance (Breit–Wigner dip) and subsequently to an "inverted" Fano resonance. Thermal effects will also be considered. In Sect. "Electronic Transport Through Mesoscopic Open Rings", after reviewing the scattering matrix approach, I will discuss the Fano resonances in electronic transport of one-dimensional (abbreviated, 1D) open rings with a short-range (abbreviated, SR) impurity in one arm and the effects of the ring-lead coupling on the resonance structure. In this Section, an interesting feature of the Fano line shape will also be discussed; namely, the systematic collapse of certain Fano resonances for special impurity positions. The effect of an Aharonov–Bohm (abbreviated, AB) magnetic flux will also briefly be considered as well as the temperature dependence of the Fano effect. Finally, in Sect. "Future Directions", some thoughts on future directions will be given.

Electronic Transport Through Quantum Wires

A relatively new view was established for the quantum theory of scattering since it became clear that the electrical linear response of open conductors could be related to its transmission and reflection properties [2,28]. In addition, experiments performed at that time independently by van Wees et al. [49] and Wharam et al. [44] have been a breakthrough in the field of quantum ballistic transport in a quantum point contact (abbreviated, QPC) in a two-dimensional electron gas (abbreviated, 2D EG). Using high-mobility $GaAs-Al_{1-x}Ga_xAs$ heterojunctions and the split-gate technique, they imposed a small constriction on the sample. A channel was obtained from this constriction by applying a negative bias to the split gate, and thus by causing the depletion of electrons beneath the gate. Thus, the portion of the 2D EG lying below the gap of the split-gate electrode remains conducting. In their experiments the width of the constriction is in the range of the Fermi wavelength λ_F, whereby quantum size effects become relevant.

In the theoretical models [44,49] initially used to explain the electron conduction and the quantization of conductance, the QPC was perceived as a uniform waveguide, and only the events in this waveguide were taken into account. Here, the word uniform refers to the confining potential, which is the same throughout the constriction. The current-transporting states are laterally confined in the waveguide, the width of which is in the range of the Fermi wavelength. Then, the transverse momenta of these states are quantized, resulting in a subband structure. In the following we confine ourselves to such uniform, infinite long waveguide (or quantum wire) at zero temperature.

Feshbach Coupled-Channel Theory

We consider a uniform 2D quantum wire as shown in Fig. 1, in which electrons are confined along the y direction (transverse direction) but are free to move along the x direction (propagation direction). In the presence of a scattering potential the Schrödinger equation describing the electron motion in the wire is

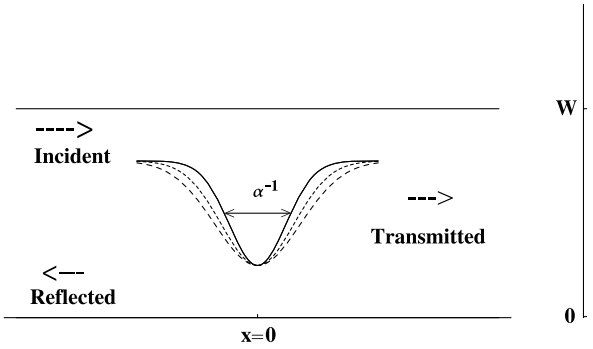
$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_c(y) + V(x,y)\right]\Psi(x,y) = E\Psi(x,y), \quad (1)$$

where $V_c(y)$ is the confining potential, $V(x,y)$ is a scattering potential, and m is the effective mass of the electron. The transverse potential $V_c(y)$, providing confinement of the electron motion along the y direction, gives rise to channel modes $\phi_n(y)$ which satisfy

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial y^2} + V_c(y)\right]\phi_n(y) = E_n\phi_n(y)\,, \quad (2)$$

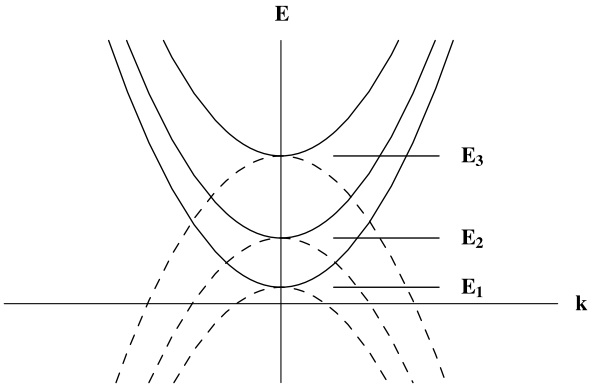
where E_n is the threshold energy for the nth mode. The wave propagation (or transport) in subband n takes place, as long as the threshold E_n (i. e., the bottom of the nth subband) is smaller than the Fermi level E (see Fig. 2). We expand the wave function $\Psi(x,y)$ of Eq. (1) in terms of the channel modes as

$$\Psi(x,y) = \sum_{n=1}^{\infty}\psi_n(x)\phi_n(y)\,. \quad (3)$$



Resonances in Electronic Transport Through Quantum Wires and Rings, Figure 1

A uniform quantum wire of width *W*. An incident electron wave is partially reflected and transmitted through a Pöschl–Teller attractive impurity $V(x,y) = (\hbar^2\gamma/2m)\,\mathrm{sech}^2(\alpha x)\upsilon(y)$, where α^{-1} is the decay length and $\upsilon(y)$ is arbitrary function. Larger values of α^{-1} represent impurities of larger size, as shown by the *dashed lines*



Resonances in Electronic Transport Through Quantum Wires and Rings, Figure 2

Dispersion relation of propagating (*solid*) and evanescent (*dashed*) states in a 2D quantum wire $E = E_n + (\hbar^2 k_n^2/2m)$, where E_n define the threshold energies (bottom of subbands). Modes with energy E_n such that $E_n < E$, where $n = 1, 2, \ldots$, are propagating along the wire while modes with $E_n > E$ (that is, $k_n = i\kappa_n$) are evanescent

Substituting Eq. (3) into Eq. (1) we obtain the following coupled-channel equations for $\psi_n(x)$:

$$(E - E_n - \hat{K})\psi_n(x) = \sum_{l=1}^{\infty} V_{nl}(x)\psi_l(x)\,, \tag{4}$$

where $\hat{K} = -(\hbar^2/2m)d^2/dx^2$ and the coupling matrix elements $V_{nl}(x)$ are given as

$$V_{nl}(x) = \int \mathrm{d}y\phi_n(y)V(x,y)\phi_l(y)\,. \tag{5}$$

We confine ourselves now to a single-subband regime, i. e., $E_1 \le E \le E_2$ where only the first mode $n = 1$ can propagate along the wire while the higher (evanescent) modes can only contribute to electron propagation due to their coupling with the first by the scattering potential. Thus all coupling matrix elements $V_{nl}(x)$ for the higher modes vanish away from the scattering region ($x \to \pm\infty$) and therefore only the first mode $n = 1$ can actually be scattered and, therefore, be found in some scattering state. It can be seen from Eq. (4) that the scattering states for the first mode are given as solutions of the equation

$$\left[\hat{K} + V_{11}(x)\right]\chi_k^{\pm}(x) = (E - E_1)\chi_k^{\pm}(x)\,, \tag{6}$$

where we denoted by $\chi_k^{+}(x)$ and $\chi_k^{-}(x)$ the wave functions that correspond to scattering states for which the incident wave comes from $-\infty$ and $+\infty$, respectively. These wave functions describe the background (nonresonant) scattering, which is the scattering in a hypothetical system in which there is no coupling to a bound state [14,30,37] of the impurity. In Eq. (6) the wave vector of the propagating mode is given as $k = [2m(E - E_1)]^{1/2}/\hbar$ while $V_{11}(x)$ is an effective 1D potential for scattering of the first mode. The scattering wave functions in the asymptotic region can be expressed as

$$\chi_k^{\pm}(x) = \begin{cases} t^{bg}\mathrm{e}^{\pm ikx} & (x \to \pm\infty) \\ \mathrm{e}^{\pm ikx} + r_{\pm}^{bg}\mathrm{e}^{\mp ikx} & (x \to \mp\infty) \end{cases}. \tag{7}$$

In Eq. (7) the upper signs correspond to incident wave from $-\infty$, while t^{bg} and $r_{\pm}^{bg}$ correspond to the background transmission and reflection amplitudes in the quantum wire. Let E be close to the energy $\tilde{E}_0$ of the bound state $\Phi_0(x)$ in the potential $V_{22}(x)$ of the uncoupled channel $n = 2$. This bound state is found as the solution of

$$\left[\hat{K} + V_{22}(x)\right]\Phi_0(x) = (\tilde{E}_0 - E_2)\Phi_0(x)\,. \tag{8}$$

Assuming no other channels exhibit bound states close to $\tilde{E}_0$, we can make the approximation of truncating the sum of Eq. (4) at $n = 2$. We then obtain the following system of equations

$$\left[E - E_1 - \hat{K} - V_{11}(x)\right]\psi_1(x) = V_{12}(x)\psi_2(x)\,, \tag{9}$$

$$\left[E - E_2 - \hat{K} - V_{22}(x)\right]\psi_2(x) = V_{21}(x)\psi_1(x)\,. \tag{10}$$

The system of coupled-channel equations [i. e., Eqs. (9) and (10)] is solved using the ansatz [15]

$$\psi_2(x) = A\Phi_0(x)\,. \tag{11}$$

Inserting this ansatz into Eq. (9) we obtain an inhomogeneous equation for $\psi_1(x)$, which can then be solved by employing the retarded Green's operator

$$\hat{G}_1 = (E - E_1 - \hat{K} - V_{11} + i0^+)^{-1} . \tag{12}$$

When $\hat{G}_1$ acts on Eq. (9), the general solution to the inhomogeneous equation can be found as

$$|\psi_1\rangle = |\chi_k^+\rangle + A\hat{G}_1 V_{12}|\Phi_0\rangle . \tag{13}$$

Then, Eq. (10) can be written as

$$A(E - E_2 - \hat{K} - V_{22})|\Phi_0\rangle = V_{21}|\chi_k^+\rangle + AV_{21}\hat{G}_1 V_{12}|\Phi_0\rangle . \tag{14}$$

We can rewrite Eq. (14) with the help of Eq. (8) as

$$A(E - \tilde{E}_0)|\Phi_0\rangle = V_{21}|\chi_k^+\rangle + AV_{21}\hat{G}_1 V_{12}|\Phi_0\rangle . \tag{15}$$

Multiplying Eq. (15) by $\langle\Phi_0|$ allows us to determine A in Eq. (11) as

$$A = \frac{\langle\Phi_0|V_{21}|\chi_k^+\rangle}{E - \tilde{E}_0 - \langle\Phi_0|V_{21}G_1V_{12}|\Phi_0\rangle} . \tag{16}$$

Using the scattering states $\chi_k^{\pm}(x)$, we can write the explicit form of the retarded Green's function in 1D as

$$G_1(x,x') = \frac{m}{i\hbar^2 k t^{bg}} \times \begin{cases} \chi_k^+(x)\chi_k^-(x') & (x > x') \\ \chi_k^+(x')\chi_k^-(x) & (x < x') \end{cases} \tag{17}$$

We then finally obtain from Eq. (13) the solution for $x \to \infty$ as

$$\psi_1(x) = \chi_k^+(x) + \frac{m}{i\hbar^2 k t^{bg}}\chi_k^+(x) \cdot \frac{\langle(\chi_k^-)^*|V_{12}|\Phi_0\rangle\langle\Phi_0|V_{21}|\chi_k^+\rangle}{E - \tilde{E}_0 - \langle\Phi_0|V_{21}G_1V_{12}|\Phi_0\rangle} . \tag{18}$$

In Eq. (18) the matrix element $\langle\Phi_0|V_{21}G_1V_{12}|\Phi_0\rangle$ is the self-energy due to the coupling of the bound state with the continuum and has in general both a real and an imaginary part.

The above formalism applies to a general scattering potential. In the following subsection we will examine a particular potential with finite range and smooth profile; namely, the Pöschl–Teller potential.

Pöschl–Teller Scattering Potential

We consider now a scattering potential of the Pöschl–Teller (abbreviated, PT) type

$$V(x,y) = \frac{\hbar^2\gamma}{2m}\,\mathrm{sech}^2(\alpha x)\upsilon(y) , \tag{19}$$

where $\upsilon(y)$ is an arbitrary function of the coordinate y, α^{-1} is the decay length, and γ sets the magnitude of the potential which is taken to be attractive ($\gamma < 0$). The longitudinal part of this potential has smooth profile as shown in Fig. 1. The matrix elements of Eq. (5) take the form

$$V_{nl}(x) = \frac{\hbar^2\gamma}{2m}\,\mathrm{sech}^2(\alpha x)\upsilon_{nl} , \tag{20}$$

where $\upsilon_{nl} = \langle\phi_n|\upsilon(y)|\phi_l\rangle$. In order to find the scattering states $\chi_k^{\pm}(x)$ we must solve Eq. (6) in the presence of an attractive potential

$$V_{11}(x) = -U_{11}\,\mathrm{sech}^2(\alpha x) , \tag{21}$$

where $U_{11} = -(\hbar^2\gamma/2m)\upsilon_{11}$. The solution proceeds in the same way as in a 1D scattering problem [27] and the asymptotic form of the wave function, as $x \to -\infty$, is written in terms of Gamma functions as

$$\chi_k^+(x) \sim \mathrm{e}^{-ikx}\frac{\Gamma(ik/\alpha)\Gamma(1 - ik/\alpha)}{\Gamma(-\eta)\Gamma(1+\eta)} + \mathrm{e}^{ikx}\frac{\Gamma(-ik/\alpha)\Gamma(1 - ik/\alpha)}{\Gamma(-ik/\alpha - \eta)\Gamma(-ik/\alpha + \eta + 1)} , \tag{22}$$

where $\eta = (1/2)\left[-1 + \sqrt{1 + (8m^*U_{11}/\alpha^2\hbar^2)}\right]$. In Eq. (22) Γ is the Gamma function. The reflection amplitude r_+^{bg} is the ratio of coefficients of the function $\chi_k^+(x)$ in Eq. (22) and due to symmetry $r_-^{bg} = r_+^{bg}$ holds. Having found the scattering states we now proceed to find the bound state of the 1D potential $V_{22}(x) = -U_{22}\,\mathrm{sech}^2(\alpha x)$ by solving Eq. (8), where $U_{22} = -(\hbar^2\gamma/2m)\upsilon_{22}$. Employing the notation $\epsilon = [-2m(\tilde{E}_0 - E_2)]^{1/2}/\hbar\alpha$ and $s = (1/2)\left[-1 + \sqrt{1 + (8mU_{22}/\alpha^2\hbar^2)}\right]$, we can bring Eq. (8) into a form that has solutions the associated Legendre polynomials $P_s^{\epsilon}(\xi)$, where $\xi = \tanh(\alpha x)$. The energy levels are then obtained from the condition $\epsilon = s - p$ as

$$\tilde{E}_p = E_2 - \frac{\hbar^2\alpha^2}{8m}\left[-(1+2p) + \sqrt{1 + \frac{8mU_{22}}{\alpha^2\hbar^2}}\right]^2 , \tag{23}$$

where $p = 0, 1, 2, \ldots$ The number of levels is determined by the condition $\epsilon > 0$ (or $p < s$) and generally depends on the depth U_{22} of the potential as well as on the inverse decay length α. In the following we will assume that U_{22} and α are such that $s \leq 1$, which implies that there is only one bound state with energy $\tilde{E}_0$. When this bound state lies in the continuum of the first subband it is transformed to a quasibound state and acquires a finite width. The normalized bound state wave function is given as $\Phi_0(x) = (\alpha/2)^{1/2}\,\mathrm{sech}(\alpha x)$.

Having found the scattering states $\chi_k^{\pm}(x)$ as well as the bound state, $\tilde{E}_0$ and Φ_0, we now proceed to calculate the matrix elements that occur in Eq. (18) and thus find the transmission amplitude. We then have

$$\begin{aligned}\langle(\chi_k^-)^*|V_{12}|\Phi_0\rangle &= \frac{\hbar^2}{2m}\gamma v_{12}\sqrt{\frac{\alpha}{2}} \\ &\quad\times\int \mathrm{d}x\left(\mathrm{e}^{-ikx}+r_-^{bg}\mathrm{e}^{ikx}\right)\mathrm{sech}^3(\alpha x) \\ &= \frac{\hbar^2}{2m}\gamma v_{12}\sqrt{\frac{\alpha}{2}}\left(I_1^*+r_-^{bg}I_1\right),\end{aligned} \tag{24}$$

where

$$\begin{aligned}I_1 &= \int_{-\infty}^{\infty}\mathrm{d}x\mathrm{e}^{ikx}\,\mathrm{sech}^3(\alpha x)\,, \\ I_1^* &= \int_{-\infty}^{\infty}\mathrm{d}x\mathrm{e}^{-ikx}\,\mathrm{sech}^3(\alpha x)\,.\end{aligned} \tag{25}$$

It can be verified that $I_1 = I_1^* = [(\alpha^2+k^2)\pi\cdot\mathrm{sech}(k\pi/2\alpha)]/2\alpha^3$. Also,

$$\langle\Phi_0|V_{21}|\chi_k^+\rangle = \frac{\hbar^2}{2m}\gamma v_{21}\sqrt{\frac{\alpha}{2}}t^{bg}I_1\,. \tag{26}$$

We calculate now the matrix element that occurs in the denominator of Eq. (18). We have

$$\begin{aligned}&\langle\Phi_0|V_{21}G_1V_{12}|\Phi_0\rangle \\ &= \frac{m}{i\hbar^2kt^{bg}}\int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{\infty}\mathrm{d}x'\Phi_0(x)V_{12}(x) \\ &\quad\cdot\Phi_0(x')V_{12}(x')\chi_k^-(x)\chi_k^+(x') \\ &\quad+\frac{m}{i\hbar^2kt^{bg}}\int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{x}\mathrm{d}x'\Phi_0(x)V_{12}(x)\Phi_0(x') \\ &\quad\cdot V_{12}(x')[\chi_k^+(x)\chi_k^-(x')-\chi_k^-(x)\chi_k^+(x')] \\ &= Q_1+Q_2-Q_3\,.\end{aligned} \tag{27}$$

Inserting the explicit expressions for the bound state Φ_0, the coupling V_{12}, and the scattering states $\chi_k^{\pm}$ into Eq. (27) we obtain

$$Q_1 = \frac{m}{i\hbar^2k}\left(\frac{\hbar^2}{2m}\right)^2\gamma^2v_{12}^2\left(\frac{\alpha}{2}\right)\left[I_1^*I_1+r_-^{bg}(I_1)^2\right], \tag{28}$$

and

$$Q_2 = \frac{m}{i\hbar^2k}\left(\frac{\hbar^2}{2m}\right)^2\gamma^2v_{12}^2\left(\frac{\alpha}{2}\right)\left[I_2+r_-^{bg}I_3\right], \tag{29}$$

where

$$I_2 = \int_{-\infty}^{\infty}\mathrm{d}x\mathrm{e}^{ikx}\,\mathrm{sech}^3(\alpha x)\int_{-\infty}^{x}\mathrm{d}x'\mathrm{e}^{-ikx'}\,\mathrm{sech}^3(\alpha x')\,, \tag{30}$$

and

$$I_3 = \int_{-\infty}^{\infty}\mathrm{d}x\mathrm{e}^{ikx}\,\mathrm{sech}^3(\alpha x)\int_{-\infty}^{x}\mathrm{d}x'\mathrm{e}^{ikx'}\,\mathrm{sech}^3(\alpha x')\,. \tag{31}$$

For Q_3 we get the same expression as for Q_2 and therefore $Q_2-Q_3=0$. Thus $\langle\Phi_0|V_{21}G_1V_{12}|\Phi_0\rangle = Q_1$. Using Eq. (28) and the results of the integrals I_1 and I_1^* we can finally write the matrix element as

$$\langle\Phi_0|V_{21}G_1V_{12}|\Phi_0\rangle = \delta - i\Gamma\,, \tag{32}$$

with

$$\begin{aligned}\delta = &-\frac{\hbar^2\gamma^2v_{12}^2(\alpha^2+k^2)^2\pi^2\,\mathrm{sech}^2(k\pi/2\alpha)}{32mk\alpha^5} \\ &\cdot\frac{cos\left[(\pi/2)\sqrt{1+(8mU_{11}/\hbar^2\alpha^2)}\right]\sinh(\pi k/\alpha)}{\sinh^2(\pi k/\alpha)+\cos^2\left[(\pi/2)\sqrt{1+(8mU_{11}/\hbar^2\alpha^2)}\right]},\end{aligned} \tag{33}$$

and

$$\begin{aligned}\Gamma = &\frac{\hbar^2\gamma^2v_{12}^2(\alpha^2+k^2)^2\pi^2\,\mathrm{sech}^2(k\pi/2\alpha)}{32mk\alpha^5} \\ &\cdot\frac{\sinh^2(\pi k/\alpha)}{\sinh^2(\pi k/\alpha)+\cos^2\left[(\pi/2)\sqrt{1+(8mU_{11}/\hbar^2\alpha^2)}\right]}.\end{aligned} \tag{34}$$

Equations (33) and (34) for δ and Γ give the shift and width respectively, that the bound state acquires. Furthermore, using Eqs. (24) and (26) we can write the numerator of Eq. (18) as

$$\frac{m}{i\hbar^2kt^{bg}}\langle(\chi_k^-)^*|V_{12}|\Phi_0\rangle\langle\Phi_0|V_{21}|\chi_k^+\rangle = \delta - i\Gamma\,. \tag{35}$$

Equation (18) can then finally be written for $x\to\infty$ as

$$\psi_1(x) = t^{bg}\mathrm{e}^{ikx}\frac{E-\tilde{E}_0}{E-\tilde{E}_0-\delta+i\Gamma}\,. \tag{36}$$

From Eq. (36) we can extract the transmission probability of the quantum wire as

$$T = |t^{bg}|^2\frac{\left(E-\tilde{E}_0\right)^2}{\left(E-E_R\right)^2+\Gamma^2}\,. \tag{37}$$

Due to the real part δ of the self-energy acquired by the bound state, a shifted quasibound state energy $E_R = \tilde{E}_0+\delta$

appears in Eq. (37). The background transmission coefficient, which enters Eq. (37), is found as

$$|t^{bg}|^2 = \frac{\sinh^2(\pi k/\alpha)}{\sinh^2(\pi k/\alpha)+\cos^2\left[(\pi/2)\sqrt{1+(8mU_{11}/\hbar^2\alpha^2)}\right]}. \tag{38}$$

Equation (37) can be transformed into the asymmetric Fano function [12] if we define reduced variables $\epsilon = (E-E_R)/\Gamma$ and $q = (E_R-\tilde{E}_0)/\Gamma = \delta/\Gamma$, where $\tilde{E}_0$ is the energy of the transmission zero and q is the asymmetry parameter of the resonance line shape. An interesting feature of Fano resonances in scattering from impurities in quantum wires is that the asymmetry parameter depends only on the characteristics of the impurity potential and not on the coupling of the quasibound state. In addition, the limit $q \to 0$ leads to a symmetric antiresonance (Breit–Wigner dip) while the opposite limit $q \to \infty$ leads to a resonance peak. We now define units as follows. For *GaAs* the value of $\hbar^2/2m \approx 0.57\,\text{eV nm}^2$. We take $\hbar^2/2m$ equal to unity and therefore the unit of the matrix element υ_{nl} can be taken to be 0.1 eV nm which then yields an energy unit of ≈ 17.7 meV and a length unit of ≈ 5.7 nm. Then, we can write the asymmetry parameter with the help of Eqs. (33) and (34) as

$$q = \frac{\delta}{\Gamma} = -\frac{\cos\left[(\pi/2)\sqrt{1+(4U_{11}/\alpha^2)}\right]}{\sinh(\pi k/\alpha)}. \tag{39}$$

The asymmetry parameter is a particularly important quantity in the detailed structure of the line shape and has been investigated experimentally [16,24] and theoretically [9].

Resonances in the Transmission Probability of a Quantum Wire

In this section we will discuss the influence of the fundamental parameters of the impurity discussed in Sect. "Electronic Transport Through Quantum Wires" on the resonance line shape. In the numerical calculations we take the electron mass to be the effective mass for *GaAs* which is 0.067 of the free electron mass.

Inversion of the Resonance Level

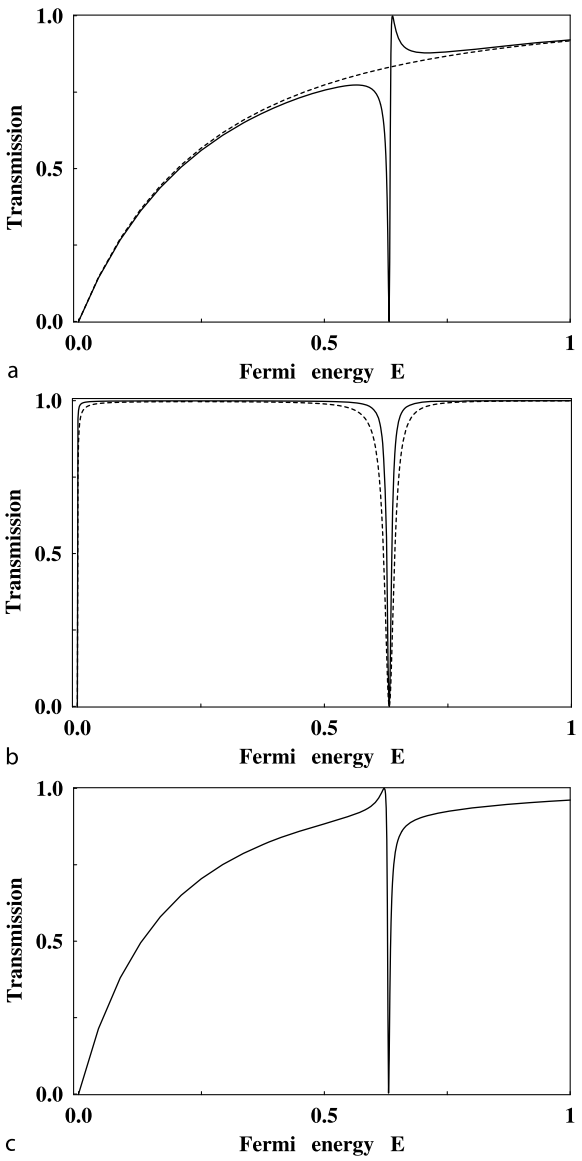
An important feature of Eq. (39) is that for certain ranges of values of U_{11}/α^2 the quantity δ (which is the difference between the resonance energy E_R and the energy $\tilde{E}_0$ of the transmission zero) can be negative, positive, or zero. This means that the resonance energy may occur before, after, or be equal to the energy of the transmission zero, which leads to an oscillatory behavior of the asymmetry parameter q. In particular, when $0 < U_{11}/\alpha^2 < 2$ we have $E_R > \tilde{E}_0$ and in this case the peak follows the dip. We call this a $0 \to 1$ type Fano resonance. However, for $2 < U_{11}/\alpha^2 < 6$ we have $E_R < \tilde{E}_0$ and the dip follows the peak. We call this a $1 \to 0$ type Fano resonance and, in this case, the resonance level appears inverted, which means that the the location of the pole is switched with the zero energy. For the interval $6 < U_{11}/\alpha^2 < 12$ we have $E_R > \tilde{E}_0$ and the Fano resonance is again of the $0 \to 1$ type. Exactly at the values of U_{11}/α^2, for which inversion of the resonance level occurs (i. e., at $U_{11}/\alpha^2 = 2, 6, \ldots$), δ and therefore q become zero. The special values, for which $\delta = 0$, are given as

$$\frac{U_{11}}{\alpha^2} = \frac{1}{4}[(2n+1)^2 - 1], \tag{40}$$

where $n = 1, 2, \ldots$ Since $q = 0$ at these special values, the transmission exhibits symmetric Breit–Wigner dips.

These effects are illustrated in Fig. 3 where we show the transmission probability through the attractive potential of Eq. (19) plotted versus the Fermi energy over the first subband, i. e., $E_1 < E < E_2$. In Fig. 3a we have chosen $U_{11} = 1.5$, $U_{22} = 1.4$, $\upsilon_{12} = 0.09$, and $\alpha = 1.6$. The transmission exhibits an asymmetric Fano resonance (solid line) of the $0 \to 1$ type while the dashed line in Fig. 3a represents the direct (nonresonant) transmission, which occurs in the decoupling limit $\upsilon_{12} = 0$ (i. e., in the limit in which there is no coupling of the bound state with the continuum of states of the first subband) for which $\delta = \Gamma = 0$. In this limit there are two scattering mechanisms: a direct scattering from the first subband and a resonant scattering from the quasibound state. When the coupling to the quasibound level becomes nonzero the interference between direct and nonresonant transmission pathways produces the asymmetric Fano line shape [23,30,37,39]. The resonance width is proportional to the square of the coupling, i. e., $\Gamma \sim \upsilon_{12}^2$, as can be seen from Eq. (34).

In Fig. 3b we have chosen, for both solid and dashed lines, $\alpha = 1$, $U_{11} = 2$, and $U_{22} = 1$ but $\upsilon_{12} = 0.095$ has been chosen for plotting the solid line and $\upsilon_{12} = 0.14$ for the dashed line. These give $U_{11}/\alpha^2 = 2$, which is one of the special values of Eq. (40), and it can be seen that the transmission exhibits a completely symmetric antiresonance (Breit–Wigner dip). This is due to the fact that the transmission zero coincides with the energy of the resonance level. Furthermore, as mentioned above, the width of the resonance increases as the intersubband coupling υ_{12} increases, which is shown by the dashed line. It's also



Resonances in Electronic Transport Through Quantum Wires and Rings, Figure 3

Transmission probability *T* versus Fermi energy *E* through an attractive impurity potential of the Pöschl–Teller type, in a 2D uniform quantum wire (see text for the numerical values of the parameters used). a An asymmetric Fano resonance of the 0 → 1 type (*solid*) and the background transmission (*dashed*). b symmetric Breit–Wigner antiresonance for one of the special values given in the text, at which the resonance energy coincides with the energy of the transmission zero. The *dashed line* corresponds to larger coupling parameter which gives rise to larger resonance width. c A 1 → 0 type Fano resonance indicating inversion of the resonance level

interesting to note that there is perfect transmission over the whole energy window, except for the very small energy interval where the antiresonance occurs. This is due to the fact that the scattering potential of Eq. (19) varies smoothly and the direct transmission $|t^{bg}|^2$ becomes unity [see Eq. (38)] whenever U_{11}/α^2 assumes the special values given in Eq. (40). In Fig. 3c, we have chosen $\alpha = 1.4$, $U_{11} = 6.3$, $U_{22} = 1.2$, and $\upsilon_{12} = 0.8$. The resonance line shape is now of the $1 \to 0$ type. The inversion of the resonance level indicates that the roles of destructive and constructive interference between the direct and resonant channels have been reversed.

The resonance level inversion discussed in the context of Fig. 3, has also been predicted in References [23] and [22] where the scattering potential was modeled by a rectangular square well. Increasing the longitudinal size of the impurity, it was found that there are some critical values of the impurity size for which the Fano resonance disappears. This occurs over an energy interval around a critical value E_c where a transformation of the quasi-bound state into a true bound state (for which $\Gamma \to 0$) occurs, giving rise to unity transmission in this interval.

We also mention that the Fano asymmetry parameter q is a particularly important quantity for the detailed structure of the line shape, and has been investigated experimentally and theoretically in the context of various types of mesoscopic systems [9,14,16,24]. In our case, q can be evaluated at the resonance energy from Eq. (39). The calculated values are $q = 0.47$, 0, and -0.43 in Figs. 3a, b, and c, respectively. Furthermore, we note that resonance phenomena in quantum wires with impurities can also lead to a connection between the line shape and the fundamental parameters of the impurity, such as its strength and size.

Temperature Dependence of the Fano Effect

So far we have considered the zero-temperature case. At finite temperatures the resonances are rapidly smeared out. It will be shown that the effect of temperature (due to thermal broadening) becomes gradually weaker with increasing size of the PT impurity given in Eq. (19).

In order to consider thermal effects, we employ the finite-temperature conductance formula

$$G(\mu, T) = G_0 \int \mathrm{d}E \left(\frac{\partial f}{\partial \mu} \right) G(E, 0) \,, \tag{41}$$

where $G(E, 0)$ is the zero-temperature conductance and f is the Fermi distribution function given by

$$f(E) = \frac{1}{\mathrm{e}^{(E-\mu)/k_\mathrm{B}T} + 1} \,. \tag{42}$$

For comparison, we also consider the case of an attractive short-range (SR) impurity. The SR impurity is modeled by the Dirac δ function along the propagation direction

$$V(x, y) = \frac{\hbar^2 \gamma}{2m} \delta(x) \upsilon(y) \,, \tag{43}$$

where, as in the case of the PT impurity, $\upsilon(y)$ is an arbitrary function of the coordinate y, and $\gamma < 0$. For this scattering potential, the bound-state energy is found to be $\tilde{E}_0 = E_2 - (U_{22}^2/4)$. The shift and width that the bound state acquires can also be found analytically as

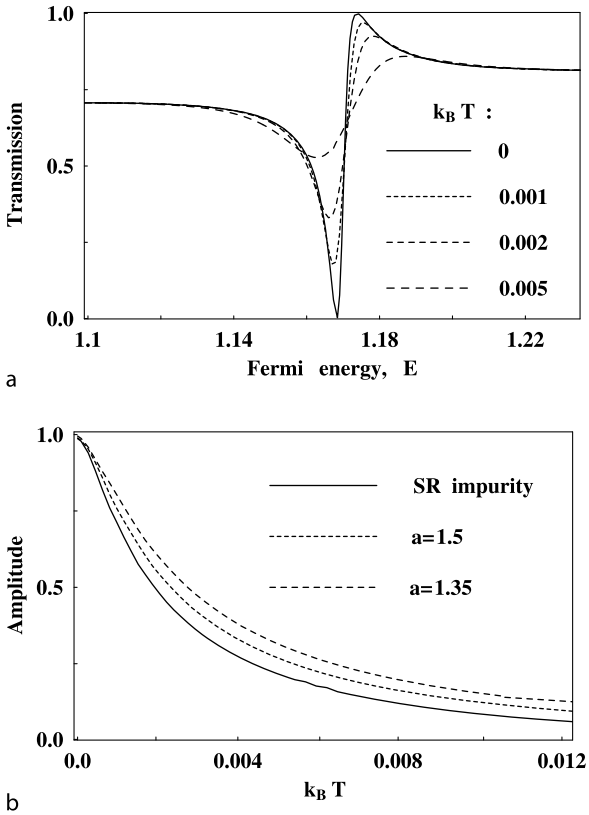
$$\delta = \frac{U_{11} U_{22} U_{12}^2}{8k^2 + 2U_{11}^2} \,, \tag{44}$$

and

$$\Gamma = \frac{k U_{12}^2 U_{22}}{4k^2 + U_{11}^2} \,. \tag{45}$$

In our energy unit (≈ 17.7 meV) that we employ in this article, $k_B T = 0.001$ corresponds to 205 mK. The temperature dependence of the Fano resonance for the SR impurity is shown in Fig. 4a. The corresponding plot in the case of a PT impurity is similar. In Fig. 4a we have chosen $U_{11} = 1$, $U_{22} = 0.9$, and $\upsilon_{12} = 0.08$. We note that the resonance structure is rapidly smeared out with increasing temperature. That is, the Fano resonance becomes gradually broader while its amplitude, $\Delta G = G_{\max} - G_{\min}$, decreases. However, this amplitude is influenced by the type and size of the impurity in a significant way. The amplitude versus $k_B T$ is plotted in Fig. 4b for both a SR and a PT impurity for two values of α, i. e., for $\alpha = 1.5$ and 1.35. The fastest decrease of the amplitude with increasing temperature occurs for a SR impurity for which $\Delta G \simeq 0.1$ at $k_B T \simeq 0.012$ (or $T \simeq 2.5$ K). However, depending on the size α^{-1} of the PT impurity, the Fano resonance persists for higher temperatures. That is, the effect of temperature becomes progressively weaker as the size of the finite-range PT impurity increases.

The above-mentioned smearing of the Fano resonance can be explained in terms of the thermal broadening, via the smooth peak in $\partial f/\partial \mu$, which obscures the resonance as $k_B T$ becomes comparable to the resonance width. In the case of a SR impurity [solid line of Fig. 4b], the amplitude of the resonance becomes half its zero-temperature value as soon as $k_B T \simeq 0.0018$ (or, equivalently, $T \simeq 380$ mK), which is comparable to the calculated resonance width $\Gamma \simeq 0.0022$. However, as has been shown elsewhere [40], in the case of a PT impurity the resonance width increases with decreasing α (i. e., with increasing size), resulting in a gradually broader resonance. Therefore, for gradually



Resonances in Electronic Transport Through Quantum Wires and Rings, Figure 4

The temperature dependence of the transmission probability T versus Fermi energy E through an attractive δ function potential, $V(x, y) = (\hbar^2 \gamma/2m)\delta(x)\upsilon(y)$, in a uniform 2D quantum wire is shown in a. In our energy unit (see text) $k_B T = 0.001$ corresponds to 205 mK. Note the smearing of the Fano resonance as the temperature increases. b Amplitude, $\Delta G = G_{\max} - G_{\min}$, of the Fano resonance versus $k_B T$ for the δ function (*solid line*) and the PT (*dashed lines*) impurities. Note the progressively weaker effect of temperature for PT impurities with gradually larger size (i. e., gradually smaller α)

larger PT impurities, the Fano resonance requires higher temperatures in order to diminish [see dashed lines of Fig. 4b]. For a PT impurity with $\alpha = 1.5$, the amplitude becomes half its zero-temperature value at $k_B T \simeq 0.0024$ (which corresponds to $T \simeq 480mK$), while the resonance width is $\Gamma \simeq 0.0026$. The above-mentioned broadening of the Fano resonance is reflected in an almost linear increase of the resonance width [31].

Electronic Transport Through Mesoscopic Open Rings

As mentioned above, resonances of the Fano type have been treated theoretically and observed experimentally in

various mesoscopic systems. One particular class of mesoscopic systems in which the Fano effect shows up, is the 1D ring connected to current leads with [41,43] or without [18] an impurity in one of its arms. In such a system, the electron motion between the junctions is treated as purely one-dimensional, i. e., no subband structure is taken into account and therefore no intersubband interaction occurs. Thus, the usual interpretation of the Fano effect as being due to quasibound states splitting off from nonpropagating subbands and the subsequent coupling with the continuum of states, does not apply in this system. Fano resonances have also been studied in various other ring geometries, such as a ring with coupling between the leads [47], and an ideal 1D double ring [48], to name a few.

The presence of an Aharonov–Bohm flux leads to further interesting effects in mesoscopic open rings. The application of an AB magnetic flux between the two electronic paths leads to the interference of the electronic wave functions, giving rise to a circulating (or persistent) current which is periodic in the magnetic flux with period Φ_0, where $\Phi_0 = hc/e$ is the flux quantum. Great effort has been devoted to the description of this effect in various open ring geometries [7,10,17,18,33,47,48]. The persistent current has also been related to the collapse of Fano resonances in Reference [41].

Most of the investigations on electronic transport in 1D open rings employ the Griffith's boundary condition (conservation of current) and continuity of wave functions at the junctions. Even though useful, the Griffith's boundary condition constitutes a significant limitation on the possible scattering effects at the junctions on the transmission probability. The reason is that all the scattering effects are included in the calculation via a single value of the ring-lead coupling strength, whereas the actual coupling strength may vary, in a continuous manner, from weak to strong. Therefore, the imposition of Griffith's boundary condition is insufficient to fully describe electron scattering at the junctions and important aspects of electronic transport are neglected [41]. After reviewing in the next section the more general S-matrix approach, which is suitable for variable coupling strength, we will illustrate the effect of this coupling on the transmission resonances. We point out that the existence of Fano resonances in 1D mesoscopic rings has been known for a long time [3]. However, recently [41,43] the dependence of the Fano resonance on the system parameters has been studied, in the zero-flux case, and revealed an important feature; namely, the collapse of a Fano resonance when the system parameters are commensurate. Particularly, there are special impurity positions for which certain Fano resonances collapse.

S-Matrix Approach

We consider the electronic transport through a ring connected to current leads in the presence of an AB magnetic flux, as shown in Fig. 5. The leads and the ring are taken to be strictly 1D, i. e., no subband structure is taken into account. We consider the current flow to be from left to right. At a junction of a lead with the ring, the three outgoing waves with amplitudes $(\alpha', \beta', \gamma')$ are related by an S matrix to the three incoming waves (α, β, γ) as

$$\mathbf{O} = S\mathbf{I}\,, \tag{46}$$

where $\mathbf{O}$ represents the outgoing and $\mathbf{I}$ the incoming waves. Current conservation and time-reversal invariance imply that S is unitary and symmetric [3], and given as

$$S = \begin{pmatrix} -(a+b) & \sqrt{\varepsilon} & \sqrt{\varepsilon} \\ \sqrt{\varepsilon} & a & b \\ \sqrt{\varepsilon} & b & a \end{pmatrix}. \tag{47}$$

Due to current conservation,

$$a = \frac{1}{2}\left(\sqrt{1-2\varepsilon} - 1\right), \tag{48}$$

$$b = \frac{1}{2}\left(\sqrt{1-2\varepsilon} + 1\right). \tag{49}$$

Therefore, the coefficients a and b are expressed as functions of the single parameter ε, where $0 \le \varepsilon \le 1/2$. For $\varepsilon = 0$, electrons from the current lead are totally reflected and thus there is no coupling between the current leads and the ring. For $\varepsilon = 1/2$, the junction is completely transparent for incoming electrons and the current lead is strongly coupled to the ring. It can be verified that the Griffith's boundary condition corresponds to $\varepsilon = 4/9$.

In the presence of an impurity and a magnetic flux Φ, the Schrödinger equation that describes the electron motion in the ring is

$$\left[-\frac{\hbar^2}{2m}\left(\frac{\partial}{\partial y} - i\frac{2\pi}{L}\frac{\Phi}{\Phi_0}\right)^2 + V_i(y)\right]\psi(y) = E\psi(y)\,, \tag{50}$$

where y is the coordinate along the ring, $\Phi_0 = hc/e$ is the flux quantum associated with a single charge e, L is the circumference of the ring, and $V_i(y)$ is the impurity potential in the upper arm (at point D of Fig. 5). The lengths of the two arms, L_1 and L_2, are taken to be equal, i. e., $L_1 = L_2 = L/2$. We consider an impurity potential modeled by a Dirac δ function

$$V_i(y) = \gamma\delta(y - y_i)\,, \tag{51}$$

where y_i is the impurity position in the arm, and γ sets the magnitude of the impurity potential, which is taken to be repulsive (i. e., $\gamma > 0$). The wave functions in both leads are written as

$$\psi_l(x) = \alpha_j e^{ikx} + \alpha'_j e^{-ikx} , \tag{52}$$

where $j = 1(2)$ for the left(right) leads, respectively, x is the coordinate along the leads, and $k = (2mE)^{1/2}/\hbar$ is the wave vector. In the ring the wave functions are expressed as

$$\psi_2(y) = e^{2\pi i(\Phi/\Phi_0)(y/L)} \left(\beta'_1 e^{iky} + \beta_1 e^{-iky}\right) , \tag{53}$$

$$\psi_3(y) = e^{2\pi i(\Phi/\Phi_0)(y/L)} e^{-i\pi(\Phi/\Phi_0)} \left(\beta_2 e^{-ik(L_1-y)} + \beta'_2 e^{ik(L_1-y)}\right) , \tag{54}$$

$$\psi_4(y) = e^{-2\pi i(\Phi/\Phi_0)(y/L)} \left(\gamma'_1 e^{iky} + \gamma_1 e^{-iky}\right) , \tag{55}$$

where $\psi_2(y)$ and $\psi_3(y)$ are the wave functions before and after the impurity, respectively, in the upper arm while $\psi_4(y)$ is the wave function in the lower arm (see Fig. 5). With the help of a transfer matrix, the amplitudes β_2 and β'_2 to the right of the impurity are expressed in terms of the amplitudes β'_1 and β_1 to the left of the impurity. The transfer matrix is found using a standard procedure, i. e., by applying the boundary conditions at the impurity. We then get

$$\begin{pmatrix} \beta_2 \\ \beta'_2 \end{pmatrix} = e^{i\pi(\Phi/\Phi_0)} T_i \begin{pmatrix} \beta'_1 \\ \beta_1 \end{pmatrix} , \tag{56}$$

where the transfer matrix T_i is given as

$$T_i = \begin{pmatrix} (1/t^*)\, e^{iq} & (r/t)\, e^{iq} e^{-2iqs_i} \\ (r^*/t^*)\, e^{-iq} e^{2iqs_i} & (1/t)\, e^{-iq} \end{pmatrix} . \tag{57}$$

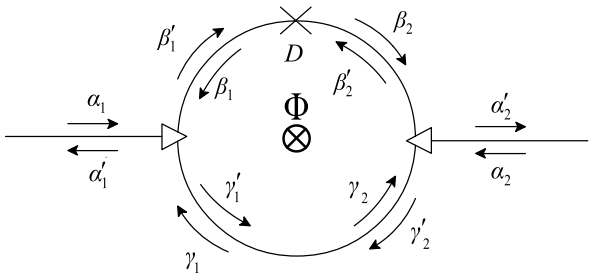
In Eq. (57), we have used a dimensionless wave vector defined as $q = kL_1$ and the dimensionless impurity position $s_i = y_i/L_1$. Note that q has also been used previously to denote the Fano asymmetry parameter but its meaning becomes clear from the particular context. With this notation, the transmission and reflection amplitudes of the impurity potential are expressed as

$$t = \frac{2iq}{2iq - V} , \tag{58}$$

and

$$r = \frac{V}{2iq - V} , \tag{59}$$

where we have used the dimensionless parameter $V = 2m\gamma L_1/\hbar^2$ to characterize the impurity strength. We note



Resonances in Electronic Transport Through Quantum Wires and Rings, Figure 5

A one-dimensional ring connected to current leads with an impurity at point D (in the upper arm) and an Aharonov–Bohm flux Φ. The $\alpha's$, $\beta's$, and $\gamma's$ represent the amplitudes of the wave functions (see text)

that an impurity with a small spatial extension d and a magnitude of $\bar{\gamma}$ can be approximated by a δ function with $\gamma = d\bar{\gamma}$ and therefore $V = 2m\bar{\gamma}dL_1/\hbar^2$. If we choose an experimentally realizable [20,21,24] arm length $L_1 = 1\,\mu$m, the effective mass for *GaAs* ($m = 0.067m_0$), and $d = 0.01\,\mu$m, the unit of the impurity parameter $V = 1$ corresponds to $\bar{\gamma} = 0.12$ me.

We consider now a wave of unit amplitude $\alpha_1 = 1$, incident from the left. To find the transmission probability $T = |\alpha'_2|^2$ of the ring we have to find the amplitudes β'_1, β'_2, γ'_1, γ'_2 under the condition $\alpha_2 = 0$. Therefore, we employ the S-matrix at the two junctions. Applying Eqs. (46) and (47) at the right junction we can express the $\gamma's$ in terms of the $\beta's$ as

$$\begin{pmatrix} \gamma'_2 \\ \gamma_2 \end{pmatrix} = T_j \begin{pmatrix} \beta_2 \\ \beta'_2 \end{pmatrix} , \tag{60}$$

where T_j is a matrix given as

$$T_j = \frac{1}{b} \begin{pmatrix} (b^2 - a^2) & a \\ -a & 1 \end{pmatrix} . \tag{61}$$

Applying Eqs. (46) and (47) at the left junction and using $\alpha_1 = 1$, we are able to express β_1, β'_1 in terms of γ_1, γ'_1 as

$$\begin{pmatrix} \beta'_1 \\ \beta_1 \end{pmatrix} = \frac{\sqrt{\varepsilon}}{b} \begin{pmatrix} b - a \\ -1 \end{pmatrix} + T_j \begin{pmatrix} \gamma_1 \\ \gamma'_1 \end{pmatrix} . \tag{62}$$

In the lower arm we insert $y = L_2$ into $\psi_4(y)$ of Eq. (55) which then yields two terms, the first being γ_2 and the second γ'_2. Eliminating γ_2 and γ'_2 from these two terms by employing Eq. (60) yields a relation between γ_1, γ'_1 and β_2, β'_2 with the help of a matrix T'_j given as

$$T'_j = \frac{1}{b} \begin{pmatrix} (b^2 - a^2)\, e^{iq} & a e^{iq} \\ -a e^{-iq} & e^{-iq} \end{pmatrix} . \tag{63}$$

Using the matrix Π defined as

$$\Pi = T_j e^{i\pi(\Phi/\Phi_0)} T'_j e^{i\pi(\Phi/\Phi_0)} T_i - 1 \,, \tag{64}$$

and employing the relation $b - a = 1$, the transmission amplitude can finally be expressed as

$$\alpha'_2 = -\frac{\varepsilon}{b^2} e^{i\pi(\Phi/\Phi_0)} \frac{p}{\det(\Pi)} \,, \tag{65}$$

where

$$p = \det(\Pi)[(\ 1 \quad 0\ ) + (\ 0 \quad 1\ )] T_i \Pi^{-1} \begin{pmatrix} 1 \\ -1 \end{pmatrix} . \tag{66}$$

The poles of the transmission amplitude, i. e., the solutions of $\det(\Pi) = 0$, determine the resonant behavior of the transmission probability. The transmission probability of the ring can then be written as

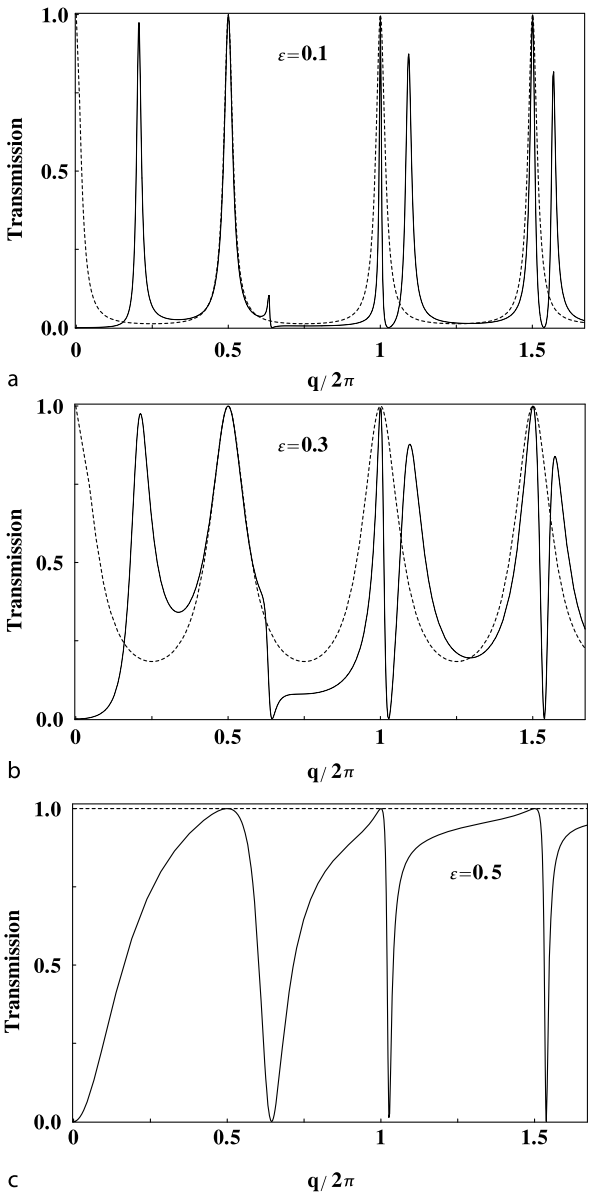
$$T = |\alpha'_2|^2 = \frac{\varepsilon^2}{b^4} \frac{|p|^2}{|\det(\Pi)|^2} \,. \tag{67}$$

Fano Resonances in the Transmission Probability

In the numerical calculations we take the impurity position at $s_i = 0.42$ and its scattering strength to be $V = 11$. We first consider the zero-flux case.

We show by the solid lines in Fig. 6 the transmission probability T versus the (dimensionless) wave vector q through a ring with an impurity in its upper arm, for three values of the coupling parameter ε. The dashed lines show the transmission probability of a ring with no impurity. In the weak coupling regime of Fig. 6a, we note that for the no impurity case sharp Breit–Wigner peaks occur at $q = n\pi$, where $n = 1, 2, \ldots$ These resonances are the result of constructive interference of the electron waves in the two arms and they become gradually broader with increasing coupling. In the strong coupling regime of Fig. 6c, the ring becomes completely transparent for incoming electrons resulting in unity transmission.

In the presence of an impurity and in the weak coupling regime of Fig. 6a we identify two types of transmission resonances; Breit–Wigner and Fano type. Most of the Breit–Wigner peaks are similar to those of the no impurity case, but note that they occur at different positions on the q axis from those of the no impurity case and their amplitudes can be less than unity. These resonances are attributed to the constructive scatterings from the junctions and the impurity. Thus, in the weak coupling regime, it's possible for electrons entering the ring to spend enhanced periods of time in the ring before being transmit-



Resonances in Electronic Transport Through Quantum Wires and Rings, Figure 6

Transmission probability T versus dimensionless wave vector $q/2\pi$, where $q = kL_1$, through a one-dimensional ring with equal arms ($L_1 = L_2$) connected to current leads, for three values of the coupling parameter ε. The *dashed lines* represent the case in which there is no impurity. The *solid lines* represent the case in which there is an impurity, $V_i(y) = (\hbar^2 V/2mL_1)\delta(y - y_i)$, in the upper arm (see Fig. 5). We have employed the value $V = 11$ for the impurity strength, where V is defined in the text, and $s_i = 0.42$, where $s_i = y_i/L_1$ and $L_1 = 1\ \mu$m. Depending on the coupling, there may be both Breit–Wigner and Fano line shapes. In the strong coupling regime, $\varepsilon = 0.5$, only the Fano resonances survive. Note that while varying the coupling the positions of the transmission zeros and ones remain fixed

ted. For these electrons the interference is constructive and the Breit–Wigner resonances are narrow.

When the coupling increases to $\varepsilon = 0.3$, scattering effects at the junctions become weaker, resulting in larger transmissivity and larger widths of the Breit–Wigner resonances. The larger resonance widths correspond to larger imaginary parts of the associated poles. In addition, in the limit $\varepsilon \to 0.5$ the poles move away from the real axis resulting in the disappearance of the Breit–Wigner resonances [41], as shown in Fig. 6c.

The Fano resonances occur at $q \approx n\pi$, and in fact their transmission ones occur exactly at $q = n\pi$. These resonances are not much sensitive to the perturbation caused by the junctions, and the transmission probability exhibits Fano line shapes even in the strong coupling regime, as shown in Fig. 6c. Furthermore, increasing the coupling does not affect the positions of the transmission zeros and ones, but note that the asymmetry parameters change. This is reflected in a modification of the resonance width q_i, where q_i is the imaginary part of the associated pole in the complex q plane, with increasing coupling. For example, for the Fano resonance at $q \approx 2\pi$, the difference ξ between the transmission zero q_0 and one q_1, i. e., $\xi = q_0 - q_1$, has the value $\xi = 0.06782$ and remains constant for any value of the coupling parameter ε. However, the values of the associated pole are $q = 6.291 - i0.0182$ for $\varepsilon = 0.1$, whereas for $\varepsilon = 0.3$ the pole location is at $q = 6.312 - i0.039$ and for $\varepsilon = 0.5$ at $q = 6.351 - i0.0302$. To further describe this in the framework of Fano's asymmetric function [12] written in terms of the wave vector q [see Eq. (69)], we note that ξ can be expressed [41] as $\xi = q_i\left(|\tilde{q}| + 1/|\tilde{q}|\right)$, where q_i is the resonance width and $\tilde{q}$ the asymmetry parameter. We see that, as the asymmetry parameter $\tilde{q}$ varies, the width q_i also varies in such a way that ξ remains constant.

Thus, in the absence of impurity, the transmission exhibits Breit–Wigner resonances in the weak coupling regime. However, increasing V from zero in this regime, leads to shifting of the Breit–Wigner resonances on the q axis, but also results in a transmission zero, thus giving rise also to a Fano line shape.

Tuning of the Fano Resonance

Recently [41] am interesting feature of the Fano line shape in 1D open rings was discovered; namely, when the impurity is located at some special positions in the arm, certain Fano resonances collapse. Even though it is independent of the ring-lead coupling, we will illustrate this feature in the strong coupling regime $\varepsilon = 0.54$. In this regime, an analytic expression for the transmission amplitude can be found,

$$\alpha_2' = -\mathrm{e}^{-i\pi\Phi/\Phi_0} \cdot \frac{4iq(\mathrm{e}^{-iq} - \mathrm{e}^{iq}) - V(\mathrm{e}^{iq} + \mathrm{e}^{-iq}) + V(\mathrm{e}^{iD_iq} + \mathrm{e}^{-iD_iq})}{4iq - 4iq\mathrm{e}^{-2iq} - V\mathrm{e}^{-2iQ_iq} - V\mathrm{e}^{-2is_iq} + 2V\mathrm{e}^{-2iq}}, \tag{68}$$

where $D_i = 1 - 2s_i$, $Q_i = 1 - s_i$ and we have also employed the relation $t/r = 2iq/V$. In the vicinity of a Fano resonance, the transmission amplitude α_2' can be expressed in the asymmetric Fano form [30]. Expanding the numerator and denominator of Eq. (68) in a Taylor series around a transmission zero, q_0, and keeping only linear terms we get

$$\alpha_2' = t_d \frac{q - q_0 - \mathrm{Re}(\Sigma) + \delta}{q - q_0 - \Sigma}, \tag{69}$$

where $t_d = \eta_1/\eta_2$, $\delta = \lambda_1/\eta_1 = \widetilde{q}q_i$, and $\Sigma = \lambda_2/\eta_2 = \Delta - iq_i$ with

$$\eta_1 = 8q\cos q_0 + 2V\sin q_0 - 2VD_i\sin(q_0D_i), \tag{70}$$

$$\eta_2 = 8q\mathrm{e}^{-2iq_0} - 2iVQ_i\mathrm{e}^{-2iq_0Q_i} - 2is_iV\mathrm{e}^{-2iq_0s_i} + 4iV\mathrm{e}^{-2iq_0}, \tag{71}$$

$$\lambda_1 = 8q\sin q_0 - 2V\cos q_0 + 2V\cos(q_0D_i), \tag{72}$$

$$\lambda_2 = 4iq - 4iq\mathrm{e}^{-2iq_0} - V\mathrm{e}^{-2iq_0Q_i} - V\mathrm{e}^{-2iq_0s_i} + 2V\mathrm{e}^{-2iq_0}. \tag{73}$$

In Eq. (69), δ and Σ can be decomposed as $\delta = \tilde{q}q_i$ and $\Sigma = \Delta - iq_i$. Equation (69) will be used for the description of the collapse of Fano resonances. It will be shown that, for particular impurity positions, the expression for Σ vanishes which further indicates that the resonance width shrinks to zero and the Fano line shape collapses.

Let us consider an impurity located at s_i, where s_i is the ratio of two integers,

$$s_i = \frac{r}{n}, \tag{74}$$

where n is fixed and $r = 1, 2, \ldots, n$. This means that $0 \le s_i \le 1$. For an impurity located at one of the special positions given by Eq. (74), the expression for Σ takes the form

$$\Sigma = \frac{4iq - 4iq\mathrm{e}^{-2iq_0} - V\mathrm{e}^{-2iq_0(1-s_i)} - V\mathrm{e}^{-2iq_0s_i} + 2V\mathrm{e}^{-2iq_0}}{8q\mathrm{e}^{-2iq_0} - 2iV(1-s_i)\mathrm{e}^{-2iq_0(1-s_i)} - 2is_iV\mathrm{e}^{-2iq_0s_i} + 4iV\mathrm{e}^{-2iq_0}}. \tag{75}$$

Now let $q_0 \to \mu n\pi$, where n is the same (fixed) integer as that in Eq. (74) and $\mu = 1, 2, \ldots$. For these values of q_0, we note that the numerator of Σ in Eq. (75) vanishes which implies that the widths q_i, of all Fano resonances which occur at integer multiples of $n\pi$, that is at $q \simeq n\pi, 2n\pi, 3n\pi, \ldots$, shrink to zero and therefore collapse. Note that δ also vanishes. In this case, the transmission amplitude α_2' is equal to the nonresonant transmission amplitude t_d and given as $\alpha_2' = t_d = -[1 + (iV/4q)]^{-1}$. To illustrate this behavior, we plot in Fig. 7 the exact transmission probability from Eq. (68) versus the wave vector q, for three values of the impurity position (i. e., for $s_i = 1/2$, 1/3, and 1/4). For comparison, we also plotted the approximate transmission probability with α_2' given by Eq. (69) in Fig. 7c in the vicinity of $q \approx 3\pi$, for the same impurity parameters as those used for plotting the exact expression. We note that there is good agreement with the exact line shape. In Fig. 7 the value of the impurity strength is taken to be $V = 12$.

In agreement with the above discussion related to Eqs. (74) and (75), we note in Fig. 7a that placing the impurity at $s_i = 1/2$, which means $n = 2$ in Eq. (74), causes simultaneous collapse of all Fano resonances at $q \approx 2\pi, 4\pi, 6\pi, \ldots$. Furthermore, as shown in Fig. 7b, an impurity placed at $s_i = 1/3$, which means $n = 3$ in Eq. (74) causes simultaneous collapse of all Fano resonances at $q \approx 3\pi, 6\pi, \ldots$. For $s_i = 1/4$, collapse of all Fano resonances at $q \approx 4\pi, 8\pi, \ldots$ occurs, as shown in Fig. 7c. The above-mentioned collapsing behavior has also been described in terms of the pole structure of the transmission amplitude, in the complex q plane, as a function of the impurity position. It has been shown [41] that as the impurity approaches, for example $s_i = 1/2$, the pole associated with $q = 2\pi$ approaches the real axis and, exactly at $s_i = 1/2$ the pole is located on the real axis of the q plane implying that the resonance width becomes zero.

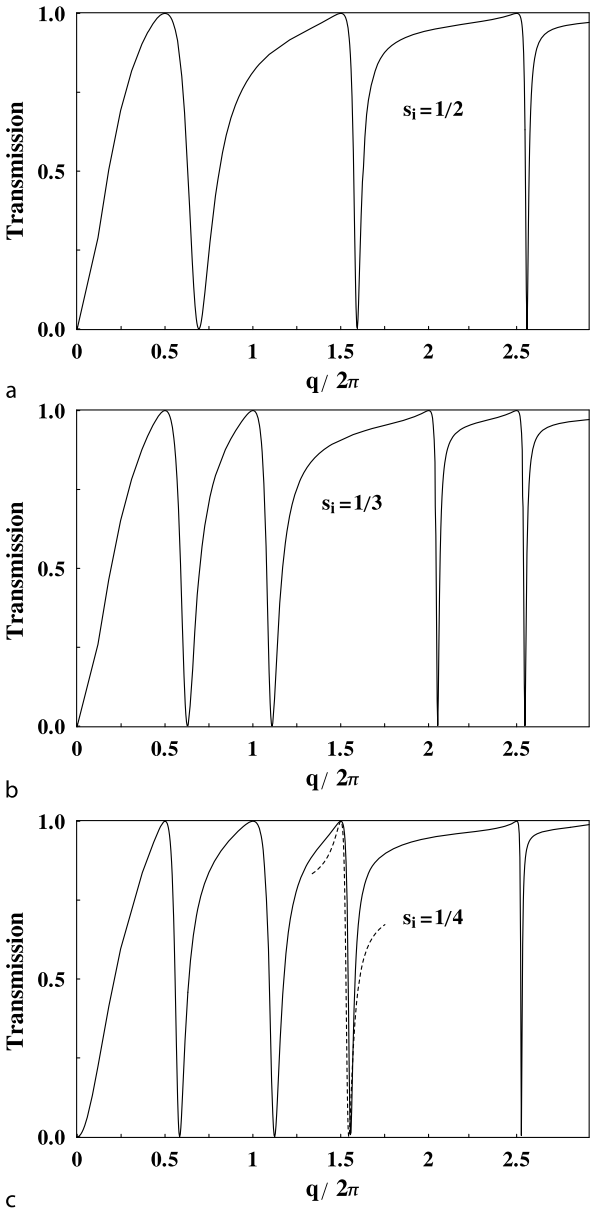
We also note that we can write Eq. (74) as $s_i = r\pi/n\pi = r\pi/q_n$, where $q_n = n\pi$ are the wave numbers at which the Fano resonances occur. This means that, given the impurity position s_i, enables us to find all Fano profiles that collapse from

$$q = \frac{r\pi}{s_i} . \tag{76}$$

For example, at $s_i = 1/3$, all Fano profiles that occur at

$$q = \frac{r\pi}{(\frac{1}{3})} = 3r\pi = 0, 3\pi, 6\pi, \ldots , \tag{77}$$

collapse. At the extreme positions $s_i = 0$ and 1, all Fano resonances collapse. This tunability of the Fano reso-



Resonances in Electronic Transport Through Quantum Wires and Rings, Figure 7

Transmission probability T versus dimensionless wave vector $q/2\pi$, where $q = kL_1$, through a 1D ring with equal arms connected to current leads with an impurity of strength $V = 12$ in the upper arm. For the impurity position we have chosen $s_i = 1/2$, 1/3, and 1/4 in parts a, b, and c respectively. In a, all Fano resonances at $q \approx \mu 2\pi = 2\pi, 4\pi, 6\pi, \ldots$, collapse. In b and c, the Fano resonances collapse at $q \approx \mu 3\pi$ and $q \approx \mu 4\pi$, respectively, where $\mu = 1, 2, 3, \ldots$ The *dashed line* in c represents the approximate expression given in Eq. (69)

nances with respect to the impurity position may provide useful means for the design of electronic nanodevices. In this context the impurity potential may be provided by a movable scanning tunneling microscope tip. Note also that this tunability is independent of the scattering strength V of the impurity.

Further interesting issues related to open 1D rings are how the ring-lead coupling, impurity strength and position affect the persistent current [18,32,34,41,45,46,47,48]. Even though we will not go into details here, we would like to mention that the persistent current versus magnetic flux increases dramatically whenever the impurity approaches one of the values given in Eq. (74), i. e., $s_i \to r/n$ and, exactly at these values, the persistent current diverges [41]. This behavior can be related to the motion of the poles of the transmission amplitude in the complex q plane as a function of magnetic flux, taking into account the fact that the persistent current arises near these poles [32,33,34,41,48].

The effect of an AB magnetic flux on the Fano resonance structure is also interesting. It has been shown [41] that increasing the AB flux encircled by the ring results in additional phase changes of the electron waves in the upper and lower arms, leading to modification of the Fano interference pattern at the right junction. Applying gradually the magnetic flux causes weaker destructive and constructive interference and consequently transmission zeros and ones do not exist anymore. As a result of these interference effects, the amplitudes of the resonances usually decrease and become gradually distorted. These effects repeat periodically with period Φ_0, where $\Phi_0 = hc/\mathrm{e}$ is the flux quantum.

Temperature Dependence of the Fano Resonances

Until now we discussed the transmission resonances at zero temperature. As in the case of quantum wires discussed in Sect. "Temperature Dependence of the Fano Effect", to consider thermal effects we should employ the finite-temperature conductance formula, Eq. (41). In the integral of this equation we first make a change of variables from energy E to the dimensionless wave vector q as

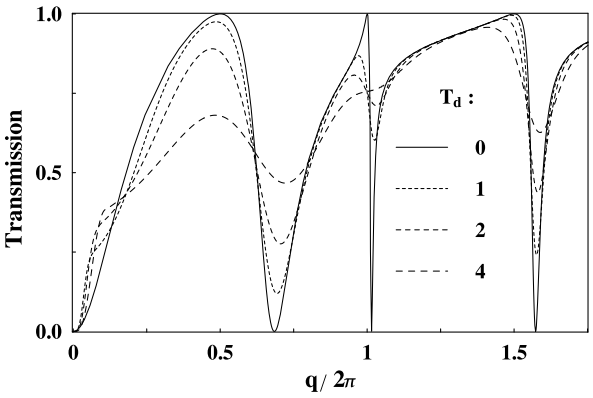
$$E = E^{\star} q^2 \,, \tag{78}$$

where $E^{\star} = \hbar^2/2mL_1^2$. In the following we define a dimensionless parameter $T_d = k_{\mathrm{B}}T/E^{\star}$ which is a measure of temperature. Then, the unit of this parameter $T_d = 1$ corresponds to $T = 6.59\,\mathrm{mK}$. After some manipulations, the transmission probability of Eq. (41) can be written as

$$\frac{G(q, T_d)}{G_0} = \frac{1}{2T_d} \int q' |\alpha_2'(q')|^2 \cosh^{-2}\left(\frac{q'^2 - q^2}{2T_d}\right) \mathrm{d}q' \,. \tag{79}$$

Restricting ourselves to the strong coupling limit $\varepsilon = 0.5$, we plot in Fig. 8 the transmission probability versus the wave vector q for several values of the parameter T_d, i. e., for $T_d = 0, 1, 2, 4$. In Fig. 8 we have taken the strength of the impurity $V = 11$ and its position $s_i = 0.47$. Note the strong temperature dependence of the narrower Fano resonances, especially of the resonance that occurs at $q \approx 2\pi$. The effect of temperature is relatively weaker for the wider resonances. For high enough temperatures the resonance structure diminishes.

The origin of this rapid smearing of the resonances is again the thermal broadening, via the smooth peak in $\partial f/\partial\mu$, which obscures the Fano resonance as $k_{\mathrm{B}}T$ becomes comparable to the resonance width. For the parameters chosen in the plot of Fig. 8, the amplitude of the Fano resonance at $q \approx 2\pi$ becomes half its zero-temperature value as soon as $T_d = 0.35$, which corresponds to $T \simeq 2.3\,\mathrm{mK}$ while the resonance width is $q_i \simeq 0.11$ (which corresponds to $T \simeq 0.27\,\mathrm{mK}$). This resonance diminishes at $T_d \approx 3.4$. However, the first resonance at $q \approx \pi$, which is much broader than the sec-



Resonances in Electronic Transport Through Quantum Wires and Rings, Figure 8

Temperature dependence of the transmission probability T versus dimensionless wave vector $q/2\pi$, where $q = kL_1$, through a 1D ring with equal arms connected to current leads. We have chosen $s_i = 0.47$ and $V = 11$ for the impurity position and strength, respectively. We have used the values $T_d = 0, 1, 2, 4$, where $T_d = k_{\mathrm{B}}T/E^{*}$ and $E^{*} = \hbar^2/2mL_1^2$. The unit of the parameter $T_d = 1$ corresponds to 6.59 mK. Note the rapid smearing of the Fano resonances (especially of the narrow ones) as the temperature increases

ond, as well as the third one at $q \approx 3\pi$ require higher temperatures in order to diminish. Particularly, the amplitude of the third resonance becomes half its zero-temperature value as soon as $T_d = 1.6$, which corresponds to $T \simeq 10\,\mathrm{mK}$, while the resonance width is $q_i \approx 0.27$ which corresponds to $T \simeq 0.65\,\mathrm{mK}$. The above-mentioned smearing of the Fano resonances has also been discussed in the context of various other mesoscopic systems [16,24,25,26,29].

Future Directions

In the past two decades advances in microfabrication have made possible the confinement of electrons in a conductor with lateral extension of 100 nm and less, resulting in narrow quantum wires, constrictions, and quantum dots. Consequently, it has been possible to engineer device potentials which vary over the length scale such that the electron motion is ballistic or quasiballistic at low temperatures. The ultrasmall size of these structures greatly eliminates electron scattering by defects and therefore one can obtain high mobility conducting channels. In these mesoscopic systems, electronic transport is governed by quantum mechanics rather than classical mechanics. At very low temperatures, the scattering by phonons is suppressed and the phase coherence length of the electron (i. e., the length over which the electron can be considered to be in a pure state) becomes larger than the system dimensions. As a result, the mesoscopic system becomes essentially an electron waveguide where the electronic transport properties are mainly determined by the impurity configuration, the geometrical characteristics of the conductor, and the principles of quantum mechanics. Convincing demonstrations of the quantum transport regime came from experiments on thin-metal or semiconductor films. Given the phase coherence of the electrons throughout the sample, several ideas for active quantum devices (such as, transistors, switches, etc.) have been proposed and are now under investigation. These are the quantum analogs of optical or microwave devices.

My intention in writing this article was to review certain important aspects of resonant electronic transport in mesoscopic systems. The resonant behavior of the transmission probability through such systems has actually been an issue which stimulated a great deal of experimental and theoretical research over the last two decades. The Fano function has been shown to arise as the most general resonance line shape, provided that two scattering channels – a resonant and a nonresonant one interfere. One important aspect of resonant electronic transport, which has actually been a problem common to many branches of physics, is to extract lifetimes of quasibound states from resonances in transmission coefficients or scattering cross sections. This requires knowledge of the resonance widths, which have been discussed in this article for two systems: i) a uniform quantum waveguide with a "smooth" finite size impurity, and ii) a 1D ring connected to current leads with an impurity in one arm. These resonances can also provide us with information about the parameters of the impurity potential. An interesting future experiment would be the determination of the impurity parameters from knowledge of the resonance line shapes.

The effect of temperature is also important since it destroys the Fano resonance structure. This depends on the size of the impurity on the quantum wire and also on the resonance width. The effect of temperature on the Fano resonances becomes stronger for impurities with smaller size while narrow resonances tend to diminish faster with increasing temperature.

An interesting feature is also the systematic collapse of certain Fano line shapes in electronic transport through a 1D ring with a short-range impurity. The collapse occurs for special values of the impurity position in the arm. As mentioned in this article, the tunability of the Fano resonances with respect to the impurity position could possibly prove useful for the design and engineering of electronic nanodevices. Even though it has not been discussed in this article, this tunability could further be useful for the generation of strong persistent currents [41].

Bibliography

Primary Literature

1. Adair RK, Bockelman CK, Peterson RE (1949) Experimental corroboration of the theory of neutron resonance scattering. Phys Rev 76:308–308
2. Büttiker M (1988) Symmetry of electrical-conduction. IBM J Res Dev 32:317–334
3. Büttiker M, Imry Y, Azbel MY (1984) Quantum oscillations in one-dimensional normal-metal rings. Phys Rev A 30:1982–1989
4. Bagwell PF (1990) Evanescent modes and scattering in quasi-one-dimensional wires. Phys Rev B 41:10354–10371
5. Banjamin C, Jayannavar AM (2003) Features in evanescent Aharonov–Bohm interferometry. Phys Rev B 68:085325-1–085325-6
6. Bardarson JH, Magnusdottir I, Gudmundsdottir G, Tang CS, Manolescu A, Gudmundsson (2004) Coherent electronic transport in a multimode quantum channel with Gaussian-type scatterers. Phys Rev B 70:245308-1–245308-10
7. Benjamin C, Jayannavar AM (2001) Current magnification effect in mesoscopic systems at equilibrium. Phys Rev B 64:233406-1–233406-4

8. Cerdeira F, Fjeldly TA, Cardona M (1973) Effect of free carriers on zone-center vibrational modes in heavily doped p-type Si. II. Optical modes. Phys Rev B 8:4734–4745
9. Clerk AA, Waintal X, Brouwer PW (2001) Fano resonances as a probe of phase coherence in quantum dots. Phys Rev Lett 86:4636–4639
10. Deo PS, Moskalets MV (2000) Features of level broadening in a ring-stub system. Phys Rev B 61:10559–10562
11. Faist J, Capasso F, Sirtori C, West KW, Pfeiffer LN (1997) Controlling the sign of quantum interference by tunneling from quantum wells. Nature 390:589–591
12. Fano U (1961) Effects of configuration interaction on intensities and phase shifts. Phys Rev 124:1866–1878
13. Fano U, Cooper JW (1965) Line profiles in the far-uv absorption spectra of the rare gases. Phys Rev 137:A1364–A1379
14. Feshbach H (1958) Unified theory of nuclear reactions. Ann Phys 5:357–390
15. Friedrich H (1990) Theoretical Atomic Physics. Springer, Berlin
16. Göres J, Goldhaber-Gordon D, Heemeyer S, Kastner MA, Shtrikman H, Mahalu D, Meirav U (2000) Fano resonances in electronic transport through a single-electron transistor. Phys Rev B 62:2188–2194
17. Jayannavar AM, Deo PS (1994) Persistent currents and conductance of a metal loop connected to electron reservoirs. Phys Rev B 49:13685–13690
18. Jayannavar AM, Deo PS (1994) Persistent currents in the presence of a transport current. Phys Rev B 51:10175–10178
19. Johnson AC, Marcus CM, Hanson MP, Gossard AC (2004) Coulomb-modified Fano resonance in a one-lead quantum dot. Phys Rev Lett 93:106803-1–106803-4
20. Keyser UF, Borck S, Haug RJ, Bichler M, Abstreiter G, Wegscheider W (2002) Aharonov–Bohm oscillations of a tuneable quantum ring. Semicond Sci Technol 17:L22–L24
21. Keyser UF, Fühner C, Borck S, Haug RJ, Bichler M, Abstreiter G, Wegscheider W (2003) Kondo effect in a few-electron quantum ring. Phys Rev Lett 90:196601-1–196601-4
22. Kim CS, Roznova ON, Satanin AM, Stenberg VB (2002) Interference of quantum states in electronic waveguides with impurities. JETP 94:992–1007
23. Kim CS, Satanin AM, Joe YS, Cosby RM (1999) Resonant tunneling in a quantum waveguide: Effect of a finite-size attractive impurity. Phys Rev B 60:10962–10970
24. Kobayashi K, Aikawa H, Katsumoto S, Iye Y (2002) Tuning of the Fano effect through a quantum dot in an Aharonov–Bohm interferometer. Phys Rev Lett 88:256806-1–256806-4
25. Kobayashi K, Aikawa H, Sano A, Katsumoto S, Iye Y (2003) Mesoscopic Fano effect in a quantum dot embedded in an Aharonov–Bohm ring. Phys Rev B 68:235304-1–235304-8
26. Kobayashi K, Aikawa H, Sano A, Katsumoto S, Iye Y (2004) Fano resonance in a quantum wire with a side-coupled quantum dot. Phys Rev B 70:035319-1–035319-6
27. Landau LD, Lifshitz EM (1965) Quantum Mechanics. Pergamon, New York
28. Landauer R (1987) Electrical transport in open and closed systems. Z Phys 68:217–228
29. Lee M, Bruder C (2006) Spin filter using a semiconductor quantum ring side coupled to a quantum wire. Phys Rev B 73:085315-1–085315-5
30. Nöckel JU, Stone AD (1994) Resonance line shapes in quasi-one-dimensional scattering. Phys Rev B 50:17415–17432
31. Olendski O, Mikhailovska (2003) Fano resonances of a curved waveguide with an embedded quantum dot. Phys Rev B 67:035310-1–035310-13
32. Orellana PA, Pacheco M (2005) Persistent current magnification in a double quantum-ring system. Phys Rev B 71:235330-1–235330-6
33. Pareek TP, Deo PS, Jayannavar AM (1995) Effect of impurities on the current magnification in mesoscopic open rings. Phys Rev B 52:14657–14663
34. Park W, Hong J (2004) Analysis of coherent current flows in the multiply connected open Aharonov–Bohm rings. Phys Rev B 69:035319-1–035319-7
35. Ryu CM, Cho SY (1998) Phase evolution of the transmission coefficient in an Aharonov–Bohm ring with Fano resonance. Phys Rev B 58:3572–3575
36. Satanin AM, Joe YS (2005) Fano interference and resonances in open systems. Phys Rev B 71:205417-1–205417-12
37. Tekman E, Bagwell PF (1993) Fano resonances in quasi-one-dimensional electron waveguides. Phys Rev B 48:2553–2559
38. Tekman E, Ciraci S (1990) Ballistic transport through a quantum point contact: Elastic scattering by impurities. Phys Rev B 42:9098–9103
39. Vargiamidis V, Polatoglou HM (2005) Conductance of a quantum wire with a Gaussian impurity potential and variable cross-sectional shape. Phys Rev B 71:075301-1–075301-16
40. Vargiamidis V, Polatoglou HM (2005) Resonances in electronic transport through a quantum wire with impurities and variable cross-sectional shape. Phys Rev B 72:195333-1–195333-12
41. Vargiamidis V, Polatoglou HM (2006) Fano resonance and persistent current in mesoscopic open rings: Influence of coupling and Aharonov–Bohm flux. Phys Rev B 74:235323-1–235323-14
42. Vargiamidis V, Polatoglou HM (2007) Temperature dependence of the Fano effect in quantum wires with short- and finite-range impurities. Phys Rev B 75:153308-1–153308-4
43. Voo KK, Chu CS (2005) Fano resonance in transport through a mesoscopic two-lead ring. Phys Rev B 72:165307-1–165307-9
44. Wharam DA, Thornton TJ, Newbury R, Pepper M, Ahmed H, Frost JEF, Hasko DG, Peacock DC, Ritchie DA, Jones GAC (1988) One-dimensional transport and the quantization of the ballistic resistance. J Phys C 21:L209–L214
45. Wu HC, Guo Y, Chen XY, Gu BL (2003) Giant persistent current in a quantum ring with multiple arms. Phys Rev B 68:125330-1–125330-5
46. Wunsch B, Chudnovskiy A (2003) Quasistates and their relation to the Dicke effect in a mesoscopic ring coupled to a reservoir. Phys Rev B 68:245317-1–245317-9
47. Xiong YS, Liang XT (2004) Fano resonance and persistent current of a quantum ring. Phys Lett A 330:307–312
48. Yi J, Wei JH, Hong J, Lee SI (2001) Giant persistent currents in the open Aharonov–Bohm rings. Phys Rev B 65:033305-1–033305-4
49. van Wees BJ, van Houten H, Beenakker CWJ, Williamson JG, Kouwenhoven LP, van der Marel D, Foxon CT (1988) Quantized conductance of point contacts in a two-dimensional electron gas. Phys Rev Lett 60:848–850

Books and Reviews

Beenakker CWJ, van Houten H (1991) Quantum transport in semiconductor nanostructures. In: Ehrenreich H, Turnbull D (eds) Solid State Physics, vol 44. Academic Press, New York

Datta S (1995) Electronic Transport in Mesoscopic Systems. Cambridge University Press, Cambridge
Economou EN (1983) Green's Functions in Quantum Physics. Springer, Heidelberg
Imry Y (1986) Physics of mesoscopic systems. In: Grinstein G, Mazenko G (eds) Directions in Condensed Matter Physics. World Scientific Press, Singapore

Reversible Cellular Automata

Kenichi Morita
Hiroshima University, Higashi-Hiroshima, Japan

Article Outline

Glossary

Cellular automaton A cellular automaton (CA) is a system consisting of a large (theoretically, infinite) number of finite automata, called cells, which are connected uniformly in a space. Each cell transits its state depending on the states of itself and the cells in its neighborhood. Thus the state transition of a cell is specified by a local function. Applying the local function to all the cells in the space synchronously, the transition of a configuration (i. e., a whole state of the cellular space) is induced. Such a transition function is called a global function. A CA is regarded as a kind of dynamical system that can deal with various kinds of spatio-temporal phenomena.

Cellular automaton with block rules A CA with block rules was proposed by Margolus [18], and it is often called a CA with Margolus-neighborhood. The cellular space is divided into infinitely many blocks of the same size (in the two-dimensional case, e. g., 2×2). A local transition function consisting of "block rules", which is a mapping from a block state to a block state, is applied to all the blocks in parallel. At the next time step, the block division pattern is shifted by some fixed amount (e. g., to the north-east direction by one cell), and the same local function is applied to them. This model of CA is convenient to design a reversible CA. Because, if the local transition function is injective, then the resulting CA is reversible.

Partitioned cellular automaton A partitioned cellular automaton (PCA) is a framework for designing a reversible CA. It is a subclass of a usual CA where each cell is partitioned into several parts, whose number is equal to the neighborhood size. Each part of a cell has its own state set, and can be regarded as an output port to a specified neighboring cell. Depending only on the corresponding parts (not on the entire states) of the neighboring cells, the next state of each cell is determined by a local function. We can see that if the local function is injective, then the resulting PCA is reversible. Hence, a PCA makes it feasible to construct a reversible CA.

Reversible cellular automaton A reversible cellular automaton (RCA) is defined as a one whose global function is injective (i. e., one-to-one). It can be regarded as a kind of a discrete model of reversible physical space. It is in general difficult to construct an RCA with a desired property such as computation-universality. Therefore, the frameworks of a CA with Margolus neighborhood, a partitioned cellular automaton, and others are often used to design RCAs.

Universal cellular automaton A CA is called computationally universal, if it can compute any recursive function by giving an appropriate initial configuration. Equivalently, it is also defined as a CA that can simulate a universal Turing machine. Universality of RCAs can be proved by simulating other systems such as arbitrary (irreversible) CAs, reversible Turing machines, reversible counter machines, and reversible logic elements and circuits, which have already been known to be universal.

Definition of the Subject

Reversible cellular automata (RCAs) are defined as cellular automata (CAs) with an injective global function. Every configuration of an RCA has exactly one previous configuration, and thus RCAs are "backward deterministic" CAs. The notion of reversibility originally comes from physics. It is one of the fundamental microscopic physical laws of Nature. In this sense, an RCA is thought as an abstract model of a physically reversible space as well as a computing model. It is very important to investigate how computation can be carried out efficiently and elegantly in a sys-

tem having reversibility. This is because future computing devices will surely become those of a nanoscale size.

In this article, we mainly discuss on the properties of RCAs from the computational aspects. In spite of the strong constraint of reversibility, RCAs have very rich ability of computing. We can see that even very simple RCAs have universal computing ability. We can also recognize, in some reversible cellular automata, computation is carried out in a very different manner from conventional computing systems, thus they may give new ways and concepts for future computing.

Introduction

Problems related to injectivity and surjectivity of global functions of CAs were first studied by Moore [22] and Myhill [33] in the Garden-of-Eden problem. A Garden-of-Eden configuration is such that it can exist only at time 0. Therefore, if a CA has such a configuration, then its global function is not surjective, and vice versa. They proved the following Garden-of-Eden theorem: A CA has a Garden-of-Eden configuration, if and only if it has an "erasable configuration". After that, many researchers studied on injectivity and surjectivity of global functions more generally [1,19,20,34]. In particular, Richardson [34] showed that if a CA is injective, then it is surjective.

Toffoli [36] first studied reversible (i. e., injective) CAs from the computational viewpoint. He showed that every k-dimensional irreversible CA can be simulated by a $k+1$-dimensional RCA. Hence, a two-dimensional RCA has universal computing ability. Since then, extensive studies on RCAs have been done until now.

After the pioneering work of Bennett [3] on reversible Turing machines, several models of reversible computing were proposed besides RCAs. They are, for example, reversible logic circuits [8,26,37], Billiard Ball Model of computing [8], and reversible counter machines [25]. Most of these models have close relation to physical reversibility. In fact, reversible computing plays an important role when considering inevitable power dissipation in computing [3,4,5,16,38]. It is also one of the basis of quantum computing (see e. g., [9]) because an evolution of a quantum system is a reversible process.

In this article, we discuss how RCAs can have universal computing ability, and how simple they can be. Since reversibility is one of the fundamental microscopic properties of physical systems, it is important to investigate whether we can use such physical mechanisms directly for computation. An RCA is a useful framework to formalize and investigate these problems. Since this article is not an exhaustive survey, many interesting topics related to RCAs, such as complexity of RCA [35], relations to quantum CA (e. g., [41]), etc., are omitted.

An outline of the following sections is as follows. In Sect. "Reversible Cellular Automata", we give basic definitions on RCAs, and design methods for obtaining RCAs. In Sect. "Simulating Irreversible Cellular Automata by Reversible Ones", it is shown how irreversible CAs are simulated by RCAs. In Sect. "1-D Universal Reversible Cellular Automata", two kinds of computation-universal one-dimensional RCAs are shown. In Sect. "2-D Universal Reversible Cellular Automata", several universal two-dimensional RCAs with simple local functions are shown. In Sect. "Future Directions", we discuss future directions and open problems as well as some other problems on RCAs not given in the previous sections.

Reversible Cellular Automata

Formal Definitions

We first give definitions on conventional cellular automata, and then their reversibility.

Definition 1 A *deterministic k-dimensional (k-D) m-neighbor cellular automaton (CA)* is a system defined by

$$A = (\mathbb{Z}^k, Q, (n_1, \ldots, n_m), f, \#) ,$$

where $\mathbb{Z}$ is the set of all integers (hence $\mathbb{Z}^k$ is the set of all k-dimensional points with integer coordinates at which cells are placed), Q is a non-empty finite set of states of each cell, $(n_1, \ldots, n_m)$ is an element of $(\mathbb{Z}^k)^m$ called a neighborhood ($m = 1, 2, \ldots$), $f\colon Q^m \to Q$ is a local function, and $\# \in Q$ is a quiescent state satisfying $f(\#, \ldots, \#) = \#$.

A *k-dimensional configuration* over the set Q is a mapping $\alpha\colon \mathbb{Z}^k \to Q$. Let $\mathrm{Conf}_k(Q)$ denote the set of all k-dimensional configurations over Q, i. e., $\mathrm{Conf}_k(Q) = \{\alpha \mid \alpha\colon \mathbb{Z}^k \to Q\}$. If k is understood, we write it by $\mathrm{Conf}(Q)$. We say that a configuration α is *finite* iff the set $\{x \mid x \in \mathbb{Z}^k \wedge \alpha(x) \neq \#\}$ is finite. Otherwise, it is called *infinite*.

The global function $F\colon \mathrm{Conf}(Q) \to \mathrm{Conf}(Q)$ of A is defined as the one that satisfies the following formula.

$$\forall \alpha \in \mathrm{Conf}(Q),\ x \in \mathbb{Z}^k\colon$$
$$F(\alpha)(x) = f(\alpha(x + n_1), \ldots, \alpha(x + n_m))$$

Definition 2 Let $A = (\mathbb{Z}^k, Q, (n_1, \ldots, n_m), f, \#)$ be a CA. (1) A is called an *injective CA* iff F is one-to-one. (2) A is called an *invertible CA* iff there is a CA $A' =$

$(\mathbb{Z}^k, Q, N', f', \#)$ that satisfies the following condition:

$$\forall \alpha, \beta \in \mathrm{Conf}(Q)\colon F(\alpha) = \beta \quad \text{iff} \quad F'(\beta) = \alpha\ ,$$

where F and F' are the global functions of A and A', respectively.

The following theorem can be derived from the results independently proved by Hedlund [10], and Richardson [34].

Theorem 3 (Hedlund [10] and Richardson [34]) *Let A be a CA. A is injective iff it is invertible.*

By the above theorem, we see the notions of injectivity and invertibility are equivalent. Henceforth, we use the terminology "reversible CA" (RCA) for such a CA, instead of injective CA or invertible CA, because an RCA is regarded as an analog of physically reversible space. (Note that, in some other computing models such as Turing machines, counter machines, and logic circuits, injectivity is trivially equivalent to invertibility, if they are suitably defined. Therefore, for these models, we can directly define reversibility without introducing the notions of injectivity and invertibility.)

How Can We Find RCAs?

The class of RCAs is a special subclass of CAs. Therefore there arises a problem how we can find or construct RCAs with some desired property. It is in general hard to do so if we use the conventional framework of CAs. Because, the following result is shown by Kari [14] for the two-dimensional case (hence it also holds for higher dimensional CAs).

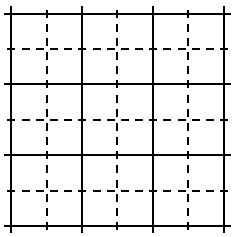
Theorem 4 (Kari [14]) *The problem whether a given two-dimensional CA is reversible is undecidable.*

For the case of one-dimensional CA, Amoroso and Patt [2] showed it is decidable.

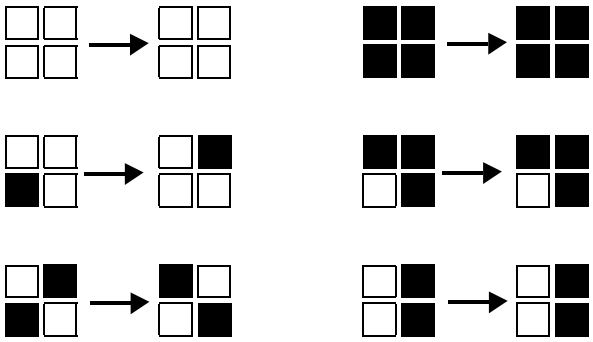
Theorem 5 (Amoroso and Patt [2]) *There is an algorithm to test whether a given one-dimensional CA is reversible or not.*

There are also several studies on enumerating all reversible one-dimensional CAs (e. g., [6,23]). But, it is generally difficult to find RCAs with specific properties such as computation-universality, even for the one-dimensional case.

In order to make it feasible to design an RCA, several methods have been proposed until now. They are, for example, CAs with block rules [18,38], partitioned CAs [28], CAs with second order rules [18,38,39], and others (see e. g., [38]). Here, we describe the first two methods in some detail.



Reversible Cellular Automata, Figure 1
A cellular space with the Margolus neighborhood



Reversible Cellular Automata, Figure 2
Block rules for the Margolus RCA [18]. (Rotation-symmetry is assumed here. Hence, rules obtained by rotating both sides of a rule are also included.)

Cellular Automata with Block Rules Margolus [18] proposed an interesting variant of a CA, by which he composed a computation-universal two-dimensional two-state reversible CA. In his model, all the cells are grouped into "blocks" of size 2×2 as shown in Fig. 1. A specific example of a transformation specified by "block rules" is shown in Fig. 2. This CA evolves as follows: At time 0 the local transformation is applied to every solid line block, then at time 1 to every dotted line block, and so on, alternately. Since this local transformation is one-to-one, the global function of the CA is also one-to-one. Such a neighborhood is called Margolus neighborhood.

One can obtain reversible CAs, by giving one-to-one block rules. However, CAs with Margolus neighborhood are not conventional CAs, because each cell should know the relative position in a block and the parity of time besides its own state.

Related to this topic, Kari [15] showed that every one- and two-dimensional RCA can be represented by a block permutations and translations.

Partitioned Cellular Automata The method of using partitioned cellular automata (PCA) has some similarity to the one that uses block rules. However, resulting re-

versible CAs are in the framework of conventional CA (in other words, a PCA is a special case of a CA). In addition, flexibility of neighborhood is rather high. Shortcomings of PCA is that, in general, the number of states per cell becomes large.

Definition 6 A *deterministic k-dimensional m-neighbor* partitioned cellular automaton *(PCA)* is a system defined by

$$P = (\mathbb{Z}^k, (Q_1, \ldots, Q_m), (n_1, \ldots, n_m), f, (\#_1, \ldots, \#_m)),$$

where $\mathbb{Z}$ is the set of all integers, Q_i $(i = 1, \ldots, m)$ is a non-empty finite set of states of the i-th part of each cell (thus the state set of each cell is $Q = Q_1 \times \cdots \times Q_m$), $(n_1, \ldots, n_m) \in (\mathbb{Z}^k)^m$ is a neighborhood, $f: Q \to Q$ is a local function, and $(\#_1, \ldots, \#_m) \in Q$ is a quiescent state satisfying $f(\#_1, \ldots, \#_m) = (\#_1, \ldots, \#_m)$. (In general, the states $\#_1, \ldots, \#_m$ may be different from each other. However, by renaming the states in each part appropriately, we can identify the states $\#_1, \ldots, \#_m$ as representing the same state #. In what follows, we often assume so, and write the quiescent state by $(\#, \ldots, \#)$.)

The notion of a finite (or infinite) configuration is defined similarly as in CA. Let $p_i: Q \to Q_i$ be the projection function such that $p_i(q_1, \ldots, q_m) = q_i$ for all $(q_1, \ldots, q_m) \in Q$. The global function F: Conf$(Q) \to$ Conf(Q) of P is defined as the one that satisfies the following formula.

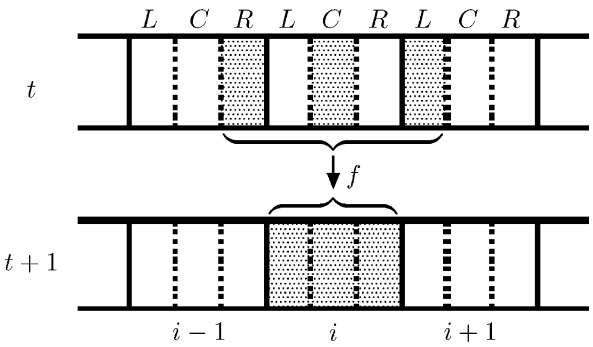
$$\forall \alpha \in \mathrm{Conf}(Q), x \in \mathbb{Z}^k:$$
$$F(\alpha)(x) = f(p_1(\alpha(x + n_1)), \ldots, p_m(\alpha(x + n_m)))$$

By the above definition, a one-dimensional radius 1 (3-neighbor) PCA P_{1d} can be defined as follows.

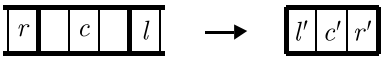
$$P_{1d} = (\mathbb{Z}, (L, C, R), (1, 0, -1), f, (\#, \#, \#))$$

Each cell is divided into three parts, i. e., left, center, and right parts, and their state sets are L, C, and R. The next state of a cell is determined by the present states of the left part of the right-neighbor cell, the center part of this cell, and the right part of the left-neighbor cell (not depending on the whole three parts of the three cells). Figure 3 shows its cellular space, and how the local function f works.

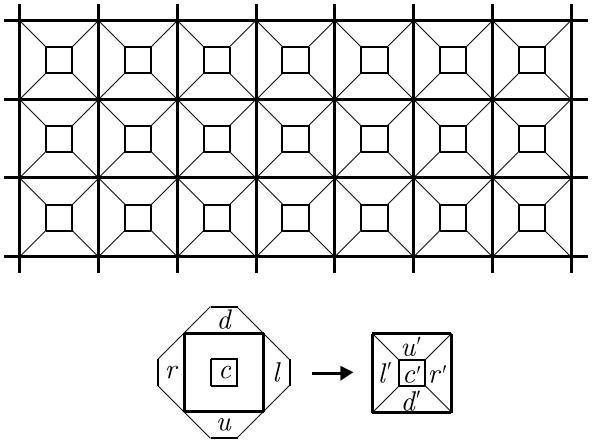
Let $(l, c, r), (l', c', r') \in L \times C \times R$. If $f(l, c, r) = (l', c', r')$, then this equation is called a local rule (or simply a rule) of the PCA P_{1d}, and it is sometimes written in a pictorial form as shown in Fig. 4. Note that, in the pictorial representation, the arguments of the lefthand side of $f(l, c, r) = (l', c', r')$ appear in a reverse order.



Reversible Cellular Automata, Figure 3
Cellular space of a one-dimensional 3-neighbor PCA P_{1d}, and its local function f



Reversible Cellular Automata, Figure 4
A pictorial representation of a local rule $f(l, c, r) = (l', c', r')$ of a one-dimensional 3-neighbor PCA P_{1d}



Reversible Cellular Automata, Figure 5
Cellular space of a two-dimensional 5-neighbor PCA P_{2d}, and its local rule

Similarly, a two-dimensional PCA P_{2d} with Neumann-like neighborhood is defined as follows.

$$P_{2d} = (\mathbb{Z}^2, (C, U, R, D, L), ((0, 0), (0, -1), (-1, 0), (0, 1), (1, 0)), f, (\#, \#, \#, \#, \#))$$

Figure 5 shows the cellular space of P_{2d}, and a pictorial representation of a rule $f(c, u, r, d, l) = (c', u', r', d', l')$.

Let $P = (\mathbb{Z}^k, (Q_1, \ldots, Q_m), (n_1, \ldots, n_m), f, (\#_1, \ldots, \#_m))$ be a k-dimensional PCA, and F be its global function. It is easy to show the following proposition (a proof for

the one-dimensional case given in [28] can be extended to higher dimensions).

Proposition 7 *The local function f is one-to-one, iff the global function F is one-to-one.*

It is also easy to see that the class of PCAs is a subclass of CAs. More precisely, the following proposition is derived by extending the domain of the local function of P.

Proposition 8 *For any k-dimensional m-neighbor PCA P, we can obtain a k-dimensional m-neighbor CA A whose global function is identical with that of P.*

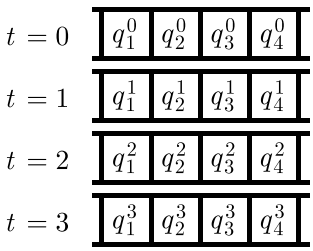
By above, if we want to construct an RCA, it is sufficient to give a PCA whose local function f is one-to-one. This makes a design of an RCA feasible.

Simulating Irreversible Cellular Automata by Reversible Ones

Toffoli [36] first showed that for every irreversible CA there exists a reversible one that simulates the former by increasing the dimension by one. From this result, computation-universality of two-dimensional RCA is derived, since it is easy to embed a Turing machine in a (irreversible) one-dimensional CA.

Theorem 9 (Toffoli [36]) *For any k-dimensional (irreversible) CA A, we can construct a $k+1$-dimensional RCA A' that simulates A in real time.*

Although Toffoli's proof is rather complex, the idea of the proof is easily implemented by using a PCA. Here we explain it informally. Consider a one-dimensional 3-neighbor irreversible CA A that evolves as in Fig. 6. Then, we can construct a two-dimensional reversible PCA P that simulates A as shown in Fig. 7. The configuration of A is kept in some row of P. A state of each cell of A is stored in left, center, and right parts of a cell in P in triplicate. By this, each cell of P can compute the next state of the corresponding cell of A correctly. At the same time, the previous states of the cell and the left and right neighbor cells (which were used to compute the next state) are put downward as a "garbage" signal to keep P reversible. In other words, the additional dimension is used to record all the past history of the evolution of A. In this way, P can simulate A reversibly.



Reversible Cellular Automata, Figure 6
An example of an evolution in an irreversible one-dimensional CA *A*

As for one-dimensional CA with finite configuration, reversible simulation is possible without increasing the dimension.

Theorem 10 (Morita [24]) *For any one-dimensional (irreversible) CA A with finite configurations, we can construct a one-dimensional RCA A' that simulates A (but not in real time).*

1-D Universal Reversible Cellular Automata

Simulating Reversible Turing Machines by 1-D RCAs

It is possible to prove computation-universality of one-dimensional RCAs by constructing RCAs that can simulate reversible Turing machines directly. Here, we first give definitions on reversible Turing machines, and then show how they can be simulated by RCAs.

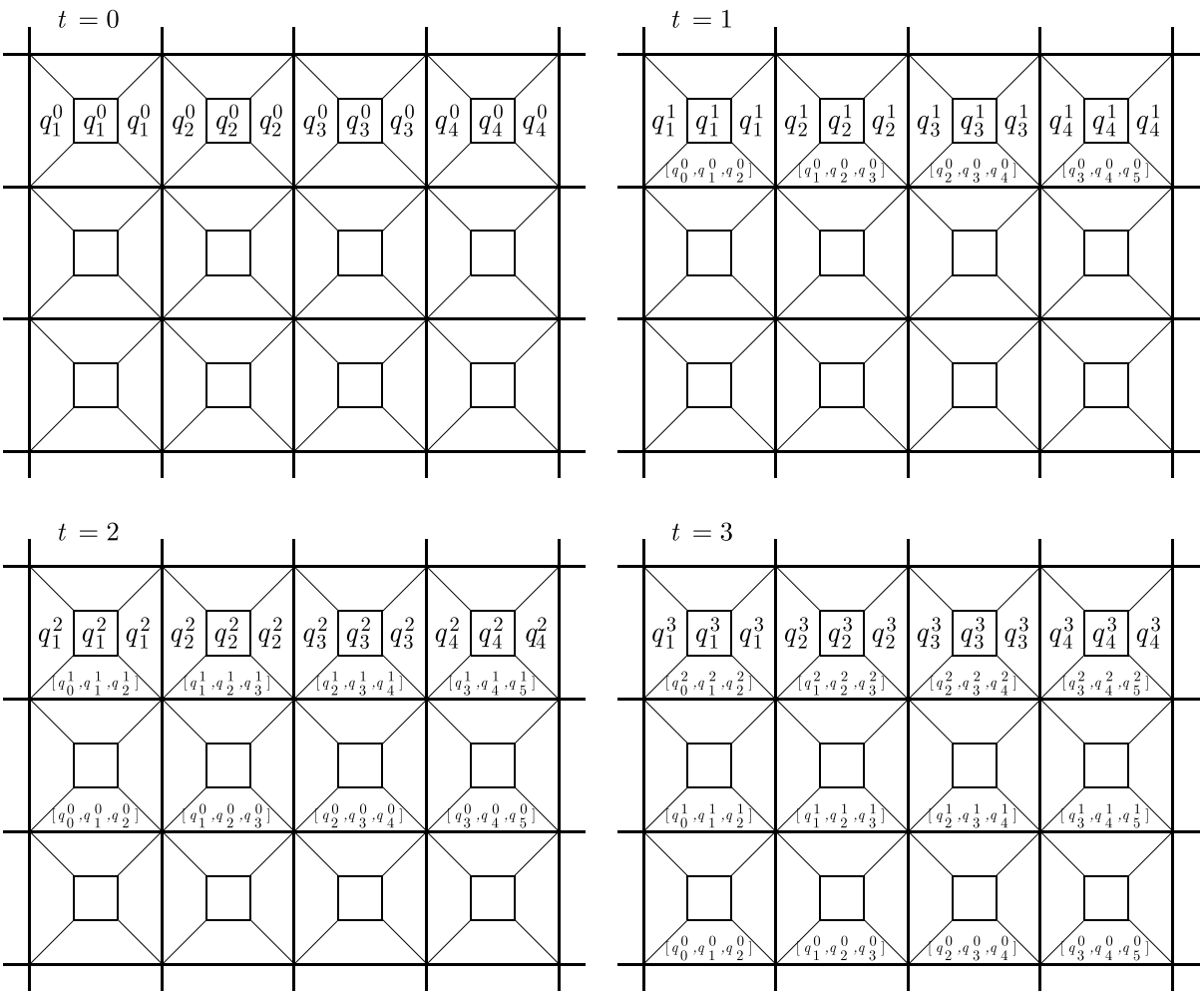
Bennett [3] showed a nice construction method of a reversible Turing machine that simulates a given irreversible Turing machine, and never leaves garbage signals on its tape at the end of computation. We now give a definition of a one-tape Turing machine and its reversible version (a multi-tape reversible Turing machine can be also defined similarly). It is convenient to use quadruple formalism [3] of a Turing machine to define a reversible one.

Definition 11 A one-tape *Turing machine* (TM) is defined by

$$T = (Q, S, q_0, q_a, q_r, s_0, \delta)\,,$$

where Q is a non-empty finite set of states, S is a non-empty finite set of symbols, q_0 is an initial state ($q_0 \in Q$), q_a is an accepting state ($q_a \in Q$), q_r is a rejecting state ($q_r \in Q$), s_0 is a special blank symbol ($s_0 \in S$), and δ is a move relation which is a subset of $(Q \times S \times S \times Q) \cup (Q \times \{/\} \times \{-, 0, +\} \times Q)$. Each element of δ is called a quadruple, and either of the form $[q_i, s, s', q_j] \in (Q \times S \times S \times Q)$ or $[q_i, /, d, q_j] \in (Q \times \{/\} \times \{-, 0, +\} \times Q)$. The symbols "$-$", "0", and "$+$" denote left-shift, zero-shift, and right-shift, respectively. $[q_i, s, s, q_j]$ means that if T reads the symbol s in the state q_i, then write s' and go to the state q_j. $[q_i, /, d, q_j]$ means that if T is in the state q_i, then shift the head to the direction d and go to the state q_j.

T is called deterministic iff the following statement holds for any pair of distinct quadruples $[p_1, b_1, c_1, q_1]$



Reversible Cellular Automata, Figure 7
Simulating the irreversible CA *A* in Fig. 6 by a two-dimensional reversible PCA *P*

and $[p_2, b_2, c_2, q_2]$.

If $p_1 = p_2$, then $b_1 \neq / \wedge b_2 \neq / \wedge b_1 \neq b_2$.

On the other hand, T is called reversible iff the following statement holds for any pair of distinct quadruples $[p_1, b_1, c_1, q_1]$ and $[p_2, b_2, c_2, q_2]$.

If $q_1 = q_2$, then $b_1 \neq / \wedge b_2 \neq / \wedge c_1 \neq c_2$.

Hereafter, we consider only deterministic Turing machines. The next theorem shows computation-universality of a reversible three-tape Turing machine.

Theorem 12 (Bennett [3]) *For any (irreversible) one-tape Turing machine, there is a reversible three-tape Turing machine which simulates the former.*

It is also shown in [31] that for any irreversible one-tape TM, there is a reversible one-tape two-symbol TM which simulates the former. In fact, to prove computation-universality of a one-dimensional reversible PCA, it is convenient to simulate a reversible one-tape TM.

Theorem 13 (Morita and Harao [28]) *For any reversible one-tape Turing machine T, there is a one-dimensional reversible PCA P that simulates the former.*

We show how P simulates T (the method given below is slightly modified from the one in [28]). Let $T = (Q, S, q_0, q_a, q_r, s_0, \delta)$ be a reversible one-tape TM. We assume that q_0 does not appear as a fourth element in any quintuple in δ (because we can always construct such a reversible TM from an irreversible one [31]). A reversible PCA $P = (\mathbb{Z}, (L, C, R), (1, 0, -1), f, (\#, s_0, \#))$ that simu-

lates T is as follows. The state sets L, C, and R are:

$$L = R = Q \cup \{*, \#\}, \quad C = S \cup (Q \times S).$$

The local function f is as below. Note that, in (3)–(6), if $p \in \{q_0\}$ then $x = *$, else $x = \#$. Likewise, if $q \in \{q_a, q_r\}$ then $y = *$, else $y = \#$.

(1) For every $s \in S$,

$$f(\#, s, \#) = (\#, s, \#),$$
$$f(\#, s, *) = (\#, s, *),$$
$$f(*, s, \#) = (*, s, \#).$$

(2) For every $q \in \{q_0, q_a, q_r\}$ and $s \in S$,

$$f(\#, [q, s], \#) = (\#, [q, s], \#).$$

(3) For every $p, q \in Q$ and $s, t \in S$, if $[p, s, t, q] \in \delta$,

$$f(\#, [p, s], x) = (y, [q, t], \#).$$

(4) For every $p, q \in Q$ and $s, t \in S$, if $[p, /, -, q] \in \delta$,

$$f(\#, [p, s], x) = (q, s, \#),$$
$$f(q, t, \#) = (y, [q, t], \#).$$

(5) For every $p, q \in Q$ and $s \in S$, if $[p, /, 0, q] \in \delta$,

$$f(\#, [p, s], x) = (y, [q, s], \#).$$

(6) For every $p, q \in Q$ and $s, t \in S$, if $[p, /, +, q] \in \delta$,

$$f(\#, [p, s], x) = (\#, s, q),$$
$$f(\#, t, q) = (y, [q, t], \#).$$

We can see that the right-hand side of each rule differs from that of any other rule, because T is deterministic and reversible. The rules in (3)–(6) are for simulating T step by step. If the initial computational configuration of T is

$$\cdots s_0 t_1 \cdots q_0 t_i \cdots t_n s_0 \cdots$$

then set P to the following configuration.

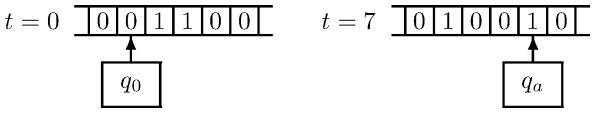
$$\ldots, (\#, s_0, \#), (\#, t_1, \#), \ldots, (\#, t_{i-2}, \#), (\#, t_{i-1}, *),$$
$$(\#, [q_0, t_i], \#), (\#, t_{i+1}, \#), \ldots, (\#, t_n, \#), (\#, s_0, \#), \ldots$$

The simulation process starts when a right-moving signal $*$ meets a center state of the form $[q_0, t_i]$. It is easily seen that, from this configuration, P can correctly simulates T by the rules in (3)–(6). If T becomes a halting state q_a or q_r, then a left-moving signal $*$ is emitted, and the final computational configuration of T is kept in all the successive configurations of P. Note that P itself cannot halt (i. e., it cannot keep exactly the same configuration after reaching it from a different configuration), because P is reversible.

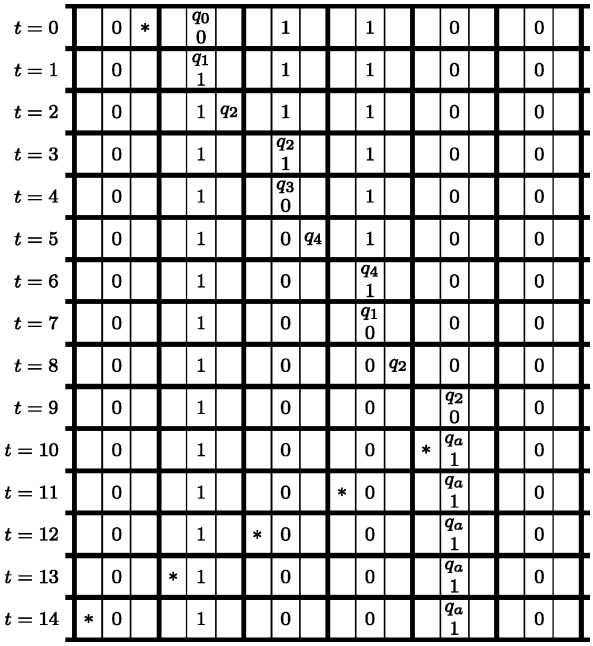
Example 14 Consider a reversible TM $T_{\text{parity}} = (Q, \{0, 1\}, q_0, q_a, q_r, 0, \delta)$, where $Q = \{q_0, q_1, q_2, q_3, q_4, q_a, q_r\}$, and δ is as below.

$$\delta = \{[q_0, 0, 1, q_1], [q_1, /, +, q_2], [q_2, 0, 1, q_a],$$
$$[q_2, 1, 0, q_3], [q_3, /, +, q_4], [q_4, 0, 1, q_r],$$
$$[q_4, 1, 0, q_1]\}.$$

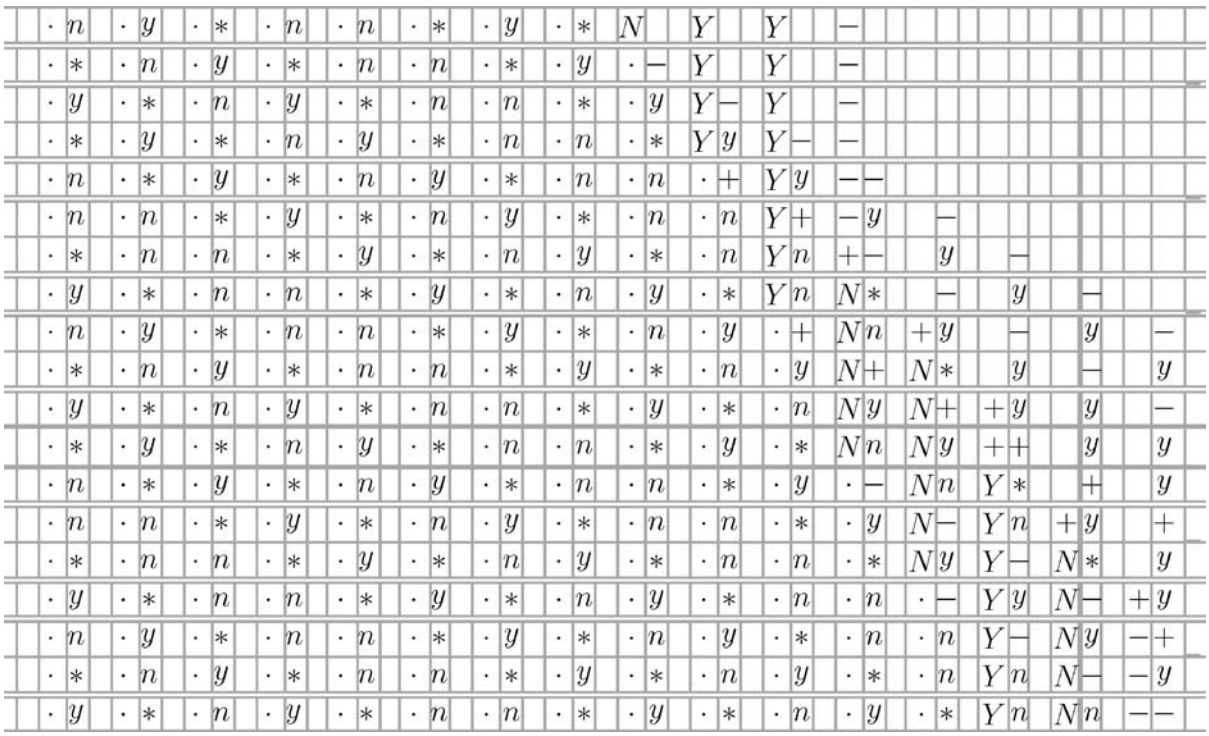
For a given unary number n on the tape, T_{parity} checks if n is even or odd. If it is even, then T_{parity} halts in the accepting state q_a, otherwise halts in the rejecting state q_r. All the symbols read by T_{parity} are complemented. See Fig. 8. A simulation process of T_{parity} by a reversible PCA P_{parity}, which is constructed by the method described above, is shown in Fig. 9.



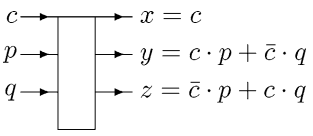
Reversible Cellular Automata, Figure 8
The initial and the final computational configuration of T_{parity} for a given unary input 11



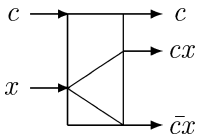
Reversible Cellular Automata, Figure 9
Simulating T_{parity} by a one-dimensional reversible PCA P_{parity}. The state # is indicated by a blank



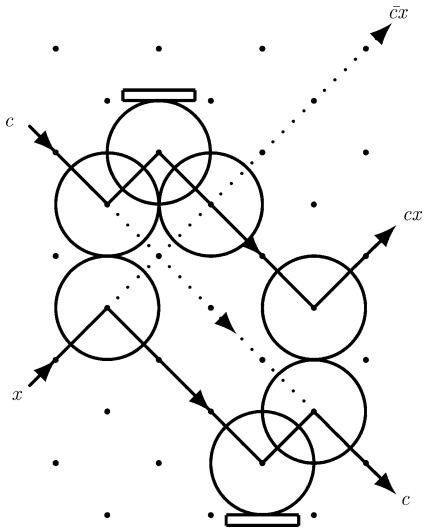
Reversible Cellular Automata, Figure 10
Simulating the CTAG C_0 by the reversible PCA P_{36} [27]



Reversible Cellular Automata, Figure 11
A Fredkin gate [8]



Reversible Cellular Automata, Figure 12
A switch gate [8]

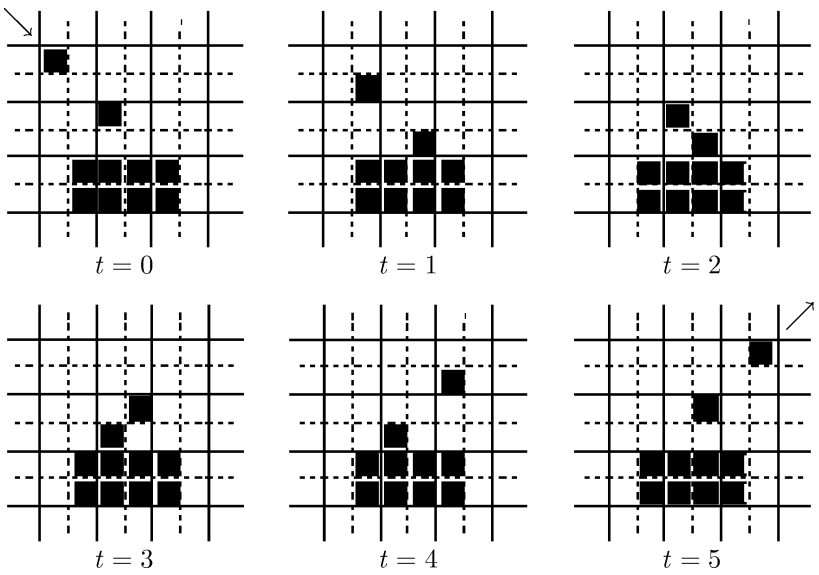


Reversible Cellular Automata, Figure 13
A switch gate realized in the Billiard Ball Model [8]

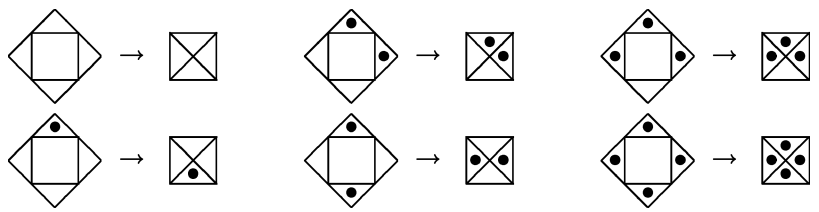
Simulating Cyclic Tag Systems by 1-D RCAs

From Theorem 12 we can see the existence of a universal reversible TM, and thus from Theorem 13 there is a one-dimensional universal RCA. Then the following problem arises: How simple can it be? To get a universal RCA with a small number (say several dozen) of states, we need another useful framework of a universal system.

A cyclic tag system (CTAG) is proposed by Cook [7] to show universality of the elementary cellular automaton of rule 110. As we shall see, it is also useful for constructing simple universal RCAs.



Reversible Cellular Automata, Figure 14
Reflection of a ball by a reflector in the Margolus' RCA [18]



Reversible Cellular Automata, Figure 15
The local function of the 2^4-state rotation-symmetric reversible PCA S_1 [30]

Definition 15 A *cyclic tag system* (CTAG) is defined by $C = (k, \{Y, N\}, (p_0, \ldots, p_{k-1}))$, where k ($k = 1, 2, \ldots$) is the length of a cycle (i. e., period), $\{Y, N\}$ is the (fixed) alphabet, and $(p_0, \ldots, p_{k-1}) \in (\{Y, N\}^*)^k$ is a k-tuple of production rules.

A pair (v, m) is called an *instantaneous description* (ID) of C, where $v \in \{Y, N\}^*$ and $m \in \{0, \ldots, k-1\}$. m is called a *phase* of the ID. A transition relation $\Rightarrow$ on the set of IDs is defined as follows. For any $(v, m), (v', m') \in \{Y, N\}^* \times \{0, \ldots, k-1\}$,

$$
\begin{aligned}
&(Yv, m) \Rightarrow (v', m') \\
&\quad \text{iff } [m' = m + 1 \bmod k] \wedge [v' = vp_m], \\
&(Nv, m) \Rightarrow (v', m') \\
&\quad \text{iff } [m' = m + 1 \bmod k] \wedge [v' = v] \,.
\end{aligned}
$$

A sequence of IDs $(v_0, m_0), (v_1, m_1), \ldots$ is called a *computation starting from* $v \in \{Y, N\}^*$ iff $(v_0, m_0) = (v, 0)$ and $(v_i, m_i) \Rightarrow (v_{i+1}, m_{i+1})(i = 0, 1, \ldots)$. (In what follows, we write a computation by $(v_0, m_0) \Rightarrow (v_1, m_1) \Rightarrow \cdots$.)

A CTAG is a variant of a classical tag system (see e. g., [21]), where production rules are applied cyclically. If the first symbol of a host (i. e., rewritten) string is Y, then it is removed and a specified string at that phase is attached to the end of the host string. If it is N, then it is simply removed and no string is attached.

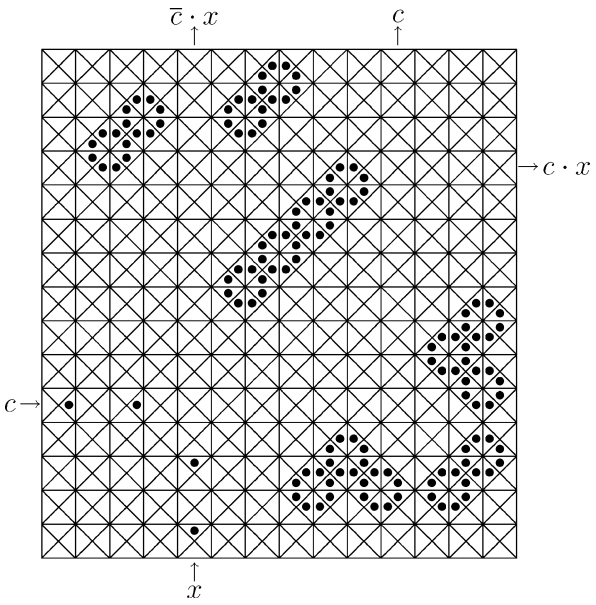
Example 16 Let us consider the following CTAG.

$$C_0 = (3, \{Y, N\}, (Y, NN, YN)) \,.$$

If we give NYY to C_0 as an initial string, then

$$
\begin{aligned}
(NYY, 0) &\Rightarrow (YY, 1) \Rightarrow (YNN, 2) \\
&\Rightarrow (NNYN, 0) \Rightarrow (NYN, 1) \Rightarrow (YN, 2)
\end{aligned}
$$

is an initial segment of a computation starting from NYY.



Reversible Cellular Automata, Figure 16
A switch gate realized in the reversible PCA S_1 [30]

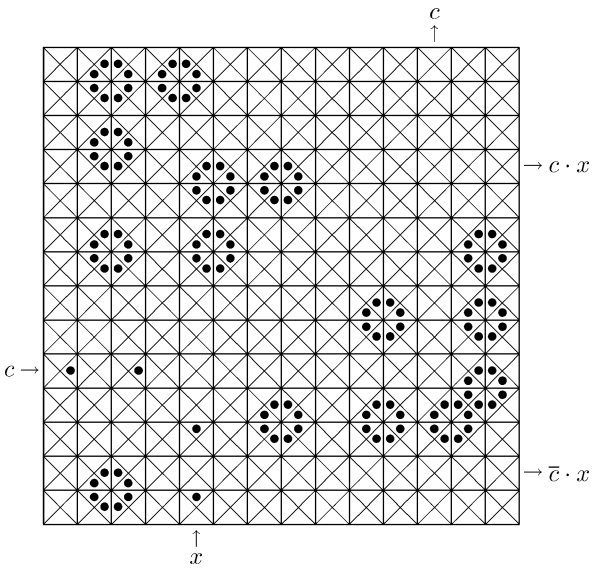
Minsky [21] proved that a 2-tag system, which is a special class of classical tag systems, is universal. The following theorem shows the universality of a CTAG.

Theorem 17 (Cook [7]) *For any 2-tag system, there is a CTAG that simulates the former.*

It was shown by Morita [27] that there are universal one-dimensional RCAs that can simulate any CTAG.

Theorem 18 (Morita [27]) *There is a 36-state one-dimensional reversible PCA P_{36} that can simulate any CTAG on infinite (leftward-periodic) configurations.*

Theorem 19 (Morita [27]) *There is a 98-state one-dimensional reversible PCA P_{98} that can simulate any CTAG on finite configurations. (Note: it can also handle halting of a CTAG.)*



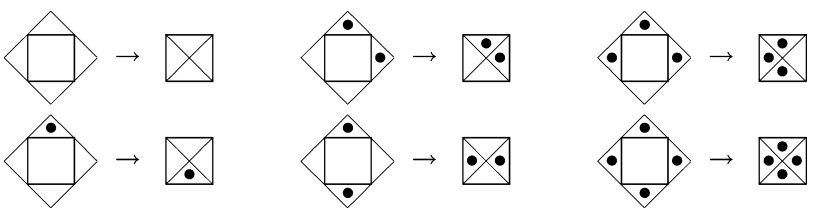
Reversible Cellular Automata, Figure 18
A switch gate realized in the reversible PCA S_1 [30]

The reversible PCA P_{36} in Theorem 18 is given below.

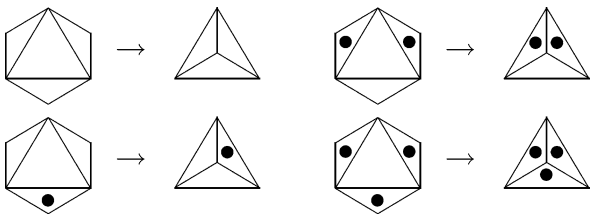
$$P_{36} = (\mathbb{Z}, (\{\#\}, \{\#, Y, N, +, -, \bullet\}, \{\#, y, n, +, -, *\}), (1, 0, -1), f, (\#, \#, \#)) ,$$

where f is defined as follows.

$$\begin{aligned}
f(\#, c, \#) &= (\#, c, \#) && \text{for } c \in \{\#, Y, N, +, -, \bullet\} \\
f(\#, c, r) &= (\#, c, r) && \text{for } c \in \{\#, Y, N\}, \\
& && \text{and } r \in \{y, n, +, -\} \\
f(\#, -, r) &= (\#, -, r) && \text{for } r \in \{y, n\} \\
f(\#, \bullet, r) &= (\#, \bullet, r) && \text{for } r \in \{y, n, *\} \\
f(\#, Y, *) &= (\#, \bullet, +) \\
f(\#, N, *) &= (\#, \bullet, -) \\
f(\#, c, r) &= (\#, r, c) && \text{for } c, r \in \{+, -\} \\
f(\#, +, y) &= (\#, Y, *) \\
f(\#, +, n) &= (\#, N, *) \\
f(\#, \#, *) &= (\#, +, y)
\end{aligned}$$



Reversible Cellular Automata, Figure 17
The local function of the 2^4-state rotation-symmetric reversible PCA S_2 [30]



Reversible Cellular Automata, Figure 19
The local function of the 2^3-state rotation-symmetric reversible triangular PCA T_1 [12]

It is easy to see that f is injective. Note that since the state set of the left part of a cell of P_{36} has only one element #, it is actually a two-neighbor PCA.

Consider the CTAG C_0 in Example 16. The computation in C_0 starting from NYY is simulated in P_{36} as shown in Fig. 10. The initial string NYY is placed in the center parts of some consecutive cells of P_{36}. The production rules (Y, NN, YN) in C_0 are given in the reverse order in the right parts of consecutive cells. The state $*$ is used to delimit the production rules. Since the production rules are applied cyclically, infinite copies of the state sequence "$ny * nn * y*$" should be given to the left half of the cellular space of P_{36}. Figure 10 shows an initial segment $(NYY, 0) \Rightarrow (YY, 1) \Rightarrow (YNN, 2) \Rightarrow (NNYN, 0) \Rightarrow (NYN, 1) \Rightarrow (YN, 2)$ of the computation in C_0.

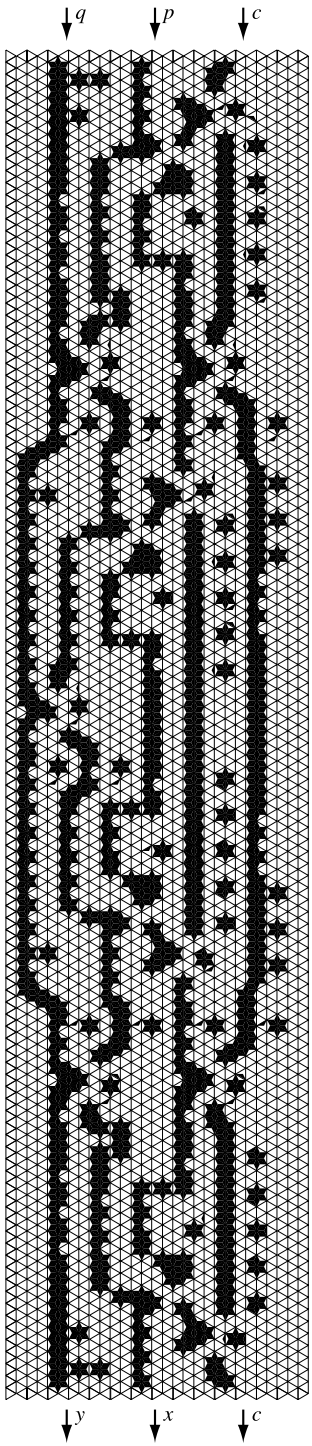
2-D Universal Reversible Cellular Automata

Simulating Reversible Logic Gates by 2-D RCAs

A set of logic elements is called *logically universal*, if any sequential machine (i. e., finite automaton with outputs) can be composed using only elements in the set. Since a finite-state control and tape cells of a Turing machine are in fact sequential machines, we can construct any Turing machine by using these elements. In this section, we show several computation-universal two-dimensional RCAs in which universal reversible logic elements are embedded.

A Fredkin gate [8] is a reversible (i. e., its logical function is one-to-one) and bit-conserving (i. e., the number of 1's is conserved between inputs and outputs) logic gate shown in Fig. 11. It has been known that any combinational logic element (especially, AND, OR, NOT, and fan-out elements) can be realized only with Fredkin gates [8]. Hence, we can construct any sequential machine with Fredkin gate and delay elements.

It is also known that a Fredkin gate can be composed of a much simpler gate called a switch gate and its inverse gate [8]. A switch gate is again a reversible and bit-conserving logic gate (Fig. 12). Furthermore, a switch gate



Reversible Cellular Automata, Figure 20
A Fredkin gate realized in the 2^3-state reversible triangular PCA T_1 [12]

is realized by the Billiard Ball Model (BBM) of computation [8]. The BBM is a kind of physical model of computation where a signal "1" is represented by an ideal ball, and logical operations and routing can be performed by their elastic collisions and reflections by reflectors. Figure 13 shows a BBM realization of a switch gate.

The 2-state RCA with Margolus neighborhood The BBM can be realized by the two-dimensional 2-state RCA proposed by Margolus [18], which has a block rules shown in Fig. 2. Figure 14 shows a reflection of a ball by a mirror in the Margolus' CA. Hence, Margolus' CA is computationally universal.

A 16-State Reversible PCA Model S_1 The model S_1 [30] is a 4-neighbor rotation- and reflection-symmetric reversible PCA model. A cell is divided into four parts, and each part has the state set $\{0, 1\}$. Its local transition rules are shown in Fig. 15. Rotated rules are omitted since it is rotation-symmetric (hence each rule can be regarded as a "rule scheme"). The states 0 and 1 are represented by a blank and a dot, respectively. The set of these rules has some similarity with that of Margolus' CA, and in fact, it can simulate the BBM in a similar manner. In S_1, a ball of the BBM is represented by two dots. Figure 16 gives a configuration of a switch gate.

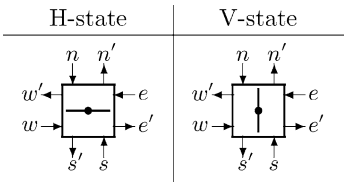
A 16-State Reversible PCA Model S_2 The second model S_2 [30] is also a 4-neighbor reversible PCA having the set of transition rules shown in Fig. 17. It is rotation-symmetric but not reflection-symmetric. In S_2, the shape of a mirror and reflection by it are different from those of S_1, i. e., only a left-turn is possible. Hence a right-turn should be realized by three left-turns. The other features are similar to S_1. Figure 18 shows a configuration of a switch gate.

An 8-State Triangular Reversible PCA Model T_1 The model T_1 [12] is an 8-state reversible PCA on a triangular grid. It is rotation-symmetric but not reflection-symmetric. Its local function is extremely simple as shown in Fig. 19. Signal routing, crossing, and delay are very complex to realize, because a kind of "wall" is necessary to make a signal go straight ahead. Thus the size of the configuration of a Fredkin gate is very large (26×220) as in Fig. 20.

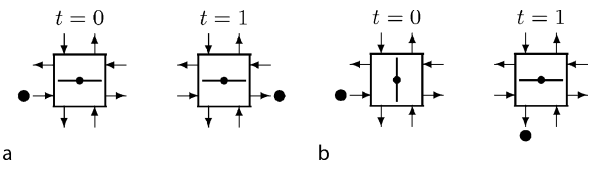
Simulating Reversible Counter Machines by 2-D RCAs

Besides reversible logic gates like a Fredkin gate, there are also universal reversible logic elements with memory. A rotary element (RE) [26] is a typical one of such elements. An RE has four input lines $\{n, e, s, w\}$ and four output lines $\{n', e', s', w'\}$, and two states called H-state and V-state shown in Fig. 21 (hence it has a 1-bit memory). All the values of inputs and outputs are either 0 or 1. Here, the input (and the output) are restricted as follows: at most one "1" appears as an input (output) at a time. The operation of an RE is undefined for the cases that signal 1s are given to two or more input lines.

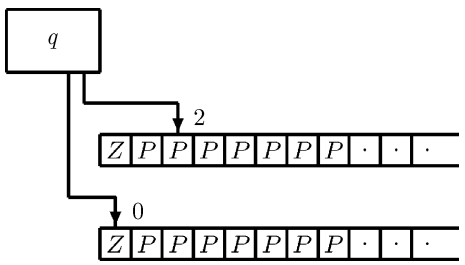
We employ the following intuitive interpretation for the operation of an RE. Signals 1 and 0 are interpreted as existence and non-existence of a particle. An RE has a "rotating bar" to control the moving direction of a particle. When no particle exists, nothing happens on the RE. If a particle comes from a direction parallel to the rotating bar, then it goes out from the output line of the opposite side (i. e., it goes straight ahead) without affecting the direction of the bar (Fig. 22a). If a particle comes from a direction orthogonal to the bar, then it makes a right



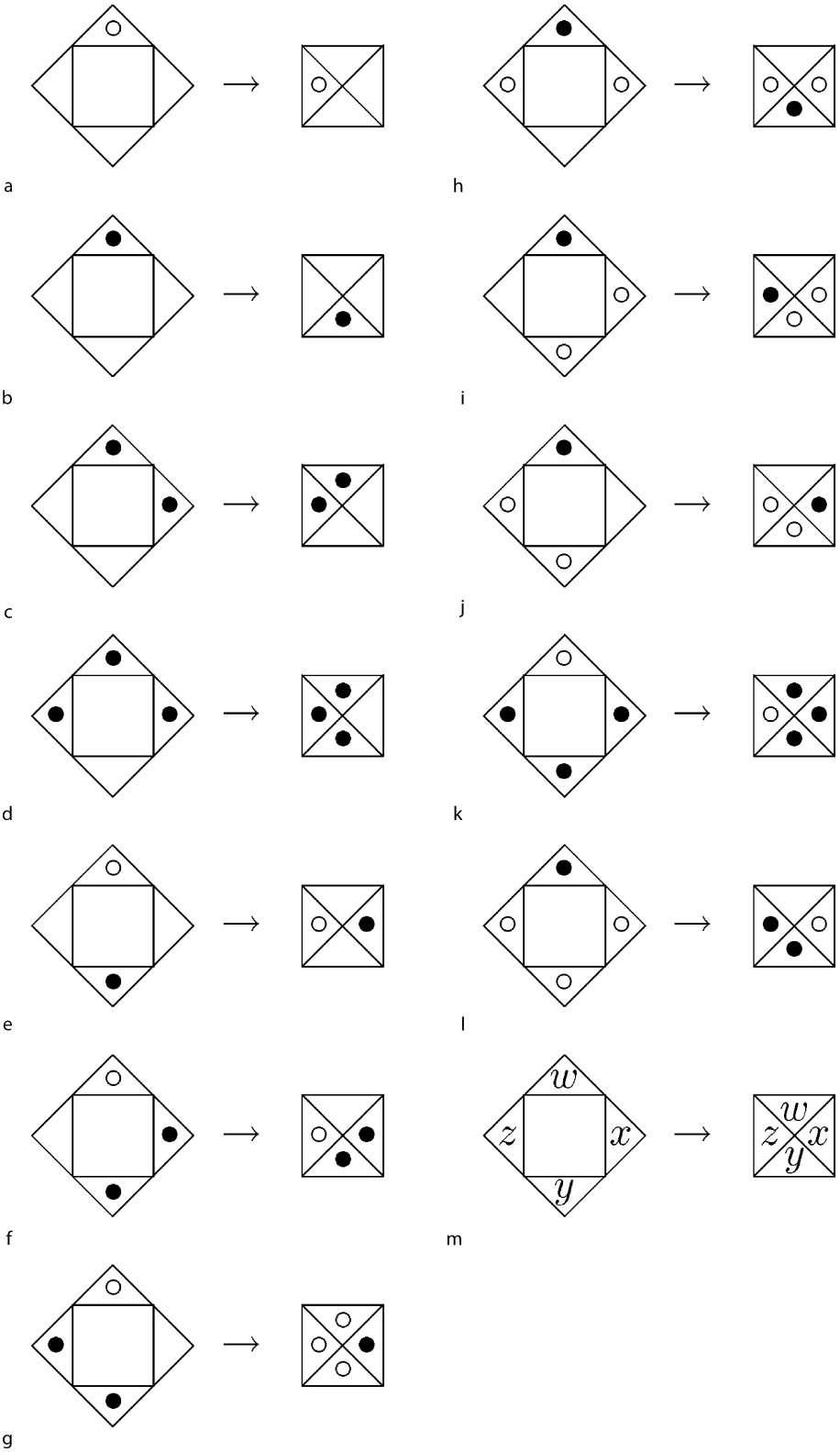
Reversible Cellular Automata, Figure 21
Two states of a rotary element (RE)



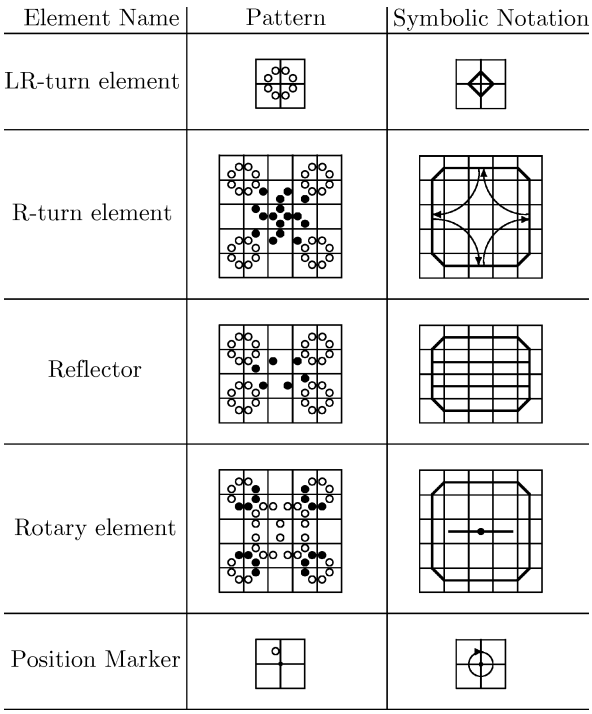
Reversible Cellular Automata, Figure 22
Operations of an RE: a the parallel case, and b the orthogonal case



Reversible Cellular Automata, Figure 23
A counter machine with two counters



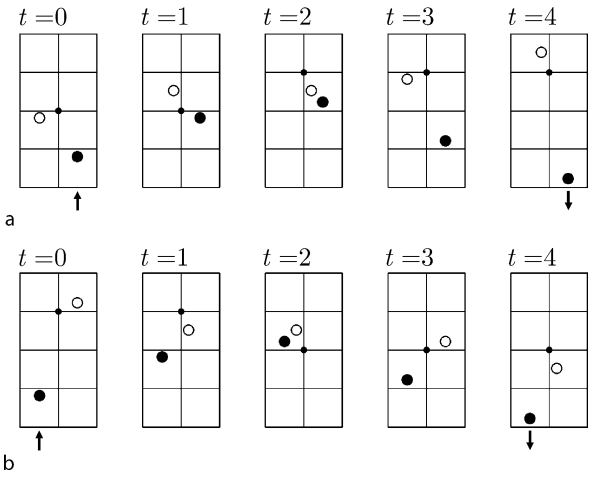
Reversible Cellular Automata, Figure 24
The local function of the 3^4-state rotation-symmetric reversible PCA P_3 [32]. The rule scheme (m) represents 33 rules not specified by (a)–(l), where $w, x, y, z \in \{blank, ○, ●\}$



Reversible Cellular Automata, Figure 25
Basic elements realized in the reversible cellular space of P_3 [32]

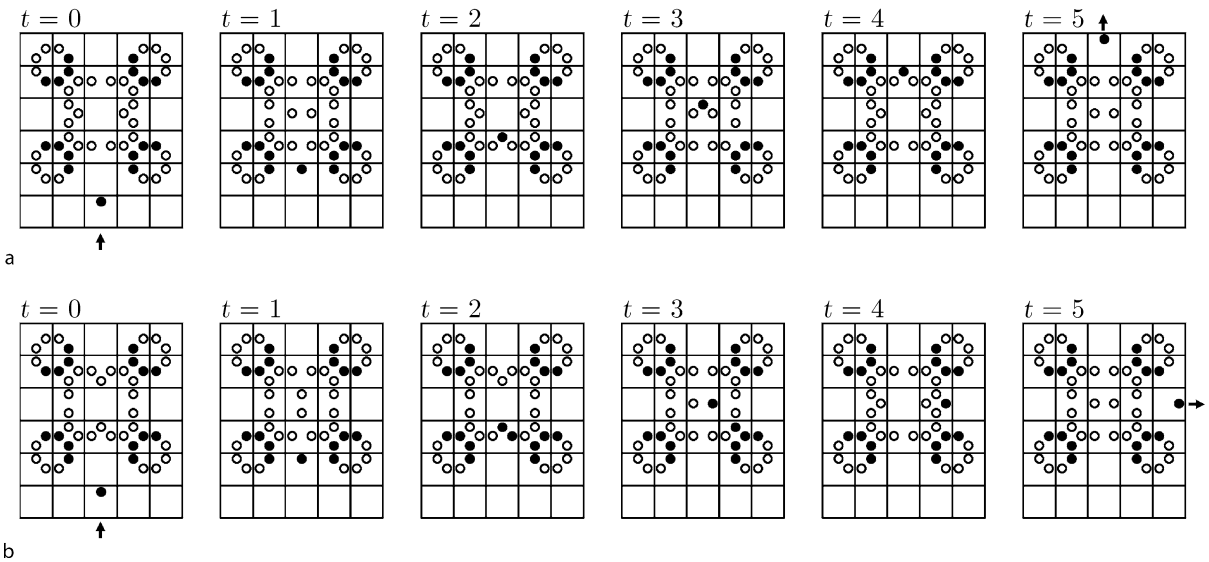
turn, and rotates the bar by 90 degrees counterclockwise (Fig. 22b). It is clear its operation is reversible.

It has been shown that any reversible two-counter machine can be implemented in a quite simple way by using

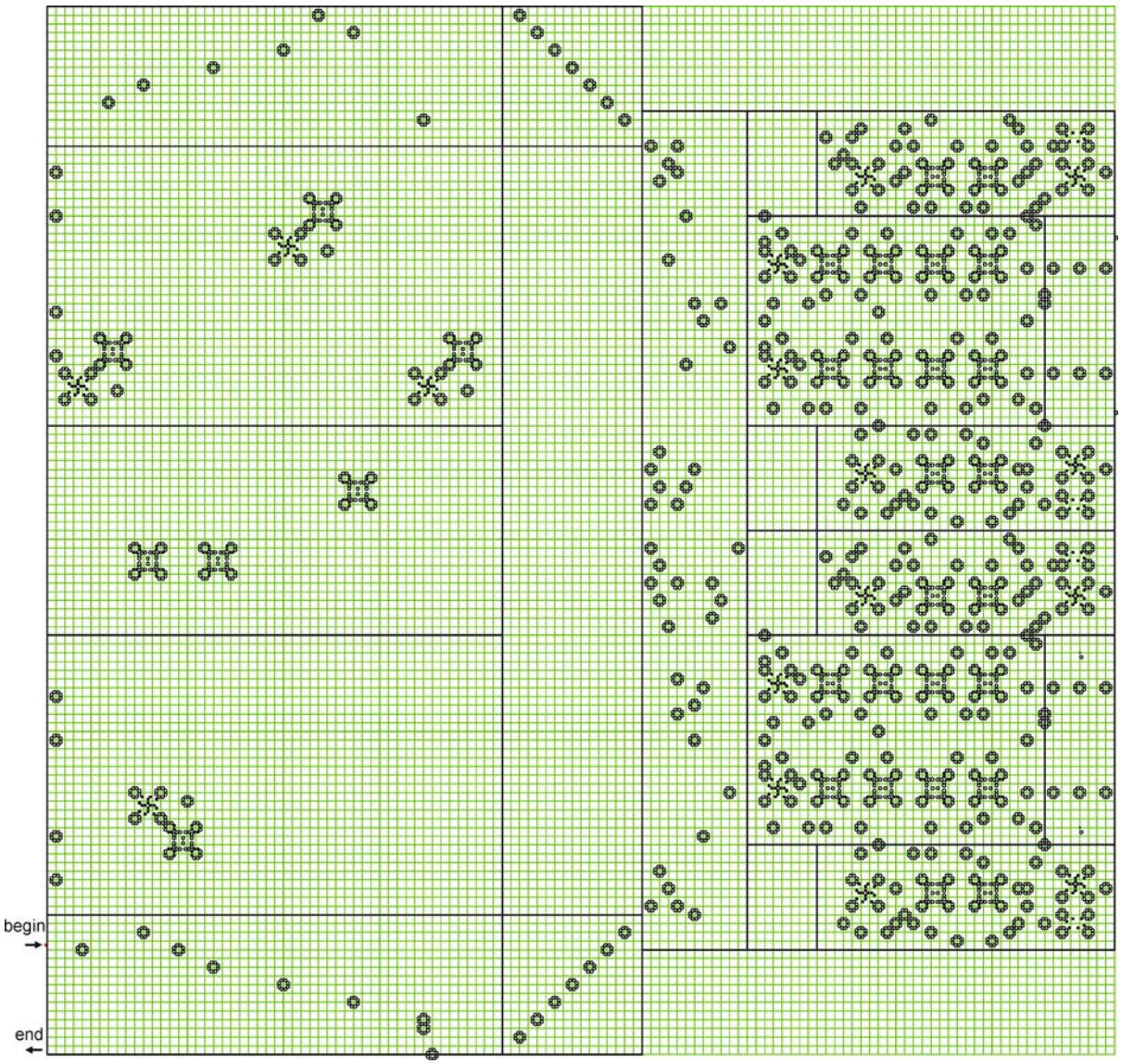


Reversible Cellular Automata, Figure 27
Pushing and pulling operations to a position marker in the reversible PCA P_3

REs and some additional elements [32]. Since a reversible two-counter machine is known to be universal [25], such a reversible PCA is also universal. A counter machine (CM) is a simple computation model consisting of a finite number of counters and a finite-state control [21]. In [25] a CM is defined as a kind of multi-tape Turing machine whose heads are read-only ones and whose tapes are all blank except the leftmost squares as shown in Fig. 23 (P is a blank symbol). This definition is convenient for giving the notion of reversibility on a CM.



Reversible Cellular Automata, Figure 26
Operations of an RE in the reversible PCA P_3: a the parallel case, and b the orthogonal case



Reversible Cellular Automata, Figure 28
An example of a reversible counter machine, which computes the function $2x + 2$, embedded in the reversible PCA P_3 [32]

It is well known that a CM with two counters is computation-universal [21]. This result also holds even if the reversibility constraint is added.

Theorem 20 (Morita, [25]) *For any Turing machine T, there is a deterministic reversible CM with two counters M that simulates T.*

A 3^4-State Reversible PCA Model P_3 Any reversible CM with two counters is embeddable in the model P_3 with the local function shown in Fig. 24 [32]. In P_3, five kinds of signal processing elements shown in Fig. 25 can be realized. Here, a single ● acts as a signal. An LR-turn element, an R-turn element, and a reflector are used for signal routing. Figure 26 shows the operations of an RE in the P_3 space. A position marker is used to keep a head position of a CM, and realized by a single ○, which rotates clockwise at a certain position by the rule (a) in Fig. 24. Figure 27 shows the pushing and pulling operations of a position marker. Figure 28 shows an example of a whole configuration for a reversible CM with two counters embedded in the P_3 space. In this model, no conventional logic elements

like AND, OR and NOT is used. Computation is simply carried out by a single signal that interacts with REs and position markers.

Future Directions

In this section, we discuss future directions and open problems as well as topics not dealt with in the previous sections.

How Simple Can Universal RCAs Be?

We have seen that there are many kinds of simple RCAs having computation-universality. These RCAs with least number of states known so far are summed up as follows.

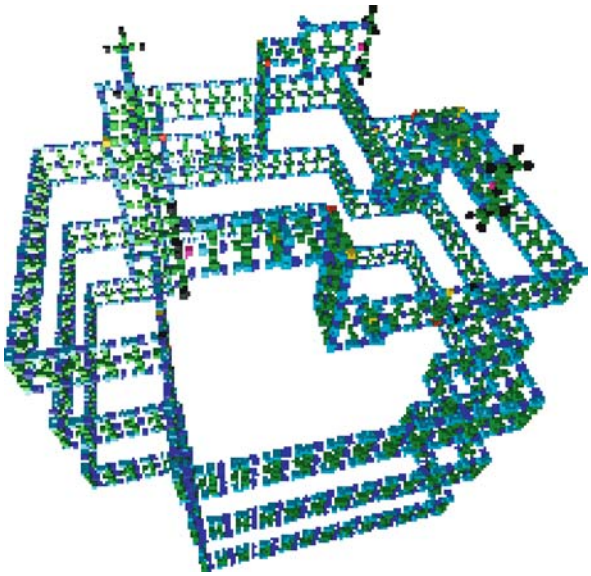
- One-dimensional case:
  Finite configuration: 98-state reversible PCA [27].
  Infinite configuration: 36-state reversible PCA [27].
- Two-dimensional case:
  Finite configuration: 81-state reversible PCA [32].
  Infinite configuration: 2-state RCA with Margolus neighborhood [18].
  8-state reversible PCA on triangular grid [12].
  16-state reversible PCA on rectangular grid [30].

We think the number of states for universal RCA can be reduced much more for each case of the above. Although the framework of PCA is useful for designing an RCA of a standard type, the number of states become relatively large because the state set is the direct product of the sets of the states of the parts. Hence we shall need some other technique to find a universal RCA with a small number of states.

How Can We Realize RCAs in Reversible Physical Systems?

This is a very difficult problem. At present there is no good solution. The Billiard Ball Model [8] is an interesting idea, but it is practically impossible to implement it perfectly. Instead of using a mechanical collision of balls, at least some quantum mechanical reversible phenomena should be used.

Furthermore, if we want to implement a CA in a real physical system, the following problem arises. In a CA, both time and space are discrete, and all the cells operate synchronously. On the other hand, in a real system, time and space are continuous, and no synchronizing clock is assumed beforehand. Hence, we need some novel theoretical framework for dealing with such problems.



Reversible Cellular Automata, Figure 29
Self-reproduction of a pattern in a three-dimensional RCA [13]

Self-Reproduction in RCAs

Von Neumann first invented a self-reproducing cellular automata by using his famous 29-state CA [40]. In his model, the size of a self-reproducing pattern is quite huge, because the pattern has both computing and self-reproducing abilities. Later, Langton [17] created a very simple self-reproducing CA by removing the condition that the pattern need not have computation-universality.

It was shown that self-reproduction of the Langton's type is possible in two- or three-dimensional reversible PCA [13,29]. Figure 29 shows a self-reproducing pattern in a three-dimensional reversible PCA [13]. But, it is left for the future study to design a simple and elegant RCA in which objects with computation-universality can reproduce themselves.

Firing Squad Synchronization in RCAs

It is also possible to solve the firing squad synchronization problem using RCAs. Imai and Morita [11] gave a 99-state reversible PCA that synchronize an array of n cells in $3n$ time steps. Though it seems possible to give an optimal time solution, i. e., a $(2n-2)$-step solution, its concrete design has not yet done.

Bibliography

Primary Literature

1. Amoroso S, Cooper G (1970) The Garden of Eden theorem for finite configurations. Proc Amer Math Soc 26:158–164

2. Amoroso S, Patt YN (1972) Decision procedures for surjectivity and injectivity of parallel maps for tessellation structures. J Comput Syst Sci 6:448–464
3. Bennett CH (1973) Logical reversibility of computation. IBM J Res Dev 17:525–532
4. Bennett CH (1982) The thermodynamics of computation. Int J Theor Phys 21:905–940
5. Bennett CH, Landauer R (1985) The fundamental physical limits of computation. Sci Am 253:38–46
6. Boykett T (2004) Efficient exhaustive listings of reversible one dimensional cellular automata. Theor Comput Sci 325:215–247
7. Cook M (2004) Universality in elementary cellular automata. Complex Syst 15:1–40
8. Fredkin E, Toffoli T (1982) Conservative logic. Int J Theor Phys 21:219–253
9. Gruska J (1999) Quantum Computing. McGraw-Hill, London
10. Hedlund GA (1969) Endomorphisms and automorphisms of the shift dynamical system. Math Syst Theor 3:320–375
11. Imai K, Morita K (1996) Firing squad synchronization problem in reversible cellular automata. Theor Comput Sci 165:475–482
12. Imai K, Morita K (2000) A computation-universal two-dimensional 8-state triangular reversible cellular automaton. Theor Comput Sci 231:181–191
13. Imai K, Hori T, Morita K (2002) Self-reproduction in three-dimensional reversible cellular space. Artifici Life 8:155–174
14. Kari J (1994) Reversibility and surjectivity problems of cellular automata. J Comput Syst Sci 48:149–182
15. Kari J (1996) Representation of reversible cellular automata with block permutations. Math Syst Theor 29:47–61
16. Landauer R (1961) Irreversibility and heat generation in the computing process. IBM J Res Dev 5:183–191
17. Langton CG (1984) Self-reproduction in cellular automata. Physica 10D:135–144
18. Margolus N (1984) Physics-like model of computation. Physica 10D:81–95
19. Maruoka A, Kimura M (1976) Condition for injectivity of global maps for tessellation automata. Inf Control 32:158–162
20. Maruoka A, Kimura M (1979) Injectivity and surjectivity of parallel maps for cellular automata. J Comput Syst Sci 18:47–64
21. Minsky ML (1967) Computation: Finite and Infinite Machines. Prentice-Hall, Englewood Cliffs
22. Moore EF (1962) Machine models of self-reproduction, Proc Symposia in Applied Mathematics. Am Math Soc 14:17–33
23. Mora JCST, Vergara SVC, Martinez GJ, McIntosh HV (2005) Procedures for calculating reversible one-dimensional cellular automata. Physica D 202:134–141
24. Morita K (1995) Reversible simulation of one-dimensional irreversible cellular automata. Theor Comput Sci 148:157–163
25. Morita K (1996) Universality of a reversible two-counter machine. Theor Comput Sci 168:303–320
26. Morita K (2001) A simple reversible logic element and cellular automata for reversible computing. In: Proc 3rd Int Conf on Machines, Computations, and Universality. LNCS, vol 2055. Springer, Berlin, pp 102–113
27. Morita K (2007) Simple universal one-dimensional reversible cellular automata. J Cell Autom 2:159–165
28. Morita K, Harao M (1989) Computation universality of one-dimensional reversible (injective) cellular automata. Trans IEICE Japan E-72:758–762
29. Morita K, Imai K (1996) Self-reproduction in a reversible cellular space. Theor Comput Sci 168:337–366
30. Morita K, Ueno S (1992) Computation-universal models of two-dimensional 16-state reversible cellular automata. IEICE Trans Inf Syst E75-D:141–147
31. Morita K, Shirasaki A, Gono Y (1989) A 1-tape 2-symbol reversible Turing machine. Trans IEICE Japan E-72:223–228
32. Morita K, Tojima Y, Imai K, Ogiro T (2002) Universal computing in reversible and number-conserving two-dimensional cellular spaces. In: Adamatzky A (ed) Collision-based Computing. Springer, London, pp 161–199
33. Myhill J (1963) The converse of Moore's Garden-of-Eden theorem. Proc Am Math Soc 14:658–686
34. Richardson D (1972) Tessellations with local transformations. J Comput Syst Sci 6:373–388
35. Sutner K (2004) The complexity of reversible cellular automata. Theor Comput Sci 325:317–328
36. Toffoli T (1977) Computation and construction universality of reversible cellular automata. J Comput Syst Sci 15:213–231
37. Toffoli T (1980) Reversible computing, Automata, Languages and Programming. In: de Bakker JW, van Leeuwen J (eds) LNCS, vol 85. Springer, Berlin, pp 632–644
38. Toffoli T, Margolus N (1990) Invertible cellular automata: a review. Physica D 45:229–253
39. Toffoli T, Capobianco S, Mentrasti P (2004) How to turn a second-order cellular automaton into lattice gas: a new inversion scheme. Theor Comput Sci 325:329–344
40. von Neumann J (1966) Theory of Self-reproducing Automata. Burks AW (ed) University of Illinois Press, Urbana
41. Watrous J (1995) On one-dimensional quantum cellular automata. In: Proc 36th Symp on Foundation of Computer Science. IEEE, Los Alamitos, pp 528–537

Books and Reviews

Adamatzky A (ed) (2002) Collision-Based Computing. Springer, London
Bennett CH (1988) Notes on the history of reversible computation. IBM J Res Dev 32:16–23
Burks A (1970) Essays on Cellular Automata. University of Illinois Press, Urbana
Kari J (2005) Theory of cellular automata: A survey. Theor Comput Sci 334:3–33
Morita K (2001) Cellular automata and artificial life–computation and life in reversible cellular automata. In: Goles E, Martinez S (eds) Complex Systems. Kluwer, Dordrecht, pp 151–200
Wolfram S (2001) A New Kind of Science. Wolfram Media, Champaign

Reversible Computing

KENICHI MORITA
Hiroshima University, Higashi-Hiroshima, Japan

Article Outline

Glossary

Billiard ball model The Billiard Ball Model (BBM) is a physical model of computation proposed by Fredkin and Toffoli [19]. It consists of idealized balls and reflectors. Balls can collide with other balls or reflectors. It is a reversible dynamical system, since it is assumed that collisions are elastic, and there is no friction. Fredkin and Toffoli showed that a reversible logic gate called Fredkin gate, which is known to be logically universal, can be embedded in BBM. Hence, a universal computer can be realized in the space of BBM.

Fredkin gate The Fredkin gate is a typical reversible logic gate with 3 inputs and 3 outputs whose operation is defined by the one-to-one mapping $(c, p, q) \mapsto (c, cp + \bar{c}q, \bar{c}p + cq)$. It is known that any logic function (even if it is not a one-to-one function) can be realized by using only Fredkin gates by allowing constant inputs and garbage outputs (i. e., useless signals). Furthermore, it is shown by Fredkin and Toffoli [19] that garbage signals can be reversibly erased, and hence any logic function is realized by a garbage-less circuit composed only of Fredkin gates.

Reversible cellular automaton A cellular automaton (CA) consists of a large number of finite automata called cells interconnected uniformly, and each cell changes its state depending on its neighboring cells. A reversible cellular automaton (RCA) is a one whose global function (i. e., a transition function from the configurations to the configurations) is one-to-one. RCAs can be thought as spatio-temporal models of reversible physical systems as well as reversible computing models. (See ► Reversible Cellular Automata)

Reversible logic element A reversible logic element is a primitive from which logic circuits can be composed, and whose operation is defined by a one-to-one mapping. There are two kinds of such elements: those without memory, and with memory. Reversible elements without memory are nothing but reversible logic gates. The Fredkin gate and the Toffoli gate are well-known examples of them, which are universal. Reversible elements with 1-bit memory are also useful when constructing reversible computing systems. The rotary element is a typical one of this type, which is also known to be universal.

Reversible Turing machine A reversible Turing machine (RTM) is a "backward deterministic" Turing machine, and is a standard model of a reversible computing system. Bennett [4] showed that for any (irreversible) Turing machine there is an RTM that simulates the former and leaves no garbage information on the tape when it halts. Hence, Turing machines still have computation-universality even if the constraint of reversibility is added. Furthermore, there is a rather small universal reversible Turing machine.

Rotary element A rotary element (RE) is a reversible logic element with 1-bit memory [36]. Conceptually, it is square-shaped, and has a "rotating bar" inside. The direction of the bar is either horizontal or vertical. It has four input ports and four output ports on the four edges. If a "particle" comes from the direction parallel to the bar, then it goes straight ahead and does not affect the direction of the bar. If it comes from the direction orthogonal to the bar, then it turns rightward and the bar rotates by 90 degrees. It is known that an RE is universal, and reversible Turing machines can be built very concisely by using only REs.

Definition of the Subject

A reversible computing system is defined as a "backward deterministic" system, where every computational configuration (state of the whole system) has exactly one previous configuration. Hence, a backward computation can be performed by its "inverse" computing system. Though its definition is rather simple, it is closely related to physical reversibility. The study of reversible computing originated from an investigation of heat generation (or energy dissipation) in computing systems. It deals with problems how computation can be carried out efficiently and elegantly in reversible computing models, and how such systems can be implemented in reversible physical systems. Since reversibility is one of the fundamental microscopic physical laws of Nature, and future computing systems will surely become in a nano-scale size, it is very important to investigate these problems.

In this article, we discuss several models of reversible computing from the viewpoint of computing theory (hence, we do not discuss problems of physical implementation in detail). We deal with several reversible computing models of different levels: from an element level to a system level. They are the Billiard Ball Model – an idealized reversible physical model, reversible logic elements and circuits, reversible Turing machines, and a few others. We see these models are related each other, i. e., reversible Turing machines are composed of reversible logic

elements, and reversible logic elements are implemented in the Billiard Ball Model. We can also see even very simple reversible systems have very rich computing ability in spite of the constraint of reversibility. Such computing models will give new insights for the future computing systems.

Introduction

The definitions of reversible computing systems are rather simple. They are the systems whose computation process can be traced in the reverse direction in a unique way from the end of the computation to the start of it. In other words, they are "backward deterministic" computing models (of course, a more precise definition should be given for each model of reversible computing). But, if we explain reversible computing systems only in this way, it is hard to understand why they are interesting. Their importance lies in the fact that it is closely related to physical reversibility.

In [24], Landauer argued the relation between logical irreversibility and physical irreversibility. He pointed out that any irreversible logical operation, such as a simple erasure of an information from the memory, or a simple merge of two paths of a program, is associated with physical irreversibility, and hence it necessarily causes heat generation. In particular, if 1 bit of information is erased, at least $kT \ln 2$ of energy will be dissipated, where k is the Boltzmann constant, and T is the absolute temperature. This is called "Landauer's principle". (Note that besides simple erasure and simple merge, there are reversible erasure and reversible merge, which are not irreversible operations, and are useful for constructing garbage-less reversible Turing machine as we shall see in Sect. "Universality of RTMs and Garbage-Less Computation"). After the work of Landauer, various studies related to "thermodynamics of computation" appeared [4,5,6,7,9,22]. There is also an argument on computation from the physical viewpoint by Feynman [16], which contains an idea leading to quantum computing (see e. g., [12,15,20]).

Bennett [4] studied reversible Turing machine from the viewpoint of Landauer's principle. He proved that for any irreversible Turing machine there is a reversible Turing machine that simulates the former and leaves no garbage information on the tape. The notion of garbage-less computation is important, because disposal of a garbage information is actually equivalent to erasure of the information, and hence leads to energy dissipation. The result of Bennett suggests that any computation can be performed efficiently in a physically reversible system.

Since then, various kinds of reversible computing models have been proposed and investigated: reversible logic elements and circuits [19,36,47,48], the Billiard Ball Model [19] – an idealized physical model of computation, reversible cellular automata [46,49], reversible counter machines [34], and so on. It has been known that all these models have *computation-universality*, i. e., any (irreversible) Turing machine can be simulated in these frameworks.

In this article, we discuss reversible computing from the theoretical point of view (though there are also studies on hardware implementation issues). We show how they can have computation-universality, and how different they are from conventional models. We can see that in some cases, computation in reversible systems is carried out in a very different fashion from that of traditional irreversible computing models.

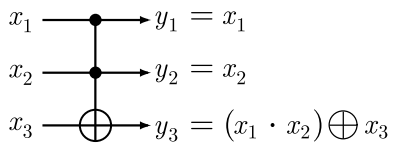
An outline of this article is as follows. In Sect. "Reversible Logic Gates and Circuits", reversible logic gates and circuits are discussed. In particular, the Fredkin gate, its logical universality, garbage-less logic circuits, and its realization in the Billiard Ball Model are explained. In Sect. "Reversible Logic Elements with Memory", reversible logic elements with memory are discussed. A rotary element (RE) is introduced, and its logical universality is shown. In Sect. "Reversible Turing Machines", reversible Turing machines (RTMs) are discussed. A converting method to a garbage-less RTM, a simple realization of an RTM by REs, and a universal RTM are shown. In Sect. "Other Models of Reversible Computing", other reversible models of computing, which are not dealt with in the previous sections, are briefly discussed. In Sect. "Future Directions", future directions and open problems are given.

Reversible Logic Gates and Circuits

A reversible logic gate is a many-input many-output gate that realizes a one-to-one logical (i. e., Boolean) function. Early research on such gates is found in the study of Petri [42]. Later, Toffoli [47,48], and Fredkin and Toffoli [19] studied them from the standpoint of reversible computing, i. e., how they can be realized in a reversible physical system, and how they are related to power dissipation in computation. In this section, we discuss issues of logical universality of reversible logic gates, garbage-less logic circuits, and a relation to the Billiard Ball Model of computation, along the line of the reference [19].

Reversible Logic Gates

Definition 1 An *m-input n-output logic gate with a domain D* is defined as an element that realizes a function $D \to \{0,1\}^n$, where $D \subseteq \{0,1\}^m$. When $D = \{0,1\}^m$, we



Reversible Computing, Figure 1
The generalized AND/NAND gate of order 3 called the Toffoli gate

simply call it an *m-input n-output* logic gate. It is called a *reversible logic gate*, if the function $D \to \{0,1\}^n$ is one-to-one.

It is easy to see that the traditional 1-input 1-output logic gate NOT is reversible, and the 2-input 1-output logic gates AND and OR are irreversible. An example of an m-input n-output reversible logic gate with a domain $D \neq \{0,1\}^m$ will appear in Sect. "Billiard Ball Model (BBM)".

Definition 2 (Toffoli [47,48]) The *generalized AND/NAND gate* of order n is a reversible logic gate that realizes the function

$$\theta^{(n)}: (x_1, \ldots, x_{n-1}, x_n) \mapsto (x_1, \ldots, x_{n-1}, (x_1 \ldots x_{n-1}) \oplus x_n) .$$

We can easily verify that $\theta^{(n)}$ is one-to-one, and in fact $(\theta^{(n)})^{-1} = \theta^{(n)}$. The gate that realizes $\theta^{(2)}$ is sometimes called the *controlled NOT (CNOT)*, which is an important gate in quantum computing (see e. g., [20]). In particular, it is shown that all unitary operations are expressed as compositions of CNOT gates and 1-bit quantum gates [3]. The gate for $\theta^{(3)}$ is called the *Toffoli gate* (Fig. 1).

Definition 3 (Fredkin and Toffoli [19]) The *Fredkin gate* is a 3-input 3-output reversible logic gate that realizes the function

$$\varphi: (c, p, q) \mapsto (c, c \cdot p + \bar{c} \cdot q, \bar{c} \cdot p + c \cdot q) .$$

It is also easy to verify that φ is one-to-one and $\varphi^{-1} = \varphi$. The operations of the Fredkin gate are depicted in Fig. 2.

If $c = 1$ then the inputs p and q are connected to x and y in parallel, while if $c = 0$ then the outputs are exchanged.

In addition to reversibility, the Fredkin gate has bit-conserving property: the total number of 1's in the output lines is the same as that of the input lines (note that the Toffoli gate is not so). This property is an analogue of the conservation law of mass, energy, or momentum in physics. Fredkin and Toffoli [19] proposed a design theory of reversible logic circuits composed of Fredkin gates and unit wires (which have unit-time delay) (Fig. 3). In their theory, a circuit composed of these elements must satisfy the following conditions: every output of a Fredkin gate can be connected to an input of a unit wire (not directly to another Fredkin gate), and every output of a Fredkin gate or a unit wire can be connected to at most one input of another element (i. e., fan-out is not allowed).

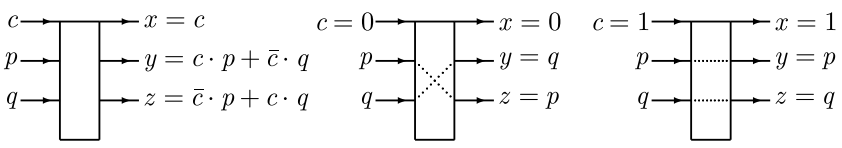
Definition 4 A set E of logic elements is called *logically universal*, if any logic function can be realized by a circuit composed only of the elements in E.

It is well known that, e. g., the set {AND, NOT} is logically universal. If we add a delay element (i. e., a memory) to a logically universal set, we can construct any sequential machine, and thus build a universal computer from them (as an infinite circuit).

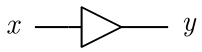
We can see that AND, NOT, and fan-out can be realized by Fredkin gates [19] as shown in Fig. 4, if we allow to supply constant inputs 0's and 1's, and allow garbage (useless) outputs besides the true inputs and outputs. Hence, the set {Fredkin gate} is logically universal. It is also possible to show logical universality of the set {Toffoli gate} (this is left as an exercise for the reader).

Theorem 1 *The sets {Fredkin gate} and {Toffoli gate} are logically universal.*

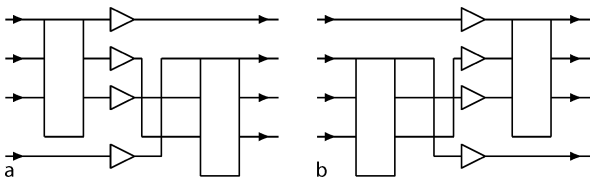
On the other hand, Toffoli [47] showed any 2-input 2-output reversible gate is composed only of CNOTs and NOTs. Since it is easy to see that {CNOT, NOT} is not logically universal, we can conclude no 2-input 2-output reversible gate is logically universal.



Reversible Computing, Figure 2
The Fredkin gate and its operations



Reversible Computing, Figure 3
A unit wire, which acts as a delay element ($y(t) = x(t-1)$)



Reversible Computing, Figure 6
a A reversible logic circuit, and b its inverse circuit

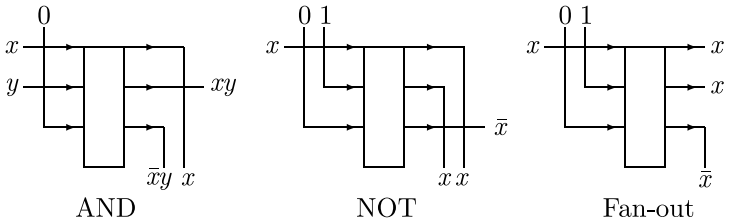
Garbage-Less Logic Circuits Composed of Fredkin Gates

As seen from Theorem 1, any logic function $f\colon \{0,1\}^m \to \{0,1\}^n$ (even if it is not one-to-one) can be realized by a circuit Φ composed of Fredkin gates and unit wires, by allowing constant inputs $\boldsymbol{c}$ and garbage outputs $\boldsymbol{g}$ (Fig. 5). Disposing the garbage $\boldsymbol{g}$ outside of the circuit is, very roughly speaking, analogous to heat generation in an actual circuit. Fredkin and Toffoli [19] showed a method to obtain a garbage-less circuit to compute the function f. This method is a logic circuit version of the garbage-less reversible Turing machine by Bennett [4], which will be discussed in Sect. "Universality of RTMs and Garbage–Less Computation".

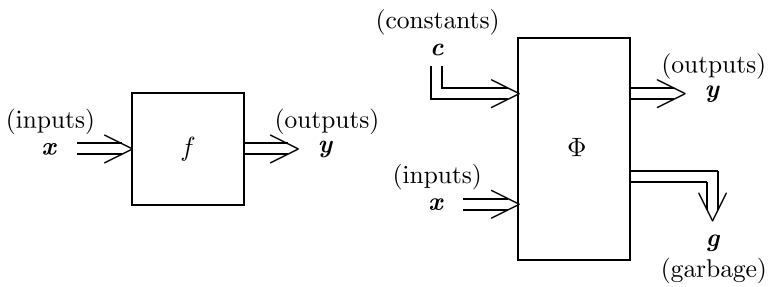
First, the notion of an *inverse circuit* Φ^{-1} of a given reversible logic circuit Φ composed of Fredkin gates and unit wires is introduced. Φ^{-1} is obtained as follows: first take its mirror image, and then exchange inputs and outputs of each element. Figure 6 shows an example. We assume, Φ has no feedback loop (thus it realizes some combinatorial logic function f), and all the delays between inputs and outputs are the same. Then, Φ^{-1} computes the inverse function f^{-1}, because Φ^{-1} "undoes" the computation performed by Φ (note that the inverse of the Fredkin gate is itself). If we connect Φ and Φ^{-1} in series as in Fig. 7, garbage signals are erased reversibly, and we get again constants, which can be used in the next computation of the logic function. However, the true outputs $\boldsymbol{y}$ are also converted to the inputs $\boldsymbol{x}$. By inserting fan-out (i. e., copying) circuits between Φ and Φ^{-1}, we can obtain the outputs $\boldsymbol{y}$ without producing the garbage signals $\boldsymbol{g}$ (Fig. 8).

Theorem 2 (Fredkin and Toffoli, [19]) *For any logic function $f\colon \boldsymbol{x} \mapsto \boldsymbol{y}$, there is a reversible logic circuit composed of Fredkin gates and unit wires that computes the mapping $f'\colon (\boldsymbol{c}, \boldsymbol{x}, \mathbf{0}, \mathbf{1}) \mapsto (\boldsymbol{c}, \boldsymbol{x}, \boldsymbol{y}, \bar{\boldsymbol{y}})$, and produces no garbage signal.*

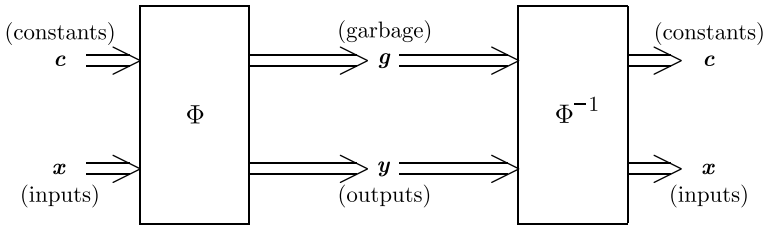
Although the circuit in Fig. 8 does not give garbage signals $\boldsymbol{g}$, it produces $\boldsymbol{x}$ and $\bar{\boldsymbol{y}}$ besides the true output $\boldsymbol{y}$, which may be regarded as a kind of a garbage (though the amount



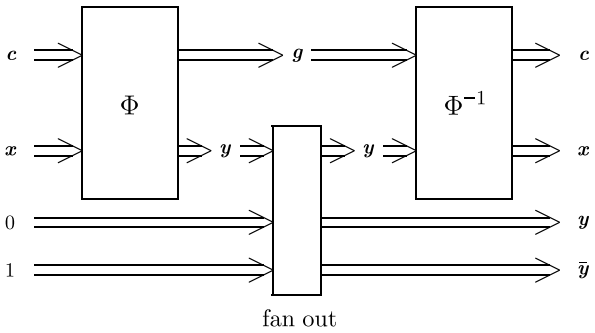
Reversible Computing, Figure 4
Implementing AND, NOT, and fan-out by Fredkin gates



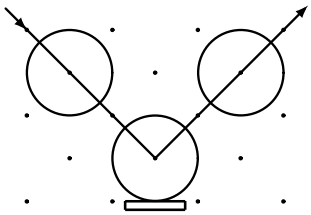
Reversible Computing, Figure 5
Embedding a logic function f in a reversible logic circuit Φ



Reversible Computing, Figure 7
Reversibly erasing garbage signals by the inverse circuit Φ^{-1}



Reversible Computing, Figure 8
A garbage-less reversible logic circuits that computes a logic function



Reversible Computing, Figure 9
The Billiard Ball Model (BBM)

is generally smaller than that of $\boldsymbol{g}$). Especially, when we construct a circuit with memories (i. e., a sequential circuit), this kind of garbage is produced at every time step. In this sense, the above circuit is an "almost" garbage-less one. However, if we consider only "reversible" sequential machines (defined in Sect. "Definitions on Reversible Logic Elements with Memory"), then completely garbage-less circuits can be obtained. This is discussed in Sect. "Garbage-less Construction of RSMs by REs".

Billiard Ball Model (BBM)

Fredkin and Toffoli [19] proposed an interesting reversible physical model of computing called the Billiard Ball Model (BBM). It consists of idealized balls and reflectors (Fig. 9), and balls can collide with other balls or reflectors. We assume collisions are elastic, and there is no friction. Hence, there is no energy dissipation inside of this model.

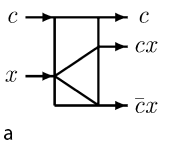
Fredkin and Toffoli showed that any circuit composed of Fredkin gates and unit wires can be embedded in BBM. Therefore, a universal computer can be constructed in the space of BBM. To show the Fredkin gate is realizable in BBM, we need other reversible logic gates called the *switch gate* and its inverse. The switch gate is a 2-input 3-output gate shown in Fig. 10a, which realizes the one-to one logic function $(c, x) \mapsto (c, cx, \bar{c}x)$. The inverse switch gate is a 3-input 2-output gate whose logical function is defined only on the set $\{(0,0,0),(0,0,1),(1,0,0),(1,1,0)\}$ as shown in Fig. 10b, and is the inverse function of the switch gate. Figure 11 shows how the Fredkin gate can be built from switch gates and inverse switch gates (this circuit is due to R. Feynman and A. Ressler [19]). Finally, the switch gate is realized in BBM as in Fig. 12. Therefore, the Fredkin gate is also implemented in BBM.

Margolus showed that BBM can be realized in a reversible cellular automaton [28]. There are also studies related to BBM and cellular automata [14,35].

Reversible Logic Elements with Memory

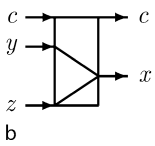
In the classical design theory of logic circuits, generally two sorts of logic elements are supposed. They are logic gates (such as AND, OR, NOT, etc.) and a delay element (i. e., a memory like a flip-flop). Its design technique is also divided into two phases: implement Boolean functions as combinatorial logic circuits consisting of gates, and then construct sequential machines from combinatorial logic circuits and delay elements. In this way, the entire process of constructing digital circuits is systematized, and various methods of circuit minimization can be applied (especially for the first phase).

An approach of Sect. "Reversible Logic Gates and Circuits", which uses reversible logic gates and delay elements, is also a method along this line. However, logic el-



c	x	c	cx	$\bar{c}x$
0	0	0	0	0
0	1	0	0	1
1	0	1	0	0
1	1	1	1	0

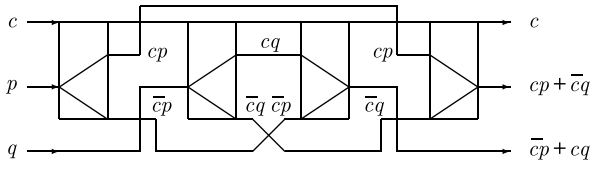
a



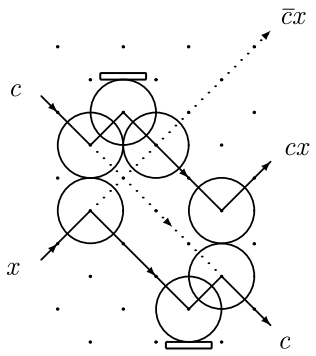
c	y	z	c	x
0	0	0	0	0
0	0	1	0	1
1	0	0	1	0
1	1	0	1	1

b

Reversible Computing, Figure 10
a The switch gate, and b the inverse switch gate



Reversible Computing, Figure 11
The Fredkin gate built from two switch gates and two inverse switch gates



Reversible Computing, Figure 12
A realization of the switch gate in BBM

ements with memory are often more useful in reversible computing. A rotary element (RE) [36] is a typical example of such elements with 1-bit memory. It has 4 input lines and 4 output lines (exactly stated, it has 4 input symbols and 4 output symbols), and its operation is intuitively very simple. In this section, we discuss logical universality of an RE, and also show that any reversible sequential machine can be realized by REs as a completely garbage-less circuit.

In Sect. "Constructing RTMs by REs", we shall also see that by appropriately designing circuits composed only of REs, a clock signal can be eliminated. We should note that a clock is necessary as far as we use "gates", because two or more incoming signals must be synchronized to perform the gate operation. In particular, we shall see any reversible sequential machines and Turing machines can be realized very concisely by such circuits.

Definitions on Reversible Logic Elements with Memory

Since a reversible logic elements with memory is a kind of a reversible sequential machine (RSM), we first give definitions on a sequential machine of Mealy type (i. e., a finite automaton with an output), and its reversibility.

Definition 5 A *sequential machine* (SM) is defined by

$$M = (Q, \Sigma, \Gamma, q_0, \delta) ,$$

where Q is a finite non-empty set of states, Σ and Γ are finite non-empty sets of input and output symbols, respectively, and $q_0 \in Q$ is an initial state. $\delta\colon Q \times \Sigma \to Q \times \Gamma$ is a mapping called a *move function*. A variation of an SM $M = (Q, \Sigma, \Gamma, \delta)$, where no initial state is specified, is also called an SM for convenience.

M is called a *reversible sequential machine* (RSM) if δ is one-to-one (hence $|\Sigma| \le |\Gamma|$). An RSM is "reversible" in the sense that, from the present state and the output of M, the previous state and the input are determined uniquely.

A reversible logic elements with memory (RLEM) is nothing but an RSM with small numbers of states and input/output symbols, from which reversible computers can be built. However, since there are infinitely many RSMs, we should restrict candidates of RLEMs somehow. Here, we consider only RLEMs with 2 states (i. e., $|Q| = 2$) and with k input/output symbols (i. e., $|\Sigma| = |\Gamma| = k$) for $k = 2, 3, 4$.

We consider, e. g., a 2-state 4-symbol RLEM $M = (Q, \Sigma, \Gamma, \delta)$. Here, we fix the state set as $Q = \{q_0, q_1\}$, and the input and output alphabets as $\Sigma = \{a, b, c, d\}$ and $\Gamma = \{w, x, y, z\}$, respectively. The move function δ is as follows.

$$\delta\colon \{q_0, q_1\} \times \{a, b, c, d\} \to \{q_0, q_1\} \times \{w, x, y, z\}$$

Since δ is one-to-one, it is specified by a permutation from the set

$$\{(q_0, w), (q_0, x), (q_0, y), (q_0, z), \\ (q_1, w), (q_1, x), (q_1, y), (q_1, z)\} .$$

Reversible Computing, Table 1
An example of a pair of equivalent 2-state 3-symbol RLEMs

Present state	Input		
	a	b	c
q_0	q_0x	q_1x	q_1y
q_1	q_0y	q_1z	q_0z
(a) RLEM No. 3-61			

Present state	Input		
	a'	b'	c'
q'_0	q'_0y'	q'_1z'	q'_1y'
q'_1	q'_0x'	q'_1x'	q'_0z'
(b) RLEM No. 3-235			

Hence, there are $8! = 40320$ RLEMs for $k = 4$. They are numbered by $0, \ldots, 40\,319$ in the lexicographic order of permutations. 2-state 2-symbol and 3-symbol RLEMs are also numbered in this way [38]. To indicate k-symbol RLEM, the prefix "k-" is attached to the serial number.

Example 1 Table 1 shows two RLEMs No. 3-61 and No. 3-235 specified by the following permutations (in Table 1(b), each symbol has a prime ($'$) for the convenience of the later argument).

$$\text{No. } 3-61: \quad ((q_0,x),(q_1,x),(q_1,y),(q_0,y),(q_1,z),(q_0,z))$$
$$\text{No. } 3-235: \quad ((q_0,y),(q_1,z),(q_1,y),(q_0,x),(q_1,x),(q_0,z))$$

There are many kinds of 2-state RLEMs even if we limit $k = 2, 3, 4$, but we can regard two RLEMs are "equivalent" if one can be obtained by renaming the states and the input/output symbols of the other. This is formalized as follows.

Definition 6 Let $M_1 = (Q_1, \Sigma_1, \Gamma_1, \delta_1)$ and $M_2 = (Q_2, \Sigma_2, \Gamma_2, \delta_2)$ be two LEMs. M_1 and M_2 are called *equivalent* (denoted by $M_1 \sim M_2$), if there exist one-to-one onto mappings $f: Q_1 \to Q_2$, $g: \Sigma_1 \to \Sigma_2$, and $h: \Gamma_1 \to \Gamma_2$ that satisfy

$$\forall\, q \in Q_1,\ \forall\, s \in \Sigma_1\ [\delta_1(q,s) = \psi(\delta_2(f(q), g(s)))]\ ,$$

where $\psi: Q_2 \times \Gamma_2 \to Q_1 \times \Gamma_1$ is defined as follows.

$$\forall\, q \in Q_2,\ \forall\, t \in \Gamma_2\ [\psi(q,t) = (f^{-1}(q), h^{-1}(t))]\ .$$

We can see the two RLEMs No. 3-61 and No. 3-235 in Example 1 are equivalent under the following one-to-one onto mappings.

$$f(q_0) = q'_1, \quad f(q_1) = q'_0$$
$$g(a) = b', \quad g(b) = a', \quad g(c) = c'$$
$$h(x) = x', \quad h(y) = z', \quad h(z) = y'$$

Reversible Computing, Table 2
The move function δ_{RE} of a rotary element RE

Present state	Input			
	n	e	s	w
H-state:	w'	w'	e'	e'
V-state:	s'	n'	n'	s'

The total numbers of 2-state 2-, 3-, and 4-symbol RLEMs are $4! = 24$, $6! = 720$, and $8! = 40\,320$, respectively. But, the numbers of essentially different RLEMs are relatively few. It has been shown that the numbers of equivalence classes of 2-state 2-, 3-, and 4-symbol RLEMs are 8, 24 and 82, respectively [38].

A Rotary Element (RE) and Its Logical Universality

A *rotary element* (RE) [36] is a 2-state 4-symbol RLEM defined as follows.

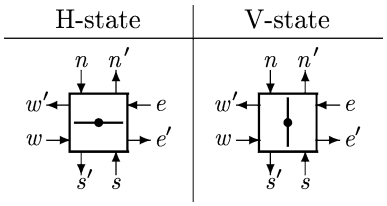
$$\text{RE} = (\{\boxed{-\bullet-}, \boxed{\bullet}\}, \{n, e, s, w\}, \{n', e', s', w'\}, \delta_{\text{RE}})$$

The move function δ_{RE} is shown in Table 2.

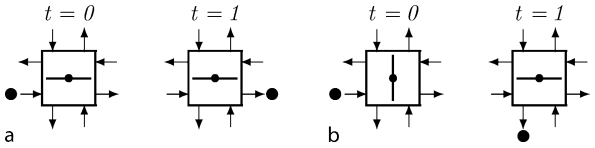
An RE can be understood by the following intuitive interpretation. It has two states called H-state ($\boxed{-\bullet-}$) and V-state ($\boxed{\bullet}$), and four input lines $\{n, e, s, w\}$ and four output lines $\{n', e', s', w'\}$ corresponding to the input and output alphabets (Fig. 13). All the values of input and output lines are either 0 or 1, i. e., $(n, e, s, w), (n', e', s', w') \in \{0, 1\}^4$. However, we restrict the domain of their values as $\{(1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1)\}$, i. e., exactly one "1" appears among the four input (output, respectively) lines at a time, when an input is given (an output is produced). The operation of an RE is left undefined for the cases that signal 1's are given to two or more input lines. Signals 1 and 0 are interpreted as existence and non-existence of a particle. We can interpret that an RE has a "rotating bar" to control the moving direction of a particle. When no particle exists, nothing happens on the RE. If a particle comes from a direction parallel to the rotating bar, then it goes out from the output line of the opposite side (i. e., it goes straight ahead) without affecting the direction of the bar (Fig. 14a). If a particle comes from a direction orthogonal to the bar, then it makes a right turn, and rotates the bar by 90 degrees counterclockwise (Fig. 14b).

Theorem 3 (Morita [36]) *The set {RE} is logically universal.*

This theorem is proved by giving a circuit composed only of REs that simulates a Fredkin gate. Figure 15 shows one such circuit [38]. (A circuit given in [36] uses 16 REs to



Reversible Computing, Figure 13
Two states of a rotary element (RE)



Reversible Computing, Figure 14
Operations of an RE: a the parallel case, and b the orthogonal case

simulate a Fredkin gate.) In Fig. 15, small triangles represent delay elements, where the number written inside of them indicates its delay time. In this circuit, we assume all the elements works synchronously. Note that an RE itself can operate as a delay element.

Garbage-Less Construction of RSMs by REs

We can show any RSM can be realized as a completely garbage-less RE circuit. (We can also construct a circuit of similar property from Fredkin gates [33].)

Theorem 4 (Morita [37]) *For any reversible sequential machine M, there is a garbage-less circuit C that realizes M, and is composed only of REs. C is completely garbage-less in the sense that it has no extra input/output lines other than true input/output lines (hence it is a "closed" circuit except the true input/output).*

Reversible Computing, Table 3
The move function of the RE-column

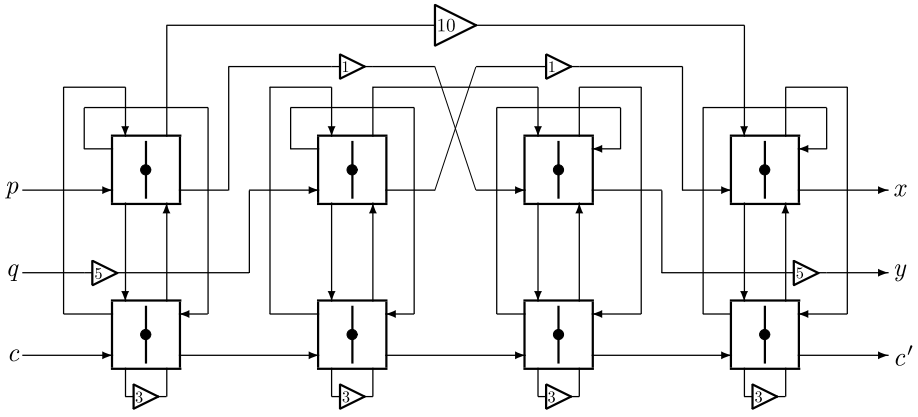
State x	Input l_i	Input r_i
(marked)	l_i'	l_i'
(unmarked)	r_i'	r_i'

Reversible Computing, Table 4
A move function δ_1 of an RSM M_1 (an example)

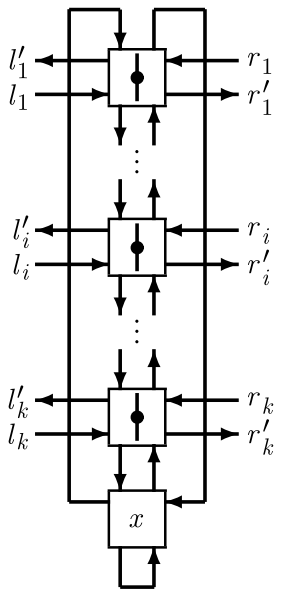
State	Input a_1	Input a_2
q_1	$q_2 b_1$	$q_3 b_2$
q_2	$q_2 b_2$	$q_1 b_1$
q_3	$q_1 b_2$	$q_3 b_1$

We show a construction method of a circuit C by an example. First, we consider a circuit composed of $k+1$ REs shown in Fig. 16. It is called an *RE-column*. In a resting state, each RE in the RE-column is in the vertical state except the bottom one indicated by x. If x is in the horizontal state, the RE-column is called in the *marked* state, otherwise *unmarked* state. It can be regarded as if it is a 2-state RSM with $2k$ input symbols $\{l_1, \ldots, l_k, r_1, \ldots, r_k\}$ and $2k$ output symbols $\{l_1', \ldots, l_k', r_1', \ldots, r_k'\}$, though it has many transient states. Its move function is shown in Table 3.

We now consider an example of an RSM $M_1 = (\{q_1, q_2, q_3\}, \{a_1, a_2\}, \{b_1, b_2\}, \delta_1)$, where its move function δ_1 is shown in Table 4. Prepare three RE-columns with $k = 2$, each of which corresponds to each state of M_1, and connect them as shown in Fig. 17. If M_1 is in



Reversible Computing, Figure 15
A realization of a Fredkin gate out of 8 REs and delay elements



Reversible Computing, Figure 16
An RE-column

the state q_1, then only the first RE-column is set to the marked state. When an input signal comes from, e. g., the input line a_2, then the signal first goes out from the line l'_2 of the first RE-column, setting this column to the unmarked state. Then, this signal enters the third column from the line r_2. This makes the third RE-column marked, and finally goes out from the output line b_2, which realizes $\delta_1(q_1, a_2) = (q_3, b_2)$.

In Fig. 17, the circuit is designed so that the delay between the inputs and outputs is constant. However, if there is no need to do so, delay elements can be removed. In such a case, even if each RE operates asynchronously, the circuit works correctly.

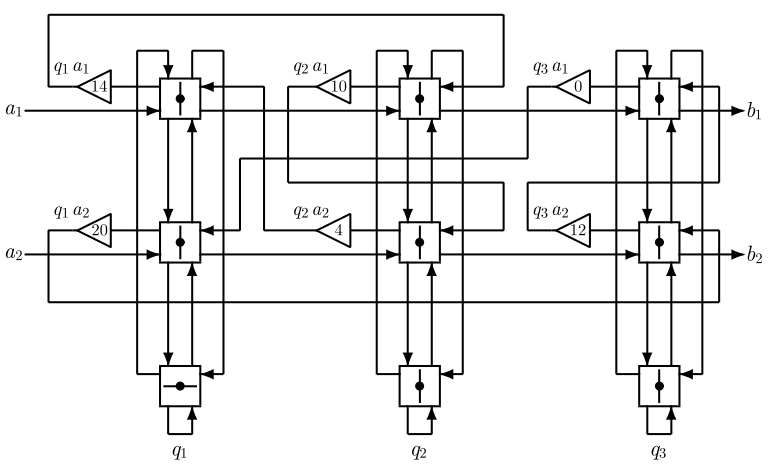
Reversible Turing Machines

A reversible Turing machine (RTM) is a standard model of reversible computing as in the case of traditional (i. e., irreversible) computation theory. The notion of an RTM first appeared in the paper of Lecerf [25], where unsolvability of the halting problem and some related problems on RTMs are shown. Bennett [4] showed that every irreversible Turing machine (TM) can be simulated by an RTM with three tapes without leaving garbage information on the tapes when it halts. He also pointed out the significance of RTMs, since they are closely related to physical reversibility and the problem of energy dissipation in computing process.

In this section, after giving definitions on RTMs, we explain the method of Bennett for converting a given irreversible TM into an equivalent 3-tape RTM. We then discuss the problems how RTMs can be built by reversible logic elements, and how a small universal RTM can be constructed.

Definitions on Reversible Turing Machines (RTMs)

We first introduce a quadruple formulation of RTMs according to Bennett [4]. This formulation makes it easy to define an "inverse" Turing machine for a given RTM. The inverse RTM, which "undoes" the computation performed by the original RTM, plays a key role for garbage-less computation. We also give a quintuple formulation of RTMs



Reversible Computing, Figure 17
A garbage-less circuit made of REs and delay elements that simulates M_1

compatible with the quadruple formulation. This is because most classical TMs are defined in quadruple form, and thus it makes convenient to compare with them.

Definition 7 A *1-tape Turing machine in the quadruple form* is defined by

$$T_4 = (Q, S, q_0, q_f, s_0, \delta) ,$$

where Q is a non-empty finite set of states, S is a non-empty finite set of symbols, q_0 is an initial state ($q_0 \in Q$), q_f is a final (halting) state ($q_f \in Q$), s_0 is a special blank symbol ($s_0 \in S$). δ is a move relation, which is a subset of $(Q \times S \times S \times Q) \cup (Q \times \{/\} \times \{-, 0, +\} \times Q)$. Each element of δ is a quadruple, and either of the form $[p, s, s', q] \in (Q \times S \times S \times Q)$ or $[p, /, d, q] \in (Q \times \{/\} \times \{-, 0, +\} \times Q)$. The symbols "$-$", "0", and "$+$" stand for "left-shift", "zero-shift", and "right-shift", respectively. $[p, s, s', q]$ means if T_4 reads the symbol s in the state p, then write s' and go to the state q. $[p, /, d, q]$ means if T_4 is in p, then shift the head to the direction d and go to the state q.

T_4 is called a *deterministic Turing machine* iff the following condition holds for any pair of distinct quadruples $[p_1, b_1, c_1, q_1]$ and $[p_2, b_2, c_2, q_2]$ in δ.

$$\text{If } p_1 = p_2 , \text{ then } b_1 \neq / \wedge b_2 \neq / \wedge b_1 \neq b_2 .$$

Note that, in what follows, we consider only deterministic Turing machines. So, we omit the word "deterministic" henceforth.

On the other hand, T_4 is called a *reversible Turing machine* (RTM) iff the following condition holds for any pair of distinct quadruples $[p_1, b_1, c_1, q_1]$ and $[p_2, b_2, c_2, q_2]$ in δ.

$$\text{If } q_1 = q_2 , \text{ then } b_1 \neq / \wedge b_2 \neq / \wedge c_1 \neq c_2 .$$

It is easy to extend the above definition to multi-tape RTMs. For example, a 2-tape TM is defined by

$$T = (Q, (S_1, S_2), q_0, q_f, (s_{1,0}, s_{2,0}), \delta) .$$

A quadruple in δ is either of the form $[p, [s_1, s_2], [s'_1, s'_2], q] \in (Q \times (S_1 \times S_2) \times (S_1 \times S_2) \times Q)$ or of the form $[p, /, [d_1, d_2], q] \in (Q \times \{/\} \times \{-, 0, +\}^2 \times Q)$. Determinism and reversibility are defined similarly as above, namely, T is reversible iff the following condition holds for any pair of distinct quadruples $[p_1, X_1, [b_1, b_2], q_1]$ and $[p_2, X_2, [c_1, c_2], q_2]$ in δ.

$$\text{If } q_1 = q_2 , \text{ then } X_1 \neq / \wedge X_2 \neq / \wedge [b_1, b_2] \neq [c_1, c_2] .$$

Definition 8 A *1-tape Turing machine in the quintuple form* is defined by

$$T_5 = (Q, S, q_0, q_f, s_0, \delta) ,$$

where Q, S, q_0, q_f, s_0 are the same as in Definition 7. δ is a move relation, which is a subset of $(Q \times S \times S \times \{-, 0, +\} \times Q)$. Each element of δ is a quintuple of the form $[p, s, s', d, q]$. It means if T_5 reads the symbol s in the state p, then write s', shift the head to the direction d, and go to the state q.

We say T_5 is *deterministic* iff the following condition holds for any pair of distinct quintuples $[p_1, s_1, s'_1, d_1, q_1]$ and $[p_2, s_2, s'_2, d_2, q_2]$ in δ.

$$\text{If } p_1 = p_2 , \text{ then } s_1 \neq s_2 .$$

We say T_5 is *reversible* iff the following condition holds for any pair of distinct quintuples $[p_1, s_1, s'_1, d_1, q_1]$ and $[p_2, s_2, s'_2, d_2, q_2]$ in δ.

$$\text{If } q_1 = q_2 , \text{ then } s'_1 \neq s'_2 \wedge d_1 = d_2 .$$

Proposition 1 *For any RTM T_5 in the quintuple form, there is an RTM T_4 in the quadruple form that simulates each step of the former in two steps.*

Proof Let $T_5 = (Q, S, q_0, q_f, s_0, \delta)$. We define $T_4 = (Q', S, q_0, q_f, s_0, \delta')$ as follows. Let $Q' = Q \cup \{q' \mid q \in Q\}$. The set δ' is given by the next procedure.

First, set the initial value of δ' to the empty set. Next, for each $q \in Q$ do the following operation. Let $[p_1, s_1, s'_1, d_1, q], [p_1, s_2, s'_2, d_2, q], \ldots, [p_m, s_m, s'_m, d_m, q]$ be all the quintuples in δ whose fifth element is q. Note that $d_1 = d_2 = \ldots = d_m$ holds, and $s'_1, s'_2, \ldots, s'_m$ are pair-wise distinct, because T_5 is reversible. Then, include the $m + 1$ quadruples $[p_1, s_1, s'_1, q'], [p_1, s_2, s'_2, q'], \ldots, [p_m, s_m, s'_m, q']$, and $[q', /, d_1, q]$ in δ'.

It is easy to see that T_4 has the required property. □

By Proposition 1, we see the definition of an RTM in the quintuple form is compatible with that in the quadruple form.

The converse of Proposition 1 is also easy to show. It is left for readers as an exercise. (Of course, it is very easy to construct an RTM in the quintuple form so that it simulates each quadruple of a given RTM by a single quintuple. In addition, it is also possible to simulate a consecutive pair of read/write and shift quadruples by one quintuple, and thus we can reduce the numbers of states and quintuples.)

Universality of RTMs and Garbage-Less Computation

We now discuss computation-universality of RTMs. It is actually easy to convert an irreversible TM to an RTM, because we can construct an RTM that simulates the former and records all the information which quadruples were executed by the irreversible TM by using a special history

tape. But, by this method a large amount of garbage information will be left at the end of a computation. An important point is that such a garbage can be reversibly erased, hence a garbage-less computation is possible as shown in the next theorem.

Theorem 5 (Bennett [4]) *For any (irreversible) 1-tape TM, we can construct a 3-tape RTM that simulates the former and leaves only an input string and an output string on the tapes when it halts (hence, it leaves no garbage information).*

Proof Let $T = (Q, S, q_0, q_f, 0, \delta)$ be a given irreversible TM in the quadruple form. We assume T satisfies the following conditions (it is easy to modify a given TM so that it meets the conditions).

(i) T has a rightward-infinite tape whose leftmost square always keeps the blank symbol.
(ii) When T halts, it does so at the leftmost symbol in the state q_f.
(iii) The output is given from the second square of the tape to the right when T halts.
(iv) The output string does not contain blank symbols.
(v) The initial state q_0 does not appear as the fourth element of any quadruple.
(vi) There is only one quadruple in δ whose fourth element is q_f.

Let m be the total number of quadruples contained in δ, and we assume the numbers $1, \ldots, m$ are assigned uniquely to these quadruples.

An RTM T^R that simulates T has three tapes: a working tape (for simulating the tape of T), a history tape (for recording the movement of T at each step), and an output tape (for writing an output string). It is defined as follows:

$$T^R = (Q', (S, \{0, 1, \ldots, m\}, S), q_0, p_0, (0, 0, 0), \delta')$$
$$Q' = \{q_0, q_1, \ldots, q_f\} \cup \{q'_0, q'_1, \ldots, q'_f\} \cup \{c_1, c'_1, c_2, c'_2\} \cup \{p_0, p_1, \ldots, p_f\} \cup \{p'_0, p'_1, \ldots, p'_f\}$$

Computation of T^R has three stages: (1) the forward computation stage, (2) the copying stage, and (3) the backward computation stage. The set δ' of quintuples is defined as the union $\delta_1 \cup \delta_2 \cup \delta_3$ of the sets of quintuples given below. We assume when T^R starts to move, an input string is written in the working tape, and the history and the output tapes contain only blank symbols.

(1) The quadruple set δ_1 for the forward computation stage: When T^R simulates T step by step, it records on the history tape which quadruple of T was used. By this operation T^R can be reversible. This is realized by giving the quadruple set δ_1 as follows.
(i) If the hth quadruple ($h = 1, \ldots, m$) of T is

$$[q_i, s_j, s_k, q_l] \quad (q_i, q_l \in Q, s_j, s_k \in S),$$

then include the following quadruples in δ_1.

$$[q_i, \quad /, \quad [0, +, 0], \quad q'_i]$$
$$[q'_i, \quad [s_j, 0, 0], \quad [s_k, h, 0], \quad q_l]$$

(ii) If the hth quadruple ($h = 1, \ldots, m$) of T is

$$[q_i, /, d, q_l] \quad (q_i, q_l \in Q, d \in \{-, 0, +\})$$

then include the following quadruples in δ_1.

$$[q_i, \quad /, \quad [d, +, 0], \quad q'_i]$$
$$[q'_i, \quad [x, 0, 0], \quad [x, h, 0], \quad q_l] \quad (x \in S)$$

(2) The quadruple set δ_2 for the copying stage: If the forward computation stage is completed, and T^R becomes in the state q_f, then copies the output string on the working tape to the output tape. This is realized by giving the quadruple set δ_2 as below. Here let n be the number of the quadruple that contains q_f as the fourth element.

$$[q_f, \quad [0, n, 0], \quad [0, n, 0], \quad c_1]$$
$$[c_1, \quad /, \quad [+, 0, +], \quad c_2]$$
$$[c_2, \quad [y, n, 0], \quad [y, n, y], \quad c_1] \quad (y \in S - \{0\})$$
$$[c_2, \quad [0, n, 0], \quad [0, n, 0], \quad c'_2]$$
$$[c'_1, \quad [y, n, y], \quad [y, n, y], \quad c'_2] \quad (y \in S - \{0\})$$
$$[c'_2, \quad /, \quad [-, 0, -], \quad c'_1]$$
$$[c'_1, \quad [0, n, 0], \quad [0, n, 0], \quad p_f]$$

(3) The quadruple set δ_3 for the backward computation stage: After the copying process, the backward computation stage starts in order to reversibly erase garbage information left on the history tape. This is performed by an inverse TM of the forward computing. The quadruple set δ_3 is as follows.
(i) If the hth quadruple ($h = 1, \ldots, m$) of T is

$$[q_i, s_j, s_k, q_l] \quad (q_i, q_l \in Q, s_j, s_k \in S)$$

then include the following quadruples in δ_3.

$$[p'_i, \quad /, \quad [0, -, 0], \quad p_i]$$
$$[p_l, \quad [s_k, h, 0], \quad [s_j, 0, 0], \quad p'_i]$$

(ii) If the hth quadruple ($h = 1, \ldots, m$) of T is

$$[q_i, /, d, q_l] \quad (q_i, q_l \in Q, d \in \{-, 0, +\})$$

then include the following quadruples in δ_3. Note that if $d = -$ (0, $+$, respectively), then $d' = +$ (0, $-$).

$$[p'_i, \quad /, \quad [d', -, 0], \quad p_i]$$
$$[p_l, \quad [x, h, 0], \quad [x, 0, 0], \quad p'_i] \quad (x \in S)$$

It is easy to see that T^R simulates T correctly. Reversibility of T^R is verified by checking the quadruple set δ'. In particular, the set δ_3 of quadruples is reversible since T is deterministic. □

Example 2 We consider the following irreversible TM T_{erase}, and convert it into an RTM T^R_{erase} by the method of Theorem 5.

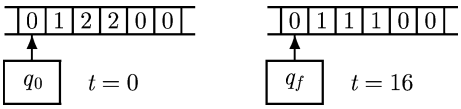
$$T_{\text{erase}} = (\{q_0, q_1, q_2, q_3, q_4, q_f\}, \{0, 1, 2\}, q_0, q_f, 0, \delta_{\text{erase}})$$

$$\begin{aligned}\delta_{\text{erase}} = \{&1 : [q_0, /, +, q_1], \quad 2 : [q_1, 0, 0, q_4],\\ &3 : [q_1, 1, 1, q_2], \quad 4 : [q_1, 2, 1, q_2],\\ &5 : [q_2, /, +, q_1], \quad 6 : [q_3, 0, 0, q_f],\\ &7 : [q_3, 1, 1, q_4], \quad 8 : [q_4, /, -, q_3]\}\,.\end{aligned}$$

When given an input string consisting of 1's and 2's, T_{erase} rewrite all the occurrences of 2's into 1's, and halts (see Fig. 18). The set of quadruples of the RTM T^R_{erase} that simulate T_{erase} is given below, and its computation process for the input 122 is shown in Fig. 19.

$$\begin{array}{llllllll}
[q_0, & /, & [+,+,0], & q'_0] & [p'_0, & /, & [-,-,0], & p_0] \\
[q'_0, & [x,0,0], & [x,1,0], & q_1] & [p_1, & [x,1,0], & [x,0,0], & p'_0] \\
[q_1, & /, & [0,+,0], & q'_1] & [p'_1, & /, & [0,-,0], & p_1] \\
[q'_1, & [0,0,0], & [0,2,0], & q_4] & [p_4, & [0,2,0], & [0,0,0], & p'_1] \\
[q'_1, & [1,0,0], & [1,3,0], & q_2] & [p_2, & [1,3,0], & [1,0,0], & p'_1] \\
[q'_1, & [2,0,0], & [1,4,0], & q_2] & [p_2, & [1,4,0], & [2,0,0], & p'_1] \\
[q_2, & /, & [+,+,0], & q'_2] & [p'_2, & /, & [-,-,0], & p_2] \\
[q'_2, & [x,0,0], & [x,5,0], & q_1] & [p_1, & [x,5,0], & [x,0,0], & p'_2] \\
[q_3, & /, & [0,+,0], & q'_3] & [p'_3, & /, & [0,-,0], & p_3] \\
[q'_3, & [0,0,0], & [0,6,0], & q_f] & [p_f, & [0,6,0], & [0,0,0], & p'_3] \\
[q'_3, & [1,0,0], & [1,7,0], & q_4] & [p_4, & [1,7,0], & [1,0,0], & p'_3] \\
[q_4, & /, & [-,+,0], & q'_4] & [p'_4, & /, & [+,-,0], & p_4] \\
[q'_4, & [x,0,0], & [x,8,0], & q_3] & [p_3, & [x,8,0], & [x,0,0], & p'_4]
\end{array}$$

$$\begin{array}{llllllll}
[q_f, & [0,6,0], & [0,6,0], & c_1] & [c'_1, & [0,6,0], & [0,6,0], & p_f] \\
[c_1, & /, & [+,0,+], & c_2] & [c'_2, & /, & [-,0,-], & c'_1] \\
[c_2, & [y,6,0], & [y,6,y], & c_1] & [c'_1, & [y,6,y], & [y,6,y], & c'_2] \\
 & & [c_2,[0,6, & 0], & [0,6, & 0], c'_2] & &
\end{array}$$



Reversible Computing, Figure 18
The initial and the final computational configuration of T_{erase} for the given input string 122

It is also possible to convert an arbitrary (irreversible) TM into an equivalent 1-tape 2-symbol RTM [31]. There are studies related to computational complexity of RTMs [8,10,21]. Bennett [8] showed the following space-time trade-off theorem for RTMs, where an interesting method of reducing space is given. Note that the RTM in Theorem 5 uses large amount of space, which is proportional to the time that the simulated TM uses.

Theorem 6 (Bennett [8]) *For any $\varepsilon > 0$ and any (irreversible) TM using time T and space S, we can construct an RTM that simulates the former using time $O(T^{1+\varepsilon})$ and space $O(S \log T)$.*

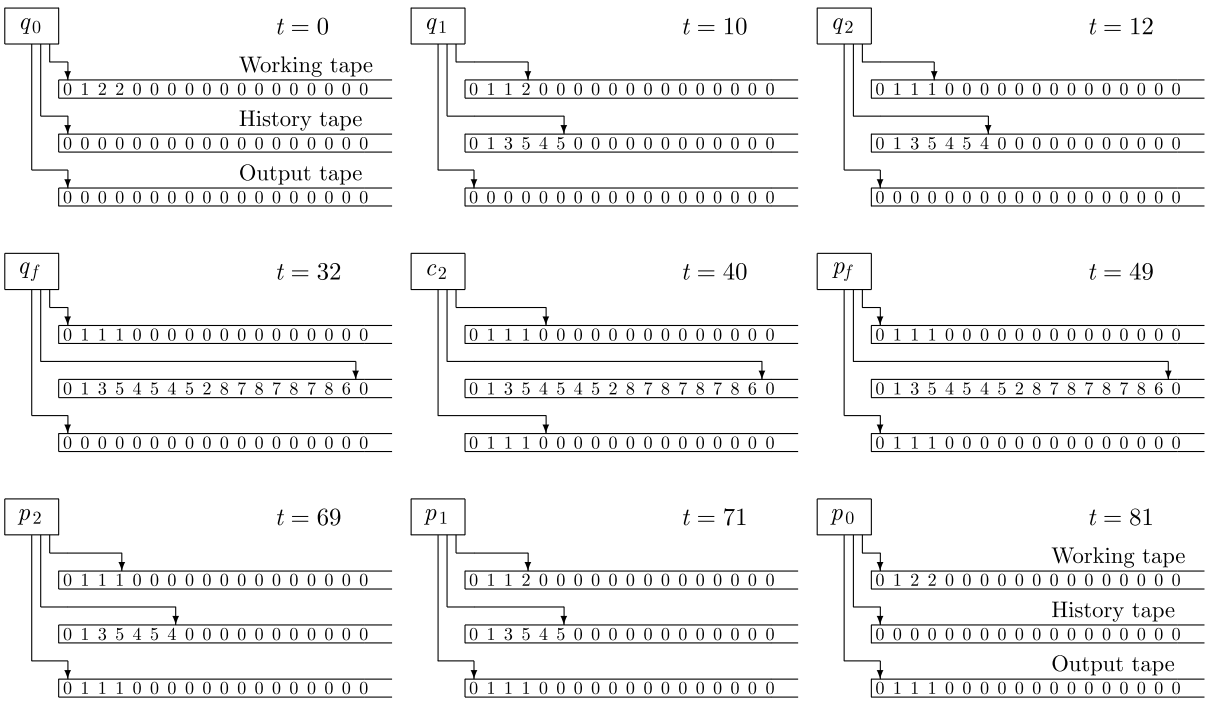
Constructing RTMs by REs

RTMs are related to many other models of reversible computing. For example, for any given RTM, there is a one-dimensional reversible cellular automaton (RCA) that simulates the RTM directly [32]. Hence, a one-dimensional RCA is computation-universal.

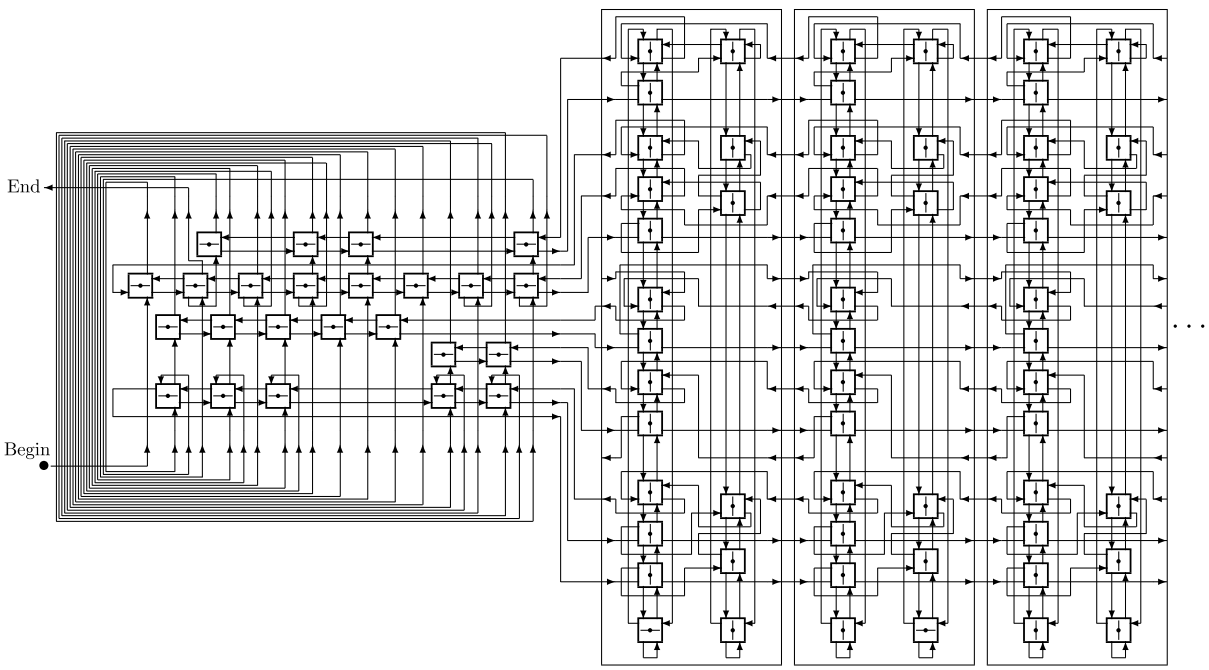
It is also possible to construct a garbage-less reversible logic circuit that simulates a given RTM. We have already seen that any reversible sequential machine can be realized as a garbage-less reversible RE circuit (Theorem 4). In [36], it is shown that an RTM can be decomposed into tape cell modules and a finite state control module, each of which is a reversible sequential machine. Therefore, any RTM can be composed of REs as a garbage-less circuit.

Figure 20 shows an example of an RTM composed only of REs [36]. There is a finite state control in the left half of this figure. To the right of it, infinitely many copies of tape cell modules are connected. (Note that, here, some ad hoc techniques are used to reduce the number of REs rather than to use the systematic method employed in Theorem 4.)

Note that this circuit is closed except the Begin and the End ports. This RTM computes the function $f(n) = 2n$ if n is given as a unary number [36]. Setting the initial states of tape cell modules, and then giving a signal from Begin port, a computing process starts. If the RTM halts, a signal goes out from the End port, leaving an answer in the tape cell modules.



Reversible Computing, Figure 19
The computation process of the RTM T_{erase} for the input string 122: the forward computation stage (from $t = 0$ to $t = 32$), the copying stage (from $t = 33$ to $t = 48$), and the backward computation stage (from $t = 49$ to $t = 81$)



Reversible Computing, Figure 20
A circuit composed only of REs that realizes an RTM [36]

It is also possible to realize RTMs by Fredkin gates and delay elements as a garbage-less circuit using the method of composing RSMs by Fredkin gates shown in [33]. If we do so, however, we need a larger number of Fredkin gates than that of REs. In addition, when designing the circuit, signal delays and timings should be very carefully adjusted, because signals must be synchronized at each gate. On the other hand, the circuit of Fig. 20 works correctly even if each RE operates asynchronously, since only one signal is moving in this circuit. In an RE, a single moving signal interacts with a rotating bar, where the latter is thought as a kind of stationary signal. Hence, its operation may be performed at any time, and there is no need of synchronizing signals at each RE. Of course, if two or more signals are moving in a circuit, some control mechanism for timing is required.

Universal RTMs

From Theorem 5, we can see the existence of a universal RTM (URTM) that can compute any recursive function. However, if we use the technique in the proof of Theorem 5, the numbers of states and symbols of a URTM will become very large.

As for classical (i. e., irreversible) universal Turing machines (UTMs), the following UTMs have been proposed so far, where UTM(m, n) denotes an m-state n-symbol UTM: a UTM(7,4) by Minsky [30], a UTM(24,2), a UTM(10,3), a UTM(7,4), a UTM(5,5), a UTM(4,6), a UTM(3,10) and a UTM(2,18) by Rogozhin [44], a UTM(19,2) by Baiocchi [2], a UTM(3,9) by Kudlek and Rogozhin [23], etc. Most of these small UTMs simulates a 2-tag system [30], which is a simple string rewriting system having computation-universality.

In the following, we describe a 17-state 5-symbol URTM (URTM(17,5)) proposed by Morita and Yamaguchi [39]. It simulates a cyclic tag system (CTAG) given by Cook [11], which is also a very simple string rewriting system with computation-universality. Since the notion of halting was not defined explicitly in the original definition of a CTAG, we use here a modified definition of a CTAG with the halting property, which can simulate a 2-tag system with the halting property [39].

Definition 9 A *cyclic tag system* (CTAG) is defined by

$$C = (k, \{Y, N\}\,,\ \ (\text{halt}, p_1, \ldots, p_{k-1}))\,,$$

where k ($k = 1, 2, \ldots$) is the length of a cycle (i. e., period), $\{Y, N\}$ is the (fixed) alphabet, and $(p_1, \ldots, p_{k-1}) \in (\{Y, N\}^*)^{k-1}$ is a $(k-1)$-tuple of production rules. A pair (v, m) is called an *instantaneous description* (ID) of C, where $v \in \{Y, N\}^*$ and $m \in \{0, \ldots, k-1\}$. m is called a *phase* of the ID. A *halting ID* is an ID of the form $(Yv, 0)(v \in \{Y, N\}^*)$. The transition relation $\Rightarrow$ is defined as follows. For any $(v, m), (v', m') \in \{Y, N\}^* \times \{0, \ldots, k-1\}$,

$$\begin{aligned}(Yv, m) \Rightarrow (v', m') \text{ iff } [m \neq 0] \wedge [m' = m + 1 \mod k] \\ \wedge [v' = vp_m]\,, \\ (Nv, m) \Rightarrow (v', m') \text{ iff } [m' = m + 1 \mod k] \\ \wedge [v' = v]\,.\end{aligned}$$

A sequence of IDs $((v_0, m_0), \ldots, (v_n, m_n))$ is called a *complete computation starting from* $v \in \{Y, N\}^*$ iff $(v_0, m_0) = (v, 0), (v_i, m_i) \Rightarrow (v_{i+1}, m_{i+1})$ $(i = 0, 1, \ldots, n-1)$, and (v_n, m_n) is a halting ID.

Example 3 Consider the CTAG $C_1 = (3, \{Y, N\}, (\text{halt}, YN, YY))$. The complete computation starting from NYY is $(NYY,\ 0) \Rightarrow (YY, 1) \Rightarrow (YYN, 2) \Rightarrow (YNYY, 0)$.

Theorem 7 (Morita and Yamaguchi, [39]) *There is a URTM(17,5).*

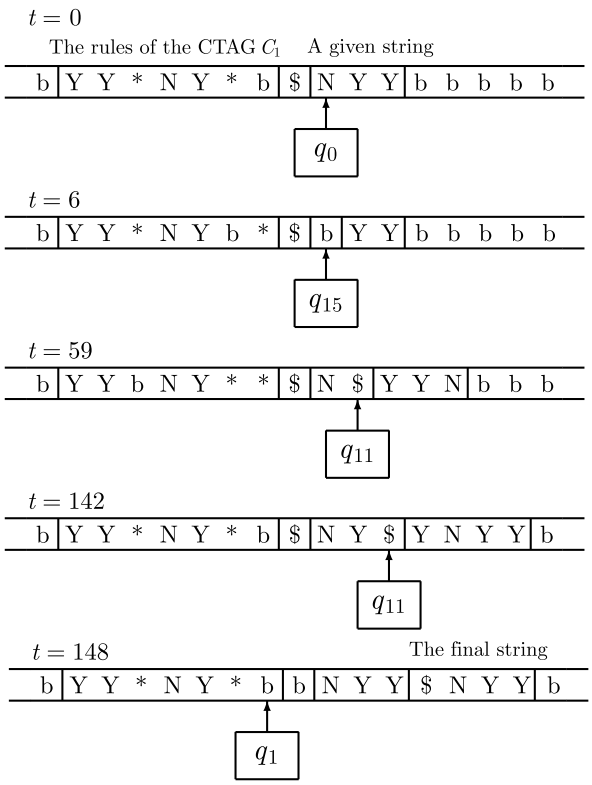
A URTM(17,5) $T_{17,5}$ in the quintuple form that simulates any CTAG is as follows [39].

$$T_{17,5} = (\{q_0, \ldots, q_{16}\}\,,\ \ \{b, Y, N, *, \$\}, q_0, b, \delta)\,,$$

where the set δ of quintuples is shown in Table 5. (Note that, in a construction of a UTM, the final state is usually omitted from the state set.) There are 67 quintuples in total. If a CTAG halts with a halting ID, then $T_{17,5}$ halts in the state q_1. If the string becomes an empty string, then it halts in the state q_2. In Table 5, it is indicated by "null".

Figure 21 shows how the CTAG C_1 with the initial string NYY in Example 3 is simulated by the URTM $T_{17,5}$. On the tape of the URTM, the production rules (halt, NY, YY) of C_1 are expressed by the reversal sequence over $\{Y, N, *\}$, i. e., $YY * YN * *$, where $*$ is used as a delimiter between rules, and "halt" is represented by the empty string. Note that in the initial tape of $T_{17,5}$ $(t = 0)$, the rightmost $*$ is replaced by b. This indicates that the phase is 0. In general, if the phase is $i - 1$, then the ith $*$ from the right is replaced by b. This symbol b is called a "phase marker". On the other hand, the given initial string for C_1 is placed to the right of the rules, where \$ is used as a delimiter.

The IDs in the complete computation $(NYY, 0) \Rightarrow (YY, 1) \Rightarrow (YYN, 2) \Rightarrow (YNYY, 0)$ of C_1 appear in the computational configurations of $T_{17,5}$ at $t = 0, 6, 59$ and 142, respectively. The symbol \$ in the final string $(t = 148)$ should be regarded as Y.



Reversible Computing, Figure 21
Simulating the CTAG C_1 in Example 3 by the URTM $T_{17,5}$

Reversible Computing, Table 5
The set of quintuples of the URTM $T_{17,5}$ [39]

	b	Y	N	$*$	$\$$
q_0	$\$ - q_2$	$\$ - q_1$	$b - q_{13}$		
q_1	halt	$Y - q_1$	$N - q_1$	$* + q_0$	$b - q_1$
q_2	$* - q_3$	$Y - q_2$	$N - q_2$	$* - q_2$	null
q_3	$b + q_{12}$	$b + q_4$	$b + q_7$	$b + q_{10}$	
q_4	$Y + q_5$	$Y + q_4$	$N + q_4$	$* + q_4$	$\$ + q_4$
q_5	$b - q_6$				
q_6	$Y - q_3$	$Y - q_6$	$N - q_6$	$* - q_6$	$\$ - q_6$
q_7	$N + q_8$	$Y + q_7$	$N + q_7$	$* + q_7$	$\$ + q_7$
q_8	$b - q_9$				
q_9	$N - q_3$	$Y - q_9$	$N - q_9$	$* - q_9$	$\$ - q_9$
q_{10}		$Y + q_{10}$	$N + q_{10}$	$* + q_{10}$	$\$ + q_{11}$
q_{11}		$Y + q_{11}$	$N + q_{11}$	$* + q_{11}$	$Y + q_0$
q_{12}		$Y + q_{12}$	$N + q_{12}$	$* + q_{12}$	$\$ - q_3$
q_{13}	$* - q_{14}$	$Y - q_{13}$	$N - q_{13}$	$* - q_{13}$	$\$ - q_{13}$
q_{14}	$b + q_{16}$	$Y - q_{14}$	$N - q_{14}$	$b + q_{15}$	
q_{15}	$N + q_0$	$Y + q_{15}$	$N + q_{15}$	$* + q_{15}$	$\$ + q_{15}$
q_{16}		$Y + q_{16}$	$N + q_{16}$	$* + q_{16}$	$\$ - q_{14}$

It is easy to see that $T_{17,5}$ is reversible by checking the set of quintuples shown in Table 5 according to the definition of an RTM. Intuitively, its reversibility is guaranteed from the fact that no information is erased in the whole simulation process. Furthermore, every branch of the program caused by reading the symbol Y or N is "merged reversibly" by writing the original symbol. For example, the states q_{11} and q_{15} transit to the same state q_0 by writing Y and N, respectively, using the quintuples $[q_{11}, \$, Y, +, q_0]$ and $[q_{15}, b, N, +, q_0]$.

Other Models of Reversible Computing

There are also several models of reversible computing that are not dealt with in the previous sections. Here, we discuss a few of them briefly.

A *reversible cellular automaton* (RCA) is an important model, because it can deal with reversible spatio-temporal phenomena. In fact, it can be thought as an abstract model of a reversible space. So far, a lot of interesting results and properties on RCAs have been shown (see e. g., [35,49]). Some of them are described in ► Reversible Cellular Automata.

A counter machine (CM) is a simple model of a computing consisting of a read-only input tape, a finite number of counters, and a finite state control. It is known that a CM with only two counters has computation-universality [30]. A *reversible counter machine* (RCM) is a backward deterministic version of a CM. An RCM with only two counters is known to be computation-universal [34], though it is a very simple model. This is useful to show universality of other reversible systems.

A *reversible finite automaton* (RFA) is also a backward deterministic version of a finite automaton. Pin [43] studied this model, and characterized its language accepting ability from the formal language theory. Note that an RSM in Sect. "Definitions on Reversible Logic Elements with Memory" is an RFA augmented by an output mechanism.

Future Directions

How Can We Realize Reversible Computers as a Hardware?

Although we have discussed reversible computing mainly from the standpoint of computation theory, it is a very important problem how reversible computers can be realized as a hardware. So far, there have been many interesting attempts from engineering side: e. g., implementing reversible logic as electrically controlled switches [29], c-MOS implementation of reversible logic gates and circuits [13], adiabatic circuits for reversible

computer [17,18], and synthesis of reversible logic circuits [1,45]. However, ultimately, reversible logic elements and computers should be implemented in atomic or molecular level. It is plausible that some nano-scale phenomena can be used as primitives of reversible computing, because the microscopic physical law is reversible. Of course, finding such solution is very difficult, but it is an interesting and challenging problem left for future investigations.

How Simple Can Reversible Computers Be?

To find very simple reversible logic elements with universality is an important problem from both theoretical and practical viewpoints. We may assume, in general, hardware implementation will become easier, if we can find much simpler logical primitives from which reversible computers can be built. As we have discussed in Sect. "Reversible Logic Gates", the Fredkin gate and the Toffoli gate are logically universal gates with minimum number of inputs and outputs. On the other hand, as for reversible logic elements with memory (RLEMs) discussed in Sect. "Reversible Logic Elements with Memory", there are 14 2-state 3-symbol logically universal RLEMS [38,40], which have a less number of symbols than an RE. However, it is an open problem whether there is a single 2-state 2-symbol RLEM that is logically universal (but it is known that there is a set of two 2-state 2-symbol RLEMs that is logically universal [27]). It is also an interesting problem to find much smaller universal reversible Turing machines than $T_{17,5}$ in Sect. "Universal RTMs".

Novel Architectures for Reversible Computers

If we try to construct more realistic computers from reversible logic primitives, we shall need new design theories suited for it. For example, a circuit composed of REs shown in Fig. 20 has a very different feature from traditional logic circuits made of gates. There will surely be many other possibilities of new design techniques that cannot be imagined from the classical theory of logic circuits. To realize efficient reversible computers, development of novel design methods and architectures are necessary.

Asynchronous and Continuous Time Reversible Models

There are still other problems between reversible physical systems and reversible computing systems. The first one is asynchronous systems versus synchronous ones. When building a computing system with a huge number of elements, it is preferable if it is realized as an asynchronous system, because if a clock is eliminated, then power dissipation is reduced, and each element can operate at any time independent to other elements [26,41]. However, it is a difficult problem to define reversibility in asynchronous systems properly. The second problem is continuous time versus discrete time. Natural physical systems are continuous (at least they seem so), while most reversible computing models are defined as discrete systems. To bridge a gap between reversible physical systems and reversible computing models, and to implement the latter in the former systems, we shall need further investigations on these problems.

Bibliography

Primary Literature

1. Al-Rabadi AN (2004) Reversible Logic Synthesis. Springer, Berlin
2. Baiocchi C (2001) Three small universal Turing machines. 3rd Int Conference on Machines, Computations, and Universality. LNCS, vol 2055. Springer, Berlin, pp 1–10
3. Barenco A, Bennett CH, Cleve R, DiVincenzo DP, Margolus N, Shor P, Sleator T, Smolin J, Weinfurter H (1995) Elementary gates for quantum computation. Phys Rev A 52:3457–3467
4. Bennett CH (1973) Logical reversibility of computation. IBM J Res Dev 17:525–532
5. Bennett CH (1982) The thermodynamics of computation. Int J Theoret Phys 21:905–940
6. Bennett CH, Landauer R (1985) The fundamental physical limits of computation. Sci Am 253:38–46
7. Bennett CH (1987) Demons, engines, and the second law. Sci Am 257:108–117
8. Bennett CH (1989) Time/space trade-offs for reversible computation. SIAM J Comput 18:766–776
9. Bennett CH (2003) Notes on Landauer's principle, reversible computation, and Maxwell's Demon. Stud Hist Philos Mod Phys 34:501–510
10. Buhrman H, Tromp J, Vitanyi P (2001) Time and space bounds for reversible simulation. J Phys A 34:6821–6830
11. Cook M (2004) Universality in elementary cellular automata. Complex Syst 15:1–40
12. Deutsch D (1985) Quantum theory, the Church–Turing principle and the universal quantum computer. Proc R Soc Lond A 400:97–117
13. De Vos A, Desoete B, Adamski A, Pietrzak P, Sibinski M, Widerski T (2000) Design of reversible logic circuits by means of control gates. PATMOS 2000. LNCS vol 1918. Springer, Berlin, pp 255–264
14. Durand-Lose J (2002) Computing inside the billiard ball model. In: Adamatzky A (ed) Collision-based Computing. Springer, Londeon, pp 135–160
15. Feynman RP (1982) Simulating physics with computers. Int J Theoret Phys 21:467–488
16. Feynman RP (1996) Hey AJG, Allen RW (eds) Feynman lectures on computation. Perseus Books, Massachusetts
17. Frank M, Vieri C, Ammer MJ, Love N, Margolus NH, Knight TE (1998) A scalable reversible computer in silicon. In: Calude CS,

Casti J, Dinneen MJ (eds) Unconventional Models of Computation. Springer, Singapore, pp 183–200
18. Frank M (1999) Reversibility for Efficient Computing. Ph D Dissertation, MIT, Cambridge
19. Fredkin E, Toffoli T (1982) Conservative logic. Int J Theoret Phys 21:219–253
20. Gruska J (1999) Quantum Computing. McGraw-Hill, London
21. Jacopini G, Mentrasti P, Sontacchi G (1990) Reversible Turing machines and polynomial time reversibly computable functions. SIAM J Disc Math 3:241–254
22. Keys RW, Landauer R (1970) Minimal energy dissipation in logic. IBM J Res Dev 14:152–157
23. Kudlek M, Rogozhin Y (2002) A universal Turing machine with 3 states and 9 symbols. In: Kuich w, Rozenberg G, Salomaa A (eds) Developments in Language Theory (DLT 2001). LNCS vol 2295. Springer, Berlin, pp 311–318
24. Landauer R (1961) Irreversibility and heat generation in the computing process. IBM J Res Dev 5:183–191
25. Lecerf Y (1963) Machines de Turing réversibles – Reursive insolubilité en $n \in N$ de l'équation $u = \theta^n u$, où θ est un isomorphisme de codes. Comptes Rendus Hebd Séances L'acad Sci 257:2597–2600
26. Lee J, Peper F, Adachi S, Mashiko S (2004) On reversible computation in asynchronous systems. In: Hida T, Saito k, Si S (eds) Quantum Information and Complexity. World Scientific, Singapore, pp 296–320
27. Lee J, Peper F, Adachi S, Morita K (2004) An asynchronous cellular automaton implementing 2-state 2-input 2-output reversed-twin reversible elements, Proc Eighth Int Conf on Cellular Automata for Research and Industry, Yokohama, Sep 2008, (to appear)
28. Margolus N (1984) Physics-like model of computation. Physica 10D:81–95
29. Merkle RC (1993) Reversible electronic logic using switches. Nanotechnology 4:20–41
30. Minsky ML (1967) Computation: Finite and Infinite Machines. Prentice-Hall, Englewood Cliffs
31. Morita K, Shirasaki A, Gono Y (1989) A 1-tape 2-symbol reversible Turing machine. Trans IEICE Japan E-72:223–228
32. Morita K, Harao M (1989) Computation universality of one-dimensional reversible (injective) cellular automata. Trans IEICE Japan E-72:758–762
33. Morita K (1990) A simple construction method of a reversible finite automaton out of Fredkin gates, and its related problem. Trans IEICE Japan E-73:978–984
34. Morita K (1996) Universality of a reversible two-counter machine. Theoret Comput Sci 168:303–320
35. Morita K (2001) Cellular automata and artificial life – computation and life in reversible cellular automata. In: Goles E, Martinez S (eds) Complex Systems. Kluwer Academic Publishers, Dordrecht, pp 151–200
36. Morita K (2001) A simple reversible logic element and cellular automata for reversible computing. In: Proc 3rd Int Conf on Machines, Computations, and Universality. LNCS, vol 2055. Springer, Berlin, pp 102–113
37. Morita K (2003) A new universal logic element for reversible computing. In: Martin-Vide C, Mitrana V (eds) Grammars and Automata for String Processing. Taylor & Francis, London, pp 285–294
38. Morita K, Ogiro T, Tanaka K, Kato H (2005) Classification and universality of reversible logic elements with one-bit memory. In: Proc 4th Int Conf on Machines, Computations, and Universality. LNCS, vol 3354. Springer, Berlin, pp 245–256
39. Morita K, Yamaguchi Y (2007) A universal reversible Turing machine. In: Proc 5th Int Conf on Machines, Computations, and Universality. LNCS, vol 4664. Springer, Berlin, pp 90–98
40. Ogiro T, Kanno A, Tanaka K, Kato H, Morita K (2005) Nondegenerate 2-state 3-symbol reversible logic elements are all universal. Int J Unconv Comput 1:47–67
41. Peper F, Lee J, Adachi S, Mashiko S (2003) Laying out circuits on asynchronous cellular arrays: a step towards feasible nanocomputers. Nanotechnology 14:469–485
42. Petri CA (1967) Grundsätzliches zur Beschreibung diskreter Prozesse. In: Proc 3rd Colloquium über Automatentheorie. Birkhäuser, Basel, pp 121–140
43. Pin JE (1992) On reversible automata. In: Proc Latin American Symp on Theoretical Informatics. LNCS vol 583. Springer, Berlin, pp 401–416
44. Rogozhin Y (1996) Small universal Turing machines. Theor Comput Sci 168:215–240
45. Shende VV, Prasad AK, Markov IL, Hayes JP (2003) Synthesis of reversible logic circuits. IEEE Trans Computer-Aided Des Integr Circuits Syst 22:710–722
46. Toffoli T (1977) Computation and construction universality of reversible cellular automata. J Comput Syst Sci 15:213–231
47. Toffoli T (1980) Reversible computing, Automata, Languages and Programming. In: LNCS, vol 85. Springer, Berlin, pp 632–644
48. Toffoli T (1981) Bicontinuous extensions of invertible combinatorial functions. Math Syst Theory 14:12–23
49. Toffoli T, Margolus N (1990) Invertible cellular automata: a review. Physica D 45:229–253

Books and Reviews

Adamatzky A (ed) (2002) Collision-Based Computing. Springer, London
Bennett CH (1988) Notes on the history of reversible computation. IBM J Res Dev 32:16–23
Milburn GJ (1998) The Feynman Processor. Perseus Books, Reading
Vitanyi P (2005) Time, space, and energy in reversible computing. In: Proc 2005 ACM Int Conf on Computing Frontiers, Ischia, pp 435–444

Robotic Networks, Distributed Algorithms for

Francesco Bullo[1], Jorge Cortés[2], Sonia Martínez[2]
[1] Department of Mechanical Engineering, University of California, Santa Barbara, USA
[2] Department of Mechanical and Aerospace Engineering, University of California, San Diego, USA

Article Outline

Glossary

Cooperative control In recent years, the study of groups of robots and multi-agent systems has received a lot of attention. This interest has been driven by the envisioned applications of these systems in scientific and commercial domains. From a systems and control theoretic perspective, the challenges in cooperative control revolve around the analysis and design of distributed coordination algorithms that integrate the individual capabilities of the agents to achieve a desired coordination task.

Distributed algorithm In a network composed of multiple agents, a coordination algorithm specifies a set of instructions for each agent that prescribe what to sense, what to communicate and to whom, how to process the information received, and how to move and interact with the environment. In order to be scalable, coordination algorithms need to rely as much as possible on local interactions between neighboring agents.

Complexity measures Coordination algorithms are designed to enable networks of agents achieve a desired task. Since different algorithms can be designed to achieve the same task, performance metrics are necessary to classify them. Complexity measures provide a way to characterize the properties of coordination algorithms such as completion time, cost of communication, energy consumption, and memory requirements.

Averaging algorithms Distributed coordination algorithms that perform weighted averages of the information received from neighboring agents are called averaging algorithms. Under suitable connectivity assumptions on the communication topology, averaging algorithms achieve agreement, i. e., the state of all agents approaches the same value. In certain cases, the agreement value can be explicitly determined as a function of the initial state of all agents.

Leader election In leader election problems, the objective of a network of processors is to elect a leader. All processors have a variable "leader" initially set to unknown. The leader-election task is solved when only one processor has set the variable "leader" to `true`, and all other processors have set it to `false`.

LCR algorithm The classic Le Lann–Chang–Roberts (LCR) algorithm solves the leader election task on a static network with the ring communication topology. Initially, each agent transmits its unique identifier to its neighbors. At each communication round, each agent compares the largest identifier received from other agents with its own identifier. If the received identifier is larger than its own, the agent declares itself a non-leader, and transmits it in the next communication round to its neighbors. If the received identifier is smaller than its own, the agent does nothing. Finally, if the received identifier is equal to its own, it declares itself a leader. The LCR algorithm achieves leader election with linear time complexity and quadratic total communication complexity, respectively.

Agree-and-pursue algorithm Coordination algorithms for robotic networks combine the features of distributed algorithms for networks of processors with the sensing and control capabilities of the robots. The agree-and-pursue motion coordination algorithm is an example of this fusion. Multiple robotic agents moving on a circle seek to agree on a common direction of motion while at the same achieving an equally-spaced distribution along the circle. The agree-and-pursue algorithm achieves both tasks combining ideas from leader election on a changing communication topology with basic control primitives such as "follow your closest neighbor in your direction of motion."

Definition of the Subject

The study of distributed algorithms for robotic networks is motivated by the recent emergence of low-power, highly-autonomous devices equipped with sensing, communication, processing, and control capabilities. In the near future, cooperative robotic sensor networks will perform critical tasks in disaster recovery, homeland security, and environmental monitoring. Such networks will require efficient and robust distributed algorithms with guaranteed quality-of-service. In order to design coordination algorithms with these desirable capabilities, it is necessary to develop new frameworks to design and formalize the operation of robotic networks and novel tools to analyze their behavior.

Introduction

Distributed algorithms are a classic subject of study for networks composed of individual processors with communication capabilities. Within the automata-theoretic literature, important research topics on distributed algo-

rithms include the introduction of mathematical models and precise specifications for their behavior, the formal assessment of their correctness, and the characterization of their complexity.

Robotic networks have distinctive features that make them unique when compared with networks of static processors. These features include the operation under ad-hoc dynamically changing communication topologies and the complexity that results from the combination of continuous- and discrete-time dynamics. The spatially-distributed character of robotic networks and their dynamic interaction with the environment make the classic study of distributed algorithms, typically restricted to static networks, not directly applicable.

This chapter brings together distributed algorithms for networks of processors and for robotic networks. The first part of the chapter is devoted to a formal discussion about distributed algorithms for a synchronous network of processors. This treatment serves as a brief introduction to important issues considered in the literature on distributed algorithms such as network evolution, task completion, and complexity notions. To illustrate the ideas, we consider the classic Le Lann–Chang–Roberts (LCR) algorithm, which solves the leader election task on a static network with the ring communication topology. Next, we present a class of distributed algorithms called averaging algorithms, where each processor computes a weighted average of the messages received from its neighbors. These algorithms can be described as linear dynamical systems, and their correctness and complexity analysis has nice connections with the fields of linear algebra and Markov chains.

The second part of the chapter presents a formal model for robotic networks that explicitly takes into account communication, sensing, control, and processing. The notions of time, communication, and space complexity introduced here allow us to characterize the performance of coordination algorithms, and rigorously compare the performance of one algorithm versus another. In general, the computation of these notions is a complex problem that requires a combination of tools from dynamical systems, control theory, linear algebra, and distributed algorithms. We illustrate these concepts in three different scenarios: the agree-and-pursue algorithm for a group of robots moving on a circle, aggregation algorithms that steer robots to a common location, and deployment algorithms that make robots optimally cover a region of interest. In each case, we report results on the complexity associated to the achievement of the desired task. The chapter ends with a discussion about future research directions.

Distributed Algorithms on Networks of Processors

Here we introduce a synchronous network as a group of processors with the ability to exchange messages and perform local computations. What we present is a basic classic model studied extensively in the distributed algorithms literature. Our treatment is directly adopted with minor variations from the texts [57] and [74].

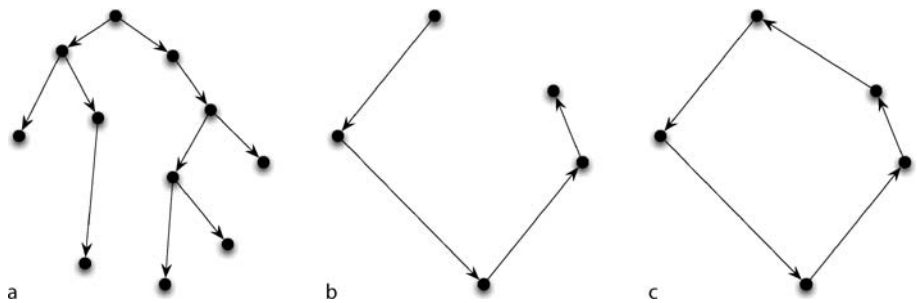
Physical Components and Computational Models

Loosely speaking, a synchronous network is a group of processors, or nodes, that possess a local state, exchange messages among neighbors, and compute an update to their local state based on the received messages. Each processor alternates the two tasks of exchanging messages with its neighboring processors and of performing a computation step.

Let us begin by providing some basic definitions. A *directed graph* [22], in short *digraph*, of order n is a pair $G = (V, E)$ where V is a set with n elements called *vertices* (or sometimes *nodes*) and E is a set of ordered pair of vertices called *edges*. In other words, $E \subseteq V \times V$. We call V and E the *vertex set* and *edge set*, respectively. For $u, v \in V$, the ordered pair (u, v) denotes an edge *from* u *to* v. The vertex u is called an *in-neighbor* of v, and v is called an *out-neighbor* of u. A *directed path* in a digraph is an ordered sequence of vertices such that any two consecutive vertices in the sequence are a directed edge of the digraph. A vertex of a digraph is *globally reachable* if it can be reached from any other vertex by traversing a directed path. A digraph is *strongly connected* if every vertex is globally reachable.

A *cycle* in a digraph is a non-trivial directed path that starts and ends at the same vertex. A digraph is *acyclic* if it contains no cycles. In an acyclic graph, every vertex with no in-neighbors is named *source*, and every vertex with no out-neighbors is named *sink*. A *directed tree* is an acyclic digraph with the following property: there exists a vertex, called the root, such that any other vertex of the digraph can be reached by one and only one path starting at the root. A *directed spanning tree*, or simply a *spanning tree*, of a digraph is a subgraph that is a directed tree and has the same vertex set as the digraph. A *directed chain* is a directed tree with exactly one source and one sink. A *directed ring digraph* is the cycle obtained by adding to the edge set of a chain a new edge from its sink to its source. Figure 1 illustrates these notions.

The physical component of a synchronous network S is a digraph (I, E_{cmm}), where



Robotic Networks, Distributed Algorithms for, Figure 1
From left to right, directed tree, chain, and ring digraphs

(i) $I = \{1, \ldots, n\}$ is called the *set of unique identifiers (UIDs)*, and
(ii) E_{cmm} is a set of directed edges over the vertices $\{1, \ldots, n\}$, called the communication links.

The set E_{cmm} models the topology of the communication service among the nodes: for $i, j \in \{1, \ldots, n\}$, processor i can send a message to processor j if the directed edge (i, j) is present in E_{cmm}. Note that, unlike the standard treatments in [57] and [74], we do not assume the digraph to be strongly connected; the required connectivity assumption is specified on a case by case basis.

Next, we discuss the state and the algorithms that each processor possesses and executes, respectively. By convention, we let the superscript $[i]$ denote any quantity associated with the node i. A *distributed algorithm* $\mathcal{DA}$ for a network S consists of the sets:

(i) $\mathbb{A}$, a set containing the `null` element, called the *communication alphabet*; elements of $\mathbb{A}$ are called *messages*;
(ii) $W^{[i]}$, $i \in I$, called the *processor state sets*;
(iii) $W_0^{[i]} \subseteq W^{[i]}$, $i \in I$, sets of *allowable initial values*;

and of the maps:

(i) $\text{msg}^{[i]} : W^{[i]} \times I \to \mathbb{A}$, $i \in I$, called *message-generation functions*;
(ii) $\text{stf}^{[i]} : W^{[i]} \times \mathbb{A}^n \to W^{[i]}$, $i \in I$, called *state-transition functions*.

If $W^{[i]} = W$, $\text{msg}^{[i]} = \text{msg}$, and $\text{stf}^{[i]} = \text{stf}$ for all $i \in I$, then $\mathcal{DA}$ is said to be *uniform* and is described by a tuple $(\mathbb{A}, W, \{W_0^{[i]}\}_{i \in I}, \text{msg}, \text{stf})$.

Now, with all elements in place, we can explain in more detail how a synchronous network executes a distributed algorithm. The *state* of processor i is a variable $w^{[i]} \in W^{[i]}$, initially set equal to an allowable value in $W_0^{[i]}$. At each time instant $\ell \in \mathbb{Z}_{\geq 0}$, processor i sends to each of its out-neighbors j in the communication digraph (I, E_{cmm}) a message (possibly the `null` message) computed by applying the message-generation function $\text{msg}^{[i]}$ to the current values of its state $w^{[i]}$ and to the identity j. Subsequently, but still at time instant $\ell \in \mathbb{Z}_{\geq 0}$, processor i updates the value of its state $w^{[i]}$ by applying the state-transition function $\text{stf}^{[i]}$ to the current value of its state $w^{[i]}$ and to the messages it receives from its in-neighbors. Note that, at each round, the first step is transmission and the second one is computation.

We conclude this section with two sets of remarks. We first discuss some aspects of our communication model that have a large impact on subsequent development. We then collect a few general comments about distributed algorithms on networks.

Remark 1 (Aspects of the communication model)

(i) The network S and the algorithm $\mathcal{DA}$ are referred to as *synchronous* because the communications between all processors takes place at the same time for all processors.
(ii) Communication is modeled as a so-called "point to point" service: a processor can specify different messages for different out-neighbors and knows the processor identity corresponding to any incoming message.
(iii) Information is exchanged between processors as messages, i. e., elements of the alphabet $\mathbb{A}$; the message `null` indicates no communication. Messages might encode logical expressions such as `true` and `false`, or finite-resolution quantized representations of integer and real numbers.
(iv) In some uniform algorithms, the messages between processors are the processors' states. In such cases, the corresponding communication alphabet is $\mathbb{A} = W \cup \{\texttt{null}\}$ and the message generation function $\text{msg}_{\text{std}}(w, j) = w$ is referred to as the *standard message-generation function*.

Remark 2 (Advanced topics: Control structures and failures)

(i) Processors in a network have only partial information about the network topology. In general, each processor only knows its own UID, and the UID of its in- and out-neighbors. Sometimes we assume that the processor knows the network diameter. In some cases [74], actively running networks might depend upon "control structures," i. e., structures that are computed at initial time and are exploited in subsequent algorithms. For example, routing tables might be computed for routing problems, "leader" processors might be elected and tree structures might be computed and represented in a distributed manner for various tasks, e. g., coloring or maximal independent set problems. We present some sample algorithms to compute these structures below.

(ii) A key issue in the study of distributed algorithms is the possible occurrence of failures. A network might experience intermittent or permanent communication failures: along given edges a `null` message or an arbitrary message might be delivered instead of the intended value. Alternatively, a network might experience various types of processor failures: a processor might transmit only `null` messages (i. e., the msg function returns `null` always), a processor might quit updating its state (i. e., the stf function neglects incoming messages and returns the current state value), or a processor might implement arbitrarily modified msg and stf functions. The latter situation, in which completely arbitrary and possibly malicious behavior is adopted by faulty nodes, is referred to as a Byzantine failure in the distributed algorithms literature.

Complexity Notions

Here we begin our analysis of the performance of distributed algorithms. We introduce a notion of algorithm completion and, in turn, we introduce the classic notions of time, space, and communication complexity.

We say that an algorithm *terminates* when only `null` messages are transmitted and all processors states become constants.

Remark 3 (Alternative termination notions)

(i) In the interest of simplicity, we have defined evolutions to be unbounded in time and we do not explicitly require algorithms to actually have termination conditions, i. e., to be able to detect when termination takes place.

(ii) It is also possible to define the termination time as the first instant when a given problem or task is achieved, independently of the fact that the algorithm might continue to transmit data subsequently.

The notion of time complexity measures the time required by a distributed algorithm to terminate. More specifically, the *(worst-case) time complexity* of a distributed algorithm $\mathcal{DA}$ on a network S, denoted $\mathrm{TC}(\mathcal{DA}, S)$, is the maximum number of rounds required by the execution of $\mathcal{DA}$ on S among all allowable initial states until termination.

Next, we quantify memory and communication requirements of distributed algorithms. From an information theory viewpoint [35], the information content of a memory variable or of a message is properly measured in bits. On the other hand, it is convenient to use the alternative notions of "basic memory unit" and "basic message." It is customary [74] to assume that a "basic memory unit" or a "basic message" contains $\log(n)$ bits so that, for example, the information content of a robot identifier $i \in \{1, \ldots, n\}$ is $\log(n)$ bits or, equivalently, one "basic memory unit." Note that elements of the processor state set W or of the alphabet set $\mathbb{A}$ might amount to multiple basic memory units or basic messages; the `null` message has zero cost. Unless specified otherwise, the following definitions and examples are stated in terms of basic memory unit and basic messages.

(i) The *(worst-case) space complexity* of a distributed algorithm $\mathcal{DA}$ on a network S, denoted by $\mathrm{SC}(\mathcal{DA}, S)$, is the maximum number of basic memory units required by a processor executing the $\mathcal{DA}$ on S among all processors and among all allowable initial states until termination.

(ii) The *(worst-case) communication complexity* of a distributed algorithm $\mathcal{DA}$ on a network S, denoted by $\mathrm{CC}(\mathcal{DA}, S)$, is the maximum number of basic messages transmitted over the entire network during the execution of $\mathcal{DA}$ among all allowable initial states until termination.

Remark 4 (Space complexity conventions) By convention, each processor knows its identity, i. e., it requires $\log(n)$ bits to represent its unique identifier in a set with n distinct elements. We do not count this cost in the space complexity of an algorithm.

We conclude this section by discussing ways of quantifying time, space, and communication complexity. The idea, borrowed from combinatorial optimization, is to adopt asymptotic "order of magnitude" measures. Formally, complexity bounds will be expressed with respect

to the Bachman–Laundau symbols O, Ω, and Θ. Let us be more specific. In the following definitions, f denotes a function from $\mathbb{N}$ to $\mathbb{R}$.

(i) We say that an algorithm has time complexity *of order* $\Omega(f(n))$ *over some network* if, for all n, there exists a network of order n and initial processor values such that the time complexity of the algorithm is greater than a constant factor times $f(n)$.
(ii) We say that an algorithm has time complexity *of order* $O(f(n))$ *over arbitrary networks* if, for all n, for all networks of order n and for all initial processor values the time complexity of the algorithm is lower than a constant factor times $f(n)$.
(iii) We say that an algorithm has time complexity *of order* $\Theta(f(n))$ if its time complexity is of order $\Omega(f(n))$ over some network and $O(f(n))$ over arbitrary networks at the same time.

We use similar conventions for space and communication complexity.

In many cases the complexity of an algorithm will typically depend upon the number of nodes of the network. It is therefore useful to present a few simple facts about these functions now. Over arbitrary digraphs $S = (I, E_{\text{cmm}})$ of order n, we have

$$\text{diam}(S) \in \Theta(n), \quad |E_{\text{cmm}}(S)| \in \Theta(n^2) \quad \text{and} \quad \text{radius}(v, S) \in \Theta(\text{diam}(S)) ,$$

where v is any node of S.

Remark 5 Numerous variations of these definitions are possible. Even though we will not pursue them here, let us provide some pointers.

(i) In the definition of lower bound, consider the logic quantifier describing the role of the network. The lower bound statement is "existential" rather than "global," in the sense that the bound does not hold for all graphs. As discussed in [74], it is possible to define also "global" lower bounds, i. e., lower bounds over all graphs, or lower bounds over specified classes of graphs.
(ii) The complexity notions introduced above focus on the worst-case situation. It is also possible to define *expected* or *average* complexity notions, where one might be interested in characterizing, for example, the average number of rounds required or the average number of basic messages transmitted over the entire network during the execution of an algorithm among all allowable initial states until termination.
(iii) It is possible to define complexity notions for problems, rather than algorithms, by considering, for example, the worst-case optimal performance among all algorithms that solve the given problem, or over classes of algorithms or classes of graphs.

Leader Election

We formulate here a classical problem in distributed networks and summarize its complexity measures.

Problem 6 (Leader election) *Assume that all processors of a network have a state variable, say* `leader`*, initially set to* `unknown`*. We say that a leader is elected when one and only one processor has the state variable set to* `true` *and all others have it set to* `false`*. Elect a leader.*

This is a task that is a bit more global in nature. We display here a solution that requires individual processors to know the diameter of the network, denoted by diam(S), or an upper bound on it.

[Informal description] At each communication round, each agent sends to all its neighbors the maximum UID it has received up to that time. This is repeated for diam(S) rounds. At the last round, each agent compares the maximum received UID with its own, and declares itself a leader if they coincide, or a non-leader otherwise.

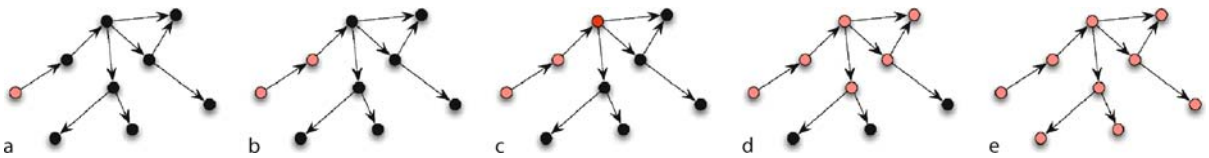
The algorithm is called FLOODMAX: the maximum UID in the network is transmitted to other agents in an incremental fashion. At the first communication round, agents that are neighbors of the agent with the maximum UID receive the message from it. At the next communication round, the neighbors of these agents receive the message with the maximum UID. This process goes on for diam(S) rounds to ensure that every agent receives the maximum UID. Note that there are networks for which all agents receive the message with the maximum UID in fewer communication rounds than diam(S). The algorithm is formally stated as follows.

`Synchronous Network:` $S = (\{1, \ldots, n\}, E_{\text{cmm}})$

`Distributed Algorithm: FLOODMAX`

`Alphabet:` $\mathbb{A} = \{1, \ldots, n\} \cup \{\texttt{null}\}$

`Processor State:` $w =$ (`my-id`, `max-id`, `leader`, `round`), where



Robotic Networks, Distributed Algorithms for, Figure 2
Execution of the FLOODMAX algorithm. The diameter of the network is 4. In the leftmost frame, the agent with the maximum UID is colored in red. After 4 communication rounds, its message has been received by all agents

```
my-id ∈ {1, ..., n},
        initially: my-id[i] = i  for all i
max-id ∈ {1, ..., n},
        initially: max-id[i] = i  for all i
leader ∈ {false, true, unknown},
        initially: leader[i] = unknown  for all i
round ∈ {0, 1, ..., diam(S)},
        initially: round[i] = 0  for all i
```

```
function msg(w, i)
 1: if {round < diam(S)} then
 2:    return max-id
 3: else
 4:    return null
```

```
function stf(w, y)
 1: new-id := max{max-id, largest identifier in y}
 2: case
 3:    round < diam(S):   new-lead :=
                                  unknown
 4:    round = diam(S) AND max-id = my-id:
                          new-lead := true
 5:    round = diam(S) AND max-id > my-id:
                         new-lead := false
 6: return (my-id, new-id, new-lead,
                                  round + 1)
```

Figure 2 shows an execution of the FLOODMAX algorithm.

The properties of the algorithm are characterized in the following lemma. A complete analysis of this algorithm, including modifications to improve the communication complexity, is discussed in [Section 4.1 in 1].

Lemma 7 (Complexity upper bounds for the FLOODMAX algorithm) *For a network S containing a spanning tree, the* FLOODMAX *algorithm has communication complexity in $O(\mathrm{diam}(S)|E_{cmm}|)$, time complexity equal to* $\mathrm{diam}(S)$*, and space complexity in $\Theta(1)$.*

A simplification of the FLOODMAX algorithm leads to the Le Lann–Chang–Roberts (LCR) algorithm for leader election in rings, see [Chapter 3.3 in 1], that we describe next.[1]. The LCR algorithm runs on a ring digraph and does not require the agents to know the diameter of the network.

[Informal description] At each communication round, if the agent receives from its in-neighbor a UID that is larger than the UIDs received earlier, then the agent records the received UID and forwards it to the out-neighbor during the following communication round. (Agents do not record the number of communication rounds.) When the agent with the maximum UID receives its own UID from a neighbor, it declares itself the leader.

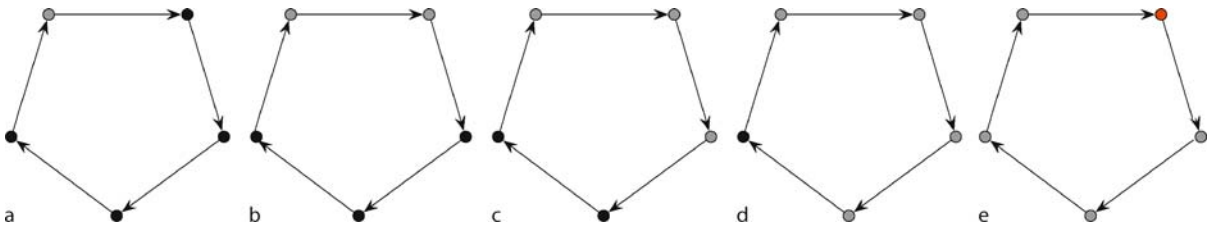
The algorithm is formally stated as follows.

```
Synchronous Network: ring digraph
Distributed Algorithm: LCR
Alphabet: 𝔸 = {1, ..., n} ∪ {null}
Processor State: w = (my-id, max-id,
leader, snd-flag), where
 my-id ∈ {1, ..., n},
        initially: my-id[i] = i  for all i
 max-id ∈ {1, ..., n},
        initially: max-id[i] = i  for all i
 leader ∈ {true, false, unknown},
        initially: leader[i] = unknown  for all i
 snd-flag ∈ {true, false},
        initially: snd-flag[i] = true  for all i
```

```
function msg(w, i)
 1: if {snd-flag = true} then
 2:    return max-id
 3: else
 4:    return null
```

```
function stf(w, y)
  1:  case
```

[1]Note that the description of the LCR algorithm given here is slightly different from the classic one as presented in [57].



Robotic Networks, Distributed Algorithms for, Figure 3
Execution of the LCR algorithm. In the leftmost frame, the agent with the maximum UID is colored in *red*. After 5 communication rounds, this agent receives its own UID from its in-neighbor and declares itself the leader

```
2:    (y contains only null msgs) OR
            (largest identifier in y < my-id):
3:    new-id := max-id
4:    new-lead := leader
5:    new-snd-flag := false
6:    (largest identifier in y = my-id):
7:    new-id := max-id
8:    new-lead := true
9:    new-snd-flag := false
10:   (largest identifier in y > my-id):
11:   new-id := largest identifier in y
12:   new-lead := false
13:   new-snd-flag := true
14: return (my-id, new-id, new-lead,
                                new-snd-flag)
```

Figure 3 shows an execution of the LCR algorithm. The properties of the LCR algorithm can be characterized as follows [57].

Lemma 8 (Complexity upper bounds for the LCR algorithm) *For a ring network S of order n, the* LCR *algorithm has communication complexity in* $\Theta(n^2)$*, time complexity equal to n, and space complexity in* $\Theta(1)$.

Averaging Algorithms

This section provides a brief introduction to a special class of distributed algorithms called averaging algorithms. The synchronous version of averaging algorithms can be modeled within the framework of synchronous networks. In an averaging algorithm, each processor updates its state by computing a weighted linear combination of the state of its neighbors. Computing linear combinations of the initial states of the processors is one the most basic computations that a network can implement. Averaging algorithms find application in optimization, distributed decision-making, e. g., collective synchronization, and have a long and rich history, see e. g., [29,30,45,65,97,98]. The richness comes from the vivid analogies with physical processes of diffusion, with Markov chain models, and the theory of positive matrices developed by Perron and Frobenius, see e. g., [19,20,44].

Averaging algorithms are defined by stochastic matrices. For completeness, let us recall some basic linear algebra definitions. A matrix $A \in \mathbb{R}^{n\times n}$ with entries a_{ij}, $i, j \in \{1, \ldots, n\}$, is

(i) *nonnegative* (resp., *positive*) if all its entries are nonnegative (resp., positive);
(ii) *row-stochastic* (or *stochastic* for brevity) if it is nonnegative and $\sum_{j=1}^{n} a_{ij} = 1$, for all $i \in \{1, \ldots, n\}$; in other words, A is row-stochastic if

$$A\mathbf{1}_n = \mathbf{1}_n;$$

where $\mathbf{1}_n = (1, \ldots, 1)^T \in \mathbb{R}^n$.
(iii) *doubly stochastic* if it is row-stochastic and column-stochastic, where we say that A is *column-stochastic* if $\mathbf{1}_n^T A = \mathbf{1}_n^T$;
(iv) *irreducible* if, for any nontrivial partition $J \cup K$ of the index set $\{1, \ldots, n\}$, there exists $j \in J$ and $k \in K$ such that $a_{jk} \neq 0$.

We are now ready to introduce the class of averaging algorithms. The *averaging algorithm* associated to a sequence of stochastic matrices $\{F(\ell) \mid \ell \in \mathbb{Z}_{\geq 0}\} \subset \mathbb{R}^{n\times n}$ is the discrete-time dynamical system

$$w(\ell + 1) = F(\ell) \cdot w(\ell)\,, \quad \ell \in \mathbb{Z}_{\geq 0}\,. \tag{1}$$

In the literature, averaging algorithms are also often referred to as agreement algorithms or as consensus algorithms.

Averaging algorithms are naturally associated with weighted digraphs, i. e., digraphs whose edges have weights. More precisely, a *weighted digraph* is a triplet $G = (V, E, A)$ where $V = \{v_1, \ldots, v_n\}$ and E are a digraph and where $A \in \mathbb{R}_{\geq 0}^{n\times n}$ is a *weighted adjacency matrix* with the following properties: for $i, j \in \{1, \ldots, n\}$, the entry $a_{ij} > 0$ if (v_i, v_j) is an edge of G, and $a_{ij} = 0$ otherwise. In other words, the scalars a_{ij}, for all $(v_i, v_j) \in E$,

are a set of weights for the edges of G. Note that edge set is uniquely determined by the weighted adjacency matrix and it can be therefore omitted. The *weighted out-degree matrix* $D_{\text{out}}(G)$ and the *weighted in-degree matrix* $D_{\text{in}}(G)$ are the diagonal matrices defined by

$$D_{\text{out}}(G) = \text{diag}\,(A\mathbf{1}_n)\,, \quad \text{and} \quad D_{\text{in}}(G) = \text{diag}\left(A^T\mathbf{1}_n\right).$$

The weighted digraph G is *weight-balanced* if $D_{\text{out}}(G) = D_{\text{in}}(G)$. Given a nonnegative $n \times n$ matrix A, its *associated weighted digraph* is the weighted digraph with nodes $\{1, \ldots, n\}$, and weighted adjacency matrix A. The unweighted version of this weighted digraph is called the *associated digraph*. The following statements can be proven:

(i) if A is stochastic, then its associated digraph has weighted out-degree matrix equal to I_n;
(ii) if A is doubly stochastic, then its associated weighted digraph is weight-balanced and additionally both in-degree and out-degree matrices are equal to I_n;
(iii) A is irreducible if and only if its associated weighted digraph is strongly connected.

Next, we characterize the convergence properties of averaging algorithms. Let us introduce a useful property of collections of stochastic matrices. Given $\alpha \in]0, 1]$, the set of *non-degenerate matrices with respect to* α consists of all stochastic matrices F with entries f_{ij}, for $i, j \in \{1, \ldots, n\}$, satisfying

$$f_{ii} \in [\alpha, 1]\,, \quad \text{and} \quad f_{ij} \in \{0\} \cup [\alpha, 1] \;\text{ for }\; j \neq i\,.$$

Additionally, the sequence of stochastic matrices $\{F(\ell) \mid \ell \in \mathbb{Z}_{\geq 0}\}$ is *non-degenerate* if these exists $\alpha \in]0, 1]$ such that $F(\ell)$ is non-degenerate with respect to α for all $\ell \in \mathbb{Z}_{\geq 0}$.

We now state the following convergence result from [65].

Theorem 9 (Convergence for time-dependent stochastic matrices) *Let $\{F(\ell) \mid \ell \in \mathbb{Z}_{\geq 0}\} \subset \mathbb{R}^{n\times n}$ be a non-degenerate sequence of stochastic matrices. For $\ell \in \mathbb{Z}_{\geq 0}$, let $G(\ell)$ be the unweighted digraph associated to $F(\ell)$. The following statements are equivalent:*

(i) the set $\text{diag}(\mathbb{R}^n)$ *is uniformly globally attractive for the associated averaging algorithm, that is, every evolution of the averaging algorithm at any time ℓ_0, approaches the set* $\text{diag}(\mathbb{R}^n)$ *in the following time-uniform manner:*

for all $\ell_0 \in \mathbb{Z}_{\geq 0}$, for all $w_0 \in \mathbb{R}^n$, and for all neighborhoods W of $\text{diag}(\mathbb{R}^n)$*, there exists a single $\tau_0 \in \mathbb{Z}_{\geq 0}$ such that the evolution $w : [\ell_0, +\infty[\to \mathbb{R}^n$ defined by $w(\ell_0) = w_0$, takes value in W for all times $\ell \geq \ell_0 + \tau_0$.*

(ii) there exists a duration $\delta \in \mathbb{N}$ such that, for all $\ell \in \mathbb{Z}_{\geq 0}$, the digraph

$$G(\ell + 1) \cup \cdots \cup G(\ell + \delta)$$

contains a globally reachable vertex.

Distributed Algorithms for Robotic Networks

This section describes models and algorithms for groups of robots that process information, sense, communicate, and move. We refer to such systems as robotic networks. In this section we review and survey a few modeling and algorithmic topics in robotic coordination; earlier versions of this material were originally presented in [15,59,60,61].

The section is organized as follows. First, we present the physical components of a network, that is, the mobile robots and the communication service connecting them. We then present the notion of control and communication law, and how a law is executed by a robotic network. We then discuss complexity notions for robotic networks. As an example of these notions, we introduce a simple law, called the agree-and-pursue law, which combines ideas from leader election algorithms and from cyclic pursuit (i. e., a game in which robots chase each other in a circular environment). We then consider in some detail algorithms for two basic motion coordination tasks, namely aggregation and deployment. We briefly formalize these problems and provide some basic algorithms for these two tasks.

Robotic Networks and Complexity

The global behavior of a robotic network arises from the combination of the local actions taken by its members. Each robot in the network can perform a few basic tasks such as sensing, communicating, processing information, and moving according to it. The many ways in which these capabilities can be integrated make a robotic network a versatile and, at the same time, complex system. The following robotic network model provides the framework to formalize, analyze, and compare distinct distributed behaviors.

We consider *uniform networks of robotic agents* defined by a tuple $S = (I, \mathcal{R}, E_{\text{cmm}})$ consisting of a set of unique identifiers $I = \{1, \ldots, n\}$, a collection of control systems $\mathcal{R} = \{R^{[i]}\}_{i\in I}$, with $R^{[i]} = (X, U, X_0, f)$, and a map E_{cmm} from X^n to the subsets of $I \times I$ called the *communication edge map*. Here, (X, U, X_0, f) is a control system with state space $X \subset \mathbb{R}^d$, input space U, set of allowable initial states $X_0 \subset X$, and system dynamics $f : X \times U \to X$. An edge between two identifiers in E_{cmm}

implies the ability of the corresponding two robots to exchange messages. A *control and communication law* for S consists of the sets:

(i) $\mathbb{A}$, called the *communication language*, whose elements are called *messages*;
(ii) W, set of values of some *processor variables* $w^{[i]} \in W$, $i \in I$, and $W_0 \subseteq W$, subset of *allowable initial values*. These sets correspond to the capability of robots to allocate additional variables and store sensor or communication data;

and the maps:

(iii) $\mathrm{msg} : X \times W \times I \to \mathbb{A}$, called *message-generation function*;
(iv) $\mathrm{stf} : X \times W \times \mathbb{A}^n \to W$, called *state-transition function*;
(v) $\mathrm{ctl} : X \times W \times \mathbb{A}^n \to U$, called *control function*.

To implement a control and communication law each robot performs the following sequence or cycle of actions. At each instant $\ell \in \mathbb{Z}_{\geq 0}$, each robot i communicates to each robot j such that (i, j) belongs to $E_{\mathrm{cmm}}(x^{[1]}, \ldots, x^{[n]})$. Each robot i sends a message computed by applying the message-generation function to the current values of $x^{[i]}$ and $w^{[i]}$. After a negligible period of time, robot i resets the value of its logic variables $w^{[i]}$ by applying the state-transition function to the current value of $w^{[i]}$, and to the messages $y^{[i]}(\ell)$ received at ℓ. Between communication instants, i. e., for $t \in [\ell, \ell + 1)$, robot i applies a control action computed by applying the control function to its state at the last sample time $x^{[i]}(\ell)$, the current value of $w^{[i]}$, and to the messages $y^{[i]}(\ell)$ received at ℓ.

Remark 10 (Algorithm properties and congestion models)

(i) In our present definition, all robots are identical and implement the same algorithm; in this sense the control and communication law is called *uniform* (or anonymous). If $W = W_0 = \emptyset$, then the control and communication law is *static* (or memoryless) and no state-transition function is defined. It is also possible for a law to be *time-independent* if the three relevant maps do not depend on time. Finally, let us also remark that this is a synchronous model in which all robots share a common clock.
(ii) Communication and physical congestion affect the performance of robotic networks. These effects can be modeled by characterizing how the network parameters vary as the number of robots becomes larger. For example, in an ad hoc networks with n uniformly randomly placed nodes, it is known [42] that the maximum-throughput communication range $r(n)$ of each node decreases as the density of nodes increases; in d dimensions the appropriate scaling law is $r(n) \in \Theta((\log(n)/n)^{1/d})$. As a second example, it is reasonable to assume that, as the number of robots increase, so should the area available for their motion. An alternative convenient approach is the one taken by [88], where robots' safety zones decrease with decreasing robots' speed. This suggests that, in a fixed environment, individual nodes of a large ensemble have to move at a speed decreasing with n, and in particular, at a speed proportional to $n^{-1/d}$.

Next, we establish the notion of coordination task and of task achievement by a robotic network. Let S be a robotic network and let $\mathcal{W}$ be a set. A *coordination task* for S is a map $\mathcal{T} : X^n \times \mathcal{W}^n \to \{\mathtt{true}, \mathtt{false}\}$. If $\mathcal{W}$ is a singleton, then the coordination task is said to be *static* and can be described by a map $\mathcal{T} : X^n \to \{\mathtt{true}, \mathtt{false}\}$. Additionally, let CC a control and communication law for S.

(i) The law CC is *compatible* with the task $\mathcal{T} : X^n \times \mathcal{W}^n \to \{\mathtt{true}, \mathtt{false}\}$ if its processor state take values in $\mathcal{W}$, that is, if $W^{[i]} = \mathcal{W}$, for all $i \in I$.
(ii) The law CC *achieves* the task $\mathcal{T}$ if it is compatible with it and if, for all initial conditions $x_0^{[i]} \in X_0$ and $w_0^{[i]} \in W_0$, $i \in I$, there exists $T \in \mathbb{Z}_{\geq 0}$ such that the network evolution $\ell \mapsto (x(\ell), w(\ell))$ has the property that $\mathcal{T}(x(\ell), w(\ell)) = \mathtt{true}$ for all $\ell \geq T$.

In control-theoretic terms, achieving a task means establishing a convergence or stability result. Beside this key objective, one might be interested in efficiency as measured by required communication service, required control energy or by speed of completion. We focus on the latter notion.

(i) The *(worst-case) time complexity to achieve* $\mathcal{T}$ *with* CC *from* $(x_0, w_0) \in X_0^n \times W_0^n$ is

$$\mathrm{TC}(\mathcal{T}, CC, x_0, w_0) = \inf\{\ell \mid \mathcal{T}(x(k), w(k)) = \mathtt{true}\,, \text{ for all } k \geq \ell\}\,,$$

where $\ell \mapsto (x(\ell), w(\ell))$ is the evolution of (S, CC) from the initial condition (x_0, w_0);
(ii) The *(worst-case) time complexity to achieve* $\mathcal{T}$ *with* CC is

$$\mathrm{TC}(\mathcal{T}, CC) = \sup\left\{\mathrm{TC}(\mathcal{T}, CC, x_0, w_0) \mid (x_0, w_0) \in X_0^n \times W_0^n\right\}.$$

Some ideas on how to define meaningful notions of space and communication complexity are discussed in [59]. In the following discussion, we describe three coordination algorithms, which have been cast into this modeling framework and whose time complexity properties have been analyzed.

Agree-and-Pursue Algorithm

We begin our list of distributed algorithms with a simple law that is related to leader election algorithms, see Sect. "Leader Election", and to cyclic pursuit algorithms as studied in the control literature. Despite the apparent simplicity, this example is remarkable in that it combines a leader election task (in the processor states) with a uniform robotic deployment task (in the physical state), arguably two of the most basic tasks in distributed algorithms and cooperative control, respectively.

We consider n robots $\{\theta^{[1]}, \ldots, \theta^{[n]}\}$ in $\mathbb{S}^1$, moving along on the unit circle with angular velocity equal to the control input. Each robot is described by the tuple $(\mathbb{S}^1, [-u_{\max}, u_{\max}], \mathbb{S}^1, (0, \boldsymbol{e}))$, where $\boldsymbol{e}$ is the vector field on $\mathbb{S}^1$ describing unit-speed counterclockwise rotation. We assume that each robot can sense its own position and can communicate to any other robot within distance r along the circle. These data define the uniform robotic network S_{circle}.

[Informal description] The processor state consists of `dir` (the robot's direction of motion) taking values in $\{\texttt{c}, \texttt{cc}\}$ (meaning clockwise and counterclockwise) and `max-id` (the largest UID received by the robot, initially set to the robot's UID) taking values in I. At each communication round, each robot transmits its position and its processor state. Among the messages received from the robots moving towards its position, each robot picks the message with the largest value of `max-id`. If this value is larger than its own value, the agent resets its processor state with the selected message. Between communication rounds, each robot moves in the clockwise or counterclockwise direction depending on whether its processor state `dir` is `c` or `cc`. Each robot moves k_{prop} times the distance to the immediately next neighbor in the chosen direction, or, if no neighbors are detected, k_{prop} times the communication range r.

For this network and this law there are two tasks of interest. First, we define the *direction agreement task* $\mathcal{T}_{\texttt{dir}} : (\mathbb{S}^1)^n \times W^n \to \{\texttt{true}, \texttt{false}\}$ by

$$\mathcal{T}_{\texttt{dir}}(\theta, w) = \begin{cases} \texttt{true}, & \text{if } \texttt{dir}^{[1]} = \cdots = \texttt{dir}^{[n]}, \\ \texttt{false}, & \text{otherwise,} \end{cases}$$

where $\theta = (\theta^{[1]}, \ldots, \theta^{[n]})$, $w = (w^{[1]}, \ldots, w^{[n]})$, and $w^{[i]} = (\texttt{dir}^{[i]}, \texttt{max-id}^{[i]})$, for $i \in I$. Furthermore, for $\varepsilon > 0$, we define the static *equidistance task* $\mathcal{T}_{\varepsilon\text{-eqdstnc}} : (\mathbb{S}^1)^n \to \{\texttt{true}, \texttt{false}\}$ to be `true` if and only if

$$\left| \min_{j \neq i} \text{dist}_{\text{c}}\left(\theta^{[i]}, \theta^{[j]}\right) - \min_{j \neq i} \text{dist}_{\text{cc}}\left(\theta^{[i]}, \theta^{[j]}\right) \right| < \varepsilon, \quad \text{for all } i \in I.$$

In other words, $\mathcal{T}_{\varepsilon\text{-eqdstnc}}$ is true when, for every agent, the distance to the closest clockwise neighbor and to the closest counterclockwise neighbor are approximately equal.

An implementation of this control and communication law is shown in Fig. 4. As parameters we select $n = 45$, $r = 2\pi/40$, $u_{\max} = 1/4$ and $k_{\text{prop}} = 7/16$. Along the evolution, all robots agree upon a common direction of motion and, after suitable time, they reach a uniform distribution. A careful analysis based on invariance properties and Lyapunov functions allows us to establish that, under appropriate conditions, indeed both tasks are achieved by the agree-and-pursue law [59].

Theorem 11 (Time complexity of agree-and-pursue law) *For $k_{prop} \in]0, \frac{1}{2}[$, in the limit as $n \to +\infty$ and $\varepsilon \to 0^+$, the network S_{circle} with $u_{max}(n) \geq k_{prop} r(n)$, the law $CC_{\text{AGREE \& PURSUE}}$, and the tasks $\mathcal{T}_{dir}$ and $\mathcal{T}_{\varepsilon\text{-eqdstnc}}$ together satisfy:*

(i) $\text{TC}(\mathcal{T}_{dir}, CC_{\text{AGREE \& PURSUE}}) \in \Theta(r(n)^{-1})$;

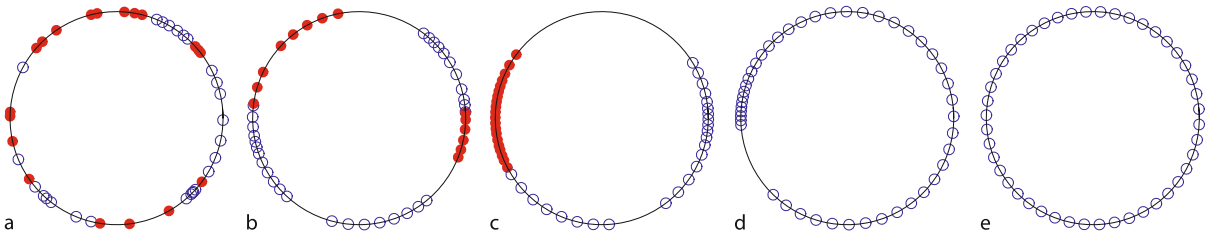
(ii) if $\delta(n)$ is lower bounded by a positive constant as $n \to +\infty$, then

$$\text{TC}(\mathcal{T}_{\varepsilon\text{-eqdstnc}}, CC_{\text{AGREE \& PURSUE}}) \in \Omega(n^2 \log(n\varepsilon)^{-1}),$$
$$\text{TC}(\mathcal{T}_{\varepsilon\text{-eqdstnc}}, CC_{\text{AGREE \& PURSUE}}) \in O(n^2 \log(n\varepsilon^{-1})).$$

If $\delta(n)$ is upper bounded by a negative constant, then $CC_{\text{AGREE \& PURSUE}}$ does not achieve $\mathcal{T}_{\varepsilon\text{-eqdstnc}}$ in general.

Finally we compare these results with the complexity result known for the leader election problem.

Remark 12 (Comparison with leader election) Let us compare the agree-and-pursue control and communication law with the classical Le Lann–Chang–Roberts (LCR) algorithm for leader election discussed in Sect. "Leader Election". The leader election task consists of electing a unique agent among all robots in the network; it is therefore different from, but closely related to, the coordination



Robotic Networks, Distributed Algorithms for, Figure 4
The AGREE & PURSUE law. *Disks* and *circles* correspond to robots moving counterclockwise and clockwise, respectively. The initial positions and the initial directions of motion are randomly generated. The five pictures depict the network state at times 0, 9, 20, 100, and 800

task $\mathcal{T}_{\mathtt{dir}}$. The LCR algorithm operates on a static network with the ring communication topology, and achieves leader election with time and total communication complexity, respectively, $\Theta(n)$ and $\Theta(n^2)$. The agree-and-pursue law operates on a robotic network with the $r(n)$-disk communication topology, and achieves $\mathcal{T}_{\mathtt{dir}}$ with time and total communication complexity, respectively, $\Theta(r(n)^{-1})$ and $O(n^2 r(n)^{-1})$. If wireless communication congestion is modeled by $r(n)$ of order $1/n$ as in Remark 10, then the two algorithms have identical time complexity and the LCR algorithm has better communication complexity. Note that computations on a possibly disconnected, dynamic network are more complex than on a static ring topology.

Aggregation Algorithms

The rendezvous objective (also referred to as the gathering problem) is to achieve agreement over the location of the robots, that is, to steer each agent to a common location. An early reference on this problem is [2]; more recent references include [26,32,52,53]. We consider two scenarios which differ in the robots' communication capabilities and the environment in which the robots move. First [26], we consider the problem of rendezvous for robots equipped with *range-limited communication* in obstacle-free environments. In this case, each robot is capable of sensing its position in the Euclidean space $\mathbb{R}^d$ and can communicate it to any other robot within a given distance r. This communication service is modeled by the r-disk graph, in which two robots are neighbors if and only if their Euclidean distance is less than or equal to r. Second [36], we consider *visually-guided robots*. Here the robots are assumed to move in a nonconvex simple polygonal environment Q. Each robot can sense, within line of sight, any other robot as well as the distance to the boundary of the environment. The relationship between the robots can be characterized by the so-called visibility graph: two robots are neighbors if and only if they are mutually visible to each other.

In both scenarios, the rendezvous problem cannot be solved with distributed information unless the robots' initial positions form a connected communication graph. Arguably, a good property of any rendezvous algorithm is that of maintaining connectivity between robots. This connectivity-maintenance objective is interesting on its own. It turns out that this objective can be achieved through local constraints on the robots' motion. Motion constraint sets that maintain connectivity are designed in [2,36] by exploiting the geometric properties of disk and visibility graphs.

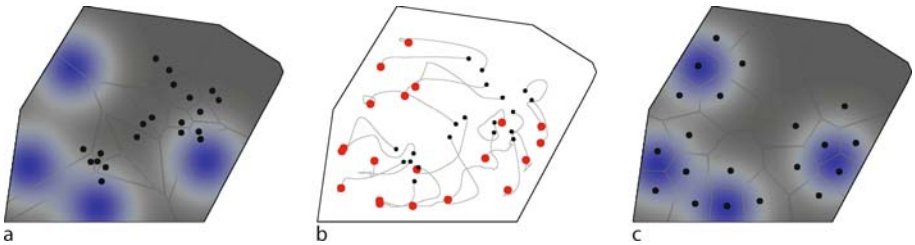
These discussions lead to the following algorithm that solves the rendezvous problems for both communication scenarios. The robots execute what is known as the *Circumcenter Algorithm*; here is an informal description. Each robot iteratively performs the following tasks:

```
1: acquire neighbors' positions
2: compute connectivity constraint set
3: move toward the circumcenter of
   the point set comprised of its
   neighbors and of itself, while
   remaining inside the connectivity
   constraint set.
```

One can prove that, under technical conditions, the algorithm does achieve the rendezvous task in both scenarios; see [26,36]. Additionally, when $d = 1$, it can be shown that the time complexity of this algorithm is $\Theta(n)$; see [60].

Deployment Algorithms

The problem of deploying a group of robots over a given region of interest can be tackled with the following simple heuristic. Each robot iteratively performs the following tasks:



Robotic Networks, Distributed Algorithms for, Figure 5
Deployment algorithm for the area-coverage problem. Each of the 20 robots moves toward the centroid of its Voronoi cell. This strategy corresponds to the network following the gradient of $\mathcal{H}_{\text{ave}}$. Areas of the polygon with greater importance are colored darker. Figures a and c show, respectively, the initial and final locations, with the corresponding Voronoi partitions. Figure b illustrates the gradient descent flow

```
1: acquire neighbors' positions
2: compute own dominance region
3: move towards the center of own
   dominance region
```

This short description can be made accurate by specifying what notions of dominance region and of center are to be adopted. In what follows we mention two examples and refer to [24,25,27,37] for more details.

First, we consider the *area-coverage deployment problem* in a convex environment Q. The objective is to maximize the area within close range of the mobile nodes. This models a scenario in which the nodes are equipped with some sensors that take measurements of some physical quantity in the environment, e. g., temperature or concentration. Assume that certain regions in the environment are more important than others and describe this by a density function ϕ. This problems leads to the coverage performance metric

$$\mathcal{H}_{\text{ave}}(p_1,\dots,p_n) = \int_Q \min_{i\in\{1,\dots,n\}} f(\|q-p_i\|)\phi(q)\mathrm{d}q = \sum_{i=1}^{n}\int_{V_i} f(\|q-p_i\|)\phi(q)\mathrm{d}q\,.$$

Here p_i is the position of the ith node, f measures the performance of an individual sensor, and $\{V_1,\dots,V_n\}$ is the Voronoi partition of the environment Q generated by the positions $\{p_1,\dots,p_n\}$. If we assume that each node obeys a first-order dynamical behavior, then a simple gradient scheme can be easily implemented in a spatially-distributed manner. Figure 5 shows an implementation of this gradient scheme. Following the gradient of $\mathcal{H}_{\text{ave}}$ corresponds, in the algorithm described above, to defining (1) the dominance regions to be the Voronoi cells generated by the robots, and (2) the center of a region to be the centroid of the region (if $f(x) = x^2$). Because the closed-loop system is a gradient flow for the cost function, performance is locally, continuously optimized. As a special case, when the environment is a segment and $\phi = 1$, the time complexity of the algorithm can be shown to be $O(n^3 \log(n\varepsilon^{-1}))$, where ε is an accuracy threshold below which we consider the task accomplished.

Second, we consider the problem of deploying *to maximize the likelihood of detecting a source*. For example, consider devices equipped with acoustic sensors attempting to detect a sound-source (or, similarly, antennas detecting RF signals, or chemical sensors localizing a pollutant source). For a variety of criteria, when the source emits a known signal and the noise is Gaussian, we know that the optimal detection algorithm involves a matched filter, that detection performance is a function of signal-to-noise-ratio, and, in turn, that signal-to-noise ratio is inversely proportional to the sensor-source distance. In this case, the appropriate cost function is

$$\mathcal{H}_{\text{worst}}(p_1,\dots,p_n) = \max_{q\in Q}\min_{i\in\{1,\dots,n\}} f(\|q-p_i\|) = \max_{q\in V_i} f(\|q-p_i\|)\,,$$

and a greedy motion coordination algorithm is for each node to move toward the circumcenter of its Voronoi cell. A detailed analysis [24] shows that the detection likelihood is inversely proportional to the circumradius of each node's Voronoi cell, and that, if the nodes follow the algorithm described above, then the detection likelihood increases monotonically as a function of time.

Bibliographical Notes

In this section we present a necessarily incomplete discussion of some relevant literature that we have not yet mentioned in the previous sections.

First, we review some literature on emergent and self-organized swarming behaviors in biological groups with distributed agent-to-agent interactions. Interesting

dynamical systems arise in biological networks at multiple levels of resolution, all the way from interactions among molecules and cells [62] to the behavioral ecology of animal groups [66]. Flocks of birds and schools of fish can travel in formation and act as one unit (see [72]), allowing these animals to defend themselves against predators and protect their territories. Wildebeest and other animals exhibit complex collective behaviors when migrating, such as obstacle avoiding, leader election, and formation keeping (see [41,89]). Certain foraging behaviors include individual animals partitioning their environment into nonoverlapping zones (see [7]). Honey bees [85], gorillas [93], and whitefaced capuchins [13] exhibit synchronized group activities such as initiation of motion and change of travel direction. These remarkable dynamic capabilities are achieved apparently without following a group leader; see [7,13,41,66,72,85,93] for specific examples of animal species and [21,28] for general studies.

With regards to distributed motion coordination algorithms, much progress has been made on pattern formation [9,46,86,94], flocking [67,96], self-assembly [49], swarm aggregation [38], gradient climbing [75], cyclic pursuit [14,58,90], vehicle routing [56,87], and connectivity maintenance problems [83,103].

Much research has been devoted to distributed task allocation problems. The work in [39] proposes a taxonomy of task allocation problems. In papers such as [1,40,63,84], advanced heuristic methods are developed, and their effectiveness is demonstrated through simulation or real world implementation. Distributed auction algorithms are discussed in [18,63] building on the classic works in [10,11]. A distributed MILP solver is proposed in [1]. A spatially distributed receding-horizon scheme is proposed in [33,73]. There has also been prior work on target assignment problems [6,91,102]. Target allocation for vehicles with nonholonomic constraints is studied in [76,81,82].

References with a focus on robotic networks include the survey in [99], the text [5] on behavior-based robotics, and the recent special issue [4] of the IEEE Transaction on Robotics and Automation. An important contribution towards a network model of mobile interacting robots is introduced in [94]. This model consists of a group of identical "distributed anonymous mobile robots" characterized as follows: no explicit communication takes place between them, and at each time instant of an "activation schedule," each robot senses the relative position of all other robots and moves according to a pre-specified algorithm. Communication complexity for control and communication algorithms A related model is presented in [32], where as few capabilities as possible are assumed on the agents, with the objective of understanding the limitations of multi-agent networks. A brief survey of models, algorithms, and the need for appropriate complexity notions is presented in [79]. Recently, a notion of communication complexity for control and communication algorithms in multi-robot systems is analyzed in [48], see also [50].

Finally, with regards to linear distributed algorithms we mention the following references, on top of the ones discussed in Sect. "Averaging Algorithms". Various results are available on continuous-time consensus algorithms [34,54,55,64,70,77], consensus over random networks [43,95,100], consensus algorithms for general functions [8,23], connections with the heat equation and partial difference equation [31], convergence in time-delayed and asynchronous settings [3,12], quantized consensus problems [47,80], applications to distributed signal processing [69,92,101], characterization of the convergence rates and time complexity [16,17,51,71]. Finally, two recent surveys are [68,78].

Future Directions

Robotic networks incorporate numerous subsystems. Their design is challenging because they integrate heterogeneous hardware and software components. Additionally, the operation of robotic networks is subject to computing, energy, cost, and safety constraints. The interaction with the physical world and the uncertain response from other members of the network are also integral parts to consider in the management of robotic networks. Traditional centralized approaches are not valid to satisfy the scalability requirements of these systems. Thus, in order to successfully deploy robotic and communication networks, it is necessary to expand our present knowledge about how to integrate and efficiently control them. The present chapter has offered a glimpse into these problems. We have presented verification and complexity tools to evaluate the cost of cooperative strategies that achieve a variety of tasks.

Networks of robotic agents are an example of the class of complex, networked systems which pervades our world. Understanding how to design robotic swarms requires the development of new fundamental theories that can explain the behavior of general networks evolving with time. Some of the desired properties of these systems are robustness, ease of control, predictability with time, guaranteed performance, and quality of service. Self-organization and distributed management would also allow for the minimal supervision necessary for scalability. However, devising systems that meet all these criteria is not an easy task. A small deviation by an agent or a change in certain parameters may produce a dramatic change in the

overall network behavior. Relevant questions that pertain these aspects are currently being approached from a range of disciplines such as systems and control, operations research, random graph theory, statistical physics, and game theory over networks. Future work will undoubtedly lead to a cross-fertilization of these and other areas that will help design efficient robotic sensor networks.

Acknowledgments

This material is based upon work supported in part by NSF CAREER Award CMS-0643679, NSF CAREER Award ECS-0546871, and AFOSR MURI Award FA9550-07-1-0528.

Bibliography

1. Alighanbari M, How JP (2006) Robust decentralized task assignment for cooperative UAVs. In: AIAA conference on guidance, navigation and control, Keystone, CO, August 2006
2. Ando H, Oasa Y, Suzuki I, Yamashita M (1999) Distributed memoryless point convergence algorithm for mobile robots with limited visibility. Trans IEEE Robotics Autom 15(5):818–828
3. Angeli D, Bliman PA (2006) Stability of leaderless discrete-time multi-agent systems. Math Control Signal Syst 18(4):293–322
4. Arai T, Pagello E, Parker LE (2002) Guest editorial: Advances in multirobot systems. Trans IEEE Robotics Autom 18(5):655–661
5. Arkin RC (1998) Behavior-Based Robotics. MIT Press, Cambridge, MA
6. Arslan G, Marden JR, Shamma JS (2007) Autonomous vehicle-target assignment: A game theoretic formulation. J ASME Dyn Syst Meas Control 129(5):584–596
7. Barlow GW (1974) Hexagonal territories. Anim Behav 22:876–878
8. Bauso D, Giarré L, Pesenti R (2006) Nonlinear protocols for optimal distributed consensus in networks of dynamic agents. Syst Control Lett 55(11):918–928
9. Belta C, Kumar V (2004) Abstraction and control for groups of robots. Trans IEEE Robotics 20(5):865–875
10. Bertsekas DP, Castañón DA (1991) Parallel synchronous and asynchronous implementations of the auction algorithm. Parallel Comput 17:707–732
11. Bertsekas DP, Castañón DA (1993) Parallel primal-dual methods for the minimum cost flow problem. Comput Optim Appl 2(4):317–336
12. Blondel VD, Hendrickx JM, Olshevsky A, Tsitsiklis JN (2005) Convergence in multiagent coordination, consensus, and flocking. In: IEEE conference on decision and control and european control conference, Seville, Spain, December 2005, pp 2996–3000
13. Boinski S, Campbell AF (1995) Use of trill vocalizations to coordinate troop movement among whitefaced capuchins – a 2nd field-test. Behaviour 132:875–901
14. Bruckstein AM, Cohen N, Efrat A (1991) Ants, crickets, and frogs in cyclic pursuit. Technical Report CIS 9105, Center for Intelligent Systems, Technion, Haifa, Israel. http://www.cs.technion.ac.il/tech-reports
15. Bullo F (2006) Notes on multi-agent motion coordination: Models and algorithms. In: Antsaklis PJ, Tabuada P (eds) Network Embedded Sensing and Control, Proceedings of NESC'05 Worskhop. Lecture Notes in Control and Information Sciences, vol 331. Springer, New York, pp 3–8
16. Cao M, Morse AS, Anderson BDO (2008) Reaching a consensus in a dynamically changing environment – convergence rates, measurement delays and asynchronous events. J SIAM Control Optim 47(2):601–623
17. Carli R, Fagnani F, Speranzon A, Zampieri S (2008) Communication constraints in the average consensus problem. Automatica 44(3):671–684
18. Castañón DA, Wu C (2003) Distributed algorithms for dynamic reassignment. In: IEEE conference on decision and control. Maui, HI, December 2003, pp 13–18
19. Chatterjee S, Seneta E (1977) Towards consensus: Some convergence theorems on repeated averaging. J Appl Probab 14(1):89–97
20. Cogburn R (1984) The ergodic theory of Markov chains in random environments. Z Wahrscheinlichkeitstheorie Geb 66(1):109–128
21. Conradt L, Roper TJ (2003) Group decision-making in animals. Nature 421(6919):155–158
22. Cormen TH, Leiserson CE, Rivest RL, Stein C (2001) Introduction to Algorithms, 2nd edn. Press MIT, Cambridge, MA
23. Cortés J (2008) Distributed algorithms for reaching consensus on general functions. Automatica 44(3):726–737
24. Cortés J, Bullo F (2005) Coordination and geometric optimization via distributed dynamical systems. J SIAM Control Optim 44(5):1543–1574
25. Cortés J, Martínez S, Bullo F (2005) Spatially-distributed coverage optimization and control with limited-range interactions. Control ESAIM Optim Calc Var 11:691–719
26. Cortés J, Martínez S, Bullo F (2006) Robust rendezvous for mobile autonomous agents via proximity graphs in arbitrary dimensions. Trans IEEE Autom Control 51(8):1289–1298
27. Cortés J, Martínez S, Karatas T, Bullo F (2004) Coverage control for mobile sensing networks. IEEE Trans Autom Control 20(2):243–255
28. Couzin ID, Krause J, Franks NR, Levin SA (2005) Effective leadership and decision-making in animal groups on the move. Nature 433:513–51
29. Cybenko G (1989) Dynamic load balancing for distributed memory multiprocessors. J Parallel Distrib Comput 7:279–301
30. DeGroot MH (1974) Reaching a consensus. J Am Stat Assoc 69(345):118–121
31. Ferrari-Trecate G, Buffa A, Gati M (2006) Analysis of coordination in multi-agent systems through partial difference equations. IEEE Trans Autom Control 51(6):1058–1063
32. Flocchini P, Prencipe G, Santoro N, Widmayer P (2005) Gathering of asynchronous oblivious robots with limited visibility. Theor Comput Sci 337(1–3):147–168
33. Frazzoli E, Bullo F (2004) Decentralized algorithms for vehicle routing in a stochastic time-varying environment. In: IEEE conference on decision and control. Paradise Island, Bahamas, December 2004. pp 3357–3363
34. Freeman RA, Yang P, Lynch KM (2006) Stability and convergence properties of dynamic average consensus estimators.

In: IEEE conference on decision and control. San Diego, CA, December 2006, pp 398–403
35. Gallager RG (1968) Information Theory and Reliable Communication. Wiley, New York
36. Ganguli A, Cortés J, Bullo F (2005) On rendezvous for visually-guided agents in a nonconvex polygon. In: IEEE conference on decision and control and european control conference, Seville, December 2005, pp 5686–5691
37. Ganguli A, Cortés J, Bullo F (2006) Distributed deployment of asynchronous guards in art galleries. In: American control conference, Minneapolis, June 2006, pp 1416–1421
38. Gazi V, Passino KM (2003) Stability analysis of swarms. IEEE Trans Autom Control 48(4):692–697
39. Gerkey BP, Mataric MJ (2004) A formal analysis and taxonomy of task allocation in multi-robot systems. Int J Robotics Res 23(9):939–954
40. Godwin MF, Spry S, Hedrick JK (2006) Distributed collaboration with limited communication using mission state estimates. In: American control conference. Minneapolis, MN, June 2006, pp 2040–2046
41. Gueron S, Levin SA (1993) Self-organization of front patterns in large wildebeest herds. J Theor Bio 165:541–552
42. Gupta P, Kumar PR (2000) The capacity of wireless networks. Trans IEEE Inf Theory 46(2):388–404
43. Hatano Y, Mesbahi M (2005) Agreement over random networks. Trans IEEE Autom Control 50:1867–1872
44. Horn RA, Johnson CR (1985) Matrix analysis. Cambridge University Press, Cambridge, UK
45. Jadbabaie A, Lin J, Morse AS (2003) Coordination of groups of mobile autonomous agents using nearest neighbor rules. Trans IEEE Autom Control 48(6):988–1001
46. Justh EW, Krishnaprasad PS (2004) Equilibria and steering laws for planar formations. Syst Control Lett 52(1):25–38
47. Kashyap A, Başar T, Srikant R (2007) Quantized consensus. Automatica 43(7):1192–1203
48. Klavins E (2003) Communication complexity of multi-robot systems. In: Boissonnat JD, Burdick JW, Goldberg K, Hutchinson S (eds) Algorithmic foundations of robotics V, vol 7. Springer tracts in advanced robotics. Springer, Berlin
49. Klavins E, Ghrist R, Lipsky D (2006) A grammatical approach to self-organizing robotic systems. Trans IEEE Autom Control 51(6):949–962
50. Klavins E, Murray RM (2004) Distributed algorithms for cooperative control. Pervasive IEEE Comput 3(1):56–65
51. Landau HJ, Odlyzko AM (1981) Bounds for eigenvalues of certain stochastic matrices. Linear Algebra Appl 38:5–15
52. Lin J, Morse AS, Anderson BDO (2004) The multi-agent rendezvous problem: An extended summary. In: Kumar V, Leonard NE, Morse AS (eds) Proceedings of the 2003 block island workshop on cooperative control, ser. Lecture notes in control and information sciences, vol 309. Springer, New York, pp 257–282
53. Lin Z, Broucke M, Francis B (2004) Local control strategies for groups of mobile autonomous agents. IEEE Trans Autom Control 49(4):622–629
54. Lin Z, Francis B, Maggiore M (2005) Necessary and sufficient graphical conditions for formation control of unicycles. Trans IEEE Autom Control 50(1):121–127
55. Lin Z, Francis B, Maggiore M (2007) State agreement for continuous-time coupled nonlinear systems. J SIAM Control Optim 46(1):288–307
56. Lumelsky VJ, Harinarayan KR (1997) Decentralized motion planning for multiple mobile robots: the cocktail party model. Auton Robots 4(1):121–135
57. Lynch NA (1997) Distributed algorithms. Morgan Kaufmann Publishers, San Mateo, CA
58. Marshall JA, Broucke ME, Francis BA (2004) Formations of vehicles in cyclic pursuit. Trans IEEE Autom Control 49(11):1963–1974
59. Martínez S, Bullo F, Cortés J, Frazzoli E (2007) On synchronous robotic networks – Part I: Models, tasks and complexity. Trans IEEE Autom Control 52(12):2199–2213
60. Martinez S, Bullo F, Cortés J, Frazzoli E (2007) On synchronous robotic networks – Part II: Time complexity of rendezvous and deployment algorithms. Trans IEEE Autom Control 52(12):2214–2226
61. Martínez S, Cortés J, Bullo F (2007) Motion coordination with distributed information. Control IEEE Syst Mag 27(4):75–88
62. Miller MB, Bassler BL (2001) Quorum sensing in bacteria. Annu Rev Microbiol 55:165–199
63. Moore BJ, Passino KM (2007) Distributed task assignment for mobile agents. Trans IEEE Autom Control 52(4):749–753
64. Moreau L (2004) Stability of continuous-time distributed consensus algorithms, Preprint. Available at http://xxx.arxiv.org/math.OC/0409010
65. Moreau L (2005) Stability of multiagent systems with time-dependent communication links. Trans IEEE Autom Control 50(2):169–182
66. Okubo A (1986) Dynamical aspects of animal grouping: swarms, schools, flocks and herds. Advances in Biophysics 22:1–94
67. Olfati-Saber R (2006) Flocking for multi-agent dynamic systems: Algorithms and theory. Trans IEEE Autom Control 51(3):401–420
68. Olfati-Saber R, Fax JA, Murray RM (2007) Consensus and cooperation in networked multi-agent systems. Proc IEEE 95(1):215–233
69. Olfati-Saber R, Franco E, Frazzoli E, Shamma JS (2006) Belief consensus and distributed hypothesis testing in sensor networks. In: Antsaklis PJ, Tabuada P (eds) Network embedded sensing and control. Proceedings of NESC'05 worskhop. Lecture notes in control and information sciences, vol 331. Springer, New York, pp 169–182
70. Olfati-Saber R, Murray RM (2004) Consensus problems in networks of agents with switching topology and time-delays. Trans IEEE Autom Control 49(9):1520–1533
71. Olshevsky A, Tsitsiklis JN (2007) Convergence speed in distributed consensus and averaging. J SIAM Control Optim
72. Parrish JK, Viscido SV, Grunbaum D (2002) Self-organized fish schools: an examination of emergent properties. Biol Bull 202:296–305
73. Pavone M, Frazzoli E, Bullo F (2007) Decentralized algorithms for stochastic and dynamic vehicle routing with general target distribution. In: IEEE conference on decision and control, New Orleans, LA, December 2007, pp 4869–4874
74. Peleg D (2000) Distributed computing. A locality-sensitive approach. Monographs on discrete mathematics and applications. SIAM, Philadelphia, PA
75. P Ögren, Fiorelli E, Leonard NE (2004) Cooperative control of mobile sensor networks: Adaptive gradient climbing in a dis-

tributed environment. Trans IEEE Autom Control 49(8):1292–1302
76. Rathinam S, Sengupta R, Darbha S (2007) A resource allocation algorithm for multi-vehicle systems with non holonomic constraints. Trans IEEE Autom Sci Eng 4(1):98–104
77. Ren W, Beard RW (2005) Consensus seeking in multi-agent systems under dynamically changing interaction topologies. Trans IEEE Autom Control 50(5):655–661
78. Ren W, Beard RW, Atkins EM (2007) Information consensus in multivehicle cooperative control: Collective group behavior through local interaction. Control IEEE Syst Mag 27(2):71–82
79. Santoro N (2001) Distributed computations by autonomous mobile robots. In: Pacholski L, Ruzicka P (eds) SOFSEM 2001: Conference on current trends in theory and practice of informatics (Piestany, Slovak Republic). Lecture notes in computer science, vol 2234. Springer, New York, pp 110–115
80. Savkin AV (2004) Coordinated collective motion of groups of autonomous mobile robots: Analysis of Vicsek's model. Trans IEEE Autom Control 49(6):981–982
81. Savla K, Bullo F, Frazzoli E (2007) Traveling Salesperson Problems for a double integrator. Trans IEEE Autom Control, To appear
82. Savla K, Frazzoli E, Bullo F (2008) Traveling Salesperson Problems for the Dubins vehicle. Trans IEEE Autom Control 53(9) To appear
83. Savla K, Notarstefano G, Bullo F (2007) Maintaining limited-range connectivity among second-order agents. J SIAM Control Optim, To appear
84. Schumacher C, Chandler PR, Rasmussen SJ, Walker D (2003) Task allocation for wide area search munitions with variable path length. In: American control conference, Denver, CO, pp 3472–3477
85. Seeley TD, Buhrman SC (1999) Group decision-making in swarms of honey bees. Behav Eco Sociobiol 45:19–31
86. Sepulchre R, Paley D, Leonard NE (2007) Stabilization of planar collective motion: All-to-all communication. Trans IEEE Autom Control 52:811–824
87. Sharma V, Savchenko M, Frazzoli E, Voulgaris P (2005) Time complexity of sensor-based vehicle routing. In: Thrun S, Sukhatme G, Schaal S, Brock O (eds) Robotics: Science and systems. Press MIT, Cambridge, MA, pp 297–304
88. Sharma V, Savchenko M, Frazzoli E, Voulgaris P (2007) Transfer time complexity of conflict-free vehicle routing with no communications. Int J Robotics Res 26(3):255–272
89. Sinclair AR (1977) The african buffalo, study a of resource limitation of population. The University of Chicago Press, Chicago, IL
90. Smith SL, Broucke ME, Francis BA (2005) A hierarchical cyclic pursuit scheme for vehicle networks. Automatica 41(6):1045–1053
91. Smith SL, Bullo F (2007) Monotonic target assignment for robotic networks. Trans IEEE Autom Control, Submitted, To appear
92. Spanos DP, Olfati-Saber R, Murray RM (2005) Approximate distributed Kalman filtering in sensor networks with quantifiable performance. In: Symposium on information processing of sensor networks (IPSN). Los Angeles, CA, April 2005. pp 133–139
93. Stewart KJ, Harcourt AH (1994) Gorillas vocalizations during rest periods – signals of impending departure. Behaviour 130:29–40
94. Suzuki I, Yamashita M (1999) Distributed anonymous mobile robots: Formation of geometric patterns. J SIAM Comput 28(4):1347–1363
95. Tahbaz-Salehi A, Jadbabaie A (2008) Consensus over random networks. Trans IEEE Autom Control 53(3):791–795
96. Tanner HG, Jadbabaie A, Pappas GJ (2007) Flocking in fixed and switching networks. Trans IEEE Autom Control 52(5):863–868
97. Tsitsiklis JN (1984) Problems in decentralized decision making and computation. Dissertation, Laboratory for Information and Decision Systems, MIT, Nov. Technical Report LIDS-TH-1424
98. Tsitsiklis JN, Bertsekas DP, Athans M (1986) Distributed asynchronous deterministic and stochastic gradient optimization algorithms. Trans IEEE Autom Control 31(9):803–812
99. Uny Y Cao, Fukunaga AS, Kahng A (1997) Cooperative mobile robotics: Antecedents and directions. Auton Robots 4(1):7–27
100. Wu CW (2006) Synchronization and convergence of linear dynamics in random directed networks. Trans IEEE Autom Control 51(7):1207–1210
101. Xiao L, Boyd S, Lall S (2005) A scheme for asynchronous distributed sensor fusion based on average consensus. In: Symposium on information processing of sensor networks (IPSN). Los Angeles, CA, April 2005. pp 63–70
102. Zavlanos M, Pappas G (2007) Dynamic assignment in distributed motion planning with local information. In: American control conference. New York, July 2007, pp 1173–1178
103. Zavlanos MM, Pappas GJ (2005) Controlling connectivity of dynamic graphs. In: IEEE conference on decision and control and european control conference. Seville, Spain, December 2005, pp 6388–6393

Rough and Rough-Fuzzy Sets in Design of Information Systems

Theresa Beaubouef[1], Frederick Petry[2]
[1] Southeastern Louisiana University, Hammond, USA
[2] Naval Research Lab, Stennis Space Center, Mississippi, USA

Article Outline

Glossary

Rough sets Rough set theory is a technique for dealing with uncertainty and for identifying cause-effect relationships in databases. It is based on a partitioning of some domain into equivalence classes and the defining of lower and upper approximation regions based on this partitioning to denote certain and possible inclusion in the rough set.

Fuzzy sets Fuzzy set theory is another technique for dealing with uncertainty. It is based on the concept of measuring the degree of inclusion in a set through the use of a membership value. Where elements can either belong or not belong to a regular set, with fuzzy sets elements can belong to the set to a certain degree with zero indicating not an element, one indicating complete membership, and values between zero and one indicating partial or uncertain membership in the set.

Information theory Information theory involves the study of measuring the information content of a signal. In databases information theoretic measures can be used to measure the information content of data. Entropy is one such measure.

Database A collection of data and the application programs that make use of this data for some enterprise is a database.

Information system An information system is a database enhanced with additional tools that can be used by management for planning and decision making.

Data mining Data mining involves the discovery of patterns or rules in a set of data. These patterns generate some knowledge and information from the raw data that can be used for making decisions. There are many approaches to data mining, and uncertainty management techniques play a vital role in knowledge discovery.

Definition of the Subject

Databases and information systems are ubiquitous in this age of information and technology. Computers have revolutionized the way data can be manipulated and stored, allowing for very large databases with sophisticated capabilities. With so much money and manpower invested in the design and daily use of these systems, it is imperative that they be as correct, secure, and as adaptable to the changing needs of the enterprise as possible. Therefore it is important to understand the design and implementation of such systems and to be able to utilize all their capabilities.

Scientists and business executives alike know the value of information. The challenge has been to produce relevant information for an ever changing uncertain world from data and facts stored on computers and archival devices. These data are considered to be exact, certain, factual values. The real world, however, is uncertain, inexact, and fraught with errors. It is a challenge, then, to extract useful and relevant information from ordinary databases. Uncertainty management techniques such as rough and fuzzy sets can help.

Introduction

Databases are recognized for their ability to store and update data in an efficient manner, providing reliability and the elimination of data redundancy. The relational database model, in particular, has well-established mechanisms built into the model for properly designing the database and maintaining data integrity and consistency. Data alone, however, are only facts. What is needed is information. Knowledge discovery attempts to derive information from the pure facts, discovering high level regularities in the data. It is defined as the nontrivial extraction of implicit, previously unknown, and potentially useful information from data [24,27].

An innovative technique in the field of uncertainty and knowledge discovery is based on rough sets. Rough set theory, introduced and further developed mathematically by [39], provides a framework for the representation of uncertainty. It has been used in various applications such as the rough querying of crisp data [6], uncertainty management in databases [11], the mining of spatial data [8], and improved information retrieval [48]. These techniques may readily be extended for use with object-oriented, spatial, and other complex databases, and may be integrated with additional data mining techniques for a comprehensive knowledge discovery approach.

Rough Sets and Rough-Fuzzy Sets

Rough set theory, introduced by Pawlak [38], is a technique for dealing with uncertainty and for identifying cause-effect relationships in databases. An extensive theory for rough sets and their properties has been developed and they have become a well established approach for the management of uncertainty in a variety of applications. Rough sets involve the following:

U is the *universe*, which cannot be empty,

R is the *indiscernibility relation*, or equivalence relation,

$A = (U, R)$ an ordered pair, is called an *approximation space*,

$[x]_R$ denotes the equivalence class of R containing x, for any element x of U,

Elementary sets in A – the equivalence classes of R,

Definable set in A – any finite union of elementary sets in A.

Given an approximation space defined on some universe U that has an equivalence relation R imposed upon it, U is partitioned into equivalence classes called elementary sets that may be used to define other sets in A. A rough set X, where $X \subseteq U$, can be defined in terms of the definable sets in A by the following:

Lower approximation of X in A is the set

$\underline{R}X = \{x \in U \mid [x]_R \subseteq X\}$

Upper approximation of X in A is the set

$\overline{R}X = \{x \in U \mid [x]_R \cap X \neq \emptyset\}$.

$\mathrm{POS}_R(X) = \underline{R}X$ denotes the R-positive region of X, or those elements which certainly belong to the rough set. The R-negative region of X, $\mathrm{NEG}_R(X) = U - \overline{R}X$, contains elements which do not belong to the rough set, and the boundary or R-borderline region of X, $\mathrm{BN}_R(X) = \overline{R}X - \underline{R}X$, contains those elements which may or may not belong to the set. X is R-definable if and only if $\underline{R}X = \overline{R}X$. Otherwise, $\underline{R}X \neq \overline{R}X$ and X is rough with respect to R. A *rough set in A* is the group of subsets of U with the same upper and lower approximations.

Fuzzy set theory [53] is another approach for managing uncertainty. It has been around for a few years longer than rough sets, and also has well developed theory, properties, and applications. Applications involving fuzzy logic are diverse and plentiful, ranging from fuzzy control systems in industry to fuzzy logic in databases.

Because there are advantages to both fuzzy set and rough set theories, several researchers have studied various ways of combining the two theories [22,30,36] Others have investigated the interrelations between the two theories [19,40,52]. A similar approach to our rough-fuzzy set is the fuzzy rough set in [31]. That approach is more in the spirit of functional analysis, however. Fuzzy sets and rough sets are not equivalent, but complementary.

It has been shown in [52] that rough sets can be expressed by a fuzzy membership function $\mu \rightarrow \{0, 0.5, 1\}$ to represent the negative, boundary, and positive regions. In this model, all elements of the lower approximation, or positive region, have a membership value of one. Those elements of the boundary region are assigned a membership value of 0.5. Elements not belonging to the rough set have a membership value of zero. Rough set definitions of union and intersection can be modified so that the fuzzy model satisfies all the properties of rough sets [7]. This allows a rough set to be expressed as a fuzzy set.

We integrate fuzziness into the rough set model in order to quantify levels of roughness in boundary region areas through the use of fuzzy membership values. Therefore, we do not require membership values of elements of the boundary region to equal 0.5, but allow them to range from zero to one, noninclusive. Additionally, the union and intersection operators for fuzzy rough sets are comparable to those for ordinary fuzzy sets, where MIN and MAX are used to obtain membership values of redundant elements.

Let U be a *universe*, X a rough set in U.

Definition A *fuzzy rough set* Y in U is a membership function $\mu_Y(x)$ which associates a grade of membership from the interval [0,1] with every element of U where

$$\mu_Y(\underline{R}X) = 1, \quad \mu_Y(U - \overline{R}X) = 0, \text{ and } 0 < \mu_Y(\overline{R}X - \underline{R}X) < 1.$$

Definition The *union* of two fuzzy rough sets A and B is a fuzzy rough set C where

$$C = \{x | x \in A \text{ OR } x \in B\}, \text{ where } \mu_C(x) = \mathrm{MAX}[\mu_A(x), \mu_B(x)].$$

Definition The *intersection* of two fuzzy rough sets A and B is a fuzzy rough set C where

$$C = \{x | x \in A \text{ AND } x \in B\}, \text{ where } \mu_C(x) = \mathrm{MIN}[\mu_A(x), \mu_B(x)].$$

Rough Relational Database

The rough relational database model [13] is an extension of the standard relational database model of Codd [23]. It captures all the essential features of rough sets theory including indiscernibility of elements denoted by equivalence classes and lower and upper approximation regions for defining sets which are indefinable in terms of the indiscernibility.

Every attribute domain is partitioned by some equivalence relation designated by the database designer or user. Within each domain, those values that are considered indiscernible belong to an equivalence class. This information is used by the query mechanism to retrieve information based on equivalence with the class to which the value belongs rather than equality, resulting in less critical wording of queries.

Recall is also improved in the rough relational database because rough relations provide *possible* matches to the query in addition to the *certain* matches which are obtained in the standard relational database. This is accomplished by using set containment in addition to equality of

attributes in the calculation of lower and upper approximation regions of the query result.

The rough relational database has several features in common with the ordinary relational database. Both models represent data as a collection of *relations* containing *tuples*. These relations are sets. The tuples of a relation are its elements, and like elements of sets in general, are unordered and nonduplicated. A tuple $\mathbf{t}_i$ takes the form $(d_{i1}, d_{i2}, \ldots, d_{im})$, where d_{ij} is a *domain value* of a particular *domain set* D_j. In the ordinary relational database, $d_{ij} \in D_j$. In the rough database, however, as in other non-first normal form extensions to the relational model [34,44], $d_{ij} \subseteq D_j$, and although it is not required that d_{ij} be a singleton, $d_{ij} \neq \emptyset$. Let $P(D_i)$ denote the powerset $(D_i) - \emptyset$.

Definition A *rough relation* R is a subset of the set cross product $P(D_1) \times P(D_2) \times \cdots \times P(D_m)$.

A rough tuple $\mathbf{t}$ is any member of R, which implies that it is also a member of $P(D_1) \times P(D_2) \times \cdots \times P(D_\mathrm{m})$. If $\mathbf{t}_\mathrm{i}$ is some arbitrary tuple, then $\mathbf{t}_\mathrm{i} = (d_{i1}, d_{i2}, \ldots, d_{im})$ where $d_{ij} \subseteq D_\mathrm{j}$. A tuple in this model differs from that of ordinary databases in that the tuple components may be sets of domain values rather than single values. The set braces are omitted from singletons for notational simplicity.

Let $[d_{xy}]$ denote the equivalence class to which d_{xy} belongs. When d_{xy} is a set of values, the equivalence class is formed by taking the union of equivalence classes of members of the set; if $d_{xy} = \{c_1, c_2, \ldots, c_n\}$, then $[d_{xy}] = [c_1] \cup [c_2] \cup \cdots \cup [c_n]$.

Definition Tuples $\mathbf{t}_i = (d_{i1}, d_{i2}, \ldots, d_{im})$ and $\mathbf{t}_k = (d_{k1}, d_{k2}, \ldots, d_{km})$ are *redundant* if $[d_{ij}] = [d_{kj}]$ for all $j = 1, \ldots, m$.

In the rough relational database, redundant tuples are removed in the merging process since duplicates are not allowed in sets, the structure upon which the relational model is based.

There are two basic types of relational operators. The first type arises from the fact that relations are considered sets of tuples. Therefore, operations which can be applied to sets also apply to relations. The most useful of these for database purposes are *set difference*, *union*, and *intersection*. Operators which do not come from set theory, but which are useful for retrieval of relational data are *select*, *project*, and *join*.

In the rough relational database, relations are rough sets as opposed to ordinary sets. Therefore, new rough operators ($-$, $\cup$, $\cap$, σ, π, $\bowtie$), which are comparable to the standard relational operators, must be developed for the rough relational database. Moreover, a mechanism must exist within the database to mark tuples of a rough relation as belonging to the lower or upper approximation of that rough relation. Properties of the rough relational operators can be found in [13].

Information Theory

In communication theory, Shannon [45] introduced the concept of entropy which was used to characterize the information content of signals. Since then, variations of these information theoretic measures have been successfully applied to applications in many diverse fields. In particular, the representation of uncertain information by entropy measures has been applied to all areas of databases, including fuzzy database querying [17], data allocation [25], classification in rule-based systems [42], and measuring uncertainty in rough and fuzzy rough relational databases [12].

In fuzzy set theory the representation of uncertain information measures has been extensively studied [14,20,29]. So this paper relates the concepts of information theory to rough sets and compares these information theoretic measures to established rough set metrics of uncertainty. The measures are then applied to the rough relational database model [13]. Information content of both stored relational schemas and rough relations are expressed as types of rough entropy.

Rough set theory [38] inherently models two types of uncertainty. The first type of uncertainty arises from the indiscernibility relation that is imposed on the universe, partitioning all values into a finite set of equivalence classes. If every equivalence class contains only one value, then there is no loss of information caused by the partitioning. In any coarser partitioning, however, there are fewer classes, and each class will contain a larger number of members. Our knowledge, or information, about a particular value decreases as the granularity of the partitioning becomes coarser.

Uncertainty is also modeled through the approximation regions of rough sets where elements of the lower approximation region have total participation in the rough set and those of the upper approximation region have uncertain participation in the rough set. Equivalently, the lower approximation is the *certain* region and the boundary area of the upper approximation region is the *possible* region.

Pawlak [41] discusses two numerical characterizations of imprecision of a rough set X: *accuracy* and *roughness*. Accuracy, which is simply the ratio of the number of elements in the lower approximation of X, $\underline{R}X$, to the number of elements in the upper approximation of the rough

set X, $\overline{R}X$, measures the degree of completeness of knowledge about the given rough set X. It is defined as a ratio of the two set cardinalities as follows:

$$\alpha_R(X) = \text{card}(\underline{R}X)/\text{card}(\overline{R}X), \text{ where } 0 \leq \alpha_R(X) \leq 1.$$

The second measure, roughness, represents the degree of incompleteness of knowledge about the rough set. It is calculated by subtracting the accuracy from 1: $\rho_R(X) = 1 - \alpha_R(X)$.

These measures require knowledge of the number of elements in each of the approximation regions and are good metrics for uncertainty as it arises from the boundary region, implicitly taking into account equivalence classes as they belong wholly or partially to the set. However, accuracy and roughness measures do not necessarily provide us with information on the uncertainty related to the granularity of the indiscernibility relation for those values that are totally included in the lower approximation region. For example:

Let the rough set X be defined as follows: $X = \{A11, A12, A21, A22, B11, C1\}$ with lower and upper approximation regions defined as

$$\underline{R}X = \{A11, A12, A21, A22\} \text{ and}$$
$$\overline{R}X = \{A11, A12, A21, A22, B11, B12, B13, C1, C2\}.$$

These approximation regions may result from one of several partitionings. Consider, for example, the following indiscernibility relations:

$$A_1 = \{[A11, A12, A21, A22], [B11, B12, B13], [C1, C2]\},$$
$$A_2 = \{[A11, A12], [A21, A22], [B11, B12, B13], [C1, C2]\},$$
$$A_3 = \{[A11], [A12], [A21], [A22], [B11, B12, B13], [C1, C2]\}.$$

All three of the above partitionings result in the same upper and lower approximation regions for the given set X, and hence the same accuracy measure ($4/9 = .444$) since only those classes belonging to the lower approximation region were re-partitioned. It is obvious, however, that there is more uncertainty in A_1 than in A_2, and more uncertainty in A_2 than in A_3. Therefore, a more comprehensive measure of uncertainty is needed.

We derive such a measure from techniques used for measuring entropy in classical information theory. Countless variations of the classical entropy have been developed, each tailored for a particular application domain or for measuring a particular type of uncertainty. Our rough entropy is defined such that we may apply it to rough databases. We define the entropy of a rough set X as follows:

Definition The *rough entropy* $E_r(X)$ of a rough set X is calculated by

$$E_r(X) = -(\rho_R(X))[\Sigma Q_i \log(P_i)] \quad \text{for } i = 1, \ldots n \text{ equivalence classes.}$$

The term $\rho_R(X)$ denotes the roughness of the set X. The second term is the summation of the probabilities for each equivalence class belonging either wholly or in part to the rough set X. There is no ordering associated with individual class members. Therefore the probability of any one value of the class being named is the reciprocal of the number of elements in the class. If c_i is the cardinality of, or number of elements in, equivalence class i and all members of a given equivalence class are equal, $P_i = 1/c_i$ represents the probability of one of the values in class i. Q_i denotes the probability of equivalence class i within the universe. Q_i is computed by taking the number of elements in class i and dividing by the total number of elements in all equivalence classes combined. The entropy of the sample rough set X, $E_r(X)$, is given below for each of the possible indiscernibility relations A_1, A_2, and A_3.

$$\text{Using } A_1 : -(5/9)[(4/9)\log(1/4) + (3/9)\log(1/3) + (2/9)\log(1/2)] = .274$$
$$\text{Using } A_2 : -(5/9)[(2/9)\log(1/2) + (2/9)\log(1/2) + (3/9)\log(1/3) + (2/9)\log(1/2)] = .20$$
$$\text{Using } A_3 : -(5/9)[(1/9)\log(1) + (1/9)\log(1) + (1/9)\log(1) + (1/9)\log(1) + (3/9)\log(1/3) + (2/9)\log(1/2)] = .048$$

From the above calculations it is clear that although each of the partitionings results in identical roughness measures, the entropy decreases as the classes become smaller through finer partitionings.

Entropy and the Rough Relational Database

The basic concepts of rough sets and their information-theoretic measures carries over to the rough relational database model [13]. Recall that in the rough relational database all domains are partitioned into equivalence classes and relations are not restricted to first normal form. We therefore have a type of rough set for each attribute of a relation. This results in a rough relation, since any tuple

having a value for an attribute that belongs to the boundary region of its domain is a tuple belonging to the boundary region of the rough relation.

There are two things to consider when measuring uncertainty in databases: uncertainty or entropy of a rough relation that exists in a database at some given time and the entropy of a relation schema for an existing relation or query result. We must consider both since the approximation regions only come about by set values for attributes in given tuples. Without the extension of a database containing actual values, we only know about indiscernibility of attributes. We cannot consider the approximation regions.

We define the entropy for a rough relation schema as follows:

Definition The *rough schema entropy* for a rough relation schema S is

$$E_s(S) = -\Sigma_j[\Sigma Q_i \log(P_i)] \text{ for } i = 1, \ldots n; j = 1, \ldots, m$$

where there are n equivalence classes of domain j, and m attributes in the schema $R(A_1, A_2, \ldots, A_m)$.

This is similar to the definition of entropy for rough sets without factoring in roughness since there are no elements in the boundary region (lower approximation = upper approximation). However, because a relation is a cross product among the domains, we must take the sum of all these entropies to obtain the entropy of the schema. The schema entropy provides a measure of the uncertainty inherent in the definition of the rough relation schema taking into account the partitioning of the domains on which the attributes of the schema are defined.

We extend the schema entropy $E_s(S)$ to define the entropy of an actual rough relation instance $E_R(R)$ of some database D by multiplying each term in the product by the roughness of the rough set of values for the domain of that given attribute.

Definition The *rough relation entropy* of a particular extension of a schema is

$$E_R(R) = -\Sigma_j D\rho_j(R)[\Sigma DQ_i \log(DP_i)] \quad \text{for } i = 1, \ldots n; j = 1, \ldots, m$$

where $D\rho_j(R)$ represents a type of database roughness for the rough set of values of the domain for attribute j of the relation, m is the number of attributes in the database relation, and n is the number of equivalence classes for a given domain for the database.

We obtain the $D\rho_j(R)$ values by letting the non-singleton domain values represent elements of the boundary region, computing the original rough set accuracy and subtracting it from one to obtain the roughness. DQ_i is the probability of a tuple in the database relation having a value from class i, and DP_i is the probability of a value for class i occurring in the database relation out of all the values which are given.

Information theoretic measures again prove to be a useful metric for quantifying information content. In rough sets and the rough relational database, this is especially useful since in ordinary rough sets Pawlak's measure of roughness does not seem to capture the information content as precisely as our rough entropy measure.

In rough relational databases, knowledge about entropy can either guide the database user toward less uncertain data or act as a measure of the uncertainty of a data set or relation. As rough relations become larger in terms of the number of tuples or attributes, the automatic calculation of some measure of entropy becomes a necessity. Our rough relation entropy measure fulfills this need.

Rough Fuzzy Relational Database

The fuzzy rough relational database, as in the ordinary relational database, represents data as a collection of *relations* containing *tuples*. Because a relation is considered a set having the tuples as its members, the tuples are unordered. In addition, there can be no duplicate tuples in a relation. A tuple $\mathbf{t}_i$ takes the form $(d_{i1}, d_{i2}, \ldots, d_{im}, d_{i\mu})$, where d_{ij} is a *domain value* of a particular *domain set* D_j and $d_{i\mu} \in D_\mu$, where D_μ is the interval [0,1], the domain for fuzzy membership values. In the ordinary relational database, $d_{ij} \in D_j$. In the fuzzy rough relational database, except for the fuzzy membership value, however, $d_{ij} \subseteq D_j$, and although d_{ij} is not restricted to be a singleton, $d_{ij} \neq \emptyset$. Let $P(D_i)$ denote any non-null member of the powerset of D_i.

Definition A *fuzzy rough relation* R is a subset of the set cross product $P(D_1) \times P(D_2) \times \cdots \times P(D_m) \times D_\mu$.

For a specific relation, R, membership is determined semantically. Given that D_1 is the set of names of nuclear/chemical plants, D_2 is the set of locations, and assuming that RIVERB is the only nuclear power plant that is located in VENTRESS,

(RIVERB, VENTRESS, 1)

(RIVERB, OSCAR, .7)

(RIVERB, ADDIS, 1)

(CHEMO, VENTRESS, .3)

are all elements of $P(D_1) \times P(D_2) \times D_\mu$. However, only the element (RIVERB, VENTRESS, 1) of those

listed above is a member of the relation R(PLANT, LOCATION, μ), which associates each plant with the town or community in which it is located. A *fuzzy rough tuple* $\mathbf{t}$ is any member of R. If $\mathbf{t}_i$ is some arbitrary tuple, then $\mathbf{t}_i = (d_{i1}, d_{i2}, \ldots, d_{im}, d_{i\mu})$ where $d_{ij} \subseteq D_j$ and $d_{i\mu} \in D_\mu$.

Definition An *interpretation* $\alpha = (a_1, a_2, \ldots, a_m, a_\mu)$ of a fuzzy rough tuple $\mathbf{t}_i = (d_{i1}, d_{i2}, \ldots, d_{im}, d_{i\mu})$ is any value assignment such that $a_j \in d_{ij}$ for all j.

The interpretation space is the cross product $D_1 \times D_2 \times \cdots \times D_m \times D_\mu$, but is limited for a given relation R to the set of those tuples which are valid according to the underlying semantics of R. In an ordinary relational database, because domain values are atomic, there is only one possible interpretation for each tuple $\mathbf{t}_i$. Moreover, the interpretation of $\mathbf{t}_i$ is equivalent to the tuple $\mathbf{t}_i$. In the fuzzy rough relational database, this is not always the case.

Let $[d_{xy}]$ denote the equivalence class to which d_{xy} belongs. When d_{xy} is a set of values, the equivalence class is formed by taking the union of equivalence classes of members of the set; if $d_{xy} = \{c_1, c_2, \ldots, c_n\}$, then $[d_{xy}] = [c_1] \cup [c_2] \cup \cdots \cup [c_n]$.

Definition Tuples $\mathbf{t}_i = (d_{i1}, d_{i2}, \ldots, d_{in}, d_{i\mu})$ and $\mathbf{t}_k = (d_{k1}, d_{k2}, \ldots, d_{kn}, d_{k\mu})$ are *redundant* if $[d_{ij}] = [d_{kj}]$ for all $j = 1, \ldots, n$.

If a relation contains only those tuples of a lower approximation, i. e., those tuples having a μ value equal to one, the interpretation α of a tuple is unique. This follows immediately from the definition of redundancy. In fuzzy rough relations, there are no redundant tuples. The merging process used in relational database operations removes duplicate tuples since duplicates are not allowed in sets, the structure upon which the relational model is based.

Tuples may be redundant in all values except μ. As in the union of fuzzy rough sets where the maximum membership value of an element is retained, it is the convention of the fuzzy rough relational database to retain the tuple having the higher μ value when removing redundant tuples during merging. If we are supplied with identical data from two sources, one certain and the other uncertain, we would want to retain the data that is certain, avoiding loss of information.

Recall that the rough relational database is in non-first normal form; there are some attribute values that are sets. Another definition, which will be used for upper approximation tuples, is necessary for some of the alternate definitions of operators to be presented. This definition captures redundancy between elements of attribute values that are sets:

Definition Two sub-tuples $X = (d_{x1}, d_{x2}, \ldots, d_{xm})$ and $Y = (d_{y1}, d_{y2}, \ldots, d_{ym})$ are *roughly-redundant*, $\approx_{\mathbf{R}}$, if for some $[p] \subseteq [d_{xj}]$ and $[q] \subseteq [d_{yj}]$, $[p] = [q]$ for all $j = 1, \ldots, m$.

In order for any database to be useful, a mechanism for operating on the basic elements and retrieving specified data must be provided. The concepts of redundancy and merging play a key role in the operations defined.

We must first design our database using some type of semantic model. We use a variation of the entity-relationship diagram that we call a fuzzy-rough E-R diagram. This diagram is similar to the standard E-R diagram in that entity types are depicted in rectangles, relationships with diamonds, and attributes with ovals. However, in the fuzzy-rough model, it is understood that membership values exist for all instances of entity types and relationships. Attributes which allow values where we want to be able to define equivalences are denoted with an asterisk (*) above the oval. These values are defined in the indiscernibility relation, which is not actually part of the database design, but inherent in the fuzzy-rough model.

Our fuzzy-rough E-R model [7] is similar to the second and third levels of fuzziness defined by Zvieli and Chen [54]. However, in our model, all entity and relationship occurrences (second level) are of the fuzzy type so we do not mark an 'f' beside each one. Zvieli and Chen's third level considers attributes that may be fuzzy. They use triangles instead of ovals to represent these attributes. We do not introduce fuzziness at the attribute level of our model in this paper, only roughness, or indiscernibility, and denote those attribute with the '*'. From the Fuzzy-Rough E-R diagram, we design the structure of the fuzzy rough relational database. If we have a priori information about the types of queries that will be involved, we can make intelligent choices that will maximize computer resources.

We next formally define the fuzzy rough relational database operators and discuss issues relating to the real-world problems of data representation and modeling. We may view indiscernibility as being modeled through the use of the indiscernibility relation, imprecision through the use of non-first normal form constructs, and degree of uncertainty and fuzziness through the use of tuple membership values, which are given as the value for the μ attribute in every fuzzy rough relation.

Fuzzy Rough Relational Operators

In [13], we defined several operators for the rough relational algebra. We now define similar operators for the fuzzy rough relational database as in [5]. Recall that for all of these operators the indiscernibility relation is used

for equivalence of attribute values rather than equality of values.

Difference The fuzzy rough relational difference operator is very much like the ordinary difference operator in relational databases and in sets in general. It is a binary operator that returns those elements of the first argument that are not contained in the second argument.

In the fuzzy rough relational database, the difference operator is applied to two fuzzy rough relations and, as in the rough relational database, indiscernibility, rather than equality of attribute values, is used in the elimination of redundant tuples. Hence, the difference operator is somewhat more complex. Let X and Y be two union compatible fuzzy rough relations.

Definition The *fuzzy rough difference*, $\boldsymbol{X}$ - $\boldsymbol{Y}$, between X and Y is a fuzzy rough relation T where

$$T = \{t(d_1, \ldots, d_n, \mu_i) \in X | t(d_1, \ldots, d_n, \mu_i) \notin Y\} \cup \{t(d_1, \ldots, d_n, \mu_i) \in X | t(d_1, \ldots, d_n, \mu_j) \in Y \text{ and } \mu_i > \mu_j\}.$$

The resulting fuzzy rough relation contains all those tuples which are in the lower approximation of X, but not redundant with a tuple in the lower approximation of Y. It also contains those tuples belonging to upper approximation regions of both X and Y, but which have a higher μ value in X than in Y. For example, let X contain the tuple **(MODERN, 1)** and Y contain the tuple **(MODERN, .02)**. It would not be desirable to subtract out certain information with possible information, so X - Y yields **(MODERN, 1)**.

Union Because relations in databases are considered as sets, the union operator can be applied to any two union-compatible relations to result in a third relation which has as its tuples all the tuples contained in either or both of the two original relations. The union operator can be extended to apply to fuzzy rough relations. Let X and Y be two union compatible fuzzy rough relations.

Definition The *fuzzy rough union* of X and Y, $\mathbf{X} \cup \mathbf{Y}$ is a fuzzy rough relation T where

$$T = \{t | t \in X \text{ OR } t \in Y\} \quad \text{and} \quad \mu_T(t) = \text{MAX}[\mu_X(t), \mu_Y(t)].$$

The resulting relation T contains all tuples in either X or Y or both, merged together and having redundant tuples removed. If X contains a tuple that is redundant with a tuple in Y except for the μ value, the merging process will retain only that tuple with the higher μ value.

Intersection The fuzzy rough intersection, another binary operator on fuzzy rough relations, can be defined similarly.

Definition The *fuzzy rough intersection* of X and Y, $\mathbf{X} \cap \mathbf{Y}$ is a fuzzy rough relation T where

$$T = \{t | t \in X \text{ AND } t \in Y\} \quad \text{and} \quad \mu_T(t) = \text{MIN}[\mu_X(t), \mu_Y(t)].$$

In intersection, the MIN operator is used in the merging of equivalent tuples having different μ values and the result contains all tuples that are members of both of the original fuzzy rough relations.

Definition The *fuzzy rough intersection* of X and Y, $\mathbf{X} \cap_{\mathbf{A}} \mathbf{Y}$ is a fuzzy rough relation T where

$$T = \{t | t \in X, \text{ and } \exists s \in Y | t \approx_R s\} \cup \{s | s \in Y, \text{ and } \exists t \in X | s \approx_R t\} \text{ and } \mu_T(t) = \text{MIN}[\mu_X(t), \mu_Y(t)].$$

Select The select operator for the fuzzy rough relational database model, $\boldsymbol{\sigma}$, is a unary operator which takes a fuzzy rough relation X as its argument and returns a fuzzy rough relation containing a subset of the tuples of X, selected on the basis of values for a specified attribute. The operation $\boldsymbol{\sigma}_{\mathbf{A}=\mathbf{a}}(X)$, for example, returns those tuples in X where attribute A is equivalent to the class $[\mathbf{a}]$. In general, select returns a subset of the tuples that match some selection criteria.

Let R be a relation schema, X a fuzzy rough relation on that schema, A an attribute in R, $\mathbf{a} = \{a_i\}$ and $\mathbf{b} = \{b_j\}$, where $a_i, b_j \in \text{dom}(A)$, and $\cup_x$ is interpreted as "the union over all x".

Definition The *fuzzy rough selection*, $\boldsymbol{\sigma}_{\mathbf{A}=\mathbf{a}}(X)$, of tuples from X is a fuzzy rough relation Y having the same schema as X and where

$$Y = \{t \in X | \cup_i [a_i] \subseteq \cup_j [b_j]\},$$

and $a_i \in \mathbf{a}$, $b_j \in t(A)$, and where membership values for tuples are calculated by multiplying the original membership value by

$$\text{card}(\mathbf{a})/\text{card}(\mathbf{b})$$

where card(x) returns the cardinality, or number of elements, in x.

Assume we want to retrieve those elements where **CITY = "ADDIS"** from the following fuzzy rough tuples:

(**ADDIS** **1**)
({**ADDIS, LOTTIE, BRUSLY**} .**7**)
(**OSCAR** **1**)
({**ADDIS, JACKSON**} .**9**).

The result of the selection is the following:

(**ADDIS** **1**)
({**ADDIS, LOTTIE, BRUSLY**} .**23**)
({**ADDIS, JACKSON**} .**45**),

where the μ for the second tuple is the product of the original membership value .7 and 1/3.

Project Project is a unary fuzzy rough relational operator. It returns a relation that contains a subset of the columns of the original relation. Let X be a fuzzy rough relation with schema A, and let B be a subset of A. The fuzzy rough projection of X onto B is a fuzzy rough relation Y obtained by omitting the columns of X which correspond to attributes in A - B, and removing redundant tuples. Recall the definition of redundancy accounts for indiscernibility, which is central to the rough sets theory and that higher μ values have priority over lower ones.

Definition The *fuzzy rough projection* of X onto B, $\boldsymbol{\pi}_{\mathbf{B}}(\mathbf{X})$, is a fuzzy rough relation Y with schema $Y(B)$ where

$$Y(B) = \{t(B) | t \in X\}.$$

Join Join is a binary operator that takes related tuples from two relations and combines them into single tuples of the resulting relation. It uses common attributes to combine the two relations into one, usually larger, relation. Let $X(A_1, A_2, \ldots, A_m)$ and $Y(B_1, B_2, \ldots, B_n)$ be fuzzy rough relations with **m** and **n** attributes, respectively, and $AB = C$, the schema of the resulting fuzzy rough relation T.

Definition The *fuzzy rough join*, $\mathbf{X} \bowtie_{\text{<JOIN CONDITION>}} \mathbf{Y}$, of two relations X and Y, is a relation $T(C_1, C_2, \ldots, C_{m+n})$ where

$$T = \{t | \exists t_X \in X, t_Y \in Y \text{ for } t_X = t(A), t_Y = t(B)\},$$

and where

$$t_X(A \cap B) = t_Y(A \cap B), \mu = 1 \tag{1}$$

$$t_X(A \cap B) \subseteq t_Y(A \cap B) \text{ or } t_Y(A \cap B) \subseteq t_X(A \cap B), \quad \mu = \text{MIN}(\mu_X, \mu_Y) \tag{2}$$

<JOIN CONDITION> is a conjunction of one or more conditions of the form **A** = **B**.

Only those tuples which resulted from the "joining" of tuples that were both in lower approximations in the original relations belong to the lower approximation of the resulting fuzzy rough relation. All other "joined" tuples belong to the upper approximation only (the boundary region), and have membership values less than one. The fuzzy membership value of the resultant tuple is simply calculated as in [18] by taking the minimum of the membership values of the original tuples. Taking the minimum value also follows the logic of [37], where in joins of tuples with different levels of information uncertainty, the resultant tuple can have no greater certainty than that of its least certain component.

Functional Dependencies

A functional dependency can be defined as in [23] through the use of a universal database relation concept. Let $R = \{A_1, A_2, \ldots, A_n\}$ be a universal relation schema describing a database having n attributes. Let X and Y be subsets of R. A functional dependency between the attributes of X and Y is denoted by $X \rightarrow Y$. This dependency specifies the constraint that for any two tuples of an instance r of R, if they agree on the X attribute(s) they must agree on their Y attributes(s): if $t_1[X] = t_2[X]$, then it must be true that $t_1[Y] = t_2[Y]$. Tuples that violate the constraint cannot be inserted into the database. The rough functional dependency is based on the rough relational database model. The classical notion of functional dependency for relational databases does not naturally apply to the rough relational database, since all the "roughness" would be lost. In the rough querying of crisp data [6], however, the data is stored in the standard relational model having ordinary functional dependencies imposed upon it and rough relations result only from querying; they are not a part of the database design in which the designer imposes constraint upon relation schemas. Rough functional dependencies for the rough relational database model are defined as follows [9]:

Definition A *rough functional dependency*, $X \rightarrow Y$, for a relation schema R exists if for all instances $T(R)$,

(1) for any two tuples t, $t' \in \underline{R}T$,

$$redundant(t(X), t'(X)) \Rightarrow redundant(t(Y), t\prime(Y))$$

(2) for any two tuples $s, s' \in \overline{R}T$,

$$roughly\text{-}redundant(s(X), s'(X)) \Rightarrow roughly\text{-}redundant(s(Y), s\prime(Y)).$$

Y is roughly functional dependent on X, or X roughly functionally determines Y, whenever the above definition holds. This implies that constraints can be imposed on a rough relational database schema in a rough manner that will aid in integrity maintenance and the reduction of update anomalies without limiting the expressiveness of the inherent rough set concepts.

It is obvious that the classical functional dependency for the standard relational database is a special case of the rough functional dependency; indiscernibility reduces to simple equality and part (2) of the definition is unused since all tuples in relations in the standard relational model belong to the lower approximation region of a similar rough model.

The first part of the definition of rough functional dependency compares with that of fuzzy functional dependencies discussed in [46], where adherence to Armstrong's axioms was proven. The results apply directly in the case of rough functional dependencies when only the lower approximation regions are considered. It is also necessary to show that axioms hold for upper approximations.

Rough Functional Dependencies Satisfy Armstrong's Axioms

Proof

(1) **Reflexive**

If $Y \subseteq X \subseteq U$, then

$$redundant(t(X), t'(X)) \Rightarrow redundant(t(Y), t'(Y)),$$

and $$roughly\text{-}redundant(t(X), t'(X)) \Rightarrow roughly\text{-}redundant(t(Y), t'(Y)).$$

Hence,

$$X \rightarrow Y.$$

(2) **Augmentation**

If $Z \subseteq U$ and the rough functional dependency $X \rightarrow Y$ holds, then

$$redundant(t(XZ), t'(XZ)) \Rightarrow redundant(t(YZ), t'(YZ)),$$

and $$roughly\text{-}redundant(t(XZ), t'(XZ)) \Rightarrow roughly\text{-}redundant(t(YZ), t'(YZ)).$$

Hence,

$$XZ \rightarrow YZ.$$

(3) **Transitive**

If the rough functional dependencies $X \rightarrow Y$ and $Y \rightarrow Z$ hold, then

$$redundant(t(X), t'(X)) \Rightarrow redundant(t(Z), t'(Z)),$$

and $$roughly\text{-}redundant(t(X), t'(X)) \Rightarrow roughly\text{-}redundant(t(Z), t'(Z)).$$

Hence,

$$X \rightarrow Z.$$

□

Hence, rough functional dependencies satisfy Armstrong's axioms. Given a set of rough functional dependencies, the complete set of rough functional dependencies can be derived using Armstrong's axioms. The rough functional dependency, therefore, is an important formalism for design in the rough relational database.

Fuzzy and rough set techniques integrated into the underlying data model result in databases that can more accurately represent real world enterprises since they incorporate uncertainty management directly into the data model itself. This is useful as is for obtaining greater information through the querying of rough and fuzzy databases. Additional benefits may be realized when they are used in the process of data mining.

Rough Set Modeling of Spatial Data

Many of the problems associated with data are prevalent in all types of database systems. Spatial databases and GIS contain descriptive as well as positional data. The various forms of uncertainty occur in both types of data, so many of the issues apply to ordinary databases as well, such as integration of data from multiple sources, time-variant data, uncertain data, imprecision in measurement, inconsistent wording of descriptive data, and "binning" or grouping of data into fixed categories, also are employed in spatial contexts.

First consider an example of the use of rough sets in representing spatially related data. Let U = {tower, stream, creek, river, forest, woodland, pasture, meadow} and let the equivalence relation R be defined as follows:

$$R^* = \{[\text{tower}], [\text{stream, creek, river}], [\text{forest, woodland}], [\text{pasture, meadow}]\}.$$

Given some set X = {tower, stream, creek, river, forest, pasture}, we would like to define it in terms of its lower and

upper approximations:

$$\underline{R}X = \{\text{tower, stream, creek, river}\}, \text{ and}$$
$$\overline{R}X = \{\text{tower, stream, creek, river, forest, woodland, pasture, meadow}\}.$$

A *rough set in A* is the group of subsets of U with the same upper and lower approximations. In the example given, the rough set is

{{tower, stream, creek, river, forest, pasture}
{tower, stream, creek, river, forest, meadow}
{tower, stream, creek, river, woodland, pasture}
{tower, stream, creek, river, woodland, meadow}}.

Often spatial data is associated with a particular grid. The positions are set up in a regular matrix-like structure and data is affiliated with point locations on the grid. This is the case for raster data and for other types of non-vector type data such as topography or sea surface temperature data. There is a tradeoff between the resolution or the scale of the grid and the amount of system resources necessary to store and process the data. Higher resolutions provide more information, but at a cost of memory space and execution time.

If we approach these data issues from a rough set point of view, it can be seen that there is indiscernibility inherent in the process of gridding or rasterizing data. A data item at a particular grid point in essence may represent data near the point as well. This is due to the fact that often point data must be mapped to the grid using techniques such as nearest-neighbor, averaging, or statistics. The rough set indiscernibility relation may be set up so that the entire spatial area is partitioned into equivalence classes where each point on the grid belongs to an equivalence class. If the resolution of the grid changes, then, in fact, this is changing the granularity of the partitioning, resulting in fewer, but larger classes.

The approximation regions of rough sets are beneficial whenever information concerning spatial data regions is accessed. Consider a region such as a forest. One can reasonably conclude that any grid point identified as FOREST that is surrounded on all sides by grid points also identified as FOREST is, in fact, a point represented by the feature FOREST. However, consider points identified as FOREST that are adjacent to points identified as MEADOW. Is it not possible that these points represent meadow area as well as forest area but were identified as FOREST in the classification process? Likewise, points identified as MEADOW but adjacent to FOREST points may represent areas that contain part of the forest. This uncertainty maps naturally to the use of the approximation regions of the rough set theory, where the lower approximation region represents certain data and the boundary region of the upper approximation represents uncertain data. It applies to spatial database querying and spatial database mining operations.

If we force a finer granulation of the partitioning, a smaller boundary region results. This occurs when the resolution is increased. As the partitioning becomes finer and finer, finally a point is reached where the boundary region is non-existent. Then the upper and lower approximation regions are the same and there is no uncertainty in the spatial data as can be determined by the representation of the model.

In [50] Worboys models imprecision in spatial data based on the resolution at which the data is represented, and for issues related to the integration of such data. This approach relies on the issue of indiscernibility – a core concept for rough sets – but does not carry over the entire framework and is just described as "reminiscent of the theory of rough sets" [51]. Ahlqvist and colleagues [2] used a rough set approach to define a rough classification of spatial data and to represent spatial locations. They also proposed a measure for quality of a rough classification compared to a crisp classification and evaluated their technique on actual data from vegetation map layers. They considered the combination of fuzzy and rough set approaches for reclassification as required by the integration of geographic data. Another research group in a mapping and GIS context [49] have developed an approach using a rough raster space for the field representation of a spatial entity and evaluated it on a classification case study for remote sensing images. In [16] Bittner and Stell consider K-labeled partitions, which can represent maps, and then develop their relationship to rough sets to approximate map objects with vague boundaries. Additionally they investigate stratified partitions, which can be used to capture levels of details or granularity such as in consideration of scale transformations in maps, and extend this approach using the concepts of stratified rough sets.

Data Mining in Rough Databases

Association rules capture the idea of certain data items commonly occurring together and have been often considered in the analysis of a "marketbasket" of purchases. For example, a delicatessen retailer might analyze the previous year's sales and observe that of all purchases 30% were of *both* cheese *and* crackers and, for any of the sales that included cheese, 75% also included crackers. Then it is pos-

sible to conclude a rule of the form:

Cheese → Crackers

This rule is said to have a 75% degree of confidence and a 30% degree of support This particular form of data is largely based on the a priori algorithm developed by Agrawal [1]. Let a database of possible data items be [28]

$$D = \{d_1, d_2, \ldots d_n\}$$

and the relevant set of transactions (sales, query results, etc.) are

$$T = \{T_1, T_2, \ldots\}$$

where $T_i \subseteq D$. We are interested in discovering if there is a relationship between two sets of items (called itemsets) X_j, X_k; X_j, $X_k \subseteq D$. For such a relationship to be determined, the entire set of transactions in T must be examined and a count made of the number of transactions containing these sets, where a transaction T_i contains X_m if $X_m \subseteq T_i$. This count, called the support count of X_m, $SC_T(X_m)$, will be appropriately modified in the case of rough sets.

There are then two measures used in determining rules: the percentage of T_i's in T that

1. Contain both X_j and X_k (i. e. $X_j \cup X_k$) – called the *support* s
2. If T_i contains X_j then T_i also contains X_k – called the *confidence* c.

The support and confidence can be interpreted as probabilities:

1. s – Prob $(X_j \cup X_k)$ and
2. c – Prob $(X_k | X_j)$

We assume the system user has provided minimum values for these in order to generate only sufficiently interesting rules. A rule whose support and confidence exceeds these minimums is called a strong rule.

The result of a query is then:

$$T = \{\underline{R}T, \overline{R}T\}$$

and so we must take into account the lower approximation, $\underline{R}T$, and upper approximation, $\overline{R}T$, results of the rough query in developing the association rules.

Recall that in order to generate frequent itemsets, we must count the number of transactions T_j that support an itemset X_j. In the ordinary data mining algorithm one simply counts the occurrence of a value as 1 if in the set, or 0 if not. But now since the query result T is a rough set, we must modify the support count SC_T. So we define the rough support count, RSC_T, for the set X_j, to count differently in the upper and lower approximations:

$$RSC_T(X_j) = \sum_i W_{T_i}(X_j); X_j \subseteq T_i$$

where

$$W(X_j) = \begin{cases} 1 \text{ if } T_i \in \underline{R}T \\ a \text{ if } T_i \in \overline{R}T, 0 < a < 1. \end{cases}$$

The value, a, can be a subjective value obtained from the user depending on relative assessment of the roughness of the query result T. For the data mining example of the next section we chose a neutral default value of $a = \frac{1}{2}$. Note that $W(X_j)$ is included in the summation only all of the values of the itemset X_j are included in the transaction, i. e. it is a subset of the transaction.

Finally to produce the association rules from the set of relevant data T retrieved from the spatial database, we must consider the frequent itemsets. For the purposes of generating a rule such as $X_j\, X_k$ we can now extend the approach to rough support and confidence as follows:

$$RS = RSC_T(X_j \cup X_k)/|T|$$
$$RC = RSC_T(X_j \cup X_k)/RSC_T(X_j)\,.$$

In the spatial data mining area there have only been a few efforts using rough sets. In the research described in [8] approaches for attribute induction knowledge discovery [43] in rough spatial data are investigated. In [15] Bittner considers rough sets for spatio-temporal data and how to discover characteristic configurations of spatial objects focusing on the use of topological relationships for characterizations. In a survey of uncertainty-based spatial data mining Shi et al. [47] provide a brief general comparison of fuzzy and rough set approaches for spatial data mining.

Future Directions

There are several other approaches to uncertainty representation that may be more suitable for certain applications. Type-2 fuzzy have been of considerable recent interest [35]. In these as opposed to ordinary fuzzy sets in which the underlying membership functions are crisp, here the membership function are themselves fuzzy. Intuitionistic sets introduced by Atanassov [3,4] are another generalization of a fuzzy set. Two characteristic functions are used for capturing both the ordinary idea of degree of membership in the intuitionistic set as well as the degree of

non-membership of elements in the set and can be used in database design [10]. Related to the concepts introduced by rough sets is the idea of granularity for managing complex data by abstraction using information granules as discussed by Lin [32,33]. A granular set approach has also been introduced [30] which is a set and a number of disjoint subsets that constitute a semi-partition. Some prior database research on ordered relations [26], although not presented in the context of uncertainty of data, may provide some approaches to extend our work in this area. A main emphasis for future work is the incorporation of some of these research topics into mainstream database, GIS commercial products and semi-structured data on the semantic web.

Acknowledgment

The authors would like to thank the Naval Research Laboratory's Base Program, Program Element No. 0602435N for sponsoring this research.

Bibliography

Primary Literature

1. Agrawal R, Imielinski T, Swami A (1993) Mining Association Rules between sets of items in large databases. Proc of the 1993 ACM-SIGMOD International Conference on Management of Data. ACM Press, New York, NY, pp 207–216
2. Ahlqvist O, Keukelaar J, Oukbir K (2000) Rough classification and accuracy assessment. Int J Geogr Inf Sci 14:475–496
3. Atanassov K (1986) Intuitionistic Fuzzy Sets. Fuzzy Sets Syst 20:87–96
4. Atanassov K (2000) Intuitionistic Fuzzy Sets; Theory and Applications. Physica Verlag, Heidelberg
5. Beaubouef T, Petry F (1994) Fuzzy Set Quantification of Roughness in a Rough Relational Database Model. Proc. Third IEEE International Conference on Fuzzy Systems, Orlando, Florida, pp 172–177
6. Beaubouef T, Petry F (1994) Rough Querying of Crisp Data in Relational Databases. Proc. Third Int.l Workshop on Rough Sets and Soft Computing (RSSC'94), San Jose, California, pp 368–375
7. Beaubouef T, Petry F (2000) Fuzzy Rough Set Techniques for Uncertainty Processing in a Relational Database. Int J Intell Syst 15:389–424
8. Beaubouef T, Petry F (2002) A Rough Set Foundation for Spatial Data Mining Involving Vague Regions. Proc.FUZZ-IEEE'02, Honolulu, Hawaii, pp 767–772
9. Beaubouef T, Petry F (2004) Rough Functional Dependencies. 2004 Multiconferences: International Conf. On Information and Knowledge Engineering (IKE'04), Las Vegas, pp 175–179
10. Beaubouef T, Petry F (2007) Intuitionistic Rough Sets for Database Applications. In: Peters JF et al (eds) Transactions on Rough Sets VI. LNCS, vol 4374. Springer, Berlin, pp 26–30
11. Beaubouef T, Petry F (2007) Rough Sets: A Versatile Theory for Approaches To Uncertainty Management in Databases. Rough Computing: Theories, Technologies and Applications, Idea Group, Inc.
12. Beaubouef T, Petry F, Arora G (1998) Information-Theoretic Measures of Uncertainty for Rough Sets and Rough Relational Databases. Inf Sci 109:185–195
13. Beaubouef T, Petry F, Buckles B (1995) Extension of the Relational Database and its Algebra with Rough Set Techniques. Comput Intell 11:233–245
14. Bhandari D, Pal NR (1993) Some New Information Measures for Fuzzy Sets. Inf Sci 67:209–228
15. Bittner T (2000) Rough sets in spatio-temporal data mining. Proceedings of International Workshop on Temporal, Spatial and Spatio–Temporal Data Mining. Springer, Berlin-Heidelberg, pp 89–104
16. Bittner T, Stell J (2003) Stratified Rough Sets and Vagueness. In: Kuhn W, Worboys M, Timpf S (eds) Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science International Conference (COSIT'03). Springer, Berlin, pp 286–303
17. Buckles B, Petry F (1983) Information-Theoretical Characterization of Fuzzy Relational Databases. IEEE Trans. Syst Man Cybern 13:74–77
18. Buckles BP, Petry F (1985) Uncertainty models in information and database systems. J Inf Sci 11:77–87
19. Chanas S, Kuchta D (1992) Further remarks on the relation between rough and fuzzy sets. Fuzzy Sets Syst 47:391–394
20. de Luca A, Termini S (1972) A Definition of a Nonprobabilistic Entropy in the Setting of Fuzzy Set Theory. Inf Control 20:301–312
21. Dubois D, Prade H (1987) Twofold Fuzzy Sets and Rough Sets–Some Issues in Knowledge Representation. Fuzzy Sets Syst 23:3–18
22. Dubois D, Prade H (1992) Putting Rough Sets and Fuzzy Sets Together. In: Slowinski R (ed) Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic Publishers, Boston
23. Elmasri R, Navathe S (2004) Fundamentals of Database Systems, 4th edn. Addison Wesley, Boston
24. Frawley W, Piatetsky-Shapiro G, Matheus C (1991) Knowledge Discovery in Databases: An Overview. In: Piatetsky-Shapiro G, Frawley W (eds) Knowledge Discovery in Databases. AAAI, MIT Press, Cambridge, pp 1–27
25. Fung KT, Lam CM (1980) The Database Entropy Concept and its Application to the Data Allocation Problem. INFOR 18(4):354–363
26. Ginsburg S, Hull R (1983) Order Dependency in the Relational Model. Theor Comput Sci 26:146–195
27. Han J, Cai Y, Cercone N (1992) Knowledge Discovery in Databases: An Attribute–Oriented Approach, Proc. 18th VLDB Conf., Vancouver, Brit. Columbia, pp 547–559
28. Han J, Kamber M (2006) Data Mining: Concepts and Techniques, 2nd ed. Morgan Kaufman, San Diego, CA
29. Klir GJ, Folger TA (1988) Fuzzy Sets, Uncertainty, and Information. Prentice Hall, Englewood Cliffs NJ
30. Ligeza A (2002) Granular Sets and Granular Relation. Intelligent Information Systems, Physica Verlag, pp 331–340
31. Lin TY (1992) Topological and Fuzzy Rough Sets. In: Slowinski R (ed) Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory. Kluwer Academic Publishers, Boston, pp 287–304

32. Lin TY (1997) Granular Computing: From Rough Sets and Neighborhood Systems to Information Granulation and Computing in Words, European Congress on Intelligent Techniques and Soft Computing, September 8–12, pp 1602–1606
33. Lin TY (1999) Granular computing: Fuzzy Logic and Rough Sets, In: Zadeh L, Kacprzyk J (eds) Computing with Words in Information/ Intelligent Systems. Physica-Verlag, Heidelberg, pp 183–200
34. Makinouchi A (1977) A Consideration on normal form of not-necessarily normalized relation in the relational data model. Proc. of the 3rd Int. Conf. VLDB, pp 447–453
35. Mendel J, John R (2002) Type-2 Fuzzy Sets Made Simple. IEEE Trans on Fuzzy Sets 10:117–127
36. Nanda S, Majumdar S (1992) Fuzzy rough sets. Fuzzy Sets Syst 45:157–160
37. Ola A, Ozsoyoglu G (1993) Incomplete Relational Database Models Based on Intervals. IEEE Trans Knowl Data Eng 5:293–308
38. Pawlak Z (1982) Rough Sets. Int J Comput Inform Sci 11:341–356
39. Pawlak Z (1984) Rough Sets. Int J Man-Machine Stud 21:127–134
40. Pawlak Z (1985) Rough Sets and Fuzzy Sets. Fuzzy Sets Syst 17:99–102
41. Pawlak Z (1991) Rough Sets: Theoretical Aspects of Reasoning About Data. Kluwer Academic Publishers, Norwell, MA
42. Quinlan JR (1986) Induction of Decision Trees. Mach Learn 1:81–106
43. Raschia G, Mouaddib N (2002) SAINTETIQ:a fuzzy set-based approach to database summarization. Fuzzy Sets Syst 129:137–162
44. Roth M, Korth H, Batory D (1987) SQL/NF: A query language for non-1NF databases. Infor Syst 12:99–114
45. Shannon CL (1948) The Mathematical Theory of Communication. Bell Syst Tech J 27
46. Shenoi S, Melton A, Fan L (1992) Functional Dependencies and Normal Forms in the Fuzzy Relational Database Model. Infor Sci 60:1–28
47. Shi W, Wang S, Li D, Wang X (2003) Uncertainty-based Spatial Data Mining, Proceedings of Asia GIS Association, Wuhan, China, pp 124–35
48. Srinivasan P (1991) The importance of rough approximations for information retrieval. Int J Man-Mach Stud 34:657–671
49. Wang S, Li D, Shi W, Wang X (2002) Rough Spatial Description, International Archives of Photogrammetry and Remote Sensing, XXXII, Commission II, pp 503–510
50. Worboys M (1998) Computation with imprecise geospatial data. Comput Environ Urban Syst 22:85–106
51. Worboys M (1998) Imprecision in Finite Resolution Spatial Data. Geoinformatica 2:257–280
52. Wygralak M (1989) Rough Sets and Fuzzy Sets–Some Remarks on Interrelations. Fuzzy Sets Syst 29:241–243
53. Zadeh L (1965) Fuzzy Sets. Infor Control 8:338–353
54. Zvieli A, Chen P (1986) Entity–Relationship Modeling and Fuzzy Databases. Proc. of International Conference on Data Engineering, pp 320–327

Books and Reviews

Aczel J, Daroczy Z (1975) On Measures of Information and their Characterization. Academic Press, New York

Angryk R, Petry F (2007) Attribute-Oriented Fuzzy Generalization in Proximity and Similarity-based Relation Database Systems. Int J Intell Syst 22:763–781

Arora G, Petry F, Beaubouef T (1997) Information Measure of Type β Under Similarity Relations. Sixth IEEE International Conference on Fuzzy Systems Barcelona, pp 857–862

Arora G, Petry F, Beaubouef T (2001) A Note On New Parametric Measures of Information for Fuzzy Sets. J Comb Infor Syst Sci 26:167–174

Beaubouef T, Petry F (2001) Vague Regions and Spatial Relationships: A Rough Set Approach. Fourth International Conference on Computational Intelligence and Multimedia Applications, Yokosuka City, Japan, pp 313–318

Beaubouef T, Petry F (2001) Vagueness in Spatial Data: Rough Set and Egg-Yolk Approaches, 14th International Conference on Industrial & Engineering Applications of Artificial Intelligence, pp 367–373

Beaubouef T, Petry F (2003) Rough Set Uncertainty in an Object Oriented Data Model, Intelligent Systems for Information Processing. In: Bouchon B, Meunier L, Foulloy, Yager R (eds) From Representation to Applications. Elsevier, Amsterdam, pp 37–46

Beaubouef T, Petry F (2005) Normalization in a Rough Relational Database, International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing, pp 257–265

Beaubouef T, Petry F (2005) Representation of Spatial Data in an OODB Using Rough and Fuzzy Set Modeling, Soft Comput J 9:364–373

Beaubouef T, Petry F (2007) An Attribute-Oriented Approach for Knowledge Discovery in Rough Relational Databases, Proc FLAIRS'07, pp 507–508

Beaubouef T, Ladner R, Petry F (2004) Rough Set Spatial Data Modeling for Data Mining. Int J Intell Syst 19:567–584

Beaubouef T, Petry F, Arora G (1998) Information Measures for Rough and Fuzzy Sets and Application to Uncertainty in Relational Databases. In: Pal S, Skowron A (eds) Rough-Fuzzy Hybridization: A New Trend in Decision-Making. Springer, Singapore, pp 200–214

Beaubouef T, Petry F, Ladner R (2007) Spatial Data Methods and Vague Regions: A Rough Set Approach. Appl Soft Comput J 7:425–440

Buckles B, Petry F (1982) Security and Fuzzy Databases, Procs 1982 IEEE Int Conf on Cybernetics and Society, pp 622–625

Codd E (1970) A relational model of data for large shared data banks. Comm ACM 13:377–387

Ebanks B (1983) On Measures of Fuzziness and their Representations. J Math Anal Appl 94:24–37

Grzymala-Busse J (1991) Managing Uncertainty in Expert Systems. Kluwer Academic Publishers, Boston

Han J, Nishio S, Kawano H, Wang W (1998) Generalization-Based Data Mining in Object-Oriented Databases Using an Object-Cube Model. Data Knowl Eng 25:55–97

Havrda J, Charvat F (1967) Quantification Methods of Classification Processes: Concepts of Structural α Entropy. Kybernetica 3:149–172

Kapur J, Kesavan H (1992) Entropy Optimization Principles with Applications. Academic Press, New York

Slowinski R (1992) A Generalization of the Indiscernibility Relation for Rough Sets Analysis of Quantitative Information. 1st Int. Workshop on Rough Sets: State of the Art and Perspectives, Poland, pp 41–48

Rough Set Data Analysis

SHUSAKU TSUMOTO
Department of Medical Informatics, Faculty of Medicine, Shimane University, Shimane, Japan

Article Outline

Introduction

Rule-induction methods are classified into two categories, induction of deterministic rules and of probabilistic ones [4,5,7,10].

On one hand, deterministic rules are described as *if-then* rules, which can be viewed as propositions. From the set-theoretical point of view, a set of examples supporting the conditional part of a deterministic rule, denoted by C, is a subset of a set whose examples belong to the consequence part, denoted by D. That is, the relation $C \subseteq D$ holds and deterministic rules are supported only by positive examples in a data set.

On the other hand, probabilistic rules are *if-then* rules with probabilistic information [10].

When a classical proposition will not hold for C and D, C is not a subset of D but closely overlapped with D. That is, the relations $C \cap D \neq \phi$ and $|C \cap D|/|C| \geq \delta$ will hold in this case, where the threshold δ is the degree of closeness of overlapping sets, which will be given by domain experts. (For more information, see Sect. "Definition of Rules")

Thus, probabilistic rules are supported by a large number of positive examples and a few negative examples.

The common feature of both deterministic and probabilistic rules is that they deduce their consequence positively if an example satisfies their conditional parts. We call the reasoning by these rules *positive reasoning*.

However, medical experts use not only positive reasoning but also *negative reasoning* for selection of candidates, which is represented as *if-then* rules whose consequences include negative terms.

For example, when a patient who complains of headache does not have a throbbing pain, migraine should not be suspected with a high probability. Thus, negative reasoning also plays an important role in cutting the search space of a differential diagnosis process [10]. Thus, medical reasoning includes both positive and negative reasoning, though conventional rule-induction methods do not reflect this aspect. This is one of the reasons medical experts have difficulty in interpreting induced rules, and interpreting rules for a discovery procedure does not proceed easily. Therefore, negative rules should be induced from databases in order not only to induce rules reflecting experts' decision processes, but also to induce rules that will be easier for domain experts to interpret, both of which are important to enhance the discovery process done by the cooperation of medical experts and computers.

In this chapter, first the characteristics of medical reasoning are discussed and then two kinds of rules, positive rules and negative rules, are introduced as a model of medical reasoning. Interestingly, from the set-theoretic point of view, sets of examples supporting both rules correspond to the lower and upper approximations in rough sets [5]. On the other hand, from the viewpoint of propositional logic, both positive and negative rules are defined as classical propositions or deterministic rules with two probabilistic measures, classification accuracy, and coverage.

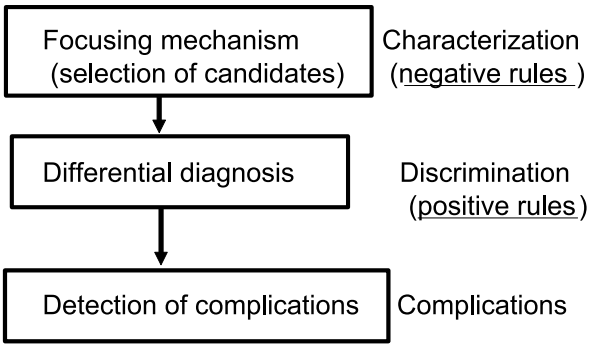
Second, two algorithms for induction of positive and negative rules are introduced, defined as search procedures using accuracy and coverage as evaluation indices.

Finally, the proposed method is evaluated on several medical databases. The experimental results show that the induced rules correctly represent experts' knowledge. In addition, several interesting patterns are discovered.

Focusing Mechanism

One of the characteristics in medical reasoning is a focusing mechanism, which is used to select the final diagnosis from many candidates [10,11]. For example, in differential diagnosis of headache, more than 60 diseases will be checked by present history, physical examinations, and laboratory examinations. In diagnostic procedures, a candidate is excluded if a symptom necessary to diagnose is not observed.

This style of reasoning consists of the following two processes: exclusive reasoning and inclusive reasoning. Relations of this diagnostic model with another diagnostic model are discussed in [12]. The diagnostic procedure proceeds as follows (Fig. 2): First, exclusive reasoning ex-



Rough Set Data Analysis, Figure 1
Illustration of focusing mechanism

cludes a disease from candidates when a patient does not have a symptom that is necessary to diagnose that disease.

Second, inclusive reasoning suspects a disease in the output of the exclusive process when a patient has symptoms specific to a disease.

These two steps are modeled as two kinds of rules, negative rules (or exclusive rules) and positive rules; the former corresponds to exclusive reasoning, the latter to inclusive reasoning.

In the next two sections, these two rules are represented as special kinds of probabilistic rules.

Definition of Rules

Rough Sets

In the following sections, we use the following notation introduced by Grzymala–Busse and Skowron [8], based on rough set theory [5].

These notations are illustrated by a small data set shown in Table 1, which includes symptoms exhibited by six patients who complained of headache.

Let U denote a nonempty finite set called the universe and A denote a nonempty, finite set of attributes, i. e., $a: U \to V_a$ for $a \in A$, where V_a is called the domain of a, respectively. Then a decision table is defined as an information system, $A = (U, A \cup \{d\})$. For example, Table 1 is an information system with $U = \{1, 2, 3, 4, 5, 6\}$ and $A = \{\text{age, location, nature, prodrome, nausea, } M1\}$ and $d = \text{class}$. For location $\in A$, V_{location} is defined as {occular, lateral, whole}.

The atomic formulas over $B \subseteq A \cup \{d\}$ and V are expressions of the form $[a = v]$, called descriptors over **B**, where $a \in B$ and $v \in V_a$. The set $F(B, V)$ of formulas over **B** is the least set containing all atomic formulas over B and closed with respect to disjunction, conjunction, and negation. For example, [location = occular] is a descriptor of B.

For each $f \in F(B, V)$, f_A denotes the meaning of f in A, i. e., the set of all objects in U with property f, defined inductively as follows:

1. If f is of the form $[a = v]$, then $f_A = \{s \in U | a(s) = v\}$
2. $(f \wedge g)_A = f_A \cap g_A$; $(f \vee g)_A = f_A \vee g_A$; $(\neg f)_A = U - f_a$.

For example, $f = [\text{location} = \text{whole}]$ and $f_A = \{2, 4, 5, 6\}$. As an example of a conjunctive formula, $g = [\text{location} = \text{whole}] \wedge [\text{nausea} = \text{no}]$ is a descriptor of U and f_A is equal to $g_{\text{location, nausea}} = \{2, 5\}$.

Classification Accuracy and Coverage

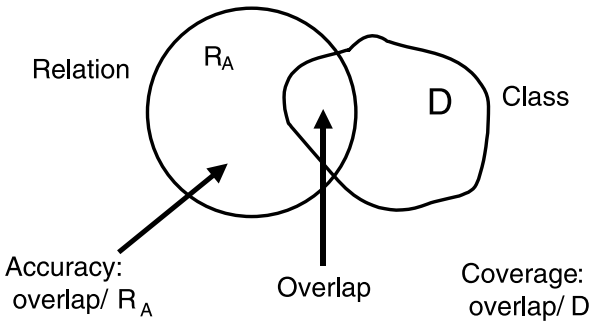
Definition of Accuracy and Coverage By use of the preceding framework, classification accuracy and coverage, or true positive rate are defined as follows.

Definition 1 Let R and D denote a formula in $F(B, V)$ and a set of objects that belong to a decision d. Classification accuracy and coverage (true positive rate) for $R \to d$

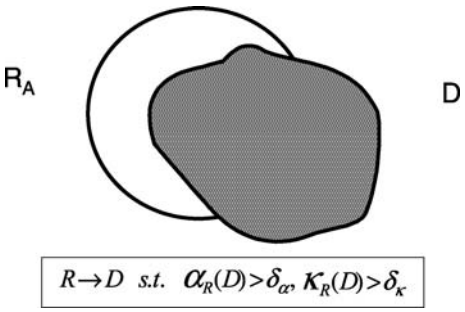
Rough Set Data Analysis, Table 1
An example of a data set

No.	Age	Location	Nature	Prodrome	Nausea	M1	Class
1	50–59	occular	persistent	no	no	yes	m.c.h.
2	40–49	whole	persistent	no	no	yes	m.c.h.
3	40–49	lateral	throbbing	no	yes	no	migra
4	40–49	whole	throbbing	yes	yes	no	migra
5	40–49	whole	radiating	no	no	yes	m.c.h.
6	50–59	whole	persistent	no	yes	yes	psycho

M1: tenderness of M1; m.c.h.: muscle contraction headache; migra: migraine; psycho: psychological pain



Rough Set Data Analysis, Figure 2
Venn diagram of accuracy and coverage



Rough Set Data Analysis, Figure 3
Venn diagram for probabilistic rules

is defined as:

$$\alpha_R(D) = \frac{|R_A \cap D|}{|R_A|} (= P(D|R)) \text{ and}$$
$$\kappa_R(D) = \frac{|R_A \cap D|}{|D|} (= P(R|D)),$$

where $|S|$, $\alpha_R(D)$, $\kappa_R(D)$, and $P(S)$ denote the cardinality of a set S, a classification accuracy of R as to classification of D, and coverage (a true positive rate of R to D), and probability of S, respectively.

Figure 2 depicts the Venn diagram of relations between accuracy and coverage. Accuracy views the overlapped region $|R_A \cap D|$ from the meaning of a relation R. On the other hand, coverage views the overlapped region from the meaning of a concept D.

In the preceding example, when R and D are set to [nau = yes] and [class = migraine], $\alpha_R(D) = 2/3 = 0.67$ and $\kappa_R(D) = 2/2 = 1.0$.

It is notable that $\alpha_R(D)$ measures the degree of the sufficiency of a proposition, $R \to D$, and that $\kappa_R(D)$ measures the degree of its necessity. For example, if $\alpha_R(D)$ is equal to 1.0, then $R \to D$ is true. On the other hand, if $\kappa_R(D)$ is equal to 1.0, then $D \to R$ is true. Thus, if both measures are 1.0, then $R \leftrightarrow D$.

Probabilistic Rules

By use of accuracy and coverage, a probabilistic rule is defined as:

$$R \overset{\alpha,\kappa}{\to} d \quad \text{s.t.} \quad R = \wedge_j [a_j = v_k], \quad \alpha_R(D) \geq \delta_\alpha \quad \text{and} \quad \kappa_R(D) \geq \delta_\kappa .$$

If the thresholds for accuracy and coverage are set to high values, the meaning of the conditional part of probabilistic rules corresponds to the highly overlapped region. Figure 3 depicts the Venn diagram of probabilistic rules with highly overlapped region.

This rule is a kind of probabilistic proposition with two statistical measures, which is an extension of Ziarko's variable precision model (VPRS) [15].[1]

It is also notable that both a positive rule and a negative rule are defined as special cases of this rule, as shown in the next sections.

Positive Rules

A positive rule is defined as a rule supported by only positive examples, the classification accuracy of which is equal to 1.0.

It is notable that the set supporting this rule corresponds to a subset of the lower approximation of a target concept, which is introduced in rough sets [5]. Thus, a positive rule is represented as:

$$R \to d \quad \text{s.t.} \quad R = \wedge_j [a_j = v_k], \quad \alpha_R(D) = 1.0 .$$

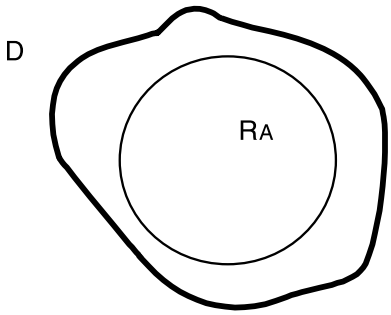
Figure 4 shows the Venn diagram of a positive rule. As shown in this figure, the meaning of R is a subset of that of D. This diagram is exactly equivalent to the classic proposition $R \to d$.

In the preceding example, one positive rule of m.c.h. (muscle contraction headache) is:

$$[\text{nausea} = \text{no}] \to \text{m.c.h.} \quad \alpha = 3/3 = 1.0 .$$

This positive rule is often called a deterministic rule. However, we use the term, positive (deterministic) rules, because a deterministic rule supported only by negative examples, called a negative rule, is introduced in the next section.

[1] This probabilistic rule is also a kind of *rough modus ponens* [6]



Rough Set Data Analysis, Figure 4
Venn diagram of positive rules

Negative Rules

Before defining a negative rule, let us first introduce an exclusive rule, the contrapositive of a negative rule [10]. An exclusive rule is defined as a rule supported by all the positive examples, the coverage of which is equal to 1.0. That is, an exclusive rule represents the necessity condition of a decision.

It is notable that the set supporting an exclusive rule corresponds to the upper approximation of a target concept, which is introduced in rough sets [5]. Thus, an exclusive rule is represented as:

$$R \to d \quad \text{s.t.} \qquad R = \vee_j [a_j = v_k] \,, \quad \kappa_R(D) = 1.0 \,.$$

Figure 5 shows the Venn diagram of an exclusive rule. As shown in this figure, the meaning of R is a superset of that of D. This diagram is exactly equivalent to the classic proposition $d \to R$.

In the preceding example, the exclusive rule of m.c.h. is:

$$[M1 = \text{yes}] \vee [\text{nau} = \text{no}] \to \text{m.c.h.} \quad \kappa = 1.0 \,,$$

From the viewpoint of propositional logic, an exclusive rule should be represented as:

$$d \to \vee_j [a_j = v_k] \,,$$

because the condition of an exclusive rule corresponds to the necessity condition of conclusion d. Thus, it is easy to see that a negative rule is defined as the contrapositive of an exclusive rule:

$$\wedge_j \neg [a_j = v_k] \to \neg d \,,$$

which means that if a case does not satisfy any attribute value pairs in the condition of a negative rule, then we can exclude a decision d from candidates.

For example, the negative rule of m.c.h. is:

$$\neg [M1 = \text{yes}] \wedge \neg [\text{nausea} = \text{no}] \to \neg \text{m.c.h.}$$

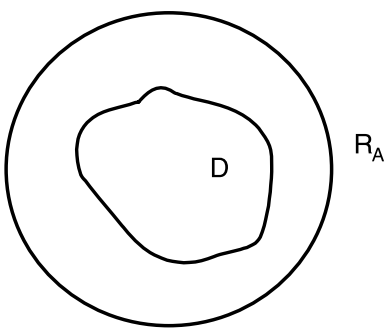
In summary, a negative rule is defined as:

$$\wedge_j \neg [a_j = v_k] \to \neg d \quad \text{s.t.} \quad \forall [a_j = v_k] \kappa_{[a_j = v_k]}(D) = 1.0 \,,$$

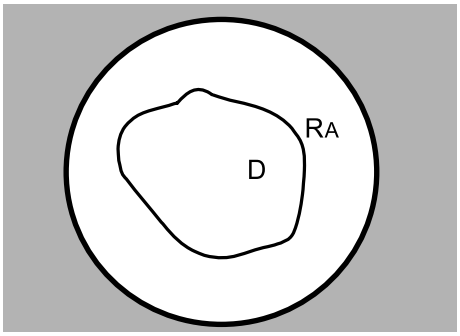
where D denotes a set of samples that belong to a class d. Figure 6 shows the Venn diagram of a negative rule. As shown in this figure, it is notable that this negative region is the "positive region" of "negative concept."

Negative rules should also be included in a category of deterministic rules, because their coverage, a measure of negative concepts, is equal to 1.0. It is also notable that the set supporting a negative rule corresponds to a subset of negative region, which is introduced in rough sets [5].

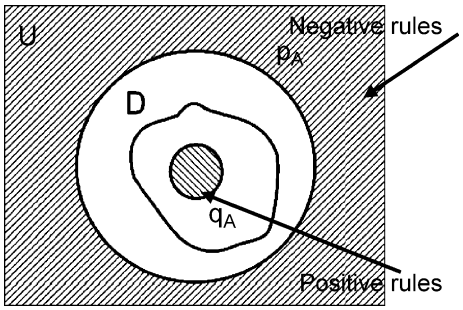
In summary, positive and negative rules correspond to positive and negative regions defined in rough sets. Figure 7 shows the Venn diagram of those rules.



Rough Set Data Analysis, Figure 5
Venn diagram of exclusive rules



Rough Set Data Analysis, Figure 6
Venn diagram of negative rules



Rough Set Data Analysis, Figure 7
Venn diagram of defined rules

Algorithms for Rule Induction

The contrapositive of a negative rule, an exclusive rule, is induced as an exclusive rule by the modification of the algorithm introduced in PRIMEROSE-REX [10], as shown in Fig. 8. This algorithm works as follows.

(1) First it selects a descriptor $[a_i = v_j]$ from the list of attribute-value pairs, denoted by L.
(2) Then it checks whether this descriptor overlaps with a set of positive examples, denoted by D.
(3) If so, this descriptor is included in a list of candidates for positive rules and the algorithm checks whether its coverage is equal to 1.0. If the coverage is equal to 1.0, then this descriptor is added to $R_e r$, the formula for the conditional part of the exclusive rule of D.
(4) Then $[a_i = v_j]$ is deleted from the list L. This procedure, from (1) to (4), will continue unless L is empty.
(5) Finally, when L is empty, this algorithm generates negative rules by taking the contrapositive of induced exclusive rules.

On the other hand, positive rules are induced as inclusive rules by the algorithm introduced in PRIMEROSE-REX [10], as shown in Fig. 9. For induction of positive rules, the threshold of accuracy and coverage is set to 1.0 and 0.0, respectively.

This algorithm works in the following way.

(1) First it substitutes L_1, which denotes a list of formulas composed of only one descriptor, with the list L_{er} generated by the former algorithm shown in Fig. 9.

```
procedure Exclusive and Negative Rules;
  var
    L : List;
      /* A list of elementary attribute-value pairs */
  begin
    L := P0;
    /* P0: A list of elementary attribute-value pairs given in a database */
    while (L ≠ {}) do
      begin
        Select one pair [ai = vj] from L;
        if ([ai = vj]A ∩ D ≠ ϕ) then do /* D: positive examples of a target class d */
          begin
          Lir := Lir + [ai = vj];    /* Candidates for Positive Rules */
          if (κ[ai=vj](D) = 1.0)
          then Rer := Rer ∧ [ai = vj];
            /* Include [ai = vj] into the formula of Exclusive Rule */
          end
      L := L − [ai = vj];
      end
    Construct Negative Rules:
        Take the contrapositive of Rer.
  end {Exclusive and Negative Rules};
```

Rough Set Data Analysis, Figure 8
Induction of exclusive and negative rules

```
procedure Positive Rules;
  var
    i : integer;    M, L_i : List;
  begin
    L_1 := L_ir;
    /* L_ir: A list of candidates generated by induction of exclusive rules */
    i := 1;    M := {};
    for i := 1 to n do
    /* n: Total number of attributes given
       in a database */
      begin
        while ( L_i ≠ {} ) do
          begin
            Select one pair R = ∧[a_i = v_j] from L_i;
            L_i := L_i − {R};
            if   (α_R(D) > δ_α)
              then do S_ir := S_ir + {R};
          /* Include R in a list of the Positive Rules */
            else M := M + {R};
          end
        L_{i+1} := (A list of the whole combination of the conjunction formulae in M);
      end
  end {Positive Rules};
```

Rough Set Data Analysis, Figure 9
Induction of positive rules

(2) Then until L_1 becomes empty, the following steps will continue: (a) A formula $[a_i = v_j]$ is removed from L_1. (b) Then the algorithm checks whether $\alpha_R(D)$ is larger than the threshold. (For induction of positive rules, this is equal to checking whether $\alpha_R(D)$ is equal to 1.0.) If so, then this formula is included a list of the conditional parts of positive rules. Otherwise, it will be included in M, which is used for making conjunctions.
(3) When L_1 is empty, the next list L_2 is generated from the list M.

Experimental Results

For experimental evaluation, a new system, called PRIMEROSE-REX2 (Probabilistic Rule Induction Method for Rules of Expert System ver. 2.0), was developed, where the algorithms discussed in Sect. "Algorithms for Rule Induction" were implemented.

PRIMEROSE-REX2 was applied to the following three medical domains:

(1) Headache (RHINOS domain), whose training samples consist of 52,119 samples, 45 classes, and 147 attributes;
(2) Cerebulovasular diseases (CVD), whose training samples consist of 7620 samples, 22 classes, and 285 attributes; and
(3) Meningitis, whose training samples consist of 1211 samples, 4 classes, and 41 attributes (Table 2).

For evaluation, we used the following two types of experiments. One experiment was to evaluate the predictive accuracy using the cross-validation method, which is often used in the machine-learning literature [9]. The other ex-

Rough Set Data Analysis, Table 2
Databases

Domain	Samples	Classes	Attributes
Headache	52,119	45	147
CVD	7620	22	285
Meningitis	1211	4	41

periment was to evaluate the induced rules by medical experts and to check whether these rules lead to a new discovery.

Performance of Rules Obtained

For comparison of performance, the experiments are conducted by the following four procedures. First, rules are acquired manually from experts. Second, the data sets are randomly split into new training samples and new test samples. Third, PRIMEROSE-REX2, conventional rule-induction methods, AQ15 [4] and C4.5 [7] are applied to the new training samples for rule generation. Fourth, the induced rules and rules acquired from experts are tested using new test samples. The second through fourth steps are repeated 100 times, and the average classification accuracy over 100 trials is computed. This process is a variant of repeated two-fold cross-validation, introduced in [10].

Experimental results (performance) are shown in Table 3. The first and second rows show the results obtained using PRIMROSE-REX2; the results in the first row are derived using both positive and negative rules and those in the second row are derived by only positive rules. The third row shows the results derived from medical experts. For comparison, we compare the classification accuracy of C4.5 and AQ-15, which is shown in the fourth and fifth rows.

These results show that the combination of positive and negative rules outperforms positive rules, although it is a little worse than medical experts' rules.

What Is Discovered?

Interesting Rules Are Very Few

One of the most important observations is that there were very few rules interesting or unexpected to medical experts compared to the number of rules extracted from the data sets.

Table 4 shows the mentioned results. The second column denotes the number of positive and negative rules obtained from each data set. The third column denotes the number of positive and negative rules interesting or unexpected for domain experts. For example, the first row shows that the number of induced positive rules in headache is 24,335, but the number of interesting rules are 114, which shows that only 0.47% of rules are interesting for domain experts.

This table shows that the number of interesting rules are very few: even in the case of meningitis, only 6.5% are interesting, which suggests that the interpretation part of domain experts is very hard for the knowledge discovery process.

Next, we show several examples of rules that are interesting to domain experts.

Positive Rules in Differential Diagnosis of Headache

In the domain of differential diagnosis of headache, the following interesting positive rules were found.

$$[\text{Age} < 20] \wedge [\text{History:paroxysmal}] \rightarrow \text{Common Migraine(Coverage:0.75)}$$

This rule is said to be interesting if it is compared with the following rule:

$$[\text{Age} > 40] \wedge [\text{History:paroxysmal}] \rightarrow \text{Classic Migraine(Coverage:0.72)}$$

These two rules include two parts; although the values of the attribute "Age" are different, those of the attribute "History" are the same. This suggests that the attribute "Age" is important for differential diagnosis between common migraine and classic migraine.

Negative Rules in Differential Diagnosis of Headache

In the domain of differential diagnosis of headache, the following interesting negative rule was found.

$$\neg[\text{Nature:Persistent}] \wedge \neg[\text{History:acute}] \wedge \neg[\text{History:paroxysmal}]) \wedge \neg[\text{Neck Stiffness:yes}] \rightarrow \neg\text{Common Migraine}$$

It is notable that it is difficult even for domain experts to interpret the interestingness of this rule if only this rule is shown. The domain experts pointed out that the rule is interesting when compared with the following rules.

$$[\text{Nature:Persistent}] \wedge [\text{History:acute}] \wedge [\text{Neck Stiffness:yes}] \rightarrow \text{Meningitis}$$

$$[\text{Nature:Persistent}] \wedge [\text{History:chronic}] \rightarrow \text{Brain Tumor}$$

This means that the rule includes information that is very important for differential diagnosis between common migraine and meningitis or brain tumor. Note that these two rules are very similar to the negative rule except for the negative symbols.

Rough Set Data Analysis, Table 3
Experimental results (accuracy: averaged)

Method	Headache	CVD	Meningitis
PRIMEROSE-REX2 (positive+negative)	91.3%	89.3%	92.5%
PRIMEROSE-REX2 (positive)	68.3%	71.3%	74.5%
Experts	95.0%	92.9%	93.2%
PR-REX	88.3%	84.3%	82.5%
C4.5	85.8%	79.7%	81.4%
AQ15	86.2%	78.9%	82.5%

Rough Set Data Analysis, Table 4
Number of extracted rules and interesting rules

	Induced positive rules	Interesting rules
Headache	24,335	114 (0.47%)
CVD	14,715	106 (0.72%)
Meningitis	1,922	40 (2%)

	Induced negative rules	Interesting rules
Headache	12,113	120 (0.99%)
CVD	7,231	155 (2.1%)
Meningitis	77	5 (6.5%)

Positive Rules in Meningitis

In the domain of meningitis, the following positive rules, which medical experts do not expect, are obtained.

$$[\text{WBC} < 12\,000] \wedge [\text{Sex} = \text{Female}] \wedge [\text{Age} < 40] \wedge [\text{CSF\_CELL} < 1000] \rightarrow \text{Virus(Coverage:0.91)}$$

$$[\text{Age} \geq 40] \wedge [\text{WBC} \geq 8000] \wedge [\text{Sex} = \text{Male}] \wedge [\text{CSF\_CELL} \geq 1000] \rightarrow \text{Bacteria(Coverage:0.64)}$$

The former rule means that if WBC (white blood cell count) is less than 12 000, the gender of a patient is female, the age is less than 40, and CSF_CELL (cell count of cerebulospinal fluid), then the type of meningitis is Viral. The latter means that the age of a patient is less than 40, WBC is larger than 8000, the gender is male, and CSF_CELL is larger than 1000, then the type of meningitis is Bacterial.

The most interesting points are that these rules included information about age and gender, which often seems to be unimportant attributes for differential diagnosis of meningitis. The first discovery was that women did not often suffer from bacterial infection compared with men, because such relationships between gender and meningitis has not been discussed in medical context [1]. By a close examination of the database on meningitis, it was found that most of the patients suffered from chronic diseases, such as DM, LC, and sinusitis, which are the risk factors of bacterial meningitis. The second discovery was that [age < 40] was also an important factor not to suspect viral meningitis, which also matches the fact that most old people suffer from chronic diseases. These results were also reevaluated in medical practice. Recently, the preceding two rules were checked by an additional 21 cases who suffered from meningitis (15 cases viral, 6 cases bacterial meningitis.) Surprisingly, the rules misclassified only three cases (two viral, the other bacterial), that is, the total accuracy was equal to $18/21 = 85.7\%$, and the accuracies for viral and bacterial meningitis were equal to $13/15 = 86.7\%$ and $5/6 = 83.3\%$. The reasons for misclassification are the following: a case of bacterial infection involved a patient who had a severe immunodeficiency, although he is very young. Two cases of viral infection involved patients who suffered from herpes zoster. It is notable that even those misclassified cases could be explained from the viewpoint of the immunodeficiency: that is, it was confirmed that immunodeficiency is a key factor for meningitis.

The validation of these rules is still ongoing; it will be reported in the near future.

Positive and Negative Rules in CVD

Concerning the database on CVD, several interesting rules were derived. The most interesting results were the following positive and negative rules for thalamus hemorrhage:

$$[\text{Sex} = \text{Female}] \wedge [\text{Hemiparesis} = \text{Left}] \wedge [\text{LOC:positive}] \rightarrow \text{Thalamus}$$

$$\neg[\text{Risk:Hypertension}] \wedge \neg[\text{Sensory} = \text{no}] \rightarrow \neg\text{Thalamus}$$

The former rule means that if the gender of a patient is female and he or she suffered from the left hemiparesis ([Hemiparesis=Left]) and loss of consciousness ([LOC:positive]), then the focus of CVD is thalamus. The latter rule means that if he or she suffers neither from hypertension ([Risk: Hypertension]) or sensory disturbance ([Sensory=no]), then the focus of CVD is thalamus.

Interestingly, LOC (loss of consciousness) under the condition of [Gender = Female] ∧ [Hemiparesis = Left] was found to be an important factor to diagnose thalamic damage. In this domain, any strong correlations between these attributes and others, like the database of meningitis, have not been found yet. It will be our future work to find what factor is behind these rules.

Rule Discovery as Knowledge Acquisition and Decision Support

Expert System: RH

Another point of discovery of rules is automated knowledge acquisition from databases. Knowledge acquisition is referred to as a bottleneck problem in development of expert systems [2], which has not fully been solved and is expected to be solved by induction of rules from databases. However, there are few papers that discuss the evaluation of discovered rules from the viewpoint of knowledge acquisition [12].

For this purpose, we have developed an expert system, called RH (rule-based system for headache) using the acquired knowledge. The reason for selecting the domain of headache is that earlier we developed an expert system RHINOS (rule-based headache information organizing system), which makes a differential diagnosis in headache [3]. In this system, it takes about six months to acquire knowledge from domain experts. RH consists of two parts. First, it requires inputs and applies exclusive and negative rules to select candidates (focusing mechanism). Then, it requires additional inputs and applies positive rules for differential diagnosis between selected candidates. Finally, RH outputs diagnostic conclusions.

Evaluation of RH

RH was evaluated in clinical practice with respect to its classification accuracy by using 930 patients who came to the outpatient clinic after the development of this system. Experimental results about classification accuracy are shown in Table 5. The first and second rows show the performance of rules obtained using PRIMROSE-REX2; the results in the first row are derived using both positive and negative rules and those in the second row are derived using only positive rules. The third and fourth rows show the results derived using both positive and negative rules and those by positive rules acquired directly from medical experts. These results show that the combination of positive and negative rules outperforms positive rules and gains almost the same performance as those by experts.

Rough Set Data Analysis, Table 5
Evaluation of RH (accuracy: averaged)

Method	Accuracy
PRIMEROSE-REX2 (positive and negative)	91.4% (851/930)
PRIMEROSE-REX2 (positive)	78.5% (729/930)
RHINOS (positive and negative)	93.5% (864/930)
RHINOS (positive)	82.8% (765/930)

Discussion

Hierarchical Rules for Decision Support

One of the problems with rule induction is that conventional rule-induction methods cannot extract rules that plausibly represent experts' decision processes [12]. The description length of induced rules is too short, compared to the experts' rules. (It may be observed that this length part does not contribute much to the classification performance.)

For example, rule-induction methods introduced in this chapter induced the following common rule for muscle contraction headache from databases on differential diagnosis of headache:

$$
\begin{aligned}
&[\text{location} = \text{whole}] \\
&\quad \wedge [\text{Jolt Headache} = \text{no}] \\
&\quad \wedge [\text{Tenderness of M1} = \text{yes}] \\
&\rightarrow \text{muscle contraction headache}
\end{aligned}
$$

This rule is shorter than the following rule given by medical experts:

$$
\begin{aligned}
&[\text{Jolt Headache} = \text{no}] \\
&\quad \wedge ([\text{Tenderness of M0} = \text{yes}] \\
&\quad \vee [\text{Tenderness of M1} = \text{yes}] \\
&\quad \vee [\text{Tenderness of M2} = \text{yes}]) \\
&\quad \wedge [\text{Tenderness of B1} = \text{no}] \\
&\quad \wedge [\text{Tenderness of B2} = \text{no}] \\
&\quad \wedge [\text{Tenderness of B3} = \text{no}] \\
&\quad \wedge [\text{Tenderness of C1} = \text{no}] \\
&\quad \wedge [\text{Tenderness of C2} = \text{no}] \\
&\quad \wedge [\text{Tenderness of C3} = \text{no}] \\
&\quad \wedge [\text{Tenderness of C4} = \text{no}] \\
&\rightarrow \text{muscle contraction headache}
\end{aligned}
$$

These results suggest that conventional rule-induction methods do not reflect a mechanism of knowledge acquisition of medical experts.

Typically, rules acquired from medical experts are much longer than those induced from databases, the decision attributes of which are given by the same experts. This is because rule induction methods generally search for shorter rules, compared with decision tree induction. In the case of decision tree induction, the induced trees are sometimes too deep, and in order for the trees to be useful for learning, pruning and examination by experts are required. One of the main reasons rules are short and decision trees are sometimes long is that these patterns are generated by only one criteria, such as high accuracy or high information gain. The comparative study in this section suggests that experts should acquire rules by usage of several measures. Those characteristics of medical experts' rules are fully examined not by comparing between those rules for the same class, but by comparing experts' rules with those for another class. For example, a classification rule for muscle contraction headache is given by:

$$
\begin{aligned}
&[\text{Jolt Headache} = \text{no}] \\
&\quad \wedge ([\text{Tenderness of M0} = \text{yes}] \\
&\quad \vee [\text{Tenderness of M1} = \text{yes}] \\
&\quad \vee [\text{Tenderness of M2} = \text{yes}]) \\
&\quad \wedge [\text{Tenderness of B1} = \text{no}] \\
&\quad \wedge [\text{Tenderness of B2} = \text{no}] \\
&\quad \wedge [\text{Tenderness of B3} = \text{no}] \\
&\quad \wedge [\text{Tenderness of C1} = \text{no}] \\
&\quad \wedge [\text{Tenderness of C2} = \text{no}] \\
&\quad \wedge [\text{Tenderness of C3} = \text{no}] \\
&\quad \wedge [\text{Tenderness of C4} = \text{no}] \\
&\rightarrow \text{muscle contraction headache}
\end{aligned}
$$

This rule is very similar to the following classification rule for disease of cervical spine:

$$
\begin{aligned}
&[\text{Jolt Headache} = \text{no}] \\
&\quad \wedge ([\text{Tenderness of M0} = \text{yes}] \\
&\quad \vee [\text{Tenderness of M1} = \text{yes}] \\
&\quad \vee [\text{Tenderness of M2} = \text{yes}]) \\
&\quad \wedge ([\text{Tenderness of B1} = \text{yes}] \\
&\quad \vee [\text{Tenderness of B2} = \text{yes}] \\
&\quad \vee [\text{Tenderness of B3} = \text{yes}] \\
&\quad \vee [\text{Tenderness of C1} = \text{yes}] \\
&\quad \vee [\text{Tenderness of C2} = \text{yes}] \\
&\quad \vee [\text{Tenderness of C3} = \text{yes}] \\
&\quad \vee [\text{Tenderness of C4} = \text{yes}]) \\
&\rightarrow \text{disease of cervical spine}
\end{aligned}
$$

The differences between these two rules are attribute-value pairs, from tenderness of B1 to C4. Thus, these two rules can be simplified into the following form:

$$
\begin{aligned}
a_1 \wedge A_2 \wedge \neg A_3 &\rightarrow \textit{muscle contraction headache}, \\
a_1 \wedge A_2 \wedge A_3 &\rightarrow \textit{disease of cervical spine}.
\end{aligned}
$$

The first two terms and the third one represent different reasoning. The first and second terms a_1 and A_2 are used to differentiate muscle contraction headache and disease of cervical spine from other diseases. The third term A_3 is used to make a differential diagnosis between these two diseases. Thus, medical experts first select several diagnostic candidates, which are similar to each other, from many diseases and then make a final diagnosis from those candidates.

This problem has been partially solved; Tsumoto introduced a new approach for inducing these rules in [13], as induction of hierarchical decision rules. In that paper, the characteristics of experts' rules are closely examined and a new approach to extract plausible rules is introduced, which consists of the following three procedures. First, the characterization of decision attributes (given classes) is done from databases and the classes are classified into several groups with respect to the characterization. Then two kinds of subrules, characterization rules for each group and discrimination rules for each class in the group, are induced. Finally, those two parts are integrated into one rule for each decision attribute. The proposed method was evaluated on a medical database, the experimental results of which show that induced rules correctly represent experts' decision processes.

This observation also suggests that medical experts implicitly look at the relation between rules for different concepts. Future work should discover the relations between induced rules.

Relations Between Rules

In [14], Tsumoto focuses on the characteristics of medical reasoning(focusing mechanism) and introduces three kinds of rules, positive rules, exclusive rules and total covering rules, as a model of medical reasoning, which is an extended formalization of rules defined in [12].

Interestingly, from the set-theoretic point of view, sets of examples supporting these rules correspond to the lower and upper approximations in rough sets.

Furthermore, from the viewpoint of propositional logic, both inclusive and exclusive rules are defined as classical propositions, or deterministic rules with two probabilistic measures, classification accuracy and coverage. To-

tal covering rules have several interesting relations with inclusive and exclusive rules, which reflects the characteristics of medical reasoning.

Originally, a total covering rule is defined as a set of symptoms that can be observed in at least one case of a target disease. That is, this rule is defined as a collection of attribute-value pairs whose accuracy is larger than 0:

$$R \to d \quad \text{s.t.} \quad R = \vee_j [a_j = v_k], \quad \alpha_R(D) > 0\,.$$

From the definition of accuracy and coverage, this formula can be transformed into:

$$R \to d \quad \text{s.t.} \quad R = \vee_j [a_j = v_k], \quad \kappa_R(D) > 0\,.$$

For each attribute, the attribute-value pairs form a partition of U. Thus, for each attribute, total covering rules include a covering of all the positive examples. According to this property, the preceding formula is redefined as:

$$R \to d \quad \text{s.t.} \quad R = \vee_j R(a_j),$$
$$R(a_j) = \vee_k [a_j = v_k] \quad \text{s.t.} \quad \kappa_R(D) = 1.0\,.$$

It is notable that this definition is an extension of exclusive rules and this total covering rule can be written as:

$$d \to \vee_j \vee_k [a_j = v_k] \quad \text{s.t.} \quad \kappa_{[a_j = v_k]}(D) = 1.0\,.$$

Let $S(R)$ denote a set of attribute-value pairs of rule R. For each class d, let $R_{\text{pos}}(d)$, $R_{\text{ex}}(d)$, and $R_{\text{tc}}(d)$ denote the positive exclusive rule and total covering rules, respectively. Then

$$S(R_{\text{ex}}(d)) \subseteq S(R_{\text{tc}}(d))\,,$$

because a total covering rule can be viewed as an upper approximation of exclusive rules. It is also notable that this relation will hold in the relation between the positive rule (inclusive rule) and total covering rule. That is,

$$S(R_{\text{pos}}(d)) \subseteq S(R_{\text{tc}}(d))\,.$$

Thus, the total covering rule can be viewed as an upper approximation of inclusive rules. This relation also holds when $R_{\text{pos}}(d)$ is replaced with a probabilistic rule, which shows that total covering rules are the weakest form of diagnostic rules.

In this way, rules that reflect the diagnostic reasoning of medical experts have sophisticated background from the viewpoint of set theory. Especially, the rough set framework provides a good tool for modeling such focusing mechanisms.

Our future work will investigate the relation between these three types of rules from the viewpoint of rough set theory.

Conclusions

In this chapter, the characteristics of two measures, classification accuracy and coverage, are discussed, which show that both measures are dual and that accuracy and coverage are measures of both positive and negative rules, respectively. Then an algorithm for induction of positive and negative rules is introduced.

The proposed method is evaluated on medical databases. The experimental results have demonstrated that the induced rules are able to correctly represent experts' knowledge. We also demonstrated that the method can discover several interesting patterns.

Bibliography

1. Adams RD, Victor M (1993) Principles of Neurology, 5th edn. McGraw-Hill, New York
2. Buchnan BG, Shortliffe EH (eds) (1984) Rule-Based Expert Systems. Addison-Wesley
3. Matsumura Y, Matsunaga T, Hata Y, Kimura M, Matsumura H (1988) Consultation system for diagnoses of headache and facial pain. Med Inform RHINOS 11:145–157
4. Michalski RS, Mozetic I, Hong J, Lavrac N (1986) The multi-purpose incremental learning system AQ15 and its testing application to three medical domains. Proc 5th National Conference on Artificial Intelligence. AAAI Press, Palo Alto, pp 1041–1045
5. Pawlak Z (1991) Rough Sets. Kluwer, Dordrecht
6. Pawlak Z (1998) Rough Modus Ponens. In: Proc. International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems 98, Paris
7. Quinlan JR (1993) C4.5–Programs for Machine Learning. Morgan Kaufmann, Palo Alto
8. Skowron A, Grzymala-Busse J (1994) From rough set theory to evidence theory. In: Yager R, Fedrizzi M, Kacprzyk J (eds) Advances in the Dempster-Shafer Theory of Evidence. Wiley, New York, pp 193–236
9. Shavlik JW, Dietterich TG (eds) (1990) Readings in Machine Learning. Morgan Kaufmann, Palo Alto
10. Tsumoto S, Tanaka H (1996) Automated discovery of medical expert system rules from clinical databases based on rough sets. In: Proc. 2nd International Conference on Knowledge Discovery and Data Mining 96. AAAI Press, Palo Alto, pp 63–69
11. Tsumoto S (1998) Modelling medical diagnostic rules based on rough sets. In: Polkowski L, Skowron A (eds) Rough Sets and Current Trends in Computing. Lecture Note in Artificial Intelligence, vol 1424. Springer, Berlin
12. Tsumoto S (1998) Automated extraction of medical expert system rules from clinical databases based on rough set theory. J Inf Sci 112:67–84
13. Tsumoto S (1998) Extraction of experts' decision rules from clinical databases using rough set model. Intell Data Anal 2(3)
14. Tsumoto S (2001) Medical diagnostic rules as upper approximation of rough sets. Proc FUZZ-IEEE2001
15. Ziarko W (1993) Variable precision rough set model. J Comput Syst Sci 46:39–59

Rough Sets in Decision Making

Roman Słowiński[1,2], Salvatore Greco[3], Benedetto Matarazzo[3]
[1] Institute of Computing Science, Poznan University of Technology, Poznan, Poland
[2] Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
[3] Faculty of Economics, University of Catania, Catania, Italy

Article Outline

Glossary

Multiple attribute (or multiple criteria) decision support aims at giving the decision maker (DM) a recommendation concerning a set of *objects* A (also called alternatives, actions, acts, solutions, options, candidates, ...) evaluated from multiple points of view called *attributes* (also called features, variables, criteria, objectives, ...).

Main categories of *multiple attribute (or multiple criteria) decision problems*:

- *classification*, when the decision aims at assigning each object to one of predefined classes,
- *choice*, when the decision aims at selecting the best object,
- *ranking*, when the decision aims at ordering objects from the best to the worst.

Two kinds of *classification problems* are distinguished:

- *taxonomy*, when the value sets of attributes and the predefined classes are not preference ordered,
- *ordinal classification* (also known as *multiple criteria sorting*), when the value sets of attributes and the predefined classes are preference ordered.

Two kinds of *choice problems* are distinguished:

- *discrete choice*, when the set of objects is finite and reasonably small to be listed,
- *multiple objective optimization*, when the set of objects is infinite and defined by constraints of a mathematical program.

If value sets of attributes are preference-ordered, they are called *criteria* or *objectives*, otherwise they keep the name of attributes.

Criterion is a real-valued function g_i defined on A, reflecting a worth of objects from a particular point of view, such that in order to compare any two objects $a, b \in A$ from this point of view it is sufficient to compare two values: $g_i(a)$ and $g_i(b)$.

Dominance Object a is non-dominated in set A (Pareto-optimal) if and only if there is no other object b in A such that b is not worse than a on all considered criteria, and strictly better on at least one criterion.

Preference model is a representation of a value system of the decision maker on the set of objects with vector evaluations.

Decision under uncertainty takes into account consequences of decisions that distribute over multiple states of nature with given probability. The preference order, characteristic for data describing *multiple attribute decision problems*, concerns also decision under uncertainty, where the objects correspond to acts, attributes are outcomes (gain or loss) to be obtained with a given probability, and the problem consists in ordinal classification, choice, or ranking of the acts.

Rough set in universe U is an approximation of a set based on available information about objects of U. The rough approximation is composed of two ordinary sets called *lower and upper approximation*. Lower approximation is a maximal subset of objects which, according to the available information, certainly belong to the approximated set, and upper approximation is a minimal subset of objects which, according to the available information, possibly belong to the approximated set. The difference between upper and lower approximation is called *boundary*.

Decision rule is a logical statement of the type "*if* ..., *then* ...", kursiv where the premise (condition part)

specifies values assumed by one or more condition attributes and the conclusion (decision part) specifies an overall judgment.

Definition of the Subject

Scientific analysis of decision problems aims at giving the decision maker (DM) a recommendation concerning a set of *objects* (also called alternatives, solutions, acts, actions, options, candidates, ...) evaluated from multiple points of view considered relevant for the problem at hand and called *attributes* (also called features, variables, criteria, objectives, ...). For example, a decision can concern:

1) diagnosis of pathologies for a set of patients, where patients are objects of the decision, and symptoms and results of medical tests are the attributes,
2) assignment of enterprises to classes of risk, where enterprises are objects of the decision, and financial ratio indices and other economic indicators, such as the market structure, the technology used by the enterprise and the quality of management, are the attributes,
3) selection of a car to be bought from among a given set of cars, where cars are objects of the decision, and maximum speed, acceleration, price, fuel consumption, comfort, color and so on, are the attributes,
4) ordering of students applying for a scholarship, where students are objects of the decision, and scores in different subjects are the attributes.

The following three main categories of decision problems are typically distinguished [44]:

- *classification*, when the decision aims at assigning each object to one of predefined classes,
- *choice*, when the decision aims at selecting the best objects,
- *ranking*, when the decision aims at ordering objects from the best to the worst.

Looking at the above examples, one can say that 1) and 2) are classification problems, 3) is a choice problem and 4) is a ranking problem.

The above categorization can be refined by distinguishing two kinds of classification problems: *taxonomy*, when the value sets of attributes and the predefined classes are not preference ordered, and *ordinal classification* (also known as *multiple criteria sorting*), when the value sets of attributes and the predefined classes are preference ordered [12]. In the above examples, 1) is a taxonomy problem and 2) is an ordinal classification problem. If value sets of attributes are preference ordered, they are called *criteria*, otherwise they keep the name of attributes. For example, in a decision regarding the selection of a car, its price is a criterion because, obviously, a low price is better than a high price. Instead, the color of a car is not a criterion but simply an attribute, because red is not intrinsically better than green. One can imagine, however, that also the color of a car could become a criterion if, for example, a DM would consider red better than green.

Introduction

Scientific support of decisions makes use of a more or less explicit model of the decision problem. The model relates the decision to the characteristics of the objects expressed by the considered attributes. Building such a model requires information about conditions and parameters of the aggregation of multi-attribute characteristics of objects. The nature of this information depends on the adopted methodology: prices and interest rates for cost-benefit analysis, cost coefficients in objectives and technological coefficients in constraints for mathematical programming, a training set of decision examples for neural networks and machine learning, substitution rates for a value function of multi-attribute utility theory, pairwise comparisons of objects in terms of intensity of preference for the analytic hierarchy process, attribute weights and several thresholds for ELECTRE methods, and so on (see the state-of-the-art survey [4]). This information has to be provided by the DM, possibly assisted by an analyst.

Very often this information is not easily definable. For example, this is the case of the price of many immaterial goods and of the interest rates in cost-benefit analysis, or the case of the coefficients of objectives and constraints in mathematical programming models. Even if the required information is easily definable, like a training set of decision examples for neural networks, it is often processed in a way which is not clear for the DM, such that (s)he cannot see what are the exact relations between the provided information and the final recommendation. Consequently, very often the decision model is perceived by the DM as a *black box* whose result has to be accepted because the analyst's authority guarantees that the result is "right". In this context, the aspiration of the DM to find good reasons to make decision is frustrated and rises the need for a more transparent methodology in which the relation between the original information and the final recommendation is clearly shown. Such a transparent methodology searched for has been called *glass box* [32]. Its typical representative is using a learning set of decision examples as the input preference information provided by the DM, and

it is expressing the decision model in terms of a set of "*if* ..., *then* ..." decision rules induced from the input information. From one side, the decision rules are explicitly related to the original information and, from the other side, they give understandable justifications for the decisions to be made.

For example, in case of a medical diagnosis problem, the decision rule approach requires as input information a set of examples of previous diagnoses, from which some diagnostic rules are induced, such as "if there is symptom α and the test result is β, then there is pathology γ". Each one of such rules is directly related to examples of diagnoses in the input information, where there is symptom α, test result β and pathology γ. Moreover, the DM can verify easily that in the input information there is no example of diagnosis where there is symptom α, test result β but no pathology γ.

The rules induced from the input information provided in terms of exemplary decisions represent a decision model which is transparent for the DM and enables his understanding of the reasons of his past decisions. The acceptance of the rules by the DM justifies, in turn, their use for future decisions.

The induction of rules from examples is a typical approach of artificial intelligence. This explains our interest in rough set theory [38,39] which proved to be a useful tool for analysis of vague description of decision situations [41,48]. The rough set analysis aims at explaining the values of some decision attributes, playing the role of "dependent variables", by means of the values of condition attributes, playing the role of "independent variables". For example, in the above diagnostic context, data about the presence of a pathology are given by decision attributes, while data about symptoms and tests are given by condition attributes. An important advantage of the rough set approach is that it can deal with partly inconsistent examples, i. e. cases where the presence of different pathologies is associated with the presence of the same symptoms and test results. Moreover, it provides useful information about the role of particular attributes and their subsets, and prepares the ground for representation of knowledge hidden in the data by means of "*if* ..., *then* ..." decision rules.

Classical rough set approach proposed by Pawlak [38,39] cannot deal, however, with preference order in the value sets of condition and decision attributes. For this reason, the classical rough set approach can deal with only one of four decision problems listed above – classification of taxonomy type. To deal with ordinal classification, choice and ranking, it is necessary to generalize the classical rough set approach, so as to take into account preference orders and monotonic relationships between condition and decision attributes. This generalization, called Dominance-based Rough Set Approach (DRSA), has been proposed by Greco, Matarazzo and Slowinski [12,14,15,18,21,49].

Classical Rough Set Approach to Classification Problems of Taxonomy Type

Information Table and Indiscernibility Relation

Information regarding classification examples is supplied in the form of an *information table*, whose separate rows refer to distinct *objects*, and whose columns refer to different *attributes* considered. This means that each cell of this table indicates an *evaluation* (quantitative or qualitative) of the object placed in the corresponding row by means of the attribute in the corresponding column. Formally, an information table is the 4-tuple $\boldsymbol{S} = \langle U, Q, V, v\rangle$, where U is a finite set of objects, called *universe*, $Q = \{q_1, \ldots, q_n\}$ is a finite set of attributes, V_q is a value set of the attribute q, $V = \bigcup_{q\in Q} V_q$, and $v\colon U \times Q \to V$ is a total function such that $v(x, q) \in V_q$ for each $q \in Q$, $x \in U$, called *information function*.

Therefore, each object x of U is described by a vector (string) $\mathrm{Des}_Q(x) = [v(x, q_1), \ldots, v(x, q_n)]$, called *description* of x in terms of the evaluations on the attributes from Q; it represents the available information about x. Obviously, $x \in U$ can be described in terms of any non-empty subset $P \subseteq Q$.

To every (non-empty) subset of attributes $P \subseteq Q$ there is associated an *indiscernibility relation* on U, denoted by I_P:

$$I_P = \{(x, y) \in U \times U\colon\ v(x, q) = v(y, q)\,,\ \forall q \in P\}\,.$$

If $(x, y) \in I_P$, it is said that the objects x and y are P-indiscernible. Clearly, the indiscernibility relation thus defined is an equivalence relation (reflexive, symmetric and transitive). The family of all the equivalence classes of the relation I_P is denoted by U/I_P, and the equivalence class containing object $x \in U$, by $I_P(x)$. The equivalence classes of the relation I_P are called *P-elementary sets*.

Approximations

Let $\boldsymbol{S}$ be an information table, X a non-empty subset of U and $\emptyset \neq P \subseteq Q$. The *P-lower approximation* and the *P-upper approximation* of X in $\boldsymbol{S}$ are defined, respectively, as:

$$\underline{P}(X) = \{x \in U\colon I_P(x) \subseteq X\}\,,$$

$$\overline{P}(X) = \{x \in U\colon I_P(x) \cap X \neq \emptyset\}\,.$$

The elements of $\underline{P}(X)$ are all and only those objects $x \in U$ which belong to the equivalence classes generated by the indiscernibility relation I_P, *contained* in X; the elements of $\overline{P}(X)$ are all and only those objects $x \in U$ which belong to the equivalence classes generated by the indiscernibility relation I_P, *containing at least one* object x belonging to X. In other words, $\underline{P}(X)$ is the largest union of the P-elementary sets included in X, while $\overline{P}(X)$ is the smallest union of the P-elementary sets containing X.

The *P-boundary* of X in $\mathbf{S}$, denoted by $Bn_P(X)$, is defined by

$$Bn_P(X) = \overline{P}(X) - \underline{P}(X) .$$

The following *inclusion relation* holds:

$$\underline{P}(X) \subseteq X \subseteq \overline{P}(X) .$$

Thus, in view of information provided by P, if an object x belongs to $\underline{P}(X)$, then it certainly belongs to set X, while if x belongs to $\overline{P}(X)$, then it possibly belongs to set X. $Bn_P(X)$ constitutes the "doubtful region" of X: nothing can be said with certainty about the membership of its elements to set X, using the subset of attributes P only.

Moreover, the following *complementarity relation* is satisfied:

$$\underline{P}(X) = U - \overline{P}(U - X) .$$

If the P-boundary of set X is empty, $Bn_P(X) = \emptyset$, then X is an ordinary (exact) set with respect to P, that is, it may be expressed as union of a certain number of P-elementary sets; otherwise, if $Bn_P(X) \neq \emptyset$, set X is an approximate (rough) set with respect to P and may be characterized by means of the approximations $\underline{P}(X)$ and $\overline{P}(X)$. The family of all sets $X \subseteq U$ having the same P-lower and P-upper approximations is called the *rough set*.

The *quality* of the approximation of set X by means of the attributes from P is defined as

$$\gamma_P(X) = \frac{|\underline{P}(X)|}{|X|} ,$$

such that $0 \leq \gamma_P(X) \leq 1$. The quality γ_P(X) represents the relative frequency of the objects correctly classified using the attributes from P.

The definition of approximations of a set $X \subseteq U$ can be extended to a classification, i. e. a partition $Y = \{Y_1, \ldots, Y_m\}$ of U. The P-lower and P-upper approximations of Y in $\mathbf{S}$ are defined by sets $\underline{P}(Y) = \{\underline{P}(Y_1), \ldots, \underline{P}(Y_m)\}$ and $\overline{P}(Y) = \{\overline{P}(Y_1), \ldots, \overline{P}(Y_m)\}$, respectively. The coefficient

$$\gamma_P(Y) = \frac{\sum_{i=1}^{m} |\underline{P}(Y_i)|}{|U|}$$

is called *quality of the approximation of classification* Y by set of attributes P, or in short, *quality of classification*. It expresses the ratio of all P-correctly classified objects to all objects in the system.

Dependence and Reduction of Attributes

An issue of great practical importance is the reduction of "superfluous" attributes in an information table. Superfluous attributes can be eliminated without deteriorating the information contained in the original table.

Let $P \subseteq Q$ and $p \in P$. It is said that attribute p is *superfluous* in P with respect to classification Y if $\underline{P}(Y) = \overline{(P-p)}(Y)$; otherwise, p is *indispensable* in P. The subset of Q containing all the indispensable attributes is known as the *core*.

Given classification Y, any minimal (with respect to inclusion) subset $P \subseteq Q$, such that $\underline{P}(Y) = \underline{Q}(Y)$, is called *reduct*. It specifies a minimal subset P of Q which keeps the quality of classification at the same level as the whole set of attributes, i. e. $\gamma_P(Y) = \gamma_Q(Y)$. In other words, the attributes that do not belong to the reduct are superfluous with respect to the classification Y of objects from U.

More than one reduct may exist in an information table and their intersection gives the *core*.

Decision Table and Decision Rules

In the information table describing examples of classification, the attributes of set Q are divided into *condition* attributes (set $C \neq \emptyset$) and *decision* attributes (set $D \neq \emptyset$), $C \cup D = Q$ and $C \cap D = \emptyset$. Such an information table is called a *decision table*. The decision attributes induce a partition of U deduced from the indiscernibility relation I_D in a way that is independent of the condition attributes. D-elementary sets are called *decision classes*, denoted by $Cl_t, t = 1, \ldots, m$. The partition of U into decision classes is called classification $\boldsymbol{Cl} = \{Cl_1, \ldots, Cl_m\}$. There is a tendency to reduce the set C while keeping all important relationships between C and D, in order to make decisions on the basis of a smaller amount of information. When the set of condition attributes is replaced by one of its reducts, the quality of approximation of the classification induced by the decision attributes is not deteriorating.

Since the aim is to underline the functional dependencies between condition and decision attributes, a decision table may also be seen as a set of *decision rules*. These are logical statements of the type "*if* ..., *then* ...", where the premise (condition part) specifies values assumed by one or more condition attributes (description of C-elementary sets) and the conclusion (decision part) specifies an assign-

ment to one or more decision classes. Therefore, the syntax of a rule is the following:

"*if* $v(x, q_1) = r_{q1}$ *and* $v(x, q_2) = r_{q2}$ *and* ... $v(x, q_p) = r_{qp}$, *then* x belongs to decision class Cl_{j1} *or* Cl_{j2} *or* ... Cl_{jk}",

where $\{q_1, q_2, \ldots, q_p\} \subseteq C, (r_{q1}, r_{q2}, \ldots, r_{qp}) \in V_{q1} \times V_{q2} \times \cdots \times V_{qp}$ and $Cl_{j1}, Cl_{j2}, \ldots, Cl_{jk}$ are some decision classes of the considered classification $\boldsymbol{Cl}$. If the consequence is univocal, i. e. $k = 1$, then the rule is *univocal*, otherwise it is *approximate*.

An object $x \in U$ *supports* decision rule r if its description is matching both the condition part and the decision part of the rule. The decision rule r *covers* object x if it matches the condition part of the rule. Each decision rule is characterized by its *strength*, defined as the number of objects supporting the rule. In the case of approximate rules, the strength is calculated for each possible decision class separately.

If a univocal rule is supported by objects from the lower approximation of the corresponding decision class only, then the rule is called *certain* or *deterministic*. If, however, a univocal rule is supported by objects from the upper approximation of the corresponding decision class only, then the rule is called *possible* or *probabilistic*. Approximate rules are supported, in turn, only by objects from the boundaries of the corresponding decision classes.

Procedures for generation of decision rules from a decision table use an inductive learning principle. The objects are considered as examples of classification. In order to induce a decision rule with a univocal and certain conclusion about assignment of an object to decision class X, the examples belonging to the C-lower approximation of X are called *positive* and all the others *negative*. Analogously, in case of a possible rule, the examples belonging to the C-upper approximation of X are positive and all the others negative. Possible rules are characterized by a coefficient, called *confidence*, telling to what extent the rule is consistent, i. e. what is the ratio of the number of positive examples supporting the rule to the number of examples belonging to set X according to decision attributes. Finally, in case of an approximate rule, the examples belonging to the C-boundary of X are positive and all the others negative. A decision rule is called *minimal* if removing any attribute from the condition part gives a rule covering also negative objects.

The existing induction algorithms use one of the following strategies [55]:

(a) Generation of a *minimal* representation, i. e. minimal set of rules covering all objects from a decision table,
(b) Generation of an *exhaustive* representation, i. e. all rules for a given decision table,
(c) Generation of a *characteristic* representation, i. e. a set of rules covering relatively many objects, however, not necessarily all objects from a decision table.

Explanation of the Classical Rough Set Approach by an Example

Suppose that one wants to describe the classification of basic traffic signs to a novice. There are three main classes of traffic signs corresponding to:

- Warning (W),
- Interdiction (I),
- Order (O).

Then, these classes may be distinguished by such attributes as the shape (S) and the principal color (PC) of the sign. Finally, one can consider a few examples of traffic signs, like those shown in Table 1. These are:

a) Sharp right turn,
b) Speed limit of 50 km/h,
c) No parking,
d) Go ahead.

The rough set approach is used here to build a model of classification of traffic signs to classes W, I, O on the basis of attributes S and PC. This is a typical problem of taxonomy.

One can remark that the sets of signs indiscernible by "Class" are:

$$\mathrm{W} = \{a\}_{\mathrm{Class}}, \quad \mathrm{I} = \{b, c\}_{\mathrm{Class}}, \quad \mathrm{O} = \{d\}_{\mathrm{Class}},$$

Rough Sets in Decision Making, Table 1
Examples of traffic signs described by S and PC

Traffic sign	Shape (S)	Primary color (PC)	Class
a) 	triangle	yellow	W
b) 	circle	white	I
c) 	circle	blue	I
d) 	circle	blue	O

and the sets of signs indiscernible by S and PC are as follows:

$$\{a\}_{\mathrm{S,PC}}, \quad \{b\}_{\mathrm{S,PC}}, \quad \{c,d\}_{\mathrm{S,PC}}.$$

The above elementary sets are generated, on the one hand, by decision attribute "Class" and, on the other hand, by condition attributes S and PC. The elementary sets of signs indiscernible by "Class" are denoted by $\{\cdot\}_{\mathrm{Class}}$ and those by S and PC are denoted by $\{\cdot\}_{\mathrm{S,PC}}$. Notice that $\mathrm{W} = \{a\}_{\mathrm{Class}}$ is characterized precisely by $\{a\}_{\mathrm{S,PC}}$. In order to characterize $\mathrm{I} = \{b,c\}_{\mathrm{Class}}$ and $\mathrm{O} = \{d\}_{\mathrm{Class}}$, one needs $\{b\}_{\mathrm{S,PC}}$ and $\{c,d\}_{\mathrm{S,PC}}$, however, only $\{b\}_{\mathrm{S,PC}}$ is included in $\mathrm{I} = \{b,c\}_{\mathrm{Class}}$ while $\{c,d\}_{\mathrm{S,PC}}$ has a non-empty intersection with both $\mathrm{I} = \{b,c\}_{\mathrm{Class}}$ and $\mathrm{O} = \{d\}_{\mathrm{Class}}$. It follows, from this characterization, that by using condition attributes S and PC, one can characterize class W precisely, while classes I and O can only be characterized approximately:

- Class W includes sign a certainly and possibly no other sign than a,
- Class I includes sign b certainly and possibly signs b, c and d,
- Class O includes no sign certainly and possibly signs c and d.

The terms "*certainly*" and "*possibly*" refer to the absence or presence of ambiguity between the description of signs by S and PC from the one side, and by "Class", from the other side. In other words, using description of signs by S and PC, one can say that all signs from elementary sets $\{\cdot\}_{\mathrm{S,PC}}$ included in elementary sets $\{\cdot\}_{\mathrm{Class}}$ belong certainly to the corresponding class, while all signs from elementary sets $\{\cdot\}_{\mathrm{S,PC}}$ having a non-empty intersection with elementary sets $\{\cdot\}_{\mathrm{Class}}$ belong to the corresponding class only possibly. The two sets of certain and possible signs are, respectively, the *lower* and *upper approximation* of the corresponding class by attributes S and PC:

$$\begin{aligned}
&\mathrm{lower_approx.}_{\mathrm{S,PC}}(\mathrm{W}) = \{a\}, \qquad \mathrm{upper_approx.}_{\mathrm{S,PC}}(\mathrm{W}) = \{a\},\\
&\mathrm{lower_approx.}_{\mathrm{S,PC}}(\mathrm{I}) = \{b\}, \qquad \mathrm{upper_approx.}_{\mathrm{S,PC}}(\mathrm{I}) = \{b,c,d\},\\
&\mathrm{lower_approx.}_{\mathrm{S,PC}}(\mathrm{O}) = \emptyset, \qquad \mathrm{upper_approx.}_{\mathrm{S,PC}}(\mathrm{O}) = \{c,d\}.
\end{aligned}$$

The *quality of approximation* of the classification by attributes S and PC is equal to the number of all the signs in the lower approximations divided by the number of all the signs in the table, i. e. 1/2.

Rough Sets in Decision Making, Table 2
Examples of traffic signs described by S, PC and SC

Traffic sign	Shape (S)	Primary color (PC)	Secondary color (SC)	Class
a) 	triangle	yellow	red	W
b) 	circle	white	red	I
c) 	circle	blue	red	I
d) 	circle	blue	white	O

One way to increase the quality of the approximation is to add a new attribute so as to decrease the ambiguity. Let us introduce the secondary color (SC) as a new condition attribute. The new situation is now shown in Table 2.

As one can see, the sets of signs indiscernible by S, PC and SC, i. e. the elementary sets $\{\cdot\}_{\mathrm{S,PC,SC}}$, are now:

$$\{a\}_{\mathrm{S,PC,SC}}, \quad \{b\}_{\mathrm{S,PC,SC}}, \quad \{c\}_{\mathrm{S,PC,SC}}, \quad \{d\}_{\mathrm{S,PC,SC}}.$$

It is worth noting that the elementary sets are finer than before and this enables the ambiguity to be eliminated. Consequently, the quality of approximation of the classification by attributes S, PC and SC is now equal to 1.

A natural question occurring here is to ask if, indeed, all three attributes are necessary to characterize precisely the classes W, I and O. When attribute S or attribute PC is eliminated from the description of the signs, the elementary sets $\{\cdot\}_{\mathrm{PC,SC}}$ or $\{\cdot\}_{\mathrm{S,SC}}$ are defined, respectively, as follows:

$$\{a\}_{\mathrm{PC,SC}}, \quad \{b\}_{\mathrm{PC,SC}}, \quad \{c\}_{\mathrm{PC,SC}}, \quad \{d\}_{\mathrm{PC,SC}},$$

$$\{a\}_{\mathrm{S,SC}}, \quad \{b,c\}_{\mathrm{S,SC}}, \quad \{d\}_{\mathrm{S,SC}}.$$

Using any one of the above elementary sets, it is possible to characterize (approximate) classes W, I and O with the same quality (equal to 1) as it is when using the elementary sets $\{\cdot\}_{\mathrm{S,PC,SC}}$ (i. e. those generated by the complete set of three condition attributes). Thus, the answer to the above question is that the three condition attributes are not all necessary to characterize precisely the classes W, I and O. It is, in fact, sufficient to use either PC and SC or S and SC. The subsets of condition attributes {PC, SC} and {S, SC} are called *reducts* of {S, PC, SC} because they have

this property. Note that the identification of reducts enables us to reduce attributes about the signs from the table to only those which are relevant.

Other useful information can be generated from the identification of reducts by taking their intersection. This is called the *core*. In our example, the core contains attribute SC. This tells us that it is clearly an *indispensable* attribute i. e. it cannot be eliminated from the description of the signs without decreasing the quality of the approximation. Note that other attributes from the reducts (i. e. S and PC) are *exchangeable*. If there happened to be some other attributes which were not included in any reduct, then they would be *superfluous*, i. e. they would not be useful at all in the characterization of the classes W, I and O.

If, however, column S or PC is eliminated from Table 2, then the resulting table is not a minimal representation of knowledge about the classification of the four traffic signs. Note that, in order to characterize class W in Table 2, it is sufficient to use the condition "S = triangle". Moreover, class I is characterized by two conditions ("S = circle" and "SC = red") and class O is characterized by the condition "SC = white". Thus, the minimal representation of this information system requires only four conditions (rather than the eight conditions that are presented in Table 2 with either column S or PC eliminated). This representation corresponds to the following set of *decision rules* which may be seen as classification model discovered in the data set contained in Table 2 (in the braces there are symbols of signs covered by the corresponding rule):

$$
\begin{aligned}
&\mathit{rule}\,\#1 : \mathit{if}\ \mathrm{S} = \text{triangle}, && \mathit{then}\ \mathrm{Class} = \mathrm{W} && \{a\}\\
&\mathit{rule}\,\#2 : \mathit{if}\ \mathrm{S} = \text{circle} && && \\
&\qquad \mathit{and}\ \ \mathrm{SC} = \text{red}, && \mathit{then}\ \mathrm{Class} = \mathrm{I} && \{b, c\}\\
&\mathit{rule}\,\#3 : \mathit{if}\ \mathrm{SC} = \text{white}, && \mathit{then}\ \mathrm{Class} = \mathrm{O} && \{d\}\,.
\end{aligned}
$$

This is not the only representation, because an alternative set of rules is:

$$
\begin{aligned}
&\mathit{rule}\,\#1' : \mathit{if}\ \mathrm{PC} = \text{yellow}, && \mathit{then}\ \mathrm{Class} = \mathrm{W} && \{a\}\\
&\mathit{rule}\,\#2' : \mathit{if}\ \mathrm{PC} = \text{white}, && \mathit{then}\ \mathrm{Class} = \mathrm{I} && \{b\}\\
&\mathit{rule}\,\#3' : \mathit{if}\ \mathrm{PC} = \text{blue}, && && \\
&\qquad \mathit{and}\ \ \mathrm{SC} = \text{red}, && \mathit{then}\ \mathrm{Class} = \mathrm{I} && \{c\}\\
&\mathit{rule}\,\#4' : \mathit{if}\ \mathrm{SC} = \text{white}, && \mathit{then}\ \mathrm{Class} = \mathrm{O} && \{d\}\,.
\end{aligned}
$$

It is interesting to come back to Table 1 and to ask what decision rules represent this information system. As the description of the four signs by S and PC is not sufficient to characterize exactly all the classes, it is not surprising that not all the rules will have a non-ambiguous decision. Indeed, the following decision rules can be induced:

$$
\begin{aligned}
&\mathit{rule}\,\#1'' : \mathit{if}\ \mathrm{S} = \text{triangle}, && \mathit{then}\ \mathrm{Class} = \mathrm{W} && \{a\}\\
&\mathit{rule}\,\#2'' : \mathit{if}\ \mathrm{PC} = \text{white}, && \mathit{then}\ \mathrm{Class} = \mathrm{I} && \{b\}\\
&\mathit{rule}\,\#3'' : \mathit{if}\ \mathrm{PC} = \text{blue}, && \mathit{then}\ \mathrm{Class} = \mathrm{I\ or\ O} && \{c, d\}\,.
\end{aligned}
$$

Note that these rules can be induced from the lower approximations of classes W and I, and from the set called the *boundary* of both I and O. Indeed, for exact rule #1″, the supporting example is in $\text{lower\_approx.}_{\mathrm{S,PC}}(\mathrm{W}) = \{a\}$; for exact rule #2″ it is in $\text{lower\_approx.}_{\mathrm{S,PC}}(\mathrm{I}) = \{b\}$; and the supporting examples for approximate rule #3″ are in the boundary of classes I and O, defined as:

$$
\begin{aligned}
&\text{boundary}_{\mathrm{S,PC}}(\mathrm{I})\\
&\quad = \text{upper\_approx.}_{\mathrm{S,PC}}(\mathrm{I}) - \text{lower\_approx.}_{\mathrm{S,PC}}(\mathrm{I})\\
&\quad = \{c, d\}\,,
\end{aligned}
$$

$$
\begin{aligned}
&\text{boundary}_{\mathrm{S,PC}}(\mathrm{O})\\
&\quad = \text{upper\_approx.}_{\mathrm{S,PC}}(\mathrm{O}) - \text{lower\_approx.}_{\mathrm{S,PC}}(\mathrm{O})\\
&\quad = \{c, d\}\,.
\end{aligned}
$$

As a result of the approximate characterization of classes W, I and O by S and PC, an approximate representation in terms of decision rules is obtained. Since the quality of the approximation is 1/2, exact rules (#1″ and #2″) cover one half of the examples and the other half is covered by the approximate rule (#3″). Now, the quality of approximation by S and SC, or by PC and SC, was equal to 1, so all examples were covered by exact rules (#1 to #3, or #1′ to #4′ respectively).

One can see, from this simple example, that the rough set analysis of data included in an information system provides some useful information. In particular, the following results are obtained:

- A characterization of decision classes in terms of chosen condition attributes through lower and upper approximation.
- A measure of the quality of approximation which indicates how good the chosen set of attributes is for approximation of the classification.
- The reducts of condition attributes including all relevant attributes. At the same time, superfluous and exchangeable attributes are also identified.
- The core composed of indispensable attributes.
- A set of decision rules which is induced from the lower and upper approximations of the decision classes. This constitutes a classification model for a given information system.

Dominance-Based Rough Set Approach to Ordinal Classification Problems

Dominance-Based Rough Set Approach (DRSA)

Dominance-based Rough Set Approach (DRSA) has been proposed by the authors to handle background knowledge about ordinal evaluations of objects from a universe, and about monotonic relationships between these evaluations, e. g. "the larger the mass and the smaller the distance, the larger the gravity" or "the greater the debt of a firm, the greater its risk of failure". Such a knowledge is typical for data describing various phenomena. It is also characteristic for data concerning multiple criteria decision or decision under uncertainty, where the order of value sets of condition and decision attributes corresponds to increasing or decreasing preference. In case of multiple criteria decision, the condition attributes are called *criteria*.

Let us consider a decision table including a finite universe of objects (solutions, alternatives, actions) U evaluated on a finite set of criteria $F = \{f_1, \ldots, f_n\}$, and on a single decision attribute d. The set of the indices of criteria is denoted by $I = \{1, \ldots, n\}$. Without loss of generality, $f_i : U \to \Re$ for each $i = 1, \ldots, n$, and, for all objects $x, y \in U$, $f_i(x) \geq f_i(y)$ means that "x is at least as good as y with respect to criterion f_i", which is denoted by $x \succeq_i y$. Therefore, it is supposed that $\succeq_i$ is a complete preorder, i. e. a strongly complete and transitive binary relation, defined on U on the basis of evaluations $f_i(\cdot)$. Furthermore, decision attribute d makes a partition of U into a finite number of decision classes, $\mathbf{Cl} = \{Cl_1, \ldots, Cl_m\}$, such that each $x \in U$ belongs to one and only one class Cl_t, $t = 1, \ldots, m$. It is assumed that the classes are preference ordered, i. e. for all $r, s = 1, \ldots, m$, such that $r > s$, the objects from Cl_r are preferred to the objects from Cl_s. More formally, if $\succeq$ is a *comprehensive weak preference relation* on U, i. e. if for all $x, y \in U$, $x \succeq y$ reads "x is at least as good as y", then it is supposed that

$$[x \in Cl_r,\ y \in Cl_s,\ r > s] \Rightarrow x \succ y\,,$$

where $x \succ y$ means $x \succeq y$ and *not* $y \succeq x$.

The above assumptions are typical for consideration of an ordinal classification (or multiple criteria sorting) problem. Indeed, the decision table characterized above includes examples of ordinal classification which constitute an input *preference information* to be analyzed using DRSA.

The sets to be approximated are called *upward union* and *downward union* of decision classes, respectively:

$$Cl_t^{\geq} = \bigcup_{s \geq t} Cl_s, \quad Cl_t^{\leq} = \bigcup_{s \leq t} Cl_s, \quad t = 1, \ldots, m\,.$$

The statement $x \in Cl_t^{\geq}$ reads "x belongs to at least class Cl_t", while $x \in Cl_t^{\leq}$ reads "x belongs to at most class Cl_t". Let us remark that $Cl_1^{\geq} = Cl_m^{\leq} = U$, $Cl_m^{\geq} = Cl_m$ and $Cl_1^{\leq} = Cl_1$. Furthermore, for $t = 2, \ldots, m$,

$$Cl_{t-1}^{\leq} = U - Cl_t^{\geq} \quad \text{and} \quad Cl_t^{\geq} = U - Cl_{t-1}^{\leq}\,.$$

The key idea of DRSA is representation (approximation) of upward and downward unions of decision classes, by *granules of knowledge* generated by criteria. These granules are *dominance cones* in the criteria values space.

x *dominates* y with respect to set of criteria $P \subseteq I$ (shortly, x *P-dominates* y), denoted by xD_Py, if for every criterion $i \in P$, $f_i(x) \geq f_i(y)$. The relation of P-dominance is reflexive and transitive, i. e. it is a partial preorder.

Given a set of criteria $P \subseteq I$ and $x \in U$, the granules of knowledge used for approximation in DRSA are:

- a set of objects dominating x, called *P-dominating set*, $D_P^+(x) = \{y \in U : yD_Px\}$,
- a set of objects dominated by x, called *P-dominated set*, $D_P^-(x) = \{y \in U : xD_Py\}$.

Let us recall that the dominance principle requires that an object x dominating object y on all considered criteria (i. e. x having evaluations at least as good as y on all considered criteria) should also dominate y on the decision (i. e. x should be assigned to at least as good decision class as y). Objects satisfying the dominance principle are called *consistent*, and those which are violating the dominance principle are called *inconsistent*.

The *P-lower approximation* of $Cl_t^{\geq}$, denoted by $\underline{P}\left(Cl_t^{\geq}\right)$, and the *P-upper approximation* of $Cl_t^{\geq}$, denoted by $\overline{P}\left(Cl_t^{\geq}\right)$, are defined as follows ($t = 2, \ldots, m$):

$$\underline{P}\left(Cl_t^{\geq}\right) = \left\{x \in U : D_P^+(x) \subseteq Cl_t^{\geq}\right\},$$
$$\overline{P}\left(Cl_t^{\geq}\right) = \left\{x \in U : D_P^-(x) \cap Cl_t^{\geq} \neq \emptyset\right\}\,.$$

Analogously, one can define the *P-lower approximation* and the *P-upper approximation* of $Cl_t^{\leq}$ as follows ($t = 1, \ldots, m-1$):

$$\underline{P}\left(Cl_t^{\leq}\right) = \left\{x \in U : D_P^-(x) \subseteq Cl_t^{\leq}\right\},$$
$$\overline{P}\left(Cl_t^{\leq}\right) = \left\{x \in U : D_P^+(x) \cap Cl_t^{\leq} \neq \emptyset\right\}\,.$$

The P-lower and P-upper approximations so defined satisfy the following *inclusion properties*, for all $P \subseteq I$:

$$\underline{P}\left(Cl_t^{\geq}\right) \subseteq Cl_t^{\geq} \subseteq \overline{P}\left(Cl_t^{\geq}\right), \quad t = 2, \ldots, m\,,$$

$$\underline{P}\left(Cl_t^{\leq}\right) \subseteq Cl_t^{\leq} \subseteq \overline{P}\left(Cl_t^{\leq}\right), \quad t = 1, \ldots, m-1\,.$$

The P-lower and P-upper approximations of $Cl_t^{\geq}$ and $Clt^{\leq}$ have an important *complementarity property*, according to which,

$$\underline{P}\left(Cl_t^{\geq}\right) = U - \overline{P}\left(Cl_{t-1}^{\leq}\right) \quad \text{and}$$
$$\overline{P}\left(Cl_t^{\geq}\right) = U - \underline{P}\left(Cl_{t-1}^{\leq}\right), \quad t = 2, \ldots, m,$$
$$\underline{P}\left(Cl_t^{\leq}\right) = U - \overline{P}\left(Cl_{t+1}^{\geq}\right) \quad \text{and}$$
$$\overline{P}\left(Cl_t^{\leq}\right) = U - \underline{P}\left(Cl_{t+1}^{\geq}\right), \quad t = 1, \ldots, m-1.$$

The *P-boundary* of $Cl_t^{\geq}$ and $Cl_t^{\leq}$, denoted by $Bn_P\left(Cl_t^{\geq}\right)$ and $Bn_P\left(Cl_t^{\leq}\right)$, respectively, are defined as follows:

$$Bn_P\left(Cl_t^{\geq}\right) = \overline{P}\left(Cl_t^{\geq}\right) - \underline{P}\left(Cl_t^{\geq}\right), \quad t = 2, \ldots, m,$$
$$Bn_P\left(Cl_t^{\leq}\right) = \overline{P}\left(Cl_t^{\leq}\right) - \underline{P}\left(Cl_t^{\leq}\right), \quad t = 1, \ldots, m-1.$$

Due to the above complementarity property, $Bn_P\left(Cl_t^{\geq}\right) = Bn_P\left(Cl_{t-1}^{\leq}\right)$, for $t = 2, \ldots, m$.

For every $P \subseteq C$, the *quality of approximation* of the ordinal classification $\boldsymbol{Cl}$ by a set of criteria P is defined as the ratio of the number of objects P-consistent with the dominance principle and the number of all the objects in U. Since the P-consistent objects are those which do not belong to any P-boundary $Bn_P(Cl_t^{\geq})$, $t = 2, \ldots, m$, or $Bn_P\left(Cl_t^{\leq}\right)$, $t = 1, \ldots, m-1$, the quality of approximation of the ordinal classification $\boldsymbol{Cl}$ by a set of criteria P, can be written as

$$\gamma_P(\boldsymbol{Cl}) = \frac{\left|U - \left(\bigcup_{t=2,\ldots,m} Bn_P\left(Cl_t^{\geq}\right)\right)\right|}{|U|} = \frac{\left|U - \left(\bigcup_{t=1,\ldots,m-1} Bn_P\left(Cl_t^{\leq}\right)\right)\right|}{|U|}.$$

$\gamma_P(\boldsymbol{Cl})$ can be seen as a degree of consistency of the objects from U, when P is the set of criteria and $\boldsymbol{Cl}$ is the considered ordinal classification.

Each minimal (with respect to inclusion) subset $P \subseteq C$ such that $\gamma_P(\boldsymbol{Cl}) = \gamma_C(\boldsymbol{Cl})$ is called a *reduct* of $\boldsymbol{Cl}$, and is denoted by $\text{RED}_{\boldsymbol{Cl}}$. Let us remark that for a given set U one can have more than one reduct. The intersection of all reducts is called the *core*, and is denoted by $\text{CORE}_{\boldsymbol{Cl}}$. Criteria in $\text{CORE}_{\boldsymbol{Cl}}$ cannot be removed from consideration without deteriorating the quality of approximation. This means that, in set C, there are three categories of criteria:

- *indispensable* criteria included in the core,
- *exchangeable* criteria included in some reducts, but not in the core,
- *redundant* criteria, neither indispensable nor exchangeable, and thus not included in any reduct.

The dominance-based rough approximations of upward and downward unions of decision classes can serve to induce a generalized description of objects in terms of "*if …, then …*" decision rules. For a given upward or downward union of classes, $Cl_t^{\geq}$ or $Cl_s^{\leq}$, the decision rules induced under a hypothesis that objects belonging to $\underline{P}(Cl_t^{\geq})$ or $\underline{P}\left(Cl_s^{\leq}\right)$ are positive examples, and all the others are negative, suggest a *certain* assignment to "class Cl_t or better", or to "class Cl_s or worse", respectively. On the other hand, the decision rules induced under a hypothesis that objects belonging to $\overline{P}\left(Cl_t^{\geq}\right)$ or $\overline{P}\left(Cl_s^{\leq}\right)$ are positive examples, and all the others are negative, suggest a *possible* assignment to "class Cl_t or better", or to "class Cl_s or worse", respectively. Finally, the decision rules induced under a hypothesis that objects belonging to the intersection $\overline{P}(Cl_s^{\leq}) \cap \overline{P}(Cl_t^{\geq})$ are positive examples, and all the others are negative, suggest an *approximate* assignment to some classes between Cl_s and $Cl_t (s < t)$.

In the case of preference ordered description of objects, set U is composed of examples of ordinal classification. Then, it is meaningful to consider the following five types of decision rules:

1) *certain $D_{\geq}$-decision rules*, providing lower profile descriptions for objects belonging to $\underline{P}(Cl_t^{\geq})$: if $f_{i_1}(x) \geq r_{i_1}$ and … and $f_{i_p}(x) \geq r_{i_p}$, then $x \in Cl_t^{\geq}$, $\{i_1, \ldots, i_p\} \subseteq I$, $t = 2, \ldots, m$, $r_{i_1}, \ldots, r_{i_p} \in \Re$;
2) *possible $D_{\geq}$-decision rules*, providing lower profile descriptions for objects belonging to $\overline{P}(Cl_t^{\geq})$: if $f_{i_1}(x) \geq r_{i_1}$ and … and $f_{i_p}(x) \geq r_{i_p}$, then x possibly belongs to $Cl_t^{\geq}$, $\{i_1, \ldots, i_p\} \subseteq I$, $t = 2, \ldots, m$, $r_{i_1}, \ldots, r_{i_p} \in \Re$;
3) *certain $D_{\leq}$-decision rules*, providing upper profile descriptions for objects belonging to $\underline{P}(Cl_t^{\leq})$: if $f_{i_1}(x) \leq r_{i_1}$ and … and $f_{i_p}(x) \leq r_{i_p}$, then $x \in Cl_t^{\leq}$, $\{i_1, \ldots, i_p\} \subseteq I$, $t = 1, \ldots, m-1$, $r_{i_1}, \ldots, r_{i_p} \in \Re$;
4) *possible $D_{\leq}$-decision rules*, providing upper profile descriptions for objects belonging to $\overline{P}(Cl_t^{\leq})$: if $f_{i_1}(x) \leq r_{i_1}$ and … and $f_{i_p}(x) \leq r_{i_p}$, then x possibly belongs to $Cl_t^{\leq}$, $\{i_1, \ldots, i_p\} \subseteq I$, $t = 1, \ldots, m-1$, $r_{i_1}, \ldots, r_{i_p} \in \Re$;
5) *approximate $D_{\geq\leq}$-decision rules*, providing simultaneously lower and upper profile descriptions for objects belonging to $Cl_s \cup Cl_{s+1} \cup \ldots \cup Cl_t$, without possibility of discerning to which class: if $f_{i_1}(x) \geq r_{i_1}$ and … and $f_{i_k}(x) \geq r_{i_k}$ and $f_{i_{k+1}}(x) \leq r_{i_{k+1}}$ and … and $f_{i_p}(x) \leq r_{i_p}$, then $x \in Cl_s \cup Cl_{s+1} \cup \ldots \cup Cl_t$, $\{i_1, \ldots, i_p\} \subseteq I$, $s, t \in \{1, \ldots, m\}$, $s < t$, $r_{i_1}, \ldots, r_{i_p} \in \Re$.

In the premise of a $D_{\geq\leq}$-decision rule, there can be "$f_i(x) \geq r_i$" and "$f_i(x) \leq r_i'$", where $r_i \leq r_i'$, for the same

$i \in I$. Moreover, if $r_i = r'_i$, the two conditions boil down to "$f_i(x) = r_i$".

Since a decision rule is a kind of implication, a *minimal* rule is understood as an implication such that there is no other implication with the premise of at least the same weakness (in other words, a rule using a subset of elementary conditions or/and weaker elementary conditions) and the conclusion of at least the same strength (in other words, a $D_{\geq}$- or a $D_{\leq}$-decision rule assigning objects to the same union or sub-union of classes, or a $D_{\geq\leq}$-decision rule assigning objects to the same or smaller set of classes).

The rules of type 1) and 3) represent certain knowledge extracted from data (examples of ordinal classification), while the rules of type 2) and 4) represent possible knowledge; the rules of type 5) represent doubtful knowledge, because they are supported by inconsistent objects only.

Moreover, the rules of type 1) and 3) are *exact* if they do not cover negative examples, and they are *probabilistic* otherwise. In the latter case, each rule is characterized by a *confidence ratio*, representing the probability that an object matching the premise of the rule also matches its conclusion.

Given a certain or possible $D_{\geq}$-decision rule $r \equiv$ "if $f_{i_1}(x) \geq r_{i_1}$ and … and $f_{i_p}(x) \geq r_{i_p}$, then $x \in Cl_t^{\geq}$", an object $y \in U$ *supports* r if $f_{i_1}(y) \geq r_{i_1}$ and … and $f_{i_p}(y) \geq r_{i_p}$. Moreover, object $y \in U$ supporting decision rule r is a *base* of r if $f_{i_1}(y) = r_{i_1}$ and … and $f_{i_p}(y) = r_{i_p}$. Similar definitions hold for certain or possible $D_{\leq}$-decision rules and approximate $D_{\geq\leq}$-decision rules. A decision rule having at least one base is called *robust*. Identification of supporting objects and bases of robust rules is important for interpretation of the rules in multiple criteria decision analysis perspective. The ratio of the number of objects supporting a rule and the number of all considered objects is called *relative support* of a rule. The relative support and the confidence ratio are basic characteristics of a rule, however, some *Bayesian confirmation measures* reflect much better the attractiveness of a rule [25].

A set of decision rules is *complete* if it covers all considered objects (examples of ordinal classification) in such a way that consistent objects are re-assigned to their original classes, and inconsistent objects are assigned to clusters of classes referring to this inconsistency. A set of decision rules is *minimal* if it is complete and non-redundant, i. e. exclusion of any rule from this set makes it incomplete.

Note that the syntax of decision rules induced from rough approximations defined using dominance cones, is using consistently this type of granules. Each condition profile defines a dominance cone in $\frac{p}{\uparrow V(p \leq n)}$-dimensional condition space $\Re^p (p \leq n)$, and each decision profile defines a dominance cone in one-dimensional decision space $\{1, \ldots, m\}$. In both cases, the cones are positive for $D_{\geq}$-rules and negative for $D_{\leq}$-rules.

Let also remark that dominance cones corresponding to condition profiles can originate in any point of $\Re^n$, without the risk of their being too specific. Thus, contrary to traditional granular computing, the condition space $\Re^n$ need not be discretized.

Procedures for rule induction from dominance-based rough approximations have been proposed in [17].

In [10], a new methodology for the induction of monotonic decision trees from dominance-based rough approximations of preference ordered decision classes has been proposed.

Application of DRSA to decision related problems goes far beyond ordinal classification. In [27], DRSA has been used for decision support involving multiple decision makers, and in [28], DRSA has been applied to case-based reasoning. The following sections present applications of DRSA to multiple criteria choice and ranking, to decision under uncertainty and to interactive multiobjective optimization. The surveys [24,26,29,51,52] include other applications of DRSA.

Example Illustrating DRSA in the Context of Ordinal Classification

This subsection presents a didactic example which illustrates the main concepts of DRSA. Let us consider the following ordinal classification problem. Students of a college must obtain an overall evaluation on the basis of their achievements in Mathematics, Physics and Literature. The three subjects are clearly criteria (condition attributes) and the comprehensive evaluation is a decision attribute. For simplicity, the value sets of the criteria and of the decision attribute are the same, and they are composed of three values: bad, medium and good. The preference order of these values is obvious. Thus, there are three preference ordered decision classes, so the problem belongs to the category of ordinal classification. In order to build a preference model of the jury, DRSA is used to analyze a set of exemplary evaluations of students provided by the jury. These examples of ordinal classification constitute an input preference information presented as decision table in Table 3.

Note that the dominance principle obviously applies to the examples of ordinal classification, since an improvement of a student's score on one of three criteria, with other scores unchanged, should not worsen the student's overall evaluation, but rather improve it.

Rough Sets in Decision Making, Table 3
Exemplary evaluations of students (examples of ordinal classification)

Student	Mathematics	Physics	Literature	Overall evaluation
*S*1	good	medium	bad	bad
*S*2	medium	medium	bad	medium
*S*3	medium	medium	medium	medium
*S*4	good	good	medium	good
*S*5	good	medium	good	good
*S*6	good	good	good	good
*S*7	bad	bad	bad	bad
*S*8	bad	bad	medium	bad

Rough Sets in Decision Making, Table 4
Exemplary evaluations of students excluding Literature

Student	Mathematics	Physics	Overall evaluation
*S*1	good	medium	bad
*S*2	medium	medium	medium
*S*3	medium	medium	medium
*S*4	good	good	good
*S*5	good	medium	good
*S*6	good	good	good
*S*7	bad	bad	bad
*S*8	bad	bad	bad

Observe that student *S*1 has not worse evaluations than student *S*2 on all the considered criteria, however, the overall evaluation of *S*1 is worse than the overall evaluation of *S*2. This contradicts the dominance principle, so the two examples of ordinal classification, and only those, are inconsistent. The quality of approximation of the ordinal classification represented by examples in Table 3 is equal to 0.75.

One can observe that in result of reducing the set of considered criteria, i. e. the set of considered subjects, some new inconsistencies can occur. For example, removing from Table 3 the evaluation on Literature, one obtains Table 4, where *S*1 is inconsistent not only with *S*2, but also with *S*3 and *S*5. In fact, student *S*1 has not worse evaluations than students *S*2, *S*3 and *S*5 on all the considered criteria (Mathematics and Physics), however, the overall evaluation of *S*1 is worse than the overall evaluation of *S*2, *S*3 and *S*5.

Observe, moreover, that removing from Table 3 the evaluations on Mathematics, one obtains Table 5, where no new inconsistencies occur, comparing to Table 3.

Similarly, after removing from Table 3 the evaluations

Rough Sets in Decision Making, Table 5
Exemplary evaluations of students excluding Mathematics

Student	Physics	Literature	Overall evaluation
*S*1	medium	bad	bad
*S*2	medium	bad	medium
*S*3	medium	medium	medium
*S*4	good	medium	good
*S*5	medium	good	good
*S*6	good	good	good
*S*7	bad	bad	bad
*S*8	bad	medium	bad

Rough Sets in Decision Making, Table 6
Exemplary evaluations of students excluding Physics

Student	Mathematics	Literature	Overall evaluation
*S*1	good	bad	bad
*S*2	medium	bad	medium
*S*3	medium	medium	medium
*S*4	good	medium	good
*S*5	good	good	good
*S*6	good	good	good
*S*7	bad	bad	bad
*S*8	bad	medium	bad

on Physics, one obtains Table 6, where no new inconsistencies occur, comparing to Table 3.

The fact that no new inconsistency occurs when Mathematics or Physics is removed, means that the subsets of criteria {Physics, Literature} or {Mathematics, Literature} contain sufficient information to represent the overall evaluation of students with the same quality of approximation as using the complete set of three criteria. This is not the case, however, for the subset {Mathematics, Physics}. Observe, moreover, that subsets {Physics, Literature} and {Mathematics, Literature} are minimal, because no other criterion can be removed without new inconsistencies occur. Thus, {Physics, Literature} and {Mathematics, Literature} are the reducts of the complete set of criteria {Mathematics, Physics, Literature}. Since Literature is the only criterion which cannot be removed from any reduct without introducing new inconsistencies, it constitutes the core, i. e. the set of indispensable criteria. The core is, of course, the intersection of all reducts, i. e. in our example:

$$\{\text{Literature}\} = \{\text{Physics, Literature}\} \cap \{\text{Mathematics, Literature}\}.$$

In order to illustrate in a simple way the concept of rough approximation, let us confine our anal-

ysis to the reduct {Mathematics, Literature}. Let us consider student $S4$. His positive dominance cone $D^{+}_{\{\text{Mathematics, Literature}\}}(S4)$ is composed of all the students having evaluations not worse than him on Mathematics and Literature, i. e. of all the students dominating him with respect to Mathematics and Literature. Thus,

$$D^{+}_{\{\text{Mathematics, Literature}\}}(S4) = \{S4, S5, S6\} .$$

On the other hand, the negative dominance cone of student $S4$, $D^{-}_{\{\text{Mathematics, Literature}\}}(S4)$, is composed of all the students having evaluations not better than him on Mathematics and Literature, i. e. of all the students dominated by him with respect to Mathematics and Literature. Thus,

$$D^{-}_{\{\text{Mathematics, Literature}\}}(S4) = \{S1, S2, S3, S4, S7, S8\} .$$

Similar dominance cones can be obtained for all the students from Table 6. For example, for $S2$, the dominance cones are

$$D^{+}_{\{\text{Mathematics, Literature}\}}(S2) = \{S1, S2, S3, S4, S5, S6\}$$

and

$$D^{-}_{\{\text{Mathematics, Literature}\}}(S2) = \{S2, S7\} .$$

The rough approximations can be calculated using dominance cones. Let us consider, for example, the lower approximation of the set of students having a "good" overall evaluation $\underline{P}(Cl^{\geq}_{\text{good}})$, with $P = \{\text{Mathematics, Literature}\}$. Notice that $\underline{P}(Cl^{\geq}_{\text{good}}) = \{S4, S5, S6\}$, because positive dominance cones of students $S4$, $S5$ and $S6$ are all included in the set of students with an overall evaluation "good". In other words, this means that there is no student dominating $S4$ or $S5$ or $S6$ while having an overall evaluation worse than "good". From the viewpoint of decision making, this means that, taking into account the available information about evaluation of students on Mathematics and Literature, the fact that student y dominates $S4$ or $S5$ or $S6$ is a *sufficient* condition to conclude that y is a "good" student.

The upper approximation of the set of students with a "good" overall evaluation is $\overline{P}(Cl^{\geq}_{\text{good}}) = \{S4, S5, S6\}$, because negative dominance cones of students $S4$, $S5$ and $S6$ have a nonempty intersection with the set of students having a "good" overall evaluation. In other words, this means that for each one of the students $S4$, $S5$ and $S6$, there is at least one student dominated by him with an overall evaluation "good". From the point of view of decision making, this means that, taking into account the available information about evaluation of students on Mathematics and Literature, the fact that student y dominates $S4$ or $S5$ or $S6$ is a *possible* condition to conclude that y is a "good" student.

Let us observe that for the set of criteria $P = \{\text{Mathematics, Literature}\}$, the lower and upper approximations of the set of "good" students are the same. This means that examples of ordinal classification concerning this decision class are all consistent. This is not the case, however, for the examples concerning the union of decision classes "at least medium". For this upward union the rough approximations are $\underline{P}(Cl^{\geq}_{\text{medium}}) = \{S3, S4, S5, S6\}$ and $\overline{P}(Cl^{\geq}_{\text{medium}}) = \{S1, S2, S3, S4, S5, S6\}$. The difference between $\overline{P}(Cl^{\geq}_{\text{medium}})$ and $\underline{P}(Cl^{\geq}_{\text{medium}})$, i. e. the boundary $Bn_P(Cl^{\geq}_{\text{medium}}) = \{S1, S2\}$, is composed of students with inconsistent overall evaluations, which has already been noticed above. From the viewpoint of decision making, this means that, taking into account the available information about evaluation of students on Mathematics and Literature, the fact that student y is dominated by $S1$ and dominates $S2$ is a condition to conclude that y can obtain an overall evaluation "at least medium" with some doubts.

Until now, rough approximations of only upward unions of decision classes have been considered. It is interesting, however, to calculate also rough approximations of downward unions of decision classes. Let us consider first the lower approximation of the set of students having "at most medium" overall evaluation $\underline{P}(Cl^{\leq}_{\text{medium}})$. Observe that $\underline{P}(Cl^{\leq}_{\text{medium}}) = \{S1, S2, S3, S7, S8\}$, because the negative dominance cones of students $S1$, $S2$, $S3$, $S7$, and $S8$ are all included in the set of students with overall evaluation "at most medium". In other words, this means that there is no student dominated by $S1$ or $S2$ or $S3$ or $S7$ or $S8$ while having an overall evaluation better than "medium". From the viewpoint of decision making, this means that, taking into account the available information about evaluation of students on Mathematics and Literature, the fact that student y is dominated by $S1$ or $S2$ or $S3$ or $S7$ or $S8$ is a *sufficient* condition to conclude that y is an "at most medium" student.

The upper approximation of the set of students with an "at most medium" overall evaluation is $\overline{P}(Cl^{\leq}_{\text{medium}}) = \{S1, S2, S3, S7, S8\}$, because the positive dominance cones of students $S1$, $S2$, $S3$, $S7$, and $S8$ have a nonempty intersection with the set of students having an "at most medium" overall evaluation. In other words, this means that for each one of the students $S1$, $S2$, $S3$, $S7$, and $S8$, there is at least one student dominating him with an overall evaluation "at most medium". From the viewpoint of decision making, this means that, taking into account

the available information about evaluation of students on Mathematics and Literature, the fact that student y is dominated by $S1$ or $S2$ or $S3$ or $S7$ or $S8$ is a *possible* condition to conclude that y is an "at most medium" student.

Finally, the lower and upper approximations of the set of students having a "bad" overall evaluation are $\underline{P}(Cl^{\leq}_{\text{bad}}) = \{S7, S8\}$ and $\overline{P}(Cl^{\leq}_{\text{bad}}) = \{S1, S2, S7, S8\}$. The difference between $\overline{P}(Cl^{\leq}_{\text{bad}})$ and $\underline{P}(Cl^{\leq}_{\text{bad}})$, i. e. the boundary $Bn_P(Cl^{\leq}_{\text{bad}}) = \{S1, S2\}$ is composed of students with inconsistent overall evaluations, which has already been noticed above. From the point of view of decision making, this means that, taking into account the available information about evaluation of students on Mathematics and Literature, the fact that student y is dominated by $S1$ and dominates $S2$ is a condition to conclude that y can obtain an overall evaluation "bad" with some doubts. Observe, moreover, that $Bn_P(Cl^{\geq}_{\text{medium}}) = Bn_P(Cl^{\leq}_{\text{bad}}) = \{S1, S2\}$.

Given the above rough approximations with respect to the set of criteria $P = \{\text{Mathematics, Literature}\}$, one can induce a set of decision rules representing the preferences of the jury. The idea is that evaluation profiles of students belonging to the lower approximations can serve as a base for some certain rules, while evaluation profiles of students belonging to the boundaries can serve as a base for some approximate rules. The following decision rules have been induced (between parentheses there are id's of students supporting the corresponding rule; the student being a rule base is underlined):

rule 1) if the evaluation on Mathematics is (at least) good, and the evaluation on Literature is at least medium, then the overall evaluation is (at least) good, {$\underline{S4}$, $S5$, $S6$},

rule 2) if the evaluation on Mathematics is at least medium, and the evaluation on Literature is at least medium, then the overall evaluation is at least medium, {$\underline{S3}$, $S4$, $S5$, $S6$},

rule 3) if the evaluation on Mathematics is at least medium, and the evaluation on Literature is (at most) bad, then the overall evaluation is bad or medium, {$S1$, $\underline{S2}$},

rule 4) if the evaluation on Mathematics is at least medium, then the overall evaluation is at least medium, {$\underline{S2, S3}$, $S4$, $S5$, $S6$},

rule 5) if the evaluation on Literature is (at most) bad, then the overall evaluation is at most medium, {$\underline{S1, S2, S7}$},

rule 6) if the evaluation on Mathematics is (at most) bad, then the overall evaluation is (at most) bad, {$\underline{S7, S8}$}.

Notice that rules 1)–2), 4)–7) are certain, while rule 3) is an approximate one. These rules represent knowledge discovered from the available information. In the current context, the knowledge is interpreted as a preference model of the jury. A characteristic feature of the syntax of decision rules representing preferences is the use of expressions "at least" or "at most" a value; in case of extreme values ("good" and "bad"), these expressions are put in parentheses because there is no value above "good" and below "bad".

Even if one can represent all the knowledge using only one reduct of the set of criteria (as it is the case using $P = \{\text{Mathematics, Literature}\}$), when considering a larger set of criteria than a reduct, one can obtain a more synthetic representation of knowledge, i. e. the number of decision rules or the number of elementary conditions, or both of them, can get smaller. For example, considering the set of all three criteria, {Mathematics, Physics, Literature}, one can induce a set of decision rules composed of the above rules 1), 2), 3) and 6), plus the following:

rule 7) if the evaluation on Physics is at most medium, and the evaluation on Literature is at most medium, then the overall evaluation is at most medium, {$S1$, $S2$, $\underline{S3}$, $S7$, $S8$}.

Thus, the complete set of decision rules induced from Table 3 is composed of 5 instead of 6 rules.

Once accepted by the DM, these rules represent his/her preference model. Assuming that rules 1)–7) in our example represent the preference model of the jury, it can be used to evaluate new students. For example, student $S9$ who is "medium" in Mathematics and Physics and "good" in Literature, would be evaluated as "medium" because his profile matches the premise of rule 2), having as consequence an overall evaluation at least "medium". The overall evaluation of $S9$ cannot be "good", because his profile does not match any rule having as consequence an overall evaluation "good" (in the considered example, the only rule of this type is rule 1) whose premise is not matched by the profile of $S9$.

DRSA on a Pairwise Comparison Table for Multiple Criteria Choice and Ranking Problems

Multiple Criteria Choice and Ranking Problems

Ordinal classification decisions are based on absolute evaluation of objects on multiple criteria, however, multiple criteria choice and ranking decisions are based on preference relations between objects. Decision table including examples of ordinal classification does not contain information about preference relations between objects, thus, in order to apply DRSA to multiple criteria choice and ranking problems, a different representation of the input preference information is needed.

To handle binary relations within the rough set approach, it has been proposed in [14,15] to operate on, so-called, *pairwise comparison table* (PCT), i. e. a decision table including pairs of objects for which multiple criteria evaluations and a comprehensive preference relation are known. PCT represents preference information provided by the DM in the form of decision examples (pairwise comparisons of objects).

Similarly to ordinal classification, decision examples concerning multiple criteria choice and ranking may be inconsistent with respect to the dominance principle, however, interpretation of the inconsistency is different: It occurs when preferences of a pair of objects, say (a, b), on all considered criteria are not weaker than preferences of another pair of objects, say (c, d), on these criteria, however, the comprehensive preference of object a over object b is weaker than the comprehensive preference of object c over object d.

The Pairwise Comparison Table

Similarly to the ordinal classification, let us consider a finite set of criteria $F = \{f_1, \ldots, f_n\}$, the set of their indices $I = \{1, \ldots, n\}$ and a finite universe of objects (actions, solutions, alternatives) U. For any criterion $i \in I$, let T_i be a finite set of binary relations defined on U on the basis of the evaluations of objects from U with respect to the considered criterion i, such that $\forall (x, y) \in U \times U$ exactly one binary relation $t \in T_i$ is verified; $t \in T_i$ has the meaning of a *preference relation* for a pair of objects on a particular criterion i. More precisely, given value set V_i of $i \in I$, if $v_i', v_i'' \in V_i$ are the respective evaluations of $x, y \in U$ on criterion i, and $(x, y) \in t$, with $t \in T_i$, then for each $w, z \in U$ having the same evaluations v_i', v_i'' on i, $(w, z) \in t$. For interesting applications it should be $|T_q| \geq 2$, $\forall i \in I$. Furthermore, let T_d be a set of binary relations defined on U, such that at most one binary relation $t \in T_d$ is verified $\forall (x, y) \in U \times U$; $t \in T_d$ has the meaning of a *comprehensive preference relation* for a pair of objects (comprehensive pairwise comparison).

The preference information provided by the DM, has the form of pairwise comparisons of some reference objects from $B \subseteq U$. These decision examples are presented in the pairwise comparison table (PCT), defined as information table $\boldsymbol{S}_{\text{PCT}} = \langle \hat{B}, F \cup \{d\}, T_F \cup T_d, g \rangle$, where $\hat{B} \subseteq B \times B$ is a non-empty set of *exemplary pairwise comparisons of reference objects*, $T_F = \bigcup_{i \in I} T_i, d$ is a decision corresponding to the comprehensive pairwise comparison (comprehensive preference relation), and $g\colon \hat{B} \times (F \cup \{d\}) \rightarrow T_F \cup T_d$ is a total function such that $g[(x, y), i] \in T_i$, $\forall (x, y) \in U \times U$ and $\forall i \in I$, and $g[(x, y), d] \in T_d$, $\forall (x, y) \in \hat{B}$. It follows that for any pair of reference objects $(x, y) \in \hat{B}$ there is verified one and only one binary relation $t \in T_d$. Thus, T_d induces a partition of $\hat{B}$. In fact, information table $\boldsymbol{S}_{\text{PCT}}$ can be seen as decision table, since the set of considered criteria F and decision d are distinguished.

It is assumed that the exemplary pairwise comparisons provided by the DM can be represented in terms of *graded preference relations* (for example "very weak preference", "weak preference", "strict preference", "strong preference", "very strong preference") $P_i^h\colon \forall i \in I$ and $\forall (x, y) \in U \times U$,

$$T_i = \left\{ P_i^h, h \in H_i \right\},$$

where H_i is a particular subset of the relative integers and

- $xP_i^h y, h > 0$, means that object x is preferred to object y by degree h with respect to criterion i,
- $xP_i^h y, h < 0$, means that object x is not preferred to object y by degree h with respect to criterion i,
- $xP_i^0 y$ means that object x is similar (asymmetrically indifferent) to object y with respect to criterion i.

Of course, $\forall i \in I$ and $\forall (x, y) \in U \times U$, it holds:

$$\left[xP_i^h y, h > 0 \right] \Rightarrow \left[yP_i^k x, k \leq 0 \right].$$

The set of binary relations T_d may be defined in a similar way, but $xP_d^h y$ means that object x is comprehensively preferred to object y by degree h.

Technically, the modeling of the binary relation P_i^h, i. e. the assessment of h, can be organized as follows:

- first, it is observed that criterion i is a function $f_i\colon U \rightarrow \Re$ increasing with respect to the preferences on i, for each $i = 1, \ldots, n$,
- then, for each $i = 1, \ldots, n$, it is possible to define a function $k_i\colon \Re^2 \rightarrow \Re$ which measures the *strength of the preference* (positive or negative) of x over y (e. g. $k_i[f_i(x), f_i(y)] = f_i(x) - f_i(y)$); it should satisfy the following properties $\forall x, y, z \in U$:
  - i) $f_i(x) > f_i(y) \Leftrightarrow k_i[f_i(x), f_i(z)] > k_i[f_i(y), f_i(z)]$,
  - ii) $f_i(x) > f_i(y) \Leftrightarrow k_i[f_i(z), f_i(x)] < k_i[f_i(z), f_i(y)]$,
  - iii) $f_i(x) = f_i(y) \Leftrightarrow k_i[f_i(x), f_i(y)] = 0$,
- next, the domain of k_i can be divided into intervals, using a suitable set of thresholds Δ_i, $\forall i \in I$; these intervals are numbered in such a way that $k_i[f_i(x), f_i(y)] = 0$ belongs to interval no. 0,

- the value of h in the relation $xP_i^h y$ is then equal to the number of the interval including $k_i[f_i(x), f_i(y)]$, for any $(x, y) \in U \times U$.

Actually, property iii) can be relaxed in order to obtain a more general preference model which, for instance, does not satisfy preferential independence.

To simplify the presentation, let us consider a PCT where the set T_d is composed of two binary relations defined on U:

- *x outranks y* (denotation xSy or $(x, y) \in S$), where $(x, y) \in \hat{B}$,
- *x does not outrank y* (denotation $xS^c y$ or $(x, y) \in S^c$), where $(x, y) \in \hat{B}$,

and $S \cup S^c = \hat{B}$, where "x outranks y" means "x is at least as good as y"; observe that the binary relation S is reflexive, but neither necessarily transitive nor complete. In [8], a more general PCT was considered, where the set T_d is composed of multi-graded binary relations defined on U.

Approximation by Means of Graded Dominance Relations

Let $H_P = \bigcap_{i \in P} H_i$, $\forall P \subseteq I$. Given $P \subseteq I$ and $h \in H_P$, $\forall (x, y) \in U \times U$ it is said that *x positively dominates y* by degree h with respect to criteria from P iff $xP_i^{c_i} y$ with $c_i \geq h$, $\forall i \in P$. Analogously, $\forall (x, y) \in U \times U$, *x negatively dominates y* by degree h with respect to criteria from P iff $xP_i^{c_i} y$ with $c_i \leq h$, $\forall i \in P$. Therefore, each $P \subseteq I$ and $h \in H_P$ generate two binary relations (possibly empty) on U, called *positive P-dominance by degree h* (denotation D_{+P}^h) and *negative P-dominance by degree h* (denotation D_{-P}^h), respectively. They satisfy the following conditions:

(P1) if $(x, y) \in D_{+P}^h$, then $(x, y) \in D_{+R}^h$ for each $R \subseteq P$ and $k \leq h$;
(P2) if $(x, y) \in D_{-P}^h$, then $(x, y) \in D_{-R}^h$ for each $R \subseteq P$ and $k \geq h$.

In [15], it has been proposed to approximate the outranking relation S by means of the dominance relation D_{+P}^h. Therefore, S is considered a *rough binary relation*.

The P-lower approximation of S (denotation $\underline{P}(S)$) and the P-upper approximation of S (denotation $\overline{P}(S)$) are defined, respectively, as:

$$\underline{P}(S) = \bigcup_{h \in H_P} \left\{ \left(D_{+P}^h \cap \hat{B} \right) \subseteq S \right\},$$

$$\overline{P}(S) = \bigcap_{h \in H_P} \left\{ \left(D_{+P}^h \cap \hat{B} \right) \supseteq S \right\}.$$

$\underline{P}(S)$ may be interpreted as the dominance relation D_{+P}^h having the largest intersection with $\hat{B}$ included in the outranking relation S, and $\overline{P}(S)$ as the dominance relation D_{+P}^h including S and having the smallest intersection with $\hat{B}$.

Analogously, it is possible to approximate the relation S^c by means of the dominance relation D_{-P}^h. Observe that, in general, the definitions of the approximations of S and S^c do not satisfy the condition of complementarity, i. e. it is not true, in general, that $\underline{P}(S)$ is equal to $\hat{B} - \overline{P}(S^c)$ and that $\underline{P}(S^c)$ is equal to $\hat{B} - \overline{P}(S)$. This is because S and S^c are approximated using two different relations, D_{+P}^h and D_{-P}^h, respectively. Nevertheless, the approximations thus obtained constitute a good basis for the generation of simple decision rules.

Decision Rules

It is possible to represent preferences of the DM revealed in terms of exemplary pairwise comparisons contained in a given PCT, using decision rules. Since approximations of S and S^c were made using graded dominance relations, it is possible to induce decision rules being propositions of the following type:

- D_{++}-decision rule: *if* $xD_{+P}^h y$, *then* xSy,
- D_{+-}-decision rule: *if not* $xD_{+P}^h y$, *then* $xS^c y$,
- D_{-+}-decision rule: *if not* $xD_{-P}^h y$, *then* xSy,
- D_{--}-decision rule: *if* $xD_{-P}^h y$, *then* $xS^c y$,

where P is a non-empty subset of I. Therefore, for example, a D_{++}– decision rule is a proposition of the type: "*if* x positively dominates y by degree h with respect to criteria from P, *then* x outranks y".

A constructive definition of these rules may be given, being a kind of implication supported by the existence of at least one pair of objects from $\hat{B}$ satisfying one of the four propositions listed above, and by the absence of pairs from $\hat{B}$ contradicting it. Thus, for example, if

- there exists at least one pair $(w, z) \in \hat{B}$ such that $wD_{+P}^h z$ and wSz and
- there does not exist any pair $(v, u) \in \hat{B}$ such that $vD_{+P}^h u$ and $vS^c u$,
- then "*if* $xD_{+P}^h y$, *then* xSy" is accepted as a D_{++}– decision rule.

A D_{++}-decision rule "*if* $xD_{+P}^h y$, *then* xSy" is said to be *minimal* if there does not exist any other rule "*if* $xD_{+R}^k y$, *then* xSy" such that $R \subseteq P$ and $k \leq h$. Analogous definitions hold for the other cases. In other words, a minimal decision rule is a kind of implication for which there is

no other implication whose premise is of at least the same weakness and whose consequence is of at least the same strength.

The following results show connections of the decision rules with the P-lower and P-upper approximations of S and S^c [15]:

- D_{++}-minimal decision rule "*if* $xD^h_{+P}y$, *then* xSy" is supported by pairs of objects belonging to $\underline{P}(S) = D^h_{+P} \cap \hat{B}$,
- D_{--}-minimal decision rule "*if* $xD^h_{-P}y$, *then* xS^cy" is supported by pairs of objects belonging to $\underline{P}(S^c) = D^h_{-P} \cap \hat{B}$,
- D_{+-}-minimal decision rule "*if not* $xD^h_{+P}y$, *then* xS^cy" is supported by pairs of objects belonging to $\overline{P}(S) = D^h_{+P} \cap \hat{B}$,
- D_{-+}-minimal decision rule "*if not* $xD^h_{-P}y$, *then* xSy" is supported by pairs of objects belonging to $\overline{P}(S^c) = D^h_{-P} \cap \hat{B}$.

Application of the Decision Rules and Final Recommendation

In order to obtain a recommendation in the multiple criteria choice or ranking problems with respect to a set of objects $M \subseteq U$, the decision rules induced from the approximations of S and S^c (defined with respect to reference objects from B) should be applied on set $M \times M$. The application of the rules to any pair of objects $(u, v) \in M \times M$ establishes the presence (uSv) or the absence (uS^cv) of outranking with respect to (u, v). More precisely,

- from D_{++}-decision rule "*if* $xD^h_{+P}y$ *then* xSy" and from $uD^h_{+P}v$, one concludes uSv,
- from D_{+-}-decision rule "*if not* $xD^h_{+P}y$ *then* xS^cy" and from *not* $uD^h_{+P}v$, one concludes uS^cy,
- from D_{-+}-decision rule "*if not* $xD^h_{-P}y$, *then* xSy" and from *not* $uD^h_{-P}v$, one concludes uSv,
- from D_{--}-decision rule "*if* $xD^h_{-P}y$, *then* xS^cy" and from $uD^h_{-P}v$, one concludes uS^cv.

After the application of the decision rules to each pair of objects $(u, v) \in M \times M$, one of the following four situations may occur:

- uSv and *not* uS^cv, that is *true* outranking (denotation uS^Tv),
- uS^cv and *not* uSv, that is *false* outranking (denotation uS^Fv),
- uSv and uS^cv, that is *contradictory* outranking (denotation uS^Kv),
- *not* uSv and *not* uS^cv, that is *unknown* outranking (denotation uS^Uv).

The four above situations, which together constitute the so-called four-valued outranking (see [56]), have been introduced to underline the *presence* and the *absence* of *positive* and *negative* reasons for the outranking. Moreover, they make it possible to distinguish contradictory situations from unknown ones.

The following theorem underlines the operational importance of the minimal decision rules [15]: The application of *all* the decision rules obtained for a given $\mathbf{S}_{\mathrm{PCT}}$ to a pair of objects $(u, v) \in M \times M$ results in the same outranking relations S and S^c as those obtained from the application of the *minimal* decision rules *only*. Therefore, the set of the minimal decision rules totally characterizes the preferences of the DM contained in $\mathbf{S}_{\mathrm{PCT}}$.

A final *recommendation* can be obtained upon a suitable exploitation of the presence and the absence of outranking S and S^c on M. A possible exploitation procedure consists in calculating a specific score, called *Net Flow Score*, for each object $x \in M$:

$$S_{\mathrm{nf}}(x) = S^{++}(x) - S^{+-}(x) + S^{-+}(x) - S^{--}(x)\,,$$

where

$$\begin{aligned}
S^{++}(x) &= |\{y \in M\colon \text{ there is at least one decision rule which affirms } xSy\}|\,,\\
S^{+-}(x) &= |\{y \in M\colon \text{ there is at least one decision rule which affirms } ySx\}|\,,\\
S^{-+}(x) &= |\{y \in M\colon \text{ there is at least one decision rule which affirms } yS^cx\}|\,,\\
S^{--}(x) &= |\{y \in M\colon \text{ there is at least one decision rule which affirms } xS^cy\}|\,.
\end{aligned}$$

The recommendation in multiple criteria ranking problems consists of the total preorder determined by $S_{\mathrm{nf}}(x)$ on M; in multiple criteria choice problems it consists of the object(s) $x^* \in M$ such that $S_{\mathrm{nf}}(x^*) = \max_{x \in M} S_{\mathrm{nf}}(x)$.

The procedure described above has been characterized with reference to a number of desirable properties in [13].

Approximation by Means of Multi-graded Dominance Relations

The graded dominance relation introduced above assumes a common grade of preference for all the considered criteria. While this permits a simple calculation of the approximations and of the resulting decision rules, it is lacking in precision. A dominance relation allowing a different

degree of preference for each considered criterion (multi-graded dominance) gives a far more accurate picture of the preference information contained in the pairwise comparison table $\boldsymbol{S}_{\mathrm{PCT}}$ [14,16,18].

More formally, given $P \subseteq I\,(P \neq \emptyset), (x, y), (w, z) \in U \times U, (x, y)$ is said to *dominate* (w, z) with respect to criteria from P (denotation $(x, y)D_P(w, z)$) if x is preferred to y at least as strongly as w is preferred to z with respect to each $i \in P$. Precisely, "at least as strongly as" means "by at least the same degree", i. e. $hi \geq ki$, where $hi, ki \in H_i$, $xP_i^{hi}y$ and $wP_i^{ki}z, \forall i \in P$. Let $D_{\{i\}}$ be the dominance relation confined to the single criterion $i \in P$. The binary relation $D_{\{i\}}$ is reflexive ($(x, y)D_{\{i\}}(x, y), \forall (x, y) \in U \times U$), transitive ($(x, y)D_{\{i\}}(w, z)$ and $(w, z)D_{\{i\}}(u, v)$ imply $(x, y)D_{\{i\}}(u, v), \forall (x, y), (w, z), (u, v) \in U \times U$), and complete ($(x, y)D_{\{i\}}(w, z)$ or $(w, z)D_{\{i\}}(x, y), \forall (x, y), (w, z) \in U \times U$). Therefore, $D_{\{i\}}$ is a complete preorder on $U \times U$. Since the intersection of complete preorders is a partial preorder and $D_P = \bigcap_{i \in P} D_{\{i\}}, P \subseteq I$, then the dominance relation D_P is a partial preorder on $U \times U$.

Let $R \subseteq P \subseteq I$ and $(x, y), (u, v) \in U \times U$; then the following implication holds:

$$(x, y)D_P(u, v) \Rightarrow (x, y)D_R(u, v)\,.$$

Given $P \subseteq I$ and $(x, y) \in U \times U$, let us introduce the positive dominance set (denotation $D_P^+(x, y)$) and the negative dominance set (denotation $D_P^-(x, y)$):

$$D_P^+(x, y)) = \{(w, z) \in U \times U\colon (w, z)D_P(x, y)\}\,,$$

$$D_P^-(x, y)) = \{(w, z) \in U \times U\colon (x, y)D_P(w, z)\}\,.$$

Using the dominance relation D_P, it is possible to define P-lower and P-upper approximations of the outranking relation S with respect to $P \subseteq I$, respectively, as:

$$\underline{P}(S) = \left\{(x, y) \in \hat{B}\colon D_P^+(x, y) \subseteq S\right\}\,,$$

$$\overline{P}(S) = \bigcup_{(x,y)\in S} D_P^+(x, y)\,.$$

Analogously, it is possible to define the approximations of S^c:

$$\underline{P}(S^c) = \left\{(x, y) \in \hat{B}\colon D_P^-(x, y) \subseteq S^c\right\}\,,$$

$$\overline{P}(S^c) = \bigcup_{(x,y)\in S^c} D_P^-(x, y)\,.$$

It may be proved that

$$\underline{P}(S) \subseteq S \subseteq \overline{P}(S)$$

$$\underline{P}(S^c) \subseteq S^c \subseteq \overline{P}(S^c)\,.$$

Furthermore, the following complementarity properties hold:

$$\underline{P}(S) = \hat{B} - \overline{P}(S^c), \quad \overline{P}(S) = \hat{B} - \underline{P}(S^c)\,,$$

$$\underline{P}(S^c) = \hat{B} - \overline{P}(S), \quad \overline{P}(S^c) = \hat{B} - \underline{P}(S)\,.$$

The P-boundaries (P-doubtful regions) of S and S^c are defined as

$$Bn_P(S) = \overline{P}(S) - \underline{P}(S), \quad Bn_P(S^c) = \overline{P}(S^c) - \underline{P}(S^c)\,.$$

It is easy to prove that $Bn_P(S) = Bn_P(S^c)$.

The concepts of quality of approximation, reducts and core can be extended also to the approximation of the outranking relation by multi-graded dominance relations. In particular,

$$\gamma_P = \frac{|\underline{P}(S) \cup \underline{P}(S^c)|}{|\hat{B}|}$$

defines the *quality of approximation* of S and S^c by $P \subseteq I$. It expresses the ratio of all pairs of objects $(x, y) \in \hat{B}$ correctly assigned to S and S^c by the set P of criteria, to all the pairs of objects contained in $\hat{B}$. Each minimal subset $P' \subseteq P$ such that $\gamma_{P'} = \gamma_P$ is called a *reduct* of P (denotation $RED_{PCT}(P)$). Let us remark that $\boldsymbol{S}_{\mathrm{PCT}}$ can have more than one reduct. The intersection of all reducts is called the *core* (denotation $CORE_{PCT}(P)$).

Using the approximations defined above, it is then possible to induce a generalized description of the preference information contained in a given $\boldsymbol{S}_{\mathrm{PCT}}$ in terms of suitable decision rules. The syntax of these rules is based on the concept of upward cumulated preferences (denotation $P_i^{\geq h}$) and downward cumulated preferences (denotation $P_i^{\geq h}$), having the following interpretation:

- $xP_i^{\geq h}y$ means "x is preferred to y with respect to i by at least degree h",
- $xP_i^{\leq h}y$ means "x is preferred to y with respect to i by at most degree h".

Exact definition of the cumulated preferences, for each $(x, y) \in U \times U, i \in I$ and $h \in H_i$, is the following:

- $xP_i^{\geq h}y$ if $xP_i^k y$, where $k \in H_i$ and $k \geq h$,
- $xP_i^{\leq h}y$ if $xP_i^k y$, where $k \in H_i$ and $k \leq h$.

Using the above concepts, three types of decision rules can be obtained:

1. $D_{\geq}$ – decision rules, being statements of the type:

$$\text{if } xP_{i1}^{\geq h(i1)}y \text{ and } xP_{i2}^{\geq h(i2)}y \text{ and } \ldots xP_{ip}^{\geq h(ip)}y, \text{ then } xSy,$$

where $P = \{i1, \ldots, ip\} \subseteq I$ and $(h(i1), \ldots, h(ip)) \in H_{i1} \times \cdots \times H_{ip}$; these rules are supported by pairs of objects from the P-lower approximation of S only;

2. $D_{\leq}$ – decision rules, being statements of the type:

$$\text{if } xP_{i1}^{\leq h(i1)}y \text{ and } xP_{i2}^{\leq h(i2)}y \text{ and } \ldots xP_{ip}^{\leq h(ip)}y, \text{ then } xS^c y,$$

where $P = \{i1, \ldots, ip\} \subseteq I$ and $(h(i1), \ldots, h(ip)) \in H_{i1} \times \cdots \times H_{ip}$; these rules are supported by pairs of objects from the P-lower approximation of S^c only;

3. $D_{\geq\leq}$ – decision rules, being statements of the type:

$$\text{if } xP_{i1}^{\geq h(i1)}y \text{ and } xP_{i2}^{\geq h(i2)}y \text{ and } \ldots xP_{ik}^{\geq h(ik)}y \text{ and } xP_{ik+1}^{\leq h(ik+1)}y \text{ and } \ldots xP_{ip}^{\leq h(ip)}y, \text{ then } xSy \text{ or } xS^c y,$$

where $O' = \{i1, \ldots, ik\} \subseteq I$, $O'' = \{ik+1, \ldots, ip\} \subseteq I$, $P = O' \cup O''$, O' and O'' not necessarily disjoint, $(h(i1), \ldots, h(ip)) \in H_{i1} \times H_{i2} \cdots \times H_{ip}$; these rules are supported by objects from the P-boundary of S and S^c only.

Dominance Without Degrees of Preference

The degree of graded preference considered in Subsect. "The Pairwise Comparison Table" is defined on a *quantitative* scale of the strength of preference k_i, $i \in I$. However, in many real world problems, the existence of such a quantitative scale is rather questionable. Roy [45] distinguishes the following cases:

- Preferences expressed on an *ordinal* scale: this is the case where the difference between two evaluations has no clear meaning;
- Preferences expressed on a *quantitative* scale: this is the case where the scale is defined with reference to a unit clearly identified, such that it is meaningful to consider an origin (zero) of the scale and ratios between evaluations (ratio scale);
- Preferences expressed on a *numerical non-quantitative* scale: this is an intermediate case between the previous two; there are two well-known particular cases:
  - *Interval scale*, where it is meaningful to compare ratios between differences of pairs of evaluations,
  - Scale for which a *complete preorder* can be defined on all possible pairs of evaluations.

The preference scale has also been considered within economic theory (e. g. [47]), where cardinal utility is distinguished from ordinal utility: the former deals with a strength of preference, while, for the latter, this concept is meaningless. From this point of view, preferences expressed on an ordinal scale refer to *ordinal utility*, while preferences expressed on a quantitative scale or a numerical non-quantitative scale deal with *cardinal utility*.

The strength of preference k_i and, therefore, the graded preference considered in Subsect. "Approximation by Means of Multi-graded Dominance Relations", is meaningful when the scale is quantitative or numerical non-quantitative. If the information about k_i is non-available, then it is possible to define a rough approximation of S and S^c using a specific dominance relation between pairs of objects from $U \times U$, defined on an ordinal scale represented by evaluations $f_i(x)$ on criterion i, for $x \in U$ [14]. Let us explain this latter case in more details.

Let I^{O} be the set of criteria expressing preferences on an ordinal scale, and I^{N}, the set of criteria expressing preferences on a quantitative scale or a numerical non-quantitative scale, such that $I^{\mathrm{O}} \cup I^{\mathrm{N}} = I$ and $I^{\mathrm{O}} \cap I^{\mathrm{N}} = \emptyset$. Moreover, for each $P \subseteq I$, P^{O} denotes the subset of P composed of criteria expressing preferences on an ordinal scale, i. e. $P^{\mathrm{O}} = P \cap I^{\mathrm{O}}$, and P^{N} the subset of P composed of criteria expressing preferences on a quantitative scale or a numerical non-quantitative scale, i. e. $P^{\mathrm{N}} = P \cap I^{\mathrm{N}}$. Of course, for each $P \subseteq I$, $P = P^{\mathrm{N}} \cup P^{\mathrm{O}}$ and $P^{\mathrm{O}} \cap P^{\mathrm{N}} = \emptyset$.

If $P = P^{\mathrm{N}}$ and $P^{\mathrm{O}} = \emptyset$, then the definition of dominance is the same as in the case of multi-graded dominance (Subsect. "Approximation by Means of Multi-graded Dominance Relations"). If $P = P^{\mathrm{O}}$ and $P^{\mathrm{N}} = \emptyset$, then, given $(x, y), (w, z) \in U \times U$, the pair (x, y) is said to dominate the pair (w, z) with respect to P if, for each $i \in P$, $f_i(x) \geq f_i(w)$ and $f_i(z) \geq f_i(y)$. Let $D_{\{i\}}$ be the dominance relation confined to the single criterion $i \in P^{\mathrm{O}}$. The binary relation $D_{\{i\}}$ is reflexive ($(x, y)D_{\{i\}}(x, y)$, $\forall (x, y) \in U \times U$), transitive ($(x, y)D_{\{i\}}(w, z)$ and $(w, z)D_{\{i\}}(u, v)$ imply $(x, y)D_{\{i\}}(u, v)$, $\forall (x, y), (w, z), (u, v) \in U \times U$), but non-complete (it is possible that *not* $(x, y)D_{\{i\}}(w, z)$ and *not* $(w, z)D_{\{i\}}(x, y)$ for some $(x, y), (w, z) \in U \times U$). Therefore, $D_{\{i\}}$ is a partial preorder. Since the intersection of partial preorders is also a partial preorder and $D_P = \bigcap_{i \in P} D_{\{i\}}$, $P = P^{\mathrm{O}}$, then the dominance relation D_P is also a partial preorder. If some criteria from $P \subseteq I$ express preferences on a quantitative or a numerical non-quantitative scale and others on an ordinal scale, i. e. if $P^{\mathrm{N}} \neq \emptyset$ and $P^{\mathrm{O}} \neq \emptyset$, then, given $(x, y), (w, z) \in U \times U$, the pair (x, y) is said to dominate the pair (w, z) with respect to criteria from P, if (x, y) dominates (w, z) with respect to both P^{N} and P^{O}. Since the dominance relation with respect to P^{N} is a partial preorder on $U \times U$ (because it is a multi-graded dominance) and the dominance with respect to P^{O} is also

a partial preorder on $U \times U$ (as explained above), then also the dominance D_P, being the intersection of these two dominance relations, is a partial preorder. In consequence, all the concepts introduced in the previous subsection can be restored using this specific definition of dominance relation.

Using the approximations of S and S^c based on the dominance relation defined above, it is possible to induce a generalized description of the available preference information in terms of decision rules. These decision rules are of the same type as the rules already introduced in the previous subsection, however, the conditions on criteria from I^{O} are expressed directly in terms of evaluations belonging to value sets of these criteria. Let $F_i = \{f_i(x), x \in U\}$, $i \in I^{\mathrm{O}}$. The decision rules have in this case the following syntax:

1. $D_{\geq}$-decision rule, being a statement of the type:

$$\begin{aligned} &\textit{if } xP_{i1}^{\geq h(i1)}y \textit{ and } \ldots xP_{ie}^{\geq h(ie)}y \textit{ and } f_{ie+1}(x) \geq r_{ie+1} \\ &\textit{and } f_{ie+1}(y) \leq s_{ie+1} \textit{ and } \ldots f_{ip}(x) \geq r_{ip} \\ &\textit{and } f_{ip}(y) \leq s_{qp}, \textit{ then } xSy\,, \end{aligned}$$

where $P = \{i1, \ldots, ip\} \subseteq I$, $P^{\mathrm{N}} = \{i1, \ldots, ie\}$, $P^{\mathrm{O}} = \{ie+1, \ldots, ip\}$, $(h(i1), \ldots, h(ie)) \in H_{i1} \times \cdots \times H_{ie}$ and $(r_{ie+1}, \ldots, r_{ip}), (s_{ie+1}, \ldots, s_{ip}) \in F_{ie+1} \times \cdots \times F_{ip}$; these rules are supported by pairs of objects from the P-lower approximation of S only;

2. $D_{\leq}$-decision rule, being a statement of the type:

$$\begin{aligned} &\textit{if } xP_{i1}^{\leq h(i1)}y \textit{ and } \ldots xP_{ip}^{\leq h(ip)}y \textit{ and } f_{ie+1}(x) \leq r_{ie+1} \\ &\textit{and } f_{ie+1}(y) \geq s_{ie+1} \textit{ and } \ldots f_{ip}(x) \leq r_{ip} \\ &\textit{and } f_{ip}(y) \geq s_{ip}, \textit{ then } xS^cy\,, \end{aligned}$$

where $P = \{i1, \ldots, ip\} \subseteq I$, $P^{\mathrm{N}} = \{i1, \ldots, ie\}$, $P^{\mathrm{O}} = \{ie+1, \ldots, ip\}$, $(h(i1), \ldots, h(ie)) \in H_{i1} \times \cdots \times H_{ie}$ and $(r_{ie+1}, \ldots, r_{ip}), (s_{ie+1}, \ldots, s_{ip}) \in F_{ie+1} \times \cdots \times F_{ip}$; these rules are supported by pairs of objects from the P-lower approximation of S^c only;

3. $D_{\geq\leq}$-decision rule, being a statement of the type:

$$\begin{aligned} &\textit{if } xP_{i1}^{\geq h(i1)}y \textit{ and } \ldots xP_{ie}^{\geq h(ie)}y \textit{ and } xP_{ie+1}^{\leq h(ie+1)}y \ldots \\ &\quad xP_{if}^{\leq h(if)}y \textit{ and } f_{if+1}(x) \geq r_{if+1} \\ &\quad \textit{and } f_{if+1}(y) \leq s_{if+1} \textit{ and } \ldots f_{ig}(x) \geq r_{ig} \\ &\quad \textit{and } f_{ig}(y) \leq s_{ig} \textit{ and } f_{ig+1}(x) \leq r_{ig+1} \\ &\quad \textit{and } f_{ig+1}(y) \geq s_{ig+1} \textit{ and } \ldots f_{ip}(x) \leq r_{ip} \\ &\textit{and } f_{ip}(y) \geq s_{ip}, \textit{ then } xSy \textit{ or } xS^cy\,, \end{aligned}$$

where $O' = \{i1, \ldots, ie\} \subseteq I$, $O'' = \{ie+1, \ldots, if\} \subseteq I$, $P^{\mathrm{N}} = O' \cup O''$, O' and O'' not necessarily disjoint, $P^{\mathrm{O}} = \{if+1, \ldots, ip\}$, $(h(i1), \ldots, h(if)) \in H_{i1} \times \cdots \times H_{if}$ and $(r_{if+1}, \ldots, r_{ip}), (s_{if+1}, \ldots, s_{ip}) \in F_{if+1} \times \cdots \times F_{ip}$; these rules are supported by pairs of objects from the P-boundary of S and S^c only.

Rough Sets in Decision Making, Table 7
Information table of the illustrative example

Warehouse	A_1	A_2	A_3	A_4
1	high	good	no	profit
2	medium	medium	no	loss
3	medium	medium	no	profit
4	low	medium	no	loss
5	medium	good	yes	loss
6	high	medium	yes	profit

Example Illustrating DRSA in the Context of Multiple Criteria Choice and Ranking

The following example illustrates DRSA in the context of multiple criteria choice and ranking. Six warehouses have been described by means of four criteria:

- f_1, capacity of the sales staff,
- f_2, perceived quality of goods,
- f_3, high traffic location,
- f_4, warehouse profit or loss.

The components of the information table $\boldsymbol{S}$ are: $U = \{1, 2, 3, 4, 5, 6\}$, $F = \{f_1, f_2, f_3, f_4\}$, $I = \{1, 2, 3, 4\}$, $F_1 = \{\text{high, medium, low}\}$, $F_2 = \{\text{good, medium}\}$, $F_3 = \{\text{no, yes}\}$, $F_4 = \{\text{profit, loss}\}$, the criterion $f_i(x)$, taking values $f_1(1) = \text{high}$, $f_2(1) = \text{good}$, and so on.

It is assumed that the DM accepts to express preferences with respect to criteria f_1, f_2, f_3 on a *numerical non-quantitative scale*, for which a complete preorder can be defined on all possible pairs of evaluations. According to this assumption, in order to build the PCT, as described in Subsect. "Approximation by Means of Multi-graded Dominance Relation", the DM specifies sets of possible degrees of preference; for example, $H_1 = \{-2, -1, 0, 1, 2\}$, $H_2 = \{-1, 0, 1\}$, $H_3 = \{-1, 0, 1\}$. Therefore, with respect to f_1, there are the following preference relations P_1^h:

- xP_1^2y (and $yP_1^{-2}x$), meaning that x is preferred to y with respect to f_1, if $f_1(x) = \text{high}$ and $f_1(y) = \text{low}$,
- xP_1^1y (and $yP_1^{-1}x$), meaning that x is weakly preferred to y with respect to f_1, if $f_1(x) = \text{high}$ and $f_1(y) = \text{medium}$, or $f_1(x) = \text{medium}$ and $f_1(y) = \text{low}$,
- xP_1^0y (and yP_1^0x), meaning that x is indifferent to y with respect to f_1, if $f_1(x) = f_1(y)$.

Analogously, with respect to f_2 and f_3, there are the following preference relations P_2^h and P_3^h:

- xP_2^1y (and $yP_2^{-1}x$), meaning that x is weakly preferred to y with respect to f_2, if $f_2(x) =$ good and $f_2(y) =$ medium,
- xP_2^0y (and yP_2^0x), meaning that x is indifferent to y with respect to f_2, if $f_2(x) = f_2(y)$,
- xP_3^1y (and $yP_3^{-1}x$), meaning that x is weakly preferred to y with respect to f_3, if $f_3(x) =$ yes and $f_3(y) =$ no,
- xP_3^0y (and yP_3^0x), meaning that x is indifferent to y with respect to f_3, if $f_3(x) = f_3(y)$.

As to the comprehensive preference relation, the DM considers that, given two different warehouses $x, y \in U = \{1, 2, 3, 4, 5, 6\}$, if x makes profit and y makes loss, then xSy and yS^cx. Moreover, the DM accepts xSx for each warehouse x. As to warehouses x and y, which both make profit or both make loss, the DM abstains from judging whether xSy or xS^cy. Therefore, the set of exemplary pairwise comparisons supplied by the DM is $\hat{B} = \{(1,1), (1,2), (1,4), (1,5), (2,1), (2,2), (2,3), (2,6), (3,2), (3,3), (3,4), (3,5), (4,1), (4,3), (4,4), (4,6), (5,1), (5,3), (5,5), (5,6), (6,2), (6,4), (6,5), (6,6)\}$.

At this stage, the PCT can be build as shown in Table 8.

The I-lower approximations, the I-upper approximations and the I-boundaries of S and S^c obtained by means of multi-graded dominance relations are as follows:

- $\underline{I}(S) = \{(1,2), (1,4), (1,5), (3,4), (6,2), (6,4), (6,5)\}$,
- $\overline{I}(S) = \{(1,1), (1,2), (1,4), (1,5), (2,2), (2,3), (3,2), (3,3), (3,4), (3,5), (4,4), (5,3), (5,5), (6,2), (6,4), (6,5), (6,6)\}$,
- $\underline{I}(S^c) = \{(2,1), (2,6), (4,1), (4,3), (4,6), (5,1), (5,6)\}$,
- $\overline{I}(S^c) = \{(1,1), (2,1), (2,2), (2,3), (2,6), (3,2), (3,3), (3,5), (4,1), (4,3), (4,4), (4,6), (5,1), (5,3), (5,5), (5,6), (6,6)\}$,
- $Bn_I(S) = Bn_I(S^c) = \{(1,1), (2,2), (2,3), (3,2), (3,3), (3,5), (4,4), (5,3), (5,5), (6,6)\}$.

Therefore, the quality of approximation is equal to 0.58. Moreover, there is only one reduct which is also the core, i. e. $RED_S(I) = CORE_S(I) = \{1\}$.

Finally, the following decision rules can be induced (within parentheses there are the pairs of objects supporting the rule):

- *if* $xP_1^{\geq 1}y$, *then* xSy (or, in words, *if* x is at least weakly preferred to y with respect to f_1, *then* x outranks y), $((1,2), (1,4), (1,5), (3,4), (6,2), (6,4), (6,5))$,
- *if* $xP_1^{\leq -1}y$, *then* xS^cy (or, in words, *if* y is at least weakly preferred to x with respect to f_1, *then* x does not outrank y), $((2,1), (2,6), (4,1), (4,3), (4,6), (5,1), (5,6))$,
- *if* $xP_1^{\geq 0}y$, and $xP_1^{\leq 0}y$, (i. e. *if* xP_1^0y), *then* xSy *or* xS^cy (or, in words, *if* x and y are indifferent with respect to f_1, *then* x outranks y *or* x does not outrank y), $((1,1), (2,2), (2,3), (3,2), (3,3), (3,5), (4,4), (5,3), (5,5), (6,6))$.

Rough Sets in Decision Making, Table 8
Pairwise comparison table

Pairs	P_1^h	P_2^h	P_3^h	Outranking
(1,1)	0	0	0	S
(1,2)	1	1	0	S
(1,4)	2	1	0	S
(1,5)	1	0	−1	S
(2,1)	−1	−1	0	S^c
(2,2)	0	0	0	S
(2,3)	0	0	0	S^c
(2,6)	−1	0	−1	S^c
(3,2)	0	0	0	S
(3,3)	0	0	0	S
(3,4)	1	0	0	S
(3,5)	0	−1	−1	S
(4,1)	−2	−1	0	S^c
(4,3)	−1	0	0	S^c
(4,4)	0	0	0	S
(4,6)	−2	0	−1	S^c
(5,1)	−1	0	1	S^c
(5,3)	0	1	1	S^c
(5,5)	0	0	0	S
(5,6)	−1	1	0	S^c
(6,2)	1	0	1	S
(6,4)	2	0	1	S
(6,5)	1	−1	0	S
(6,6)	0	0	0	S

Let us assume now that the DM accepts to express preferences with respect to criteria f_1, f_2, f_3 on an *ordinal scale* of preference, for which there is only information about a partial preorder on all possible pairs of evaluations. In this case, S and S^c can be approximated in the way described in Subsect. "Dominance Without Degrees of Preference", i. e. without considering degrees of preference. The I-lower approximations, the I-upper approximations and the I-boundaries of S and S^c are as follows:

- $\underline{I}(S) = \{(1,1), (1,2), (1,4), (1,5), (3,4), (4,4), (6,2), (6,4), (6,5), (6,6)\}$,
- $\overline{I}(S) = \{(1,1), (1,2), (1,4), (1,5), (2,2), (2,3), (3,2), (3,3), (3,4), (3,5), (4,4), (5,3), (5,5), (6,2), (6,4), (6,5), (6,6)\}$,

- $\underline{I}(S^c) = \{(2,1), (2,6), (4,1), (4,3), (4,6), (5,1), (5,6)\}$,
- $\overline{I}(S^c) = \{(2,1), (2,2), (2,3), (2,6), (3,2), (3,3), (3,5), (4,1), (4,3), (4,6), (5,1), (5,3), (5,5), (5,6)\}$,
- $Bn_I(S) = Bn_I(S^c) = \{(2,2), (2,3), (3,2), (3,3), (3,5), (5,3), (5,5)\}$.

Let us observe that the pairs $(1,1)$, $(4,4)$ and $(6,6)$ belong now to the I-lower approximation of S and are not contained in the I-boundaries. Therefore, the quality of approximation is equal to 0.71. Moreover, there is still only one reduct which is also the core, i. e. again $RED_S(I) = CORE_S(I) = \{1\}$.

The following decision rules are induced from the above approximations and boundaries (within parentheses there are the pairs of objects supporting the rule):

- *if* $f_1(x)$ is at least high *and* $f_1(y)$ is at most high, *then* xSy, $((1,1), (1,2), (1,4), (1,5), (6,2), (6,4), (6,5), (6,6))$,
- *if* $f_1(x)$ is at least low *and* $f_1(y)$ is at most low, *then* xSy, $((1,4), (3,4), (4,4), (6,4))$,
- *if* $f_1(x)$ is at most medium *and* $f_1(y)$ is at least high, *then* xS^cy, $((2,1), (2,6), (4,1), (4,6)(5,1), (5,6))$,
- *if* $f_1(x)$ is at most low *and* $f_1(y)$ is at least medium, *then* xS^cy, $((4,1), (4,3), (4,6))$
- *if* $f_1(x)$ is at least medium *and* $f_1(y)$ is at most medium *and* $f_1(x)$ is at most medium *and* $f_1(y)$ is at least medium, (i. e. *if* $f_1(x)$ is equal to medium *and* $f_1(y)$ is equal to medium), *then* xSy *or* xS^cy, $((2,2), (2,3), (3,2), (3,3), (3,5), (5,3), (5,5))$.

DRSA for Decision Under Uncertainty

Basic Concepts

To apply the rough set approach to decision under uncertainty, the following basic elements must be considered:

- a set $S = \{s_1, s_2, \ldots, s_u\}$ of states of the world, or simply *states*, which are supposed to be mutually exclusive and collectively exhaustive,
- an a priori probability distribution P over the states of the world: more precisely, the probabilities of states $s_1, s_2, \ldots, s_u$ are $p_1, p_2, \ldots, p_u$, respectively $(p_1 + p_2 + \cdots + p_u = 1, p_i \geq 0, i = 1, \ldots, u)$,
- a set $A = \{a_1, a_2, \ldots, a_m\}$ of *acts*,
- a set $X = \{x_1, x_2, \ldots, x_r\}$ of *outcomes* or *consequences* expressed in monetary terms $(X \subseteq \Re)$,
- a function $g: A \to X$ assigning to each pair act-state $(a_i, s_j) \in A \times S \times S$ an outcome $x \in X$,
- a set of *classes* $\boldsymbol{Cl} = \{Cl_1, Cl_2, \ldots, Cl_p\}$, such that $Cl_1 \cup Cl_2 \cup \ldots \cup Cl_p = A, Cl_r \cap Cl_q = \emptyset$ for each $r, q \in \{1, \ldots, p\}$ with $r \neq q$; the classes of $\boldsymbol{Cl}$ are preference-ordered according to the increasing order of their indices,
- a function $e: A \to \boldsymbol{Cl}$ assigning each act $a_i \in A$ to a class $Cl_j \in \boldsymbol{Cl}$.

In this context, two different types of dominance can be considered:

1) *Classical dominance*: given $a_p, a_q \in A$, a_p dominates a_q iff, for each possible state of the world, act a_p gives an outcome at least as good as act a_q; more formally, $g(a_p, s_j) \geq g(a_q, s_j)$, for each $s_j \in S$,
2) *Stochastic dominance*: given $a_p, a_q \in A$, a_p dominates a_q iff, for each outcome $x \in X$, act a_p gives an outcome at least as good as x with a probability at least as great as the probability that act a_q gives the same outcome, i. e. for all $x \in X$,

$$P[S(a_p, x)] \geq P[S(a_q, x)]$$

where, for each $(a_i, x) \in A \times X$, $S(a_i, x) = \{s_j \in S : g(a_i, s_j) \geq x\}$.

In [19], it has been shown how to apply stochastic dominance in this context. On the basis of an a priori probability distribution P, one can assign to each subset of states of the world $W \subseteq S$ $(W \neq \emptyset)$ the probability $P(W)$ that one of the states in W is verified, i. e. $P(W) = \sum_{i:s_i \in W} p_i$, and then to build up the set Π of all possible values $P(W)$, i. e.

$$\Pi = \{\pi \in [0,1] : \pi = P(W), W \subseteq S\}.$$

Let us define the following functions $z: A \times S \to \Pi$ and $z': A \times S \to \Pi$ assigning to each act-state pair $(a_i, s_j) \in A \times S$ a probability $\pi \in \Pi$, as follows:

$$z(a_i, s_j) = \sum_{r: g(a_i, s_r) \geq g(a_i, s_j)} p_r ,$$

and

$$z'(a_i, s_j) = \sum_{r: g(a_i, s_r) \leq g(a_i, s_j)} p_r .$$

Therefore, $z(a_i, s_j)$ represents the probability of obtaining an outcome whose value is at least $g(a_i, s_j)$ by act a_i. Analogously, $z'(a_i, s_j)$ represents the probability of obtaining an outcome whose value is at most $g(a_i, s_j)$ by act a_i. On the basis of function $z(a_i, s_j)$, function $\rho: A \times \Pi \to X$ can be defined as follows:

$$\rho(a_i, \pi) = \max_{j:\, z(a_i, s_j) \geq \pi} g(a_i, s_j) .$$

Thus, $\rho(a_i, \pi) = x$ means that the outcome got by act a_i is greater than or equal to x with a probability at least p (i. e. a probability p or greater). On the basis of function $z'(a_i, s_j)$, the function $\rho' : A \times \Pi \to X$ can be defined as follows:

$$\rho'(a_i, \pi) = \min_{j:\, z(a_i,s_j) \le \pi^{g(a_i,s_j)}} .$$

$\rho'(a_i, \pi) = x$ means that the outcome got by act a_i is smaller than or equal to x with a probability at least p.

Let us observe that information given by $\rho(a_i, \pi)$ and $\rho'(a_i, \pi)$ is related. In fact, if the elements of Π, $0 = \pi_{(0)}, \pi_{(1)}, \pi_{(2)}, \ldots, \pi_{(d)} = 1$ $(d = |\Pi|)$, are reordered in such a way that $0 = \pi_{(0)} \le \pi_{(1)} \le \pi_{(2)} \le \ldots \le \pi_{(d)} = 1$, then

$$\rho\left(a_i, \pi_{(j)}\right) = \rho'\left(a_i, 1 - \pi_{(j-1)}\right) .$$

Therefore, $\rho(a_i, \pi_{(j)}) \le x$ is equivalent to $\rho'(a_i, 1 - \pi_{(j-1)}) \ge x$, $a_i \in A, \pi_{(j-1)}, \pi_{(j)} \in \Pi, x \in X$. This implies that the analysis of the possible decisions can be equivalently conducted on values of either $\rho(a_i, \pi)$ or $\rho'(a_i, \pi)$. However, from the point of view of representation of results, it is interesting to consider both values $\rho(a_i, \pi)$ and $\rho'(a_i, \pi)$. The reason is that, contrary to intuition, $\rho(a_i, \pi) \le x$ is not equivalent to the statement that by act a_i the outcome is smaller than or equal to x with a probability at least π. The following example clarifies this point. Let us consider a game a with rolling a dice, in which if the result is 1, then the gain is 1, if the result is 2 then the gain is 2, and so on. Suppose, moreover, that the dice is equilibrated and thus each result is equiprobable with probability 1/6. The values of $\rho(a_i, \pi)$ for all possible values of probability are:

$$\rho(a, 1/6) = \$6, \quad \rho(a, 2/6) = \$5, \quad \rho(a, 3/6) = \$4 ,$$

$$\rho(a, 4/6) = \$3, \quad \rho(a, 5/6) = \$2, \quad \rho(a, 6/6) = \$1 .$$

Let us remark that $\rho(a, 5/6) \le \$3$ (indeed, $\rho(a, 5/6) = \$2$, and thus $\rho(a, 5/6) \le \$3$ is true), however, this is not equivalent to the statement that by act a the outcome is smaller than or equal to \$3 with a probability at least 5/6. In fact, this is false because this probability is 3/6 (related to results 1, 2 and 3). Analogously, the values of $\rho'(a, \pi)$ for all possible values of probability are:

$$\rho'(a, 1/6) = \$1, \quad \rho'(a, 2/6) = \$2, \quad \rho'(a, 3/6) = \$3 ,$$

$$\rho'(a, 4/6) = \$4, \quad \rho'(a, 5/6) = \$5, \quad \rho'(a, 6/6) = \$6 .$$

Let us remark that $\rho'(a, 5/6) \ge \$5$ (indeed, $\rho'(a, 5/6) = \$5$, and thus $\rho'(a, 5/6) \ge \$3$ is true), however, this is not equivalent to the statement that by act a the outcome is greater than or equal to 3 with a probability at least 5/6. In fact, this is false because this probability is 4/6 (related to results 3, 4, 5 and 6). Therefore, in the context of stochastic acts, an outcome expressed in positive terms refers to $\rho(a, \pi)$ giving a lower bound of an outcome ("for act a there is a probability π to gain at least $\rho(a, \pi)$"), while an outcome expressed in negative terms refers to $\rho'(a, \pi)$ giving an upper bound of an outcome ("for act a there is a probability π to gain at most $\rho'(a, \pi)$").

Given $a_p, a_q \in A$, a_p *stochastically dominates* a_q if and only if $\rho(a_p, \pi) \ge \rho(a_q, \pi)$ for each $\pi \in \Pi$. This is equivalent to the statement: given $a_p, a_q \in A$, a_p stochastically dominates a_q if and only if $\rho'(a_p, \pi) \ge \rho'(a_q, \pi)$ for each $\pi \in \Pi$.

For example, consider the game a^* with rolling a dice, in which if the result is 1, then the gain is \$7, if the result is 2 then the gain is \$6, and so on until the case in which the result is 6 and the gain is \$2. In this case game a^* stochastically dominates game a because $\rho(a^*, 1/6) = \$7$ is not smaller than $\rho(a, 1/6) = \$6$, $\rho(a^*, 2/6) = \$6$ is not smaller than $\rho(a, 2/6) = \$5$, and so on. Equivalently, game a^* stochastically dominates game a because $\rho'(a^*, 1/6) = \$2$ is not smaller than $\rho'(a, 1/6) = \$1$, $\rho'(a^*, 2/6) = \$3$ is not smaller than $\rho'(a, 2/6) = \$2$, and so on.

DRSA can be applied in the context of decision under uncertainty considering as set of objects U the set of acts A, as set of criteria (condition attributes) I the set Π, as decision attribute $\{d\}$ the classification $\boldsymbol{Cl}$, as value set of all criteria the set X, as information function f a function $\boldsymbol{f}$ such that $\boldsymbol{f}(a_i, \pi) = \rho(a_i, \pi)$ and $\boldsymbol{f}(a_i, cl) = e(a_i)$. Let us observe that due to equivalence $\rho(a_i, \pi_{(j)}) = \rho'(a_i, 1 - \pi_{(j-1)})$, one can also consider information function $\boldsymbol{f}'(a_i, \pi) = \rho'(a_i, \pi)$.

The aim of the rough set approach to preferences under uncertainty is to explain the preferences of the DM represented by the assignments of the acts from A to the classes from $\boldsymbol{Cl}$ in terms of stochastic dominance, expressed by means of function ρ. The syntax of decision rules obtained from this rough set approach is as follows:

1) $D_{\ge}$-decision rules with the following syntax: "if $\rho(a, p_{h_1}) \ge x_{h_1}$ and …, and $\rho(a, p_{h_z}) \ge x_{h_z}$, then $a \in Cl_r^{\ge}$" (i. e. "if by act a the outcome is at least x_{h_1} with probability at least p_{h_1}, and …, and the outcome is at least x_{h_z} with probability at least p_{h_z}, then $a \in Cl_r^{\ge}$") where $p_{h_1}, \ldots, p_{h_z} \in \Pi, x_{h_1}, \ldots, x_{h_z} \in X$ and $r \in \{2, \ldots, p\}$;
2) $D_{\le}$-decision rules with the following syntax: "if $\rho'(a, p_{h_1}) \le x_{h_1}$ and …, and $\rho'(a, p_{h_z}) \le x_{h_z}$, then $a \in Cl_r^{\le}$" (i. e. "if by act a the outcome is at most

x_{h_1} with probability at least p_{h_1}, and …, and the outcome is at most x_{h_z} with probability at least p_{h_z}, then $a \in Cl_r^{\leq}$") where $p_{h_1}, \ldots, p_{h_z} \in \Pi, x_{h_1}, \ldots, x_{h_z} \in X$ and $r \in \{1, \ldots, p-1\}$;

3) $D_{\geq\leq}$-decision rules with the following syntax: "if $\rho(a, p_{h_1}) \geq x_{h_1}$ and …, and $\rho(a, p_{h_z}) \geq x_{h_w}$ and $\rho'(a, p_{h_{w+1}}) \leq x_{h_{w+1}}$ and …, and $\rho'(a, p_{h_z}) \leq x_{h_z}$, then $a \in Cl_s \cup Cl_{s+1} \cup \ldots Cl_t$" (i. e. "if by act a the outcome is at least x_{h_1} with probability at least p_{h_1}, and …, and the outcome is at least x_{h_w} with probability at least p_{h_w} and the outcome is at most $x_{h_{w+1}}$ with probability at least $p_{h_{w+1}}$, and …, and the outcome is at most x_{h_z} with probability at least p_{h_z}, then $a \in Cl_s \cup Cl_{s+1} \cup \ldots Cl_t$") where $p_{h_1}, \ldots, p_{h_w}, p_{h_{w+1}}, p_{h_z} \in \Pi$, $x_{h_1}, \ldots, x_{h_z} \in X$ and $s, t \in \{1, \ldots, p\}$, such that $s < t$.

According to the meaning of $\rho(a_i, p)$ and $\rho'(a_i, p)$ discussed above, $D_{\geq}$-decision rules are expressed in terms of $\rho(a_i, p)$, $D_{\leq}$-decision rules are expressed in terms of $\rho'(a_i, p)$, and $D_{\geq\leq}$-decision rules are expressed in terms of both $\rho(a_i, p)$ and $\rho'(a_i, p)$. Let us observe that due to equivalence $\rho(a_i, \pi_{(j)}) = \rho'(a_i, 1 - \pi_{(j-1)})$, all above decision rules can be expressed equivalently in terms of values of $\rho(a_i, p)$ or $\rho'(a_i, p)$. For example, a $D_{\geq}$-decision rule $r_{\geq}(\rho)$ ="if $\rho(a, p_{h_1}) \geq x_{h_1}$ and …, and $\rho(a, p_{h_z}) \geq x_{h_z}$, then $a \in Cl_r^{\geq}$" can be expressed in terms of $\rho'(a_i, p)$ as $r_{\geq}(\rho')$ ="if $\rho'(a, p^*_{h_1}) \geq x_{h_1}$ and …, and $\rho'(a, p^*_{h_z}) \geq x_{h_z}$, then $a \in Cl_r^{\geq}$", where, if $p_{h_r} = \pi(j_r)$, then $p^*_{h_r} = 1 - \pi(j_r - 1)$, with $r = 1, \ldots, z$, and $0 = \pi(0), \pi(1), \pi(2), \ldots, \pi(|P|) = 1$ reordered in such a way that $0 = \pi(0) \leq \pi(1) \leq \pi(2) \leq \cdots \leq \pi(|P|) = 1$. Analogously, a $D_{\leq}$-decision rule $r_{\leq}(\rho')$ ="if $\rho'(a, p_{h_1}) \leq x_{h_1}$ and …, and $\rho'(a, p_{h_z}) \leq x_{h_z}$, then $a \in Cl_r^{\leq}$" can be expressed in terms of $\rho(a_i, p)$ as $r_{\leq}(\rho)$ ="if $\rho(a, p^*_{h_1}) \leq x_{h_1}$ and …, and $\rho(a, p^*_{h_z}) \leq x_{h_z}$, then $a \in Cl_r^{\leq}$, where, if $p_{h_r} = \pi(j_r)$, then $p^*_{h_r} = 1 - \pi(j_r - 1)$, with $r = 1, \ldots, z$, and $0 = \pi(0), \pi(1), \pi(2), \ldots, \pi(|P|) = 1$ reordered in such a way that $0 = \pi(0) \leq \pi(1) \leq \pi(2) \leq \cdots \leq \pi(|P|) = 1$.

Let us observe, however, that $r_{\geq}(\rho)$ is an expression much more natural and meaningful than $r_{\geq}(\rho')$, as well as $r_{\leq}(\rho')$ is an expression much more natural and meaningful than $r_{\leq}(\rho)$. Another useful remark concerns minimality of rules, related to the specific intrinsic structure of the stochastic dominance. Let us consider the following two decision rules:

- $r1 \equiv$ "if by act a the outcome is at least 100 with probability at least 0.25, then a is at least good",
- $r2 \equiv$ "if by act a the outcome is at least 100 with probability at least 0.50, then a is at least good".

$r1$ and $r2$ can be induced from the analysis of the same information table, because they involve different criteria (condition attributes). In fact, $r1$ involves attribute $\rho(a, 0.25)$ (it can be expressed as "if $\rho(a, 0.25) \geq 100$, then a is at least good"), $r2$ involves attribute $\rho(a, 0.50)$ (it can be expressed as "if $\rho(a, 0.50) \geq 100$, then a is at least good"). Considering the structure of the stochastic dominance, the condition part of rule $r1$ is the weakest. In fact, rule $r1$ requires a cumulated outcome to be at least 100 with probability of 0.25, while rule $r2$ requires the same outcome but with a greater probability, 0.5 against 0.25. Since the decision part of these two rules is the same, $r1$ is minimal among these two rules. From a practical point of view, this observation says that, if one induces decision rules using the algorithms designed for DRSA, it is necessary to further filter the obtained results in order to remove rules which are not minimal in the specific context of the DRSA analysis based on stochastic dominance.

Example Illustrating DRSA in the Context of Decision Under Uncertainty

The following example illustrates the approach. Let us consider

- a set $S = \{s_1, s_2, s_3\}$ of states of the world,
- an a priori probability distribution P over the states of the world defined as follows: $p_1 = 0.25, p_2 = 0.35, p_3 = 0.40$,
- a set $A = \{a_1, a_2, a_3, a_4, a_5, a_6\}$ of acts,
- a set $X = \{0, 10, 15, 20, 30\}$ of consequences
- a set of classes $\boldsymbol{Cl} = \{Cl_1, Cl_2, Cl_3\}$, where Cl_1 is the set of bad acts, Cl_2 is the set of medium acts, Cl_3 is the set of good acts,
- a function $g: A \to X$ assigning to each act-state pair $(a_i, s_j) \in A \times S \times S$ a consequence $x_h \in X$ and a function $e: A \to \boldsymbol{Cl}$ assigning each act $a_i \in A$ to a class $Cl_j \in \boldsymbol{Cl}$ presented in the following Table 9.

Rough Sets in Decision Making, Table 9
Acts, consequences and assignment to classes from *Cl*

	p_j	a_1	a_2	a_3	a_4	a_5	a_6
s_1	0.25	30	0	15	0	20	10
s_2	0.35	10	20	0	15	10	20
s_3	0.40	10	20	20	20	20	20
Cl		good	medium	medium	bad	medium	good

Rough Sets in Decision Making, Table 10
Acts, values of function $\rho(a_i, p)$ and assignment to classes from Cl

π	a_1	a_2	a_3	a_4	a_5	a_6
0.25	30	20	20	20	20	20
0.35	10	20	20	20	20	20
0.40	10	20	20	20	20	20
0.60	10	20	15	15	20	20
0.65	10	20	15	15	20	20
0.75	10	20	0	15	10	20
1	10	0	0	0	10	10
Cl	good	medium	medium	bad	medium	good

Table 10 shows the values of function $\rho(a_i, p)$.

Table 10 is the decision table on which the DRSA is applied. Let us give some examples of the interpretation of the values in Table 10. The column of act a_3 can be read as follow:

- the value 20 in the row corresponding to 0.25 means that the outcome is at least 20 with a probability of at least 0.25,
- the value 15 in the row corresponding to 0.65 means that the outcome is at least 15 with a probability of at least 0.65,
- the value 0 in the row corresponding to 0.75 means that the outcome is at least 0 with a probability of at least 0.75.

Analogously, the row corresponding to 0.65, can be read as follows:

- the value 10 relative to a_1, means that by act a_1 the outcome is at least 10 with a probability of at least 0.65,
- the value 20 relative to a_2, means that by act a_2 the outcome is at least 20 with a probability of at least 0.65,
- and so on.

Applying DRSA, the following upward union and downward union of classes are approximated:

- $Cl_2^{\geq} = Cl_2 \cup Cl_3$, i. e. the set of the acts at least medium,
- $Cl_3^{\geq} = Cl_3$, i. e. the set of the acts (at least) good,
- $Cl_1^{\leq} = Cl_1$, i. e. the set of the acts (at most) bad,
- $Cl_2^{\leq} = Cl_1 \cup Cl_2$, i. e. the set of the acts at most medium.

The first result of the DRSA approach was a discovery that the decision table (Table 10) is not consistent. Indeed, Table 10 shows that act a_4 stochastically dominates act a_3, however act a_3 is assigned to a better class (medium) than act a_4 (bad). Therefore, act a_3 cannot be assigned without doubts to the set of the class of the at least medium acts as well as act a_4 cannot be assigned without doubts to the set of the classes of the (at most) bad acts. In consequence, lower approximation and upper approximation of $Cl_2^{\geq}$, $Cl_3^{\geq}$ and $Cl_1^{\leq}$, $Cl_2^{\leq}$ are equal, respectively, to

- $\underline{I}\left(Cl_2^{\geq}\right) = \{a_1, a_2, a_5, a_6\} = Cl_2^{\geq} - \{a_3\}$,
- $\overline{I}\left(Cl_2^{\geq}\right) = \{a_1, a_2, a_3, a_4, a_5, a_6\} = Cl_2^{\geq} \cup \{a_4\}$,
- $\underline{I}\left(Cl_3^{\geq}\right) = \{a_1, a_6\} = Cl_3^{\geq}$,
- $\overline{I}\left(Cl_3^{\geq}\right) = \{a_1, a_6\} = Cl_3^{\geq}$,
- $\underline{I}\left(Cl_1^{\leq}\right) = \emptyset = Cl_1^{\leq} - \{a_4\}$,
- $\overline{I}\left(Cl_1^{\leq}\right) = \{a_3, a_4\} = Cl_1^{\leq} \cup \{a_3\}$,
- $\underline{I}\left(Cl_2^{\leq}\right) = \{a_2, a_3, a_4, a_5\} = Cl_2^{\leq}$,
- $\overline{I}\left(Cl_2^{\leq}\right) = \{a_2, a_3, a_4, a_5\} = Cl_2^{\leq}$.

Since there are two inconsistent acts on a total of six acts (a_3, a_4), then the quality of approximation of the ordinal classification is equal to 4/6. The second discovery was one reduct of criteria (condition attributes) ensuring the same quality of approximation as the whole set Π of probabilities: $\mathrm{RED}_{Cl} = \{0.25, 0.75, 1\}$. This means that the preferences of the DM can be explained using only the probabilities from RED_{Cl}. RED_{Cl} is also the core because no probability value in RED_{Cl} can be removed without deteriorating the quality of approximation. The third discovery was a set of minimal decision rules describing the DM's preferences (within parentheses there is a verbal interpretation of the corresponding decision rule, and the supporting acts):

1) if $\rho(a_i, 0.25) \geq 30$, then $a_i \in Cl_3^{\geq}$ (*if the probability of gaining at least 30 is at least 0.25, then act a_i is at least good*) (a_1),
2) if $\rho(a_i, 0.75) \geq 20$ and $\rho(a_i, 1) \geq 10$, then $a_i \in Cl_3^{\geq}$ (*if the probability of gaining at least 20 is at least 0.75 and the probability of gaining at least 10 is (at least) 1 (i. e. for sure the gaining is at least 10), then act a_i is at least good*) (a_6),
3) if $\rho(a_i, 1) \geq 10$, then $a_i \in Cl_2^{\geq}$ (*if the probability of gaining at least 10 is (at least) 1 (i. e. for sure the gaining is at least 10), then act a_i is at least medium*) (a_1, a_5, a_6),
4) if $\rho(a_i, 0.75) \geq 20$, then $a_i \in Cl_2^{\geq}$ (*if the probability of gaining at least 20 is at least 0.75, then act a_i is at least medium*) (a_2, a_6),
5) if $\rho(a_i, 0.25) \leq 20$ (i. e. $\rho'(a_i, 1) \leq 20$) and $\rho(a_i, 0.75) \leq 15$ (i. e. $\rho'(a_i, 0.35) \leq 15$), then $a_i \in Cl_2^{\leq}$ (*if the probability of gaining at most 20 is (at least) 1 (i. e. for sure you gain at most 20) and the probability to gain at most 15 is at least 0.35, then act a_i is at most medium*) (a_3, a_4, a_5),

6) if $\rho(a_i, 1) \leq 0$ (i. e. $\rho'(a_i, 0.25) \leq 0$), then $a_i \in Cl_2^{\leq}$ (*if the probability of gaining at most 0 is at least 0.25, then act a_i is at most medium*) (a_2, a_3, a_4),
7) if $\rho(a_i, 1) \geq 0$ and $\rho(a_i, 1) \leq 0$ (i. e. $\rho(a_i, 1) = 0$) and $\rho(a_i, 0.75) \leq 15$ ($\rho'(a_i, 0.35) \leq 15$), then $a_i \in Cl_1 \cup Cl_2$ (*if the probability of gaining at least 0 is 1, then act a_i is at most medium*) (a_2, a_3, a_4).

Minimal sets of minimal decision rules represent the most concise and non-redundant knowledge contained in Table 9 (and, consequently, in Table 10). The above minimal set of 7 decision rules uses 3 attributes (probability 0.25, 0.75 and 1) and 11 elementary conditions, i. e. 26% of descriptors from the original data table (Table 10). Of course, this is only a didactic example. Representation in terms of decision rules of larger sets of exemplary acts from real applications are more synthetic in the sense of the percentage of used descriptors from the original decision table.

Multiple Criteria Decision Analysis Using Association Rules

In multiple criteria decision analysis, the DM is often interested in relationships between attainable values of criteria. This information is particularly useful in multiobjective optimization (see [32] and Sect. "Interactive Multiobjective Optimization Using DRSA (IMO-DRSA)"). For instance, in a car selection problem, one can observe that in the set of considered cars, if the maximum speed is at least 200 km/h and the time to reach 100 km/h is at most 7 s, then the price is not less than 40,000 $ and the fuel consumption is not less than 9 liters per 100 km. These relationships are association rules whose general syntax, in case of minimization of criteria f_i, $i \in I$, is:

"if $f_{i_1}(x) \leq r_{i_1}$ and … and $f_{i_p}(x) \leq r_{i_p}$, then $f_{i_{p+1}}(x) \geq r_{i_{p+1}}$ and … and $f_{i_q}(x) \geq r_{i_q}$", where $\{i_1, \ldots, i_q\} \subseteq I, r_{i_1}, \ldots, r_{i_q} \in \Re$.

If criterion $f_i, i \in I$, should be maximized, the corresponding condition in the association rule should be reversed, i. e. in the premise, the condition becomes $f_i(x) \geq r_i$, and in the conclusion it becomes $f_i(x) \leq r_i$.

Given an association rule $r \equiv$ "if $f_{i_1}(x) \leq r_{i_1}$ and … and $f_{i_p}(x) \leq r_{i_p}$, then $f_{i_{p+1}}(x) \geq r_{i_{p+1}}$ and … and $f_{i_q}(x) \geq r_{i_q}$", an object $y \in U$ *supports* r if $f_{i_1}(y) \leq r_{i_1}$ and … and $f_{i_p}(y) \leq r_{i_p}$ and $f_{i_{p+1}}(y) \geq r_{i_{p+1}}$ and … and $f_{i_q}(y) \geq r_{i_q}$. Moreover, object $y \in U$ supporting decision rule r is a *base* of r if $f_{i_1}(y) = r_{i_1}$ and … and $f_{i_p}(y) = r_{i_p}$ and $f_{i_{p+1}}(y) = r_{i_{p+1}}$ and … and $f_{i_q}(y) = r_{i_q}$. An association rule having at least one base is called *robust*.

An association rule $r \equiv$ "if $f_{i_1}(x) \leq r_{i_1}$ and … and $f_{i_p}(x) \leq r_{i_p}$, then $f_{i_{p+1}}(x) \geq r_{i_{p+1}}$ and … and $f_{i_q}(x) \geq r_{i_q}$" holds in universe U if:

1) there is at least one $y \in U$ supporting r,
2) r is not contradicted in U, i. e. there is no $z \in U$ such that $f_{i_1}(z) \leq r_{i_1}$ and … and $f_{i_p}(z) \leq r_{i_p}$, while *not* $f_{i_{p+1}}(z) \geq r_{i_{p+1}}$ or … or $f_{i_q}(z) \geq r_{i_q}$.

Given the two association rules:

- $r_1 \equiv$"if $f_{i_1}(x) \leq r^1_{i_1}$ and … and $f_{i_p}(x) \leq r^1_{i_p}$, then $f_{i_{p+1}}(x) \geq r^1_{i_{p+1}}$ and … and $f_{i_q}(x) \geq r^1_{i_q}$",
- $r_2 \equiv$"if $f_{j_1}(x) \leq r^2_{j_1}$ and … and $f_{j_s}(x) \leq r^2_{j_s}$, then $f_{j_{s+1}}(x) \geq r^2_{j_{s+1}}$ and … and $f_{j_t}(x) \geq r^2_{j_t}$",

rule r_1 is not weaker than rule r_2, denoted by $r_1 \trianglerighteq r_2$, if:

$\alpha)$ $\{i_1, \ldots, i_p\} \subseteq \{j_1, \ldots, j_s\}$,
$\beta)$ $r^1_{i_1} \geq r^2_{i_1}, \ldots, r^1_{i_p} \geq r^2_{i_p}$,
$\gamma)$ $\{i_{p+1}, \ldots, i_q\} \supseteq \{j_{s+1}, \ldots, j_t\}$,
$\delta)$ $r^1_{j_{s+1}} \geq r^2_{j_{s+1}}, \ldots, r^1_{j_t} \geq r^2_{j_t}$.

Conditions β and δ are formulated for criteria f_i to be minimized. If criterion f_i should be maximized, the corresponding inequalities should be reversed, i. e. $r^1_i \leq r^2_i$ in condition β as well as in condition δ. Notice that $\trianglerighteq$ is a binary relation on the set of association rules, which is a partial preorder, i. e. it is reflexive (each rule is not weaker than itself) and transitive. The asymmetric part of the relation $\trianglerighteq$ is denoted by $\triangleright$, and $r_1 \triangleright r_2$ reads "r_1 is stronger than r_2".

For example, consider the following association rules:

- $r_1 \equiv$ "if the maximum speed is at least 200 km/h and the time to reach 100 km/h is at most 7 s, then the price is not less than 40,000 $ and the fuel consumption is not less than 9 liters per 100 km",
- $r_2 \equiv$ "if the maximum speed is at least 200 km/h and the time to reach 100 km/h is at most 7 s and the horse power is at least 175 kW, then the price is not less than 40,000 $ and the fuel consumption is not less than 9 liters per 100 km",
- $r_3 \equiv$ "if the maximum speed is at least 220 km/h and the time to reach 100 km/h is at most 7 s, then the price is not less than 40,000 $ and the fuel consumption is not less than 9 liters per 100 km",
- $r_4 \equiv$ "if the maximum speed is at least 200 km/h and the time to reach 100 km/h is at most 7 s, then the price is not less than 40,000 $",
- $r_5 \equiv$ "if the maximum speed is at least 200 km/h and the time to reach 100 km/h is at most 7 s, then the price

is not less than 35,000 $ and the fuel consumption is not less than 9 liters per 100 km",

- $r_6 \equiv$ "if the maximum speed is at least 220 km/h and the time to reach 100 km/h is at most 7 s and the horse power is at least 175 kW, then the price is not less than 35,000 $".

Let us observe that rule r_1 is stronger than each of the other five rules for the following reasons:

- $r_1 \rhd r_2$ for condition α) because, all things equal elsewhere, in the premise of r_2 there is an additional condition: "the horse power is at least 175 kW",
- $r_1 \rhd r_3$ for condition β) because, all things equal elsewhere, in the premise of r_3 there is a condition with a worse threshold value: "the maximum speed is at least 220 km/h" instead of "the maximum speed is at least 200 km/h",
- $r_1 \rhd r_4$ for condition γ) because, all thing equal elsewhere, in the conclusion of r_4 one condition is missing: "the fuel consumption is not less than 9 liters per 100 km",
- $r_1 \rhd r_5$ for condition δ) because, all thing equal elsewhere, in the conclusion of r_5 there is a condition with a worse threshold value: "the price is not less than 35,000 $" instead of "the price is not less than 40,000 $",
- $r_1 \rhd r_6$ for conditions α), β), γ) and δ) because all weak points for which rules r_2, r_3, r_4 and r_5 are weaker than rule r_1 are present in r_6.

An association rule r is *minimal* if there is no other rule stronger than r with respect to $\rhd$. An algorithm for induction of association rules from preference ordered data has been presented in [20].

Interactive Multiobjective Optimization Using DRSA (IMO-DRSA)

This section presents a recently proposed method for Interactive Multiobjective Optimization using Dominance-based Rough Set Approach (IMO-DRSA) [32]. Assuming that objective functions $f_{i,j=1,\ldots,n}$, are to minimized, the method is composed of the following steps.

***Step* 1.** Generate a representative sample of solutions from the currently considered part of the Pareto optimal set.

***Step* 2.** Present the sample to the DM, possibly together with association rules showing relationships between attainable values of objective functions in the Pareto optimal set.

***Step* 3.** If the DM is satisfied with one solution from the sample, then this is the most preferred solution and the procedure stops. Otherwise continue.

***Step* 4.** Ask the DM to indicate a subset of relatively "good" solutions in the sample.

***Step* 5.** Apply DRSA to the current sample of solutions classified into "good" and "others" solutions, in order to induce a set of decision rules with the following syntax "if $f_{j_1}(\mathbf{x}) \le \alpha_{j_1}$ and … and $f_{j_p}(\mathbf{x}) \le \alpha_{j_p}$, then solution $\mathbf{x}$ is good", $\{j_1, \ldots, j_p\} \subseteq \{1, \ldots, n\}$.

***Step* 6.** Present the obtained set of rules to the DM.

***Step* 7.** Ask the DM to select the decision rules most adequate to his/her preferences.

***Step* 8.** Adjoin the constraints $f_{j_1}(\mathbf{x}) \le \alpha_{j_1}, \ldots, f_{j_p}(\mathbf{x}) \le \alpha_{j_p}$ coming from the rules selected in *Step* 7 to the set of constraints imposed on the Pareto optimal set, in order to focus on a part interesting from the point of view of DM's preferences.

***Step* 9.** Go back to *Step* 1.

In a sequence of iterations, the method is exploring the Pareto optimal set of a multiobjective optimization problem or an approximation of this set. In the calculation stage (*Step* 1), any multiobjective optimization method, which finds the Pareto optimal set or its approximation, such as Evolutionary Multiobjective Optimization methods, can be used. In the dialogue stage of the method (*Step* 2 to 7), the DM is asked to select a decision rule induced from his/her preference information, which is equivalent to fixing some upper bounds for the minimized objective functions f_j.

In *Step* 1, the representative sample of solutions from the currently considered part of the Pareto optimal set can be generated using one of existing procedures, such as [33,57,58]. It is recommended to use a fine grained sample of representative solutions to induce association rules, however, the sample of solutions presented to the DM in *Step* 2 should be much smaller (about a dozen) in order to avoid an excessive cognitive effort of the DM. Otherwise, the DM would risk to give non reliable information.

The association rules presented in *Step* 2 help the DM in understanding what (s)he can expect from the optimization problem. More precisely, any association rule

"if $f_{i_1}(x) \le r_{i_1}$ and … and $f_{i_p}(x) \le r_{i_p}$, then $f_{i_{p+1}}(x) \ge r_{i_{p+1}}$ and … and $f_{i_q}(x) \ge r_{i_q}$", where $\{i_1, \ldots, i_q\} \subseteq I, r_{i_1}, \ldots, r_{i_q} \in \Re$

says to the DM that, if (s)he wants attain the values of objective functions $f_{i_1}(x) \le r_{i_1}$ and … and $f_{i_p}(x) \le r_{i_p}$, then (s)he cannot reasonably expect to obtain values of objective functions $f_{i_{p+1}}(x) < r_{i_{p+1}}$ and … and $f_{i_q}(x) < r_{i_q}$.

With respect to the ordinal classification of solutions into the two classes of "good" and "others", observe that "good" means in fact "relatively good", i. e. better than the rest. In case, the DM would refuse to classify as "good" any solution, one can ask the DM to specify some minimal requirements of the type $f_{j_1}(\mathbf{x}) \leq \alpha_{j_1}$ and … and $f_{j_p}(\mathbf{x}) \leq \alpha_{j_p}$ for "good" solutions. These minimal requirements give some constraints that can be used in *Step* 8, in the same way as the analogous constraints coming from selected decisions rules.

The rules considered in *Step* 5 have a syntax corresponding to minimization of objective functions. In case of maximization of an objective function f_j, the condition concerning this objective in the decision rule should have the form $f_j(\mathbf{x}) \geq \alpha_j$.

Remark, moreover, that the Pareto optimal set reduced in *Step* 8 by constraints $f_{j_1}(\mathbf{x}) \leq \alpha_{j_1}, \ldots, f_{j_p}(\mathbf{x}) \leq \alpha_{j_p}$ is certainly not empty if these constraints are coming from one decision rule only. Since robust rules (see the glossary Dominance-Based Rough Set Approach (DRSA)) are considered, the threshold values $\alpha_{j_1}, \ldots, \alpha_{j_p}$ are values of objective functions of some solutions from the Pareto optimal set. If $\{j_1, \ldots, j_p\} = \{1, \ldots, n\}$, i. e. $\{j_1, \ldots, j_p\}$ is the set of all objective functions, then the new reduced part of the Pareto optimal set contains only one solution $\mathbf{x}$ such that $f_1(\mathbf{x}) = \alpha_1, \ldots, f_n(\mathbf{x}) = \alpha_n$. If $\{j_1, \ldots, j_p\} \subset \{1, \ldots, n\}$, i. e. $\{j_1, \ldots, j_p\}$ is a proper subset of the set of all objective functions, then the new reduced part of the Pareto optimal set contains solutions satisfying conditions $f_{j_1}(\mathbf{x}) \leq \alpha_{j_1}$ and … and $f_{j_p}(\mathbf{x}) \leq \alpha_{j_p}$. Since the considered rules are robust, then there is at least one solution $\mathbf{x}$ satisfying these constraints. When the Pareto optimal set is reduced in *Step* 8 by constraints $f_{j_1}(\mathbf{x}) \leq \alpha_{j_1}, \ldots, f_{j_p}(\mathbf{x}) \leq \alpha_{j_p}$ coming from more than one rule, then it is possible that the resulting reduced part of the Pareto optimal set is empty. Thus, before passing to *Step* 9, it is necessary to verify if the reduced Pareto optimal set is not empty. If the reduced Pareto optimal set is empty, then the DM is required to revise his/her selection of rules. The DM can be supported in this task, by information about minimal sets of constraints $f_j(\mathbf{x}) \leq \alpha_j$ coming from the considered decision rules to be removed in order to get a non-empty part of the Pareto optimal set.

The constraints introduced in *Step* 8 are maintained in the following iterations of the procedure, however, they cannot be considered as irreversible. Indeed, the DM can come back to the Pareto optimal set considered in one of previous iterations and continue from this point. This is in the spirit of a learning oriented conception of interactive multiobjective optimization, i. e. it agrees with the idea that the interactive procedure permits the DM to learn about his/her preferences and about the "shape" of the Pareto optimal set.

Example Illustrating IMO-DRSA in the Context of Multiobjective Optimization

To illustrate the interactive multiobjective optimization procedure based on DRSA, a product mix problem is considered. There are three products: A, B, C which are produced in quantities denoted by x_A, x_B, and x_C, respectively. The unit prices of the three products are $p_A = 20$, $p_B = 30, p_C = 25$. The production process involves two machines. The production times of A, B, C on the first machine are equal to $t_{1A} = 5, t_{1B} = 8, t_{1C} = 10$, and on the second machine they are equal to $t_{2A} = 8, t_{2B} = 6$, $t_{2C} = 2$. Two raw materials are used in the production process. The first raw material has a unit cost of 6 and the quantity required for production of one unit of A, B and C is $r_{1A} = 1, r_{1B} = 2$ and $r_{1C} = 0.75$, respectively. The second raw material has a unit cost of 8 and the quantity required for production of one unit of A, B and C is $r_{2A} = 0.5, r_{2B} = 1$ and $r_{2C} = 0.5$, respectively. Moreover, the market cannot absorb a production greater than 10, 20 and 10 units for A, B and C, respectively. To decide how much of A, B and C should be produced, the following objectives have to be taken into account:

- Profit (to be maximized),
- Time (total production time on two machines – to be minimized),
- Production of A (to be maximized),
- Production of B (to be maximized),
- Production of C (to be maximized),
- Sales (to be maximized).

The above product mix problem can be formulated as the following multiobjective optimization problem:

$$\begin{aligned}
&\text{Maximize} \\
&20x_A + 30x_B + 25x_C - (1x_A + 2x_B + 0.75x_C)6 \\
&\qquad - (0.5x_A + 1x_B + 0.5x_C)8 \quad \text{[Profit]}, \\
&\text{Minimize} \quad 5x_A + 8x_B + 10x_C + 8x_A + 6x_B + 2x_C \quad \text{[Time]}, \\
&\text{Maximize} \quad x_A \quad \text{[Production of } A\text{]}, \\
&\text{Maximize} \quad x_B \quad \text{[Production of } B\text{]}, \\
&\text{Maximize} \quad x_C \quad \text{[Production of } C\text{]}, \\
&\text{Maximize} \quad 20x_A + 30x_B + 25x_C \quad \text{[Sales]},
\end{aligned}$$

Rough Sets in Decision Making, Table 11
A sample of Pareto optimal solutions proposed in the first iteration

Solution	Profit	Time	Prod. *A*	Prod. *B*	Prod. *C*	Sales	Evaluation
*s*1	165	120	0	0	10	250	∗
*s*2	172.69	130	0.769	0	10	265.38	∗
*s*3	180.38	140	1.538	0	10	280.77	good
*s*4	141.13	140	3	3	4.92	272.92	good
*s*5	148.38	150	5	2	4.75	278.75	good
*s*6	139.13	150	5	3	3.58	279.58	∗
*s*7	188.08	150	2.308	0	10	296.15	∗
*s*8	159	150	6	0	6	270	∗
*s*9	140.5	150	6	2	3.67	271.67	good
*s*10	209.25	200	6	2	7.83	375.83	∗
*s*11	189.38	200	5	5	5.42	385.42	∗
*s*12	127.38	130	3	3	4.08	252.08	∗
*s*13	113.63	120	3	3	3.25	231.25	∗

subject to:

$$x_A \leq 10\,, \quad x_B \leq 20\,, \quad x_C \leq 10$$
[Market absorption limits] ,

$$x_A \geq 0\,, \quad x_B \geq 0\,, \quad x_C \geq 0$$
[Non-negativity constraints] .

A sample of representative Pareto optimal solutions has been calculated and proposed to the DM. Observe that the considered problem is a Multiple Objective Linear Programming (MOLP) problem, and thus representative Pareto optimal solutions can be calculated using classical linear programming looking for the solutions optimizing each one of the considered objectives or fixing all the considered objective functions but one at a satisfying value, and looking for the solution optimizing the remaining objective function. The set of representative Pareto optimal solutions is shown in Table 11. Moreover, a set of potentially interesting association rules have been induced from the sample and presented to the DM. These rules represent strongly supported relationships between attainable values of objective functions. The association rules are the following (between parentheses there are id's of solutions supporting the rule):

1) if Time $\leq$ 140, then Profit $\leq$ 180.38 and Sales $\leq$ 280.77 ($s1, s2, s3, s4, s12, s13$),
2) if Time $\leq$ 150, then Profit $\leq$ 188.08 and Sales $\leq$ 296.15 ($s1, s2, s3, s4, s5, s6, s7, s8, s9, s12, s13$),
3) if $x_B \geq 2$, then Profit $\leq$ 209.25 and $x_A \leq 6$ and $x_C \leq 7.83$ ($s4, s5, s6, s9, s10, s11, s12, s13$),
4) if Time $\leq$ 150, then $x_B \leq 3$ ($s1, s2, s3, s4, s5, s6, s7, s8, s9, s12, s13$),
5) if Profit $\geq$ 148.38 and Time $\leq$ 150, then $x_B \leq 2$ ($s1, s2, s3, s5, s7, s8$),
6) if $x_A \geq 5$, then Time $\geq$ 150 ($s5, s6, s8, s9, s10, s11$),
7) if Profit $\geq$ 127.38 and $x_A \geq 3$, then Time $\geq$ 130 ($s4, s5, s6, s8, s9, s10, s11, s12$),
8) if Time $\leq$ 150 and $x_B \geq 2$, then Profit $\leq$ 148.38 ($s4, s5, s6, s9, s12, s13$),
9) if $x_A \geq 3$ and $x_C \geq 4.08$, then Time $\geq$ 130 ($s4, s5, s8, s10, , s11, s12$),
10) if Sales $\geq$ 265.38, then Time $\geq$ 130 ($s2, s3, s4, s5, s6, s7, s8, s9, s10, s11$).

Then, the DM has been asked if (s)he was satisfied with one of the proposed Pareto optimal solutions. Since his/her answer was negative, (s)he was requested to indicate a subset of relatively "good" solutions which are indicated in the "Evaluation" column of Table 11.

Taking into account the ordinal classification of Pareto optimal solutions into "good" and "others", made by the DM, twelve decision rules have been induced from the lower approximation of "good" solutions. The frequency of the presence of objectives in the premises of the rules gives a first idea of the importance of the considered objectives. These frequencies are the following:

- Profit: $\frac{4}{12}$,
- Time: $\frac{12}{12}$,
- Production of A : $\frac{7}{12}$,
- Production of B : $\frac{4}{12}$,
- Production of C : $\frac{5}{12}$,
- Sales: $\frac{5}{12}$.

The following potentially interesting decision rules were presented to the DM:

1) if Profit $\geq$ 140.5 and Time $\leq$ 150 and $x_B \geq 2$, then product mix is good ($s4$, $s5$, $s9$),
2) if Time $\leq$ 140 and $x_A \geq 1.538$ and $x_C \geq 10$, then product mix is good ($s3$),
3) if Time $\leq$ 150 and $x_B \geq 2$ and $x_C \geq 4.75$, then product mix is good ($s4$, $s5$),
4) if Time $\leq$ 140 and Sales $\geq$ 272.9167, then product mix is good ($s3$, $s4$),
5) if Time $\leq$ 150 and $x_B \geq 2$ and $x_C \geq 3.67$ and Sales $\geq$ 271.67, then product mix is good ($s4$, $s5$, $s9$).

Among these decision rules, the DM has selected rule 1) as the most adequate to his/her preferences. This rule permits to define the following constraints reducing the feasible region of the production mix problem:

- $20x_A+30x_B+25x_C-(x_A+2x_B+0.75x_C)6-(0.5x_A+x_B+0.5x_C)8 \geq 140.5$ [Profit $\geq$ 140.5],
- $5x_A + 8x_B + 10x_C + 8x_A + 6x_B + 2x_C \leq 150$ [Time $\leq$ 150],
- $x_B \geq 2$ [Production of $B \geq 2$].

These constraints have been considered together with the original constraints for the production mix problem, and a new sample of representative Pareto optimal solutions shown in Table 12 have been calculated and presented to the DM, together with the following potentially interesting association rules:

1′) if Time $\leq$ 140, then Profit $\leq$ 174 and $x_C \leq 9.33$ and Sales $\leq$ 293.33 ($s5'$, $s6'$, $s7'$, $s8'$, $s9'$, $s10'$, $s11'$, $s12'$),
2′) if $x_A \geq 2$, then $x_B \leq 3$ and Sales $\leq$ 300.83 ($s2'$, $s3'$, $s4'$, $s6'$, $s7'$, $s9'$),
3′) if $x_A \geq 2$, then Profit $\leq$ 172 and $x_C \leq 8$ ($s2'$, $s3'$, $s4'$, $s6'$, $s7'$, $s9'$),
4′) if Time $\leq$ 140, then $x_A \leq 2$ and $x_B \leq 3$ ($s5'$, $s6'$, $s7'$, $s8'$, $s9'$, $s10'$, $s11'$, $s12'$),
5′) if Profit $\geq$ 158.25, then $x_A \leq 2$ ($s1'$, $s3'$, $s4'$, $s5'$, $s6'$, $s8'$),
6′) if $x_A \geq 2$, then Time $\geq$ 130 ($s2'$, $s3'$, $s4'$, $s6'$, $s7'$, $s9'$),
7′) if $x_C \geq 7.17$, then $x_A \leq 2$ and $x_B \leq 2$ ($s1'$, $s3'$, $s5'$, $s6'$, $s8'$, $s10'$),
8′) if $x_C \geq 6$, then $x_A \leq 2$ and $x_B \leq 3$ ($s1'$, $s3'$, $s4'$, $s5'$, $s6'$, $s7'$, $s8'$, $s9'$, $s10'$, $s11'$, $s12'$),
9′) if $x_C \geq 7$, then Time $\geq$ 125 and $x_B \leq 2$ ($s1'$, $s3'$, $s5'$, $s6'$, $s8'$, $s10'$, $s11'$),
10′) if Sales $\geq$ 280, then Time $\geq$ 140 and $x_B \leq 3$ ($s1'$, $s2'$, $s3'$, $s4'$, $s5'$, $s7'$),
11′) if Sales $\geq$ 279.17, then Time $\geq$ 140 ($s1'$, $s2'$, $s3'$, $s4'$, $s5'$, $s6'$, $s7'$),
12′) if Sales $\geq$ 272, then Time $\geq$ 130 ($s1'$, $s2'$, $s3'$, $s4'$, $s5'$, $s6'$, $s7'$, $s8'$).

The DM has been asked again if (s)he was satisfied with one of the proposed Pareto optimal solutions. Since his/her answer was negative, (s)he was requested again to indicate a subset of relatively "good" solutions, which are indicated in the "Evaluation" column of Table 12.

Taking into account the ordinal classification of Pareto optimal solutions into "good" and "others", made by the DM, eight decision rules have been induced from the lower approximation of "good" solutions. The frequencies of the presence of objectives in the premises of the rules are the following:

- Profit: $\frac{2}{8}$,
- Time: $\frac{1}{8}$,
- Production of A: $\frac{5}{8}$,
- Production of B: $\frac{3}{8}$,
- Production of C: $\frac{3}{8}$,
- Sales: $\frac{2}{8}$.

The following potentially interesting decision rules were presented to the DM:

1) if Time $\leq$ 125 and $x_A \geq 1$, then product mix is good ($s11'$, $s12'$),
2) if $x_A \geq 1$ and $x_C \geq 7$, then product mix is good ($s3'$, $s6'$, $s11'$),
3) if $x_A \geq 1.5$ and $x_C \geq 6.46$, then product mix is good ($s3'$, $s4'$, $s6'$, $s12'$),
4) if Profit $\geq$ 158.25 and $x_A \geq 2$, then product mix is good ($s3'$, $s4'$, $s6'$),
5) if $x_A \geq 2$ and Sales $\geq$ 300, then product mix is good ($s3'$, $s4'$).

Among these decision rules, the DM has selected rule 4) as the most adequate to his/her preferences. This rule permits to define the following constraints reducing the Pareto optimal set of the production mix problem:

- $20x_A+30x_B+25x_C-(x_A+2x_B+0.75x_C)6-(0.5x_A+x_B+0.5x_C)8 \geq 158.25$ [Profit $\geq$ 158.25],
- $x_A \geq 2$ [Production of $A \geq 2$].

Let us observe that the first constraint is just strengthening an analogous constraint introduced in the first iteration (Profit $\geq$ 140.5).

Considering the new set of constraints, a new sample of representative Pareto optimal solutions shown in Table 13 has been calculated and presented to the DM, together with the following potentially interesting association rules:

Rough Sets in Decision Making, Table 12
A sample of Pareto optimal solutions proposed in the second iteration

Solution	Profit	Time	Prod. *A*	Prod. *B*	Prod. *C*	Sales	Evaluation
$s1'$	186.53	150	0.154	2	10	313.08	*
$s2'$	154.88	150	3	3	5.75	293.75	*
$s3'$	172	150	2	2	8	300	good
$s4'$	162.75	150	2	3	6.83	300.83	good
$s5'$	174	140	0	2	9.33	293.33	*
$s6'$	158.25	140	2	2	7.17	279.17	good
$s7'$	149	140	2	3	6	280	*
$s8'$	160.25	130	0	2	8.5	272	good
$s9'$	144.5	130	2	2	6.33	258.33	*
$s10'$	153.38	125	0	2	8.08	262.08	*
$s11'$	145.5	125	1	2	7	255	good
$s12'$	141.56	125	1.5	2	6.46	251.46	good

$1''$) if Time $\leq$ 145, then $x_A \leq 2$ and $x_B \leq 2.74$ and Sales $\leq$ 290.2 ($s2''$, $s3''$, $s4''$),
$2''$) if $x_C \geq 6.92$, then $x_A \leq 3$ and $x_B \leq 2$ and Sales $\leq$ 292.92 ($s3''$, $s4''$, $s5''$),
$3''$) if Time $\leq$ 145, then Profit $\leq$ 165.13 and $x_A \leq 2$ and $x_C \leq 7.58$ ($s2''$, $s3''$, $s4''$),
$4''$) if $x_C \geq 6.72$, then $x_B \leq 2.74$ ($s2''$, $s3''$, $s4''$, $s5''$),
$5''$) if Sales $\geq$ 289.58, then Profit $\leq$ 165.13 and Time $\geq$ 145 and $x_C \leq 7.58$ ($s1''$, $s2''$, $s3''$, $s5''$).

The DM has been asked again if (s)he was satisfied with one of the presented Pareto optimal solutions shown in Table 13 and this time (s)he declared that solution $s3''$ is satisfactory for him/her. This ends the interactive procedure.

Characteristics of the IMO-DRSA

The interactive procedure presented in Sect."Interactive Multiobjective Optimization Using DRSA (IMO-DRSA)" can be analyzed from the point of view of input and output information. As to the input, the DM gives preference information by answering easy questions related to ordinal classification of some representative solutions into two classes ("good" and "others"). Very often, in multiple criteria decision analysis, in general, and in interactive multiobjective optimization, in particular, the preference information has to be given in terms of preference model parameters, such as importance weights, substitution rates and various thresholds (see [6] for the Multiple Attribute Utility Theory and [1,5,34,46] for outranking methods; for some well-known interactive multiobjective optimization methods requiring preference model parameters, see the Geoffrion–Dyer–Feinberg method [9], the method of Zionts and Wallenius [60,61] and the Interactive Surrogate Worth Tradeoff method [2,3] requiring information in terms of marginal rates of substitution, the reference point method [58] requiring a reference point and weights to formulate an achievement scalarizing function, the Light Beam Search method [33] requiring information in terms of weights and indifference, preference and veto thresholds, being typical parameters of ELECTRE methods). Eliciting such information requires a significant cognitive effort on the part of the DM. It is generally acknowledged that people often prefer to make exemplary decisions and cannot always explain them in terms of specific parameters. For this reason, the idea of inferring preference models from exemplary decisions provided by the DM is very attractive. The output result of the analysis is the model of preferences in terms of "*if* …, *then* …" decision rules which is used to reduce the Pareto optimal set iteratively, until the DM selects a satisfactory solution. The decision rule preference model is very convenient for decision support, because it gives argumentation for preferences in a logical form, which is intelligible for the DM, and identifies the Pareto optimal solutions supporting each particular decision rule. This is very useful for a critical revision of the original ordinal classification of representative solutions into the two classes of "good" and "others". Indeed, decision rule preference model speaks the same language of the DM without any recourse to technical terms, like utility, tradeoffs, scalarizing functions and so on.

All this implies that IMO-DRSA has a transparent feedback organized in a learning oriented perspective, which permits to consider this procedure as a "glass box", contrary to the "black box" characteristic of many procedures giving final result without any clear explanation. Note that with the proposed procedure, the DM learns about the shape of the Pareto optimal set using the association rules. They represent relationships between attain-

Rough Sets in Decision Making, Table 13
A sample of Pareto optimal solutions proposed in the third iteration

Solution	Profit	Time	Prod. *A*	Prod. *B*	Prod. *C*	Sales	Evaluation
$s1''$	158.25	150	2	3.49	6.27	301.24	*
$s2''$	158.25	145	2	2.74	6.72	290.20	*
$s3''$	165.13	145	2	2	7.58	289.58	selected
$s4''$	158.25	140	2	2	7.17	279.17	*
$s5''$	164.13	150	3	2	6.92	292.93	*
$s6''$	158.25	145.3	3	2	6.56	284.02	*

able values of objective functions on the Pareto optimal set in logical and very natural statements. The information given by association rules is as intelligible as the decision rule preference model, since they speak the language of the DM and permit him/her to identify the Pareto optimal solutions supporting each particular association rule.

Thus, decision rules and association rules give an explanation and a justification of the final decision, that does not result from a mechanical application of a certain technical method, but rather from a mature conclusion of a decision process based on active intervention of the DM.

Observe, finally, that the decision rules representing preferences and the association rules describing the Pareto optimal set are based on ordinal properties of objective functions only. Differently from methods involving some scalarization (almost all existing interactive methods), in any step the proposed procedure does not aggregate the objectives into a single value, avoiding operations (such as averaging, weighted sum, different types of distance, achievement scalarization) which are always arbitrary to some extent. Observe that one could use a method based on a scalarization to generate the representative set of Pareto optimal solutions, nevertheless, the decision rule approach would continue to be based on ordinal properties of objective functions only, because the dialogue stage of the method operates on ordinal comparisons only. In the proposed method, the DM gets clear arguments for his/her decision in terms of "*if* . . . , *then* . . . " decision rules and the verification if a proposed solution satisfies these decision rules is particularly easy. This is not the case of interactive multiobjective optimization methods based on scalarization. For example, in the methods using an achievement scalarization function, it is not evident what does it mean for a solution to be "close" to the reference point. How to justify the choice of the weights used in the achievement function? What is their interpretation? Observe, instead, that the method proposed in this chapter operates on data using ordinal comparisons which would not be affected by any increasing monotonic transformation of scales, and this ensures the meaningfulness of results from the point of view of measurement theory (see e. g. [42]).

With respect to computational aspects of the method, notice that the decision rules can be calculated efficiently in few seconds only using the algorithms presented in [17,20]. When the number of objective functions is not too large to be effectively controlled by the DM (let us say seven plus or minus two, as suggested by Miller [37]), then the decision rules can be calculated in a fraction of one second. In any case, the computational effort grows exponentially with the number of objective functions, but not with respect to the number of considered Pareto optimal solutions, which can increase with no particularly negative consequence on calculation time.

Conclusions

Rough set theory is a mathematical tool for dealing with granularity of information and possible inconsistencies in the description of objects. Considering this description as an input data about a decision problem, the rough set approach permits to structure this description into lower and upper approximations, corresponding to certain and possible knowledge about the problem. Induction algorithms run on these approximations discover, in turn, certain and possible decision rules that facilitate an understanding of the DM's preferences, and that enable a recommendation concordant with these preferences.

The original version of the rough set approach, based on indiscernibility or similarity relation, and typical induction algorithms considered within machine learning, data mining and knowledge discovery, deal with data describing problems of taxonomy-type classification, i. e. problems where neither the attributes describing the objects, nor the classes to which the objects are assigned, are ordered. On the other hand, multiple criteria decision making deals with problems where descriptions (evaluations) of objects by means of attributes (criteria), as well as decisions in classification, choice and ranking problems,

are ordered. Moreover, in data describing multiple criteria decision making, there exist a monotonic relationship between conditions and decisions, like "the bigger the house, the more expensive it is". The generalization of the rough set approach and of the induction algorithms about problems in which order properties and monotonic relationships are important gave birth to the Dominance-based Rough Set Approach (DRSA) which made a breakthrough in scientific decision support.
The main features of DRSA are the following:

- preference information necessary to deal with any multiple criteria decision problem, or with decision under uncertainty, is asked to the DM just in terms of exemplary decisions,
- the rough set analysis of preference information supplies some useful elements of knowledge about the decision situation; these are: the relevance of attributes or criteria, the minimal subsets of attributes or criteria (reducts) conveying the relevant knowledge contained in the exemplary decisions, the set of indispensable attributes or criteria (core),
- DRSA can deal with preference information concerning taxonomy-type classification, ordinal classification, choice, ranking, multiobjective optimization and decision under uncertainty,
- the preference model induced from preference information structured by DRSA is expressed in a natural and comprehensible language of "*if* ..., *then* ..." decision rules,
- suitable procedures have been proposed to exploit the results of application of the decision rule preference model on a set of objects or pairs of objects in order to workout a recommendation,
- no prior discretization of quantitative condition attributes or criteria is necessary,
- heterogeneous information (qualitative and quantitative, ordered and non-ordered, nominal and ordinal, quantitative and numerical non-quantitative scales of preferences) can be processed within DRSA,
- the proposed methodology fulfills some desirable properties for both rough set approach (the approximated sets include lower approximation and are included in upper approximation, and the complementarity property is satisfied), and for multiple criteria decision analysis (the decision rule preference model is formally equivalent to the non-additive, non-transitive and non-complete conjoint measurement model, and to a more general model for preferences defined on all kinds of scales),
- the decision rule preference model resulting from DRSA is more general than all existing models of conjoint measurement, due to its capacity of handling inconsistent preferences (a new model of conjoint measurement is formally equivalent to the decision rule preference model handling inconsistencies),
- the decision rule preference model fulfills the postulate of transparency and interpretability of preference models in decision support; each decision rule can be clearly identified with those parts of the preference information (decision examples) which support the rule; the rules inform the DM in a quasi-natural language about the relationships between conditions and decisions; in this way, the rules permit traceability of the decision support process and give understandable justifications for the decision to be made,
- the proposed methodology is based on elementary concepts and mathematical tools (sets and set operations, binary relations), without recourse to any algebraic or analytical structures.

The decision rules entering the preference model have a special syntax which involves partial evaluation profiles and dominance relations on these profiles. The traditional preference models, which are the utility function and the outranking relation, can be represented in terms of equivalent decision rules. The clarity of the rule representation of preferences enables one to see the limits of these aggregation functions. Several studies [22,23,26,50] presented an axiomatic characterization of all three kinds of preference models in terms of conjoint measurement theory and in terms of a set of decision rules. The given axioms of "cancellation property" type are the weakest possible. In comparison to other studies on the characterization of aggregation functions, these axioms do not require any preliminary assumptions about the scales of criteria. A side-result of these investigations is that the decision rule preference model is the most general among the known aggregation functions.

Future Directions

The article shows that the Dominance-based Rough Set Approach is a very powerful tool for decision analysis and support. Its potential goes, however, beyond the theoretical frame considered in this article. There are many possibilities of applying DRSA to real life problems. The non-exhaustive list of potential applications includes:

Decision support in medicine: In this area there are already many interesting applications (see, e. g., [35,36,40,59]), however, they exploit the classical rough set approach; applications requiring DRSA, which handle

ordered value sets of medical signs, as well as monotonic relationship between the value of signs and the degree of gravity of a disease, are in progress;

Customer satisfaction survey: Theoretical foundations for application of DRSA in this field are available in [30], however, a fully documented application is still missing;

Bankruptcy risk evaluation: This is a field of many potential applications, as can be seen from promising results reported, e. g. in [12,53,54], however, a wider comparative study involving real data sets is needed;

Operational research problems, such as location, routing, scheduling or inventory management: These are problems formulated either in terms of classification of feasible solutions (see, e. g., [11]), or in terms of interactive multiobjective optimization, for which the IMO-DRSA [32] procedure is suitable;

Finance: This is a domain where DRSA for decision under uncertainty has to be combined with interactive multiobjective optimization using IMO-DRSA; some promising results going in this direction have been presented in [31];

Ecology: Assessment of the impact of human activity on the ecosystem is a challenging problem for which the presented methodology is suitable; the up to date applications are based on the classical rough set concept (see, e. g., [7,43]), however, it seems that DRSA handling ordinal data has a greater potential in this field.

Bibliography

1. Brans JP, Mareschal B (2005) PROMETHEE Methods. In: Figueira J, Greco S, Ehrgott M (eds) Multiple criteria decision analysis: State of the art surveys. Springer, Berlin, pp 163–195
2. Chankong V, Haimes YY (1978) The interactive surrogate worth trade-off (ISWT) method for multiobjective decision-making. In: Zionts S (ed) Multiple Criteria Problem Solving. Springer, Berlin, pp 42–67
3. Chankong V, Haimes YY (1983) Multiobjective decision making theory and methodology. Elsiever Science, New York
4. Figueira J, Greco S, Ehrgott M (eds) (2005) Multiple criteria decision analysis: State of the art surveys. Springer, Berlin
5. Figueira J, Mousseau V, Roy B (2005) ELECTRE methods. In: Figueira J, Greco S, Ehrgott M (eds) Multiple criteria decision analysis: State of the art surveys. Springer, Berlin, pp 133–162
6. Fishburn PC (1967) Methods of estimating additive utilities. Manag Sci 13(7):435–453
7. Flinkman M, Michalowski W, Nilsson S, Słowiński R, Susmaga R, Wilk S (2000) Use of rough sets analysis to classify Siberian forest ecosystem according to net primary production of phytomass. INFOR 38:145–161
8. Fortemps P, Greco S, Słowiński R (2008) Multicriteria decision support using rules that represent rough-graded preference relations. Eur J Oper Res 188:206–223
9. Geoffrion A, Dyer J, Feinberg A (1972) An interactive approach for multi-criterion optimization, with an application to the operation of an academic department. Manag Sci 19(4):357–368
10. Giove S, Greco S, Matarazzo B, Słowiński R (2002) Variable consistency monotonic decision trees. In: Alpigini JJ, Peters JF, Skowron A, Zhong N (eds) Rough sets and current trends in computing. LNAI, vol 2475. Springer, Berlin, pp 247–254
11. Gorsevski PV, Jankowski P (2008) Discerning landslide susceptibility using rough sets. Comput Environ Urban Syst 32:5365
12. Greco S, Matarazzo B, Słowiński R (1998) A new rough set approach to evaluation of bankruptcy risk. In: Zopounidis C (ed) Operational tools in the management of financial risks. Kluwer, Dordrecht, pp 121–136
14. Greco S, Matarazzo B, Słowiński R (1999) The use of rough sets and fuzzy sets in MCDM. In: Gal T, Stewart T, Hanne T (eds) Advances in multiple criteria decision making. Kluwer, Boston, pp 14.1–14.59
15. Greco S, Matarazzo B, Słowiński R (1999) Rough approximation of a preference relation by dominance relations. Eur J Oper Res 117:63–83
16. Greco S, Matarazzo B, Słowiński R (2000) Extension of the rough set approach to multicriteria decision support. INFOR 38:161–196
13. Greco S, Matarazzo B, Słowiński R, Tsoukias A (1998) Exploitation of a rough approximation of the outranking relation in multicriteria choice and ranking. In: Stewart TJ, van den Honert RC (eds) Trends in multicriteria decision making. LNEMS, vol 465. Springer, Berlin, pp 45–60
17. Greco S, Matarazzo B, Słowiński R, Stefanowski J (2001) An algorithm for induction of decision rules consistent with dominance principle. In: Ziarko W, Yao Y (eds) Rough sets and current trends in computing. LNAI, vol 2005. Springer, Berlin, pp 304–313
18. Greco S, Matarazzo B, Słowiński R (2001) Rough sets theory for multicriteria decision analysis. Eur J Oper Res 129:1–47
19. Greco S, Matarazzo B, Słowiński R (2001) Rough set approach to decisions under risk. In: Ziarko W, Yao Y (eds) Rough sets and current trends in computing. LNAI, vol 2005. Springer, Berlin, pp 160–169
20. Greco S, Matarazzo B, Słowiński R, Stefanowski J (2002) Mining association rules in preference-ordered data. In: Hacid M-S, Ras ZW, Zighed DA, Kodratoff Y (eds) Foundations of intelligent systems. LNAI, vol 2366. Springer, Berlin, pp 442–450
21. Greco S, Matarazzo B, Slowinski R (2002) Multicriteria classification. In: Kloesgen W, Zytkow J (eds) Handbook of data mining and knowledge discovery. Oxford University Press, New York, pp 318–328
22. Greco S, Matarazzo B, Słowiński R (2002) Preference representation by means of conjoint measurement & decision rule model. In: Bouyssou D, Jacquet-Lagrèze E, Perny P, Słowiński R, Vanderpooten D, Vincke P (eds) Aiding decisions with multiple criteria – essays in honor of Bernard Roy. Kluwer, Dordrecht, pp 263–313
23. Greco S, Matarazzo B, Słowiński R (2004) Axiomatic characterization of a general utility function and its particular cases in terms of conjoint measurement and rough-set decision rules. Eur J Oper Res 158:271–292
24. Greco S, Matarazzo B, Słowiński R (2004) Dominance-based Rough Set Approach to Knowledge Discovery, (I) – General Perspective, (II) – Extensions and Applications. In: Zhong N,

Liu J (eds) Intelligent technologies for information analysis. Springer, Berlin, pp 513–612
25. Greco S, Pawlak Z, Słowiński R (2004) Can Bayesian confirmation measures be useful for rough set decision rules? Eng Appl Artif Intell 17:345–361
26. Greco S, Matarazzo B, Słowiński R (2005) Decision rule approach. In: Figueira J, Greco S, Ehrgott M (eds) Multiple criteria decision analysis: State of the art surveys. Springer, Berlin, pp 507–563
27. Greco S, Matarazzo B, Słowiński R (2006) Dominance-based rough set approach to decision involving multiple decision makers. In: Greco S, Hata Y, Hirano S, Inuiguchi M, Miyamoto S, Nguyen HS, Slowinski R (eds) Rough Sets and Current Trends in Computing, RSCTC. LNAI, vol 4259. Springer, Berlin, pp 306–317
28. Greco S, Matarazzo B, Słowiński R (2006) Dominance-based rough set approach to case-based reasoning. In: Torra V, Narukawa Y, Valls A, Domingo-Ferrer J (eds) Modelling decisions for artificial intelligence. LNAI, vol 3885. Springer, Berlin, pp 7–18
29. Greco S, Matarazzo B, Słowiński R (2007) Dominance-based rough set approach as a proper way of handling graduality in rough set theory. In: Transactions on Rough Sets VII. LNCS, vol 4400. Springer, Berlin, pp 36–52
30. Greco S, Matarazzo B, Słowiński R (2007) Customer satisfaction analysis based on rough set approach. Z Betr 16(3):325–339
31. Greco S, Matarazzo B, Słowiński R (2007) Financial portfolio decision analysis using dominance-based rough set approach. Invited paper at the 22nd European Conference on Operational Research (EURO XXII), Prague, 08–11 July
32. Greco S, Matarazzo B, Słowiński R (2008) Dominance-based rough set approach to interactive multiobjective optimization. In: Branke J, Deb K, Miettinen K, Slowinski R (eds) Multiobjective optimization: Interactive and evolutionary approaches. Springer, Berlin
33. Jaszkiewicz A, Słowiński R (1999) The "Light Beam Search" approach – an overview of methodology and applications. Eur J Oper Res 113:300–314
34. Martel JM, Matarazzo B (2005) Other outranking approaches. In: Figueira J, Greco S, Ehrgott M (eds) Multiple criteria decision analysis: State of the art surveys. Springer, Berlin, pp 197–262
35. Michalowski W, Rubin S, Słowiński R, Wilk S (2003) Mobile clinical support system for pediatric emergencies. J Decis Support Syst 36:161–176
36. Michalowski W, Wilk S, Farion K, Pike J, Rubin S, Słowiński R (2005) Development of a decision algorithm to support emergency triage of scrotal pain and its implementation in the MET system. INFOR 43:287–301
37. Miller GA (1956) The magical number seven, plus or minus two: some limits in our capacity for processing information. Psychol Rev 63:81–97
38. Pawlak Z (1982) Rough Sets. Int J Comput Inf Sci 11:341–356
39. Pawlak Z (1991) Rough Sets. Kluwer, Dordrecht
40. Pawlak Z, Słowiński K, Słowiński R (1986) Rough classification of patients after highly selective vagotomy for duodenal ulcer. Int J Man-Machine Stud 24:413–433
41. Pawlak Z, Słowiński R (1994) Rough set approach to multi-attribute decision analysis. Eur J Oper Res 72:443–459
42. Roberts F (1979) Measurement theory, with applications to decision making, utility and the social sciences. Addison-Wesley, Boston
43. Rossi L, Słowiński R, Susmaga R (1999) Rough set approach to evaluation of stormwater pollution. Int J Environ Pollut 12:232–250
44. Roy B (1996) Multicriteria Methodology for Decision Aiding. Kluwer, Dordrecht
45. Roy B (1999) Decision-aiding today: What should we expect. In: Gal T, Stewart T, Hanne T (eds) Advances in Multiple Criteria Decision Making. Kluwer, Boston, pp 1.1–1.35
46. Roy B, Bouyssou D (1993) Aide Multicritère à la Décision: Méthodes et Cas. Economica, Paris
47. Shoemaker PJH (1982) The expected utility model: its variants, purposes, evidence and limitations. J Econ Lit 20:529–562
48. Słowiński R (1993) Rough set learning of preferential attitude in multi-criteria decision making. In: Komorowski J, Ras ZW (eds) Methodologies for intelligent systems. LNAI, vol 689. Springer, Berlin, pp 642–651
49. Słowiński R, Greco S, Matarazzo B (2002) Rough set analysis of preference-ordered data. In: Alpigini JJ, Peters JF, Skowron A, Zhong N (eds) Rough sets and current trends in computing. LNAI, vol 2475. Springer, Berlin, pp 44–59
50. Słowiński R, Greco S, Matarazzo B (2002) Axiomatization of utility, outranking and decision-rule preference models for multiple-criteria classification problems under partial inconsistency with the dominance principle. Control Cybern 31:1005–1035
51. Słowiński R, Greco S, Matarazzo B (2005) Rough set based decision support, chapter 16. In: Burke EK, Kendall G (eds) Search methodologies: Introductory tutorials in optimization and decision support techniques. Springer, New York, pp 475–527
52. Słowiński R, Greco S, Matarazzo B (2007) Dominance-based rough set approach to reasoning about ordinal data, keynote lecture. In: Kryszkiewicz M, Peters JF, Rybiński H, Skowron A (eds) Rough sets and intelligent systems paradigms. LNAI, vol 4585. Springer, Berlin, pp 5–11
53. Słowiński R, Zopounidis C (1995) Application of the rough set approach to evaluation of bankruptcy risk. Int J Intell Syst Acc Finance Manag 4:27–41
54. Słowiński R, Zopounidis C, Dimitras AI (1997) Prediction of company acquisition in Greece by means of the rough set approach. Eur J Oper Res 100:1–15
55. Stefanowski J (1998) On rough set based approaches to induction of decision rules. In: Polkowski L, Skowron A (eds) Rough sets in data mining and knowledge discovery, vol 1. Physica, Heidelberg, pp 500–529
56. Tsoukias A, Vincke P (1995) A new axiomatic foundation of the partial comparability theory. Theory Decis 39:79–114
57. Steuer RE, Choo E-U (1983) An interactive weighted Tchebycheff procedure for multiple objective programming. Math Program 26:326–344
58. Wierzbicki AP (1980) The use of reference objectives in multiobjective optimization. In: Fandel G, Gal T (eds) Multiple criteria decision making, theory and applications. Springer, Berlin, pp 468–486
59. Wilk S, Słowiński R, Michalowski W, Greco S (2005) Supporting triage of children with abdominal pain in the emergency room. Eur J Oper Res 160:696–709
60. Zionts S, Wallenius J (1976) An interactive programming method for solving the multiple criteria problem. Manag Sci 22:652–663
61. Zionts S, Wallenius J (1983) An interactive multiple objective linear programming method for a class of underlying nonlinear utility functions. Manag Sci 29:519–523

Rough Sets: Foundations and Perspectives

James F. Peters[1], Andrzej Skowron[2], Jarosław Stepaniuk[3]
[1] Computational Intelligence Laboratory, University of Manitoba, Winnipeg, Canada
[2] Institute of Mathematics, Warsaw University, Warsaw, Poland
[3] Computer Science, Białystok University of Technology, Białystok, Poland

Article Outline

Glossary

Approximation The replacement of mathematical objects by others that resemble them in certain respects [64].

Approximation space An approximation space is denoted by $(\mathcal{O}, \mathcal{F}, \sim_B)$, where $\mathcal{O}$ is a set of perceived objects, $\mathcal{F}$ is a set of probe functions representing object features, and $\sim_B$ is an indiscernibility relation defined relative to $B \subseteq \mathcal{F}$. This approximation space is considered fundamental because it provided a framework for the original rough set theory [37,40]. Several generalizations of this definition of approximation space have been proposed (see, e. g., [40,44,54,55,56,58,59,69]).

Attribute A quality regarded as characteristic or inherent in an object [29]. In rough set theory, an attribute a of an object x is represented by a partial function $f_a(x) = v$, where v is a value in the range of f_a. In rough set theory, the function f_a is often called an attribute [38,45].

Boundary region The B-boundary region of an approximation of a set X is denoted by $\mathrm{Bnd}_B X$ and is defined relative to a set of functions B representing features of objects in X as well as the lower approximation B_*X and the upper approximation B^*X, where

$$\mathrm{Bnd}_B X = B^*X \backslash B_*X = \{x \mid x \in B^*X \text{ and } x \notin B_*X\}.$$

Elementary set A B-class in the quotient set $X/\sim_B$.

Equivalence class Given a relation $\sim$, an equivalence class is a set denoted by $[x]$ or $[x]_\sim$ [10] in the quotient set $X/\sim$ (See Glossary item "Quotient set"), where

$$[x] = \{x' \in X \mid x \sim x'\}.$$

Equivalence relation A reflexive, symmetric and transitive relation $\sim \subseteq X \times X$. An equivalence relation $\sim$ on a set X defines a partition of X into classes.

Feature Make, form, fashion, shape (of an object) [29]. A characteristic of an object perceived by the senses or knowable by the mind [41,52]. In rough set theory, a feature f of an object x is represented by a function $\phi_f(x) = v$, where v is a value in the range of ϕ_f (e. g., $\phi_{\text{♪}}(x)$ as a measure of the tonality ♪ feature of a Chopin Mazurka x) [41,52]. The function ϕ_f is sometimes also termed an attribute [38,45].

Indiscernibility relation An equivalence relation

$$\sim_B = \{(x, x') \in X \times X | f(x) = f(x') \text{ for any } f \in B\},$$

where X denotes a set of objects, B denotes a set of functions, and $f \in B$ is a function representing a feature of an object $x \in X$. The notation used to denote an equivalence relation in rough set theory has varied widely over time. For example, $\widetilde{B}$ was originally introduced by Zdzisław Pawlak in 1981 [37]. Later, $\mathrm{Ind}(B)$ [18,30,38,66] or IND_B [45] or Ind_B [14] or IND [66] or $I(B)$ [40] or $=_B$ [16] has also been used to denote an equivalence relation on a set X defined relative to attributes of objects. In rough set theory, the equivalence relation $\sim_B$ was introduced by Zdzisław Pawlak [37].

Information granule Information granules are obtained in the process of granulation. Granulation can be viewed as a human way of achieving data compression and it plays a key role in implementing the divide-and-conquer strategy in human problem-solving. An *information granule* represents a set of objects that have descriptions matching the granule [52], e. g. elementary set $[x]_B$, lower approximation B_*X, quotient set $X/\sim_B$.

Lower approximation The B-lower approximation of a set X is denoted by B_*X and is defined relative to a set of functions B representing features of objects in X and the quotient set $X/\sim_B$, where

$$B_*X = \bigcup_{x:[x]_B \subseteq X} [x]_B.$$

Object Something perceptible to the senses or knowable by the mind [29].

Information Whatever is conveyed or represented by a particular sequence of symbols [29]. In rough set theory, information is derivable either from the patterns in a particular information table or from what can be observed in a particular approximation space [37,40].

Information system A system to represent knowledge [25,36,40]. Syntactic representation of knowledge in table form [25,45].

Partition of a non-empty set X A family of non-empty, pairwise disjoint subsets of X (called classes) such that the union of this family is equal to X.

Quotient set Set of all classes in a partition defined by an equivalence relation $\sim$ on a set X (denoted by $X/\sim$).

Rough set A set X is considered a rough set if, and only if it has a non-empty B-boundary $\mathrm{Bnd}_B X$, i. e., the B-approximation of X has a non-empty boundary.

Upper approximation The B-upper approximation of X is denoted by B^*X and is defined relative to a set of functions B representing features of objects in X and the quotient set $X/\sim_B$, where

$$B^*X = \bigcup_{x:[x]_B \cap X \neq \emptyset} [x]_B \ .$$

Definition of the Subject

Rough set theory was introduced by Zdzisław Pawlak (1926–2006) during the early 1980s. This theory has two distinguishing hallmarks: knowledge description systems introduced by Pawlak in 1973 [36] and set approximation (See Glossary Items "Approximation", "Boundary region", "Lower approximation", "Upper approximation") as a means of classifying a set [37]. Knowledge description systems represent our knowledge about sample objects in tabular form. This knowledge results from identifying a set of attributes A (apparent qualities) of the objects. A total function $\rho : X \times A \longrightarrow V_a$ maps an attribute $a \in A$ of an object $x \in X$ to a value $v \in V_a$. Our knowledge about objects $x_1, \ldots, x_n$ having attributes $a_1, \ldots, a_k$ then can be represented in table form as rows of knowledge descriptions, e. g., description of x_i as a tuple.

$$(\rho(x_i, a_1), \rho(x_i, a_2), \ldots, \rho(x_i, a_k)) \ .$$

At this point, there is a natural transition to the formulation of an indiscernibility relation $\sim_B$ that defines a partition of a set of sample objects relative to set $B \subseteq A$ (see Glossary item "Indiscernibility relation"). In many ways, this relation is both fecund and important. The fecundity (high fertility) of this relation can be seen in several ways. First, $\sim_B$ makes it possible to organize tabular representations of objects into equivalence classes. This, of course, means that now our observations can be made at the level of classes (elementary granules) instead of the more intransigent level of individual objects. Second, the majesty of $\sim_B$ can be seen in Pawlak's discovery of lower and upper descriptions of objects. These descriptions accrue naturally from lower and upper approximations of a set of objects. Third, it is then possible gauge the *roughness* of our knowledge about a set of objects by considering the size of the boundary of an approximation. The roughness of our knowledge about a set of sample objects is directly proportional to the size of the approximation boundary. So, in effect, the indiscernibility relation lead to the discovery of rough sets.

The importance of the indiscernibility relation can be seen in a number of ways, if one considers the ordinary difficulties associated with the complexity of ordinary, perceptual objects. The introduction of this relation led to the creation of an approximation space that provides a framework for perception or observation on the level of classes [33,37]. An important byproduct of the approach to approximation in rough set theory is information granulation [52]. Each selection of features of a set of sample objects leads to granulation of the information associated with the objects. For a set of sample objects X with selected features represented by a set of functions B, a rich harvest of information granules from the partition of X defined by the indiscernibility relation $\sim_B$, e. g., individual classes $[x]_B$, quotient space $X/\sim_B$, lower approximation B_*X, upper approximation B^*X, and boundary $\mathrm{Bnd}_B X$. From the introduction of knowledge description systems and approximation of our knowledge springs various forms of learning, starting with learning from examples exemplified by a number rough set toolsets released during the past decade (see, e. g., [20,50]).

Basic ideas of rough set theory and its extensions, as well as many interesting applications can be found in numerous books [50], issues of the Transactions on Rough Sets [62], special issues of other journals, proceedings of international conferences, tutorials or surveys (see, references and further readings listed in this paper and, e. g., [40] and numerous web pages such as [20,63,65]).

Introduction

This chapter gives a concise overview of some of the features of the foundations and perspectives of rough set theory. The foundations of rough set theory are rich and varied. These foundations include the study of algebras and algebraic structures associated with rough sets [2,8], entropy [7,61], lattice theory [9,23], similarity and indiscernibility [34], rough-satisfiability [14], logic [11,45],

and rough mereology [46]. In general, an overview of the mathematical foundations of rough sets is given by Lech Polkowski [45,46].

In addition to providing a fairly comprehensive inventory of some of the highlights of rough set-based research, this chapter gives a capsule view of several of the hallmarks of rough set theory, namely, approximation, approximation spaces, discovery of rough sets, and feature set reduction (reducts). For many other issues related to rough set theory and its applications, the reader is referred to the bibliography on rough sets.

This chapter has the following organization. The rough set approach to approximation spaces is given in Sect. "Approximation Spaces". This is followed in Sect. "Rough Sets" by an overview of the general notion of rough sets. The rough set approach to dimensionality reduction is briefly presented in Sect. "Dimensionality Reduction". A capsule view of the future of rough sets and their applications is given in Sect. "Future Directions".

Approximation Spaces

In rough set theory, approximations are carried out within the context of an approximation space $(\mathcal{O}, \mathcal{F}, \sim_B)$, where $\mathcal{O}$ is a set of objects, $\mathcal{F}$ is a set of functions representing object features, and $\sim_B$ is an indiscernibility relation defined relative to $B \subseteq \mathcal{F}$. This space is considered fundamental because it provided a framework for the original rough set theory [37]. It has also been observed that an approximation space is the formal counterpart of perception [33]. Approximation starts with the partition $\mathcal{O}/\sim_B$ of $\mathcal{O}$ defined by the relation $\sim_B$. Next, any set $X \subseteq \mathcal{O}$ is approximated by considering the relation between X and the classes $[x]_B \in \mathcal{O}/\sim_B, x \in \mathcal{O}$.

An approximation space $(\mathcal{O}, \mathcal{F}, \sim_B)$ defined relative to a set of perceptual objects $\mathcal{O}$, a set functions $\mathcal{F}$, and relation $\sim_B$, has the following basic framework that facilitates observations concerning sample objects.

Framework for an Approximation Space

$$
\begin{aligned}
\mathcal{O} &= \text{Set of perceived objects,}\\
\mathcal{F} &= \text{Set of probe functions objects,}\\
B &\subseteq \mathcal{F},\\
\sim_B &= \{(x, x') \in X \times X \mid f(x) = f(x') \text{ for any } f \in B\},\\
\mathcal{O}/\sim_B &= \{[x]_B \mid x \in \mathcal{O}, B \subseteq \mathcal{F}\}, B\text{-partition of } \mathcal{O},\\
[x]_B &\in \mathcal{O}/\sim_B,\\
X &\subseteq \mathcal{O}, \text{ set of objects of interest,}
\end{aligned}
$$

$$
\begin{aligned}
B_*X &= \bigcup_{x:[x]_B \subseteq X} [x]_B, B\text{-lower approximation,}\\
B^*X &= \bigcup_{x:[x]_B \cap X \neq \emptyset} [x]_B, B\text{-upper approximation,}\\
\mathrm{Bnd}_B X &= B^*X \setminus B_*X, B\text{-boundary region.}
\end{aligned}
$$

Affinities between objects of interest in the set $X \subseteq \mathcal{O}$ and classes in the quotient set $\mathcal{O}/\sim_B$ can be discovered by identifying those classes that have objects in common with X. Approximation of the set X begins by determining which elementary sets $[x]_B \in \mathcal{O}/\sim_B$ are subsets of X. This discovery process leads to the B-lower approximation of $X \subseteq \mathcal{O}$ denoted by B_*X. The B-upper approximation B^*X is a sum of equivalence classes $[x]_B \in \mathcal{O}/\sim_B$, where each class included in B^*X contains at least one object with a description that matches the description of an object in X. The lower and upper approximations of X provide a basis for defining the boundary of an approximation. Notice that B_*X is always a subset of B^*X.

Several generalizations of the classical rough set approach based on approximation spaces defined as pairs of the form $(\mathcal{O}, \mathcal{F}, \sim_B), B \subseteq F$, have been reported in the literature [44,45,54]. Among them it is worthwhile mentioning the rough set approach based on similarity (tolerance) relations and the approach to approximation of vague concepts based on the adaptive extension of approximation spaces [54] and approximations spaces that consider the nearness of objects [44].

Moreover, the approach based on inclusion functions has been generalized to the *rough mereological approach* [32,45,46]. The inclusion relation $x\mu_r y$ with the intended meaning *x is a part of y to a degree at least r* has been taken as the basic notion of the rough mereology being a generalization of the Leśniewski mereology [26]. Research on rough mereology has shown importance of another notion, namely *closeness* of compound objects (e. g., concepts). This can be defined by $x\,\mathrm{cl}_{r,r'}\,y$ if and only if $x\mu_r y$ and $y\mu_{r'}x$.

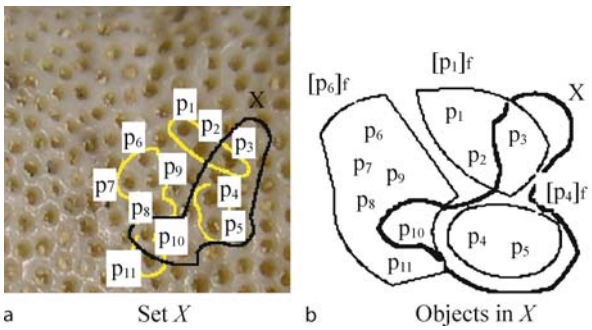
Rough mereology offers a methodology for synthesis and analysis of objects in a distributed environment of intelligent agents, in particular, for synthesis of objects satisfying a given specification to a satisfactory degree or for control in such a complex environment. Moreover, rough mereology has been used for developing the foundations of the *information granule calculi*, aiming at formalization of the Computing with Words paradigm, recently formulated by Lotfi Zadeh [67,68]. More complex information granules are defined recursively using already defined information granules and their measures of inclusion and closeness. Information granules can have complex structures like classifiers or approximation spaces. Com-

putations on information granules lead to the discovery of relevant information granules, e. g., information granules useful in compound concept approximations, pattern recognition and in machine learning. For example, families of approximation spaces labeled by some parameters are considered. By tuning such parameters according to chosen criteria (e. g., minimal description length), one can search for an optimal approximation space.

Rough Sets

In its original conception by Zdzisław Pawlak during the early 1980s, the discovery of the rough set approach to classification of sets of objects resulted from work on knowledge description systems begun by Pawlak during the early 1970s. In the rough set approach, the roughness of our knowledge of a set of objects is gauged in terms of a new approach to approximation that pivots on object descriptions commonly found in information tables also introduced by Pawlak. Approximation is carried out at the level of sets relative to our knowledge about objects of interest. That is, one approximates one set by another set. This approximation is built on a feature space for individual objects. A feature space for objects is defined relative to functions that map sample objects (pre-images) to measurements (images) that gauge the extent of our knowledge about the appearance of the objects. The basic idea is to select a set X of objects of interest and approximate it with another set that resembles X. This means that approximation is carried out at two different levels in the rough set approach.

1. **Object Level**. For a given set of objects design a set of functions $\phi = \{\phi_1, \ldots, \phi_n\}$ that represent features of sample objects of interest. The description of an object x is approximated by an n-tuple $(\phi_1(x), \ldots, \phi_n(x))$ of feature measurements.



Rough Sets: Foundations and Perspectives, Figure 1
Sample $X \subseteq \mathcal{O}$ to be approximated

2. **Set Level**. Define the partition of objects in the sample space $\mathcal{O}$ using the indiscernibility relation $\sim_B$, where $B \subseteq \phi$. Approximate a set X containing objects of interest relative to the elementary sets in $X/\sim_B$. Approximation of a set X containing objects with features represented by functions ϕ is carried out in terms of the lower approximation B_*X and the upper approximation B^*X.

Example 1 (Sample approximation of a set of coral pores) For this example, consider the set of coral pores in the coral fragment shown in Fig. 1(b). Then sets $\mathcal{O}, \mathcal{X}, \mathcal{B}$ and an approximation of X are defined in the following way:

$$\begin{aligned}
\mathcal{O} &= \{p \mid p \text{ equals a labeled coral pore in Fig. 1}\} \\
&= \{p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9, p_{10}, p_{11}\}, \\
\mathcal{F} &= \{f\}, \text{ where } f : \mathcal{O} \to \Re, \\
f(p_i) &= \text{ depth (in mm) of a coral pore } p_i \in X, \\
\mathcal{O}/\sim_B &= \{[p_1]_f, [p_4]_f, [p_6]_f\} \text{ shown in Fig. 1a}, \\
X &= \{p \mid p \text{ equals a labeled coral pore in Fig. 1b}\}, \\
&= \{p_3, p_4, p_5, p_{10}\}.
\end{aligned}$$

Hence

$$\begin{aligned}
B_*X &= [p_4]_f = \{p_4, p_5\}, \\
B^*X &= [p_1]_f \cup [p_4]_f \cup [p_6]_f, \\
\mathrm{Bnd}_B X &= [p_1]_f \cup [p_6]_f.
\end{aligned}$$

Then observe that only $[p_4]_f$ is a proper subset of X. In effect, only the class $[p_4]_f$ has a complete affinity with X, since all of the objects $[p_4]_f$ have descriptions that match the descriptions of objects in X. By contrast, classes $[p_1]_f, [p_6]_f$ have only a partial affinity with X, since there are objects in each class that have descriptions that do match the description of any object in X. Observe that the boundary $\mathrm{Bnd}_B X$ is not empty. Hence, X is an example of a rough set.

Information Tables

For computational reasons, a syntactic representation of information systems is usually given in the form of tables. An information system is represented by the pair $(\mathcal{O}, \mathcal{F})$, where $\mathcal{O}$ denotes a set of objects (a universe) and $\mathcal{F}$ denotes a set of functions representing features of objects in $\mathcal{O}$. Let $B \subseteq \mathcal{F}$. Discovering objects in the composition of a class $[x]_B, x \in \mathcal{O}$ in the partition $\mathcal{O}/\sim_B$ in the system $(\mathcal{O}, \mathcal{F})$ is accomplished by gathering together inside the class all of those objects that have matching function values. Identifying the classes in $X/\sim_B$ is greatly aided by

a table representation of $(\mathcal{O}, \mathcal{F})$. This is illustrated with the following example.

Example 2 (Sample information table) For this example, assume again that $\mathcal{O}$ is a set of words. For simplicity, we limit $\mathcal{O}$ to words from poems by John Keats (1795–1821). Consider, for example, the words in the following lines in Part 3 of Keats' ode *To Autumn* [24].

TO AUTUMN

3.

Where are the songs of spring? Aye, where are they?

Think not of them, thou has thy music too,—

...

John Keats

19 September 1819

By way of approximation of a set of conceptual objects, consider $X \subseteq \mathcal{O}$ defined as

$$\begin{aligned}\mathcal{O} &= \{x \mid \text{x is a word in Keats' poetry}\},\\ &= \{Where, are, the, songs, of, spring, \dots\},\\ \mathcal{A} &= \{f_1, f_2\},\end{aligned}$$

where

$$f_1(x) = \begin{cases} 1, & x \text{ begins with a vowel,}\\ 0, & otherwise.\end{cases}$$

$$f_2(x) = \begin{cases} 1, & x \text{ contains two or more vowels,}\\ 0, & otherwise.\end{cases}$$

The objects in this example are conceptualized as poetic words and the functions appropriately represent attributes of poetic words. On might say that a word gains in poesy by beginning with a vowel sound (taking away the leading vowel sound tends to undermine the poetic character of a word). Similarly, a repetition of vowel sounds in a word is characteristic of poetic words, especially in the poetry of John Keats. For simplicity, consider only the vowels *a, e, i, o, u*, the vowel sound *ha* in the word *has* and omit consideration of other vowel sounds. In addition, $\mathcal{O}$ is restricted to words in Keats' ode *To Autumn*, partially represented in this Example. The function values for each of the words in the fragment from Keat's ode are summarized in Table 1. Let $B = \mathcal{A}$. This leads to the following partition of $\mathcal{O}$:

$$\begin{aligned}\left[where\right]_B &= \{where, thou, music, too\}\\ [are]_B &= \{are, Aye\}\\ \left[the\right]_B &= \{the, songs, spring, they, think, not,\\ &\qquad them, has, thy\}\\ \left[of\right]_B &= \{of\}\end{aligned}$$

Rough Sets: Foundations and Perspectives, Table 1
Information System for Keats' Words

$\mathcal{O}$	f_1	f_2	d
where	0	1	1
are	1	1	1
the	0	0	0
songs	0	0	1
of	1	0	1
spring	0	0	1
Aye	1	1	1
they	0	0	1
think	0	0	1
not	0	0	1
them	0	0	1
thou	0	1	1
has	0	0	0
thy	0	0	1
music	0	1	1
too	0	1	1

Now select, for example,

$$X = \{not, are, Aye, has, of, too\}\ .$$

This choice of X leads to

$$\begin{aligned}B_*X &= [are]_B \cup [of]_B = \{are, Aye, of\}\\ B^*X &= [are]_B \cup [of]_B \cup [the]_B\ .\end{aligned}$$

In effect, the words in $[are]_B \cup [of]_B$ have an affinity to the words in the set X. In particular, notice that each of the words in the class $\left[where\right]_B$

$$\left[where\right]_B = \{where, thou, music, too\}\ ,$$

have exactly the same f_1, f_2 function values. Similarly, for the composition of the remaining three classes in the partition $X/\sim_B$. Again, for example, notice that the words in class $[the]_A$ have the same function values, namely, $f_1(x) = f(x) = 0$, where x is a word in $[the]_B$. It is a fairly straightforward task to identify the classes extractable from small tables. For large tables, it is necessary to mechanize the class extraction process.

Decision Systems and Decision Rules

Of particular interest is the extension of information systems made possible by including a function d representing what is known as a decision attribute in rough set theory. A decision system is represented by the triple $(\mathcal{O}, \mathcal{F}, d)$, where $\mathcal{O}$ denotes a set of objects (a universe),

$\mathcal{F}$ denotes a set of functions representing features of objects in $\mathcal{O}$, and $d : \mathcal{O} \longrightarrow V_d$, where V_d is a set of values representing decisions. Presenting decision procedures in a tabular form goes back at least to ancient Babylon. Tabular forms for computer programming dates back to the late 1950s. Next, tabular forms became popular in databases (see http://www.catalyst.com/products/logicgem/overview.html). It is clear from the common meaning of the verb *decide* that a decision represents a resolution or determination about something and presumes something knowable by the mind but not necessarily observable by the senses. In more general setting, one can consider a set of decisions D instead of a single decision d.

Example 3 (Sample decision system) For example, consider an extension of the information system $(\mathcal{O}, \mathcal{A})$ with a decision d for Keats' words in Example 3. Recall that $\mathcal{O}$ denotes a set of words and $\mathcal{F}$ denotes a set of features representing attributes of words in $\mathcal{O}$. By way of illustration, let d be defined as follows.

$$d(x) = \begin{cases} 1, & x \in \mathcal{O} \mid x \text{ is a part of an alliteration}, \\ 0, & \textit{otherwise}, \end{cases}$$

where *alliteration* means the occurrence of the same letter or sound at the beginning of adjacent or closely connected words. *Alliteration* is also an example of a concept (i. e., an idea for a class of objects), that provides a basis for decision-making about various words. The common sense view of a concept is the one favored in this introduction to rough sets. Considering *alliteration* to be a concept is consonant with the general meaning of a theoretical concept [1]. That is, a *theoretical concept* is a concept expressed by a theoretical term associated with an axiomatic system underlying logic, mathematics, terms drawn from natural language or theories or that constitute the vocabulary of a particular theory such as poetics or a branch of science such as physics. at odds with the notion of concept in logic, mathematics, scientific theory. Table 1 is a sample decision table.

For example, consider the first line of Part 3 of Keats' *To Autumn*:

> Where are the songs of spring? Aye, where are they?

The repetition of the s sound in *songs* and *Spring* is alliterative. Then $d(songs) = d(Spring) = 1$. The word the in the first question in *To Autumn* is not part of an alliteration in that question. Hence, $d(the) = 0$.

Let $V = \bigcup\{V_a \mid a \in C\} \cup \{V_d\}$. Atomic formulae over $B \subseteq C \cup D$ and V are expressions $a = v$ called *descriptors* (*selectors*) over B and V, where $a \in B$ and $v \in V_a$. The set $\mathcal{F}(B, V)$ of formulae over B and V is the least set containing all atomic formulae over B and V and closed with respect to the propositional connectives $\wedge$ (conjunction), $\vee$ (disjunction) and $\neg$ (negation).

By $\|\varphi\|_{\mathcal{A}}$, we denote the meaning of $\varphi \in \mathcal{F}(B, V)$ in the decision table $\mathcal{A}$ which is the set of all objects in $\mathcal{O}$ with the property φ. These sets are defined by $\|a = v\|_{\mathcal{A}} = \{x \in \mathcal{O} \mid a(x) = v\}$, $\|\varphi \wedge \varphi'\|_{\mathcal{A}} = \|\varphi\|_{\mathcal{A}} \cap \|\varphi'\|_{\mathcal{A}}$; $\|\varphi \vee \varphi'\|_{\mathcal{A}} = \|\varphi\|_{\mathcal{A}} \cup \|\varphi'\|_{\mathcal{A}}$; $\|\neg\varphi\|_{\mathcal{A}} = \mathcal{O} - \|\varphi\|_{\mathcal{A}}$ The formulae from $\mathcal{F}(C, V)$, $\mathcal{F}(d, V)$ are called *condition formulae of* $\mathcal{A}$ and *decision formulae of* $\mathcal{A}$, respectively.

Any object $x \in \mathcal{O}$ belongs to the *decision class* $\| \bigwedge_{d \in D} d = d(x)\|_{\mathcal{A}}$ of $\mathcal{A}$. All decision classes of $\mathcal{A}$ create a partition $\mathcal{O}/\mathcal{D}$ of the universe $\mathcal{O}$.

A *decision rule* for $\mathcal{A}$ is any expression of the form $\varphi \Rightarrow \psi$, where $\varphi \in \mathcal{F}(C, V), \psi \in \mathcal{F}(D, V)$, and $\|\varphi\|_{\mathcal{A}} \neq \emptyset$. Formulae φ and ψ are referred to as the *predecessor* and the *successor* of decision rule $\varphi \Rightarrow \psi$. Decision rules are often called "IF ... THEN ..." rules. Such rules are used in machine learning (see, e. g.,[13]).

Decision rule $\varphi \Rightarrow \psi$ is *true* in $\mathcal{A}$ if and only if $\|\varphi\|_{\mathcal{A}} \subseteq \|\psi\|_{\mathcal{A}}$. Otherwise, one can measure its *truth degree* by introducing some inclusion measure of $\|\varphi\|_{\mathcal{A}}$ in $\|\psi\|_{\mathcal{A}}$.

Given two unary predicate formulae $\alpha(x), \beta(x)$ where x runs over a finite set $\mathcal{O}$, Łukasiewicz [28] proposes to assign to $\alpha(x)$ the value $\text{card}(\|\alpha(x)\|)/\text{card}(\mathcal{O})$, where $\|\alpha(x)\| = \{x \in \mathcal{O} : x \text{ satisfies } \alpha\}$. The fractional value assigned to the implication $\alpha(x) \Rightarrow \beta(x)$ is then $\text{card}(\|\alpha(x) \wedge \beta(x)\|)/\text{card}(\|\alpha(x)\|)$ under the assumption that $\|\alpha(x)\| \neq \emptyset$. Proposed by Łukasiewicz, that fractional part was much later adapted by machine learning and data mining literature.

Each object x of a decision system determines a *decision rule*

$$\bigwedge_{a \in C} a = a(x) \Rightarrow \bigwedge_{d \in D} d = d(x)\,. \tag{1}$$

For any decision table $\mathcal{A} = (\mathcal{O}, C, d)$ one can consider a *generalized decision* function $\partial_A : \mathcal{O} \longrightarrow V_d$ defined by

$$\partial_A(x) = \left\{i : \exists x' \in \mathcal{O}\left[\,(x', x) \in I(A) \text{ and } d(x') = i\right]\right\}. \tag{2}$$

$\mathcal{A}$ is called *consistent (deterministic)*, if $\text{card}(\partial_A(x)) = 1$, for any $x \in \mathcal{O}$. Otherwise $\mathcal{A}$ is said to be *inconsistent (nondeterministic)*. Hence, a decision table is inconsistent if it consists of some objects with different decisions but indiscernible with respect to condition attributes. Any set consisting of all objects with the same generalized decision value is called a *generalized decision class*. Now, one can

consider certain (possible) rules (see, e. g. [17]) for decision classes defined by the lower (upper) approximations of such generalized decision classes of $\mathcal{A}$. This approach can be extend, using the relationships of rough sets with the Dempster–Shafer theory (see, e. g., [51]), by considering rules relative to decision classes defined by the lower approximations of unions of decision classes of $\mathcal{A}$.

Numerous methods have been developed for different decision rule generation that the reader can find in the literature on rough sets. Usually, one is searching for decision rules (semi) optimal with respect to some optimization criteria describing quality of decision rules in concept approximations.

In the case of searching for concept approximation in an extension of a given universe of objects (sample), the following steps are typical. When a set of rules has been induced from a decision table containing a set of training examples, they can be inspected to see if they reveal any novel relationships between attributes that are worth pursuing for further research. Furthermore, the rules can be applied to a set of unseen cases in order to estimate their classificatory power.

Dimensionality Reduction

We often face the question whether one or more attributes can be removed and still preserve the basic properties of an information system. Let us express this idea more precisely.

Let $C, D \subseteq A$, be sets of condition and decision attributes respectively. We will say that $C' \subseteq C$ is a *D-reduct* (reduct with *respect* to D) of C, if C' is a minimal subset of C such that

$$\gamma(C, D) = \gamma(C', D) \,. \tag{3}$$

The intersection of all D-reducts is called a *D-core* (core with *respect* to D). Because the core is the intersection of all reducts, it is included in every reduct. Thus, in a sense, the core is the most important subset of attributes, since none of its elements can be removed without affecting the classification power of attributes. Certainly, the geometry of reducts can be more compound. For example, the core can be empty but there can exist a partition of reducts into a few sets with non empty intersection.

Many other kinds of reducts and their approximations are discussed in the literature (see, e. g., [30,60]). For example, if one change the condition (3) to $\partial_A(x) = \partial_B(x)$, then the defined reducts are preserving the generalized decision. Other kinds of reducts are preserving, e. g.,: (i) the distance between attribute value vectors for any two objects, if this distance is greater than a given threshold, (ii) the distance between entropy distributions between any two objects, if this distance exceeds a given threshold [60], or (iii) the so called reducts relative to object used for generation of decision rules [3]. There are some relationships between different kinds of reducts. If B is a reduct preserving the generalized decision, than in B is included a reduct preserving the positive region. For mentioned above reducts based on distances and thresholds one can find analogous dependency between reducts relative to different thresholds. By choosing different kinds of reducts we select different degrees to which information encoded in data is preserved. Reducts are used for building data models. Choosing a particular reduct or a set of reducts has impact on the model size as well as on its quality in describing a given data set. The model size together with the model quality are two basic components tuned in selecting relevant data models. This is known as the minimal length principle (see, e. g., [49]). Selection of relevant kinds of reducts is an important step in building data models. It turns out that the different kinds of reducts can be efficiently computed using heuristics based, e. g., on the Boolean reasoning approach [30,53].

Summary

We are now observing a growing research interest in the foundations of rough sets that include various logical, algebraic, philosophical complexity issues of rough sets. Some relationships have already been established between rough sets and other approaches as well as a wide range of hybrid systems have been developed.

As a result, rough sets are linked with decision system modeling and analysis of complex systems, fuzzy sets, neural networks, evolutionary computing, data mining and knowledge discovery, pattern recognition, machine learning, data mining, and approximate reasoning, multicriteria decision making. In particular, rough sets are used in probabilistic reasoning, granular computing (including information granule calculi based on rough mereology), intelligent control, intelligent agent modeling, identification of autonomous systems, and process specification.

A wide range of applications of methods based on rough set theory alone or in combination with other approaches have been discovered in the following areas: acoustics, bioinformatics, business and finance, chemistry, computer engineering (e. g., data compression, digital image processing, digital signal processing, parallel and distributed computer systems, sensor fusion, fractal engineering), decision analysis and systems, economics, electrical engineering (e. g., control, signal analysis, power systems), environmental studies, digital image process-

ing, informatics, medicine, molecular biology, musicology, neurology, robotics, social science, software engineering, spatial visualization, Web engineering, and Web mining.

For further readings on rough set theory and applications the reader is referred to [40,45] and to books, special issues of journals, issues of Transactions on Rough Sets, and proceedings cited in the bibliography of this article.

Future Directions

One of the main future directions for research based on rough sets in combination with other approaches such as granular computing and other computational intelligence approaches for developing intelligent systems can be found in the framework of Wisdom Technology [21,22], for complex vague concept approximation and approximate reasoning about such concepts by agents or teams of agents searching for solutions of problems in real-life dynamically changing (distributed) environments in which these agents are operating. Such systems consist of autonomous agents operating in highly unpredictable environments and interacting with each other.

In addition, a number of new rough set-based research areas have emerged during the past several years such as various forms of learning, including hierarchical learning or analogy based reasoning [4,6,31,43,55,66], concept approximation [5], multicriteria decision making [15], entropy in information systems [7,61], rough neural computing [32], rough granular computing [52], and a near set approach to image processing [19,27].

Acknowledgments

The research has been supported by the grant from Ministry of Scientific Research and Information Technology of the Republic of Poland and by grant 185986 from the Natural Sciences and Engineering Research Council of Canada (NSERC).

Bibliography

Primary Literature

1. Audi R (1999) The Cambridge Dictionary of Philosophy, 2nd edn. Cambridge University Press, UK
2. Banerjee M, Chakraborty MK (2004) Algebras from rough sets. In: Pal SK, Polkowski L, Skowron A (eds) Rough-Neural Computing. Techniques for Computing with Words. Springer, Berlin, pp 157–184
3. Bazan J (1998) A Comparison of Dynamic and Non-dynamic Rough Set Methods for Extracting Laws from Decision Tables. In: Polkowski L, Skowron A (eds) Rough Sets in Knowledge Discovery 1: Methodology and Applications. Studies in Fuzziness and Soft Computing, vol 18. Physica, Heidelberg, pp 321–365
4. Bazan J (2008) Rough sets and granular computing in behavioral pattern identification and planning In: Pedrycz W, Skowron A, Kreinovich V (eds) Handbook of Granular Computing. Wiley, New York (in press), pp 777–799
5. Bazan J, Skowron A, Swiniarski R (2006) Rough sets and vague concept approximation: From sample approximation to adaptive learning. Transactions on Rough Sets V. Lecture Notes in Computer Science, vol 4100. Springer, Heidelberg, pp 39–62
6. Bazan J, Kruczek P, Bazan-Socha S, Skowron A, Pietrzyk JJ (2006) Automatic planning of treatment of infants with respiratory failure through rough set modeling. In: Greco S, Hato Y, Hirano S, Inuiguchi M, Miyamoto S, Nguyen HS, Słowiński R (eds) Proceedings of the 5th International Conference on Rough Sets and Current Trends in Computing, RSCTC 2006, Kobe, Japan, 6–8 November. Lecture Notes in Artificial Intelligence, vol 4259. Springer, Heidelberg, pp 418–427
7. Bianucci D, Cattaneo G, Ciucci D (2007) Entropies and coentropies of coverings with application to incomplete information systems. Fundam Inform 75:77–105
8. Cattaneo G, Ciucci D (2004) Algebraic structures for rough sets. Transactions on Rough Sets II. LNCS (Lecture Notes in Artificial Intelligence), vol 3135. Springer, Heidelberg, pp 63–101
9. Chakraborty MK, Banerjee M (2007) Rough dialogue and implication lattices. Fundam Inform 75:123–139
10. Devlin J (1993) The Joy of Sets. Fundamentals of Contemporary Set Theory. Springer, Berlin
11. Düntsch I (1997) A logic for rough sets. Theor Comput Sci 179:427–436
12. Frege G (1879) Begriffsschrift: eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. Nebert, Halle
13. Friedman JH, Hastie T, Tibshirani R (2001) The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, Heidelberg
14. Gomolińska A (2005) Rough validity, confidence, and coverage of rules in approximation spaces. Transactions on Rough Sets III. LNCS (Lecture Notes in Artificial Intelligence), vol 3400. Springer, Heidelberg, pp 57–81
15. Greco S, Matarazzo B, Słowiński R (2001) Rough sets theory for multicriteria decision analysis. Eur J Oper Res 129:1–47
16. Greco S, Słowiński R, Stefanowski J, Zurawski M (2005) Incremental vs. non-incremental rule induction for multicriteria classification. Transactions on Rough Sets II. LNCS (Lecture Notes in Artificial Intelligence), vol 3135. Springer, Heidelberg, pp 33–53
17. Grzymała-Busse JW (1992) LERS – A system for learning from examples based on rough sets In: Słowiński R (ed) Intelligent Decision Support – Handbook of Applications and Advances of the Rough Sets Theory, System Theory, Knowledge Engineering and Problem Solving, vol 11. Kluwer, Dordrecht, pp 3–18
18. Grzymała-Busse JW, Grzymała-Busse WJ (2007) An experimental comparison of three rough set approaches to missing attribute values. Transactions on Rough Sets VI. LNCS (Lecture Notes in Artificial Intelligence), vol 4374. Springer, Heidelberg, pp 31–50
19. Henry C, Peters JF (2007) Image Pattern Recognition Using Approximation Spaces and Near Sets. In: Proceedings of Eleventh International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing. Joint Rough Set Symposium.

Lecture Notes in Artificial Intelligence, vol 4482. Springer, Berlin, pp 475–482
20. Int. Rough Set Society website: www.roughsets.org
21. Jankowski A, Skowron A (2007) Logic for Artificial Intelligence: The Rasiowa – Pawlak School Perspective. In: Ehrenfeucht A, Marek V, Srebrny M (eds) Andrzej Mostowski and Foundational Studies. IOS Press, Amsterdam, pp 106–143
22. Jankowski A, Skowron A (2008) Wisdom Granular Computing. In: Pedrycz W, Skowron A, Kreinovich V (eds) Handbook of Granular Computing. Wiley, New York, pp 329–345
23. Järvinen J (2007) Lattice theory for rough sets. Transactions on Rough Sets VI. LNCS (Lecture Notes in Artificial Intelligence), vol 4374. Springer, Heidelberg, pp 400–496
24. Keats J (1820) Lamia, Isabella, The Eve of St. Agness, and Other Poems. Taylor and Hessey, London
25. Konrad E, Orłowska E, Pawlak Z (1981) Knowledge Representation Systems. Definability of Informations. Prace IPI Pan, ICS PAS Report, vol 433 April. Institute of Computer Science, Polish Academy of Sciences, Warsaw
26. Leśniewski S (1929) Grungzüge eines neuen Systems der Grundlagen der Mathematik. Fundam Math pp 14:1–81
27. Lockery D, Peters JF (2007) Robotic Target Tracking with Approximation Space-based Feedback During Reinforcement Learning, Springer best paper award. In: Proceedings of Eleventh International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing. Joint Rough Set Symposium. Lecture Notes in Artificial Intelligence, vol 4482. Springer, Berlin, pp 483–490
28. Łukasiewicz J (1913) Die logischen Grundlagen der Wahrscheinlichkeitsrechnung. In: Borkowski L (ed) Jan Łukasiewicz – Selected Works. North Holland, Amsterdam; (1970) Polish Scientific Publishers, Warsaw, pp 16–63
29. Murray JA, Bradley H, Craigie W, Onions C (1933) The Oxford English Dictionary. Oxford University Press, London
30. Nguyen HS (2006) Approximate boolean reasoning: Foundations and applications in data mining. Transactions on Rough Sets V. LNCS (Lecture Notes in Artificial Intelligence), vol 4100. Springer, Heidelberg, pp 344–523
31. Nguyen SH, Bazan J, Skowron A, Nguyen HS (2004) Layered learning for concept synthesis, Transactions on Rough Sets I. LNCS (Lecture Notes in Artificial Intelligence), vol 3100. Springer, Heidelberg, pp 187–208
32. Pal SK, Polkowski L, Skowron A (2004) Rough-Neural Computing. Techniques for Computing with Words. Springer, Berlin
33. Orłowska E (1985) Semantics of Vague Concepts. In: Dorn G, Weingartner P (eds) Foundations of Logic and Linguistics. Problems and Solutions. Plenum, London, pp 465–482
34. Orłowska E (1985) A logic of indiscernibility relations. Lecture Notes in Computer Science, vol 208. Springer, Berlin, pp 177–186
35. Pal SK, Polkowski L, Skowron A (2004) Rough-Neural Computing: Techniques for Computing with Words. Springer, Berlin
36. Pawlak Z (1973) Mathematical foundations of information retrieval. Proceedings of Symposium of Mathematical Foundations of Computer Science, 3–8 September, High Tartras, pp. 135–136. See also: Mathematical Foundations of Information Retrieval. Computation Center, Polish Academy of Sciences, Research Report CC PAS Report 101, Warsaw
37. Pawlak Z (1981) Classification of objects by means of attributes. Institute for Computer Science, Polish Academy of Sciences. Research Report CC PAS Report 429, Warsaw
38. Pawlak Z (1991) Rough Sets. Theoretical Aspects of Reasoning about Data. Kluwer, Dordrecht
39. Pawlak Z (1995) On rough derivatives, rough integrals, and rough differential equations, ICS Research Report 41/95. Institute of Computer Science, Warsaw
40. Pawlak Z, Skowron A (2007) Rudiments of rough sets. Information Sciences. Int J 177(1): 3–27; Rough sets: Some extensions. Information Sciences. Int J 177(1): 28–40; Rough sets and Boolean reasoning. Information Sciences. Int J 177(1):41–73
41. Peters JF (2008) Classification of perceptual objects by means of features. Int J Inf Technol Intell Comput 3(2):1–35
42. Peters JF, Henry C (2007) Approximation spaces in off-policy Monte Carlo learning. Engineering Applications of Artificial Intelligence. Int J Intell Real-Time Autom 20(5):667–675
43. Peters JF, Henry C, Gunderson DS (2006) Biologically-inspired approximate adaptive learning control strategies: A rough set approach. Inter J Hybrid Intell Syst 3:1–14
44. Peters JF, Skowron A, Stepaniuk J (2007) Nearness of Objects: Extension of Approximation Space Model. Fundam Inform 79(3–4):497–512
45. Polkowski L (2002) Rough Sets: Mathematical Foundations. Advances in Soft Computing. Physica, Heidelberg
46. Polkowski L, Skowron A (1996) Rough mereology: A new paradigm for approximate reasoning. J Approx Reas 15(4):333–365
47. Polkowski L, Skowron A (1998) Rough Sets in Knowledge Discovery 1: Methodology and Applications. Studies in Fuzziness and Soft Computing, vol 18. Physica, Heidelberg
48. Polkowski L, Skowron A (1998) Rough Sets in Knowledge Discovery 2: Applications, Case Studies and Software Systems. Studies in Fuzziness and Soft Computing, vol 19. Physica, Heidelberg
49. Rissanen J (1985) Minimum-description-length Principle. In: Kotz S, Johnson N (eds) Encyclopedia of Statistical Sciences. Wiley, New York, pp 523–527
50. Rough Set Database: http://rsds.univ.rzeszow.pl/
51. Skowron A, Grzymała-Busse JW (1994) From Rough Set Theory to Evidence Theory. In: Yager R, Fedrizzi M, Kacprzyk J (eds) Advances in the Dempster-Shafer Theory of Evidence. Wiley, New York, pp 193–236
52. Skowron A, Peters JF (2008) Rough-granular computing. In: Pedrycz W, Skowron A, Kreinovich V (eds) Handbook of Granular Computing. Wiley, pp 285–328
53. Skowron A, Rauszer C (1992) The Discernibility Matrices and Functions in Information Systems. In: Słowiński R (ed) Intelligent Decision Support – Handbook of Applications and Advances of the Rough Sets Theory, System Theory, Knowledge Engineering and Problem Solving, vol 11. Kluwer, Dordrecht, pp 331–362
54. Skowron A, Stepaniuk J (1996) Tolerance approximation spaces. Fundam Inform 27:245–253
55. Skowron A, Swiniarski R, Synak P (2005) Approximation spaces and information granulation. Transactions on Rough Sets III. LNCS (Lecture Notes in Artificial Intelligence), vol 3400. Springer, Heidelberg, pp 175–189
56. Skowron A, Stepaniuk J, Peters JF, Swiniarski R (2006) Calculi of approximation spaces. Fundam Inform 72(1–3):363–378
57. Stepaniuk J (1998) Approximation spaces, reducts and representatives. In: Polkowski L, Skowron A (eds) Rough Sets in Knowledge Discovery 2: Applications, Case Studies and Soft-

ware Systems, Studies in Fuzziness and Soft Computing, vol 19. Physica, Heidelberg, pp 109–126
58. Stepaniuk J (2000) Knowledge discovery by application of rough set models. In: Polkowski L, Tsumoto S, Lin TY (eds) Rough Set Methods and Applications. New Developments in Knowledge Discovery in Information Systems. Physica, Heidelberg, pp 137–233
59. Stepaniuk J (2007) Approximation Spaces in Multi Relational Knowledge Discovery, Transactions on Rough Sets VI: LNCS (Lecture Notes in Artificial Intelligence), vol 4374. Springer, Heidelberg, pp 351–365
60. Ślęzak D (1996) Approximate reducts in decision tables. Sixth International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU'1996, vol III, Granada, Spain, pp 1159–1164
61. Ślęzak D (2002) Approximate entropy reducts. Fundam Inform 53:365–387
62. Transactions on Rough Sets website: http://www.springer.com/east/home/computer/lncs?SGWID=5-164-6-99627-0, special issues of other journals, proceedings of international conferences, tutorials or surveys (see, references in this paper and, e. g., Pawlak Z, Skowron A (2007) Rudiments of rough sets. Information Sciences. Int J 177(1):3–27; Rough sets: Some extensions. Information Sciences. Int J 177(1):28–40; Rough sets and Boolean reasoning. Information Sciences. Int J 177(1):41–73, and numerous web pages (www.roughsets.org, and logic.mimuw.edu.pl))
63. University of Regina Electronic Bulletin of the Rough Set Community: http://www.cs.uregina.ca/~roughset
64. Vinogradov IM (ed) (1995) Encyclopedia of Mathematics. Kluwer, Dordrecht
65. Warsaw University Logic Group website: logic.mimuw.edu.pl
66. Wojna A (2007) Analogy-based reasoning in classifier construction. Transactions on Rough Sets IV. LNCS (Lecture Notes in Artificial Intelligence), vol 3700. Springer, Heidelberg, pp 277–374
67. Zadeh LA (2001) A new direction in AI: Toward a computational theory of perceptions. AI Mag 22(1):73–84
68. Zadeh LA (2005) Toward a generalized theory of uncertainty (GTU) – An outline. Inform Sci 171:1–40
69. Ziarko W (1993) Variable precision rough set model. J Comput Syst Sci 46:39–59

Further Readings – Books

Demri S, Orłowska E (2002) Incomplete Information: Structure, Inference, Complexity. Monographs in Theoretical Computer Science. Springer, Berlin

Dunin-Kęplicz B, Jankowski A, Skowron A, Szczuka M (eds) (2005) Monitoring, Security, and Rescue Tasks in Multiagent Systems, MSRAS'2004. Advances in Soft Computing. Springer, Heidelberg

Düntsch I, Gediga G (2000) Rough set data analysis: A road to noninvasive knowledge discovery. Methodos, Bangor

Grzymała-Busse JW (1990) Managing Uncertainty in Expert Systems. Kluwer, Norwell

Inuiguchi M, Hirano S, Tsumoto S (eds) (2003) Rough Set Theory and Granular Computing, Studies in Fuzziness and Soft Computing, vol 125. Springer, Heidelberg

Kostek B (2005) Perception-Based Data Processing in Acoustics. Applications to Music Information Retrieval and Psychophysiology of Hearing, Studies in Computational Intelligence, vol 3. Springer, Heidelberg

Lin TY, Cercone N (eds) (1997) Rough Sets and Data Mining – Analysis of Imperfect Data. Kluwer, Boston

Lin TY, Yao YY, Zadeh LA (eds) (2001) Rough Sets, Granular Computing and Data Mining. Studies in Fuzziness and Soft Computing. Physica, Heidelberg

Mitra S, Acharya T (2003) Data mining. Multimedia, Soft Computing, and Bioinformatics. Wiley, New York

Munakata T (ed) (1998) Fundamentals of the New Artificial Intelligence: Beyond Traditional Paradigms. Graduate Texts in Computer Science, vol 10. Springer, New York

Orłowska E (ed) (1997) Incomplete Information: Rough Set Analysis. Studies in Fuzziness and Soft Computing, vol 13. Physica, Heidelberg

Pal SK, Skowron A (eds) (1999) Rough Fuzzy Hybridization: A New Trend in Decision-Making. Springer, Singapore

Polkowski L, Lin TY, Tsumoto S (eds) (2000) Rough Set Methods and Applications: New Developments in Knowledge Discovery in Information Systems, Studies in Fuzziness and Soft Computing, vol 56. Physica, Heidelberg

Słowiński R (ed) (1992) Intelligent Decision Support – Handbook of Applications and Advances of the Rough Sets Theory, System Theory, Knowledge Engineering and Problem Solving, vol 11. Kluwer, Dordrecht

Zhong N, Liu J (eds) (2004) Intelligent Technologies for Information Analysis. Springer, Heidelberg

Further Readings – Transactions on Rough Sets

Peters JF, Skowron A et al (eds) (2004–2008) Transactions on Rough Sets I-VIII. (Lecture Notes in Computer Science), vols 3100, 3135, 3400, 3700, 4100, 4374, 4400, 5084. Springer, Heidelberg

Further Readings – Special Issues of Journals

Cercone N, Skowron A, Zhong N (eds) (2001) Comput Intell Int J 17(3)

Lin TY (ed) (1996) J Intell Autom Soft Comput 2(2)

Pal SK, Pedrycz W, Skowron A, Swiniarski R (eds) (2001) Special volume: Rough-neuro computing. Neurocomputing 36

Peters J, Skowron A (eds) (2001) Special issue: on a rough set approach to reasoning about data. Int J Intell Syst 16(1)

Skowron A, Pal SK (eds) (2003) Special volume: Rough sets, pattern recognition and data mining. Pattern Recognit Lett 24(6)

Słowiński R, Stefanowski J (eds) (1993) Special issue: Proceedings of the First International Workshop on Rough Sets: State of the Art and Perspectives, Kiekrz, Poznań, Poland, 2–4 September 1992. Found Comput Decis Sci 18(3–4)

Ziarko W (ed) (1995) Comput Intell Int J 11(2)

Ziarko W (ed) (1996) Fundam Inform 27(2–3)

Further Readings – Proceedings of International Conferences

Alpigini JJ, Peters JF, Skowron A, Zhong N (eds) (2002) Third International Conference on Rough Sets and Current Trends in Computing, RSCTC'2002, Malvern, PA, 14–16 October. Lecture Notes in Artificial Intelligence, vol 2475. Springer, Heidelberg

An A, Stefanowski J, Ramanna S, Butz CJ, Pedrycz W, Wang G (eds) (2007) Proceedings of the Eleventh International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, RSFDGrC 2007, Toronto, Canada, 14–16 May. Lecture Notes in Artificial Intelligence, vol 4482. Springer, Heidelberg

Greco S, Hata Y, Hirano S, Inuiguchi M, Miyamoto S, Nguyen HS, Słowiński R (eds) (2006) Proceedings of the Fifth International Conference on Rough Sets and Current Trends in Computing, RSCTC 2006, Kobe, Japan, 6–8 November. Lecture Notes in Artificial Intelligence, vol 4259. Springer, Heidelberg

Hirano S, Inuiguchi M, Tsumoto S (eds) (2001) Proceedings of International Workshop on Rough Set Theory and Granular Computing, RSTGC'2001, Matsue, Japan, 20–22 May. Bull Int Rough Set Soc 5(1–2)

Kryszkiewicz M, Peters JF, Rybiński H, Skowron A (eds) (2007) Proceedings of the International Conference on Rough Sets and Intelligent Systems Paradigms, RSEISP 2007, Warsaw, Poland, 28–30 June. Lecture Notes in Artificial Intelligence, vol 4585. Springer, Heidelberg

Lin TY, Wildberger AM (eds) (1995) Soft Computing: Rough Sets, Fuzzy Logic, Neural Networks, Uncertainty Management, Knowledge Discovery. Simulation Councils, USA

Polkowski L, Skowron A (eds) (1998) First International Conference on Rough Sets and Soft Computing, RSCTC'1998. Lecture Notes in Artificial Intelligence, vol 1424. Springer, Poland

Skowron A (ed) (1985) Proceedings of the 5th Symposium on Computation Theory. Zaborów, Poland, 1984. Lecture Notes in Computer Science, vol 208. Springer, Berlin

Skowron A, Szczuka M (eds) (2003) Proceedings of the Workshop on Rough Sets in Knowledge Discovery and Soft Computing at ETAPS, 12–13 April. Electronic Notes in Computer Science, vol 82(4). Elsevier, Netherlands, http://www.elsevier.nl/locate/entcs/volume82.html

Ślęzak D, Wang G, Szczuka M, Düntsch I, Yao Y (eds) (2005) Proceedings of the 10th International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, RSFDGrC'2005, Regina, Canada, 31 August–3 September, Part I. Lecture Notes in Artificial Intelligence, vol 3641. Springer, Heidelberg

Ślęzak D, Yao JT, Peters JF, Ziarko W, Hu X (eds) (2005) Proceedings of the 10th International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, RSFDGrC'2005, Regina, Canada, 31 August–3 September, Part II. Lecture Notes in Artificial Intelligence, vol 3642. Springer, Heidelberg

Terano T, Nishida T, Namatame A, Tsumoto S, Ohsawa Y, Washio T (eds) (2001) New Frontiers in Artificial Intelligence, Joint JSAI'2001. Workshop Post-Proceedings. Lecture Notes in Artificial Intelligence, vol 2253. Springer, Heidelberg

Tsumoto S, Kobayashi S, Yokomori T, Tanaka H, Nakamura A (eds) (1996) Proceedings of the The Fourth Internal Workshop on Rough Sets, Fuzzy Sets and Machine Discovery, University of Tokyo, Japan, 6–8 November. The University of Tokyo, Tokyo

Tsumoto S, Słowiński R, Komorowski J, Grzymała-Busse J (eds) (2004) Proceedings of the 4th International Conference on Rough Sets and Current Trends in Computing, RSCTC'2004, Uppsala, Sweden, 1–5 June. Lecture Notes in Artificial Intelligence, vol 3066. Springer, Heidelberg

Yao JT, Lingras P, Wu WZ, Szczuka M, Cercone NJ, Ślęzak D (eds) (2007) Proceedings of the Second International Conference on Rough Sets and Knowledge Technology, RSKT 2007, Toronto, Canada, 14–16 May. Lecture Notes in Artificial Intelligence, vol 4481. Springer, Heidelberg

Wang G, Liu Q, Yao Y, Skowron A (eds) (2003) Proceedings of the 9th International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, RSFDGrC'2003, Chongqing, China, 26–29 May. Lecture Notes in Artificial Intelligence, vol 2639. Springer, Heidelberg

Zhong N, Skowron A, Ohsuga S (eds) (1999) Proceedings of the 7th International Workshop on Rough Sets, Fuzzy Sets, Data Mining, and Granular-Soft Computing, RSFDGrC'99, Yamaguchi, 9–11 November. Lecture Notes in Artificial Intelligence, vol 1711. Springer, Heidelberg

Ziarko W (ed) (1994) Rough Sets, Fuzzy Sets and Knowledge Discovery: Proceedings of the Second International Workshop on Rough Sets and Knowledge Discovery, RSKD'93, Banff, Canada, 12–15 October 1993. Workshops in Computing. Springer & British Computer Society, London, Berlin

Ziarko W, Yao Y (2001) Proceedings of the 2nd International Conference on Rough Sets and Current Trends in Computing, RSCTC'2000, Banff, Canada, 16–19 October 2000. Lecture Notes in Artificial Intelligence, vol 2005. Springer, Heidelberg

Rule Induction, Missing Attribute Values and Discretization

Jerzy W. Grzymala-Busse[1,2]

[1] Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence, USA

[2] Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland

Article Outline

Glossary

Discretization Discretization is a process of converting numerical attributes into symbolic ones by splitting the numerical attribute domain into intervals. Usually discretization is conducted before the main process of rule induction, but in some rule induction algorithms, e. g., in MLEM2 (Modified LEM2), rules are induced concurrently with discretization.

LEM2 algorithm LEM2 (Learning from Examples Module, version 2) is the basic rule induction algorithm of the machine learning/data mining system LERS. LEM2, implemented for the first time in 1990, uses an idea of a local covering to induce a minimal set of minimal rules describing all data concepts.

LERS machine learning/data mining system LERS (Learning from Examples based on Rough Sets) is a rule induction system created at the University of Kansas. Its first implementation was done in Franz Lisp in 1988. This first version of LERS had only one algorithm called LEM1 (Learning form Examples Module, version 1) to induce all rules from input data.

Missing attribute values Missing attribute values frequently affect real-life data. Some attribute values are *lost* (e. g., erased), some are of the type *"do not care" conditions* (such attribute values were irrelevant for classification of the case). In most existing machine learning/data mining systems some method of handling missing attribute values is applied before the main process of rule induction. However, in MLEM2 rule induction and handling missing attribute values are conducted at the same time.

Rule induction Rule induction is understood here as an instance of supervised learning. Rule induction is one of the basic processes of acquiring knowledge (knowledge extraction) in the form of rule sets from raw data. This process is widely used in machine learning (data mining). A data set contains cases (examples) characterized by attribute values and classified as members of concepts by an expert. *Rules* are expressions of the following format:

if condition$_1$ **and** condition$_2$ **and** … **and** condition$_n$ **then** decision.

Definition of the Subject

Rule induction is a process of creating rule sets from raw data called *training data*. Such rules represent hidden and previously unknown knowledge contained in the training data. These rules may be used for successful classification of new cases that were not used for training. One of the possible applications of this methodology is rule-based expert systems. There are many documented examples of successful use of rule induction in medicine (e. g., for decision support for diagnosis), finances, military, etc.

In particular, one of the rule induction systems called LERS has proven its applicability having been used for years by NASA Johnson Space Center (Automation and Robotics Division), as a tool to develop expert systems of the type used in medical decision-making on board the International Space Station. LERS was also used to enhance facility compliance under Sections 311, 312, and 313 of Title III, the Emergency Planning and Community Right to Know. System LERS was used in other areas as well, for example, in the medical field to assess preterm labor risk for pregnant women. Currently used manual methods of assessing preterm birth have a positive predictive value of 17–38%. The data mining methods based on LERS reached positive predictive value of 59–92%. Other applications of LERS to the medical area include diagnosis of melanoma, prediction of behavior under mental retardation and analysis of animal models for prediction of self-injurious behavior. LERS has been used also in nursing, global warming, environmental protection, natural language, and data transmission. LERS can process large datasets and frequently outperforms not only other rule induction systems but also human experts.

Introduction

Data mining uses methods of machine learning to acquire knowledge from raw data. Rule induction is one of the most successful techniques of machine learning. In many application areas, such as medicine, it is essential not only to make an appropriate decision or diagnosis but also to be able to justify or explain the decision. Rules provide such explanations.

We will discuss rule induction using for that purpose one of the very successful algorithms called LEM2 (Learning from Examples Module, version 2), see for example [4]. This algorithm is used in the LERS (Learning from Examples based on Rough Sets) data mining system. LERS can process inconsistent data, i. e., data with conflicting cases, in which values for all attributes are the same yet the decision values are distinct. LERS deals with inconsistent data using a typical rough set approach: it computes lower and upper approximations for all concepts. Then it passes such approximations to LEM2.

Data are frequently incomplete, i. e., some attribute values are missing. Rule sets may be induced from incomplete data using a modified LEM2 algorithm called MLEM2. Again, MLEM2 uses a rough set approach to missing attribute values.

Additionally, many real-life data sets contain numerical attributes, i. e., attributes whose values are integers or real numbers. Numerical data need to be converted into symbolic data by a process called discretization. For a numerical attribute a few intervals are determined, usually these intervals are disjoint and together they cover the entire domain of the numerical attribute. The process of dis-

cretization may be performed before rule induction or, as it is done in MLEM2, during rule induction.

In the entire process of rule learning, handling missing attribute values, and discretization, the applied methodology is based on the same granules called *attribute-value blocks*. Similar granules were discussed in [13,14,15].

Discretization

The input examples are presented in a table, called a decision table. An example of such a table is presented as Table 1. Table 1 represents the input data set in the format of LERS [4]. Table 1 contains four variables: *Age*, *Skill*, *Experience* and *Productivity*. The first three variables are called *attributes*, the last one is called *decision*. Table 1 presents seven cases (or examples) labeled by integers $1, 2, \ldots, 7$.

Any subset of the set of all cases, defined by the same value of the decision is called a *concept*. Table 1 consists of two concepts, defined by (*Productivity*, *low*) and (*Productivity*, *high*). For example, the concept defined by (*Productivity*, *low*) is the set $\{1, 3, 4, 6\}$.

The attribute *Age* from Table 1 is numerical. The numerical attributes should be discretized before rule induction, otherwise rules induced from data with numerical attributes will be too specific. In general, during *discretization* an original range $[a, b]$ of a numerical attribute is partitioned into a set of n intervals

$$\{[a_0, a_1), [a_1, a_2), \ldots, [a_{n-1}, a_n]\},$$

where $a_0 = a$, $a_k < a_{k+1}$ for $k = 0, 1, \ldots, n-1$ and $a_n = b$. The data mining system LERS uses for discretization a number of discretization algorithms [2], including Equal Interval Width, Equal Frequency per Interval, Minimal Class Entropy, and two discretization methods based on Cluster Analysis [5,8]. As an example, let us use one of the simplest discretization methods, discretization based on the Equal Interval Width. Let us start with $n = 2$.

Rule Induction, Missing Attribute Values and Discretization, Table 1
An example of a decision table

Case	Attributes			Decision
	Age	Skill	Experience	Productivity
1	25	low	low	low
2	56	high	high	high
3	36	low	high	low
4	42	high	low	low
5	59	high	low	high
6	25	high	low	low
7	42	high	high	high

Rule Induction, Missing Attribute Values and Discretization, Table 2
An example of an inconsistently discretized decision table

Case	Attributes			Decision
	Age	Skill	Experience	Productivity
1	25..42	low	low	low
2	42..57	high	high	high
3	25..42	low	high	low
4	42..57	high	low	low
5	42..57	high	low	high
6	25..42	high	low	low
7	42..57	high	high	high

Rule Induction, Missing Attribute Values and Discretization, Table 3
An example of a discretized decision table

Case	Attributes			Decision
	Age	Skill	Experience	Productivity
1	25..36.3	low	low	low
2	47.7..59	high	high	high
3	25..36.3	low	high	low
4	36.3..47.7	high	low	low
5	47.7..59	high	low	high
6	25..36.3	high	low	low
7	36.3..47.7	high	high	high

Thus, the range $[25, 59]$ of *Age* is partitioned into two intervals of equal width: $[25, 42)$ and $[42, 59]$. We will denote the first interval by 25..42 and the second interval by 42..59. As a result, the decision table that represents the discretized data set (Table 2) is inconsistent, since cases 4 and 5 are conflicting. Therefore, the number of discretization intervals was too small. Let us try $n = 3$. The new intervals are $[25, 36.3)$, $[36.3..47.7)$, and $[47.7..59]$, denoted, respectively, by 25..36.3, 36.3..47.74, and 47.7..59. The new discretized table is presented in Table 3. This table is consistent and may be considered to be an input data set for rule induction. For more details on discretization methods see [2,6,8,10].

LEM2 Algorithm

LEM2 explores the search space of attribute-value pairs. Its input data set is a lower or upper approximation of a concept, so its input data set is always consistent. In general, LEM2 computes a local covering and then converts it into a rule set. We will quote a few definitions to describe the LEM2 algorithm [1,4,5].

The LEM2 algorithm is based on an idea of an attribute-value pair block. For an attribute-value pair

Rule Induction, Missing Attribute Values and Discretization, Table 4
An example of an inconsistent decision table

Case	Attributes			Decision
	Age	Skill	Experience	Productivity
1	25	low	low	low
2	56	high	high	high
3	36	low	high	low
4	42	high	low	low
5	59	high	low	high
6	25	high	low	low
7	42	high	high	high
8	25	high	low	high

$(a, v) = t$, a *block* of t, denoted by $[t]$, is a set of all cases from U such that for attribute a have value v. Let B be a nonempty lower or upper approximation of a concept represented by a decision-value pair (d, w). Set B *depends* on a set T of attribute-value pairs $t = (a, v)$ if and only if

$$\emptyset \neq [T] = \bigcap_{t \in T} [t] \subseteq B \,.$$

Set T is a *minimal complex* of B if and only if B depends on T and no proper subset T' of T exists such that B depends on T'. Let $\mathcal{T}$ be a nonempty collection of nonempty sets of attribute-value pairs. Then $\mathcal{T}$ is a *local covering* of B if and only if the following conditions are satisfied:

- Each member T of $\mathcal{T}$ is a minimal complex of B,
- $\bigcup_{t \in \mathcal{T}} [T] = B$, and
- $\mathcal{T}$ is minimal, i. e., $\mathcal{T}$ has the smallest possible number of members.

The procedure LEM2, based on rule induction from local coverings, is presented in Fig. 1

For a set X, $|X|$ denotes the cardinality of X.

For Table 3 the following rules are induced by LEM2:

(Age, 25..36.3) → (Productivity, low),
(Age, 36.3..47.7) & (Experience, low) → (Productivity, low),
(Age, 47.7..59) → (Productivity, high),
(Age, 36.3..47.7) & (Experience, high) → (Productivity, high).

Inconsistent Data

LERS handles inconsistencies using rough set theory, introduced by Z. Pawlak in 1982 [16,17]. In rough set theory lower and upper approximations are computed for all concepts involved in conflicts with other concepts.

Procedure LEM2
(**input**: a set B,
output: a single local covering $\mathcal{T}$ of set B);
begin
  $G := B$;
  $\mathcal{T} := \emptyset$;
  while $G \neq \emptyset$
    begin
    $T := \emptyset$;
    $T(G) := \{t | [t] \cap G \neq \emptyset\}$;
    while $T = \emptyset$ **or** $[T] \not\subseteq B$
      begin
        select a pair $t \in T(G)$ such that $|[t] \cap G|$ is maximum; if a tie occurs, select a pair $t \in T(G)$ with the smallest cardinality of $[t]$; if another tie occurs, select first pair;
        $T := T \cup \{t\}$;
        $G := [t] \cap G$;
        $T(G) := \{t | [t] \cap G \neq \emptyset\}$;
        $T(G) := T(G) - T$;
      end {while}
    for each $t \in T$ **do**
      if $[T - \{t\}] \subseteq B$
      then $T := T - \{t\}$;
    $\mathcal{T} := \mathcal{T} \cup \{T\}$;
    $G := B - \cup_{T \in \mathcal{T}} [T]$;
  end {while};
  for each $T \in \mathcal{T}$ **do**
    if $\bigcup_{S \in \mathcal{T} - \{T\}} [S] = B$ **then** $\mathcal{T} := \mathcal{T} - \{T\}$;
end {procedure}.

Rule Induction, Missing Attribute Values and Discretization, Figure 1
Procedure LEM2

Let U denote the set of all examples of the decision table and let P denote a nonempty subset of the set Q of all variables, i. e., attributes and decisions. Let P be a subset of A. An indiscernibility relation ρ on U is defined for all $x, y \in U$ by $x \rho y$ if and only if for both x and y the values for all variables from P are identical. Equivalence classes of ρ are called *elementary sets* of P. An equivalence class of ρ containing x is denoted $[x]_P$. Any finite union of elementary sets of P is called a *definable set* in P. Let X be a concept. In general, X is not a definable set in P. However, set X may be approximated by two definable sets in P, the first one is called a *lower approximation* of X in P,

denoted by $\underline{P}X$ and defined as follows

$$\{x \in U | [x]_P \subseteq X\}.$$

The second set is called an *upper approximation* of X in P, denoted by $\overline{P}X$ and defined as follows

$$\{x \in U | [x]_P \cap X \neq \emptyset\}.$$

The lower approximation of X in A is the greatest definable set in A, contained in X. The upper approximation of X in A is the least definable set in A containing X. A rough set of X is the family of all subsets of U having the same lower and the same upper approximations of X.

Table 4 is an example of inconsistent data. Cases 6 and 8 are conflicting. The lower approximation of the concept $\{1, 3, 4, 6\}$ is the set $\{1, 3, 4\}$ and the upper approximation of the same concept is the set $\{1, 3, 4, 6, 8\}$. The Table 4, discretized using the same Equal Interval Width method that was used before to Table 1, with three intervals, is presented as Table 5. The corresponding rule sets are: the certain rule set:

(Skill, low) → (Productivity, low)
(Age, 36.3..47.7) & (Experience, low) → (Productivity, low)
(Age, 47.7..59) → (Productivity, high),
(Age, 36.3..47.7) & (Experience, high) → (Productivity, high),

and the possible rule set:

(Age, 25..36.3) → (Productivity, low),
(Age, 36.3..47.7) & (Experience, low) → (Productivity, low),
(Skill, high) & (Age, 25..36.3) → (Productivity, high),
(Age, 47.7..59) → (Productivity, high),
(Age, 36.3..47.7) & (Experience, high) → (Productivity, high).

Rule Induction, Missing Attribute Values and Discretization, Table 5
An example of a discretized and inconsistent decision table

Case	Attributes			Decision
	Age	Skill	Experience	Productivity
1	25..36.3	low	low	low
2	47.7..59	high	high	high
3	25..36.3	low	high	low
4	36.3..47.7	high	low	low
5	47.7..59	high	low	high
6	25..36.3	high	low	low
7	36.3..47.7	high	high	high
8	25..36.3	high	low	high

Missing Attribute Values

For incomplete decision tables there are three important and different possibilities to define lower and upper approximations, called singleton, subset, and concept approximations [9]. Singleton lower and upper approximations were studied, for example in [11,12,19,20]. Note that similar definitions of lower and upper approximations, though not for incomplete decision tables, were studied in [15]. Additionally, note that some other rough-set approaches to missing attribute values were presented in [4] as well.

For the rest of the paper we will assume that all decision values are specified, i. e., they are not missing. Also, we will assume that lost values will be denoted by "?" and "do not care" conditions by "$*$". Additionally, we will assume that for each case at least one attribute value is specified.

For incomplete decision tables the definition of a block of an attribute-value pair must be modified in the following way:

- If for an attribute a there exists a case x such that $\rho(x, a) =$?, i. e., the corresponding value is lost, then the case x should not be included in any blocks $[(a, v)]$ for all values v of attribute a.
- If for an attribute a there exists a case x such that the corresponding value is a "do not care" condition, i. e., $\rho(x, a) = *$, then the case x should be included in blocks $[(a, v)]$ for all specified values v of attribute a.

For Table 6 the blocks of attribute-value pairs are:

$[(\text{Age}, 25)] = \{1, 6\}$,
$[(\text{Age}, 36)] = \{3\}$,
$[(\text{Age}, 42)] = \{4, 7\}$,
$[(\text{Skill}, \text{low})] = \{1, 3, 4, 7\}$,
$[(\text{Skill}, \text{high})] = \{2, 4, 5, 6, 7\}$,

Rule Induction, Missing Attribute Values and Discretization, Table 6
An example of a decision table with missing attribute values

Case	Attributes			Decision
	Age	Skill	Experience	Productivity
1	25	low	*	low
2	?	high	*	high
3	36	low	high	low
4	42	*	low	low
5	?	high	low	high
6	25	high	*	low
7	42	*	high	high

$[(\text{Experience, low})] = \{1, 2, 4, 5, 6\}$,
$[(\text{Experience, high})] = \{1, 2, 3, 6, 7\}$,

For a case $x \in U$ the *characteristic set* $K_B(x)$ is defined as the intersection of the sets $K(x, a)$, for all $a \in B$, where the set $K(x, a)$ is defined in the following way:

- If $\rho(x, a)$ is specified, then $K(x, a)$ is the block $[(a, \rho(x, a)]$ of attribute a and its value $\rho(x, a)$.
- If $\rho(x, a) =$? or $\rho(x, a) = *$ then the set $K(x, a) = U$.

For Table 6 and $B = A$,

$K_A(1) = \{1, 6\} \cap \{1, 3, 4, 7\} = \{1\}$,
$K_A(2) = \{2, 4, 5, 6, 7\}$,
$K_A(3) = \{3\} \cap \{1, 3, 4, 7\} \cap \{1, 2, 3, 6, 7\} = \{3\}$,
$K_A(4) = \{4, 7\} \cap \{1, 2, 4, 5, 6\} = \{4\}$,
$K_A(5) = \{2, 4, 5, 6, 7\} \cap \{1, 2, 4, 5, 6\} = \{2, 4, 5, 6\}$,
$K_A(6) = \{1, 6\} \cap \{2, 4, 5, 6, 7\} = \{6\}$,
$K_A(7) = \{4, 7\} \cap \{1, 2, 3, 6, 7\} = \{7\}$.

Characteristic set $K_B(x)$ may be interpreted as the set of cases that are indistinguishable from x using all attributes from B and using a given interpretation of missing attribute values. For incomplete decision tables lower and upper approximations may be defined in a few different ways. Here we suggest three different definitions of lower and upper approximations for incomplete decision tables, following [9]. Let B be a subset of the set A of all attributes. Let X be any subset of the set U of all cases. The set X is called a *concept* and is usually defined as the set of all cases defined by a specific value of the decision.

Our first definition uses a similar idea as in the previous articles on incomplete decision tables [11,12,19,20], i. e., lower and upper approximations are sets of singletons from the universe U satisfying some properties. We will call these approximations *singleton*. A singleton B-lower approximation of X is defined as follows:

$$\underline{B}X = \{x \in U | K_B(x) \subseteq X\} .$$

A singleton B-upper approximation of X is

$$\overline{B}X = \{x \in U | K_B(x) \cap X \neq \emptyset\} .$$

In our example of the decision table presented in Table 6 let us say that $B = A$. Then the singleton A-lower and A-upper approximations of the two concepts: $\{1, 3, 4, 6\}$ and $\{2, 5, 7\}$ are:

$\underline{A}\{1, 3, 4, 6\} = \{1, 3, 4, 6\}$,
$\underline{A}\{2, 5, 7\} = \{7\}$,
$\overline{A}\{1, 3, 4, 6\} = \{1, 2, 3, 4, 5, 6\}$,
$\overline{A}\{2, 5, 7\} = \{2, 4, 5, 6, 7\}$.

The next possibility is to define lower and upper approximations for incomplete decision tables using characteristic sets instead of singletons. There are two ways to do this. Using the first way, a *subset* B-lower approximation of X is defined as follows:

$$\underline{B}X = \cup\{K_B(x) | x \in U, K_B(x) \subseteq X\} .$$

A *subset* B-upper approximation of X is

$$\overline{B}X = \cup\{K_B(x) | x \in U, K_B(x) \cap X \neq \emptyset\} .$$

For the same decision table, presented in Table 6, the subset A-lower and A-upper approximations are

$\underline{A}\{1, 3, 4, 6\} = \{1, 3, 4, 6\}$,
$\underline{A}\{2, 5, 7\} = \{7\}$,
$\overline{A}\{1, 3, 4, 6\} = \{1, 2, 3, 4, 6, 7\} = U$,
$\overline{A}\{2, 5, 7\} = \{2, 4, 5, 6, 7\}$.

The second way is to modify the subset definition of lower and upper approximation by replacing the universe U from the subset definition by a concept X. A *concept* B-lower approximation of the concept X is defined as follows:

$$\underline{B}X = \cup\{K_B(x) | x \in X, K_B(x) \subseteq X\} .$$

A *concept* B-upper approximation of the concept X is defined as follows:

$$\overline{B}X = \cup\{K_B(x) | x \in X, K_B(x) \cap X \neq \emptyset\}$$
$$= \cup\{K_B(x) | x \in X\} .$$

For the decision table presented in Table 6, the concept A-upper approximations are

$\overline{A}\{1, 3, 4, 6\} = \{1, 3, 4, 6\}$,
$\overline{A}\{2, 5, 7\} = \{2, 4, 5, 6, 7\}$.

Note that for complete decision tables, all three definitions of lower approximations, singleton, subset and concept, coalesce to the same definition. Also, for complete decision tables, all three definitions of upper approximations coalesce to the same definition. This is not true for incomplete decision tables, as our example shows.

MLEM2

MLEM2, a modified version of LEM2 [7], processes numerical attributes differently than symbolic attributes. For numerical attributes MLEM2 sorts all values of a numerical attribute. Then it computes cutpoints as averages for

any two consecutive values of the sorted list. For each cutpoint q MLEM2 creates two blocks, the first block contains all cases for which values of the numerical attribute are smaller than q, the second block contains remaining cases, i. e., all cases for which values of the numerical attribute are larger than q. The search space of MLEM2 is the set of all blocks computed this way, together with blocks defined by symbolic attributes. Starting from that point, rule induction in MLEM2 is conducted the same way as in LEM2. At the very end MLEM2 simplifies rules by merging appropriate intervals for numerical attributes.

The MLEM algorithm induced the following rules from Table 6 certain rule set:

(Age, 25..39) → (Productivity, low),
(Skill, low) & (Experience, low) → (Productivity, low),
(Age, 39..42) & (Experience, high) → (Productivity, high),

and possible rule set:

(Age, 25..39) → (Productivity, low),
(Skill, low) & (Experience, low) → (Productivity, low),
(Skill, high) → (Productivity, high).

Classification System

Rule sets, induced from data sets, are used mostly to classify new, unseen cases. Such rule sets may be used in rule-based expert systems.

There are a few existing classification systems, for example, associated with rule induction systems LERS or AQ. A classification system used in LERS is a modification of the well-known bucket brigade algorithm [5,18]. In the rule induction system AQ, the classification system is based on a rule estimate of probability. Some classification systems use a decision list, in which rules are ordered, the first rule that matches the case classifies it. In this section we will concentrate on a classification system used in LERS.

The decision as to which concept a case belongs to is made on the basis of three factors: *strength*, *specificity*, and *support*. These factors are defined as follows: *strength* is the total number of cases correctly classified by the rule during training. *Specificity* is the total number of attribute-value pairs on the left-hand side of the rule. The matching rules with a larger number of attribute-value pairs are considered more specific. The third factor, *support*, is defined as the sum of products of strength and specificity for all matching rules indicating the same concept. The concept C for which the support, i. e., the following expression

$$\sum_{\substack{\text{matching rules } r \\ \text{describing } C}} \text{Strength}(r) * \text{Specificity}(r)$$

is the largest is the winner and the case is classified as being a member of C.

In the classification system of LERS, if complete matching is impossible, all partially matching rules are identified. These are rules with at least one attribute-value pair matching the corresponding attribute-value pair of a case. For any partially matching rule r, the additional factor, called *Matching_factor* (r), is computed. Matching_factor (r) is defined as the ratio of the number of matched attribute-value pairs of r with a case to the total number of attribute-value pairs of r. In partial matching, the concept C for which the following expression is the largest

$$\sum_{\substack{\text{partially matching} \\ \text{rules } r \text{ describing } C}} \text{Matching\_factor}(r) * \text{Strength}(r) * \text{Specificity}(r)$$

is the winner and the case is classified as being a member of C.

Validation

The most important performance criterion of rule induction methods is the error rate. If the number of cases is less than 100, the *leaving-one-out* method is used to estimate the error rate of the rule set. In leaving-one-out, the number of learn-and-test experiments is equal to the number of cases in the data set. During the ith experiment, the ith case is removed from the data set, a rule set is induced by the rule induction system from the remaining cases, and the classification of the omitted case by rules produced is recorded. The error rate is computed as

$$\frac{\text{total number of misclassifications}}{\text{total number of cases}}.$$

On the other hand, if the number of cases in the data set is greater than or equal to 100, the *tenfold cross-validation* should be used. This technique is similar to leaving-one-out in that it follows the learn-and-test paradigm. In this case, however, all cases are randomly re-ordered, and then a set of all cases is divided into ten mutually disjoint subsets of approximately equal size. For each subset, all remaining cases are used for training, i. e., for rule induction, while the subset is used for testing. This method is used primarily to save time at the negligible expense of accuracy.

Tenfold cross validation is commonly accepted as a standard way of validating rule sets. However, using this method twice, with different preliminary random re-ordering of all cases yields – in general – two different estimates for the error rate [5]. For large data sets (at least 1000

cases) a single application of the train-and-test paradigm may be used. This technique is also known as *holdout*. Two thirds of cases should be used for training, one third for testing.

In yet another way of validation, *resubstitution*, it is assumed that the training data set is identical with the testing data set. In general, an estimate for the error rate is here too optimistic. However, this technique is used in many applications.

Future Directions

Rule induction, a part of supervised learning, is subject to extensive research. Many possible extensions for rule induction include *boosting* or *ensemble learning*, where ensembles of classifiers vote on the final outcome during classification of a new case. *Semi-supervised learning*, where for learning is used not only the case set, pre-classified by an expert, but also a set of not classified cases, is another example.

Bibliography

1. Chan CC, Grzymala-Busse JW (1991) On the attribute redundancy and the learning programs ID3, PRISM, and LEM2. Department of Computer Science, University of Kansas, TR-91-14 20
2. Chmielewski MR, Grzymala-Busse JW (1996) Global discretization of continuous attributes as preprocessing for machine learning. Int J Approx Reason 15:319–331
3. Grzymala-Busse JW (1988) Knowledge acquisition under uncertainty – A rough set approach. J Intell Robot Syst 1:3–16
4. Grzymala-Busse JW (1992) LERS – A system for learning from examples based on rough sets. In: Slowinski R (ed) Intelligent decision support. Handbook of applications and advances of the rough set theory. Kluwer, Dordrecht, pp 3–18
5. Grzymala-Busse JW (1997) A new version of the rule induction system LERS. Fundam Inform 31:27–39
6. Grzymala-Busse JW (2002) Discretization of numerical attributes. In: Klösgen W, Zytkow J (eds) Handbook of data mining and knowledge discovery. Oxford University Press, New York, pp 218–225
7. Grzymala-Busse JW (2002) MLEM2: A new algorithm for rule induction from imperfect data. In: Proceedings of the 9th international conference on information processing and management of uncertainty in knowledge-based systems, IPMU 2002, Annecy, France, pp 243–250
8. Grzymala-Busse JW (2003) A comparison of three strategies to rule induction from data with numerical attributes. In: Proceedings of the international workshop on rough sets in knowledge discovery (RSKD 2003), in conjunction with the European joint conferences on theory and practice of software, Warsaw, pp 132–140
9. Grzymala-Busse JW (2003) Rough set strategies to data with missing attribute values. In: Workshop notes, foundations and new directions of data mining, in conjunction with the 3rd IEEE international conference on data mining, Melbourne, FL, pp 56–63
10. Grzymala-Busse JW (2007) Mining numerical data – A rough set approach. In: Proceedings of the RSEISP'2007, the international conference of rough sets and emerging intelligent systems paradigms, Warsaw, Poland. Lecture Notes in artificial intelligence, vol 4585. Springer, Berlin, pp 12–21
11. Kryszkiewicz M (1995) Rough set approach to incomplete information systems. In: Proceedings of the second annual joint conference on information sciences, pp 194–197
12. Kryszkiewicz M (1999) Rules in incomplete information systems. Inf Sci 113:271–292
13. Lin TY (1989) Chinese Wall security policy – An aggressive model. In: Proceedings of the fifth aerospace computer security application conference, Tucson, AZ, pp 286–293
14. Lin TY (1989) Neighborhood systems and approximation in database and knowledge base systems. In: Proceedings of the ISMIS-89, the fourth international symposium on methodologies of intelligent systems, Charlotte, NC, pp 75–86
15. Lin TY (1992) Topological and fuzzy rough sets. In: Slowinski R (ed) Intelligent decision support. Handbook of applications and advances of the rough set theory. Kluwer, Dordrecht, pp 287–304
16. Pawlak Z (1982) Rough Sets. Int J Comput Inf Sci 11:341–356
17. Pawlak Z (1991) Rough Sets: Theoretical aspects of reasoning about data. Kluwer, Dordrecht
18. Stefanowski J (2001) Algorithms of decision rule induction in data mining. Poznan University of Technology Press, Poznan
19. Stefanowski J, Tsoukias A (1999) On the extension of rough sets under incomplete information. In: Proceedings of the RSFDGrC'1999, 7th international workshop on new directions in rough sets, data mining, and granular-soft computing, pp 73–81
20. Stefanowski J, Tsoukias A (2001) Incomplete information tables and rough classification. Comput Intell 17:545–566