Chapter 4

CELLULAR AUTOMATA WITH ERRORS: PROBLEMS for STUDENTS of PROBABILITY

Andrei Toom
Incarnate Word College
San Antonio, Texas

Abstract

ABSTRACT. This is a survey of some problems and methods in the theory of probabilistic cellular automata. It is addressed to those students who love to learn a theory by solving problems. The only prerequisite is a standard course in probability. General methods are illustrated by examples, some of which have played important roles in the development of these methods. More than a hundred exercises and problems and more than a dozen of unsolved problems are discussed. Special attention is paid to the computational aspect; several pseudo-codes are given to show how to program processes in question.

We consider probabilistic systems with local interactions, which may be thought of as special kinds of Markov processes. Informally speaking, they describe the functioning of a finite or infinite system of automata, all of which update their states at every moment of discrete time, according to some deterministic or stochastic rule depending on their neighbors' states.

From statistical physics came the question of uniqueness of the limit as $t \to \infty$ behavior of the infinite system. Existence of at least one limit follows from the famous fixed-point theorem and is proved here as Theorem 1. Systems which have a unique limit behavior and converge to it from any initial condition may be said to forget everything when time tends to infinity, while the others remember something forever. The former systems are called *ergodic*, the latter *non-ergodic*. In statistical physics the non-ergodic systems help us understand conservation of non-symmetry in macro-objects.

Computer scientists prefer to consider finite systems. Although all non-degenerate finite systems converge, some of them converge enormously slower than others. Some converge *fast*, i.e. lose practically all the information about the initial state in a time which does not depend on their size (or grows as the logarithm of their size, according to another definition), whereas some others converge *slowly*, i.e.can retain information during time which grows exponentially in their size. The latter may help to solve the problem of designing reliable systems consisting of unreliable elements.

There is much in common between the discrete-time and continuous-time approaches to interacting processes. American readers are better acquainted with the continuous-time approach, summarized, for example, in [5]. Methods to prove ergodicity are similar in the continuous and discrete time cases. Our Theorem 2 is a simpler version of the well-known result about ergodicity [39]. The same result in continuous time is given as Theorem 4.1 in Chapter 1 of [5].

Methods to prove non-ergodicity, however, have hitherto been better developed for the discrete time case. I think this is because the discrete time case is technically simpler. For example, the theoretical problems of definition provide no essential difficulties in this case. In a fairly general case every question about the discrete-time systems can be reformulated as a question about 'hidden' independent random variables, and we use this possibility throughout this paper.

The theory of interacting systems is a relatively new branch of mathematics where efficient general methods are still in the process of development. In these circumstances it may be useful to pay special attention to individual cases which defy all known methods. Thus we speak of several more or less specific *Examples* which we call special names for reference purposes. Pluses in the following table show which cases of interaction (boxes) between examples (rows) and general themes (columns) are paid attention to here.

themes / *examples*	computer results	attractors and eroders	proofs of ergodicity	proofs of non-ergodicity	critical values	rate of convergence	chaos approximation
Percolations	+	+	+	+	+	+	+
Flattening	-	+	+	+	-	+	-
NEC-Voting	+	+	+	+	-	+	+
1-Dim. Votings	+	+	+	-	-	-	+
Windroses	+	+	+	+	-	-	+
Soldiers	+	+	+	-	-	+	-
2-Line Voting	-	+	+	-	-	-	+

We would like to share with the reader the flavor of connecting theoretical work with computer simulation. Computer results are placed in the first column of our table, as a natural prerequisite for theoretical work. In fact all of them were obtained using the Monte Carlo method, that is by generation of realizations of the process in question using random numbers. In the vein of this method we shall illustrate the functioning of our systems with pseudo-codes similar to computer programs.

In the sixties and seventies, Ilya Piatetski-Shapiro, who consulted with Roland Dobrushin and Yakov Sinai, challenged members of a seminar of Moscow mathematicians, including me, with results of computer simulations of several carefully chosen examples of interacting random processes with discrete time. He proposed that we prove the observed properties, particularly switches from ergodic to non-ergodic behavior as a result of continuous change of parameters.

[2] is the most complete survey of these developments. Some of our examples were first proposed at that seminar. Examples first proposed in [29] and [41], and described as 2-Percolation and NEC-Voting systems here, moved me to develop a method to prove non-ergodicity [30, 31, 35, 36]. One version of this result is formulated here as Theorem 3.

We prefer to present here results that are explicit and understandable rather than general. More general versions can be found in original papers to which we refer. The only prerequisite for this paper is a standard course in probability. Our potential readers are students who love to learn a theory by solving problems. The idea to write a paper with problems and examples about time-discrete interacting random processes comes from Laurie Snell.

Problems of different levels of difficulty can be found here. **Exercises** are technical; they are for students who strive to become professional mathematicians. **Problems** need more non-trivial reasoning, and some of them were or might have been considered publishable results in the past. I try to refer to papers which contain their solutions, but some have never been published. Some proofs and commentaries are given here; □ marks the end of longer ones. **Unsolved problems** are the most feasible of those whose solutions are unknown to me now. The boundaries between the three certainly are not absolute. After all, our common purpose is to turn every unsolved problem into a series of exercises.

1. SYSTEMS AND DETERMINISTIC CASE

1.1. Systems of Neighborhoods

A *system of neighborhoods* or just a *system* is defined as follows. We assume that there is a finite or countable set which we call *Space*. We call a system finite if its *Space* is finite, and infinite otherwise. The natural number m is called *memory.* (All the examples discussed here belong to the simplest case $m = 1$, but we keep in mind a possibility of interesting examples with $m > 1$.) The set $Time = \{-m, \ldots, -1,\ 0,\ 1,\ 2,\ \ldots\}$ is the set of values of the discrete time variable t. The product $V = Space \times Time$ is called *volume*, whose elements (s, t) are called *points.* Points with $t < 0$ are called *initial* and their set I is called *initial volume.* $V_t = \{(s, t),\ s \in Space\}$ is called t-volume. Any subset $W \subset V$ may be called a *subvolume.*

For every non-initial point w there is a finite sequence $N_1(w), \ldots, N_n(w)$ of points which are called its *neighbors.* The set $N(\cdot)$ of neighbors of a point is called its *neighborhood.* The number n of neighbors is the same for all non-initial points. It is essential that all the neighbors of a point (s, t) have values of their time variable less than t. Sometimes we denote $N_i(s, t) = (n_i(s, t),\ t - t_i(s, t))$ where $n_i(s, t)$ and $t_i(s, t)$ are suitable functions of s and t and $t_i(s, t) > 0$.

Thus, to specify a *system* $\mathcal{S}$ one has to choose *Space*, natural numbers m and n and n neighbors of every non-initial point: $\mathcal{S} = \{Space,\ m,\ n,\ N(\cdot)\}$.

1.1.1. Exercise. Call neighbors of a point its 1-st degree neighbors. Define a point's k-th degree neighbors as neighbors of its $(k-1)$-th degree neighbors. What is the greatest and the smallest number of a point's k-th degree neighbors in systems with a given n ?

Answer: The smallest number is 1. (Or n if we assume that a point's neighbors may not coincide.) The greatest number is n^k.

Comment: For every n there is a *Tree system* (to which we shall refer) in which every point has the greatest possible number n^k of k-th degree neighbors for all k. It can be defined as follows: *Space* is the set of vertices of an infinite directed tree, every vertex of which serves as the end point of n edges and as the starting point of one edge. Neighbors of a point (s,t) have their time components equal to $t-1$ and those space components whence edges go to s.

1.2. Standard Systems and Uniformity

Let $Z = (\ldots -1, 0, 1, \ldots)$ be the ring of integers and let Z_S be the ring of residues modulo S. Thus, Z^d is the d-dimensional discrete space and Z_S^d is the d-dimensional torus. Most of our examples belong to the following *standard* case: *Space* is Z^d or Z_S^d and there are n *neighbor vectors* $v_1, \ldots v_n \in Z^{d+1}$ such that for all non-initial (s,t) and for all $i = 1, \ldots, n$

$$N_i(s,t) = (s,t) + v_i.$$

(Time components of all the neighbor vectors are negative.) In the standard case it is sufficient to state neighbors of the origin $\mathcal{O} = (\bar{0}, 0)$ of $Space \times Time$, to define neighbors of all the points, and this is what we shall typically do.

Standard systems are a special case of uniform systems in which 'all automata are equal.' Call a one-to-one map $a : Space \mapsto Space$ a system's *automorphism* if

$$\forall\ s \in Space,\ t \geq 0,\ i = 1, \ldots, n\ :\ N_i(a(s), t) = a(N_i(s,t)).$$

Automorhisms of a system form a group $\mathcal{A}$. Call $\mathcal{A}$ transitive if

$$\forall\ s_1, s_2 \in Space\ \ \exists a \in \mathcal{A}\ :\ a(s_1) = s_2.$$

Call a system *uniform* if it has a transitive group of automorphisms and $t_i(s,t)$ do not depend on s, t. Call a uniform system *commutative* if it has a commutative transitive group of automorphisms. All the systems treated of in our paper are uniform and all except the 2-Line Voting are commutative. In theory this seems not to be very restrictive. I can mention only one theoretical result [43] where non-uniformity is essential. However, computer simulations sometimes have to be performed on a finite subvolume of Z^d, and in these non-uniform cases one has to specify what is going on at the boundaries.

In uniform systems $n_i(s,t)$ and $t_i(s,t)$ do not actually depend on t, and we shall call them $n_i(s)$ and $t_i(s)$:

$$N_i(s,t) = (n_i(s), \quad t - t_i(s)). \tag{1}$$

Call a uniform system *d-dimensional* if

$$Space = Z^d \times S = \{(i,j) \ : \ i \in Z^d, \ j \in S\} \tag{2}$$

where S is finite, and the automorphism group $\mathcal{A}$ contains the subgroup $\mathcal{Z}^d$ whose elements are shifts at any constant vector $V \in Z^d$:

$$\mathcal{Z}^d = \{z \ : \ (i,j) \mapsto (i+V, j), \ V \in Z^d\}. \tag{3}$$

1.2.1. Exercise. Prove that any standard system is d-dimensional and commutative.

1.2.2. Exercise. Prove that the Tree system, which we defined when answering the exercise 1.1.1, is commutative but non-standard.

Comment: The group of all automorphisms of a Tree system is not commutative. But it has transitive commutative subgroups.

1.2.3. Exercise. In a commutative system every map $n_i \ : \ Space \mapsto Space$ is an automorphism, where $i = 1, \ldots, n$ and $n_i(\cdot)$ is that of formula (1).

Proof: Choose some $s_0 \in Space$. We can choose $a_i \in \mathcal{A}$ such that $a_i(s_0) = n_i(s_0)$. Now let us prove that $a_i(s) = n_i(s)$ for any $s \in Space$. Take $a \in \mathcal{A}$ such that $s = a(s_0)$. Then

$$a_i(s) = a_i(a(s_0)) = a(a_i(s_0)) = a(n_i(s_0)) = n_i(a(s_0)) = n_i(s) \qquad \square$$

1.2.4. Exercise. For a commutative system prove that $\mathcal{A}$ contains such a subgroup $\mathcal{A}_0$ that for any $s, \ s' \in Space$, there is exactly one $a \in \mathcal{A}_0$ such that $a(s) = s'$.

Comment: For any $s \in Space$ denote the subgroup $\mathcal{A}_s = \{a \in \mathcal{A} \ : \ a(s) = s\}$. Generally, $\mathcal{A}_s$ may depend on s, but in the commutative case it does not, and $\mathcal{A}_0 = \mathcal{A}/\mathcal{A}_s$.

Using this exercise, we can choose a *reference element* $s_0 \in Space$ which defines the one-to-one correspondence between $\mathcal{A}_0$ and *Space* by the rule $a \leftrightarrow a(s_0)$.

1.3. States and Configurations

Let us call elements of *Space automata*. Every automaton at any time has the same finite set X_0 of possible states. We may write $X_0 = \{0, \dots, k\}$. The simplest case $X_0 = \{0, 1\}$ alone provides so many unsolved problems that we shall enjoy it as much as possible. To every point $(s, t) \in V$ there corresponds a variable $x_{s,t} \in X_0$ which represent the state of automaton s at time t. Set $X = X_V = X_0^V$ is the configuration space, whose elements are called configurations. Any configuration consists of *components*, that is, elements of A corresponding to all points. 'All a-s' will mean any configuration, all components of which equal the same element $a \in X_0$. $X_I = X_0^I$ stands for the set of *initial conditions*, that is configurations on I. Elements of the configuration space $X_t = X_0^{V_t}$ at time t are called t-configurations. Generally, to any subvolume W there corresponds its configuration space $X_W = S^W$, whose elements are called configurations on W. Also $X^m = X_0^{Space \times [1,m]}$ will be used when denoting operators.

We call a point w a *point of difference* between two configurations s and s' iff $s_w \neq s'_w$. *The set of difference* between two configurations is the set of their points of difference. Call a configuration a *finite perturbation* of another configuration if their set of difference is finite.

1.4. Deterministic Systems

To turn a system into a *deterministic system* we need to choose a set X_0 of states of every point and a *transition function* $\text{tr} : X_0^n \mapsto X_0$. This function gives the state of a non-initial point using states of its neighbors as arguments.

For any deterministic system we define a map $\text{TR} : X_I \mapsto X$ as follows: given an initial configuration x_I, first, the initial components of $\text{TR}(x_I)$ are the same as those of x_I, and, second, all the non-initial components of $y = \text{TR}(x_I)$ are determined inductively according to the formula $y_w = \text{tr}(y_{N(w)})$. We call $\text{TR}(x_I)$ the *trajectory* resulting from the initial condition x_I.

To any deterministic system $\mathcal{D}$ there corresponds a *(deterministic) operator* $D : X^m \mapsto X^m$ defined in the following way: the result of application D to a configuration $x = (x_{s,t}),\ s \in Space,\ t \in [1, m]$ has components

$$D(x)_{s,t} = \begin{cases} x_{s,t+1} & \text{if } t < m, \\ \text{tr}(x_{N(s,t)}) & \text{if } t = m. \end{cases}$$

(If $m = 1$, the upper line is not actually used.) Call a configuration $x \in X^m$ *fixed* for an operator D (and thereby for the corresponding system $\mathcal{D}$) if $D(x) = x$.

1.4.1. Exercise. Prove: If x is fixed then components of x and $\text{TR}(x)$ do not depend on t.

1.4.2. Exercise. Prove that any deterministic system $\mathcal{D}$ and operator D are uniform in the following sense: For any automorphism $\mathcal{A}$ of the system, the corresponding one-to-one maps of X and X^m commute with $\mathcal{D}$ and D respectively.

1.4.3. Problem. For any commutative d-dimensional system with *Space* given by formula (2), prove that if components of an initial state do not depend on j, then components of the resulting trajectory also do not depend on j.

Sketch of proof: From exercise 1.2.3, every n_i is an automorphism. Call $\mathcal{H} : \mathcal{A} \mapsto \mathcal{Z}^d$ the homomorphism of the commutative transitive group $\mathcal{A}$ of the system's automorphisms to the subgroup $\mathcal{Z}^d$ of shifts (3). Thus $\mathcal{H}(n_i)$ is a shift at some vector V_i. Now let $z(w)$ stand for the first coordinate of any point $w \in V$ in the (2) representation of *Space*. The statement in question is assured by the fact that the difference $z(n_i(s))-z(s)$ does not depend on s. This is because this difference equals V_i. □

This exercise shows that commutative d-dimensional systems boil down to standard ones if we apply them only to configurations whose components do not depend on j.

1.5. Pseudo-Codes

Since this article emphasizes a computer approach, we illustrate the functioning of our systems with pseudo-codes of imaginary programs which model them. These pseudo-codes do not obey strict syntactical rules of any actually existing programming language; they explain the algorithm. *Space* and the range of t are assumed to be finite whenever we write a pseudo-code. Lines of every pseudo-code are numbered for reference purposes. For example, the functioning of a deterministic system can be expressed by the following pseudo-code:

1 for $t = -m$ to -1 do
2   for all $s \in Space$ do
3     $x_{s,t} \leftarrow x_{s,t}^{initial}$
4 for $t = 0$ to $t_{max} - 1$ do
5   for all $s \in Space$ do
6     $x_{s,t} \leftarrow \mathrm{tr}(x_{N(s,t)})$

(4)

Here lines 1-3 assign the initial configuration and lines 4-6 calculate t_{max} subsequent t-configurations.

In our imagination, values of $x_{s,t}$ for all $s \in Space$ are assigned simultaneously, at one moment t. However, in the actual programming this is not essential, and practically the line 'for all $s \in Space$ do' can be implemented using d nested cycles if $Space$ is d-dimensional.

The well-known Cellular Automata (see, for example, [1, 7]) are standard deterministic systems with $m = 1$ and $Space = Z^d$.

1.5.1. Exercise. Write a pseudo-code for the well-known Game of Life, that is, a cellular automaton with $Space = Z^2$ and $X_0 = \{0, 1\}$. Here neighbors of a cell are its eight nearest neighbors. A cell in the state 0 is called 'dead' and a cell in the state 1 is called 'alive'. The transition rule is: a dead cell becomes alive at the next moment only if it has exactly three live neighbors and an alive cell stays alive only if it has two or three live neighbors.

2. MEASURES

2.1. Measures and Topology

For any configuration space X_W we shall consider the set $\mathcal{M}_W$ of probabilistic, i.e. normed measures on the σ-algebra generated by cylinder sets in the product-space X_W. Most often we deal with $\mathcal{M}_V$, which we call simply $\mathcal{M}$. Call a measure *degenerate* if it has a zero value on at least one cylinder set. Let δ_s stand for the δ-measure concentrated on one configuration s. A product-measure on any configuration space $X_W = X_0^W$ is a product of measures for coordinates. A *mixture* of several normed measures is their linear combination with non-negative coefficients whose sum equals 1. A measure is called uniform if it is fixed for any map on $\mathcal{M}$ corresponding to an automorphism.

Let the *weak topology* correspond to convergence on all the cylinder sets and let the *strong topology* correspond to convergence on all the elements of the σ-algebra generated by them. The following exercises show why the weak topology is typically used when treating infinite interactive systems:

2.1.1. Exercise. Prove that if W is finite, the weak and strong topologies on $\mathcal{M}_W$ coincide, but in the infinite case they are different.

2.1.2. Exercise. Let $B(p)$ stand for the uniform product-measure on the product-space $X = \{0, 1\}^Z$ in which every component equals 1 with probability p, independently of others. Prove that when $p \to 0$, the measure $B(p)$ tends to the measure concentrated in the configuration *'all zeroes'* in the weak topology, but not in the strong topology.

2.1.3. Exercise. Prove:

1) For any configuration space X_0^W, where X_0 is compact (e.g. finite) and W is finite or countable, $\mathcal{M}_W$ is compact in the weak topology.

2) For any configuration space X_0^W, where $|X_0| > 1$ and W is infinite, $\mathcal{M}_W$ is not compact in the strong topology.

For any two normed measures μ_1 and μ_2 on a measurable space X, call a measure on $X \times X$ *a coupling* of μ_1 and μ_2 if its marginals are μ_1 and μ_2. Coupling of measures is the simplest version of couplings. Below we shall speak of couplings of random systems [5, 4]; [16, 23] are some of the newest papers on coupling.

2.2. A Distance between Measures

As long as we are interested only in convergence vs. non-convergence, the weak topology is sufficient for us. But as soon as we want to speak about how fast systems converge, we need a distance between measures. For any coupling c of μ_1 and μ_2 on a configuration space $X_W = X_0^W$, let

$$rift(c) = \sup\{c(x_w^1 \neq x_w^2),\ w \in W\},$$

where $x^1 = (x_w^1)$ and $x^2 = (x_w^2)$ stand for the first and second components in $X^1 \times X^2 = X_W \times X_W$. Define $dist(\mu_1, \mu_2)$, i.e. the *distance* between μ_1 and μ_2, as the infimum of $rift(c)$ over all the couplings c of μ_1 and μ_2.

2.2.1. Exercise. Prove: Any value of $rift(\cdot)$ or $dist(\cdot)$ belongs to $[0, 1]$.

2.2.2. Exercise. Let $\mu_1, \mu_2 \in \mathcal{M}_W$ where $X_W = \prod_{w \in W} X_w$. Let $\mu_1|_w$ and $\mu_2|_w$ stand for projections of μ_1 and μ_2 to X_w. Prove:

$$dist(\mu_1, \mu_2) = \sup\{dist(\mu_1|_w, \mu_2|_w)\ :\ w \in W\}.$$

2.2.3. Exercise. Let S be a cylinder subset of $X_W = \{(x_w),\ w \in W\}$ indicated by a condition involving some r components $x_1, \ldots, x_r$. Prove:

$$\forall\, \mu_1, \mu_2 \in \mathcal{M}_W\ :\ |\mu_1(S) - \mu_2(S)| \leq r \cdot dist(\mu_1, \mu_2).$$

2.2.4. Problem. Prove that $dist$ is a metric, that is

1) $dist(\mu_1, \mu_2) \geq 0$.

2) $dist(\mu_1, \mu_2) = dist(\mu_2, \mu_1)$.

3) $dist(\mu_1, \mu_2) = 0 \Longleftrightarrow \mu_1 = \mu_2$.

4) $dist(\mu_1, \mu_3) \leq dist(\mu_1, \mu_2) + dist(\mu_2, \mu_3)$.

Proof: We leave proof of 1) and 2) to the reader and 3) follows from exercise 2.2.3. The following proof of 4) for the finite case was proposed by Eugene Speere. Take any μ_1, μ_2 and μ_3 on X_W. Let c_{12} be any coupling of μ_1 and μ_2 and c_{23} be any coupling of μ_2 and μ_3. All we have to do is to present a coupling c_{13} of μ_1 and μ_3 for which

$$\mathit{rift}(c_{13}) \leq \mathit{rift}(c_{12}) + \mathit{rift}(c_{23}). \tag{5}$$

Take the following measure c_{123} on $X^1 \times X^2 \times X^3 = X_W^3$:

$$c_{123}(x^1, x^2, x^3) = \begin{cases} 0 & \text{if } \mu_2(x^2) = 0, \\ c_{12}(x^1, x^2) \times c_{23}(x^2, x^3)/\mu_2(x^2) & \text{otherwise.} \end{cases}$$

Since μ_2 is a marginal of c_{23}, the sum of $c_{23}(x^2, x^3)$ over x^3 gives $\mu_2(x^2)$. Therefore the sum of $c_{123}(x^1, x^2, x^3)$ over x^3 gives $c_{12}(x^1, x^2)$. Analogously, the sum of $c_{123}(x^1, x^2, x^3)$ over x^1 gives $c_{23}(x^2, x^3)$. Thus c_{123} is a normed measure and the following is a coupling of μ_1 and μ_3:

$$c_{13}(x^1, x^3) = \sum_{x^2} c_{123}(x^1, x^2, x^3).$$

Thus defined c_{13} fits our claim (5) because

$$\forall\, w \in W \;:\; c_{13}(x^1_w \neq x^3_w) \leq c_{12}(x^1_w \neq x^2_w) + c_{23}(x^2_w \neq x^3_w). \qquad \square$$

2.2.5. Exercise. Extend this proof to the infinite case.

Call a coupling c of μ_1 and μ_2 their *fine coupling* if $\mathit{rift}(c) = \mathit{dist}(\mu_1, \mu_2)$.

2.2.6. Exercise. Let W be finite. Prove that our metric *dist* is continuous both in the weak and strong topology and that $\mathcal{M}_W$ is compact w.r.t. our metric. Prove also that in this case for every two measures there exists a fine coupling.

2.2.7. Exercise. Let W be infinite and $|X_0| > 1$. Prove that $\mathcal{M}$ is not compact with *dist* as a metric. Prove also that *dist* is not continuous in the weak topology, because there is a μ and a sequence $\mu_1, \mu_2, \ldots$ which tends to μ, such that $\mathit{dist}(\mu_n, \mu)$ does not tend to 0.

2.2.8. Unsolved problem. Is it true that for any two normed measures on X_W there exists a fine coupling ?

Comment: The difficulty lies in non-compactness of $\mathcal{M}$. However, I can not present measures for which there is no fine coupling.

2.2.9. Exercise. Prove that fine coupling may be non-unique.

Proof: For example, let some x_w certainly equal 0 in μ_1 and certainly equal 1 in μ_2. Then *rift* of any coupling of these measures equals 1, and all their couplings are fine. □

The following are some cases when fine coupling exists and sometimes is unique, even with an infinite W:

2.2.10. Exercise. If $\mu_1 = \mu_2$ then their fine coupling c_{fine} exists and is unique. Prove this and write an explicit formula for $c_{fine}(S)$ for any cylinder set $S \subset X_W \times X_W$.

2.2.11. Exercise. Let μ_1 be an arbitrary measure on X_W and let μ_2 be a δ-measure δ_y concentrated on one configuration $y \in X_W$. Then μ_1 and μ_2 have only one coupling at all. Therefore their fine coupling exists and is unique and $dist(\mu_1, \mu_2) = \sup\{\mu_1(x_w \neq y_w) \;:\; w \in W\}$.

2.2.12. Problem. Let $|W| = 1$, i.e. $X_W = X_0$, and μ_1 and μ_2 be any measures on X_W. Denote $m(a) = \min(\mu_1(a), \mu_2(a))$ for all $a \in X_0$. Then:

1)

$$\begin{aligned} dist(\mu_1, \mu_2) &= \max\{\mu_1(S) - \mu_2(S) \;:\; S \subset X_0\} \\ &= \max\{\mu_2(S) - \mu_1(S) \;:\; S \subset X_0\} \\ &= \sum_{a \in X_0} (\mu_1(a) - m(a)) = \sum_{a \in X_0} (\mu_2(a) - m(a)). \end{aligned}$$

2) The set of fine couplings of μ_1 and μ_2 is non-empty and coincides with the set of those couplings of μ_1 and μ_2 which have the form $c_{fine} = c_0 + c_1$ where

$$c_0(x_1, x_2) = \begin{cases} m(x_1) & \text{if } x_1 = x_2, \\ 0 & \text{otherwise.} \end{cases}$$

and c_1 is any measure on $X_0 \times X_0$, for which

$$\begin{cases} \forall\, x_1 \;:\; \sum_{x_2 \in X_0} c_1(x_1, x_2) = \mu_1(x_1) - m(x_1), \\ \forall\, x_2 \;:\; \sum_{x_1 \in X_0} c_1(x_1, x_2) = \mu_2(x_2) - m(x_2). \end{cases} \tag{6}$$

Comment: One example of c_1 which satisfies (6) is

$$c_1^*(x_1, x_2) = (\mu_1(x_1) - m(x_1)) \times (\mu_2(x_2) - m(x_2)),$$

and the resulting $c_{fine}^* = c_0 + c_1^*$ is Vaserstein's $*$-operation which was defined in [39] for any measurable X_0.

2.2.13. Problem. Let μ_1 and μ_2 be product-measures on the product-space X_W:

$$\mu_1 = \prod \mu_1^s, \quad \mu_2 = \prod \mu_2^s, \quad s \in Space.$$

Then any product of fine couplings of μ_1^s and μ_2^s is a fine coupling of μ_1 and μ_2.

2.2.14. Problem. Let $X_0 = \{0, \ldots, k\}$, let $X = X_0^W$, and let $\mathcal{B}(p)$ stand for the product-measure on the product-space X, every factor of which is the same measure $p = (p_0, \ldots, p_k)$ on X_0. Prove that for any $\mathcal{B}(p)$ and $\mathcal{B}(q)$ there exists a fine coupling and

$$dist(\mathcal{B}(p), \mathcal{B}(q)) = \max\{p(S) - q(S) \ : \ S \subset X_0\}.$$

Clearly, there are many ways to define a distance between measures. The distance we describe here seems to be useful at least in the present context. The facts that it is not continuous in the weak topology and that $\mathcal{M}$ is not compact seem to be an inevitable trade-off for relevance to problems about how quickly or slowly large interactive systems converge. The idea of this definition of distance was present in [39] (in the form of that definition on p. 50 which speaks of estimate α_k). However, the $*$-operation proposed in [39], generally speaking, does not provide a fine coupling. For example, the $*$-operation applied directly to any different uniform product-measures on X^W with $|X| \geq 2$ and infinite W has *rift* equal to 1.

3. RANDOM SYSTEMS

We shall define *random systems* using (every time the same) mutually independent *hidden* variables h_w for all points $w \in V$, everyone of which is uniformly distributed on $[0, 1]$; thus we have the *hidden* product-measure η on the hidden configuration space $H = [0, 1]^V$ of the hidden variables. In pseudo-codes the same idea is expressed in generating and using random numbers *rnd*. We assume that every call of *rnd* in a pseudo-code generates a random number uniformly distributed on $[0, 1]$ which is independent from all the past events, including past calls of *rnd*.

3.1. Special case: $X_0 = \{0, 1\}$

All we need to define a random system in this case is a function $p_0 : X_0^n \mapsto [0, 1]$, the conditional probability of a point to be in the state 0 given states of its neighbors. A point's probability to be in the state 1 is $p_1(\cdot) = 1 - p_0(\cdot)$. Values of $p_0(\cdot)$ and $p_1(\cdot)$ are called *transition probabilities*. If at least one of them equals 0, we call the random system *degenerate*. The following pseudo-code describes functioning of an arbitrary random system with $X_0 = \{0, 1\}$ and a finite *Space*:

```
1   for t = -m to -1 do
2       for all s ∈ Space do
3           x_{s,t} ← x_{s,t}^{initial}
4   for t = 0 to t_max - 1 do
5       for all s ∈ Space do
6           if rnd < p_0(x_{N(s,t)}) then x_{s,t} ← 0 else x_{s,t} ← 1
```

(7)

Here lines 1-3 assign initial values and lines 4-6 calculate a realization of the process during t_{max} time steps. In mathematical terms this means that we define our process as the measure on X induced by any initial measure and the hidden measure η with the following map from $X_I \times H$ to X:

$$\begin{cases} \forall\ s \in Space,\ \ t \in [-m, -1] : x_{s,t} = x_{s,t}^{initial}, \\ \forall\ s \in Space,\ \ t \in [\ \ \ 0,\ \infty) : x_{s,t} = \begin{cases} 0 & \text{if } h_{s,t} < p_0(x_{N(s,t)}), \\ 1 & \text{otherwise.} \end{cases} \end{cases} \tag{8}$$

3.2. General case: $X_0 = \{0, \ldots, k\}$

Take a function p from X_0^n to the set of normed measures on X_0 and call its values $p(a \mid b_1, \ldots, b_n)$ *transition probabilities*; they serve as conditional probabilities of $x_w = a$ if $x_{N(w)} = (b_1, \ldots, b_n)$. (Denotations $p_0(\cdot)$ and $p_1(\cdot)$, which we used in the case $X_0 = \{0, 1\}$, now are substituted by $p(0|\cdot)$ and $p(1|\cdot)$.) If at least one of the transition probabilities equals 0, we call the random system *degenerate*. As soon as parameters $p(a \mid b_1, \ldots, b_n)$ are given, we can define our process as a measure $proc(\mu_I) \in \mathcal{M}$ induced by any initial measure μ_I and the hidden measure η with the following map:

$$\begin{cases} \forall\ s \in Space,\ \ t \in [-m, -1] : x_{s,t} = x_{s,t}^{initial}, \\ \forall\ s \in Space,\ \ t \in [\ \ \ 0,\ \infty) : x_{s,t} = \\ \qquad\qquad \begin{cases} 0 & \text{if } h_{s,t} \le p(0|x_{N(s,t)}), \\ i & \text{if } \sum_{j=0}^{i-1} p(j|x_{N(s,t)}) < h_{s,t} \le \sum_{j=0}^{i} p(j|x_{N(s,t)}) \\ & \text{for all } i > 0. \end{cases} \end{cases} \tag{9}$$

3.2.1. Exercise. Write a pseudo-code for a random system with $X_0 = \{0, 1, \ldots, k\}$.

3.2.2. Exercise. Let a random system be given. Prove that a mixture of two processes is a process too.

3.2.3. Exercise. Prove that if a sequence of processes of a random system has a limit (in the weak topology), this limit is also a process for the random system.

3.2.4. Exercise. For any deterministic system with a transition function $\mathrm{tr}(\cdot)$ let us define the corresponding random system with the same *Space*, m and neighborhoods as follows:

$$p(a_w | b_{N(w)}) = \begin{cases} 1 & \text{if } a_w = \mathrm{tr}(b_{N(w)}), \\ 0 & \text{otherwise.} \end{cases}$$

Prove that for any deterministic initial condition the resulting process is concentrated in the resulting trajectory.

3.3. Deterministic Systems with Noise

Here we consider random systems which result from deterministic ones by adding random noise.

First assume $X_0 = \{0, 1\}$. In this case we need only two non-negative parameters ε and δ where $\varepsilon + \delta \le 1$. Take the process induced by the hidden measure η with the following map from H to X:

$$\begin{cases} \forall\, s \in Space,\ t \in [-m, -1] : x_{s,t} = x_{s,t}^{initial}, \\ \forall\, s \in Space,\ t \in [\ \ 0,\ \infty) : x_{s,t} = \begin{cases} 0 & \text{if } h_{s,t} < \varepsilon, \\ 1 & \text{if } h_{s,t} > 1 - \delta, \\ \mathrm{tr}(x_{N(s,t)}) & \text{otherwise.} \end{cases} \end{cases} \tag{10}$$

We shall say that this system results from the deterministic one by adding an ε-δ noise. To obtain the corresponding pseudo-code it is sufficient to add the following two lines to the pseudo-code (4):

```
h ← rnd
if h < ε then x_{s,t} ← 0 else if h > 1 − δ then x_{s,t} ← 1
```

(11)

The resulting pseudo-code (4) and (11) imitates a process in which every automaton first follows the deterministic rule, then makes a two-way random error: turns 0 with probability ε and turns 1 with probability δ. We call the ε-δ noise *symmetric* if $\varepsilon = \delta$, *biased* if $\varepsilon \ne \delta$, and *one-way* or *degenerate* if $\varepsilon = 0$ or $\delta = 0$.

3.3.1. Exercise. Show that for any $\mathrm{tr}(\cdot)$, ε, and δ there exists $p(\cdot)$ such that the process (8) is the same as the process (10).

Now the general case $X_0 = \{0, \dots, k\}$. We need non-negative parameters ε_i for all $i \in X_0$ whose sum does not exceed 1. The process is induced by any initial measure and the hidden measure η with the following map:

$$\begin{cases} \forall\, s \in Space,\ t \in [-m, -1] : x_{s,t} = x_{s,t}^{initial}, \\ \forall\, s \in Space,\ t \in [\ \ 0,\ \infty) : x_{s,t} = \begin{cases} 0 & \text{if } h_{s,t} < \varepsilon, \\ i & \text{if } \sum_{j=0}^{i-1} \varepsilon_j \le h_{s,t} < \sum_{j=0}^{i} \varepsilon_j, \\ & \text{for } i > 0, \\ \mathrm{tr}(x_{N(s,t)}) & \text{otherwise.} \end{cases} \end{cases} \tag{12}$$

To obtain the corresponding pseudo-code we may add the following lines to the pseudo-code (4):

```
h ← rnd
i ← 0; θ ← ε_0
while i < k and θ < h do
      i ← i + 1; θ ← θ + ε_i
if θ ≥ h then x_{s,t} ← i
```

(13)

The resulting pseudo-code (4) and (13) imitates a process in which every automaton first follows the deterministic rule, after which the noise turns it into the state i with probability ε_i for all $i \in A$. We call the noise *degenerate* if some $\varepsilon_i = 0$.

3.3.2. Exercise. Show that for any $\mathrm{tr}(\cdot)$ and $\varepsilon_i,\ i \in A$ there exists $p(\cdot)$ such that the process (9) is the same as the process (12).

3.4. Definitions of Ergodicity, Slow and Fast Convergence

By definition, *time-shift* of the volume V maps any (s, t) to $(s, t+1)$. *Time-shifts* $T : X \mapsto X$ and $\mathcal{T} : \mathcal{M} \mapsto \mathcal{M}$ are defined in the same vein:

$$\forall\, x \in X,\ s \in Space,\ t \in Time\ :\ T(x)_{s,t} = x_{s,t+1},$$

$$and\ \ \forall\, S \subseteq X\ :\ \mathcal{T}(\mu)(S) = \mu(T(S)).$$

3.4.1. Exercise. Prove that if μ is a process of a random system then $\mathcal{T}(\mu)$ also is a process of the same system.

Call a process μ *invariant* for a system if $\mathcal{T}(\mu) = \mu$.

Theorem 1. Any random system has at least one invariant process.

Proof: Take any process μ_1 and for all $n = 2,\ 3, \dots$ define

$$\mu_n = \frac{1}{n} \sum_{i=0}^{n-1} \mathcal{T}^n(\mu_1). \tag{14}$$

Since all $\mathcal{T}^n(\mu_1)$ are processes (exercise 3.4.1), all their mixtures are processes too (exercise 3.2.2). Since $\mathcal{M}$ is compact (exercise 2.1.3), any sequence in $\mathcal{M}$ has at least one limit point (may be more). From exercise 3.2.3 every limit point of μ_n is an invariant process. □

Definition. Say that a random system is *ergodic* if it has only one invariant process μ_{inv} and for all its processes μ the limit $\lim_{n\to\infty} \mathcal{T}^n(\mu)$ exists and equals μ_{inv}.

3.4.2. Exercise. Does exist a non-ergodic random system which has only one invariant process ?

Answer: Yes. Take a deterministic system with *Space* consisting of one element 0, with $N(0,t)$ consisting of one point $(0, t-1)$, in which $X_0 = \{0, 1\}$ and $\mathrm{tr}(a) = 1 - a$.

3.4.3. Unsolved problem. Does there exist a non-degenerate non-ergodic random system which has only one invariant process ?

3.4.4. Exercise. Prove that any non-degenerate finite random system is ergodic.

3.5. Patterns

Let us say that all those standard systems with the same d, m, and $N(\mathcal{O})$ belong to one *pattern.* In the same vein, a deterministic pattern is a quadruple $(d,\ m,\ N(\mathcal{O}),\ \mathrm{tr}(\cdot))$ and a random pattern is a quadruple $(d,\ m,\ N(\mathcal{O}),\ p(\cdot))$. This allows us to compare behavior of the infinite system and of finite systems with various S belonging to one pattern $(d,\ m,\ N(\mathcal{O}))$. Let μ_{inv} stand for the invariant process of any system which has only one invariant process. For any random pattern let us define $\rho(S,t)$ as the supremum of $dist(\mathcal{T}^t(\mu), \mu_{inv})$ over all processes μ of the finite system with $Space = Z_S^d$ that belongs to the given pattern.

Let us say that a pattern *converges fast* if there is a positive constant $q < 1$ such that

$$\forall\ S, t\ :\ \rho(S,t) \le q^t.$$

Let us say that a pattern *converges slowly* if there are $\varepsilon > 0$ and $Q > 1$ such that

$$\forall\ S, t\ :\ t \le Q^S \Longrightarrow \rho(S,t) \ge \varepsilon.$$

3.5.1. Unsolved problem. It is conjectured that any random pattern converges fast iff the infinite system is ergodic and it converges slowly iff the infinite system is non-ergodic. Is it always true ?

Comment: Note that most computer simulations refer directly to finite systems. (Typically they imitate functioning of some finite system using the Monte Carlo method.) Thus we certainly need to know (or assume) some relation between finite and infinite systems whenever we interpret results of computer simulations as telling us something about ergodicity or non-ergodicity of infinite systems.

3.6. Operators

It is possible to define ergodicity in the following equivalent way. Let $\mathcal{M}^m$ stand for the set of normed measures on X^m. To any random system there corresponds a linear *operator* $P : \mathcal{M}^m \mapsto \mathcal{M}^m$ whose definition is based on the following assumption: application of P to a δ-measure concentrated in a state $y = (y_1, \ldots, y_m) \in X^m$ is a product of the following measures pertaining to components:

$$\text{the distribution of } x_{s,t} \text{ is } \begin{cases} \delta_{y_{s,t+1}} & \text{if } t < m, \\ p(y_{N(s,t)}) & \text{otherwise.} \end{cases}$$

(If $m = 1$, the upper line is not used.) Call a measure μ *invariant* for an operator P if $P(\mu) = \mu$.

3.6.1. Exercise. Prove that any operator has at least one invariant measure.

Proof: analogous to proof of Theorem 1.

Let us say that operator P of a random system is *ergodic* if it has only one invariant measure μ_{inv} and for all μ the limit $\lim_{n\to\infty} P^n(\mu)$ exists and equals μ_{inv}.

3.6.2. Exercise. Prove that a random system is ergodic iff its operator is ergodic.

3.7. Monotony

We shall use the signs $\preceq$ and $\succeq$ to speak about (perhaps, partial) ordering of sets. For example, any set S of real numbers is said to be *naturally ordered* if $\forall\, i, j \in S : i \preceq j \iff i \leq j$. Call a function f from one ordered set to another *monotonic* if $s \preceq s' \Longrightarrow f(s) \preceq f(s')$.

Let X_0 be ordered (for example, naturally). Then any configuration space X_0^W is partially ordered by the rule $s \preceq s' \iff \forall\, i \in W \ : \ s_i \preceq s'_i$. Call a deterministic system *monotonic* if

$$\forall\, s, s' \in X_I : s \preceq s' \Longrightarrow \mathrm{TR}(s) \preceq \mathrm{TR}(s').$$

3.7.1. Exercise. Prove that a deterministic system is monotonic iff its transition function is monotonic.

For any two measures $\mu_1,\ \mu_2 \in \mathcal{M}_W$ on any configuration space X_W we say that $\mu_1 \preceq \mu_2$ iff $\mu_1 f \leq \mu_2 f$ for any monotonic function $f : X_W \mapsto R$.

Call a random system *monotonic* if

$$\forall\, \mu_1, \mu_2 \in \mathcal{M}_I \ :\ \mu_1 \preceq \mu_2 \Longrightarrow\ proc(\mu_1) \preceq proc(\mu_2).$$

Call an operator P *monotonic* if

$$\forall\, \mu_1, \mu_2 \in \mathcal{M}^m \ :\ \mu_1 \preceq \mu_2 \Longrightarrow P(\mu_1) \preceq P(\mu_2).$$

3.7.2. Exercise. Prove that a random system is monotonic iff its operator is monotonic.

3.7.3. Exercise. Prove that if a map (9) from $\mathcal{M}_I \times H$ to X is monotonic ($[0,1]$ naturally ordered), then the resulting random system is monotonic.

3.7.4. Exercise. Any order can be reversed by turning $\preceq$ into $\succeq$ and vice versa. Prove that this reversal keeps all our kinds of monotony.

3.7.5. Problem. Let A be ordered and have a smallest element $a_{\min}$ and a largest element $a_{\max}$. Thus X has the smallest state $s_{\min} =$'all $a_{\min}$' and the largest state $s_{\max} =$'all $a_{\max}$'. Then for any monotonic operator P the limits $\lim_{t\to\infty} P^t(s_{\min})$ and $\lim_{t\to\infty} P^t(s_{\max})$ exist. These limits are equal iff P is ergodic.

4. PROOFS OF ERGODICITY AND FAST CONVERGENCE

4.1. Case $X_0 = \{0, 1\}$

To prove ergodicity and fast convergence of a system, it is convenient to use its *coupling* (see [4] and [5]) which we shall first illustrate by the following pseudo-code:

```
1   for t = -m to -1 do
2       for all s ∈ Space do
3           x_{s,t} ← x_{s,t}^{initial}
4           y_{s,t} ← y_{s,t}^{initial}
5           m_{s,t} ← 0
6   for t = 0 to t_{max} - 1 do
7       for all s ∈ Space do
8           x_{s,t} ← f(x_{N(s,t)}, rnd)
9           y_{s,t} ← f(y_{N(s,t)}, rnd)
10          m_{s,t} ← min(m_{N_1(s,t)}, ..., m_{N_n(s,t)})
11          h ← rnd
12          if h < ε then
13              x_{s,t} ← 0
14              y_{s,t} ← 0
15              m_{s,t} ← 1
16          else if h > 1 - δ then
17              x_{s,t} ← 1
18              y_{s,t} ← 1
19              m_{s,t} ← 1
```

(15)

Let us first ignore all the lines which deal with the values $m_{s,t}$ and concentrate our attention on those which deal with $x_{s,t}$ and $y_{s,t}$. Then we see that we are modelling simultaneously two processes of the same system, using for them a common source of random noise. In both processes every automaton every time does the following: first, due to line 8 or 9, it assumes some value which depends in a random way on its neighbors, and second, it makes a random error, becoming 0 with probability ε due to line 13 or 14 and becoming 1 with probability δ due to line 17 or 18. (We assume that $\varepsilon + \delta \leq 1$.)

A coupling c of k random systems with a common $\mathcal{S}$ is a random system with the same $\mathcal{S}$ system of neighborhoods, $X_0(c) = X_0(1) \times X_0(2) \times \cdots \times X_0(k)$, and the transition probability which is a coupling (in the sense of coupling of measures defined in section 2.1) of their transition probabilities. Note that according to this definition the coupled systems need not to be identical; it is sufficient for them to correspond to the same $\mathcal{S}$ system of neighborhoods, that is, to have the same *Space*, m, and $N(\cdot)$. Couplings typically used in the literature are couplings of a system with itself in our terms. The pseudocode (15) describes coupling of a given system with itself and with another system which deals with the values $m_{s,t}$ (about which we comment below). We assume that the function $f : X_0^n \times [0,1] \mapsto X_0$ is chosen in such a way that each marginal process coincides with a certain given one. This assumption is substantiated below in the form of exercise 4.1.4.

4.1.1. Exercise. Prove that marginals of any process of a coupling of several random systems are processes of the coupled random systems.

Generally, any system has many different couplings. But only those, where points of difference die out fast enough for any pair of initial conditions, help to prove ergodicity. To check this, our pseudo-code assumes that at every point w we have a special *mark* m_w which may equal 0 or 1. This is to mark loss of the memory (when we are sure about it). Let us say that a point w is *marked* if $m_w = 1$. Initially all the points are unmarked (line 5) and become marked (lines 15 and 19) whenever they are assigned values which certainly do not depend on the prehistory and therefore cannot be points of difference.

Let us estimate the percentage U_t of unmarked points at time t in this process. Let k_i stand for the number of the origin's neighbors whose time coordinate equals $-i$. (Note that $k_1 + \cdots + k_m = n$.) The percentage of unmarked points which proliferate into V_t as a result of line 10 does not exceed $k_1 U_{t-1} + \cdots + k_m U_{t-m}$. Then $\varepsilon + \delta$ of them die as a result of lines 15 and 19. Thus

$$U_{-m} = \cdots = U_{-1} = 1, \qquad U_t \le (1-\varepsilon-\delta)(k_1 U_{t-1} + \ldots + k_m U_{t-m}).$$

Therefore $U_t \le \tilde{U}_t$ where

$$\tilde{U}_{-m} = \cdots = \tilde{U}_{-1} = 1, \qquad \tilde{U}_t = (1-\varepsilon-\delta)(k_1 \tilde{U}_{t-1} + \ldots + k_m \tilde{U}_{t-m}).$$

4.1.2. Exercise. Prove that $\tilde{U}_t \to 0$ iff $(1-\varepsilon-\delta)(k_1 + \cdots + k_m) < 1$, that is $1 - \varepsilon - \delta < 1/n$. Prove also that if $\tilde{U}_t$ tends to 0, then it does so exponentially.

4.1.3. Problem. For any (finite or infinite) system modelled by the pseudo-code (15), prove that dying out of unmarked points guarantees ergodicity and fast convergence.

Now to apply our results to a general process with $X_0 = \{0, 1\}$. This is done by the following:

4.1.4. Exercise. Prove that for any two processes of one random system described by the pseudo-code (7) and formula (8), there is a coupling which can be represented by the pseudo-code (15) with a suitable f and

$$\begin{aligned} \varepsilon &= \min\{p_0(x_{N(s,t)}),\ x_{N(s,t)} \in X_{N(s,t)}\}, \\ \delta &= \min\{p_1(x_{N(s,t)}),\ x_{N(s,t)} \in X_{N(s,t)}\}. \end{aligned}$$

4.1.5. Problem. Prove that if $X_0 = \{0, 1\}$ and

$$\max\{p_0(x_{N(s,t)}),\ x_{N(s,t)} \in X_{N(s,t)}\} - \min\{p_0(x_{N(s,t)}),\ x_{N(s,t)} \in X_{N(s,t)}\} < 1/n$$

then the system is ergodic and converges fast.

Comment: This follows from 4.1.2 - 4.1.4.

4.2. General case $X_0 = \{0, \ldots, k\}$

Let **P** stand for the set of functions from A^n to the set of normed measures on A. Let a δ-function $\delta_a \in \mathbf{P}$ stand for a constant function which maps any argument to the same δ-measure on A concentrated in some $a \in X_0$. To prove ergodicity of a random system, it makes sense to represent its transition probability p as a mixture of elements of **P**, in which δ-functions have the maximal possible sum of coefficients. Let $\Sigma(p)$ stand for this maximal possible sum.

4.2.1. Exercise. Prove that $\Sigma(p) = 1 - \max\{|p(B|s_1) - p(B|s_2)|,\ s_1, s_2 \in X_0^n,\ B \subset X_0\}$.

4.2.2. Exercise. Prove that if $1 - \Sigma(p) < 1/n$ then the system is ergodic and converges fast.

Based on these two exercises, the following holds:

Theorem 2. If

$$\max\{|p(B|s_1) - p(B|s_2)|,\ s_1, s_2 \in X_0^n,\ B \subset X_0\} < 1/n$$

then the random system is ergodic and converges fast.

Comment: See a similar result in [39].

5. PERCOLATION SYSTEMS

5.1. General Percolation

Since marking proved useful, it makes sense to examine it as such. This leads us to our first examples. The following pseudo-code is obtained from (15) by deleting everything which pertained to $x_{s,t}$ and $y_{s,t}$ and denoting $\theta = \varepsilon + \delta$:

```
for t = −m to −1 do
        for all s ∈ Space do
                m_{s,t} ← 0
for t = 0 to t_max − 1 do
        for all s ∈ Space do
                m_{s,t} ← min(m_{N_1(s,t)}, ..., m_{N_n(s,t)})
                if rnd < θ then m_{s,t} ← 1
```

(16)

The process described by this code can be called a *Percolation process* because of the following interpretation. Assume that some fluid is supplied to all the initial points. Every non-initial point is connected with its neighbors by pipes which pass the fluid in one direction: to the point from its neighbors. But in every non-initial point there is a tap which is closed with probability θ independently from all the other taps. Then a point is unmarked iff the fluid percolates to this point.

Lines 1-3 in this pseudo-code set the initial condition *'all zeroes'* . Instead of this, we can take an arbitrary initial condition, and thereby obtain an arbitrary *random Percolation system.* Lines 4-6 describe the *deterministic Percolation system* and line 7 adds the one-way random noise.

5.1.1. Exercise. Prove that any Percolation system is monotonic.

Note that the measure concentrated in *'all ones'* is an invariant process for any Percolation system. Therefore, ergodicity of a Percolation system amounts to a tendency to *'all ones'* from any initial state.

5.1.2. Exercise. Let P stand for the *Percolation operator*, i.e. operator of the Percolation system. Prove that any Percolation system is ergodic iff P^n(*'all zeroes'*) tends to *'all ones'* when $n \to \infty$.

Let us discuss ergodicity of a Percolation system $\mathcal{P}(\theta)$ as depending on θ with a given *Space*, m, and neighborhoods. We are interested in *critical* values θ^* which separate ergodicity from non-ergodicity of $\mathcal{P}(\theta)$. If a Percolation system is ergodic for all positive θ, we say that $\theta^* = 0$ is its only critical value.

5.1.3. Exercise. Prove that any finite Percolation system is ergodic for all $\theta > 0$.

5.1.4. Exercise. Prove that any standard Percolation system in which all the neighbor vectors are collinear is ergodic for all $\theta > 0$.

5.1.5. Exercise. Prove that $\theta_1 < \theta_2 \Longrightarrow \mathcal{P}(\theta_1) \preceq \mathcal{P}(\theta_2)$.

5.1.6. Exercise. Prove that any Percolation system $\mathcal{P}(\theta)$ has only one critical value.

5.1.7. Exercise. Let two Percolation processes $proc_1(\theta)$ and $proc_2(\theta)$ be defined with the same *Space*, m, and initial condition, and let $N_1(\mathcal{O}) \subset N_2(\mathcal{O})$. Prove that $\forall\, \theta\ :\ proc_1(\theta) \succeq proc_2(\theta)$ and that $\theta_1^* \leq \theta_2^*$.

5.1.8. Problem. Prove that a standard infinite Percolation system has $\theta^* > 0$ iff it has at least two non-collinear neighbor vectors.

Comment: Due to the preceeding exercise, it is sufficient to prove $\theta^* > 0$ for 2-Percolation (see below) which is the simplest Percolation process.

5.1.9. Unsolved problem. It is conjectured that Percolation systems are ergodic at the critical value of θ whenever it is positive. Is this true ?

Comment: Some infinite Percolation systems, including 2-Percolation, are proved to be ergodic at the critical point, see [8, 9, 12, 13, 22].

5.2. 2-Percolation

The 2-Percolation system has $d = 1$, $m = 1$, $n = 2$ and $N(0, 0) = \{(-1, -1), (0, -1)\}$. The following pseudo-code models 2-Percolation on Z_S:

```
for all s ∈ Z_S do
        x_{s,-1} ← 0
for t = 0 to t_max − 1 do
        for all s ∈ Z_S do
                x_{s,t} ← min(x_{s-1,t-1}, x_{s,t-1})
                if rnd < θ then x_{s,t} ← 1
```

(17)

This system, in its finite and infinite versions, was first introduced in [29] and received much attention. A proof of its non-ergodicity for small θ by the well-known contour method can be found in [30] or [2].

5.3. Different Rates of Convergence of Finite Percolation Systems

For Percolation systems the following special way is used to speak about how fast or slowly they converge: Let $T(\theta, S)$ stand for the mean time when a given standard finite Percolation system with $Space = Z_S$ and a given θ first reaches the state *'all ones'* , if it started from the state *'all zeroes'* .

5.3.1. Exercise. Prove that if θ is large enough (say $\theta > 1 - 1/n$), then the Percolation pattern converges fast and $T(\theta, S)$ grows logarithmically as a function of S.

5.3.2. Problem. Prove that if θ is small enough, then the Percolation pattern converges slowly and $T(\theta, S)$ grows exponentially as a function of S.

Comment: See [30] where it is proved for 2-Percolation. For other Percolation patterns it follows from monotony.

5.3.3. Unsolved problem. For any Percolation pattern the following critical values may be defined:

$\theta^*_\infty = \theta^*$	- the boundary between ergodicity and non-ergodicity of the infinite system,
θ^*_{fast}	- infimum of those values of θ with which finite systems converge fast,
θ^*_{slow}	- supremum of those values of θ with which finite systems converge slowly,
$\theta^*_{\log}$	- infimum of those values of θ for which $M(\theta, S)$ grows logarithmically,
$\theta^*_{\exp}$	- supremum of those values of θ for which $M(\theta, S)$ grows exponentially.

It is conjectured that for any Percolation pattern all of these critical values are equal. Is it true?

Comment: This conjecture is not proved (or disproved) even for 2-Percolation. However, some relations between these critical values can be stated:

5.3.4. Problem. Prove for all Percolation patterns:

$$\theta^*_{\text{slow}} \le \left\{ \begin{array}{c} \theta^*_{\exp} \\ \theta^*_\infty \end{array} \right. \le \theta^*_{\log} \le \theta^*_{\text{fast}}.$$

Comment: See [2], p. 72-84 where several critical values and relations between them are discussed.

6. NON-ERGODICITY AND SLOW CONVERGENCE

Theorem 2 has shown that any 'random enough' system is ergodic. Thus, to be non-ergodic, a system has to be 'deterministic enough'. We shall prove non-ergodicity of some systems obtained by 'spoiling' a deterministic system with a small random noise. Naturally, properties of the deterministic system are essential with such approach, and this is what we start with.

6.1. Attractors and Eroders

Let X_0 have an element called 0. Let an *island* stand for a finite perturbation of *'all zeroes'* . Call a deterministic system an *eroder* if for any initial island x_I the corresponding trajectory $\text{TR}(x_I)$ is also an island.

6.1.1. Exercise. Prove that in any eroder TR(*'all zeroes'*) = *'all zeroes'*.

We concentrated our attention on *'all zeroes'* for simplicity; generalization is easy. Say that an initial configuration s_I is an *attractor* for a deterministic system, or *attracts* it, if for any finite perturbation s'_I of s_I the trajectory

$\mathrm{TR}(s'_I)$ is a finite perturbation of $\mathrm{TR}(s_I)$. For example, a system is an eroder iff it is attracted by the initial configuration *'all zeroes'* . Note that according to our definition, any finite perturbation of an attractor is an attractor too. So it is better to speak of *bunches* of attractors, a bunch being a class of equivalence, with two initial conditions equivalent iff their set of difference is finite. Let us call a bunch an *attractor* iff all of its elements are attractors.

6.1.2. Exercise. Prove that if a bunch contains an attractor, it is an attractor.

6.1.3. Exercise. Prove that in any deterministic Percolation system the only attractors are those in the bunch of *'all zeroes'* .

6.1.4. Exercise. Can a finite deterministic system have more that one bunch of attractors ?

6.1.5. Unsolved problem. Can an infinite standard monotonic deterministic system with $X_0 = \{0, 1\}$ have a non-periodic fixed attractor ?

Comment: The answer is unknown even in the one-dimensional case with $m = 1$.

Our interest in attractors is motivated by the following consideration. It seems plausible (and sometimes it is true) that if an initial configuration is an attractor, the system remains in the vicinity of the resulting trajectory even in the presence of a small random noise. Hence it is a good idea to construct systems which are attracted by more than one bunch, and to see how they behave with a small random noise added. In the next section we give exact definitions.

6.2. Stable trajectories

As before, $\mathcal{M} = \mathcal{M}_V$ stands for the set of normed measures on $X = X_0^V$. To every initial configuration s there corresponds $\mathcal{M}^s \subset \mathcal{M}$, which consists of those measures whose projection to I is concentrated in s. To every value of the parameter $\varepsilon \in [0, 1]$ there corresponds $\mathcal{M}_\varepsilon \subset \mathcal{M}$ defined as follows: a measure μ belongs to $\mathcal{M}_\varepsilon$ if

$$\text{for all finite } W \subset V \ : \ \mu(\forall\, w \in W \ : \ x_w \neq \mathrm{tr}(x_{N(w)})) \leq \varepsilon^{|W|},$$

where $|W|$ is the cardinality of W. Finally, $\mathcal{M}_\varepsilon^s = \mathcal{M}_\varepsilon \cap \mathcal{M}^s$. An initial condition s and the resulting trajectory $\mathrm{TR}(s)$ are termed *stable* if

$$\lim_{\varepsilon \to \infty} \sup\{\mu(x_w \neq \mathrm{TR}(s)_w) : \mu \in M_\varepsilon^s, \ w \in V\} = 0.$$

This definition always makes sense: the supremum makes sense because the set $\mathcal{M}_\varepsilon^s$ is non-empty, as it contains the measure concentrated in $\mathrm{TR}(s)$, and the limit makes sense because the set $\mathcal{M}_\varepsilon^s$ decreases when ε decreases.

6.2.1. Problem. Prove that a finite system can not have more than one stable initial configuration.

To show that our definition is non-trivial, i.e. the set $\mathcal{M}_\varepsilon^s$ is rich, we propose the following exercise which constructs explicitly a multi-parametric subset of $\mathcal{M}_\varepsilon^s$.

6.2.2. Exercise. Let us have a parameter $\varepsilon_w \leq \varepsilon$ for every non-initial point w and let the hidden measure η induce a measure $\mu \in \mathcal{M}$ with the following map. For all initial w we set $x_w = s_w$ and for all non-initial w the value of x_w is defined inductively as follows: $x_w = \mathrm{tr}(x_{N(w)})$ if $\eta_w \leq 1 - \varepsilon_w$ and is assigned an arbitrary value otherwise. Prove that any measure μ defined in this way belongs to $\mathcal{M}_\varepsilon^s$. Show also that measures of this sort do not exhaust $\mathcal{M}_\varepsilon^s$.

6.2.3. Unsolved problem. Consider all infinite standard deterministic monotonic systems with $X_0 = \{0, 1, \dots, k\}$ naturally ordered. Formulate checkable equivalent criteria for their initial condition *'all zeroes'* : 1) to be an attractor, 2) to be stable.

Comment: The answer is unknown now even for the case with $m = 1$, $d = 2$ and $|X_0| = 3$. Criteria for arbitrary d, $X_0 = \{0, 1\}$, and $m = 1$ were given in [32]. The same criteria generalized for any m were given in Theorem 6 of [35] and are repeated here as Theorem 3. Criteria for $d = 1$ and arbitrary k and m were given in [20]. The monotony condition is essential; it is known that without demanding monotony, the problem of discerning eroders is algorithmically unsolvable even with $m = k = 1$ [27]. Some similar proofs of unsolvability are given in [33].

6.2.4. Exercise. Prove that if the initial condition *'all zeroes'* is stable, then it is an attractor under the conditions of problem 6.2.3. Prove also that the opposite is not true, and therefore the items 1) and 2) of the problem 6.2.3 are not equivalent to each other.

The next problem highlights the difficulty of problem 6.2.3. Let us call: *diameter* of an island - the greatest Euclidean distance between its non-zero components; *lifetime* of an initial condition y - the greatest time coordinate of non-zero components of $\mathrm{TR}(y)$. Of course, a system is an eroder iff lifetime of any initial island is finite. Finally, let us define a function $T(D)$ - the greatest lifetime of initial islands whose diameters do not exceed D.

6.2.5. Problem. Assume conditions of problem 6.2.3. Then:

1. For any system with $k = 1$ there is such $C = const$ that $T(D) \leq C \cdot (D + 1)$.
2. There is a system with $k = 2$ for which $T(D) \geq C^D - const$ where $C = const > 1$.

6.3. A theorem about attraction and stability

The following theorem treats standard deterministic systems with $Space = Z^d$. In this case the volume $V = Space \times Time$ is a half of the $d+1$-dimensional integer space. Let us consider V as a subset of the continuous space R^{d+1}. Let $conv(S)$ stand for the convex hull of any set S in R^{d+1}. For any set $S \subset R^{d+1}$ and any number k we define:

$$kS = \{ks : s \in S\}, \quad ray(S) = \bigcup \{kS : \ k \geq 0\}.$$

We call $ray(S)$ *ray* of S. Let a *zero-set* stand for any subset $z \subseteq N(\mathcal{O})$ for which

$$(\forall \ w \in z : x_w = 0) \Longrightarrow \mathrm{tr}(x_{N(\mathcal{O})}) = 0.$$

Let $\sigma(\mathcal{D})$ stand for the intersection of rays of convex hulls (in the continuous space) of all the zero-sets of $\mathcal{D}$.

Theorem 3. Let $\mathcal{D}$ be any monotonic standard deterministic system with $X_0 = \{0,1\}$ and $Space = Z^d$. The following four conditions are mutually equivalent:

1) $\mathcal{D}$ is an eroder.

2) $\sigma(\mathcal{D}) = \{\mathcal{O}\}$.

3) There exist homogeneous linear functions $L_1, \ldots, L_r : R^{d+1} \mapsto R$ (where $r \leq \nu + 2$) such that $L_1 + \cdots + L_r \equiv 0$, and for every $i = 1, \ldots, r$, the set $\{w \in N(\mathcal{O}) : L_i(w) \geq 1\}$ is a zero-set.

4) The initial condition *'all zeroes'* is stable.

From now on $r(\mathcal{D})$ stands for the minimal value of r for which condition 3) holds for an eroder $\mathcal{D}$.

6.3.1. Exercise. Prove that 2-Percolation is an eroder and find $L_1, \ldots, L_r$ with which condition 3) is fulfilled.

Answer: With $r = 2,\ L_1(s,t) = 2s - t,\ L_2(s,t) = -2s + t.$

We use Theorem 3 to prove non-ergodicity (in fact non-uniqueness of the invariant measure), in the following way.

6.3.2. Exercise. For any k construct a deterministic system $\mathcal{D}$ which has at least k different fixed attractors, for which there is $\varepsilon > 0$ such that any random system resulting from $\mathcal{D}$ by adding a random noise, all of whose parameters ε_i are less than or equal to ε, has at least k different invariant processes.

Comment: See [35].

6.4. About the proof of Theorem 3

The most difficult part to prove is that 1), 2), or 3) implies 4). This is non-trivial even for particular examples. For 2-Percolation this proof boils down to examination of a directed planar percolation which has been described in [30] or survey [2]. The case $r = 2$ (which is the smallest possible value of r) is more difficult than 2-Percolation, but still is essentially simpler than the general case. In this case 'contours' are used also, but there is no percolation interpretation. For any system with $r > 2$ there is a combinatorial general proof, but the topological objects it involves are more complicated than contours, because they ramify. The most general versions of this proof are in [35] and [36]. It seems that the best way to understand their main idea is to read [25] where the proof is reworded for the NEC-Voting (for which $r = 3$). There is another method [14] to obtain similar results, but we do not discuss it here.

Here we prove only one part of Theorem 3, namely that 2) implies 3), to show how this proof uses the theory of convex sets. Assume that $\sigma(D) = \{\mathcal{O}\}$. For any finite set $S \subset R^{d+1}$ the set $ray(conv(S))$ can be represented as an intersection of several halfspaces. (A halfspace is a set in R^{d+1}, where some homogeneous linear function is non-negative.) Apply this to zero-sets, and you have a finite list of zero-halfspaces whose intersection is $\{\mathcal{O}\}$. (A zero-halfspace is a halfspace whose intersection with $N(\mathcal{O})$ is a zero-set.) For everyone of these zero-halfspaces we introduce a linear function f_i which is non-positive on it and only on it. We know that the origin is the only point in R^{d+1} where all f_i are non-positive. This allows us to apply Theorem 21.3 of [6] (a variant of Helly's theorem) to the hyperplane $\Pi = \{(s,t) : t = -1\}$, restrictions $f_i|_\Pi$ of our functions to Π and any non-empty closed convex set $C \subseteq \Pi$. We take $C = \Pi$. Then there exist such non-negative real numbers λ_i, at most $d+1$ of which are positive, that for some $\varepsilon > 0$.

$$\forall\, w \in C : \sum_i \lambda_i f_i(w) \geq \varepsilon.$$

Since the left part is linear and bounded from below by a positive constant on C, it has to be a positive constant on C:

$$\forall\, w \in C : \sum_i \lambda_i f_i(w) = \delta = \text{const} \geq \varepsilon.$$

Hence

$$\forall\, w \in R^{d+1} : \sum_i \lambda_i f_i(w) = -\delta t.$$

Then the functions $L_i = -(t + n\lambda_i f_i/\delta)$ fit our claim. □

6.4.1. Exercise. Prove that $\sigma(D)$ equals the intersection of convex hulls of rays of *minimal* zero-sets. We call a zero-set *minimal* iff any its proper subset is not a zero-set. This fact simplifies checking of condition 2) for particular systems.

6.4.2. Exercise. Reformulate Theorem 3 as a criterion for attraction and stability of *'all ones'* .

6.4.3. Exercise. Simplify conditions 2) and 3) of Theorem 3 for the case $m = 1$.

6.4.4. Problem. Prove that item 4) implies 1), 2), and 3).

6.4.5. Problem. Prove that items 1), 2), and 3) are equivalent.

6.4.6. Problem. Formulate and prove an analogue of Theorem 3 for all standard monotonic deterministic systems with $X_0 = \{-1,\ 0,\ 1\}$. In other words, prove that in all of these systems the initial state *'all zeroes'* is stable iff it is an attractor, and give a criterion for that.

6.4.7. Problem. Generalize Theorem 3 to any d-dimensional commutative system with $X_0 = \{0, 1\}$. See [37].

6.5. Example: Flattening

This example illustrates an application of Theorem 3. It is a deterministic system with $d = 2,\ m = 1$ with ε-δ noise added. The following pseudo-code shows neighborhoods and transition probabilities of this system:

```
for all (i, j) ∈ Z_S^2 do
        x_{i,j,-1} ← x_{i,j}^{initial}
for t = 0 to t_{max} - 1 do
for all (i, j) ∈ Z_S^2 do
        x_{i,j,t} ← max(min(x_{i,j,t-1}, x_{i,j-1,t-1}),  min(x_{i-1,j,t-1}, x_{i-1,j-1,t-1}))
                if rnd < ε then x_{i,j,t} ← 0 else if rnd > δ then x_{i,j,t} ← 1
```

(18)

The infinite Flattening has the same pattern and transition probabilities and is an eroder. The condition 3) is fulfilled with $r = 2,\ L_1(i,j,t) = 2i - t,\ L_2(i,j,t) = -2i + t$.

Theorem 3 guarantees that for ε and δ small enough the infinite version of this system is non-ergodic. More specifically, Theorem 3 proves the following:

$$\forall\ (i,j) \in Z^2 : x_{i,j}^{initial} = 0 \Longrightarrow \sup_{t,\varepsilon} proc(x_{i,j,t} = 1) \to 0 \text{ when } \delta \to 0.$$

$$\forall\ (i,j) \in Z^2 : x_{i,j}^{initial} = 1 \Longrightarrow \sup_{t,\delta} proc(x_{i,j,t} = 0) \to 0 \text{ when } \varepsilon \to 0.$$

(Both of these probabilities do not actually depend on i or j.)

6.5.1. Exercise. For which values of ε and δ does Theorem 2 prove fast convergence of the finite version of Flattening ?

6.5.2. Exercise. Prove that in the Flattening systems both bunches of *'all zeroes'* and *'all ones'* are attractors.

6.5.3. Problem. Prove that although the deterministic Flattening has infinitely many fixed initial conditions, it has only two bunches of attractors: those of *'all zeroes'* and *'all ones'* .

6.5.4. Problem. Prove that the finite version of Flattening converges slowly with small enough, but positive, ε and δ.

Comment: This can be proved by a method analogous to that of [11]. The general proof is in [38].

6.5.5. Unsolved problem. Does Flattening have any attractors besides bunches of *'all zeroes'* and *'all ones'* ?

6.6. Quasi-Stability

In the section 3.4 we have defined slow convergence as a finite analogue of non-ergodicity. In the same vein we can define quasi-stability as a finite analogue of stability. Take some $a \in A$. Take a pattern in which 'all a-s' is a trajectory and call this trajectory *quasi-stable* for the given pattern if there exist $\varepsilon > 0$ and $Q > 1$ such that for the processes resulting from the initial condition 'all a-s',

$$\forall\, S, t \;:\; t \leq Q^S \Longrightarrow \mathrm{Prob}(\mathrm{x_{s,t}} \neq \mathrm{a}) \leq \varepsilon.$$

6.6.1. Exercise. Prove that if an ergodic system has two non-equivalent quasi-stable trajectories, it converges slowly.

6.6.2. Problem. Prove that *'all zeroes'* is quasi-stable for all patterns for which the conditions of Theorem 3 hold.

Comment: [11] proved this for the NEC-Voting. The general proof is in [38].

7. STANDARD VOTINGS

For any odd n, the Boolean function $Voting$ with n arguments equals 1 iff the majority of its arguments are ones. Equivalently, $Voting$ equals 0 iff the majority of its arguments are zeros. Let us consider the following systems which use $Votings$: a deterministic Voting is a deterministic system which has $Voting$ as its transition function, and a random Voting results from deterministic Voting by adding the ε-δ noise. The first computer simulations of random Votings were done in [41] for the symmetric case $\varepsilon = \delta$.

7.1. Symmetric Votings

Call a standard Voting *symmetric* if $N(\mathcal{O})$ is symmetric with respect to $ray(\bar{0}, -1)$.

7.1.1. Exercise. Prove that all deterministic symmetric Votings are not eroders and have $\sigma = ray(\bar{0}, -1)$.

7.1.2. Exercise. Prove that deterministic symmetric Votings have no fixed attractors.

7.1.3. Problem. Prove that all random symmetric Votings are ergodic if $\varepsilon = 0$ and $\delta > 0$ (or $\varepsilon > 0$ and $\delta = 0$).

Comment: A similar statement is proved as Proposition 1 in [34].

7.1.4. Exercise. For which values of ε and δ does Theorem 2 prove ergodicity of random Symmetrical Votings ?

7.1.5. Unsolved problem. Prove that all one-dimensional Votings with a non-degenerate noise are ergodic.

Comment: Computer simulations [41] of one-dimensional symmetric Votings with

$$N(\mathcal{O}) = \{(s, -1) \ : \ s = -1, \ 1, \ 0\} \tag{19}$$

showed ergodicity in any non-degenerate case. It is very plausible that in fact all non-degenerate one-dimensional Votings are ergodic. However, a rigorous proof [21] exists now only for the continuous-time analogue of (19).

Let us call a Symmetrical Voting a *Windrose* if its neighbor vectors are non-coplanar.

7.1.6. Unsolved problem. Prove that all Windroses are non-ergodic with small enough $\varepsilon = \delta$.

7.1.7. Unsolved problem. Prove ergodicity of some Windrose with some positive values of ε and δ.

Comment: Both of the last problems are suggested by computer simulations [41, 10]. However, both are unproved even for particular cases, including the following simplest two-dimensional ones:

5-Windrose has $n = 5$ and
$N(\mathcal{O}) = \{(s, -1) \ : \ s = (0,0), (-1,0), (1,0), (0,-1), (0,1)\}$.

9-Windrose has $n = 9$ and
$N(\mathcal{O}) = \{(i, j, -1) \ : \ i, j = -1, 0, 1\}$.

7.2. NEC and other Non-Symmetric Votings

One of non-symmetric votings, the NEC-Voting, has attracted special attention. Here NEC means North, East, Center. It has $N(\mathcal{O}) = \{(s,-1)\ :\ s = (0,0),(1,0),(0,1)\}$. NEC-Voting was first introduced in [41] with a symmetric noise, and the results of computer simulation showed that it is non-ergodic if $\varepsilon = \delta$ is small enough. Now this non-ergodicity is proved for any small enough ε and δ; it follows from our Theorem 3 because NEC-Voting is an eroder. Condition 3) is fulfilled with

$$r = 3, \quad L_1(i,j,t) = -3i - t, \quad L_2(i,j,t) = -3j - t, \quad L_3(i,j,t) = 3i + 3j + 2t.$$

7.2.1. Exercise. Prove that in the deterministic NEC-Voting both bunches of *'all zeroes'* and *'all ones'* are attractors.

7.2.2. Problem. Prove that although deterministic NEC-Voting has infinitely many fixed initial conditions, it has only two bunches of attractors, those of *'all zeroes'* and *'all ones'* .

7.2.3. Unsolved problem. Does NEC-Voting have any attractors besides bunches of *'all zeroes'* and *'all ones'* ?

7.2.4. Problem. Consider any Voting with $m = 1$ and $n = 3$ whose neighbor vectors are non-coplanar. Prove that it is ergodic or non-ergodic whenever the NEC-Voting with the same values of ε and δ is.

7.2.5. Problem. Use NEC-Voting to construct deterministic systems with at least m stable fixed trajectories for any natural m. In more detail: For any finite set of periodic configurations in $Space = Z^d$, propose a system for which all of the corresponding fixed configurations are stable trajectories. (You can even propose a monotonic system with this property.) See [35].

7.2.6. Problem. Take any standard Voting in which all the neighbor vectors are pairwise non-collinear. Prove that it is not an eroder iff there exists such a neighbor vector V that all the other neighbor vectors form pairs, the two vectors of each pair having the opposite V-direction projections to *Space*.

Comment: Importance of non-symmetry in a similar context was discussed in [28].

8. ONE-DIMENSIONAL CONSERVATORS

Let a *forget-me-not* stand for a non-ergodic non-degenerate process. For $d > 1$ forget-me-nots certainly exist, Flattening and NEC-Voting for example. The question whether one-dimensional forget-me-nots exist, was very intriguing for years. The *positive rates conjecture*, proposed by several authors, claimed that all non-degenerate one-dimensional random systems are ergodic (see for example [5], Chapter 4, section 3). Now this seems to be refuted: after some preliminary work [43, 24], Péter Gács proposed an elaborate construction [18], which presents a one-dimensional forget-me-not, but his construction is not yet examined sufficiently, and we shall not discuss it here.

Our purpose is more modest. Note that from the practical point of view, it is not always necessary to remember forever; it may be sufficient to keep information for a finite but long time. So it seems worthwhile to look for deterministic systems which converge slowly in presence of a small noise. With this idea in mind, the term 'conservator' and first examples of conservators were proposed in [19]. A *conservator* is a deterministic system with at least two different fixed initial attractors. The following is one of the simplest conservators.

8.1. Soldiers

This is a standard system in which

$$N(\mathcal{O}) = \{(s, -1) \ : \ s = -3, -1, \ 0, \ 1, \ 3\}.$$

Let us call every point of Z a *a soldier*. Every soldier has only two possible states, -1 and 1. The transition function equals

$$\mathrm{tr}(a_{-3}, a_{-1}, a_0, a_1, a_3) = \begin{cases} -1 & \text{if } a_0 = \ \ 1, \ \ a_1 \ = a_3 \ \ = -1, \\ \ \ 1 & \text{if } a_0 = -1, \ a_{-1} = a_{-3} = \ \ 1, \\ a_0 & \text{otherwise.} \end{cases}$$

8.1.1. Exercise. Prove that the Soldiers system is non-monotonic, whether we assume $-1 \prec 1$ or $-1 \succ 1$.

8.1.2. Exercise. Prove that the Soldiers system is symmetric in the following sense. If we define a one-to-one map $a \ : \ X \mapsto X$ by the rule $(a(x))_{s,t} = -x_{-s,t}$, then $\forall \, x_I \ : \ a(\mathrm{TR}(x_I)) = \mathrm{TR}(a(x_I))$.

8.1.3. Exercise. Prove that if Soldiers has only one invariant process μ_{inv} then

$$\mu_{inv}(x_i = -1) = \mu_{inv}(x_i = 1) = 1/2.$$

8.1.4. Problem. Prove that the Soldiers system is a conservator. In more detail: prove that both *'all ones'* and *'all minus ones'* are attractors. (From symmetry these two facts are equivalent.)

Comment: This was first proved in [19] and reinforced in [17].

8.1.5. Unsolved problem. Is it true that *'all ones'* and *'all minus ones'* are the only non-equivalent attractors for Soldiers system?

Comment: [17] proved that *'all ones'* and *'all minus ones'* are the only fixed attractors of Soldiers. See also [26].

8.1.6. Unsolved problem. Prove ergodicity of Soldiers with the one-way ε-δ noise in which $\varepsilon = 0$ or $\delta = 0$.

8.2. 2-Line Voting

This is a non-standard Voting with 3 neighbors. Here $Space = Z \times \{1, -1\}$, i.e. automata form two parallel rows and are indexed (i, j), where $i \in Z$ and $j \in \{1, -1\}$, and

$$N(i, j, t) = \{(s, t-1) \ : \ s = (i - 2j, j), (i - j, j), (i, -j)\}.$$

8.2.1. Exercise. Show that 2-Line Voting is uniform, but non-commutative.

8.2.2. Exercise. Prove that both *'all zeroes'* and *'all ones'* are attractors for 2-Line Voting.

8.2.3. Problem. Prove that 2-Line Voting is ergodic in the presence of a one-way noise, which turns zeros into ones with probability ε and never turns ones into zeros (or vice versa).

Comment: This can be proved by the method of [34]. Due to monotonicity, it it sufficient to take only *'all zeroes'* and *'all ones'* as initial states, because all the others are between them.

8.2.4. Problem. Has 2-Line Voting any other fixed attractors besides the bunches of *'all zeroes'* and *'all ones'* ?

Answer: No. To prove it, note that there are only six fixed states: 1) *'all zeroes'* ; 2) *'all ones'* ; 3) zeroes if $j = 1$, ones if $j = -1$; 4) zeroes if $j = -1$, ones if $j = 1$; 5) zeroes if i is even, ones if i is odd; 6) zeroes if i is odd, ones if i is even.

8.2.5. Unsolved problem. Has 2-Line Voting any other attractors besides the bunches of *'all zeroes'* and *'all ones'* ?

8.3. Relaxation Time

The first results of computer simulation of Soldiers (and similar systems) with a random noise were reported in [19]. If the simulation of Soldiers began in *'all ones'* (or in *'all minus ones'*), the system remained in the vicinity of this state for a long time (in fact, all the computer time which was available to the authors). Moreover, the system approached one of these states if started from a chaotic initial condition. A recent computer simulation [17] of Soldiers shows ergodicity for all non-degenerate cases (which indeed is very plausible), but it suggests some non-trivial dependence of the rate of convergence on ε.

In more detail: [17] defines the 'relaxation time' as the mean time when the percentage of ones first gets into the range $(S/2 - \sqrt{S},\ S/2 + \sqrt{S})$ if we started from the state *'all minus ones'* . This definition is convenient for all systems with $X_0 = \{-1, 1\}$ and the kind of symmetry which was shown for Soldiers in exercise 8.1.2. The computer results of [17] seem to show that the relaxation time for Soldiers does not depend on S (which corresponds to fast convergence in our sense), but is asymptotic to $\sim e^{const/\varepsilon}$ as $\varepsilon \to 0$, which is enormously greater than that for $\mathcal{D}_\equiv$ (see the next exercise). This means that the Soldiers system, although ergodic, effectively checks dissent when ε is small enough.

8.3.1. Exercise. Let $\mathcal{D}_\equiv$ stand for the identity standard deterministic system with $\mathit{Space} = Z$, in which automata never change their states. Show that the relaxation time for $\mathcal{D}_\equiv$ with an ε-δ noise is $O(1/(\varepsilon+\delta))$ when ε and δ tend to 0.

8.3.2. Unsolved problem. How do relaxation times of various Votings, especially of 2-Line Voting, behave as functions of ε when $\varepsilon \to 0$?

Conjecture: Relaxation times of one-dimensional votings with $N(\mathcal{O}) = \{(i,-1),\ i \in [-R,\ R]\}$ behave as $O((\varepsilon+\delta)^{-(R+1)})$ and the relaxation time of 2-Line Voting behaves like that of Soldiers.

8.4. Further Problems

8.4.1. Problem. Prove that there are no commutative one-dimensional monotonic conservators with $|X_0| = 2$.

Sketch of Proof: Assume the contrary and come to a contradiction. Formula (2) here becomes

$$\mathit{Space} = Z \times S = \{(i,j)\ :\ i \in Z,\ j \in S\}.$$

Define two initial configurations:

$$L(i,j) = \begin{cases} 0, & \text{if } i < k \\ 1, & \text{otherwise} \end{cases} \quad \text{and} \quad R(i,j) = \begin{cases} 1, & \text{if } i < k \\ 0, & \text{otherwise.} \end{cases}$$

Since components of L and R do not depend on j, components of $\mathrm{TR}(L)$ and $\mathrm{TR}(R)$ also do not depend on j (exercise 1.4.3). Then prove that

$$\min\{i \ : \ \mathrm{TR}(L)_{i,j,t} = 1, \ t = T\} = V_{left} \cdot T + O(1),$$

$$\min\{i \ : \ \mathrm{TR}(R)_{i,j,t} = 0, \ t = T\} = V_{right} \cdot T + O(1).$$

Here V_{left} and V_{right} are constants which may be called velocities of movement of the boundaries between zeroes and ones (the technique of velocities was developed in [20]). The following three statements complete the proof:

1) If $V_{left} < V_{right}$ then the bunch of *'all ones'* is the only attractor.

2) If $V_{left} > V_{right}$ then the bunch of *'all zeroes'* is the only attractor.

3) If $V_{left} = V_{right}$ then there are no attractors. □

Note that we did more than the problem asked; we described completely all attractors rather than only fixed ones.

Thus, to have a conservator we need either a greater-than-one dimension, or non-monotonicity, or non-commutativity, or $|X_0| > 2$. In fact, in each case there is a conservator. For greater dimensions, it is Flattening or NEC-Voting. For non-monotonity it is the Soldiers system. For non-commutativity it is the 2-Line Voting. An example for $|X_0| > 2$ is the subject of the following problem.

8.4.2. Problem. Propose a one-dimensional standard monotonic deterministic system with $|X_0| = 3$ which is a conservator.

Comment: You can take $X_0 = \{-1, \ 0, \ 1\}$ and arrange an interaction such that both *'all ones'* and *'all minus ones'* will be attractors. You can also make this system symmetric in a sense similar to Soldiers.

9. CHAOS APPROXIMATION

Let $m = 1$ and let $\mathcal{M}^1$ be the set of normed measured on X_0^{Space}. Define a map $Chaos : \mathcal{M}^1 \mapsto \mathcal{M}^1$ as follows. Any $\mu \in \mathcal{M}^1$ given, $Chaos(\mu)$ is that product-measure, whose projections to all $X_s, \ s \in Space$ coincide with those of μ. Thus, $Chaos(\mathcal{P})$, that is $Chaos$ applied after $\mathcal{P}$, is the new operator which hopefully approximates $\mathcal{P}$. The chaos approximation is an exact solution for the Tree system which we defined when answering the exercise 1.1.1. We shall illustrate the chaos approximation by two examples.

9.1. Percolations

Take any Percolation operator $\mathcal{P}(\theta)$ with $m = 1$ and take the measure concentrated in *'all zeroes'* as the initial condition. Let $P_t(\theta)$ and $\tilde{P}_t(\theta)$ stand for the percentage of ones at the t-th step in the measures $\mathcal{P}^t$(*'all zeroes'*) and

$(Chaos(\mathcal{P}))^t$(*'all zeroes'*). The sequence of $\tilde{P}_t(\theta)$ satisfies the simple iteration formula

$$\tilde{P}_{t+1}(\theta) = \theta + (1-\theta)\tilde{P}_t^n(\theta).$$

This allows us to examine the behavior of $\tilde{P}_\infty(\theta) = \lim_{t\to\infty} \tilde{P}_t(\theta)$ and to compare it with the behavior of $P_\infty(\theta) = \lim_{t\to\infty} P_t(\theta)$.

9.1.1. Exercise. Prove that when θ grows from 0 to 1, $\tilde{P}_\infty(\theta)$ starts at 0 as $\theta + o(\theta)$, grows monotonically and becomes equal to 1 at the critical value $\theta^*_{chaos} = 1 - \frac{1}{n}$.

This exercise shows that, crude as it is, the chaos approximation's behavior may be similar to that of the original process.

9.1.2. Unsolved problem. Let $\mathcal{P}$ be a Percolation operator. Let us apply *Chaos* after every T steps of $\mathcal{P}$. Does the critical point of $Chaos(\mathcal{P}^T)$ tend to that of $\mathcal{P}$ when $T \to \infty$?

9.2. Votings

It is instructive to examine the chaos approximation for random Votings. As before, $\tilde{P}_t$ stands for the percentage of ones at t-step. The iteration formula for $n = 3$ is

$$\tilde{P}_{t+1} = \varepsilon + (1-\varepsilon-\delta)(\tilde{P}_t^3 + 3\tilde{P}_t^2(1-\tilde{P}_t)).$$

Let us compare two sequences generated by this formula with two different initial conditions $\tilde{P}_0 = 0$ and $\tilde{P}_0 = 1$. We must discriminate the case when the limits of these sequences are equal, from the case when they are different.

9.2.1. Exercise. Show that these limits are equal if $\varepsilon + \delta$ is close enough to 1 and are different if ε and δ are small enough.

We see that behavior of the chaos approximation of a random Voting with 3 neighbors is similar to that of NEC-Voting, but different from one-dimensional votings, which seem to be ergodic in all non-degenerate cases.

9.2.2. Exercise. Write and examine the iteration formula for the chaos approximation of a voting with 5 neighbors and compare it with the behavior of the 5-Windrose.

Books and surveys

[1] Cellular Automata, Theory and Experiment. Amsterdam, 1990.

[2] Discrete Local Markov Systems. A. L. Toom, N. B. Vasilyev, O. N. Stavskaya, L. G. Mityushin, G. L. Kurdyumov and S. A. Pirogov. *Stochastic Cellular Systems: ergodicity, memory, morphogenesis.* Ed. by R. Dobrushin, V. Kryukov and A. Toom. Nonlinear Science: theory and applications, Manchester University Press, 1990.

[3] Richard Durrett. Lecture Notes on Particle Systems and Percolation. Wadsworth & Brooks/Cole Statistics/Probability Series, 1988.

[4] David Griffeath. Coupling Methods for Markov Processes. *Studies in Probability and Ergodic Theory; Advances in Mathematics, Supplementary Studies*, Vol. 2, pp. 1-43. Academic Press, New York, 1978.

[5] Thomas M. Liggett. Interacting Particle Systems. N.Y., Springer-Verlag, 1985.

[6] R. T. Rockafellar. Convex Analysis. Princeton Univ. Press, 1970.

[7] Theory and Application of Cellular Automata (including selected papers 1983-1986). Stephen Wolfram a.o. *Advanced series on complex systems - v. 1.* World Scientist, 1986.

Papers

[8] David J. Barsky, Geoffrey Grimmett and Charles M. Newman. Dynamic Renormalization and Continuity of the Percolation Transition in Ortants. *Spatial Stochastic Processes*, K. Alexander and J. Watkins, Eds., Birkhauser, Boston, 1991, p. 37-55.

[9] David J. Barsky, Geoffrey Grimmett and Charles M. Newman. Percolation in Half-Spaces: Equality of Critical Densities and Continuity of the Percolation Probability. *Prob. Theory and Related Fields,* 90 (1991), 1, 111-148.

[10] Charles H. Bennett and G. Grinstein. Role of Irreversibility in Stabilizing Complex and Nonergodic Behavior in Locally Interacting Discrete Systems. *Phys. Rev. Letters*, v. 55 (1985), n. 7, pp. 657-660.

[11] Piotr Berman and Janos Simon. Investigations of Fault-Tolerant Networks of Computers *ACM Symp. on Theory of Computing*, 20 (1988), 66-77.

[12] Carol Bezuidenhout and Geoffrey Grimmett. The Critical Contact Process Dies Out. *Annals of Probability*, 1990, 18, 4, 1462-82.

[13] Carol Bezuidenhout and Lawrence Gray. Critical Attractive Spin Systems. To appear in *Annals of Probability.*

[14] Maury Bramson and Lawrence Gray. A Useful Renormalization Argument. *Random walks, Brownian motion and interacting particle systems.* Festschrift for F. Spitzer. Ser. Progress in Probability, v. 28. Rick Durrett & Harry Kesten, Eds. Birkhauser, 1991, pp. 113-152.

[15] Maury Bramson and Claudia Neuhauser. Survival of One Dimensional Cellular Automata under Random Perturbations. To appear in *Annals of Probability*, 1994.

[16] Chen, Mu Fa. On coupling of jump processes. *Chinese Ann. Math.* Ser B 12, 1991, n. 4, pp. 385-399.

[17] Paula Gonzaga de Sa and Christian Maes. The Gacs-Kurdyumov-Levin Automaton Revisited. *Journal of Statistical Physics*, v. 67, 1992, p. 507.

[18] Péter Gács. Reliable Computation with Cellular Automata. *Journal of Computer and System Sciences*, V. 32, N. 1 (February 1986), pp. 15-78.

[19] Péter Gács, George Kurdjumov and Leonid Levin. One-Dimensional Homogeneous Media, which Erode Finite Islands. *Problems of Information Transm.*, V. 14 (1978), N. 3, pp. 223-226.

[20] Gregory Galperin. One-Dimensional Local Monotone Operators With Memory. *Soviet Math. Dokl.*, V. 17 (1976), N. 3, pp.688-692.

[21] Lawrence F. Gray. The Positive Rates Problem for Attractive Nearest Neighbor Spin Systems on Z. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, V. 61 (1982), pp. 389-404.

[22] G. Grimmett and J. M. Marstrend. The supercritical phase of percolation is well-behaved. *Proc. of Royal Soc. of London*, ser. A, v. 430, 1990, p. 439-457.

[23] I. G. Kopylov. Some properties of a Markovian coupling. *Moscow Univ. Comput. Math. Cybernet.* 1991, n. 4, p. 58-63.

[24] George Kurdyumov. An example of a nonergodic one-dimensional homogeneous random medium with positive transition probabilities. *Soviet Math. Dokl.*, V. 19 (1978), N. 1, pp. 211-214.

[25] Joel L. Lebowitz, Christian Maes and Eugene R. Speer. Statistical Mechanics of Probabilistic Cellular Automata. *J. Stat. Phys.* 59 (1990), 1-2, 117-68.

[26] Wentian Li. Non-Local Cellular Automata. 1991 Lectures in Complex Systems, SFI Studies in the Sciences of Complexity. Lect. Vol. IV, Eds. L. Nadel & D. Stein, Addison-Wesley, 1992.

[27] N. V. Petri. The unsolubility of the problem of discerning of annuling iterative nets. *Researches in the Theory of Algorithms and Mathematical Logic.* Moscow, Nauka, 1979 (in Russian).

[28] Nicholas Pippenger. Symmetry in Self-Correcting Cellular Automata. To appear in *Journal of Computer and System Science.*

[29] O. N. Stavskaya and I. I. Piatetski-Shapiro. [Uniform networks of spontaneously active elements.] *Problemy Kibernetiki*, V. 20 (1968), pp. 91-106 (in Russian).

[30] Andrei Toom. A family of uniform nets of formal neurons. *Soviet Math. Dokl.*, V. 9 (1968), N. 6, pp. 1338-1341.

[31] Andrei Toom. Non-ergodic multidimensional systems of automata. *Problems of Information Transmission*, v. 10 (1974), pp. 239-246.

[32] Andrei Toom. Monotonic Binary Cellular Automata. *Problems of Information Transmission*, V. 12 (1976), N. 1, pp. 33-37.

[33] Andrei Toom and Leonid Mityushin. Two Results Regarding Non-computability for Univariate Cellular automata. *Problems of Information Transmission*, V. 12 (1976), N. 2, pp. 135-140.

[34] Andrei Toom. Unstable Multicomponent Systems. *Problems of Information Transmission*, V. 12 (1976), N. 3, pp. 220-225.

[35] Andrei Toom. Stable and Attractive Trajectories in Multicomponent Systems. *Multicomponent Random Systems, (R.L.Dobrushin, Ya.G.Sinai, Eds.)* Advances in Probability and Related Topics, V. 6 (1980), Dekker, pp. 549-575.

[36] Andrei Toom. [Estimations for measures that describe the behavior of random systems with local interaction. *Interactive Markov Processes and their Application to the Mathematical Modelling of Biological Systems*]. Pushchino, 1982, pp. 21-33 (in Russian).

[37] Andrei Toom. On Reliable Simulation in Real Time. *Preprints of the 1-st World Congress of the Bernoulli Society for Mathematical Statistics and Probability Theory*, 1986, v. 2, p. 676.

[38] Andrei Toom. On Critical Phenomena in Interacting Growth Systems. Part I: General. Part II: Bounded Growth. To appear in *Journal of Statistical Physics*, 1994.

[39] Leonid Vaserstein. Markov Processes over Denumerable Products of Spaces, Describing Large Systems of Automata. *Problems of Information Transmission*, v. 5 (1969), N. 3, pp. 47-52.

[40] Leonid Vaserstein and Andrei Leontovitch. Invariant Measures of Certain Markov Operators Describing a Homogeneous Random Medium. *Problems of Information Transmission*, V. 6 (1970), N. 1, pp. 61-69.

[41] N. B. Vasilyev, M. B. Petrovskaya and I. I. Piatetski-Shapiro. Simulation of voting with random errors. *Automatica i Telemekhanika*, V. 10 (1969), pp. 103-107 (in Russian).

[42] N. B. Vasilyev and I. I. Pyatetskii-Shapiro. The Classification of One-Dimensional Homogeneous Networks. *Problems of Information Transmission*, V. 7 (1971), N. 4, pp.340-346.

[43] Boris S. Zirelson. Non-Uniform Local Ineraction Can Produce 'Far-Range Order' in a One-Dimensional System. *Theor. Probability Appl.*, 21 (3), 681-683, 1976.