Ergod. Th. & Dynam. Sys. (2007), **27**, 545–565
doi:10.1017/S0143385706000782
© *2007 Cambridge University Press*
Printed in the United Kingdom

Damage spreading and μ-sensitivity on cellular automata

BRUNO MARTIN†
Bouygues SA, 32 av. Hoche, 75008 Paris, France
(e-mail: bmartin@bouygues.com)

(*Received* 1 *August* 2000 *and accepted in revised form* 12 *June* 2003)

Abstract. We show relations between new notions on cellular automata based on probabilistic dynamics: almost everywhere sensitivity to initial conditions for Besicovitch pseudo-distance, damage spreading (which measures the information (or damage) propagation), and the destruction of the initial configuration information. Through natural examples, we illustrate the relations between these formal definitions.

0. *Introduction*

A radius-r one-dimensional cellular automaton (CA) is an infinite succession of identical finite-state machines (indexed by $\mathbb{Z}$) called *cells*. Each finite-state machine is in a state and these states change simultaneously according to a local transition function: the following state of the machine is related to its own state as well as to the states of its $2r$ neighbors. A configuration of an automaton is a function which associates a state to each cell. We can thus define a global transition function from the set of all the configurations into itself which associates the following configuration after one computation step.

An evolution of a one-dimensional cellular automaton is usually represented by a space-time diagram. Being given an initial configuration, we represent in $\mathbb{Z} \times \mathbb{N}$ the successive cellular automaton configurations.

A lot of articles have proposed classifications of cellular automata [**6, 19**] but the prime reference is still Wolfram's empirical classification [**20**] which has resisted numerous attempts of formalization [**8**]. The classification of Gilman [**11**] is interesting because it is not a classification of CA, but a classification of couples (CA, measure on its configuration set). This choice, not motivated in the paper, seems interesting because we will illustrate on an example that the intuitive Wolfram classification depends on a measure, that is, a way of choosing a random configuration. Actually, due to their local interactions, CA are often used to simulate physics phenomena and many of them, for instance fluid flow, present a non-chaotic or chaotic behavior depending on some parameters (here the fluid speed).

† This paper was written while the author was working at the Ecole Normale Supérieure de Lyon, 46 allée d'Italie, 69007 Lyon, France.

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

Two very different ways have been investigated to find a definition of chaos for cellular automata that fits with our intuition. On the one hand, many people started from the mathematical definitions of chaos for dynamical systems and adapted them to cellular automata. By using the usual product topology on $\{0, 1\}^{\mathbb{Z}}$, the shift is necessarily chaotic, and for many possible applications of CA, like road traffic for example, we see that these definitions are maladjusted. The problem is that, even if the product topology is shift-invariant, the associated distance is not and it focuses on the neighborhood of an (arbitrary) origin. To correct this, we would like to consider simultaneously all the cells and thus introduce another topology such that the distance between two configurations is the 'density' of differences between them. The problem is that it is mathematically impossible to define a shift-invariant metric on $Q^{\mathbb{Z}}$. Anyway, Formenti introduced the Besicovitch topology on CA which is based on a shift-invariant pseudo-metric [**2, 7**]. For this topology, the phase space is not locally compact, and thus all the topological dynamics results become wrong or at least have to be proved again. On the other hand, starting from the physicist's approach of chaos, that is high sensitivity to initial conditions, Bagnoli *et al* [**1**] proposed to measure chaos experimentally through Lyapunov exponents, that is, roughly, to evaluate the damage spreading speed when a single cell is modified.

In this paper, we study probabilistic dynamics of CA. Like Gilman, we consider a measure on the configuration set and aim at characterizing the CA evolution from almost all initial configurations. In this context, we will formalize both approaches of the previous paragraph and study their relationships. On the one hand, we will define the almost everywhere sensitivity to initial conditions for Besicovitch topology and prove that the additive CA are almost everywhere sensitive (Proposition 2.2). On the other hand, inspired by Lyapunov exponents, we will define the damage spreading which is a function that, given t, returns the average number of modifications after t computation steps caused by a single cell modification on the initial condition. The damage spreading allows one to partition the set of couples (CA, measure) according to whether the function tends to zero, is bounded or not. We prove (Proposition 2.1) that the additive CA have non-bounded damage spreading. This notion appears to be more precise than Lyapunov exponents, but we will see that, while the average number of modifications has to be non-bounded to observe the almost everywhere sensitivity to initial conditions (Proposition 3.2), the condition is not sufficient. To understand the meaning of a damage spreading that tends to zero, we will also be interested in the definitions of Bμ-attracting sets and Dμ-attracting sets [**17**] which allow one to point out homogenization respectively to uniform and periodical configurations and prove (Proposition 3.1) that when the damage spreading tends to zero then the CA evolves to a set of uniform configurations. We will see (Corollary 2.1) that such a homogenization destroys asymptotically all the information contained in the initial configuration but that some CA that destroy the information of its initial configuration may have non-bounded damage spreading and even be almost everywhere sensitive to initial conditions.

We will begin with the definitions of cellular automata, Besicovitch topology and Borel measures. In a second section, we will formalize damage spreading, almost everywhere sensitivity and μ-attracting sets. We will then study the relations that exist between the classes we have defined.

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

1. *Definitions*

1.1. *Cellular automata.* For simplicity, we only consider one-dimensional CA in this paper. However, all the concepts we introduce are topological and it seems that there is no problem to extend them to higher dimensional CA.

Notation. We denote by Q a finite set of states.

If $c \in Q^{\mathbb{Z}}$ is a configuration, its state at position i will be denoted by $c(i)$ or c_i. If q is a state, q^* is the uniform configuration (i.e. for all $i \in \mathbb{Z}$, $q^*(i) = q_i^* = q$).

Let $\sigma : Q^{\mathbb{Z}} \longrightarrow Q^{\mathbb{Z}}$ be the shift toward the left: $(\sigma(c))(i) = c(i+1)$.

Definition. A *one-dimensional cellular automaton* is a couple $(Q^{\mathbb{Z}}, F)$ where $Q^{\mathbb{Z}}$ is the set of configurations and $F : Q^{\mathbb{Z}} \longrightarrow Q^{\mathbb{Z}}$ is a shift-invariant (i.e. $F(\sigma(c)) = \sigma(F(c))$), continuous (for the product topology) function named the *global transition function*.

Remark. We can associate to any shift-invariant continuous function F a radius r and a local transition function $\delta : Q^{2r+1} \longrightarrow Q$ such that $F = F_\delta$ where $(F_\delta(c))(i) = \delta(c(i-r), \ldots, c(i), \ldots, c(i+r))$ (see [**13**]).

Definition. An *elementary cellular automaton* (ECA) is a radius-1 two-state (usually 0 and 1) one-dimensional cellular automaton.

For ECA, we will use Wolfram's notation: they are represented by an integer between 0 and 255 such that the transition function of the CA number i whose writing in base 2 is $i = \overline{a_7 a_6 a_5 a_4 a_3 a_2 a_1 a_0}^2$ satisfies

$$\begin{aligned} \delta_i(0,0,0) &= a_0, \quad \delta_i(1,0,0) = a_4, \\ \delta_i(0,0,1) &= a_1, \quad \delta_i(1,0,1) = a_5, \\ \delta_i(0,1,0) &= a_2, \quad \delta_i(1,1,0) = a_6, \\ \delta_i(0,1,1) &= a_3, \quad \delta_i(1,1,1) = a_7. \end{aligned}$$

Let us remark that CA with different numbers may have the same behavior by exchanging the states 0 and 1, for instance $184 = \overline{10111000}^2$ and $226 = \overline{11100010}^2$. If r is a rule number, we will denote by $\bar{r}$ the rule after exchanging the states and by $\overleftrightarrow{r}$ the rule which has a symmetric behavior (see [**5**] for more details).

We will speak about the cellular automaton $120 = (\{0,1\}^{\mathbb{Z}}, F_{\delta_{120}})$ or equivalently of the rule 120.

In the general definition of additive CA due to Wolfram, an additive CA is a CA whose set of states is a group and that satisfies the superposition principle $(\delta(x+x', y+y', z+z') = \delta(x,y,z) + \delta(x',y',z'))$. These CA are very interesting to provide examples because their behaviors obey algebraic rules adapted to a formal study while their space-time diagrams appear complicated. We will use here, like in [**17, 19**], the following strictly more restrictive definition.

Definition. An additive CA is a one-dimensional CA whose state set is $\mathbb{Z}/n\mathbb{Z}$ and whose transition function is of the form

$$\delta_{+,n}(x_{-1}, x_0, x_1) = x_0 + x_1 (\mathrm{mod}\ n).$$

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

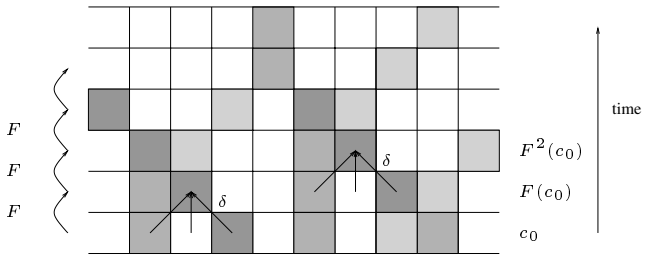


FIGURE 1. Space-time diagram example.

1.2. *Space-time diagrams.* Given an initial configuration, the orbit of this configuration may be represented by a space-time diagram. A configuration is a function from $\mathbb{Z}$ to the finite set of states Q which can be represented by a line of squares whose color indicates the state of the cell. By representing the next configuration above the previous one, we obtain a space-time diagram (thus times goes up, cf. Figure 1). Note that, vertically, we observe, for one cell, the successive states it takes.

As the phase space (the set of all orbits) cannot be represented canonically, space-time diagrams are an important tool in CA study. To try to understand the phase space through space-time diagrams which represent only one orbit, a common idea is to observe the space-time diagram of a random configuration. This is one of the basic ideas of the classifications proposed by Wolfram [**20**], Ishii [**14**] or Gilman [**11**]. The initial configurations of all the space-time diagrams of this paper are issued from a random process.

1.3. *The Besicovitch topology.* The most natural topology on CA configuration sets is the product topology. The problem is that the associated distance emphasizes what is happening close to the origin while in many applications of CA all the cells have the same importance. Thus, the adaptation of the mathematical notions of chaos to CA for the product topology are not adapted: the shift is necessarily chaotic, that is not adapted to car traffic simulation for example. Thus the idea is to define the distance between two configurations as the 'density' of differences between the configurations. The problem is that it has proved to be impossible to define a shift-invariant non-trivial distance on $Q^{\mathbb{Z}}$. Even if no canonical topology can be defined with this idea, it is natural to attempt to define a topology based on a pseudo-distance using an approximation of the 'density'. The Besicovitch topology (which is shift-invariant) is based on this idea, the formula of the 'density' of a subset A of $\mathbb{Z}$ being

$$\limsup_{l \longrightarrow +\infty} \frac{\#([-l, l] \cap A)}{2l + 1} \quad \text{where \# denotes the cardinality.}$$

We have to use a lim sup because (of course) generally the limit does not exist. Thus, for the configurations such that the lim sup will differ from the lim inf, the result is not really natural. This definition uses an origin but is independent of it.

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

There exist other shift-invariant pseudo-distances (see [**2, 15**]). Formenti suggests the Besicovitch point of view in the CA context because of the nicer properties of the induced topology. To save time, we will adopt the Besicovitch topology too since the main results related to CA are already proved, although the results of this paper are not specific to the Besicovitch topology. We will actually use definitions based on probability and be interested in almost everywhere (a.e.) sensitivity with respect to some measure. The point is that these topologies differ on null measure sets. A general study of Besicovitch-like topologies has been done in [**15**] and an interesting question would be to determine those that behave like the Besicovitch one and what happens for the other ones.

Remark. The main reason we point this out here is that there are many ways to extend Besicovitch in higher dimensional grids. The extension of Besicovitch pseudo-distance on $Q^{\mathbb{Z}^n}$ is

$$d(x, y) = \limsup_{l \longrightarrow +\infty} \frac{\#\{i \in B(0, l) \subset \mathbb{Z}^n \mid x_i \neq y_i\}}{\#B(0, l)}$$

where $B(0, l)$ is a ball centered at the origin and of radius l in $\mathbb{Z}^n$ for an arbitrary chosen distance on $\mathbb{Z}^n$, for instance $\bar{d}_1((a_1, \ldots, a_n), (b_1, \ldots, b_n)) = |b_1 - a_1| + \cdots + |b_n - a_n|$ or $\bar{d}_2((a_1, \ldots, a_n), (a_1, \ldots, a_n)) = ((b_1 - a_1)^2 + \cdots + (b_n - a_n)^2)^{1/2}$. Of course different distances may change the notion of sensitivity to initial conditions, but all the usual ones are equivalent for our purpose because they differ on a set of null measure. Let us notice that, when we choose equivalent distances on $\mathbb{Z}^n$, we obtain the same topology on $Q^{\mathbb{Z}^n}$.

Let us define the Besicovitch pseudo-distance, which induces a shift-invariant topology on the quotiented space.

Definition. The *Besicovitch pseudo-distance* on $Q^{\mathbb{Z}}$ is given by

$$d(x, y) = \limsup_{l \longrightarrow +\infty} \frac{\#\{i \in [-l, l] \mid x_j \neq y_j\}}{2l + 1}.$$

It is not a distance since obviously the pseudo-distance between two configurations equal everywhere except on a finite number of cells is null. If we consider two equal configurations except at the cells 2^n, there are an infinite number of differences, but the pseudo-distance is still null.

Properties. Besicovitch topology is the topology on $Q^{\mathbb{Z}}/\sim$ induced by d, where $x \sim y \iff d(x, y) = 0$. This topology is metric, pathwise connected, infinite dimensional, complete, and neither separable nor locally compact. Furthermore, $x \sim y \Longrightarrow F(x) \sim F(y)$ and the transition function of a CA is a continuous map from $Q^{\mathbb{Z}}/\sim$ into itself [**2**].

1.4. *Measures on the configuration set.*

Notation. Let Q be a finite alphabet with at least two letters. Q^n is the set of words of length n on Q. The cylinder of $u \in Q^p$ at position $k \in \mathbb{Z}$ is the set

$$[u]_k = \{x \in Q^{\mathbb{Z}} \mid x_{[k,k+p-1]} = u\}.$$

Definition. A Borel probability measure μ is *σ-invariant* if, for every measurable set Y, $\mu(\sigma^{-1}(Y)) = \mu(Y)$.

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

Definition. A σ-invariant measure is *σ-ergodic* if, for every invariant measurable set Y ($\sigma(Y) = Y$), either $\mu(Y) = 0$ or $\mu(Y) = 1$.

Fact. Let μ be a σ-invariant measure. The following properties are equivalent:

(1) μ is σ-weakly mixing;
(2) $\mu \times \nu$ is σ-weakly mixing for any σ-weakly mixing measure ν;
(3) $\mu \times \nu$ is σ-ergodic for any σ-ergodic measure ν.

Bernoulli measures are the most simple. The idea is that all the cells are independent, and the probability of each state q is given by a constant p_q. Their simplicity is the reason why we will use them in all our examples while some propositions hold for σ-ergodic or σ-weakly mixing measure. However, the main results (cf. §3) only concern Bernoulli measures. While some of them deal with μ-sensitivity, which only makes sense for Bernoulli measures, it may be possible to prove some Propositions (2.2 and 3.1) for other σ-weakly mixing measures, for instance Markov measures (with correlations over a finite number of cells) or measures such that the correlation between two states decreases exponentially with their distance. Obviously, to study specific rules, like number preserving rules, Bernoulli measures are less interesting and we may want to consider other σ-weakly mixing measures. But, practically, these other measures will often be Bernoulli measures after a grouping operation [**19**].

Definition. A *Bernoulli measure* is defined by a strictly positive probability vector $(p_q)_{q\in Q}$ with $\sum_{q\in Q} p_q = 1$ and if $u = u_0 \dots u_{n-1} \in Q^n$, $\mu[u_0 \dots u_{n-1}] = p_{u_0} \dots p_{u_{n-1}}$.

For a two-state CA, the Bernoulli measures will be denoted by μ_ρ where $\rho = p_1 = 1 - p_0$ is the probability for a state to be 1.

We will use the following classical result: *the Bernoulli measures are σ-weakly mixing (thus σ-ergodic).*

On Figure 2 one sees various space-time diagrams of a very simple traffic model (the CA T) based on the rule 184 but with two different models of cars. The principle is quite simple: a car stays at the same position if there is a car in front of it and it moves toward the right if the next position is free. The system seems 'chaotic' when the density ρ of cars is greater than or equal to 0.5 because of traffic jams, but not 'chaotic' elsewhere. Saying that this CA is chaotic or not does not make sense since it will depend on its utilization: whether it is used for traffic jam or for fluid traffic simulation. Its average behavior makes no sense while we do not explain what is a random configuration, that is, which measure we take on its configuration set. If we assume that the cars repartition is initially uniform and that we have the same number of red and blue cars, we will consider the Bernoulli measures μ_ρ^* such that the probability to find a blue car in a cell is $\rho/2$ and equal to the probability to find a red car, while the probability that there is no car is $1 - \rho$. Now, it is possible to say (see below) that this CA is μ_ρ^*-almost everywhere sensitive to initial conditions when $\rho \geq 1/2$, while it is μ_ρ^*-almost never sensitive to initial conditions otherwise. If it is important to take into account the fact that a lot of people take their cars at the same time to go to work, other measures allow one to model a non-uniform repartition.

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

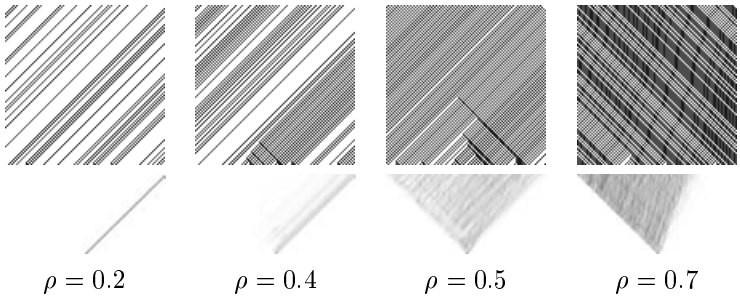


FIGURE 2. The CA T is a very simple traffic model that changes behaviors depending on the density of cars on the road (because of traffic jams). Below the space-time diagrams (time goes toward the top), we see with a grey level the space-time repartition of the average number of alterations induced by the modification of the middle cell. We see that the average number of alterations is bounded for small density of cars (less than $1/2$) and non-bounded otherwise.

2. *Some classification tools on CA*

2.1. *Damage spreading.* Inspired by Lyapunov exponents [**1**], we define the damage spreading of a CA via a measure. The main difference is that we count the 'effective' damages induced by a single cell modification. For instance, if the cell modification leads to two alterations after $t-1$ steps, such that each of these alterations would change one state at time t but the action of both leads this state to remain the same, then rather than counting two modifications like in the Lyapunov exponents, we count zero modification because the state did not change. It appears that the rule 210 (see Figure 4) has a Lyapunov exponent higher than 1 (thus the number of modifications is exponential if we may count a cell many times) while its damage spreading (the average number of different cells) is bounded. Let us now define this formally.

Definition. Let $(Q^{\mathbb{Z}}, F)$ be a CA. Let μ be a shift-invariant measure on the set of configurations. If $c \in Q^{\mathbb{Z}}$, we define

$$c_{p\leftarrow s} \left| \begin{array}{l} \mathbb{Z} \longrightarrow Q \\ i \longmapsto \begin{cases} c(i) & \text{if } i \neq p \\ s & \text{otherwise,} \end{cases} \end{array} \right.$$

that is, the configuration whose state of the cell p is changed into s. Let us now define the dependence coefficients $\alpha_{\mu,(Q^{\mathbb{Z}},F),t,p}$ which indicate the probability (according to μ) that the state of the cell 0 after t computation steps changes when we change the state at position p to q with probability $\pi_q = \mu([q])$. Formally we write

$$\alpha_{\mu,(Q^{\mathbb{Z}},F),t,p} = \sum_{q\in Q} \pi_q \mu(\{c \in Q^{\mathbb{Z}} \mid (F^t(c))_0 \neq (F^t(c_{p\leftarrow q}))_0\}).$$

Remarks. When no confusion about the measure and the cellular automaton is possible, we will simply denote the dependence coefficients by $\alpha_{t,p}$.

If r is the radius of the CA, $p > r \times t \Longrightarrow \alpha_{t,p} = 0$ because the cell 0 state after t computation steps is independent of the cell p of the initial configuration.

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

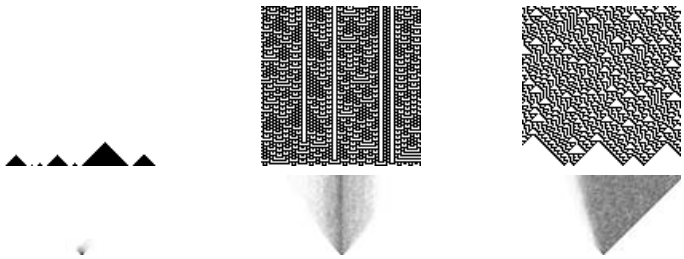


FIGURE 3. It is easy to prove that the rule 128 belongs to $[\mathcal{D} \xrightarrow{\mu}{}^{+} 0]$ and 109 to $[\mathcal{D} \xrightarrow{\mu}{}^{+} a]$ (in the second case we observe that 00 is a blocking word [**16, 3**]). Numerical experiment lets us think that 30 belongs to $[\mathcal{D} \xrightarrow{\mu}{}^{+} +\infty]$. The bottom diagrams represent with gray level the probability of each cell to be affected by the modification of the middle cell.

Definition. The *damage spreading* of a cellular automaton $(Q^{\mathbb{Z}}, F)$ according to a measure μ is the infinite sequence of positive real numbers

$$\left(\sum_{p\in\mathbb{Z}} \alpha_{\mu,(Q^{\mathbb{Z}},F),t,p}\right)_{t\in\mathbb{N}}.$$

This sequence is well defined for all t thanks to the previous remark: for any given t all except finitely many of the $(\alpha_{\mu,(Q^{\mathbb{Z}},F),t,p})_{p\in\mathbb{Z}}$ are equal to zero.

This notion allows one to define the class $[\mathcal{D} \xrightarrow{\mu}{}^{+} 0]$ of CA whose damage spreading tends to zero, the class $[\mathcal{D} \xrightarrow{\mu}{}^{+} +\infty]$ of CA whose damage spreading is not bounded (its lim sup is $+\infty$) and the class $[\mathcal{D} \xrightarrow{\mu}{}^{+} a]$ when the damage spreading lim sup is a finite non-zero value. Figure 3 shows an example in each class. Obviously, these three classes define a partition of the set of CA.

PROPOSITION 2.1. *The additive CA have non-bounded damage spreading, and they are in the class* $[\mathcal{D} \xrightarrow{\mu}{}^{+} +\infty]$ *for any non-trivial shift-invariant measure.*

We need the two following proofs (due to Allombert) before we can prove the proposition.

LEMMA 2.1. *Let p be a prime number, let a and n be positive integers and let $k < p^n$. Then*

$$\binom{ap^n - 1}{k} = \pm 1 \pmod{p}.$$

Proof. In $\mathrm{F}_p[X] = \mathbb{Z}/p\mathbb{Z}[X]$, we have

$$\begin{aligned}(X-1)^{ap^n-1} &= \frac{(X-1)^{ap^n}}{X-1}\\ &= \frac{(X^{p^n}-1)^a}{X-1} \qquad \text{(because } p \text{ is prime)}\\ &= \frac{(X^{p^n}-1)(X^{p^n}-1)^{a-1}}{X-1}.\end{aligned}$$

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

Furthermore

$$\frac{X^{p^n}-1}{X-1}=\sum_{j=0}^{p^n-1}X^j$$

and

$$(X^{p^n}-1)^{a-1}=(-1)^{a-1}+X^{p^n}P(X)$$

where $P(X)\in \mathrm{F}_p[X]$. Thus

$$(X-1)^{ap^n-1}=\sum_{j=0}^{p^n-1}(-1)^{a-1}X^j+X^{p^n}Q(X)$$

where $Q(X)\in \mathrm{F}_p[X]$ and the first p^n coefficients are ± 1. We complete the proof by remarking that the absolute values of the coefficients of $(X-1)^{ap^n}$ are the binomial coefficients. □

LEMMA 2.2. *For $N\geq 2$, for all M, there exists m such that for all $0<k<M$ we have*

$$\binom{m}{k}\in(\mathbb{Z}/N\mathbb{Z})^* \quad \textit{(i.e. it is invertible} \bmod N\textit{)}.$$

Proof. For each p which divides N, p prime, let us find e_p such that $p^{e_p}\geq M$. Define

$$m=-1+\prod_{p|N}p^{e_p}.$$

We apply Lemma 2.1 for p, $n=e_p$ and $a=\prod_{q|N,q\neq p}q^{e_q}$. Thus for all $k<M\leq p^n$

$$\binom{m}{k}=\pm 1 \pmod p.$$

So $\binom{m}{k}$ is invertible modulo any prime divisor of N, thus modulo N. □

Proof of Proposition 2.1. This is a direct consequence of Lemma 2.2 by using the well known dipolynomial representation of additive CA. Let $((\mathbb{Z}/\mathbb{Z}_n)^{\mathbb{Z}}, F)$ be an additive CA. By additivity, the evolution of any configuration c when we change the cell 0 to the state q is the superposition of the evolutions from c and from $0^*_{0\leftarrow q}$ (i.e. the configuration such that all the states are 0 except the state at the origin which is q). Formally, for all t, $F^t(c_{0\leftarrow q})=F^t(c)+F^t(0^*_{0\leftarrow q})$ where $+$ is the group operation. The dipolynomial representation says that $(F^t(0^*_{0\leftarrow q}))_i$ (the state of the cell at position i of the configuration after t time steps from a null configuration except in the cell 0) is q times the coefficient of degree i of $(1+X)^t \bmod n$. From Lemma 2.2, whatever $q\neq 0$, for all M, there exists m such that, for all $0<k<M$, $(F^m(0^*_{0\leftarrow q}))_k\neq 0$. Thus for any non-trivial measure, the additive CA are in $[\mathcal{D}\overset{\mu}{\to}+\infty]$. □

2.2. *μ-sensitivity to initial conditions.* The main reason for the introduction of μ-sensitivity is the study of some particular cases. On the one hand, experimentally, one can think that the rule 120 is sensitive to initial conditions starting from 'randomly generated' configurations but not for all configurations. On the other hand, we can prove

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

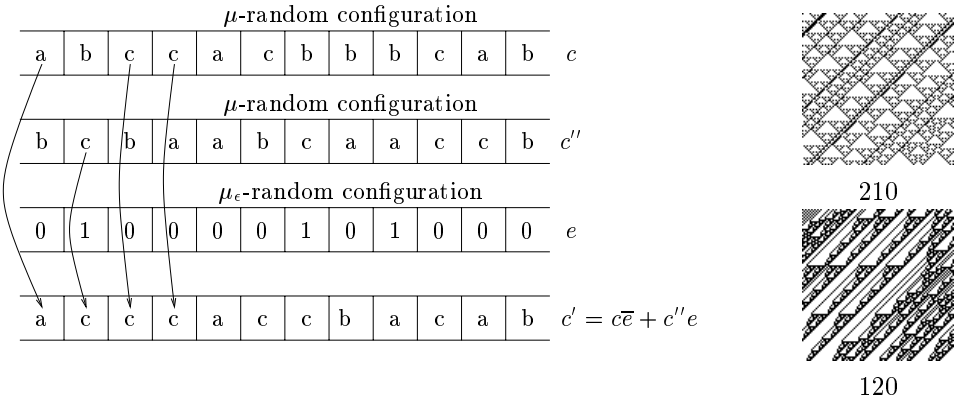


FIGURE 4. The configuration $c' = c\overline{e} + c''e$ is the configuration whose state at a given position is equal to the corresponding state of c when the corresponding state of e is equal to 0 and to the corresponding state of c'' otherwise.

that the rule 210 is not sensitive and not equicontinuous (i.e. is sensitive to initial conditions on a subset of its configuration set) because of the configuration 0^* (see Figure 4). Thus, in the same class, the elements of which are neither sensitive nor equicontinuous, we would have two rules with very different behaviors. The idea is to say that 120 is almost everywhere sensitive, while 210 is almost never sensitive. Note that while it is possible to prove that some CA like 210 are almost never sensitive but not equicontinuous, the question whether any almost everywhere sensitive CA is sensitive is open.

Let us recall the classical definition of sensitivity to initial conditions.

Definition. A CA $(Q^{\mathbb{Z}}, F)$ is *sensitive* to initial conditions for a pseudo-distance d if there exists a constant $M > 0$ such that, for all $\epsilon > 0$ and for all configurations c, there exists a configuration c' with $d(c, c') < \epsilon$ and an integer n such that $d(F^n(c), F^n(c')) \geq M$.

To define the μ-sensitivity to initial conditions, we could just replace 'for all configurations c' by 'for μ-almost all configurations c' in the sensitivity definition. This is the most natural definition. The problem is that this definition is not on the same level as the damage spreading definition and entropy consideration: while the last two notions only depend on the behavior of the CA for one measure, the natural μ-sensitivity depends on the behavior of the CA on a neighborhood of a measure. In order to obtain more significant comparison between these notions, we introduce a more restrictive definition as follows.

Definition. Let μ be a Bernoulli measure on $Q^{\mathbb{Z}}$. Define on the space $Q^{\mathbb{Z}} \times Q^{\mathbb{Z}} \times \{0, 1\}^{\mathbb{Z}}$ endowed with the measure $\mu \times \mu \times \mu_\epsilon$ the random variable X_n by $(c, c'', e) \longmapsto d(F^n(c), F^n(c'))$ where $c' = c\overline{e} + c''e$ is the configuration whose state at position i is equal to c_i when e_i is equal to 1 and to c''_i otherwise (see Figure 4). We observe that $d(F^n(c), F^n(c'))$ is shift-invariant and thus the values X_n are a.e. constant for $\mu \times \mu \times \mu_\epsilon$ which is σ-ergodic because μ and μ_ϵ are weakly mixing. Thus X_n is a.e. equal to its mean value $\overline{X_n}(\epsilon)$.

A CA is said to be *μ-sensitive* if there exists $M > 0$ such that for all $\epsilon > 0$ there exists n such that $\overline{X_n}(\epsilon) \geq M$. The class of μ-sensitive CA is denoted by $[\mu - s]$.

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

Remarks. If $c' = c\overline{e} + c''e$, where the repartition law of c and c'' (respectively e) is given by μ (respectively μ_ϵ), it is easy to check that the repartition law of c' is given by μ. Note that this is only true for Bernoulli measures and it is the reason why the μ-sensitivity definition is only relevant for Bernoulli measures.

This definition implies that there exists M such that, for μ-almost all configurations c and for all $\epsilon > 0$, there exist c' and n with $d(c, c') < \epsilon$ and $d(F^n(c), F^n(c')) \geq M$.

If d was a distance for the product topology, the previous result would imply the sensitivity to initial conditions. The point is that Bernoulli measures are of full support (i.e. the non-empty open sets have a non-null measure), but this is not the case on $Q^{\mathbb{Z}}/\sim$ with Besicovitch topology.

The set of configurations 3-uples (c, c'', e) such that if $c' = c\overline{e} + c''e$ there exists n satisfying $d(F^n(c), F^n(c')) \geq M$ is obviously shift-invariant in $(Q \times Q \times \{0, 1\})^{\mathbb{Z}}$. As Bernoulli measures are weakly mixing, $\mu \times \mu \times \mu_\epsilon$ is σ-ergodic, thus the set measure is either 1 or 0. So a CA is either μ-sensitive or 'μ-almost never sensitive to initial conditions': for any η there exists ϵ such that, if we build c, c' as before, for any n, $d(F^n(c), F^n(c')) \leq \eta$ $(\mu \times \mu \times \mu_\epsilon)$-almost everywhere. In other words, for all n, $\lim_{\epsilon \to 0} \overline{X_n}(\epsilon) = 0$.

The μ-sensitivity makes sense because we saw that some CA are not μ-sensitive (Propositions 3.2 and 3.3 will give further examples but we can notice that obviously the rule 0, the shifts and the identity are not) and we will prove that the additive CA are μ-sensitive.

PROPOSITION 2.2. *The additive CA are μ-sensitive for any non-trivial Bernoulli measure μ.*

Proof. Let $c, c'' \in (\mathbb{Z}/n\mathbb{Z})^{\mathbb{Z}}$ be two configurations, e a configuration in $\{0, 1\}^{\mathbb{Z}}$ and $c' = c\overline{e} + c''e$. The ith cell of $F^k(c')$ is different from the corresponding cell of $F^k(c)$ if and only if $[F^k(e(c'' - c))]_i \neq 0$ because $c' = c\overline{e} + c''e = (c - ce) + c''e = c + e(c'' - c)$. The number of invertible elements of the line t of the Pascal triangle modulo n is not bounded because of Lemma 2.2. As $\mathbb{Z}/n\mathbb{Z}^*$ is finite, there exists $f \in \mathbb{Z}/n\mathbb{Z}^*$ such that the number x_t of elements of the line t of the Pascal triangle equal to f modulo n is not bounded. We write $[F^t(e(c'' - c))]_0 = r_t + f S_t$ with

$$S_t = \sum_{\{i \mid \binom{t}{i} = f \pmod n\}} e_i (c''_i - c_i) \pmod n$$

and

$$r_t = \sum_{g \in \mathbb{Z}/n\mathbb{Z}, g \neq f} g \sum_{\{i \mid \binom{t}{i} = g \pmod n\}} e_i (c''_i - c_i).$$

Now, we want to calculate the probability that the sum S_t is equal to $g \in \mathbb{Z}/n\mathbb{Z}$. To do that, we can consider the random process over $\mathbb{Z}/n\mathbb{Z}$ starting from 0 that, at each step, adds $h \in \mathbb{Z}/n\mathbb{Z} \mod n$ with probability p'_h where $p'_0 = \epsilon p_0 + 1 - \epsilon$ and $p'_h = \epsilon p_h$ when $h \neq 0$. We remark first that it is a Markov process because the new 'state' (element of $\mathbb{Z}/n\mathbb{Z}$) only depends on the previous one. We observe that the probability $\pi_g(t)$ that S_t is equal to $g \in \mathbb{Z}/n\mathbb{Z}$ is exactly the probability that after x_t

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

steps the Markov process is in g (one step of the Markov process corresponds to one term of the sum and the terms of the sum are independent because the measures are Bernoulli). Let us remark that this probability can be expressed using the development of $(p'_0 + p'_1 X + p'_2 X^2 + \cdots + p'_{n-1} X^{n-1})^{x_t} = \sum_{i=0}^{(n-1)x_t} a_i X^i$: $\pi_g(t)$ is the sum of the coefficients of X^k for $k \in g + n\mathbb{Z}$, that is, $\sum_{k=g \pmod n} a_k$. This probability is thus given by the formula

$$\pi_g(t) = \frac{\sum_{k=0}^{n-1} \xi^{-kg} (p'_0 + \xi^k p'_1 + \cdots + (\xi^{n-1})^k p'_{n-1})^{x_t}}{n}$$

with $\xi = e^{2i\pi/n}$. Obviously, all the terms of the sum, except when $k = 0$, tend to 0 when $x_t \longrightarrow +\infty$ because $|p'_0 + \xi^k p'_1 + \cdots + (\xi^{n-1})^k p'_{n-1}| < 1$ and the first term ($k = 0$) is always 1 since $p'_0 + \cdots + p'_{n-1} = 1$. Thus the probability that S_t is equal to $g \in \mathbb{Z}/n\mathbb{Z}$ tends to $1/n$ on a sub-sequence t_k such that $x_{t_k} \longrightarrow +\infty$. By compactness, on a sub-sequence of t_k (that we still call t_k), the probability $\pi'_h(t_k)$ that $r_{t_k} = h \in \mathbb{Z}/n\mathbb{Z}$ converges to some value π'_h for every h. Now we remark that, for each t, S_t and r_t are independent because they depend on different cells and μ and μ_ϵ are Bernoulli. Thus the probability that $[F^{t_k}(e(c'' - c))]_0 = i \in \mathbb{Z}/n\mathbb{Z}$ is

$$\sum_{h+fg=i \pmod n} \pi'_h(t_k)\pi_g(t_k),$$

which tends to $1/n$ when k tends to $+\infty$. For any $M < 1 - 1/n$ (note that M does not depend on ϵ) there exists t such that the cells at position 0 of $F^t(c)$ and $F^t(c')$ are different with a probability greater than M for $\mu \times \mu \times \mu_\epsilon$.

Since the Besicovitch pseudo-distance X_t of $F^t(c)$ and $F^t(c')$ is shift-invariant, we already know that it is a.e. constant by ergodicity, thus equal to its mean value $\overline{X_t}(\epsilon)$, which, by the bilateral version of Birkhoff's theorem, is equal to $\mu \times \mu \times \mu_\epsilon([F^t(c)]_0 \neq [F^t(c')]_0)$. This concludes the proof since, whatever ϵ, this value is greater than M for some t. □

Remark. Note that we actually prove that $\overline{X_t}(\epsilon)$ is greater than M for infinitely many values, thus $\limsup \overline{X_t}(\epsilon) \geq M$. Furthermore, if n is prime, then it is easy to prove that, for all k, $\overline{X_{n^k}}(\epsilon) = \overline{X_1}(\epsilon) \leq 2\epsilon$ and thus $\liminf \overline{X_t}(\epsilon) \leq 2\epsilon$.

2.2.1. *μ-attracting sets.* μ-attracting sets are defined in [**17**]. In that article, Kůrka and Maass study the links between attracting and μ-attracting sets for different topologies.

Definition. A *sub-shift* is any subset $\Sigma \subseteq Q^{\mathbb{Z}}$ which is σ-invariant and closed in the product topology. The language $L(\Sigma)$ of a sub-shift $\Sigma \subseteq Q^{\mathbb{Z}}$ is the set of words of Σ. A sub-shift is of finite type (SFT) if there exists a positive integer p called order, such that, for all $c \in Q^{\mathbb{Z}}$,

$$c \in \Sigma \iff \forall i \in \mathbb{Z}, c_{[i,i+p-1]} \in L(\Sigma).$$

Definition. For an SFT $\Sigma \subseteq Q^{\mathbb{Z}}$ of order p and $x \in Q^{\mathbb{Z}}$, we define the *upper density* of Σ-defects in x by

$$d_D(x, \Sigma) = \limsup_{l \longrightarrow +\infty} \frac{\#\{i \in [-l, l] \mid x_{[i,i+p-1]} \notin L(\Sigma)\}}{2l + 1}.$$

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

Notation. When d is a pseudo-distance, we can naturally define the pseudo-distance from an element x to a set as the infimum of the pseudo-distance between x and the elements of the set:

$$d(x, \Sigma) = \inf_{y \in \Sigma} d(x, y).$$

Let us notice that $d_D(x, \Sigma)$ is not associated to any pseudo-distance because in the general case $d_D(x, \{y\}) \neq d_D(y, \{x\})$.

Definition. Let μ be a σ-ergodic measure. Then a sub-shift Σ is Bμ-*attracting* (respectively Dμ-*attracting*) if

$$\mu\left(\left\{c \in Q^{\mathbb{Z}} \,\middle|\, \lim_{n \to +\infty} \mathbf{d}(F^n(c), \Sigma) = 0\right\}\right) = 1$$

when $\mathbf{d}$ is the Besicovitch pseudo-distance d (respectively $\mathbf{d} = d_D$).

Remark. Note that for a given n the value $d(F^n(c), \Sigma)$ is shift-invariant and thus a.e. constant. Furthermore, if a sub-shift of finite type is Bμ-attracting then it is Dμ-attracting [**17**].

LEMMA 2.3. *If μ is a σ-weakly mixing measure and Σ is a* Bμ*-attracting sub-shift then the configurations c_0 such that $\mu(\{c \in Q^{\mathbb{Z}} \mid \liminf d(F^n(c), c_0) = 0\}) > 0$ are such that $d(c', q_0^*) = 0$ for some q_0. Thus when this set is not empty (for instance when Σ is finite) then we can build a minimal* Bμ*-attracting sub-shift composed only by uniform configurations.*

Proof. Let c', c'' be two configurations and $n \in \mathbb{N}$. The random variable $Y_n(c, c') = d(F^n(c'), F^n(c''))$ is $\sigma \times \sigma$ invariant and thus a.e. constant for $\mu \times \mu$ as μ is σ-weakly mixing. The a.e. value of $d(F^n(c'), F^n(c''))$ is equal to the mean value $\overline{Y_n}$. By hypothesis, there exists c_0 such that $\mu(\{c \in Q^{\mathbb{Z}} \mid \liminf d(F^n(c), c_0) = 0\}) > 0$. Let $\Omega(n, \epsilon) = \{c \mid d(F^n(c), c_0) \leq \epsilon\}$. For all $\epsilon > 0$ and $N \geq 0$ there exists $n \geq N$ such that $\mu(\Omega(n, \epsilon)) > 0$. As $d(F^n(c'), F^n(c'')) \leq d(F^n(c'), c_0) + d(c_0, F^n(c''))$, one has $Y_n(c, c') \leq 2\epsilon$ if $(c, c') \in \Omega(n, \epsilon) \times \Omega(n, \epsilon)$, which is a subset of non-null measure. Thus $\overline{Y_n} \leq \epsilon$ because Y_n is a.e. constant, thus $\liminf \overline{Y_n} = 0$.

Suppose by refutation that there is no $q \in Q$ such that $d(c_0, q^*) = 0$. Let $\delta = \min\{d(c_0, q^*) \mid q \in Q\} > 0$. Let $N \geq 0$ and $n \geq N$ such that $\mu(\Omega(n, \delta/2)) > 0$. For $c' \in \Omega(n, \delta)$ and $q \in Q$, one has $d(F^n(c'), q^*) \geq d(c_0, q^*) - d(F^n(c'), c_0) \geq \delta/2 > 0$. By applying bilateral Birkhoff's ergodic theorem to $\mathbf{1}_{[F^n(c')]_0 \neq q}$ we have $d(F^n(c'), q^*) = \mu([F^n(c)]_0 \neq q) \geq \delta/2 > 0$ because μ is σ-ergodic.

By applying again Birkhoff's ergodic theorem to $\mathbf{1}_{[F^n(c')]_0 \neq [F^n(c'')]_0}$, as $\mu \times \mu$ is σ-ergodic because μ is σ-weakly mixing, we have $\overline{Y_n} = \mu \times \mu([F^n(c')]_0 \neq [F^n(c'')]_0)$. Moreover, if $q_1, q_2 \in Q$ with $q_1 \neq q_2$, one has

$$\begin{aligned}\mu \times \mu([F^n(c')]_0 \neq [F^n(c'')]_0) &\geq \mu \times \mu([F^n(c')]_0 = q_1, [F^n(c'')]_0 = q_2) \\ &= \mu([F^n(c')]_0 = q_1)\mu([F^n(c'')]_0 = q_2) \\ &\geq \delta^2/4.\end{aligned}$$

This is a contradiction since $\liminf \overline{Y_n} = 0$. □

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

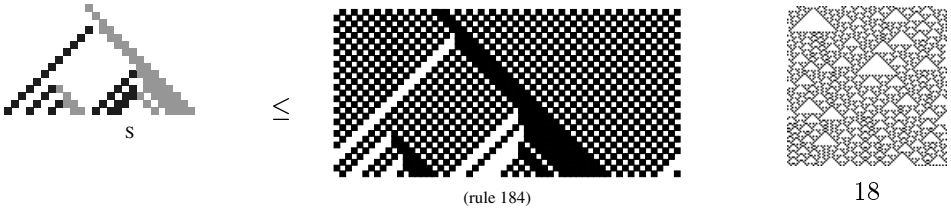


FIGURE 5. The CA $\mathcal{S}$ has three states: one is going to the right, one is going to the left (the third one is B, the blank one, into which the other states may move) and when they meet, both are annihilated. When μ is a measure such that the numbers of states going to the left and to the right have the same probability of presence on the initial configuration, then the uniform configuration composed by the blank state is a Bμ-attracting (thus Dμ-attracting) set. As $\mathcal{S}$ is a sub-automaton of 184^2 where 184^2 is the CA 184 whose states are grouped two by two (see [**18**]), we have the same result on 184^2 for the measure so that $(1, 0)$ has probability 0 and the probabilities of $(1, 1)$ and $(0, 0)$ are equal. This measure on $(\{0, 1\} \times \{0, 1\})^{\mathbb{Z}}$ corresponds to a non-shift-invariant measure on $\{0, 1\}^{\mathbb{Z}}$, and actually 184 has no finite Bμ-attracting set when μ is a non-trivial shift-invariant measure. The point is that for $\mu_{1/2}$ (so that particles going to the left and to the right have the same probability), the sub-shift $\{(01)^*, (10)^*\}$ is Dμ-attracting, but asymptotically the configurations tend to be at pseudo-distance $1/2$ of both configurations. The rule 18 could be an example of a CA with Dμ-attracting sets of non-finite type.

The definition of Bμ-attracting sets is very natural: a set is Bμ-attracting if, for almost all configurations (with respect to μ), the (Besicovitch) pseudo-distance between the successive configurations and the sub-shift tends to 0. We saw in Lemma 2.3 that, when the Bμ-attracting set is finite, almost all evolutions tend to uniform configurations when μ is a σ-weakly mixing measure. The rule 128 (see Figure 3) or the CA $\mathcal{S}$ and 184^2 for the uniform measure μ_u (see Figure 5) have a Bμ_u-attracting set. The definition of Dμ-attracting sets is more topological because it only depends on the language $L(\Sigma)$ and does not take care of how many times a pattern (a word of $L(\Sigma)$) appears. Thus, finite Dμ-attracting sets are more powerful and allow one to point out a homogenization process to periodic configurations and not only to uniform configurations. For instance, as proved in [**17**], the sub-shift $\{(01)^*, (10)^*\}$ is B$\mu_{1/2}$-attracting for the rule 184. When we observe the rule 18, we would like to extend the Dμ-attracting set definition to non-finite type subset to be able to say that $\{c \in \{0, 1\}^{\mathbb{Z}} \mid \forall i, c(2i) = 0\} \cup \{c \in \{0, 1\}^{\mathbb{Z}} \mid \forall i, c(2i+1) = 0\}$ is a Dμ-attracting set. Note that this CA has (infinite) Bμ-attracting (thus Dμ-attracting) sets of finite type like $\{c \in \{0, 1\}^{\mathbb{Z}} \mid \forall i, c_{[i,i+1]} \neq 11\}$ but that this is less significant on 18 evolution.

In the following, we only use Dμ-attracting sets to point out a homogenization to periodical configurations. In this case, we see that the whole information of the initial configurations is erased; formally, the metric entropy of the successive configurations tends to zero.

Definition. Let $(Q^{\mathbb{Z}}, F)$ be a CA and μ a σ-ergodic measure. The metric entropy of its configuration set after t computation steps is the metric entropy of σ for the measure $F^t(\mu)$. It can be defined as

$$h_\mu^{(t)}((Q^{\mathbb{Z}}, F)) = \lim_{n \to +\infty} \frac{-\sum_{u \in Q^n} p_u^{(t)} \log(p_u^{(t)})}{n} = \inf_{n \to +\infty} \frac{-\sum_{u \in Q^n} p_u^{(t)} \log(p_u^{(t)})}{n},$$

with the usual convention $0 \times \log(0) = 0$ and where $p_u^{(t)} = \mu(F^{-t}([u]))$.

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

Definition. The class of CA that erases all their initial configuration information is denoted by $[h_\mu \to 0]$ and formally defined as follows: a CA is in $[h_\mu \to 0]$ if and only if

$$\lim_{t\to+\infty} h_\mu^{(t)}((Q^{\mathbb{Z}}, F)) = 0.$$

Notation. Let us denote $L_n(\Sigma) = Q^n \cap L(\Sigma)$ and $\overline{L(\Sigma)} = Q^n \backslash L(\Sigma)$.

Definition. The topological entropy of a sub-shift Σ is

$$h_\Sigma(\sigma) = \lim_{n\longrightarrow+\infty} \frac{\log(\#\{u \in Q^n \mid \exists c \in \Sigma, c_{[0,n-1]} = u\})}{n} = \lim_{n\longrightarrow+\infty} \frac{\log(\#L_n(\Sigma))}{n}.$$

Recall the variational principle [**9**] which says that the topological entropy is the supremum of the metric entropies. This implies that the metric entropy is less than or equal to the topological entropy. We use this result to prove the following proposition.

PROPOSITION 2.3. *If a CA has a* Dμ*-attracting set* Σ *then*

$$\limsup_{t\to+\infty} h_\mu^{(t)}((Q^{\mathbb{Z}}, F)) \le h_\Sigma(\sigma).$$

COROLLARY 2.1. *If a CA has a* Dμ*-attracting set* Σ *of null topological entropy (i.e.* Σ *is a finite union of shift-periodic orbits), then it is in* $[h_\mu \to 0]$.

Notation. We denote

$$d_D^{(k)}(x, \Sigma) = \limsup_{l\longrightarrow+\infty} \frac{\#\{i \in [-l, l] \mid x_{[i,i+k-1]} \notin L(\Sigma)\}}{2l+1}.$$

LEMMA 2.4. *Let* Σ *be an SFT of order* p. *If* $k \ge p$, *then* $d_D^{(k)}(x, \Sigma) \le (k-p)d_D(x, \Sigma)$.

Proof. As Σ is an SFT, $x_{[i,i+k-1]} \in L(\Sigma)$ if and only if, for all j, $0 \le j \le k-p$, $x_{[j,j+p-1]} \in L(\Sigma)$. Thus for each $i \in [-l, l]$ such that $x_{[i,i+p-1]} \notin L(\Sigma)$, we have $x_{[j,i+k-1]} \notin L(\Sigma)$ for $i-k+p \le j \le i$. Thus in the worst case, $\#\{i \in [-l, l] \mid x_{[i,i+k-1]} \notin L(\Sigma)\} \le (k-p)\#\{i \in [-l, l] \mid x_{[i,i+p-1]} \notin L(\Sigma)\}$. We deduce that $d_D^{(k)}(x, \Sigma) \le (k-p)d_D(x, \Sigma)$. □

Proof of Proposition 2.3. By definition,

$$\lim_{t\to+\infty} h_\mu^{(t)}((Q^{\mathbb{Z}}, F)) = \lim_{t\to+\infty} \inf_n \frac{-\sum_{u\in Q^n} p_u^{(t)} \log(p_u^{(t)})}{n},$$

where $p_u^{(t)} = \mu(F^{-t}([u]))$.

Let us prove that this limit is bounded by $h_\Sigma(\sigma)$ plus a quantity that goes to 0.

For all $\epsilon > 0$, by definition of $h_\Sigma(\sigma)$, there exists k such that

$$\frac{\log(\#L_k(\Sigma))}{k} \le h_\Sigma(\sigma) + \epsilon.$$

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

In this case, we have

$$\begin{aligned}\inf_n \frac{-\sum_{u\in Q^n} p_u^{(t)} \log(p_u^{(t)})}{n} &\le \frac{-\sum_{u\in Q^k} p_u^{(t)} \log(p_u^{(t)})}{k}\\ &= \frac{-\sum_{u\in L_k(\Sigma)} p_u^{(t)} \log(p_u^{(t)})}{k} + \frac{-\sum_{u\in \overline{L_k(\Sigma)}} p_u^{(t)} \log(p_u^{(t)})}{k}\\ &\le \frac{\log(\#L_k(\Sigma))}{k} + \frac{-\sum_{u\in \overline{L_k(\Sigma)}} p_u^{(t)} \log(p_u^{(t)})}{k}\\ &\le h_\Sigma(\sigma) + \epsilon + \frac{-\sum_{u\in \overline{L_k(\Sigma)}} p_u^{(t)} \log(p_u^{(t)})}{k}.\end{aligned}$$

Now, to deal with the last term, we prove that $\sum_{u\in \overline{L_k(\Sigma)}} p_u^{(t)} \le \epsilon$. The first step is to prove that $\sum_{u\in \overline{L_k(\Sigma)}} p_u^{(t)} = d_D^{(k)}(F^t(c), \Sigma)$.

Let us define

$$f \left| \begin{array}{l} Q^{\mathbb{Z}} \longrightarrow \mathbb{R} \\ \quad c \longmapsto \begin{cases} 1 & \text{if } c_{[0,k-1]} \in \overline{L_k(\Sigma)} \\ 0 & \text{otherwise} \end{cases} \end{array} \right.$$

and $g_t = f \circ F^t$. We have

$$\int g_t \, d\mu = \sum_{u\in \overline{L_k(\Sigma)}} p_u^{(t)}.$$

We can apply bilateral Birkhoff's ergodic theorem with μ and σ: if

$$S_n(g_t)(c) = \frac{1}{2n+1} \sum_{i=-n}^{n} g_t \circ \sigma^i(c),$$

then for μ-almost all c, $\lim_{n\to\infty} S_n(g_t)(c) = \int g_t \, d\mu$:

$$\lim_{n\to\infty} S_n(g_t)(c) = \lim_{n\to\infty} \frac{1}{2n+1} \#\{i \in [-n, n] \mid F^t(c)_{[i,i+k-1]} \notin L(\Sigma)\} = d_D^{(k)}(F^t(c), \Sigma).$$

Thus, for μ-almost all c,

$$d_D^{(k)}(F^t(c), \Sigma) = \int g_t \, d\mu = \sum_{u\in \overline{L_k(\Sigma)}} p_u^{(t)}.$$

We have that $d_D^{(k)}(F^t(c), \Sigma)$ is a shift-invariant random variable, therefore μ-a.e. equal to a constant C_t. As Σ is Dμ-attracting, for any $\epsilon > 0$, there is M such that $t > M$ implies $C_t \le \epsilon$.

Thus, for μ-almost all c and $t > M$, we have $\sum_{u\in \overline{L_k(\Sigma)}} p_u^{(t)} \le \epsilon$.

Furthermore, $-\sum p_u^{(t)} \log(p_u^{(t)})$ is maximal when all $p_u^{(t)}$ are equal. Thus in the worst case they are equal to $\epsilon/|\overline{L_k(\Sigma)}|$ and we have

$$\begin{aligned}&\inf_{n\to+\infty} \frac{-\sum_{u\in Q^n} p_u^{(t)} \log(p_u^{(t)})}{n}\\ &\quad \le h_\Sigma(\sigma) + \epsilon + \limsup_{n\to+\infty} |\overline{L_k(\Sigma)}| \left(\frac{-1}{n} \frac{\epsilon}{|\overline{L_k(\Sigma)}|} \log\left(\frac{\epsilon}{|\overline{L_k(\Sigma)}|} \right) \right)\end{aligned}$$

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

$$\leq h_\Sigma(\sigma) + \epsilon + \limsup_{n\to+\infty} \frac{-\epsilon}{n} \log\left(\frac{\epsilon}{|\overline{L_k(\Sigma)}|}\right) \leq h_\Sigma(\sigma) + \epsilon + \limsup_{n\to+\infty} \frac{-\epsilon}{n} \log\left(\frac{\epsilon}{|Q|^n}\right)$$

$$\leq h_\Sigma(\sigma) + \epsilon + \limsup_{n\to+\infty} \frac{-\epsilon}{n} \log(\epsilon) + \epsilon \log(|Q|) \leq h_\Sigma(\sigma) + \epsilon + \epsilon \log(|Q|).$$

This inequality holds for all $t \geq M$, thus, as $\epsilon + \epsilon \log(|Q|) \longrightarrow 0$,

$$\limsup_t h_\mu^{(t)}((Q^{\mathbb{Z}}, F)) \leq h_\Sigma(\sigma).$$

This completes the proof of Proposition 2.3. □

3. *Relations between damage spreading, μ-a.e. sensitivity and the existence of a finite* Bμ*-attracting set*

In this section, we always assume that μ is a Bernoulli measure of coefficients $(p_q)_{q\in Q}$.

PROPOSITION 3.1. *If the μ-damage spreading of $(Q^{\mathbb{Z}}, F)$ tends to 0 then there exists a set of uniform configurations which is* Bμ*-attracting.*

Proof. Let Q_∞ be the subset of states that appear after an arbitrary number of computation steps. Let c and c' be two configurations. Let us fix $t > 0$. We consider the state at position 0 after t computation steps. Let us imagine that we change the states from index $-rt$ to rt of c to the states of c' one by one: let $c^0 = c$ and $c^{i+1} = c^i{}_{-rt+i \leftarrow c'(-rt+i)}$. The probability that $F^t(c^i)_0 \neq F^t(c^{i+1})_0$ is $\alpha_{t,-rt+i}$. The repartition law of c^i is given by μ because the cells are independent. Thus, the average number of alterations during this process of the cell 0, t computation steps later, is $\sum_p \alpha_{t,p} \longrightarrow_t 0$ so the probability that the cell always stays in the same state (i.e. for all k, $0 \leq k \leq rt + 1$, $[F^t(c^k)]_0 = [F^t(c)]_0$) is $1 - \sum \alpha_{t,p} \longrightarrow 1$. Obviously, $F^t(c^{2rt+1})_0 = F^t(c')_0$; thus, for each t, there exists q such that $\mu(\{c \mid [F^t(c)]_0 = q\}) \geq 1 - \sum \alpha_{t,p}$. We deduce that for each ϵ there exists M such that, for $t \geq M$, there is a state $q \in Q_\infty$ (q depends on t) such that $\mu(\{c \mid [F^t(c)]_0 = q\}) \geq 1 - \epsilon$. From Birkhoff's ergodic theorem, with probability 1, $\mu(\{c \mid [F^t(c)]_0 \neq q\}) = d(F^t(c), q^*)$ because μ is σ-ergodic (we apply Birkhoff's theorem to the function $f_t(c) = 1$ if $[F^t(c)]_0 \neq q$ and 0 otherwise). Thus $d(F^t(c), \{q^* \mid q \in Q_\infty\}) \leq \epsilon$ for all $t \geq M$. Consequently $\{q^* \mid q \in Q_\infty\}$ is Bμ-attracting. □

Reciprocally, there are CA with finite Bμ-attracting sets that are not in $[\mathcal{D} \overset{\mu}{\to}{}^+ 0]$. For instance, the rule $\mathcal{S}$ (see Figure 5) with three states (a blank one B into which a state $0l$ goes to the left and a state r goes to the right, and the collision of two particles leads to their annihilation) is in $[\mathcal{D} \overset{\mu}{\to}{}^+ a]$ (for all non-trivial μ) because it is obvious that the modification of one cell of the initial configuration changes exactly one state at each time step. Thus $\{B^*\}$ is a B$\mu_{1/2}$-attracting set of $\mathcal{S}$ because it is easy to prove that the number of non-B states decreases as $1/\sqrt{t}$ using random walk results (this kind of proof in a more general case is to be found in [**10**]).

In addition, we will describe later a CA that experimentally seems to be in $[\mathcal{D} \overset{\mu}{\to}{}^+ +\infty]$ but tends to a uniform configuration.

Let us now investigate the relations between damage spreading and μ-sensitivity (Figure 6).

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

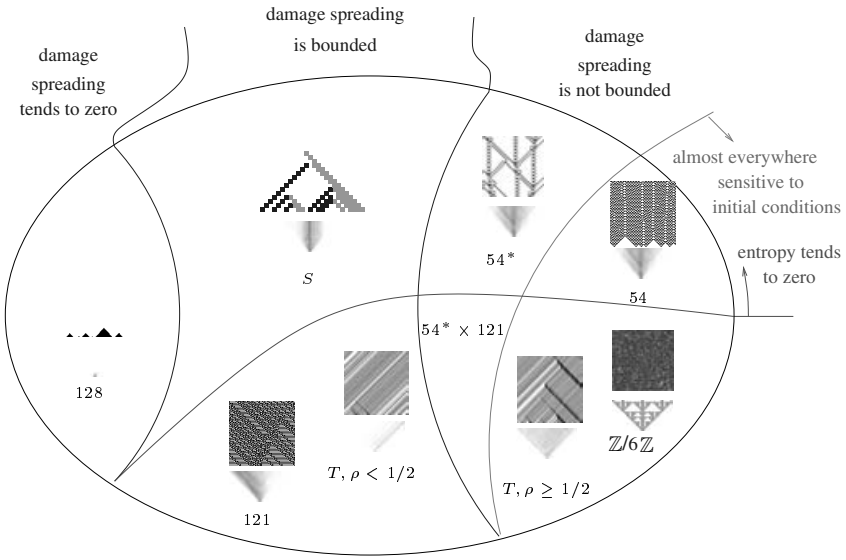


FIGURE 6. Relations between damage spreading, μ-sensitivity and the existence of Bμ-attracting sets when μ is a Bernoulli measure.

PROPOSITION 3.2. *Being μ-sensitive implies that there is non-bounded damage spreading (i.e. $[\mu - s] \subseteq [\mathcal{D} \overset{\mu}{\to}^{+} +\infty]$).*

Proof. When the damage spreading of a CA $(Q^{\mathbb{Z}}, F)$ is bounded, let us prove that, for all η, there is ϵ_0 such that $\epsilon \leq \epsilon_0$ implies that for all n, $d(F^n(c), F^n(c')) \leq \eta$ almost everywhere for $\mu \times \mu \times \mu_\epsilon$ on the set of triplets $(c, c'', e) \in Q^{\mathbb{Z}} \times Q^{\mathbb{Z}} \times \{0, 1\}^{\mathbb{Z}}$ when $c' = c\overline{e} + c''e$. Let M be a bound on the damage spreading of $(Q^{\mathbb{Z}}, F)$. Let us imagine, like in the proof of Proposition 3.1, that we change the states from index $-rt$ to rt of c to the states of c' one by one: let $c^0 = c$ and $c^{i+1} = c^i{}_{-rt+i \leftarrow c'(-rt+i)}$. Obviously $F^t(c^{2rt+1})_0 = F^t(c')_0$. This time, we change a state to q with probability ϵp_q and keep the current state otherwise (with probability $(1 - \epsilon)$); thus, the $\mu \times \mu \times \mu_\epsilon$-probability that $F^t(c^i)_0 \neq F^t(c^{i+1})_0$ is $\epsilon\alpha_{t,-rt-i}$. During this experience, the average number of alterations of a state after t steps of computation is $\epsilon \sum \alpha_{t,p} \leq \epsilon M$, which tends to 0 when $\epsilon \longrightarrow 0$. The probability that the state never changes is at least $1 - \epsilon \sum \alpha_{t,p} \geq 1 - \epsilon M$, and thus $\mu \times \mu \times \mu_\epsilon(\{(c, c'', e) \mid F^t(c)_0 = F^t(c\overline{e} + c''e)_0\}) \geq 1 - \epsilon M$. We use the σ-ergodicity of $\mu \times \mu \times \mu_\epsilon$ (because μ and μ_ϵ are σ-weakly mixing) and Birkhoff's ergodic theorem to prove that, with probability 1, $d(F^t(c), F^t(c')) = \mu \times \mu \times \mu_\epsilon(\{(c, c'', e) \mid F^t(c)_0 \neq F^t(c')_0\}) \leq \epsilon M$. So for ϵ small enough ($\epsilon < \eta/M$), the pseudo-distance between $F^t(c)$ and $F^t(c')$ is less than or equal to η with probability 1. □

It is experimentally easy to observe but it seems difficult to prove that 184 is $\mu_{1/2}$-sensitive (see Figure 7). Actually, the overall evolution of 184 leads to bigger and bigger homogeneous regions of (01)* and (10)*, but when we take two generic initial configurations, there is no reason that these regions match. Thus, asymptotically, the pseudo-distance between the configurations after many computation steps is 1/2.

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

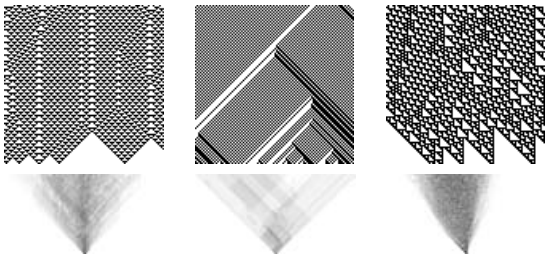


FIGURE 7. Experimentally, 54 and 184 for $\mu_{1/2}$ are in $[\mathcal{D} \xrightarrow{\mu} +\infty] \cap [h_\mu \to 0]$; we guess that 110 has the same behavior.

It seems that the rule 54 also belongs to $[\mu - s] \cap [h_\mu \to 0]$ (see Figure 8). Actually, the disappearance of g_e particles is an irreversible phenomenon as proved in [**18**], which will occur more and more rarely when the particles g_e become fewer and fewer but the number of g_e particles tends to 0 as confirmed by experiments [**4, 12**]. Then if any interaction between w and g_0 particles occurs, one particle disappears; we can think that the sub-shift $\{(0001)^*, (1110)^*\}$ is $\mathrm{D}\mu_\rho$-attracting for $0 < \rho < 1$.

Proposition 3.2 raises a natural question: Are the CA with unbounded damage spreading μ-sensitive?

It seems that the answer is no. Let us consider a CA 54* that simulates the rule 54 but in a uniform background. Such a CA can be formally defined thanks to Hanson and Crutchfield's filter [**12**] which is a CA but only on 'valid' configurations. Here, we consider any extension of this CA for a measure such that a generic configuration is correct with probability 1. Experimentally, the particle number decreases like $t^{-1/2}$ and thus tends to 0 so that this CA for well chosen measures tends to a uniform configuration. Due to the slow particle number decreasing, this CA seems to have non-bounded damage growth: with a non-null probability, a single perturbation creates or suppresses one or more particles. At each interaction of a defect particle with a particle, the defect is duplicated; the probability of such collisions is linked to the density of particles, and thus the average number of defects should look like $\int x^{-1/2} = 2\sqrt{x}$, which tends to $+\infty$ when $x \longrightarrow +\infty$. This CA seems to have non-bounded damage growth (that is experimentally observed). Furthermore this CA does not seem to be μ-sensitive because of the following proposition.

PROPOSITION 3.3. *Let μ be a Bernoulli measure and $(Q^{\mathbb{Z}}, F)$ be a CA with a finite $\mathrm{B}\mu$-attracting set (i.e. it tends to a set Σ of uniform configurations μ-almost everywhere). Then $(Q^{\mathbb{Z}}, F)$ is not μ-sensitive.*

Proof. First prove that, for every constant $\alpha > 0$, there is M such that $n > M$ implies $d(F^n(c), F^n(c')) \leq \alpha$ for $\mu \times \mu \times \mu_\epsilon$-almost every triplet (c, c'', e) where $c' = c\overline{e} + c''e$. Let $q^* \in \Sigma$, the random variable $Z_{n,q^*}(c) = d(q^*, F^n(c))$ is a.e. constant (q^* is uniform) and a.e. equal to its mean value $\overline{Z_{n,q^*}}$; thus, by hypothesis,

$$\lim_{n \to +\infty} \min_{c_0 \in \Sigma} \overline{Z_{n,q^*}} = 0.$$

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

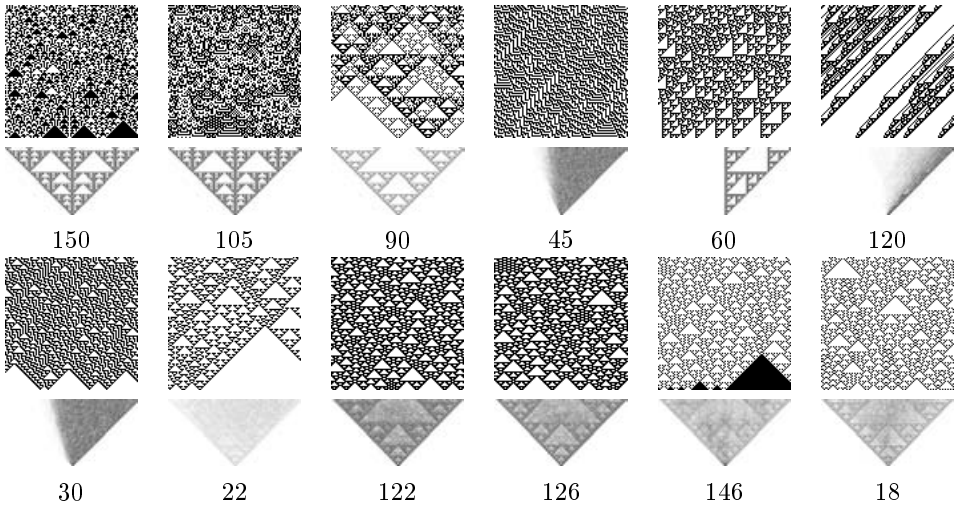


FIGURE 8. All the 'chaotic behaviors' (belonging to $[\mu - s]\backslash[h_\mu \to 0]$ for all non-trivial Bernoulli measures) among ECA (experimental results).

Thus there is M and a sequence (c'_n) with $c'_n \in \Sigma$ such that, for μ-almost all configurations c, $n > M$ implies that $d(c'_n, F^n(c)) \leq \alpha/2$. Actually, the inequality holds for almost all $c' = c\overline{e} + c''e$ because the probability distribution law of c' is also μ (first point of the Remarks preceding Proposition 2.2). Now, by triangular inequality, $\mu \times \mu \times \mu_\epsilon$-almost surely $d(F^n(c), F^n(c')) \leq d(q^*, F^n(c)) + d(q^*, F^n(c')) \leq \alpha$.

This allows one to bound the distance between two orbits after M iterations. Now, to bound it on the first M iterations, we use the fact that F and thus F^n act in a uniformly continuous way with respect to the pseudo-distance d (elementary proof in [**2**]). Thus there is $\epsilon > 0$ such that $d(c, c') \leq \epsilon$ implies that, for all $n \leq M$, we have $d(F^n(c), F^n(c')) \leq \alpha$.

Let c, c'' be two configurations of $Q^{\mathbb{Z}}$ and e a configuration of $\{0, 1\}^{\mathbb{Z}}$; we consider as before $c' = c\overline{e} + c''e$. Almost everywhere for $\mu \times \mu \times \mu_\epsilon$, we have $d(c, c') \leq \epsilon$ (second point of the Remarks preceding Proposition 2.2). Thus for $n \leq M$, we have $d(F^n(c), F^n(c')) \leq \alpha$ and for $n > M$, we have $d(F^n(c), F^n(c')) \leq \alpha$ too. We conclude that $(Q^{\mathbb{Z}}, F)$ is not μ-sensitive. □

Open questions. A lot of open questions remain, among them the following:

(1) to find examples in or to prove the emptiness of $[\mu - s]$ with a Bμ-attracting set and examples of CA in $[h_\mu \to 0]$ with no Bμ-attracting set and not in $[\mu - s]$;
(2) the generalization for more general measures of the propositions;
(3) to find a 'good' definition that separates the identity from 118 in $[\mathcal{D} \overset{\mu}{\to}^{+} a]$.

Acknowledgements. The author wants to thank B. Allombert, E. Formenti and A. Maass for helpful conversations and useful comments. The author is greatly indebted to S. Ruette for explaining to him the basic notions of ergodic theory and for her careful reading (and attentive proof checking) of the text.

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press

REFERENCES

[1] F. Bagnoli, R. Rechtman and S. Ruffo. Lyapunov exponents for cellular automata. *Lectures on Thermodynamics and Statistical Mechanics*. Eds. M. López de Haro and C. Varea. World Scientific, Singapore, 1994.

[2] F. Blanchard, E. Formenti and P. Kůrka. Cellular automata in the Cantor, Besicovitch and Weyl topological spaces. *Complex Systems* **11**(2) (1999), 107–123.

[3] F. Blanchard and A. Maass. Dynamical behaviour of Coven's cellular automata. *Theoret. Comput. Sci.* **163** (1996), 291–302.

[4] N. Boccara, J. Nasser and M. Roger. Particle-like structures and their interactions in spatio-temporal patterns generated by one-dimensional deterministic cellular-automata rules. *Phys. Rev.* A **44** (1991), 866–875.

[5] G. Cattaneo, E. Formenti, L. Margara and G. Mauri. Transformation of the one-dimensional cellular automata rule space. *Parallel Comput.* **23** (1997), 1593–1611.

[6] G. Cattaneo, E. Formenti, L. Margara and G. Mauri. On the dynamical behavior of chaotic cellular automata. *Theoret. Comput. Sci.* **217** (1999), 31–51.

[7] G. Cattaneo, E. Formenti, L. Margara and J. Mazoyer. Shift invariant distance on $S^{\mathbb{Z}}$ with non trivial topology. *Proceedings of MFCS '97*. Springer, Berlin, 1997.

[8] K. Čulik, J. Pach and S. Yu. On the limit set of cellular automata. *SIAM J. Comput.* **18** (1989), 831–842.

[9] M. Denker, C. Grillenberger and K. Sigmund. *Ergodic Theory on Compact Spaces (Lecture Notes in Mathematics, 527).* Springer, Berlin, 1976.

[10] H. Fuks. Exact results for deterministic cellular automata traffic models. *Phys. Rev.* E **60** (1999) 197–202. http://xxx.lanl.gov/abs/comp-gas/9902001.

[11] R. H. Gilman. Classes of linear automata. *Ergod. Th. & Dynam. Sys.* **7** (1987), 105–118.

[12] J. E. Hanson and J. P. Crutchfield. Computational mechanics of cellular automata: an example. *Physica* D **103** (1997) 169–189.

[13] G. A. Hedlund. Endomorphism and automorphism of the shift dynamical system. *Math. Syst. Theory* **3** (1969), 320–375.

[14] S. Ishii. Measure theoretic approach to the classification of cellular automata. *Discrete Appl. Math.* **39** (1992), 125–136.

[15] V. Kanovei and M. Reeken. On Ulam's problem of approximation of non-exact homomorphisms. *Preprint*, 2000.

[16] P. Kůrka. Languages, equicontinuity and attractors in cellular automata. *Ergod. Th. & Dynam. Sys.* **217** (1997), 417–433.

[17] P. Kůrka and A. Maass. Stability of subshifts in cellular automata. *Preprint*, 1999.

[18] B. Martin. A group interpretation of particles generated by one dimensional cellular automaton, 54 Wolfram's rule. *Internat. J. Modern Phys.* C **11**(1) (2000), 101–123.

[19] J. Mazoyer and I. Rapaport. Inducing an order on cellular automata by a grouping operation. *STACS '98 (Lecture Notes in Computer Science, 1373).* Springer, Berlin, 1998, pp. 116–127.

[20] S. Wolfram. Universality and complexity in cellular automata. *Physica* D **10** (1984), 1–35.

https://doi.org/10.1017/S0143385706000782 Published online by Cambridge University Press