Damage Spreading and μ-sensitivity on Cellular Automata

Bruno Martin

LIP, ENS de Lyon, 46, allée d'Italie, 69364 Lyon Cedex 07, France,
`Bruno.Martin@ens-lyon.fr`,
WWW home page: `http://www.ens-lyon.fr/~bmartin`

Abstract. We show relations between new notions on cellular automata based on topological and measure-theoretical concepts: 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 links between these formal definitions and Wolfram's empirical classification.

Introduction

A radius-r unidimensional 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 the states of its $2r$ neighbors. A configuration of an automaton is the function which associates to each cell a state. We can thus define a global transition function from the set of all the configurations into itself which associates the following configuration after one step of computation.

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

Recently, a lot of articles proposed classifications of cellular automata [13, 6] but the reference is still Wolfram's empirical classification [15] which has resisted numerous attempts of formalization [14]. The classification of Gilman [7] is interesting because it is not a classification of CAs, 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's classification depends on a measure, that is a way to choose a random configuration. Actually, due to their local interactions, CAs 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).

Recently, 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 CAs, like roadways traffic for example, we see that these definitions are maladjusted. That is why Formenti introduced Besicovitch topology [4, 2]. For this topology, the phase space is not locally compact, thus all the mathematical results become wrong or at least have to be proved again. On the other hand, starting from the physicist approach of chaos, that is high sensitivity to initial conditions, Bagnoli et al. [1] propose to measure chaos experimentally through Lyapunov exponents, that is, roughly, to evaluate the damage spreading speed when a single cell is modified.

In this article, we will formalize both approaches to study their relationships. First, we will define the almost everywhere sensitivity to initial conditions for Besicovitch topology and then we will partition the CAs whether the default number in average tends to zero, is bounded or not. We will also be interested in the definitions of Bμ-attracting sets and Dμ-attracting sets [10]. Let us notice that all the definitions depend on a measure.

We will begin by the definitions of cellular automata, Besicovitch topology and Bernoulli measure. 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 defined. The last section speaks about the links with Wolfram's classification.

The extended version with the proofs is to be found as research report available by FTP [11].

1 Definitions

1.1 Cellular automata

For simplicity, we will only consider unidimensional CAs 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 CAs.

Definition 1. *A* **radius-r unidimensional cellular automaton** *is a couple (Q,δ) where Q is a finite set of states and $\delta : Q^{2r+1} \longrightarrow Q$ is a transition function. A configuration $c \in Q^{\mathbb{Z}}$ of (Q,δ) is a function from $\mathbb{Z}$ into Q and its global transition function $G_\delta : Q^{\mathbb{Z}} \longrightarrow Q^{\mathbb{Z}}$ is such that $(G_\delta(c))(i) = \delta(c(i-r),...,c(i),...,c(i+r))$.*

Notation 1 *Let us define*

$$\delta^{m+2r\to m} \left| \begin{array}{l} Q^{m+2r} \longrightarrow Q^m \\ x \longmapsto y \end{array} \right.$$

such that for all $i \in \{1,...,m\}$, $y_i = \delta(x_{i-r},...,x_i,...,x_{i+r})$.

Definition 2. *An* **Elementary Cellular Automaton (ECA)** *is a radius-1 two states (usually* 0 *and* 1*) unidimensional cellular automaton.*

For ECAs, 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 $i = \overline{a_7a_6a_5a_4a_3a_2a_1a_0}^2$ satisfies:

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

Let us remark that CAs with different numbers may have the same behavior by switching 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 $\overline{r}$ the rule after exchanging the states and $\overleftrightarrow{r}$ the rule which has a symmetric behavior (see [5] for more details).

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

In the general definition of additive CAs due to Wolfram, an additive CA is a CA that satisfies the superposition principle ($\delta(x+x', y+y', z+z') = \delta(x,y,z) + \delta(x',y',z')$). These CAs are very interesting to provide examples because their behavior obey algebraic rules adapted to a formal study while their space-time diagrams appear complicated. We will use here, like in [13, 10], a more restrictive definition:

Definition 3. *We will call additive CA a unidimensional CA whose state set is* $\mathbb{Z}/n\mathbb{Z}$ *and whose transition function is of the form:*

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

1.2 Besicovitch topology

The most natural topology on CA configuration sets is the product topology. The problem is that this topology emphasizes what is happening closed to the origin while in many applications of CAs all the cells have the same importance. Thus, the adaptation of the mathematical notions of chaos to CAs for the product topology are not adapted: the shift is necessarily chaotic, that is not adapted to car traffic simulation for example. To propose more satisfying definitions of chaos, Formenti introduced Besicovitch pseudo-distance, that induce a shift invariant topology on the quotiented space:

Definition 4. *The Besicovitch pseudo-metric on* $Q^{\mathbb{Z}}$ *is given by*

$$d(c, c') = \limsup_{l \longrightarrow +\infty} \frac{\#\{i \in [-l, l] | x_i \neq y_i\}}{2l+1} \quad \textit{where \#denotes the cardinality.}$$

Property 1. $Q^{\mathbb{Z}}$ quotiented by the relation $x \sim y \iff d(x,y) = 0$ with Besicovitch topology is metric, path-wise connected, infinite dimensional, complete, neither separable nor locally compact [2]. Furthermore, $x \sim y \Longrightarrow G_\delta(x) \sim G_\delta(y)$ and the transition function of a CA is a continuous map from $Q^{\mathbb{Z}}/\sim$ into itself.

Remark 1. Actually, the results of this paper are not specific to Besicovitch topology, but are true for a wide class of topologies including Weil one. A general study of Besicovitch like topologies has been done in [9] and an interesting question would be to determine those that behave like Besicovitch one and what happens for the other ones. The only reason we point this out here is that there are many ways to extend Besicovitch in higher dimensional grids: the extention of Besicovitch pseudo-metric on $Q^{\mathbb{Z}^n}$ is

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

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 $\overline{d}_1(a_1, ..., a_n) = |a_1| + ... + |a_n|$ or $\overline{d}_2(a_1, ..., a_n) = (a_1^2 + ... + a_n^2)^{1/2}$. Of course different distances give different topologies but all of them are equivalent for our purpose because they differ on a null measure set. This is the only difficulty to extend the unidimensional concept to higher dimensional CAs because the definitions of measures and ergodicity given in the following section exist for any dimensional space.

1.3 Measure on the configuration set

Notation 2 *Let Q be a finite alphabet with at least two letters. $Q^+ = \cup_{n \geq 1} Q^n$ is the set of finite words on Q. The i^{th} coordinate $x(i)$ of a point $x \in Q^{\mathbb{Z}}$ will also be denoted x_i and $x_{[j,k]} = x_j ... x_k \in Q^{k+1-j}$ is the segment of x between indices j and k. The cylinder of $u \in Q^p$ at position $k \in \mathbb{Z}$ is the set*

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

Let σ be the shift toward the left: $\sigma(c)_i = c_{i+1}$ (i.e. the rule number 85).

A Borel probability measure is a nonnegative function μ defined on Borel sets. It is given by its values on cylinders, satisfies $\mu(Q^{\mathbb{Z}}) = 1$, and for every $u \in Q^+$, $k \in \mathbb{Z}$,

$$\sum_{q \in Q} \mu[uq]_k = \mu[u]_k \text{ and } \sum_{q \in Q} \mu[qu]_k = \mu[u]_{k+1}$$

Definition 5. *A measure μ is σ-invariant if $\mu[u]_k$ does not depend on k (and will thus be denoted $\mu[u]$).*

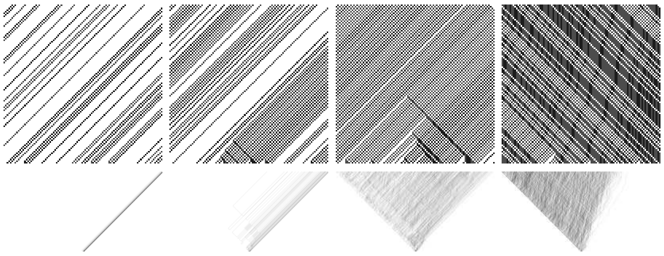
Definition 6. *A σ-invariant measure is σ-ergodic if for every invariant measurable set Y ($\sigma(Y) = Y$), either $\mu(Y) = 0$ or $\mu(Y) = 1$.*

Bernoulli measures are the most simple, this is the reason why we will use them in all our examples while the definitions and probably most of the theorems remain true for other σ-ergodic 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-ergodic measures. But, practically, these other measures will often be Bernoulli measures after a grouping operation.

Definition 7. *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* $u = u_0...u_{n-1} \in Q^n$, $\mu[u_0...u_{n-1}] = p_{u_0}...p_{u_{n-1}}$.

We will use the following classical result: *the Bernoulli measures are σ-ergodic.*

For 2-states CAs, the Bernoulli measures will be denoted μ_ρ where $\rho = p_1 = 1 - p_0$ is the probability for a state to be 1.



$\rho = 0.2$ $\rho = 0.4$ $\rho = 0.5$ $\rho = 0.7$

Fig. 1. The CA T is a very simple traffic model based on the rule 184 but with two different models of cars. The system seems "chaotic" when the density ρ of cars is greater than or equal to 0.5 because of the traffic jams, but not "chaotic" else. 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.

On the figure 1, we see a very simple example of CA that changes of behavior depending on the density of cars on the railway. 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 since 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 else. 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 to modelize a non uniform repartition.

2 Some classification tools on CAs

2.1 Damage spreading

Inspired by Lyapunov exponents [1], we will 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, 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 2 modifications like in the Lyapunov exponents, we count 0 modification because the state did not change. It appears that the rule 210 (see figure 3) 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 8. *Let μ be a σ-ergodic measure on the set of configurations, let $\mathcal{A}$ be a CA and $\mathcal{C}_\mathcal{A}$ its configuration set. If $c \in \mathcal{C}_\mathcal{A}$, we will define*

$$c_{p\leftarrow s} \left| \begin{array}{l} \mathbb{Z} \longrightarrow Q_\mathcal{A} \\ x \longmapsto \begin{cases} c(x) & \text{if } x \neq p \\ s & \text{else} \end{cases} \end{array} \right.$$

that is the configuration whose state of the cell p is changed to s. Let us now define the dependence coefficients $\alpha_{\mu,\mathcal{A},t,p}$ 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 π_q. Formally:

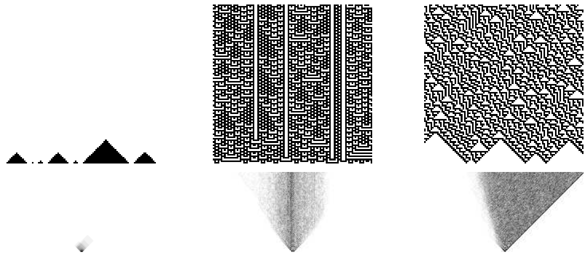
$$\alpha_{\mu,\mathcal{A},t,p} = \sum_{q\in Q_\mathcal{A}} \pi_q \mu(\{c \in \mathcal{C}_\mathcal{A} \mid (G^t_{\delta_\mathcal{A}}(c))_0 \neq (G^t_{\delta_\mathcal{A}}(c_{p\leftarrow q}))_0)$$

Remark 2.
- When no confusion about the measure and the cellular automaton is possible, we will simply denote the dependence coefficients $\alpha_{t,p}$.
- If r is the radius of the CA, $p > r \times t \implies \alpha_{t,p} = 0$ because the cell 0 state after t computation steps is independent of the cell p of the initial configuration.

Definition 9. *The damage spreading of a cellular automaton $\mathcal{A}$ according to a measure μ is the infinite sequence of positive real number*

$$\left(\sum_{p\in\mathbb{Z}} \alpha_{\mu,\mathcal{A},t,p}\right)_{t\in\mathbb{N}}$$

This sequence is well defined for all t thanks to the previous remark: the $\alpha_{\mu,\mathcal{A},t,p}$ are almost all equal to zero.



128 109 30

Fig. 2. The rule 128 belongs to $[\mathcal{D} \overset{\mu}{\rightarrow}^{+} 0]$, 109 to $[\mathcal{D} \overset{\mu}{\rightarrow}^{+} a]$ and 30 to $[\mathcal{D} \overset{\mu}{\rightarrow}^{+} +\infty]$. The bottom diagrams represent with grey level the probability of each cell to be affected by the modification of the middle cell.

This notion allows to define the class $[\mathcal{D} \overset{\mu}{\rightarrow}^{+} 0]$ of CAs whose damage spreading tends to zero, the class $[\mathcal{D} \overset{\mu}{\rightarrow}^{+} +\infty]$ of CAs whose damage spreading is not bounded (its limit sup tends to $+\infty$) and the class $[\mathcal{D} \overset{\mu}{\rightarrow}^{+} a]$ when the damage spreading limit sup tends to a finite non zero value. The figure 2 shows an example in each class. Obviously, these 3 classes define a partition of the set of CAs.

Theorem 1. *The additive CAs have non bounded damage spreading, they are in* $[\mathcal{D} \overset{\mu}{\rightarrow}^{+} +\infty]$ *(see [11] for the proof).*

2.2 μ-almost everywhere sensitivity to initial conditions

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

Definition 10. *A CA 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 $d(G^n_{\delta_\mathcal{A}}(c), G^n_{\delta_\mathcal{A}}(c')) \geq M$.*

The main reason of the introduction of μ-almost everywhere sensitivity is the study of some particular cases. On the one hand, the rule 120 appears sensitive to initial conditions but it seems that there exist some very artificial configurations that stop all the information transfer so that actually the rule 120 is not sensitive. On the other hand, the rule 210 does not appear to be sensitive, but is not equicontinuous (i.e. is sensitive to initial conditions on a subset of its configuration set) because of the configurations 0^* (see figure 3). Thus, in the same class, the elements of which are neither sensitive nor equicontinuous, we have two rules with very different behaviors. The idea is to say that 120 is almost everywhere sensitive, while 210 is almost never sensitive.

To define the almost everywhere sensitivity to initial conditions, we could just replace “for all configurations c” by “for μ-almost all configurations c” in the sensitivity definition. We will give a more restrictive (because of the first point of the next remark) definition so that a CA that is not μ-almost everywhere sensitive, is "μ-almost never" sensitive to initial conditions (see the third point of the next remark). Furthermore, because of the kind of proof we want to do, it is not more difficult to prove the μ-almost everywhere sensitivity for this definition.

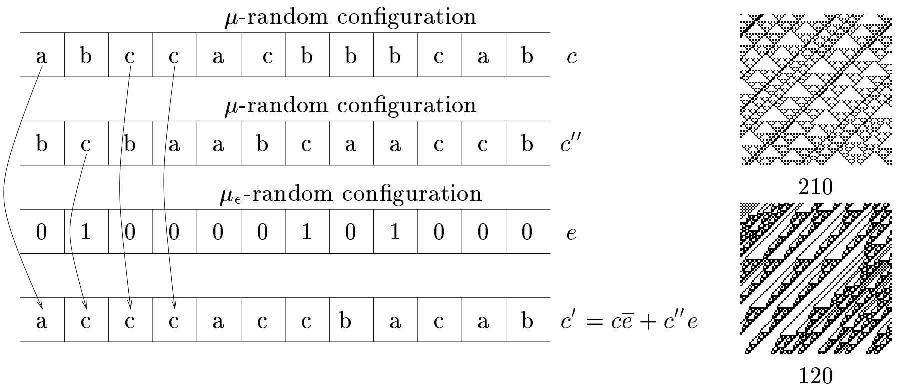


Fig. 3. The configuration $c' = c\overline{e} + c''e$ is the configuration whose state at a given position is equal to the corresponding state of c when the corresponding state of e is equal to 0 and to the corresponding state of c'' else. Let us remark that, due to the great number law, with probability 1, $d(c, c') = \epsilon d(c, c'') \leq \epsilon$.

Definition 11. *A CA is μ-almost everywhere sensitive to initial conditions (for Besicovitch pseudo-distance) if there exists $M > 0$ such that for all $\epsilon_0 > 0$, there exists $\epsilon < \epsilon_0$ such that if c and c'' are two μ-random configurations, if e is a μ_ϵ random configuration and if $c' = c\overline{e} + c''e$ 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 1 and to the corresponding state of c'' else (see figure 3), then with probability 1 (for $\mu\times\mu\times\mu_\epsilon$) there exists n such that $d(G^n_{\delta_\mathcal{A}}(c), G^n_{\delta_\mathcal{A}}(c')) \geq M$.*

Remark 3.
- 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 $d(G^n_{\delta_\mathcal{A}}(c), G^n_{\delta_\mathcal{A}}(c')) \geq M$;
- With 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 open sets have a non null measure), but it is not the case on $Q^\mathbb{Z}/\sim$ with Besicovitch topology;
- The set of configurations 3-uplets (c, c'', e) such that if $c' = c\overline{e} + c''e$ there exists n so that $d(G^n_{\delta_\mathcal{A}}(c), G^n_{\delta_\mathcal{A}}(c')) \geq M$ is obviously shift invariant on $(Q \times Q \times \{0,1\})^\mathbb{Z}$. As $\mu\times\mu\times\mu_\epsilon$ is σ-ergodic, thus the set measure is either 1 or 0. So a CA is either μ-almost everywhere sensitive to initial conditions or "μ-almost never sensitive to initial conditions": for any η there exists ϵ such that

if we build c, c' as usual, for any n, $d(G^n_{\delta_{\mathcal{A}}}(c), G^n_{\delta_{\mathcal{A}}}(c')) \leq \eta$ $\mu \times \mu \times \mu_\epsilon$-almost everywhere.

The μ-almost everywhere sensitivity to initial conditions makes sense because we saw that some CAs are not (obviously the rule 0 is not) and we will prove that the additive CAs are:

Theorem 2. *The additive CAs are μ-almost everywhere sensitive to initial conditions for any non trivial Bernoulli measure μ (see [11] for the proof).*

μ-attracting sets μ-attracting sets have been defined in [10]. In this article, P. Kůrka and A. Maass study the links between attracting and μ-attracting sets for different topologies.

Definition 12. *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 factors 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 13. *For a SFT $\Sigma \subseteq Q^{\mathbb{Z}}$ of order p and $x \in Q^{\mathbb{Z}}$, define the density of Σ-defects in x by*

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

Notation 3 *When d is a pseudo-distance, we can naturally define the pseudo-distance from an element x to a set as the inf 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 generally $d_D(x, \{y\}) \neq d_D(y, \{x\})$.

Definition 14. *Let μ be σ-ergodic measure, a sub-shift Σ is $B\mu$-attracting (resp. $D\mu$-attracting) if*

$$\mu(\{c \in Q^{\mathbb{Z}} \mid \lim_{n \to +\infty} \mathbf{d}(G^n_\delta(c), \Sigma) = 0\}) = 1$$

when $\mathbf{d} = d(= d_B), d_D$ respectively.

Remark 4.
- As the set $\{c \in Q^{\mathbb{Z}} \mid lim_{n \to +\infty} \mathbf{d}(f^n(c), \Sigma) = 0\}$ is shift invariant, its measure is either 0 or 1.
- From a Bμ-attracting sub-shift, we can easily extract a minimal Bμ-attracting sub-shift by considering the configurations c' of Σ such that $\mu(\{c \in Q^{\mathbb{Z}} \mid \liminf d_B(f^n(c), c') = 0\}) = 1$. And when μ is a Bernoulli measure, this implies that c' is uniform (i.e. of the form q^* for $q \in Q$).

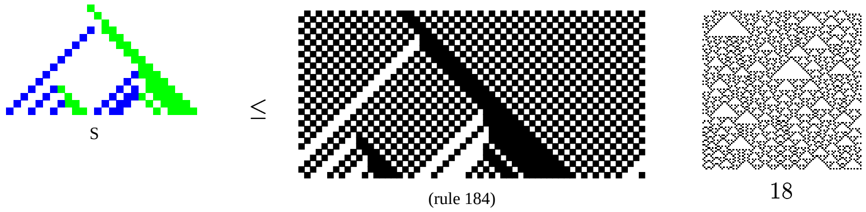


Fig. 4. $\{B^*\}$ is a Bμ-attracting (thus Dμ-attracting) set of the CA $\mathcal{S}$ which has 3 states, one is going to the right, one is going to the left (the third one is B, the blank one, into which the other state may move) and when they meet, both are annihilated. When μ is a measure such that the number 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 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 [12]), we have the same result on 184^2 for the measure so that $(1,0)$ has probability 0 and the probability 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 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 seems to be another example of CA with Dμ-attracting sets, but no Bμ-attracting set.

– Furthermore, if a sub-shift is Bμ-attracting, then it is Dμ-attracting.

The definition of Bμ-attracting sets is very natural: a set is Bμ-attracting if from almost all configurations (w.r.t. μ), the (Besicovitch) pseudo-distance between the successive configurations and the sub-shift tends to 0. We saw in the remark that in this case, almost all evolutions tend to uniform configurations when μ is a Bernoulli measure. The rule 128 (see figure 2) or the CA $\mathcal{S}$ and 184^2 for some measure μ (see figure 4) have a Bμ-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, Dμ-attracting sets are more powerful and allow to point out homogenization process to periodical configurations and not only to uniform configurations. For instance, as proved in [10], the sub-shift $\{(01)^*,(10)^*\}$ is B$\mu_{1/2}$-attracting for the rule 184. As noticed in [10], 18 seems to be an example of CA with a Dμ-attracting set which is not of finite type: $\{c\in\{0,1\}^{\mathbb{Z}}|\forall i, c(2i)=0\}\cup\{c\in\{0,1\}^{\mathbb{Z}}|\forall i, c(2i+1)=0\}$.

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

Definition 15. *Let $(\mathcal{A}, \delta)$ be a CA and μ a σ-ergodic measure, the metric entropy of its configuration after t computation steps is defined as follow:*

$$S_\mu^{(t)}(\mathcal{A}) = \lim_{n \to +\infty} \frac{-\sum_{u \in Q^n} p_u^{(t)} ln(p_u^{(t)})}{n}$$

with the usual convention $0 \times log(0) = 0$ and where p_u is the probability of apparition of the pattern u in the configuration c:

$$p_u^{(t)} = G_\delta^t(\mu)([u]_0)$$

where the notation $f(\mu)$ represents, as usual, the measure defined by $f(\mu)(X) = \mu(f^{-1}(X))$. In mathematical terminology, $S_\mu^{(t)}(\mathcal{A})$ is the metric entropy of σ for the measure $G_\delta^t(\mu)$.

Definition 16. *The class of CAs that erase all their initial configuration information will be denoted $[S_\mu \to 0]$ and formally defined as follows: a CA is in $[S_\mu \to 0]$ if and only if*

$$S_\mu^{(t)}(\mathcal{A}) \longrightarrow_{t \to +\infty} 0.$$

Theorem 3. *If a CA has a Dμ-attracting set of null topological entropy, then it is in $[S_\mu \to 0]$ (see [11] for the proof).*

Definition 17. *The topological entropy of a sub-shift Σ is*

$$S_\Sigma(\sigma) = \lim_{n \longrightarrow +\infty} \frac{ln(\#\{u \in Q^n \mid \exists c \in \Sigma, c_{[0,|u|]} = u\})}{n}.$$

3 Relations between damage spreading, μ-ae sensitiveness and the existence of a Bμ-attracting set

In this section, we always assume that μ is a Bernoulli measure.

Theorem 4. *If $\mathcal{A}$ μ-damage spreading tends to 0 then there exists a set of uniform configurations which is Bμ-attracting (see [11] for the proof).*

Reciprocally, there are CAs with Bμ-attracting sets that are not in $[\mathcal{D} \overset{\mu}{\to}^+ 0]$: the rule $\mathcal{S}$ (see figure 4) with 3 states (a blank one B into which one state l goes to the left and a state r goes to the right, the collision of two particles leads to their annihilation) is in $[\mathcal{D} \overset{\mu}{\to}^+ a]$. 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-almost everywhere sensitiveness. The idea is to take a μ-random configuration c. We then build c' from c: for all the cells, the state of c' is equal to the state of c with probability $1 - \eta$, is equal to q with probability ηp_q. With probability 1, the Besicovitch pseudo-distance $d(c, c')$ between c and c' is $\eta d(c, c'') \leq \eta$. So we can prove the following theorem.

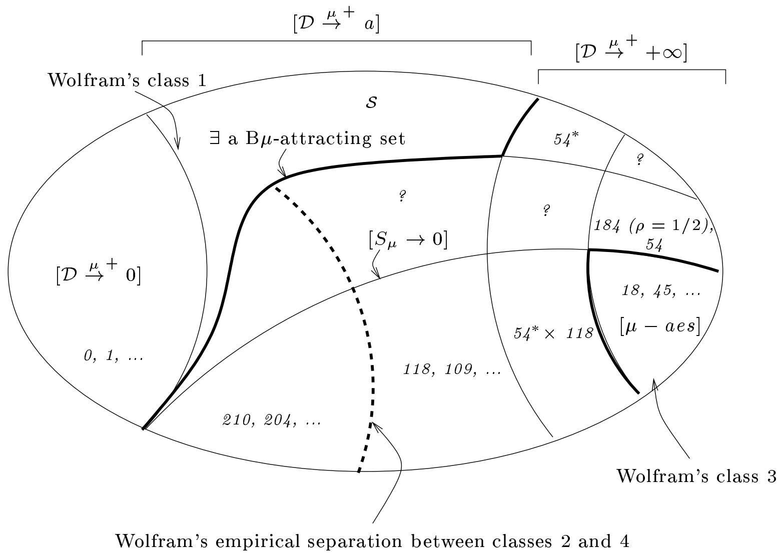


Fig. 5. Relations between damage spreading, μ-ae sensitiveness and the existence of Bμ-attracting sets when μ is a Bernoulli measure

Theorem 5. *Being μ-almost everywhere sensitive to initial conditions implies to have non bounded damage spreading (i.e. $[\mu - aes] \subseteq [\mathcal{D} \overset{\mu}{\rightarrow}^{+} +\infty]$) (see [11] for the proof).*

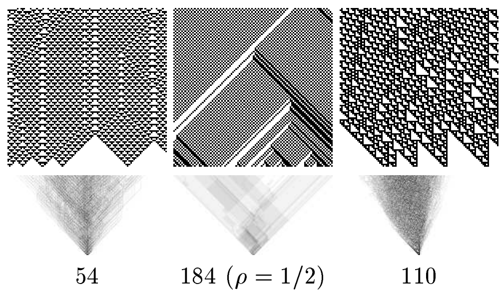


Fig. 6. 54 and 184 for $\mu_{1/2}$ are in $[\mathcal{D} \overset{\mu}{\rightarrow}^{+} +\infty] \cap [S_\mu \to 0]$, we do not know where is 110

It is experimentally easy to observe but it seems difficult to prove that 184 is $\mu_{1/2}$-almost everywhere sensitive to initial conditions. To understand what happens, let us consider a random walker on $\mathbb{Z}$ starting at 0 and reading the initial configuration from the cell 0 toward the right (resp. left). When he reads

1 he goes toward the right and when he reads 0, toward the left. If we change the cell 0 to 1 (resp. to 0), it generates at least one modification on all the configurations. This modification moves to the right when the random walker is on $\mathbb{Z}_+^*$, to the left when it is in $\mathbb{Z}_-^*$ and do not move if the random walker is on 0. Sometimes, because of this modification, a whole region of background shifts and all the cells of this region are changed. 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 $1/2$.

It seems that the rule 54 also belongs to $[\mu - aes] \cap [S_\mu \to 0]$. Actually, g_e particles disappearance is an irreversible phenomenon as proved in [12] that 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 [3, 8] experiments. 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 $0 < \rho < 1$.

The previous theorem raises a natural question: are the CAs with unbounded damage spreading μ-almost everywhere sensitive to initial conditions?

It seems that 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 [8] 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, 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 almost everywhere sensitive to initial conditions because of the following theorem:

Theorem 6. *Let A be a CA that almost everywhere tends to a uniform configuration, A is not almost everywhere sensitive to initial conditions (see [11] for the proof).*

The question to know whether there are almost everywhere sensitive CAs with a $\mathrm{B}\mu$-attracting set is open.

4 Links with Wolfram's classification

The CAs in $[\mu - aes] \backslash [S_\mu \to 0]$ practically match on ECAs with the CAs that Wolfram put in his class 3, they are represented on the figure 7. It is not sure that this remains true for more complicated (with more states, in higher dimension

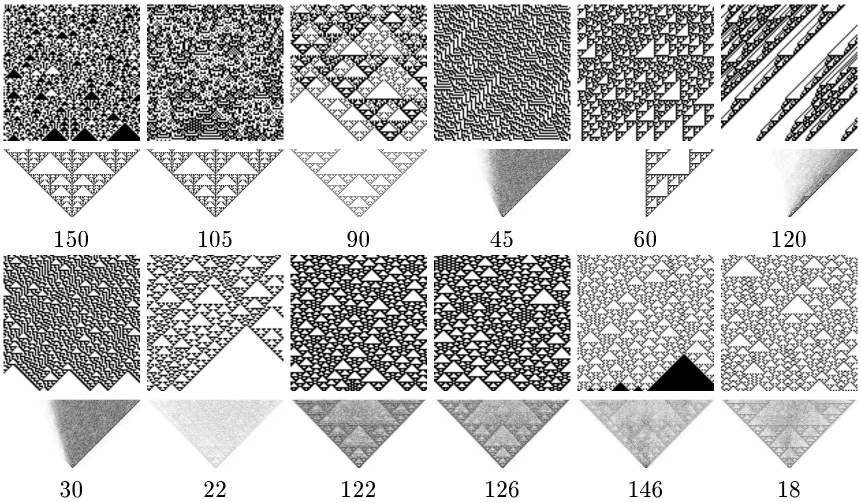


Fig. 7. All the chaotic behaviors among 2 states unidimensional CAs

or with a bigger neighborhood) CAs because they may present a lot of different behaviors depending on the initial configuration. Anyway, if we assume that this is a good formalization, Wolfram's observation that "the value of a particular site depends on an ever-increasing number of initial site values" is proved. Furthermore, we know that this condition is not sufficient to imply chaoticity.

Wolfram's class 4 on ECAs seems to split into two parts, the class $[\mu - aes] \cap [S_\mu \to 0]$ and a part of the class $[\mathcal{D} \overset{\mu}{\to}{}^{+} a] \backslash [S_\mu \to 0]$. But the definition of this class (existence of particles) or the conjecture of the universality of its elements let us think that this class is completely independent of our criteria. The main point is that it is easy to build universal CAs in any non empty class since there are some in the class $[\mathcal{D} \overset{\mu}{\to}{}^{+} 0]$ (a CA is usually universal on a null measure set).

The evolution of CAs that have Bμ-attracting sets tends μ_ρ-almost everywhere to a set of uniform configurations. Their behaviors look like the behaviors of Wolfram's class 1 CAs that "evolve after a finite number of time steps from almost all initial states to a unique homogeneous state, in which all sites have the same value". Actually, it seems natural to think that the rule 128 (see figure 2) which erases a succession of n states 1 in $n/2$ time steps is in the class 1 and this is confirmed by the examples of class 1 CAs given by Wolfram. But, if 0^* is obviously a Bμ-attracting set, with probability 1, the evolution does not converge to 0^* after a finite number of time. The point is that the probability to find a sequence of n successive 1 on the configuration is 1 whatever n. Next, Wolfram writes that "their evolution completely destroys any information on the initial state". We proved this fact for CAs with Bμ-attracting sets, but we also saw that complicated rules like 184 for $\rho = 1/2$ completely destroy any information on the initial state.

The big problem would be to find a way to split the CAs whose damage spreading is bounded but does not tend to 0 in such a way that rules like 118 or 109 are not together with the identity. If the behavior of the damage growth of a CA is almost independent of the measure we take, we can separate the CAs whose damage spreading is uniformly bounded from the others. This would separate the identity from 118 and 109, but the rule 210 would be in the second case, that was not really expected.

Conclusion and open questions

One of the main conclusion of this article is that our intuitive property of chaos does not allow to split the set of CAs into two classes because some of them may have a chaotic or non chaotic behavior depending on the way to choose a random configuration. In addition, we see that Besicovitch topology that has been specifically introduced in CAs to express an intuitive notion of chaos is effectively a very interesting notion. Note that the introduction of a measure is very helpful to deal with the too wide class of neither sensitive nor equicontinuous CAs. In this article, we introduce a new Lyapunov like notion that allows to measure information diffusion. This notion appears more precise than the other ones but still not enough to ensure a chaotic behavior. Finally, the introduced notions allow to formalize some of Wolfram's observations and thus to prove how relevant they are.

A lot of open questions remains, among them

- to find examples in or to prove the emptiness of $[\mu - aes]$ with a Bμ-attracting set and of CAs in $[S_\mu \to 0]$ with no Bμ-attracting set and not in $[\mu - aes]$.
- the generalization for more general measures (with exponentially decreasing correlations) of the theorems.
- to find a "good" definition that separates the identity from 118 in $[\mathcal{D} \overset{\mu}{\to}{}^{+} a]$.

References

1. F. Bagnoli, R. Rechtman, and S. Ruffo. Lyapunov exponents for cellular automata. In M. Lpez de Haro and C. Varea, editors, *Lectures on Thermodynamics ans Statistical Mechanics*. World Scientific, 1994.
2. F. Blanchard, E. Formenti, and P. Kůrka. Cellular automata in the Cantor, Besicovitch and Weil topological spaces. *Complex Systems*, to appear 1999.
3. 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. *Physical Review A*, 44:866–875, 1991.
4. C. Cattaneo, E. Formenti, L. Margara, and J. Mazoyer. Shift invariant distance on $s^{\mathbb{Z}}$ with non trivial topology. In *Proceedings of MFCS' 97*. Springer Verlag, 1997.

5. G. Cattaneo, E. Formenti, L. Margara, and G. Mauri. Transformation of the one-dimensional cellular automata rule space. *Parallel Computing*, 23:1593–1611, 1997.
6. G. Cattaneo, E. Formenti, L. Margara, and G. Mauri. On the dynamical behavior of chaotic cellular automata. *Theoretical Computer Science*, 217:31–51, 1999.
7. R. H. Gilman. Classes of linear automata. *Ergod. Th. & Dynam. Sys.*, 7:105–118, 1987.
8. J. E. Hanson and J. P. Crutchfield. Computational mechanics of cellular automata: an example. *Physica D*, 103:169–189, 1997.
9. V. Kanovei and M. Reeken. On Ulam's problem of approximation of non-exact homomorphisms. *preprint*, 2000.
10. P. Kůrka and A. Maass. Stability of subshifts in cellular automata. *preprint*, 1999.
11. B. Martin. Damage spreading and μ-sensitivity on CA - extended to proofs. Research Report RR2000-04, LIP, ENS Lyon, 46 allée d'Italie - 69 364 Lyon Cedex 07, 2000. `ftp://ftp.ens-lyon.fr/pub/LIP/Rapports/RR/RR2000/RR2000-04.ps.Z`.
12. B. Martin. A group interpretation of particles generated by one dimensional cellular automaton, 54 wolfram's rule. *Int. Journ. of Mod. Phys. C*, 11(1):101–123, 2000.
13. J. Mazoyer and I. Rapaport. Inducing an order on cellular automata by a grouping operation. In *STACS'98*, volume 1373, pages 116–127. Lecture Notes in Computer Science, 1998.
14. K. Čulik, J. Pach, and S. Yu. On the limit set of cellular automata. *SIAM Journal on Computing*, 18:831–842, 1989.
15. S. Wolfram. Universality and complexity in cellular automata. *Physica D*, 10:1–35, 1984.