Cellular automata:
dynamics, simulations, traces

Pierre Guillon

September 2008

1 Introduction

Understanding the emergence of complexity out of simple basic elements is a fundamental issue in various scientific fields: cellular biology, cognitive science, fluid mechanics, chemical turbulences, crystal formation, social dynamics, computer networks... These problems, abstracted from their particular modeled systems, were joined into what is now called the theory of complex systems.

They led John von Neumann, motivated by the autoreproducibility question and inspired by Stanisław Ulam, to define the first cellular automaton in the late forties. Merely formalized as a discrete space divided into cells whose states evolve in a discrete time according to their closest neighbors, it exhibits strange evolutions, such as patterns reproducing themselves indefinitely. This duality was popularized in the seventies by John Conway's game of life. The emergence of computers would soon allow anyone to program it, and nevertheless admit that the overall behavior could turn out to be very complex, motivating Stephen Wolfram's classification of the visual aspects of cellular automata in the eighties.

But what does one exactly mean by complexity? This notion was the subject of many formalization attempts. First, the computing power, already suggested by von Neumann, was formalized in terms of Turing-equivalence, for instance for the game of life [BCG82], and gave rise to some algorithmic issues [Fis65], for language recognition [SI72], or for very peculiar problems [Moo64], which give evidence on what kind of processes could be performed with the model.

Moreover, the computing power has also been studied for cellular automata with respect to each other. This approach has led to the notions of "cellular" simulations and "intrinsic" universality, whose premises could be seen in [Ban70, AČI87], before formalizations in [MI94] and mostly [Rap97, Oll02, The05]. The orders induced by the different kinds of simulations are still not very well understood.

On the other hand, cellular automata have been studied in terms of predictability with respect to other computing models – Turing machines. Most of their long-term properties have been proved undecidable since the works of Jarkko Kari in the nineties [Kar90].

On top of that, cellular automata have joined the theory of dynamical systems thanks to the 1969 characterization by Hedlund, Curtis and Lyndon [Hed69] as continuous maps over configurations that commute with the translations of cells, the configuration space being endowed with the Cantor topology that makes it perfect, compact and totally disconnected. This branch of the study has resulted in many contributions, involving equicontinuity [Gil87], attractors [Hur90], measure [Ish92]....

In 1997, Petr Kůrka proposes the modifications of two topological classifications in [Ků97], and compares them both with a third one, based on the sequences of states that are successively taken by some single cell – or some finite group of cell – during the evolution of the cellular automaton. The principle of studying such "traces" of dynamical systems through a given partition of the space probably sprang from the study of geodesic streams by Hadamard at the end of the nineteenth century, and was given its name of "symbolic dynamics" by the eponymous book by Morse and Hedlund [MH38]. In the dawn of the Internet era, a reference book by Douglas Lind and Brian Marcus [LM95] stressed that symbolic dynamics find its most promising applications in code theory.

If one sees the cellular automaton as modeling some physical phenomenon, then the letter sequences it produces may represent the measure of the phenomenon through some device with some precision. Hence, it seems relevant to study the behavior of these trace systems according to that of the global one. Topologically speaking, they are linked by a factorization, *i.e.* reading a letter of the infinite word corresponds to applying one step of the cellular automaton. To each trace can be associated the language of its finite patterns. It was noted in [Gil88] that these languages were always context-sensitive; their complexity led to the Kůrka classification, that can be applied to any system on Cantor sets. Nevertheless, the restriction to one-dimensional cellular automata presents an important advantage, stressed in [dL06a]: studying the trace with respect to neighbors of the center cell gives significant information about the cellular automaton itself. In particular, the regularity of the language associated allows to decide many long-term properties [dL06b].

Our thesis is in line with the above-mentioned works on the link between cellular automata and symbolic dynamics. We try to inspect which properties of cellular automata are transmitted to simulated cellular automata, to traces, and conversely. We also try to see if one single trace, for instance corresponding to neighbors of the central cell, is enough to deduce the corresponding property, in the general case, or in some subcases from the symbolic classification. In a second part, we investigate sufficient conditions for some set of infinite letters to be the trace of some cellular automaton. Physically speaking, we try to find back whether some phenomenon can correspond to some observation we made in the model. Finally, a third part is dedicated to decidability issues; in particular, it is shown that, being given a cellular automaton, nothing can be said about the long-term aspect of its trace.

2 Dynamical systems and cellular automata

A **dynamical system** (DS) is a couple (X, F) where X is a *compact metric* space and $F : X \to X$ a *continuous* function. A **symbolic system** (SS) is a dynamical system on some *totally disconnected* space, *i.e.* some space that admits arbitrarily thin clopen partitions.

Configurations space. If $X = A^{\mathbb{M}}$, where A is a finite alphabet and $\mathbb{M} = \mathbb{Z}$ or $\mathbb{M} = \mathbb{N}$, one can endow X with the product of the discrete topology on A, which corresponds to the distance:

$$\mathbb{d} : \begin{array}{ccc} A^{\mathbb{M}} \times A^{\mathbb{M}} & \to & \mathbb{R}_+ \\ (x,y) & \mapsto & 2^{-\min_{x_i \neq y_i} |i|} \end{array} .$$

Note $\langle k \rangle = \{ i | \, |i| \leq k \}$ and, for $x \in A^{\mathbb{M}}$ and $k, l \in \mathbb{M}$ with $k \leq l$, $x_{[\![k,l[\![} = x_k \dots x_{l-1}$. **Cylinders** $[u] = \left\{ x \middle| \, x_{\langle k \rangle} = u \right\}$, where $u \in A^{\langle k \rangle}$ and $k \in \mathbb{N}$, form a base of clopens.

Shifts. The **shift** is a DS defined for $x \in A^{\mathbb{M}}$ and $i \in A^{\mathbb{M}}$ by $\sigma(x)_i = x_{i+1}$. A **onesided (resp. twosided) subshift** is a closed σ-invariant (resp. strongly) subset of $A^{\mathbb{N}}$ (resp. $A^{\mathbb{Z}}$). A subshift Σ is **sofic** if its language $\left\{ u \in A^+ \middle| \, \exists x \in \Sigma, i \in \mathbb{M}, x_{[\![i,i+|u|[\![} = u \right\}$ is *regular*. It is of **finite type** (SFT) if it admits some *finite* **forbidden language** $F \subset A^+$ such that $\Sigma = \left\{ z \in A^{\mathbb{M}} \middle| \, \forall i \in \mathbb{M}, u \in F, x_{[\![i,i+|u|[\![} \neq u \right\}$. It is of **order** $k \in \mathbb{N}^*$ if $F \subset A^k$.

Morphisms. A **morphism** between two DS (X, F) and (Y, G) is a function $\Phi : X \to Y$ such that $\Phi F = G\Phi$. If surjective, it is a **factorization** and (Y, G) is a **factor** of (X, F); if bijective, it is a **conjugacy** and (X, F) and (Y, G) are **conjugate**. A **cellular morphism** between two CA restrictions (Λ, F) and (Σ, G) is a *morphism* which is simultaneously a morphism between subshifts iterates (Λ, σ^n) and $(\Sigma, \sigma^{n'})$ for some $n, n' \in \mathbb{N}^*$. Similarly, we will speak of **cellular factorizations**, **conjugacies**, **factors**, **conjugates**.

Cellular automata. A **cellular automaton** (CA) is a morphism F of some shift $(A^{\mathbb{M}}, \sigma)$ into itself. Equivalently, from the Hedlund theorem, there is some **radius** $r \in \mathbb{N}$ and some **local rule** $f : A^{\langle r \rangle} \to A$ such that for all configurations $x \in A^{\mathbb{M}}$ and all cells $i \in \mathbb{M}$, $F(x)_i = f(x_{i+\langle r \rangle})$; $d = |\langle r \rangle|$ will be called the **diameter**.

Simulations. A **simulation** between two DS (X, F) and (Y, G) is a *factorization* between the subsystem (X', F^m) and the iterate $(Y, G^{m'})$ for some invariant $X' \subset X$ and some $m, m' \in \mathbb{N}^*$. It is called **direct** if $m = 1$, **total** if $m' = 1$, **complete** if $X' = X$, **exact** if it is injective. A **cellular simulation** between two CA (X, F) and (Y, G) is a *simulation* which is simultaneously a simulation between the iterates of subshifts (Λ, σ^n) and $(\Sigma, \sigma^{n'})$ for some $n, n' \in \mathbb{N}^*$. It is **block** if its image is a full shift, **context-free** if its radius is 0.

Universality. A DS is **universal** with respect to some family of DS if it simulates any system of that family. This definition led to the study of intrinsically universal CA in [Oll02, The05]. In the context of sofic subshifts, we can prove that *universality* with respect to all subshifts is equivalent to *uncountability*, to *nonzero entropy*, and to existence of some *infinite transitive subsystem*. This class of subshifts will be especially prominent in section 5.

3 Traces

If $\mathcal{P}$ is a partition of some space X and $x \in X$ a point, then we denote $\mathcal{P}(x) \in \mathcal{P}$ the unique subset such that $x \in \mathcal{P}(x)$. The **trace** of some symbolic system (X, F) with respect to some clopen partition $\mathcal{P}$ is the function:

$$T_F^{\mathcal{P}} : \begin{array}{rcl} X & \to & \mathcal{P}^{\mathbb{N}} \\ x & \mapsto & (\mathcal{P}(F^j(x)))_{j\in\mathbb{N}} \end{array} .$$

It is a factorization of the system (X, F) into the **trace subshift** $(\tau_F^{\mathcal{P}} = T_F^{\mathcal{P}}(X), \sigma)$. Conversely, every *factorization* of the system into a subshift can be written as a *trace* application.

Column factors. If $X = A^{\mathbb{M}}$, we can restrict the study of *trace* subshifts to **column factors**, *i.e.* traces with respect to cylinder partitions, which represent the sequences of states taken by the central cell of a given configuration:

$$T_F^{\langle k\rangle} : \begin{array}{rcl} A^{\mathbb{M}} & \to & (A^{\langle k\rangle})^{\mathbb{N}} \\ x & \mapsto & (F^j(x)_{\langle k\rangle})_{j\in\mathbb{N}} \end{array} .$$

Let $\tau_F^{\langle k\rangle} = T_F^{\langle k\rangle}(A^{\mathbb{M}})$, τ_F denote $\tau_F^{\langle 0\rangle}$, and τ_F^* be the twosided subshift $\left\{ (x_0^j)_{j\in\mathbb{Z}} \middle| \forall j \in \mathbb{Z}, x^{j+1} = F(x^j) \right\}$, called the **bitrace**. All factor subshifts are factors of some column factor. Moreover, the SS can be essentially rebuilt from the sequence of its column factors, since it is conjugate to their **limit extension**, *i.e.* the least DS of which they are all factor.

Diametral factor. If F is a CA of radius $r \in \mathbb{N}$ and $k > r$, then by shift-invariance, $\tau_F^{\langle k\rangle}$ is the **overlap** of the **diametral factor** $\tau_F^{\langle r\rangle}$, *i.e.* the biggest subshift Σ on $A^{\langle k\rangle}$ such that each projection $\pi_{i+\langle r\rangle}(\Sigma) = \tau_F^{\langle r\rangle}$ (up to indices change), for $|i| \leq k - r$. This idea was first suggested in [dL06a]. This simple characterization of CA allows to rebuild from the diametral factor each wider trace, hence the CA itself. It can help to deduce some properties of the CA directly from that of its diametral factor.

Symbolic classification. The complexity of the trace subshifts can be used as a measure of the complexity of the SS itself. This was the purpose of Kůrka's classification, presented in [Ků97], and which can be refined a little by involving systems finite type. This class comprehends a little more that systems whose trace subshifts are of finite type, since it is invariant by conjugacy and a generalization of SFT. Its definition requires the following one.

We define a **base** of factor subshifts of a SS $(A^{\mathbb{M}}, F)$ as a sequence $(\Sigma_i)_{i\in\mathbb{N}}$ of onesided subshifts image of a family $(\Phi_i : X \to \Sigma_i)_{i\in\mathbb{N}}$ of factorizations. We have already seen that column factors form one. Now, we can define the following hierarchy in four classes.

1. **Equicontinuous SS**: all (or a base of) trace subshifts are finite. For CA, it is even equivalent to the finiteness of *some* trace subshift, thus to *preperiodicity*.

2. **SS of finite type**: a base of trace subshifts are *SFT*. This condition implies the *shadowing property*. We even suspect an equivalence, like in the case of subshifts. For CA, it is sufficient to have a *diametral factor* of *finite type*, but not necessary.

3. **Sofic SS**: all (or a base of) trace subshifts are sofic. For CA, it is equivalent to the soficness of the *diametral factor* (generalization of a result in [BM96]).

4. **All SS**: this class could also be refined, for instance according to well-know classifications on languages – if they appear to be relevant.

Similarly to the Weiss theorem for subshifts, we prove that any *sofic* SS is a *factor* of some SS of *finite type*. In the context of CA, as proved by di Lena, any CA restriction which is of *finite type* is a *cellular factor* of some *sofic* CA restriction. Nevertheless, it seems difficult to adapt the construction to get a (complete) CA.

4 Topological dynamics and traces

We inspect many behaviors of CA and DS and their consequences on their trace and on their simulated systems. We are interested in *immediate* properties (surjectivity, openness, injectivity), *simple* properties (nilpotency, periodicity and their non-uniform versions, nilpotency over periodic configurations), *transitivity-like* properties (mixingnesses, specification, nonwanderingness), *chain* properties (chain-transitivity, shadowing property), *equicontinuity* properties (almost equicontinuity, sensitivity), *expansivity* (and positive expansivity). We then inspect two characteristic sets: the *limit set* $\Omega_F = \bigcup_{j\in\mathbb{N}} F^j(X)$ and the *ultimate set* $\omega_F = \bigcup_{x\in X} \bigcap_{J\in\mathbb{N}} \overline{\{F^j(x) | j > J\}}$ of DS (X, F), from which are defined Ω-nilpotency, Ω-periodicity, ω-nilpotency and ω-nilpotency according to whether the restriction of F on each of these subsets is null or periodic, respectively. Note that the *trace subshift* $T_F(\Omega_F)$ of the *limit system* is the onesided subshift corresponding to the bitrace τ_F^*.

The questions we ask are:

- `Sim`: Is the property preserved by DS *simulation*; if not is it by some restricted simulation?
- `SimC`: Is the property preserved by CA *cellular simulation*?
- $\rightarrow \tau$: Is the property true for the *factor subshifts* of some SS if it is true for the SS itself?
- $\tau\tau \rightarrow$: Is the property true for some SS if it is true for all of its *factor subshifts*?
- $\tau \rightarrow$: Is the property true for some CA if it is true for *some particular column factor*?
- σ: Is the property preserved by CA composition with the *shift*?

We chose to separate the study of the two kinds of simulations in order to emphasize which were basic topological properties and what was the contribution of the cellular environment. It results in a better understanding of the needs in the restriction on the concept of cellular simulation, to make it the most relevantly powerful. We note of course that a positive answer to `Sim` implies one to its subcases `SimC` and $\rightarrow \tau$, but more can be expressed from the specific actions of CA and from the the particularity of the trace factorizations, respectively. Concerning question σ, most answers are well-known [Sab06, ADF07], but recalled here in order to inform, combined with `SimC`, about directional simulations used in [Oll02, The05], *i.e.* cellular simulation up to some shift iterate. The results summarized here correspond in a great part to paragraphs "…et simulation" (`Sim`) at the end of each concerned section in , "…et simulation" (`SimC`, σ) and "…et trace" ($\rightarrow \tau$, $\tau\tau \rightarrow$, $\tau \rightarrow$) at the end of each concerned section in .

	`Sim`	`SimC`	$\to \tau$	$\tau\tau \to$	$\tau \to$	σ
surjectivity	complete	block	✓	✓	✓	✓
injectivity	exact	context-free	$\perp$	✓	1	✓
openness	exact	context-free	$\perp$	$\perp$	$\perp$	✓
nilpotency, w.nilpotency	✓	✓	✓	$\perp$	1	✓
preperiodicity	✓	✓	✓	$\perp$	1	$\perp$
w.preperiodicity	✓	✓	✓	$\perp$	$\perp$	$\perp$
nilpotency/periodic	exact	✓	?	✓	1	✓
nonwanderingness	direct complete	block	✓	✓	✓	✓
transitivity, w.mixingness	direct complete	complete	✓	✓	?	$\perp$
mixingness	complete	complete	✓	✓	?	$\perp$
s.transitivity	direct complete	direct complete	✓	$\perp$	$\perp$	✓
specification	complete	complete	✓	$\perp$	$\perp$	✓
chain-transitivity	direct complete	complete	✓	✓	?	$\perp$
shadowing	exact complete	exact complete	$\perp$	✓	d	$\perp$
equicontinuity	exact	✓	✓	✓	1	$\perp$
almost equicontinuity	exact complete	context-free complete	$\perp$	✓	r	$\perp$
sensitivity	exact complete	exact complete	$\leq \varepsilon$	$\perp$	$\perp$	$\perp$
pos.expansivity	exact	exact	trivial	$\perp$	$\perp$	$\perp$
expansivity	exact	exact	$\perp$	$\perp$	$\perp$	$\perp$
Ω-nilpotency	✓	✓	✓	$\perp$	1	✓
Ω-periodicity	✓	✓	✓	$\perp$	$\perp$	$\perp$
ω-nilpotency	✓	✓	✓	✓	1	✓
ω-periodicity	✓	✓	✓	$\perp$	$\perp$	$\perp$
finite type	exact complete	exact complete	$\perp$	✓	d	$\perp$
soficness	complete	complete	✓	✓	d	$\perp$

Keys. ✓ stands for a positive answer, $\perp$ for a negative one. In columns `Sim` and `SimC`, we indicate the restrictions needed to get a positive answer, if any. In column $\tau \to$, we indicate the least (as far as we know) width of column factor which would imply that of the CA, r standing for its radius, d for its diameter, ✓ for some computable number in terms of d and $|A|$.

ω-nilpotency, which consists in the convergence of all orbits towards the same limit point, appears to be a very robust notion. Its equivalence, in the context of CA, with nilpotency, weak nilpotency, Ω-nilpotency, and trace nilpotency was published in [GR08]. They are many other questions rising, especially about ω-limit sets; for instance, can it be non-full for some surjective CA, or what does ω-periodicity really represent?

Most open questions in the table above concern implications of the properties of the diametral factor of CA; we can also mention, in that area, that it is unknown whether the topological entropy of CA is equal to that of some of their column factor, as it is the case for onesided or expansive CA. Another important class of issues, partly addressed in [The05], is to obtain minimal restrictions of cellular simulations that preserve a given property, such as immediate properties, expansivities, transitivities. Finally, we can notice a last question mark, isolated, in the third column, which can be expressed as follows: can a CA which is nilpotent over periodic configurations admit some periodic trace word?

5 Traceable subshifts

We have seen how some properties of CA could be transmitted to their trace. An interesting "reverse" problem would be to find an adequate CA being given a potential trace subshift. We can see that this is not possible for all subshifts. For instance, a trace subshift always contains some **deterministic** subshift, *i.e.* $\left\{ (\xi^j(a))_{j\in\mathbb{N}} \middle| a \in A \right\}$ for some function $\xi : A \to A$. This in not the case for subshifts like $\{(001)^\infty, (010)^\infty, (100)^\infty\}$.

In [CFG07], we inspected this **traceability** property for sofic subshifts, and reached some sufficient conditions: all *DDC SFT* and *universal DDC sofic* subshifts are *traceable* by some CA, where a **DDC** subshift stands for a subshift which contains some *deterministic* subshift $\left\{ (\xi^j(a))_{j\in\mathbb{N}} \middle| a \in A \right\}$ and some *periodic* word w^∞ such that $w \in A^* \setminus \xi(A)^*$. We give here a sketch of proof a little different from [CFG07], that will allow further results.

Polytraceability.

Any *SFT of order* 2 is *traceable* by some onesided CA.

We are interested in the subproblem of polytraceability. The **polytrace** of some CA F on some alphabet $B \subset A^k$ is the the union $\overset{\circ}{\tau}_F = \bigcup_{0\le l<k} \pi_l(\tau_F)$, where $\pi_l((z_i^j)_{\substack{0\le i<k\\ j\in\mathbb{N}}}) = (z_l^j)_{j\in\mathbb{N}}$.

Any *SFT* is the *polytrace* of some onesided CA.

To deal with more complex subshifts, we use the characterizations of universality. A subshift is **polyuniversal** if any other subshift is factor of some subshift $\Sigma' \subset \Sigma$.

Any *universal sofic* subshift Σ is *polyuniversal*.

Any *universal sofic* subshift is the **polytrace** of some onesided CA.

Partial traces.

Now it remains to simulate our CA on $B \subset A^h$ by some CA on A, transforming the polytrace into a trace. First we are looking for some CA restriction to some SFT (not over the whole space $A^{\mathbb{Z}}$). This is possible with a restriction. A subset $B \subset A^h$ is called p**-freezing** if $\forall i \in [\![1,p]\!], A^iW \cap WA^i = \emptyset$. If G is a CA on some $\left\lfloor \frac{h}{2} \right\rfloor$*-freezing* alphabet $B \subset A^h$, then we can build some CA restriction $\boxtimes_h G$ on some SFT, whose *trace* $\tau_{\boxtimes_h G}$ is the *polytrace* $\overset{\circ}{\tau}_G$ of G.

But very few alphabets are freezing; we can impose freezingness by juxtaposing a particular subset, called a border. A **border** for the subset $B \subset A^h$ is some finite DS $(\Upsilon \subset A^l, \delta_\Upsilon)$, where Υ is $\left\lfloor \frac{k+l}{2} \right\rfloor$-*freezing.*

If G is a CA on some alphabet $B \subset A^k$ and $(\Upsilon \subset A^l, \delta_\Upsilon)$ a *border* for B, then we can build some CA restriction F on some SFT, whose *trace* τ_F is $\overset{\circ}{\tau}_G \cup \left\{ (\delta_\Upsilon^j(b))_{j\in\mathbb{N}} \middle| b \in \Upsilon \right\}$.

A first example of *border* are 10^k.

If G is a CA on some alphabet $B \subset A^k$ such that $0^\infty, 1^\infty \in \overset{\circ}{\tau}_G$ and $(\Upsilon \subset A^l, \delta_\Upsilon)$ a *border* for B, then we can build some CA restriction F on some SFT, whose *trace* τ_F is $\overset{\circ}{\tau}_G$.

A second example of *border* is $\left\{ a^{k+|u|}\overline{u}ua^{|u|} \middle| a \in A, u \ne a^{|u|} \right\}$.

If G is a CA on some alphabet $B \subset A^k$ such that $u^\infty \in \overset{\circ}{\tau}_G$, u being nonuniform, and $(\Upsilon \subset A^l, \delta_\Upsilon)$ a *border* for B, then we can build some CA restriction F on some SFT, whose *trace* τ_F is $\overset{\circ}{\tau}_G$.

All sofic subshifts contain either 0^∞ and 1^∞, or u^∞ for some nonuniform u, except weakly nilpotent ones. Nilpotent ones can easily be seen as traces of CA restrictions, whereas other weakly nilpotent are not traceable.

As a result, all *sofic polytraces*, especially *SFT* and *universal sofic* subshifts, are the *traces* of some CA restrictions to some SFT.

Traces.

> We now want to be able to extend our CA to the whole space $A^{\mathbb{Z}}$ without creating invalid trace words. We need a border which cannot appear from scratch; it can be adapted from the previous one as soon as we add the DDC condition.
>
> If G is a onesided CA on alphabet $B \subset A^k$, $\xi : A \to A$ and $(\Upsilon, \delta_\Upsilon)$ a *border* for B such that $\Upsilon \subset \xi(A)^k(\xi(A)^k)^C$, then we can build some CA $F : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ whose *trace* τ_F is $\mathring{\tau}_G \cup \left\{ (\delta_\Upsilon^j(b))_{j\in\mathbb{N}} \middle| b \in \Upsilon \right\} \cup \left\{ (\xi^j(a))_{j\in\mathbb{N}} \middle| a \in A \right\}$.
>
> As a result, all *DDC* subshifts which are the *polytrace* of some onesided CA are *traceable.*

We get our result: all *DDC SFT* and all *DDC universal* sofic subshifts are *traceable.*

Because of the equivalence between ω-nilpotency and nilpotency, non-nilpotent trace subshifts of CA cannot be ω-nilpotent, *i.e.* with a finite number of nonzero letters. There is still a gap between simple trace subshifts (SFT) and complex ones (universal), that is not very well understood. The main open question in this area is to complete the characterization of traceable sofic subshifts. Subsequently, one shall try to understand the non-sofic case, which seems to need a very different approach than the one we used, based on some finite automaton simulation. We can add a subsidiary question: which subshifts are bitraces (or limit traces) of CA?

Bitraces.

> When considering the limit trace, we can perform a finite number of invalid steps and then follow the given trace. In particular we can erase invalid words in the first application of the CA. Hence, if we consider the polytrace of some CA on some alphabet $B \subset A^k$, it has the same limit than the polytrace of some CA on alphabet A^k, by applying a previous "erasing" step.
>
> The previous point makes it impossible for borders to juxtapose an invalid zone of configuration; hence we do not have to destroy borders any more. We can use the more general border $\left\{ ab^{k+1} \middle| \forall j \in \mathbb{N}, \xi^j(a) \neq \xi^j(b) \right\}$, if $\xi : A \to A$.
>
> If G is a onesided CA on alphabet A^k, $\xi : A \to A$ and $(\Upsilon \subset A^l, \xi^{(l)})$ a *border* for A^k, then we can build a CA $F : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ whose trace τ_F is $\mathring{\tau}_G \cup \left\{ (\xi^j(a))_{j\in\mathbb{N}} \middle| a \in A \right\}$.
>
> If G is a onesided CA on alphabet $B \subset A^k$ whose bitrace $\mathring{\tau}_G^*$ is sofic and contains a *deterministic* subshift, then there is a CA F on alphabet A whose *bitrace* τ_F^* is $\mathring{\tau}_G^*$.

As a consequence, all *SFT* and all *universal sofic* subshifts are the *bitrace* of some CA as soon as they contain some *deterministic* subshift.

6 Decidability

Lots of problems on the long-term behavior of CA have been shown undecidable, such as *nilpotency* [Kar92], *quasiequicontinuity* [DFV03], or any nontrivial property (Rice-like theorem) of the *limit set* [Kar94]. More recently, properties on the complexity of the *canonical factors* were proved undecidable [dL06b]; hence, both the equicontinuity and the language classifications presented in [Ků97] are undecidable – except, perhaps, the mysterious property of positive expansivity.

The techniques developed in the previous section are based on the construction of borders to simulate CA on some alphabet $B \subset A^k$ by CA on alphabet A. Hence, they shall help to adapt some undecidability proofs on CA in the general case to undecidability proofs with a given alphabet. This kind of results are often much more difficult; for instance it is still an open question whether the Rice-like theorem on limit sets can be adapted somehow to the case of fixed alphabet, sidestepping the decidability of surjectivity.

Our constructions can, for instance, help to regain the undecidability of *nilpotency* of CA on alphabet $\{0, 1\}$, proved in [DFV03]. Above all, it allows to prove a *Rice-like* theorem on *bitrace* subshifts of CA on alphabet $\{0, 1\}$. Indeed, we can reduce the nilpotency problem of **spreading** CA, *i.e.* that admit a state 0 which spreads in the configuration as soon as it appears. The idea is to build a double product of some CA which has a given trace subshift, a shift CA and some CA of which we wonder if it is nilpotent. When the last one gets to 0, it erases the shift, so that the trace subshift is –nearly – that of the first one; otherwise the trace subshift is a full shift. This gave rise to the article [CG07], since which the rather unclear erasing condition has been formalized.

This result can be expressed in terms of tilings – of which CA space-time diagrams are a particular case: for instance, being given local constraints on tiles, it is not decidable whether we can extend any line of tiles into a valid tiling of the plane.

An interesting question would be to get a similar result on diametral factor, or being given CA of bounded radius. This would require a completely different construction than ours, which took advantage of an unbounded neighborhood to store the encoding. Nevertheless, it seems hopeless to have such as strong statement on diametral factors: the existence of a uniform word in the diametral factor can clearly be read directly from the local rule.

7 Generalizations... and restrictions

We have restricted our study to one dimension; one can wonder what happens when studying two-dimensional CA, *i.e.* continuous self-maps of $A^{\mathbb{M}^2}$ that commute with both the vertical and the horizontal shifts. Actually, most topological results prevail, except those that involve a juxtaposition of words, such as the equivalence between non-sensitivity and quasiequicontinuity. Another great difference lies in decidability questions: the undecidability of the domino problem [Ber66] implies that of the simplest properties in two-dimensional CA, such as injectivity or surjectivity. An approach towards better understanding of this dimension-dependence would be to study restrictions of two-dimensional CA on some subshifts, such as sand automata, which were defined in that context in [DGM08b, DGM08a].

Another important perspective lies in the directional classifications, as presented in [Sab08, ADF07, DdLFM08]. The idea is to avoid the chaotic impact of the shift in the Cantor topology by studying dynamics of CA up to some shift composition. Especially, the symbolic classifications of directional classes is not yet well understood. This issue consists in the study of the action of some subnetwork of the continuous action $\mathbb{N} \times \mathbb{M}$ over the configuration space corresponding to simultaneous applications of the CA and shifts. Playing upon the structure of the monoid $\mathbb{M}$ itself seems another possible generalization, but with very different perspectives; recall that the case of $\mathbb{N}$ is far from being fully understood.

References

[AČI87] Jürgen Albert and Karel Čulik II. A simple universal cellular automaton and its one-way and totalistic version. *Complex Systems*, 1(1):1–16, 1987.

[ADF07] Luigi Acerbi, Alberto Dennunzio, and Enrico Formenti. Shifting and lifting of cellular automata. In Cooper et al. [CLS07], pages 1–10.

[Ban70] Edwin Roger Banks. Universality in cellular automata. In *Symposium on Switching and Automata Theory*, IEEE, pages 194–215, Santa Monica, Californie, 1970.

[BCG82] Elwyn R. Berlekamp, John H. Conway, and Richard K. Guy. *Winning Ways for your mathematical plays*, volume 2. Academic Press, 1982.

[Ber66] Robert Berger. The undecidability of the domino problem. *Memoirs of the American Mathematical Society*, 66:72, 1966.

[BM96] François Blanchard and Alejandro Maass. Dynamical behaviour of Coven's aperiodic cellular automata. *Theoretical Computer Science*, 163:391–302, 1996.

[CFG07] Julien Cervelle, Enrico Formenti, and Pierre Guillon. Sofic trace of a cellular automaton. In Cooper et al. [CLS07], pages 152–161.

[CG07] Julien Cervelle and Pierre Guillon. Towards a Rice theorem on traces of cellular automata. In Ludek Kučera and Antonín Kučera, editors, 32^{nd} *International Symposium on the Mathematical Foundations of Computer Science*, volume 4708 of *Lecture Notes in Computer Science*, pages 310–319, Český Krumlov, République Tchèque, August 2007. Springer-Verlag.

[CLS07] S. Barry Cooper, Benedikt Löwe, and Andrea Sorbi, editors. *Computation and Logic in the Real World, 3rd Conference on Computability in Europe (CiE'07)*, volume 4497 of *Lecture Notes in Computer Science*, Sienne, Italie, June 2007. Springer-Verlag.

[DdLFM08] Alberto Dennunzio, Pietro di Lena, Enrico Formenti, and Luciano Margara. Classification of directional dynamics for additive cellular automata. In Durand [Dur08], pages 40–53.

[DFV03] Bruno Durand, Enrico Formenti, and Georges Varouchas. On undecidability of equicontinuity classification for cellular automata. In Michel Morvan and Éric Rémila, editors, *Discrete Models for Complex Systems (DMCS'03)*, volume AB of *DMTCS Proc.*, pages 117–128. Discrete Mathematics and Theoretical Computer Science, June 2003.

[DGM08a] Alberto Dennunzio, Pierre Guillon, and Benoît Masson. Stable dynamics of sand automata. In Giorgio Ausiello, Juhani Karhumäki, Giancarlo Mauri, and Luke Ong, editors, 5th *IFIP International Conference on Theoretical Computer Science (TCS'08)*, volume 273 of *International Federation for Information Processing*, pages 157–169, Milan, Italie, September 2008. Springer, Boston.

[DGM08b] Alberto Dennunzio, Pierre Guillon, and Benoît Masson. Topological properties of sand automata as cellular automata. In Durand [Dur08], pages 216–227.

[dL06a] Pietro di Lena. *Decidable and Computational properties of Cellular Automata.* PhD thesis, Université de Bologne, Italie, December 2006.

[dL06b] Pietro di Lena. Decidable properties for regular cellular automata. In Gonzalo Navarro, Leopoldo E. Bertossi, and Yoshiharu Kohayakawa, editors, 4th *IFIP International Conference on Theoretical Computer Science (TCS'06)*, volume 209 of *International Federation for Information Processing*, pages 185–196, Santiago, Chili, August 2006. Springer-Verlag.

[Dur08] Bruno Durand, editor. *Journées Automates Cellulaires, 1st Symposium on Cellular Automata (JAC'08)*, Uzès, France, April 2008. MCCME Publishing House, Moscou.

[Fis65] Patrick C. Fischer. Generation of primes by a one-dimensional real-time iterative array. *Journal of the American Mathematical Society*, 12(3):388–394, July 1965.

[Gil87] Robert H. Gilman. Classes of linear automata. *Ergodic Theory & Dynamical Systems*, 7:105–118, 1987.

[Gil88] Robert H. Gilman. Notes on cellular automata. manuscript, 1988.

[GR08] Pierre Guillon and Gaétan Richard. Nilpotency and limit sets of cellular automata. In Edward Ochmański and Jerzy Tyszkiewicz, editors, 33rd *International Symposium on the Mathematical Foundations of Computer Science (MFCS'08)*, volume 5162 of *Lecture Notes in Computer Science*, pages 375–386, Toruń, Pologne, August 2008. Springer-Verlag.

[Hed69] Gustav Arnold Hedlund. Endomorphisms and automorphisms of the shift dynamical system. *Mathematical Systems Theory*, 3:320–375, 1969.

[Hur90] Mike Hurley. Attractors in cellular automata. *Ergodic Theory & Dynamical Systems*, 10:131–140, 1990.

[Ish92] Shin'ichirou Ishii. Measure theoretic approach to the classification of cellular automata. *Discrete Applied Mathematics*, 39(2):125–136, 1992.

[Kar90] Jarkko Kari. *Decision problems concerning Cellular Automata.* PhD thesis, Université de Turku, Finlande, 1990.

[Kar92] Jarkko Kari. The nilpotency problem of one-dimensional cellular automata. *SIAM Journal on Computing*, 21(3):571–586, 1992.

[Kar94] Jarkko Kari. Rice's theorem for the limit sets of cellular automata. *Theoretical Computer Science*, 127(2):229–254, 1994.

[Ků97] Petr Kůrka. Languages, equicontinuity and attractors in cellular automata. *Ergodic Theory & Dynamical Systems*, 17:417–433, 1997.

[LM95] Douglas Lind and Brian Marcus. *An Introduction to Symbolic Dynamics and Coding.* Cambridge University Press, 1995.

[MH38] Marston Morse and Gustav Arnold Hedlund. Symbolic dynamics. *American Journal of Mathematics*, 60(4):815–866, October 1938.

[MI94] Bruno Martin I. A universal cellular automaton in quasilinear time and its S-m-n form. *Theoretical Computer Science*, 123(2):199–237, 1994.

[Moo64] Edward F. Moore. The firing squad synchronisation problem. In Addison-Wesley, editor, *Sequential machines, Selected papers*, pages 213–214, 1964.

[Oll02] Nicolas Ollinger. *Automates cellulaires: structures.* PhD thesis, École Normale Supérieure de Lyon, December 2002.

[Rap97] Ivan Rapaport. *Inducing an order on cellular automata by a grouping operation.* PhD thesis, École Normale Supérieure de Lyon, June 1997.

[Sab06] Mathieu Sablik. *Étude de l'action conjointe d'un automate cellulaire et du décalage: une approche topologique et ergodique.* PhD thesis, Université de Provence, July 2006.

[Sab08] Mathieu Sablik. Directional dynamics for cellular automata: A sensitivity to initial condition approach. *Theoretical Computer Science*, 400(1–3):1–18, 2008.

[SI72] Alvy Ray Smith III. Real-time language recognition by one-dimensional cellular automata. *Journal of Computer and System Sciences*, 6:233–253, 1972.

[The05] Guillaume Theyssier. *Automates cellulaires: un modèle de complexités.* PhD thesis, École Normale Supérieure de Lyon, December 2005.