Symposium on Theoretical Aspects of Computer Science 2010 (Nancy, France), pp. 155-166
www.stacs-conf.org

ULTIMATE TRACES OF CELLULAR AUTOMATA

JULIEN CERVELLE [1] AND ENRICO FORMENTI [2] AND PIERRE GUILLON [3]

[1] Université Paris-Est, LACL, EA 4219, 61 Av du Général de Gaulle, 94010 Créteil Cedex, France
E-mail address: julien.cervelle@univ-paris-est.fr

[2] Laboratoire I3S, Université de Nice-Sophia Antipolis, 2000, Rte des Lucioles - Les Algorithmes - bât. Euclide B - BP 121, 06903 Sophia Antipolis Cedex, France
E-mail address: enrico.formenti@unice.fr

[3] DIM - CMM, UMI CNRS 2807, Universidad de Chile, Av. Blanco Encalada 2120, 8370459 Santiago, Chile
E-mail address: pguillon@dim.uchile.cl

Abstract. A cellular automaton (CA) is a parallel synchronous computing model, which consists in a juxtaposition of finite automata (cells) whose state evolves according to that of their neighbors. Its trace is the set of infinite words representing the sequence of states taken by some particular cell. In this paper we study the ultimate trace of CA and partial CA (a CA restricted to a particular subshift). The ultimate trace is the trace observed after a long time run of the CA. We give sufficient conditions for a set of infinite words to be the trace of some CA and prove the undecidability of all properties over traces that are stable by ultimate coincidence.

Introduction

Cellular automata are a formal computing model known to display many different dynamical behaviors, from the most simple like nilpotency or equicontinuity to the more complex ones like transitivity, mixing or expansivity. These different behaviors together with their ability to capture many features of natural phenomena increase their popularity in the computer sciencists, mathematicians and physicians communities.

A cellular automaton consists in finite state automata (cells) distributed on a regular lattice (or more generally, on any graph). Each cell updates its state depending on the states of a fixed finite number of neighboring cells. This dependency is given by a local rule which is common to all cells.

In this paper, we resume our study of traces of cellular automata, that is to say the sequence of states taken by one particular cell. The main motivation for this work is to study the way scientists deduce general laws from experiments. They proceed by making

1998 ACM Subject Classification: F.1.1 Models of Computation; F.4.3 Formal Languages.
Key words and phrases: discrete dynamical systems, cellular automata, symbolic dynamics, sofic systems, formal languages, decidability.
Thanks to the Projet Blanc ANR *EMC* and the Comité *ECOS-Sud*.

27th Symposium on Theoretical Aspects of Computer Science, Nancy, 2010
Editors: Jean-Yves Marion, Thomas Schwentick
Leibniz International Proceedings in Informatics (LIPIcs), Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Germany
Digital Object Identifier: 10.4230/LIPIcs.STACS.2010.2451

experimental observations using a finite number of observation variables (*i.e.* a trace in the context of CA). From these observations, they conjecture the mathematical law that rules the whole phenomenon. If this law is verified by (almost all) observations, then the scientist concludes that this is the way the phenomenon behaves, until contradicted by new experiments.

However, one also needs formal results ensuring the correctness of the procedure. Indeed, can any observed trace be generated by a CA? How "large" should a trace be to ensure correct reconstruction of the CA local rule?

The notion of trace for a CA has been studied in [CFG07, CG07]. In this paper, we proceed with two generalizations: partial traces and ultimate traces. A partial trace is the trace of a CA restricted to a particular subshift. This kind of trace is motivated by the fact that there are some experiments where not all initial configurations are admissible: some local constraints have to be respected (e.g. a sand grain cannot be above an empty cell or two positively charged particles cannot be too close to one another *etc.*). The ultimate trace is the trace for the long term behavior *i.e.* when the transient part of the phenomenon is neglected, which is often the case in experimental sciences.

The notion of trace is strictly connected with the concept of symbolic factor. Recall that given a CA $(A^{\mathbb{Z}}, F)$, the system $(B^{\mathbb{N}}, G)$ is a (symbolic) *factor* of $(A^{\mathbb{Z}}, F)$, if there exists a continuous surjection $\varphi : A^{\mathbb{Z}} \to B^{\mathbb{N}}$ such that $\varphi \circ F = G \circ \varphi$. Studying the dynamics of factors is often simpler than studying the original system. Indeed, traces are special cases of factor systems. They were introduced as a form of "back-ingeneering" tool to lift properties of factors to CA. Along this research direction, in Section 5, we prove a Rice's theorem for traces. This is an improvement of a similar result in [CG07], in the sense that it is more "natural" and covers more properties than the previous one.

The paper is organized into three parts. Section 1 recalls main definitions concerning cellular automata and symbolic dynamics. Sections 2 to 4 concern new results about traces. Section 5 presents a Rice-like theorem for traces.

1. Definitions

Let id denote the identity map. If F is a function on a set X, denote $F_{|Y}$ its restriction to some subset $Y \subset X$. If F and G are functions on sets X and Y, then $F \times G$ will denote the function on the cartesian product $X \times Y$ which maps any (x, y) to $(f(x), g(y))$.

Configurations. A *configuration* is a bi-infinite sequence of letters, that is an element of $A^{\mathbb{Z}}$. The set $A^{\mathbb{Z}}$ of configurations is the *phase space*. For integers i, j, denote $[i, j]$ the set $\{i, \ldots, j\}$, $[i, j[$ the set $[i, j-1]$, etc…For $x \in A^{\mathbb{Z}}$ and $I = \{i_0, \ldots, i_k\} \subset \mathbb{N}$, $i_0 < \cdots < i_k$, note $x_I = x_{i_0} \ldots x_{i_k}$. Moreover, for a word u, we note $u \sqsubset x$ if u is a factor of x, that is if there exists i and j such that $u = x_{[i,j]}$. If $u \in A^+$, $|u|$ denotes its length, and $x = u^\infty$ [resp. $x = {}^\infty u^\infty$] is the infinite word [resp. configuration] such that $x_{[i,i+|u|[} = u$ for any i in $\mathbb{N}$ [resp. $\mathbb{Z}$]. A word or a configuration is *uniform* if it is made of a single repeated letter. If $L \subset A^k$ and $k \in \mathbb{N} \setminus \{0\}$, we shall also note ${}^\infty L^\infty$ the set of configurations x such that $x_{[ki,(k+1)i[}$ is in L for all $i \in \mathbb{Z}$. Note that we shall assimilate the sets $A^{\mathbb{Z}} \times B^{\mathbb{Z}}$ and $(A \times B)^{\mathbb{Z}}$, for alphabets A, B.

Topology. We endow the phase space with the *Cantor topology.* A base for open sets is given by cylinders: for $j, k \in \mathbb{N}$ and a finite set W of words of length j, we will note $[W]_k$ the *cylinder* $\left\{ w \in A^{\mathbb{Z}} \middle| w_{[k,k+j[} \in W \right\}$. $[W]_k^C$ is the complement of $[W]_k$.

Cellular automata. A (one-dimensional) *cellular automaton* is a parallel synchronous computation model (A, m, d, f) consisting of cells distributed over a regular lattice indexed by $\mathbb{Z}$. Each cell $i \in \mathbb{Z}$ has a state x_i in the finite alphabet A, which evolves depending on the state of their neighbors $x_{[i-m,i-m+d[}$ according to the *local rule* $f : A^d \to A$. The integers $m \in \mathbb{Z}$ and $d > 0$ are the *anchor* and the *diameter* of the CA, respectively. If the anchor is 0, the automaton is said to be *one-sided*. In this case, a cell is only updated according to its state and the ones of its right neighbors. The *global function* of the CA (or simply the CA) is $F : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ such that $F(x)_i = f(x_{[i-m,i-m+d[})$ for every $x \in A^{\mathbb{Z}}$ and $i \in \mathbb{Z}$. The *space-time diagram* of initial configuration $x \in A^{\mathbb{Z}}$ is the sequence of the configurations $(F^j(x))_{j\in\mathbb{N}}$. When the neighborhood of the CA is symmetrical, instead of speaking of anchor and diameter, we shall simply give a *radius*. A CA of radius $r \in \mathbb{N} \setminus \{0\}$, has r for anchor and $2r + 1$ for diameter.

Shifts and subshifts. The *twosided shift* [resp. *onesided shift*], denoted σ, is a particular CA global function defined by $\sigma(x)_i = x_{i+1}$ for every $x \in A^{\mathbb{Z}}$ and $i \in \mathbb{Z}$ [resp. $x \in A^{\mathbb{N}}$ and $i \in \mathbb{N}$] . According to the Hedlund theorem [Hed69], the global functions of CA are exactly the continuous self-maps of $A^{\mathbb{Z}}$ commuting with the twosided shift.

A *twosided subshift* Σ is a closed subset of $A^{\mathbb{Z}}$ with $\sigma(\Sigma) = \Sigma$. A *onesided subshift* Σ is a closed subset of $A^{\mathbb{N}}$ with $\sigma(\Sigma) \subset \Sigma$. We simply speak about the *shift* or *subshifts* when the context allows to understand if it is twosided or onesided.

The language of Σ is $\mathcal{L}(\Sigma) = \left\{ w \in A^* \middle| \exists z \in \Sigma, w \sqsubset z \right\}$ and characterizes Σ, since $\Sigma = \left\{ z \in A^{\mathbb{N}} \middle| \forall w \sqsubset z, w \in \mathcal{L}(\Sigma) \right\}$. For $k \in \mathbb{N}$, denote $\mathcal{L}_k(\Sigma) = \mathcal{L}(\Sigma) \cap A^k$.

A subshift Σ is *sofic* if $\mathcal{L}(\Sigma)$ is a regular language, or equivalently if Σ is the set of labels of infinite paths in some edge-labeled graph. In this case, such a graph is called a *graph of* Σ.

A subshift is characterized by its language $\mathcal{F} \subset A^*$ of *forbidden words*, *i.e.* such that $\Sigma = \left\{ z \in A^{\mathbb{N}} \middle| \forall u \in \mathcal{F}, u \not\sqsubset z \right\}$. A subshift is of *finite type* (SFT for short) if its language of forbidden words is finite. It is a k-SFT (for $k \in \mathbb{N}$) if it has a set of forbidden words of length k. For $\Sigma \subset A^{\mathbb{Z}}$, define $\mathcal{O}_\sigma(\Sigma) = \bigcup_{i\in\mathbb{Z}} \sigma^i(\Sigma)$.

Partial cellular automata. A *partial CA* is the restriction of some CA to some twosided subshift.

Subshift projections. If $B \subset A^k$ is an alphabet and $0 \le q < k$, then the q^{th} *projection* of an infinite word $x \in B^{\mathbb{N}}$ is noted $\pi_q(x) \in A^{\mathbb{N}}$ and defined by $\pi_q(x)_j = a_q$ when $x_j = (a_0, \dots, a_{k-1})$. If Σ is a subshift on B, we also note $\pi(\Sigma) = \bigcup_{0\le q<k} \pi_q(\Sigma)$, which is a subshift on A.

2. Tracebility

Definition 2.1 (Traceability)**.** A subshift $\Sigma \subset A^{\mathbb{N}}$ is *traceable* if there exists a CA F on alphabet A whose *trace* $\tau_F = \left\{ (F^j(x)_0)_{j\in\mathbb{N}} \middle| x \in A^{\mathbb{Z}} \right\}$ is Σ. In this case, we say that F *traces* Σ. If F can be computed effectively from data D, we say that Σ is traceable *effectively from* D. In this notion, D can be any mathematical objet, possibly infinite, provided it has a finite representation (SFT, sofic subshifts, regular languages, CA). In this case, it means one of these representations.

Deterministic subshifts. Given some $\xi : A \to A$, we call *deterministic subshift* the subshift $\mathcal{O}_\xi = \left\{ (\xi^j(a))_{j\in\mathbb{N}} \in A^{\mathbb{Z}} \middle| a \in A \right\}$. The following proposition comes from an easy remark on the evolution of uniform configurations – see Example 4.2 for a subshift which is not traceable.

Proposition 2.2 ([CFG07])**.** *Any traceable subshift $\Sigma \subset A^{\mathbb{N}}$ contains a deterministic subshift $\mathcal{O}_\xi$ for some $\xi : A \to A$.*

Nilpotent subshifts. A subshift $\Sigma \subset A^{\mathbb{N}}$ is 0-*nilpotent* (or simply *nilpotent*) if $0 \in A$ and there is some $j \in \mathbb{N}$ such that $\sigma^j(\Sigma)$ is the singleton $\{0^\infty\}$. It is *weakly nilpotent* if there is some state $0 \in A$ such that for every infinite word $z \in \Sigma$, there is some $j \in \mathbb{N}$ such that $\sigma^j(z) = 0^\infty$. Note that a sofic subshift is weakly nilpotent if and only if it admits a unique periodic infinite word, which is uniform.

The following gives another necessary condition for being the trace of a CA.

Theorem 2.3 ([GR08])**.** *A traceable subshift cannot be weakly nilpotent without being nilpotent.*

Polytraceability. When performing some "back-engineering" from a trace over an alphabet A, *i.e.* when trying to deduce from the trace which CA could have produced it, it is sometimes easier to design a CA over an alphabeth $B \subseteq A^k$ (for some integer k). Being stacked one atop the other, letters of B can be seen as columns of letters of A. In the constructions, the first column is used to produce *all* the elements of Σ and the other columns are used to store elements that help to simulate all possible paths along some graph of Σ. This idea leads to the following notion.

Definition 2.4 (Polytraceabilty)**.** A subshift $\Sigma \subset A^{\mathbb{N}}$ is *polytraceable* if there exists a CA F of anchor 0 and diameter 2 on alphabet $B \subset A^k$ for some k whose *polytrace* $\mathring{\tau}_F = \bigcup_{0\le i<k} \pi_i(\tau_F)$ is Σ. In this case, we say that F *polytraces* Σ. If, furthermore, $B = A^k$, we say that the subshift is *totally polytraceable*. If F and B can be computed effectively from data D, we say that Σ is (totally) polytraceable *effectively from* D.

Note that a polytrace cannot be weakly nilpotent without being nilpotent, otherwise it would alors be the case of the corresponding trace. On the other hand, it need not contain a deterministic subshift.

Theorem 2.5 ([CFG07])**.** *Any subshift Σ which is either of finite type or sofic uncountable is polytraceable effectively from Σ.*

CDD subshifts. A sufficient condition for traceability can be given with the help of the following definition. A subshift $\Sigma \subset A^{\mathbb{N}}$ has *cycle distinct from deterministic* property (CDD) if it contains some deterministic subshift $\mathcal{O}_\xi$ and some periodic infinite word w^∞ such that w contains one letter not in $\xi(A)$. We say that Σ is a CDD subshift.

Lemma 2.6 ([CFG07])**.** *Let $\xi : A \to A$ and $\Sigma \subset A^{\mathbb{N}}$ a polytraceable subshift containing a periodic word w^∞, with $w \in A^+ \setminus \xi(A)^+$. Then $\Sigma \cup \mathcal{O}_\xi$ is traceable effectively from ξ, w and a CA polytracing Σ.*

This lemma, together with Theorem 2.5, gives the following result.

Theorem 2.7 ([CFG07])**.** *Any CDD subshift which is either of finite type or sofic uncountable is traceable effectively from the subshift.*

3. Partial traceability

We already discussed about partial traceability in the introduction. Here is the formal definition.

Definition 3.1 (Partial traceability)**.** A subshift Σ is *partially traceable* if there exists a partial CA F on an SFT Γ whose *trace* $\tau_F = \left\{ (F^j(x)_0)_{j\in\mathbb{N}} \middle| x \in \Gamma \right\}$ is Σ. In this case, we say that F *partially traces* (or simply *traces*) Σ. If F and some graph of Γ can be computed effectively from data D, we say that Σ is partially traceable *effectively from D*.

Assume that Σ is polytraced by some CA $G : B^{\mathbb{Z}} \to B^{\mathbb{Z}}$, with $B \subset A^h$ and $h \in \mathbb{N} \setminus \{0\}$ – for instance obtained from Theorem 2.5. We simulate it by a partial CA F on some SFT Λ in order to get a partial trace instead of a polytrace. This is a kind of *ungrouping* operation that splits *macrocells* (on B) into independent cells (on A).

Ungrouping. The *ungrouping* operation represents a standard encoding of configurations of $B^{\mathbb{Z}}$, with $B \subset A^h$ and $h \in \mathbb{N} \setminus \{0\}$, into configurations of $A^{\mathbb{Z}}$ and it is defined as follows

$$\boxplus_h : \begin{array}{rcl} B^{\mathbb{Z}} & \to & A^{\mathbb{Z}} \\ x & \mapsto & y \text{ such that } \forall i \in \mathbb{Z}, y_{[hi,h(i+1)[} = x_i \ . \end{array}$$

We need to be able to perform this encoding locally, we add some constraints to the alphabet B. Indeed, define the twosided subshift $\Lambda = \mathcal{O}_\sigma(\boxplus_h(B^{\mathbb{Z}})) = \bigcup_{0\le i<h} \sigma^i(\boxplus_h(B^{\mathbb{Z}}))$. We want this union to be disjoint, in order to know, for any configuration of Λ, up to which shift it can be considered a sequence of macrocells. For this purpose, we add a *freezing* condition to B as follows.

Freezingness. A set $W \subset A^h$ is *p-freezing*, with $p, h \in \mathbb{N}$, if $\forall i \in [1,p]\,, A^i W \cap W A^i \neq \emptyset$, *i.e.* words from W cannot overlap on $h-p$ letters or more.

When p is sufficiently large, we obtain the following property.

Proposition 3.2. *Let $W \subset A^h$ be $\left\lfloor \frac{h}{2} \right\rfloor$-freezing, with $h \in \mathbb{N}$. Then W^2 is $(h-1)$-freezing; $\Lambda = \bigcup_{0\le i<h} \sigma^i(\boxplus_h(W^{\mathbb{Z}}))$ is a disjoint union and an SFT.*

If G is a CA of radius 1 on alphabet $B \subset A^h$, we can define its *h-ungrouped* partial CA $\boxtimes_h G$ on the subshift $\Lambda = \mathcal{O}_\sigma(\boxplus_h(B^{\mathbb{Z}}))$, of radius $2h-1$ and local rule:

$$f : \begin{array}{rcl} \mathcal{L}_{4h-1}(\Lambda) & \to & A \\ w & \mapsto & g(u^{-1},u^0,u^1)_i \text{ if } \left\{ \begin{array}{l} w \in A^{h-1-i}u^{-1}u^0u^1A^i \\ u^{-1},u^0,u^1 \in B \\ i \in [0,h[\ . \end{array}\right. \end{array}$$

Proposition 3.3. *Let $B \subset A^h$ be $\left\lfloor \frac{h}{2} \right\rfloor$-freezing, and G a CA on alphabet B, of radius 1 and local rule $g : A^3 \to A$. Then the ungrouped CA $\boxtimes_h G$ is well defined and its trace is $\overset{\circ}{\tau}_G$.*

Proof. The local rule f as defined above is not ambiguous since the shift i is unique by Proposition 3.2. By construction, $f(A^{h-1}u^{-1}u^0u^1A^{h-1}) = g(u^{-1}u^0u^1)$, hence by a recurrence on $j \in \mathbb{N}$, we see that if $i \in [0, h[$ and $x \in \sigma^i(\boxplus_h(B^{\mathbb{Z}}))$, then $\forall k \in \mathbb{Z}, \boxtimes_h G^j(x)_0 = G^j((x_{[kh-i,(k+1)h-i[})_{i\in\mathbb{Z}})_i$. As a result, $\tau_{\boxtimes_h G} = \bigcup_{0\le i<h} \pi_i(\tau_G)$. ∎

Borders. The freezing condition is very restrictive, but any alphabet can be modified in such way to satisfy this property, thanks to a suitable juxtaposition to some freezing set of words. Formally, a *border* for $B \subset A^k$, with $k \in \mathbb{N} \setminus \{0\}$, is a couple $(\Upsilon, \delta_\Upsilon)$, where $\Upsilon \subset A^l$ is $\lfloor \frac{k+l}{2} \rfloor$-freezing, and δ_Υ is a function from Υ into itself. From the latter, seen as the local rule, we define the CA $\Delta_\Upsilon : \Upsilon^{\mathbb{Z}} \to \Upsilon^{\mathbb{Z}}$ of radius 0 whose polytrace is $\bigcup_{0\le i<l} \pi_i(\mathcal{O}_{\delta_\Upsilon})$.

Borders will be used to separate words representing letters of B in an non-ambiguous way.

Proposition 3.4. *Let G be a CA on alphabet $B \subset A^k$ and $(\Upsilon \subset A^l, \delta_\Upsilon)$ a border for B. Then, the ungrouped CA $F = \boxtimes_{k+l}(\Delta_\Upsilon \times G)$ on the SFT $\Lambda = \mathcal{O}_\sigma({}^\infty(\Upsilon B)^\infty)$ is well defined and its trace is $\mathring{\tau}_G \cup \mathring{\tau}_{\Delta_\Upsilon}$.*

Proof. If $\Upsilon \subset A^l$ is $\lfloor \frac{k+l}{2} \rfloor$-freezing, then we can see that so is ΥB. Hence, Proposition 3.3 can be applied to $\Delta_\Upsilon \times G$, seen as a CA on alphabet ΥB. ∎

In the following, we describe a first example of borders.

Corollary 3.5. *Let Σ be a polytraceable subshift which contains two distinct uniform infinite words 0^∞ et 1^∞. Then, Σ is partially traceable effectively from a polytracing CA and these two words.*

Proof. Define $\Upsilon^k_{(0,1)} = \{10^k\}$. Note that $\Upsilon^k_{(0,1)}$ is k-freezing so $(\Upsilon^k_{(0,1)}, \mathrm{id})$ is a border. Applying Proposition 3.4, as $\mathring{\tau}_{\Delta_{\Upsilon^k_{(0,1)}}} = \{0^\infty, 1^\infty\}$, we get that Σ is partially traceable. ∎

Dynamical borders. In the case where the polytraceable subshift does not contain two uniform infinite words, we must find another condition to get a freezing alphabet. Assume it contains some periodic non-uniform infinite word u^∞. We note $\overline{u} = u_{|u|-1} \dots u_0$ the reverse of u and $\gamma^i(u)$ the i^{th} rotation $u_{[i,|u|[}u_{[0,i[}$ of u, for $0 \le i < |u|$. Then the following represents a border: let $\Upsilon^k_u = \left\{ u_i^{k+3|u|}\overline{\gamma^i(u)}\gamma^i(u)u_i^{|u|} \middle| 0 \le i < |u| \right\} \subset A^{k+6|u|}$, and $\delta_{\Upsilon^k_u} : a^{k+3|u|}v\overline{v}a^{|u|} \mapsto v_1^{k+3|u|}\gamma(v)\overline{\gamma(v)}v_1^{|u|}$.

Proposition 3.6 ([CFG10])**.** Υ^k_u *is $(k + 3|u|)$-freezing.*

Corollary 3.7. *Let Σ be a polytraceable subshift which contains a periodic infinite word ${}^\infty u^\infty$ of smallest period $|u| > 1$. Then, Σ is partially traceable effectively from a polytracing CA and u.*

Proof. It is sufficient to apply Proposition 3.4 to the border $(\Upsilon^k_u, \delta_{\Upsilon^k_u})$. We can see that $\mathring{\tau}_{\Delta_{\Upsilon^k_u}} = \mathcal{O}_\sigma(u^\infty)$, which allows to obtain a CA $F : \Lambda \to \Lambda$ such that $\tau_F = \mathring{\tau}_G$. ∎

Actually, the only sofic subshifts which are not concerned by the two previous constructions are the nilpotent ones.

Lemma 3.8 ([CFG10])**.** *Any nilpotent subshift is partially traceable effectively from the subshift.*

The following gives an example of subshift which is nilpotent, hence partially traceable, but not traceable.

Example 3.9 ([CFG10])**.** No CA traces the subshift $\mathcal{O}_\sigma((\lambda+1+01+001+21)0^\infty)$.

Putting things together, we get the following important results.

Proposition 3.10. *Any polytraceable sofic subshift is partially traceable effectively from a polytracing CA.*

Proof. It is known that any sofic subshift Σ admits some periodic infinite word u^∞, and that it is unique only if Σ is weakly nilpotent. In this case, as the projection of some trace, it is nilpotent by Theorem 2.3, and Lemma 3.8 allows to conclude. If there are several distinct periodic infinite words among which one is non-uniform, then we can apply Corollary 3.7; otherwise there are several uniform periodic words and we can apply Corollary 3.5. ∎

The previous proposition, together with Theorem 2.5, gives the following – note that the SFT are partially traceable directly from the definition.

Corollary 3.11. *Any uncountable sofic subshift is partially traceable effectively from it.*

4. Ultimate traceability

In this section we consider traces of CA up to ultimate coincidence, *i.e.* assimilating any two subshifts that are different in only a finite number of cells.

One of the difficulties in making traces (Theorem 2.7), avoided in partial traces, was to deal with "invalid" configurations, not in $\mathcal{O}_\sigma(\boxplus_h(B^\mathbb{Z}))$. At location of "errors" (*i.e.* sites where a pattern of the configuration is not a pattern of $\boxplus_h(B^\mathbb{Z})$), instead of applying the simulating rule, we apply a default rule. However, once one of these rules is chosen, the cell must keep using it forever in order to stay in the "right" subshift.

The possibility of initially altering some cells of the subshift simplifies the problem. Indeed, it allows us to build borders in one round and remove all the "errors" in the initial configuration. We say that two subshifts Γ and Σ *ultimately coincide* if there exists some generation $J\in\mathbb{N}$ such that $\sigma^J(\Gamma)=\sigma^J(\Sigma)$.

Definition 4.1 (Ultimately traceable)**.** A subshift Σ is *ultimately traceable* if there is a CA G such that τ_G ultimately coincides with Σ. If F and J can be computed effectively from data D, we say that Σ is ultimately traceable *effectively from* D.

Note that any ultimately traceable subhift is a subsystem of some traceable subshift, and by Proposition 2.2 contains some deterministic subshift, but which may not involve all the letters of the alphabet.

Example 4.2. Consider the subshift $\Sigma=\mathcal{O}_\sigma((001)^\infty)$. It is an SFT. It is thus polytraceable, but not ultimately traceable since it does not admit any deterministic subshift.

The proof of the following proposition can be found in the online version.

Proposition 4.3 ([CFG10]). *Let $\Sigma \subset A^{\mathbb{N}}$ be a totally polytraceable subshift which contains some non-nilpotent deterministic subshift $\mathcal{O}_\xi$, $\xi : A \to A$. Then Σ is traceable effectively from a polytracing CA and ξ.*

With respect to Lemma 2.6 two additional hypotheses – first, that the subshift is totally polytraceable and, second, that the deterministic subshift is not nilpotent – help get rid of the complex CDD condition, and therefore to get a more precise result about ultimate traces.

Lemma 4.4. *If Σ is a polytraceable subshift, then there exists a subshift $\tilde{\Sigma}$ such that $\sigma(\Sigma) = \sigma(\tilde{\Sigma})$, totally polytraceable effectively from a polytracing CA.*

Proof. Let G be a CA polytracing Σ. Let $\psi : A^k \to B$ be a projection such that $\psi_{|B} = \mathrm{id}$; it can be seen as the local rule of some CA Ψ of radius 0. Define $\tilde{G} = G\Psi$. By construction, we can see that $\tilde{G}_{|B^{\mathbb{Z}}} = G$ and that $\tilde{G}((A^k)^{\mathbb{Z}}) = G(B^{\mathbb{Z}}) \subset B^{\mathbb{Z}}$, *i.e.* since the second time step the two traces coincide. ∎

Proposition 4.5. *Let $\Sigma \subset A^{\mathbb{N}}$ be a polytraceable sofic subshift that contains some deterministic subshift $\mathcal{O}_\xi$, with $\xi : A' \to A'$ and $A' \subset A$. Then Σ ultimately coincides with some subshift $\tilde{\Sigma}$ which is traceable effectively from a polytracing CA, Σ and ξ.*

Proof. Let G be a CA on $B \subset A^k$ polytracing Σ, $k \in \mathbb{N} \setminus \{0\}$. Should we replace Σ by the corresponding $\tilde{\Sigma}$ of Lemma 4.4, we can assume that $B = A^k$.

- If Σ is weakly nilpotent, then, by Theorem 2.3, it is nilpotent, *i.e.* there is some $J \in \mathbb{N}$ such that $\sigma^J(\Sigma) = \{^\infty 0^\infty\}$, property which can be effectively tested from Σ; any nilpotent CA has a trace which ultimately coincides.
- If $\mathcal{O}_\xi$ is not nilpotent, then Proposition 4.3 can be applied to build a CA whose trace will be the polytrace of G.
- Suppose $\mathcal{O}_\xi$ is nilpotent, *i.e.* there is some $J \in \mathbb{N}$ and some state $0 \in A$ such that $\xi^J(A') = \{0\}$; we define:
$$\xi' : \begin{array}{ccc} A & \to & A \\ a & \mapsto & 0 \end{array}.$$
Since the trace $\tau_{\tilde{G}}$ is not weakly nilpotent, it contains some periodic infinite word w^∞, with $w \in A^+ \setminus 0^+ = A^+ \setminus \xi'(A)^+$. Hence, we can apply Lemma 2.6 to build a CA $\tilde{G} : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ such that $\tau_{\tilde{G}} = \mathring{\tau}_G \cup \mathcal{O}_{\xi'}$. As a result, $\sigma(\tau_{\tilde{G}}) = \sigma(\mathring{\tau}_G) \cup \{^\infty 0^\infty\} = \sigma(\mathring{\tau}_G)$. ∎

Corollary 4.6. *Any SFT containing some deterministic subshift and any uncountable sofic subshift containing some deterministic subshift is ultimately traceable effectively from it.*

Here is an example of subshift which is not traceable, but ultimately traceable.

Example 4.7 ([CFG07]). The subshift $\Sigma = \{0^\infty, (01)^\infty, (10)^\infty\}$ is an SFT and contains some deterministic subshift, but is not traceable.

The previous corollary is not an equivalence: there are countable sofic ultimately traceable subshifts which are not SFT.

Example 4.8 ([CFG07]). The subshift $(0^*1 + 1^*)0^\infty$ is sofic, numerable, of infinite type, but traceable.

The study of the ultimate trace of some CA F is related to that of the limit trace, that is the set $\bigcap_{j\in\mathbb{N}} \sigma^j(\tau_F)$ of traces of configurations which can appear arbitrarily late. In particular, we can see that a surjective subshift which ultimately coincides with the trace of some CA is its limit trace. If it is sofic, the converse is true.

The bitrace of some CA F is the set of its "biorbits":

$$\tau_F^* = \left\{ (x_0^j)_{j\in\mathbb{Z}} \middle| \forall j \in \mathbb{Z}, x^j \in A^{\mathbb{Z}} \text{ and } F(x^j) = x^{j+1} \right\}.$$

We can see that it is the twosided subshift with the same language than the limit trace. As a consequence, we get the following.

Corollary 4.9. *Any onesided surjective subshift containing some deterministic subshift which is either of finite type or uncountable sofic is the limit trace of some stable CA. Any twosided subshift containing some deterministic subshift which is either of finite type or uncountable sofic is the bitrace of some stable CA.*

5. Undecidability

Let F a CA of diameter d, anchor m, local rule f on alphabet A. A state $0 \in A$ is 0-*spreading* if $d > 1$ and for all $u \in A^d$ such that $0 \sqsubset u$, we have $f(u) = 0$. The CA F is *spreading* if it is s-spreading for some $s \in A$.

The CA F is 0-*nilpotent* (or simply *nilpotent*) if there exists a $J > 0$ such that $F^J(A^{\mathbb{Z}}) = {}^\infty 0^\infty$. The proof technique developed in [Kar92] allows to prove the following.

Theorem 5.1. *The problem whether a spreading CA F is nilpotent is undecidable.*

In the sequel, we use the spreading state to control the evolution of another CA, generalizing the construction used in [CG07].

Consider two CA F_1 and F_2 of local rules f_1 and f_2 on (disjoint) alphabets A_1 and A_2. Without loss of generality, assume that they have the same diameter d and anchor m. Let $A = A_1 \cup A_2$ and $\varphi : A \to A_1$ a projection such that $\varphi_{|A_1} = \text{id}$. Let N and N_2 be two CA with the same diameter d and anchor m, local rules n, n_2, and alphabets B and $B_2 \subset B$, with $0 \in B_2$ being spreading for N_2. We build the CA H of same diameter d and anchor m, alphabet $A \times B$ and local rule:

$$h : \begin{array}{ccl} (A\times B)^d & \to & A \times B \\ (a_i, b_i)_{-m\le i<d-m} & \mapsto & \left| \begin{array}{ll} (f_2(a), n_2(b)) & \text{if } a \in A_2^d \text{ and } b \in (B_2 \setminus \{0\})^d \ , \\ (f_1 \circ \varphi(a), n(b)) & \text{otherwise .} \end{array} \right. \end{array}$$

Starting from a configuration in $(A_2 \times B_2)^{\mathbb{Z}}$, the CA simulates independently F_2 and N_2 (first part of the rule) until one 0 appears; at that moment they both change their rules; this change can happen only once for each cell, since from then the letters of the left component remain in A_1; hence the two components simulate F_1 and N respectively (second part).

The following notions and lemma will help us understand the dynamics of this CA. A set $U \subset A^k$, with $k \in \mathbb{N} \setminus \{0\}$ is *spreading* if $F([U]_1) \subset [U]_0 \cap [U]_1$ or $F([U]_0) \subset [U]_0 \cap [U]_1$. If F is a CA on alphabet A and $A' \subset A$, then we say that F is (globally) A'-*mortal* if $\forall x \in A^{\mathbb{Z}}, \exists i \in \mathbb{Z}, \exists j \in \mathbb{N}, F^j(x)_i \in A'$.

Lemma 5.2. *If F is a CA on alphabet A and $A' \subset A$ is spreading, then F is A'-mortal if and only if $\exists J \in \mathbb{N}, \forall x \in A^{\mathbb{Z}}, \forall i \in \mathbb{Z}, \forall j \geq J, F^j(x)_i \in A'$.*

Proof. Suppose F is A'-mortal. By compacity, there is some $J \in \mathbb{N}$ and some radius $I \in \mathbb{N}$ such that $\forall x \in A^{\mathbb{Z}}, \exists i \in [-I, I], F^J(x)_i \in A'$. If A' is left-spreading, we obtain thanks to a trivial recurrence, $\forall x \in A^{\mathbb{Z}}, F^{J+2I}(x)_{-I} \in A'$. Thanks to uniformity and shift-invariance, we obtain the stated result. The right-spreading case is symetric. ∎

Lemma 5.3.

- *If N_2 is nilpotent, then there is some $J \in \mathbb{N}$ such that $\pi_0(H^J((A \times B)^{\mathbb{Z}})) \subset A_1^{\mathbb{Z}}$ and then, on $H^J((A \times B)^{\mathbb{Z}})$, H behaves like $F_1 \times N$.*
- *Otherwise, there is a subshift $\Lambda \subset B_2^{\mathbb{Z}}$ such that $\pi_0 \circ H_{|A_2^{\mathbb{Z}} \times \Lambda} = F_2 \circ \pi_0$.*

Proof.

- Suppose N_2 is nilpotent. From the definition of H, no orbit implies always the first part of the rule: H is $A_1 \times B$-mortal. Moreover we can see that $A_1 \times B$ is spreading for H. Thanks to Lemma 5.2, H remains ultimately on the alphabet $A_1 \times B$.
- Otherwise, there exists, thanks to Lemma 5.2, some configuration $x \in B_2^{\mathbb{Z}}$ such that $\forall i \in \mathbb{Z}, \forall j \in \mathbb{N}, N_2^j(x)_i \neq 0$; the subshift $\Lambda = \overline{\mathcal{O}_\sigma(\mathcal{O}_N(x))}$ is such that $A_2^{\mathbb{Z}} \times \Lambda$ is H-invariant and its first column is F_2. ∎

Since they are reduced to the nilpotency of the spreading CA N_2, the two cases presented are recursively inseparable, provided that they are disjoint.

Properties of ultimate polytraces. As for the conditions of traceability, polytraces represent here a useful intermediary tool.

Let G a CA on alphabet $\{0,1\}$ and N a CA on alphabet $\{0,1\}$ of radius 0 and locale rule $\xi : \{0,1\} \to \{0,1\}$ such that $\mathcal{O}_\xi \subset \tau_G$. We build the alphabets $A_1 = \{(a,a,b) | a, b \in \{0,1\}\}$ and $A_2 = \{0,1\}^3 \setminus A_1$, as well as the CA $F_1 = (N \times N \times G)_{|A_1}$, $F_2 = (\sigma \times \sigma \times G)_{|A_2}$. We can apply Lemma 5.3 to the CA H built as above from F_1, F_2, N, and any 0-spreading CA N_2 on alphabet $\{0,1\}$.

The product is here composed of four layers. The fourth one controls the whole behavior thanks to its spreading state 0. The third one simulates G independently. When the two first ones are distinct, they simulate full shifts (whose trace is $\{0,1\}^{\mathbb{N}}$) that hide the trace of G. As soon as some 0 appears in the last layer, they stop, unify and then apply ξ, which is contained in τ_G.

In the end of the section, we consider that H is built from G, N and N_2, the CA F_1 and F_2 being defined as above.

Lemma 5.4.

- *If N_2 is nilpotent, then $\mathring{\tau}_H$ ultimately coincides with τ_G.*
- *Otherwise, $\mathring{\tau}_H = \{0,1\}^{\mathbb{N}}$.*

Proof.

- Thanks to Lemma 5.3, if N_2 is nilpotent, then the first three components of H and $(N \times N \times G)_{|A_1}$ ultimately coincide, the trace of the last component being ultimately included in τ_N. Considering that the polytrace of $(N \times N \times G)_{|A_1}$ is $\tau_G \cup \tau_N$ and that, by hypothesis, $\tau_N \subset \tau_G$ the polytrace of H ultimately coincides with τ_G.
- Otherwise, there exists a subshift Λ such that the partial CA $H_{|A_2^{\mathbb{Z}} \times \Lambda}$ admits as first three projections $(\sigma \times \sigma \times G)_{|A_2^{\mathbb{Z}}}$. The first projection of the trace is $\{0,1\}^{\mathbb{N}}$, since for any infinite word a, there is another word b distinct in every cell ($\forall i \in \mathbb{N}$, $a_i \neq b_i$); hence the trace τ_H contains and therefore is $\{0,1\}^{\mathbb{N}}$. ∎

Properties of traces. As in the previous section, we are now going to simulate CA on alphabets with several components to transform the result on polytraces into a result on traces.

Lemma 5.5. *Let G a non-nilpotent onesided CA whose trace is not $\{0,1\}^{\mathbb{N}}$. The set of CA on alphabet $\{0,1\}$ whose trace is $\{0,1\}^{\mathbb{N}}$ is recursively inseparable from the set of CA on alphabet $\{0,1\}$ whose trace ultimately coincides with τ_G.*

Proof. Let N_2 a onesided 0-spreading CA.

- Suppose that the trace τ_G contains some non-nilpotent deterministic subshift $\mathcal{O}_\xi$, with $\xi : \{0,1\} \to \{0,1\}$. ξ can be seen as the local rule of the CA N. Build CA H as before. From Proposition 4.3, H can be transformed into some CA F on alphabet $\{0,1\}$ such that $\tau_F = \mathring{\tau}_H$.
- If the trace τ_G does not contain any non-nilpotent deterministic subshift, then, as it is still non-nilpotent, it contains some periodic infinite word w^∞, $w \in \{0,1\}^*$, $w \notin 0^*$. We can define the null CA $N = \overline{0}$ on $\{0,1\}^{\mathbb{N}}$ of local rule $\xi' : a \mapsto 0$ and define H as before. Remark that w^∞ and 0^∞ are in the trace of H, hence we can apply Lemma 2.6 to build a CA F on alphabet $\{0,1\}$ such that $\tau_F = \mathring{\tau}_H$.

In both cases, Lemma 5.4 gives that if N_2 is 0-nilpotent, then τ_F ultimately coincides with τ_G, otherwise $\tau_F = \{0,1\}^{\mathbb{N}}$. As F is computable from G, were the two cases separable, Theorem 5.1 would be contradicted. ■

From the remark that some CA traces are not equal to the full shift, we can see that this behavior is undecidable. But the previous lemma also infers other nontrivial properties of traces.

A property $\mathcal{P}$ over subshifts is *stable by ultimate coincidence* if for any subshifts Σ and Γ which ultimately coincide, we have $\Sigma \in \mathcal{P} \iff \Gamma \in \mathcal{P}$.

Theorem 5.6. *Let $\mathcal{P}$ be a property over subshifts which:*

(1) is satisfied by the trace subshift of some CA over alphabet $\{0,1\}$, but not all;
(2) is stable by ultimate coincidence.

Then, the problem

> **Instance:** *a CA G on alphabet $\{0,1\}$.*
> **Question:** *does τ_G satisfy property $\mathcal{P}$?*

is undecidable.

Proof. Let $\mathcal{P}$ be such a property and assume that $\{0,1\}^{\mathbb{N}}$ does not satisfy $\mathcal{P}$, should we take the complement. If $\mathcal{P}$ is only satisfied by nilpotent subshifts, then thanks to stability by ultimate coincidence, it is equivalent either to 0-nilpotency, to 1-nilpotency or to nilpotency, which are all undecidable by Theorem 5.1. Otherwise, $\mathcal{P}$ is satisfied by the trace τ_G of some non-nilpotent CA G. Would an algorithm decide $\mathcal{P}$, it would allow to separate the trace τ_G to $\{0,1\}^{\mathbb{N}}$ among traces over alphabet $\{0,1\}$ up to ultimate coincidence, contradicting Lemma 5.5. ■

This result includes in particular the so-called "nilpotent-stable" properties defined in [CG07], such as fullness, finiteness, ultimate periodicity, soficness, finite type, inclusion of a particular word as a factor. It also includes nilpotency, as well as all properties of the trace of the limit system $(\bigcap_{J\in\mathbb{N}} F^j(A^{\mathbb{Z}}), F)$ of CA F, as stated in [Gui08]. Moreover, it can be easily adapted to larger traces, *i.e.* taking the states of a central group of cells of each configuration. We can also see that this theorem implies the undecidability of all properties of any line projection of two-dimensional SFT (tilings respecting local constraints).

6. Conclusions

In our study of CA traces, we have reached two kinds of important results. On the one hand, we provided sufficient conditions for a subshift to be a polytrace, a trace, a partial trace, an ultimate trace. On the other hand, we proved the undecidability of nearly all properties over ultimate traces. Going beyond undecidability, when it is clear that the trace has been generated by CA, it would be interesting to study which ones, and with which minimal radius.

Remark that the contructions used in the paper build CA with a very large radius. It would be interesting to study the traces produced by cellular automata of a given fixed radius. This is not a so great limitation in complexity, since elementary CA (binary alphabet, radius 1) already present rich different behaviors. In particular, a deeper study of the so-called "canonical factors", *i.e.* traces which width is the radius of the CA, could be fundamental to fully understand this notion.

Another interesting research direction consists in trying to adapt or find some refinement of Kůrka's language classification ([Kůr97]) to the case of traces or ultimate traces. This would provide an interesting link between the complexity of the dynamics of CA and the (language) complexity of its traces.

References

[CFG07] Julien Cervelle, Enrico Formenti, and Pierre Guillon. Sofic trace of a cellular automaton. In S. Barry Cooper, Benedikt Löwe, and Andrea Sorbi, editors, *Computation and Logic in the Real World, 3^{rd} Conference on Computability in Europe (CiE07)*, volume 4497 of *Lecture Notes in Computer Science*, pages 152–161, Siena, Italy, June 2007. Springer-Verlag.

[CFG10] Julien Cervelle, Enrico Formenti, and Pierre Guillon. Ultimate trace of cellular automata. Technical report, http://arxiv.org/abs/1001.0251, January 2010.

[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 *LNCS*, pages 310–319, Český Krumlov, Czech Republic, august 2007. Springer-Verlag.

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

[Gui08] Pierre Guillon. *Automates cellulaires : dynamiques, simulations, traces*. PhD thesis, Université Paris-Est, November 2008.

[Hed69] Gustav Arnold Hedlund. Endomorphism and automorphism of the shift dynamical system. *Math. Sys. Theory*, 3:320–375, 1969.

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

[Kůr97] Petr Kůrka. Languages, equicontinuity and attractors in cellular automata. *Erg. Th. & Dyn. Sys.*, 17:417–433, 1997.