HAL
open science

Clandestine Simulations in Cellular Automata

Pierre Guillon, Pierre-Etienne Meunier, Guillaume Theyssier

To cite this version:

Pierre Guillon, Pierre-Etienne Meunier, Guillaume Theyssier. Clandestine Simulations in Cellular Automata. 2010. hal-00521624v1

HAL Id: hal-00521624

https://hal.science/hal-00521624v1

Preprint submitted on 28 Sep 2010 (v1), last revised 13 Nov 2010 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CLANDESTINE SIMULATIONS IN CELLULAR AUTOMATA

P. GUILLON [1], P.-E. MEUNIER [2], AND G. THEYSSIER [2]

[1] Department of Mathematics, University of Turku,
20014 Turku, Finland
E-mail address, P. Guillon: piegui@utu.fi

[2] LAMA (CNRS, Université de Savoie),
Campus Scientifique 73376 Le Bourget-du-Lac Cedex, France
E-mail address, P.-E. Meunier: pierreetienne.meunier@univ-savoie.fr
E-mail address, G. Theyssier: guillaume.theyssier@univ-savoie.fr

Abstract. This paper studies two kinds of simulation between cellular automata: simulations based on factor and simulations based on sub-automaton. We show that these two kinds of simulation behave in two opposite ways with respect to the complexity of attractors and factor subshifts. On the one hand, the factor simulation preserves the complexity of limits sets or column factors (the simulator CA must have a higher complexity than the simulated CA). On the other hand, we show that any CA is the sub-automaton of some CA with a simple limit set (NL-recognizable) and the sub-automaton of some CA with a simple column factor (finite type). As a corollary, we get intrinsically universal CA with simple limit sets or simple column factors. Hence we are able to 'hide' the simulation power of any CA under simple dynamical indicators.

Introduction

Since the introduction of the model in the 40s, cellular automata have been studied both as dynamical systems and as a computational model. In both aspects, they can show very complex behaviors, be it through their topological dynamics [Kůr03] or through their ability to compute [vN66, Oll03]. As such, they constitute a good model to tackle one of the major question of natural computing: what kind of dynamical behavior allows to support computation, and reciprocally, what does the ability to compute imply on the dynamical behavior of a system?

In this paper, we focus on asymptotic dynamics of cellular automata (notion of limit set [ČPY89]), and on their unprecise observation (notion of column factors [Kůr97]). Intuitively, the limit set is the set of configuration that can appear arbitrarily far in the evolution of the system, and the column factor is the set of sequences of states that a cell (or group of cells) can take in a valid orbit of the

2000 ACM Subject Classification: F.1.1; F.4.3.

Key words and phrases: Cellular automata, Simulation, Limit sets, Trace, Universality.

Research partially supported by projects CONICYT-ANILLO ACT88 (Chile), FONDECYT 1090156 (Chile), ANR EMC NT09 555297 (France) and Academy of Finland project 131558.

system. These notions have been intensively studied in the literature as indicators of the dynamical complexity of cellular automata [Hur90, Kar94, BM97, CFG10].

Concerning limit sets of cellular automata, the initial (wrong) intuition was that a universal CA should necessarily have a non-recursive limit set (such a statement appears in [ČPY89]). Later, a Turing-complete CA with a simple limit set was constructed in [GMM93]. This result gives a first hint concerning the absence of correlation between the complexity of the limit set and the ability to handle computations. However, the construction goes through a slow simulation of two-register machines where registers are encoded in unary. Therefore, this results is *extrinsic* to the model of cellular automata and shows only that any behavior of register machines can be embedded into a cellular automaton with a simple limit set.

Here, we aim at exploring *intrinsic* versions of the same question: what can be the limit set complexity of a CA able to simulate any other CA? More precisely, we consider two flavors of simulations: one using factor (continuous projection) of a simulator CA onto the simulated CA, and the other using the local uniform injection (sub-automaton) of a simulated CA into a simulator CA. Such simulation relations were extensively studied in [DMOT10b], and the second flavor is the one giving rise to the notion of intrinsic universality [Oll02, DMOT10b]. The main point of the present paper is that factor simulations preserve limit set complexity whereas sub-automaton simulations do not. We also show that the same phenomenon appears for the complexity of column factors. Our main results are two embedding theorems showing that there exist an intrinsically universal CA with an NL-recognizable limit set and an intrinsically universal CA with column factors of finite type. The first result solves an open problem of [The05].

1. Definitions

Let us note $\mathbb{N}_+ = \mathbb{N} \setminus \{0\}$. If $i, j \in \mathbb{Z}$, we note $[\![i, j]\!]$ the interval of integers k such that $i \leq k \leq j$, $[\![i, j[\![= [\![i, j-1]\!]$, $]\!]i, j[\![= [\![i+1, j-1]\!]$ and so on. Consider a fixed finite *alphabet* A. If $x \in A^{\mathbb{Z}}$ and $[\![i, j]\!] \subset \mathbb{Z}$, then we note $x_{[\![i,j]\!]} \in A^{j-i+1}$ the *pattern* corresponding to the sequence of letters $x_i \dots x_j$ (and similarly for other kinds of intervals).

If $U \subset A^k$ for some $k \in \mathbb{N}$ and $i \in \mathbb{Z}$, the *cylinder* $[U]_i$ will denote the set of configurations $x \in A^{\mathbb{Z}}$ such that $x_{[\![i,i+k[\![} \in U$. We also note $[U] = [U]_0$. If $a \in A$, let ${}^{\infty}a^{\infty}$ be the configuration $x \in A^{\mathbb{Z}}$ such that $x_i = a$ for any $i \in \mathbb{Z}$.

A *dynamical system* is a pair (X, F) where X is a compact space and F is a continuous self-map of X. We can then study iterations F^t for any *generation* $t \in \mathbb{N}$.

Let $\mathbb{M}$ stand either for $\mathbb{N}$ or for $\mathbb{Z}$. The *shift* map σ is defined for any $z \in A^{\mathbb{M}}$ and any $i \in \mathbb{M}$ by $\sigma(x)_i = x_{i+1}$. A *subshift* is a dynamical system (Σ, σ), or simply Σ, that is a subset of $A^{\mathbb{M}}$ which is closed under the usual Tychonoff topology and invariant by the shift map. Equivalently, it is the set $\left\{ z \in A^{\mathbb{M}} \middle| \forall [\![i, j]\!] \subset \mathbb{M}, z_{[\![i,j]\!]} \notin \mathcal{F} \right\}$ of infinite words that avoid the finite patterns from some given family $\mathcal{F}$. If this family can be taken finite, then Σ is called a subshift of *finite type* (SFT). If it can be taken among words of length $k \in \mathbb{N}_+$, it has *order* k. The *language* $\mathcal{L}(\Sigma)$ of a subshift Σ is the set $\left\{ z_{[\![i,j]\!]} \middle| z \in \Sigma \text{ and } [\![i, j]\!] \subset \mathbb{M} \right\}$ of patterns appearing in the infinite words of Σ. If this language is regular, then we say that Σ is a *sofic* subshift. Equivalently thanks to the Weiss theorem [Wei73], it is obtained from an SFT by a letter-to-letter projection.

A *cellular automaton* (CA) is a dynamical system $(A^{\mathbb{Z}}, F)$ which commutes with the shift, *i.e.* $\sigma F = F\sigma$. Equivalently, thanks to the Curtis-Hedlund-Lyndon theorem, it is obtained by some *local map* $f : A^{2r+1} \to A$ for some *radius* $r \in \mathbb{N}$, *i.e.* for any $x \in A^{\mathbb{Z}}$ and $i \in \mathbb{Z}$, $F(x)_i = f(x_{[\![i-r,i+r]\!]})$. F admits a *spreading* state $0 \in A$ if it admits a local rule $f : A^{2r+1}$ with $r \in \mathbb{N}$ and $f(u) = 0$ whenever there exists $i \in [\![0, 2r]\!]$ such that $u_i = 0$.

1.1. Simulations

The central notion studied in this paper is that of simulation between dynamical systems and more specifically cellular automata. We will distinguish two families of simulation relations based on the following notions.

Definition 1.1 (Factors and sub-systems)**.**

- A *factor map* between two dynamical systems (X, F) and (Y, G) is a continuous onto map $\Phi : X \to Y$ such that $\Phi F = G\Phi$. In that case, we say that G is a *factor* of F.
- A *sub-system* of a dynamical system (X, F) is a dynamical system of the form (Y, G) where $Y \subseteq X$ (Y is closed) and $F(Y) \subseteq Y$.

Note that a factor map Φ between two subshifts Σ and Γ respect the Curtis-Hedlund-Lyndon theorem: there exist a *radius* $r \in \mathbb{N}$ and a *local rule* $\phi : \mathcal{L}_{2r+1}(\Sigma)$ such that $\Phi(x)_i = \phi(x_{[\![i-r,i+r]\!]})$ for any $x \in \Sigma$ and any $i \in \mathbb{Z}$. We say that the factor map between two subshifts is a *coloring* if the radius can be taken $r = 0$.

If (X, F) and (Y, G) are cellular automata, we say that (X, F) *simulates* (Y, G) *by factor* if there is a factor map Φ from (X, F) onto (Y, G) which is also a factor map from (X, σ_X) onto (Y, σ_Y). Besides, when (X, F) is a cellular automaton and Y is a full-shift included in X, then (Y, F) is a *sub-automaton* of (X, F).

These two relations (factor and sub-system) are restrictive since they don't allow entropy to increase: a factor or a sub-system of a given system has always a lower entropy. As a consequence, they don't support universality (existence of a system able to simulate any other) since it is not difficult to build systems of arbitrarily large entropy. Hence, in the literature, other ingredients were introduced to obtain richer notions of simulation: for instance, in cellular automata, operations of space and time rescaling added to the notion of sub-automaton lead to a notion of simulation supporting universality [DMOT10a, DMOT10b].

In this paper, such kind of spatio-temporal transformations are not considered explicitly for two reasons:

- our results about factor simulation (Section 2) involve properties (complexity of the subshifts) which are invariant by space and time rescaling, hence the results still hold when considering rescaling as part of the simulation relation;
- our results about sub-system simulation (Section 3) are of the form "any CA is the sub-automaton of a CA with some given property", which is actually the most general we can get, and remain true when replacing "sub-automaton" by more general notions of sub-system (such as sub-system, or simulated system in the sense of [DMOT10b]).

1.2. Complexity of limit sets and column factors

Many different points of view have been adopted to study the complexity of CA. We here use symbolic dynamics, and consider the complexity, as subshifts, of the limit set on the one hand, or the column factors on the other hand, as representing the actual complexity of the CA.

The *limit set* of the dynamical system is the nonempty closed subset $\Omega_F = \bigcap_{t\in\mathbb{N}} F^t(X)$. Its *limit system* is the maximal surjective subsystem, (Ω_F, F). It basically represent the asymptotic dynamics of the system.

With respect to CA limit sets, it is not difficult to see that the corresponding language is always corecursively enumerable (it is an *effective* subshift). However, there are known examples of non-recursive ones [Hur90]. Moreover, simple additional remarks [Taa07, Kůr03, Nas95] give the following hierarchy:
F injective $\Rightarrow F_{|\Omega_F}$ injective $\Rightarrow \Omega_F$ is an SFT $\Rightarrow F$ is stable $\Rightarrow \Omega_F$ is sofic $\Rightarrow \dots$

The *column factor* of a dynamical system $(A^{\mathbb{Z}}, F)$ upon interval $[\![i,j]\!] \subset \mathbb{Z}$ is the set $\tau_F^{[\![i,j]\!]} = T_F^{[\![i,j]\!]}(A^{\mathbb{Z}})$, where $\begin{array}{rccc} T_F^{[\![i,j]\!]}: & A^{\mathbb{Z}} & \to & (A^{[\![i,j]\!]})^{\mathbb{N}} \\ & x & \mapsto & (F^t(x)_{[\![i,j]\!]})_{t\in\mathbb{N}} \end{array}$. It is a factor subshift since $\sigma T_F^{[\![i,j]\!]} = T_F^{[\![i,j]\!]} F$. It can represent an observation of the system made by a measuring device with a finite precision (that cannot see cells which are far away). It can be seen that any factor subshift is essentially a factor of some column factor [Kůr97].

In the case of CA, shift-invariance allows to consider only the central column factors $\tau_F^{[\![-k,k]\!]}$ for radius $k \in \mathbb{N}_+$. It is known [Gil88] that these CA column factors always have a context-sensitive language; they may actually be strictly context-sensitive. In [Kůr97], strong links with topological notions are stated: finite column factors is equivalent to equicontinuity, SFT column factors imply the shadowing property which in turn imply sofic column factors.

2. Factor simulations

The two hierarchies that we have just defined, based on subshift classifications, are then very robust to factor simulations, as we can see in this section. Taking the vocabulary of order theory, we say that a class $\mathcal{C}$ of systems is *an ideal for factor simulation* if, whenever (X, F) is a factor of (Y, G), we have: $(Y, G) \in \mathcal{C} \Rightarrow (X, F) \in \mathcal{C}$.

Proposition 2.1. *If Φ is a factor map from (X, F) onto (Y, G), then $\Omega_G = \Phi(\Omega_F)$.*

Proof. For any $t \in \mathbb{N}$, $\Phi F^t(X) = G^t(Y)$. First note that $\Phi(\bigcap_{t\in\mathbb{N}} F^t(X))$ is included in $\bigcap_{t\in\mathbb{N}} \Phi F^t(X) = \bigcap_{t\in\mathbb{N}} G^t(Y)$. Conversely, let $y \in \bigcap_{t\in\mathbb{N}} G^t(Y)$ and, for $t \in \mathbb{N}$, $X_t = \Phi^{-1}(y) \cap F^t(X)$. Note that X_t is closed (since Φ and F are continuous, and X is compact) and nonempty (since Φ is onto). By compactness, the intersection $\bigcap_{t\in\mathbb{N}} X_t = \Phi^{-1}(y) \cap \Omega_F$ is not empty, *i.e.* $y \in \Phi(\Omega_F)$. We have proven that $\Phi(\Omega_F) = \Omega_G$. ∎

Moreover, we can note that $\Omega_{F^n} = \Omega_F$, since if $n \in \mathbb{N}_+$, then $F^{nt}(X) = \bigcap_{(n-1)t<j\leq nt} F^j(X)$. In the following we will no longer mention time and space rescaling, which does not essentially alter the results.

Corollary 2.2. *The class of stable systems is an ideal for factor simulation.*

Proof. If Φ is a factor map from (X, F) onto (Y, G), and $\Omega_F = F^t(X)$ for some $t \in \mathbb{N}$, then $\Omega_G = \Phi(\Omega_F) = \Phi F^t(X) = G^t(X)$. ∎

We can also derive the two following results.

Proposition 2.3. *The class of CA whose limit set (resp. factor subshifts) is sofic (resp. context-free, context-sensitive, recursive) is an ideal for factor simulation.*

Proof. It is known that each of the corresponding classes of subshifts is preserved by factor maps. Proposition 2.1 states that the limit set of the simulated CA is a factor of the limit set of the simulating CA.
Moreover, note, thanks to the transitivity of the notion of factor, that a factor subshift of the simulated CA is also a factor subshift of the simulating CA. ∎

Proposition 2.4. *The class of injective dynamical systems is an ideal for coloring.*

Proof. Let Φ be a factor map based on some radius-0 local rule $\phi : A \to B$, from a dynamical system (Σ, F) onto another one (Γ, G), where $\Sigma \subset A^{\mathbb{Z}}$ and $\Gamma \subset B^{\mathbb{Z}}$ and there exists some map $F^{-1} : F(\Sigma) \to \Sigma$ such that FF^{-1} is the identity. By surjectivity of Φ, for any letter $b \in B$, there exists a letter, denoted abusively $\phi^{-1}(b) \in A$ such that $\phi(\phi^{-1}(b)) = b$. If Φ^{-1} represents the parallel application of ϕ^{-1} to $B^{\mathbb{Z}}$, we obtain that $\Phi\Phi^{-1}$ is the identity of Γ. We can define the map $G^{-1} = \Phi F^{-1}\Phi^{-1}$, in such a way that GG^{-1} is the identity of Γ. ∎

As far as we know though, it is unknown whether the class of injective CA is still an ideal for factor simulation.

Corollary 2.5. *The class of dynamical systems which are reversible over the limit set is an ideal for coloring.*

Proof. By Proposition 2.1 the two limit systems are linked by a coloring, and respect the hypotheses of Proposition 2.4. ∎

From the previous, we can see that any universal CA for factor simulation (up to rescaling) must have a non-recursive limit set and strictly context-sensitive factor subshifts (and of course must not be injective), but the existence of such a CA is still open.

3. Sub-system simulations

We will see in this section that, contrary to factor simulations, sub-system simulations allow to hide the complexity of the simulated CA into the simulator CA.

3.1. Hiding the column factors

If G is a CA over alphabet C of radius $r > 0$, local rule g, then let us define the CA $\tilde{G}$ over alphabet $D = \{-1, 0, 1\} \times C$ with the same radius r as G and local rule:

$$\begin{array}{rrcl} \tilde{g} : & D^{2r+1} & \to & D \\ & (\varepsilon_{-r}, c_{-r}) \dots (\varepsilon_r, c_r) & \mapsto & \left| \begin{array}{ll} (\varepsilon_0, c_{\varepsilon_0}) & \text{if } \varepsilon_0 \neq 0 \ ; \\ (0, g(c_{-r} \dots c_r)) & \text{otherwise.} \end{array} \right. \end{array}$$

Clearly, G is a sub-automaton of $\tilde{G}$ (up to state renaming) corresponding to the sub-alphabet $\{0\} \times C$.

Theorem 3.1. *Any cellular automaton G is a sub-automaton of some CA whose column factors are SFT of order* 2.

Proof. Let us take a CA G over alphabet C, of radius 1. Let $\tilde{G}$ defined as above over alphabet $\tilde{C} = \{-1, 0, 1\} \times C$, $k \in \mathbb{N}_+$ and Σ its column factor of width k. Of course, Σ is included in its 2-*approximation*

$$\mathcal{A}_2(\Sigma) = \left\{ z = (z^t)_{t\in\mathbb{N}} \in (\tilde{C}^k)^{\mathbb{N}} \middle| \forall t \in \mathbb{N}, \exists x^t \in [z^t] \cap \tilde{G}^{-1}([z^{t+1}]) \right\}.$$

Let us show that they are actually equal. Let $z = (z^t)_{t\in\mathbb{N}} \in (\tilde{C}^k)^{\mathbb{N}}$ be such that for any $t \in \mathbb{N}$ there exists some configuration $x^t \in [z^t]$ with $\tilde{G}(x^t) \in [z^{t+1}]$, and x defined by:

$$x_i = \left| \begin{array}{ll} x_i^0 & \text{if } 0 \leq i < k \\ (-1, c) & \text{if } i < 0 \text{ and } x_{-1}^{-i-1} = (\varepsilon, c) \\ (1, c) & \text{if } i \geq k \text{ and } x_k^{i-k} = (\varepsilon, c) \ . \end{array} \right.$$

An inductive application of the local rule gives that for any $t \in \mathbb{N}$, we have:

$$\tilde{G}^t(x)_i = \left| \begin{array}{ll} x_i^t & \text{if } 0 \leq i < k \\ (-1, c) & \text{if } i < 0 \text{ and } x_{-1}^{t-i-1} = (\varepsilon, c) \\ (1, c) & \text{if } i \geq k \text{ and } x_k^{t+i-k} = (\varepsilon, c) \ . \end{array} \right.$$

In particular, $z = T_{\tilde{G}}^k(x) \in \Sigma$. We have proven that Σ is an SFT of order 2. If G does not have radius 1, then it is easy to widen the radius of $\tilde{G}$ (and increase the speed of the shifts) to get the same result. ∎

3.2. Hiding the limit set

The main result of this section is based on the existence of a firing-squad CA with specific properties expressed by Lemma 3.2. We actually refer to the firing-squad CA defined in [Kar94], that we denote by S, and prove additional properties in Section 4. This CA admits a so-called *firing* state γ and a spreading state κ. Let r_S the radius of S, s its local rule, Q its state set, and $Q' = Q \setminus \{\kappa, \gamma\}$. Consider the set X_S of configurations having an infinite history avoiding κ and γ:

$$X_S = \left\{ y \in Q^{\mathbb{Z}} \middle| \exists (y^t)_{t\in\mathbb{Z}} \in (Q'^{\mathbb{Z}})^{\mathbb{Z}}, y^0 = y \text{ and } \forall t \in \mathbb{Z}, S(y^t) = y^{t+1} \right\}$$

Lemma 3.2. *S satisfies the following:*

(1) *For any $j \in \mathbb{N}$, there is some configuration $z \in Q^{\mathbb{Z}}$ such that $S^j(z) = {}^{\infty}\gamma^{\infty}$ and $\forall i \in \mathbb{Z}, t < j, S^t(z)_i \notin \{\gamma, \kappa\}$;*
(2) *$\Omega_S \cap [\gamma] \subset \{\kappa, \gamma\}^{\mathbb{Z}}$;*
(3) *X_S is NL-recognizable;*

Proof. Properties (1) and (2) are proven in [Kar94]. Property (3) is given by Proposition 4.10 below. ∎

This construction allows to state the following theorem, proven at the end of the subsection.

Theorem 3.3. *Any CA is a sub-automaton of some CA whose limit set is NL-recognizable.*

By the existence of intrinsically universal CA for a simulation containing the sub-automaton relation (see for instance [Oll03]) and the transitivity of simulations, we can directly derive the following.

Corollary 3.4. *There exists an intrinsically universal CA whose limit set is NL-recognizable.*

The idea of the construction is the following: given some CA F over alphabet A, we add an extra (firing-squad) component which is able to generate any configuration of $A^{\mathbb{Z}}$ arbitrarily far in the future. The complexity of the limit set of F is thus completely flooded into the full-shift $A^{\mathbb{Z}}$. All the technical difficulty is to control the contribution of the additional component to the final limit set.

Let F be a CA of radius r_F, local rule f over alphabet A with a spreading state $0 \in A$. We define a CA $\Delta_{F,S}$ of local rule $\delta_{F,S}$ defined on alphabet $C = A \sqcup (A \times Q)$ with radius $r = \max(r_F, r_S)$ by:

$$\delta_{F,S} : c \mapsto \begin{cases} f(a_{-r_F} \dots a_{r_F}) & \text{if } c = a_{-r} \dots a_r \in A^{2r+1} \ ; \quad (1) \\ a_0 & \text{if } c = (a_{-r}, \gamma) \dots (a_r, \gamma) \in (A \times \{\gamma\})^{2r+1} \ ; \quad (2) \\ (a_0, s(b_{-r_S} \dots b_{r_S})) & \text{if } c = (a_{-r}, b_{-r}) \dots (a_r, b_r) \in (A \times Q')^{2r+1} \ ; \quad (3) \\ 0 & \text{otherwise.} \quad (4) \end{cases}$$

Basically, this CA freezes the first component while applying the firing squad on the second component until some firing state appears, which then frees this second component and starts the application of F. When the configuration is not coherent, or when κ appears, 0 begins to spread. Clearly, F is a sub-automaton of $\Delta_{F,S}$.

The structure of the corresponding limit set will be given by the following lemmas.

Lemma 3.5. $A^{\mathbb{Z}} \subset \Omega_{\Delta_{F,S}}$.

Proof. Let $x \in A^{\mathbb{Z}}$ and $j \in \mathbb{N}_+$. From Point 1 of Lemma 3.2, there is some configuration $z \in Q^{\mathbb{Z}}$ such that $F^{j-1}(z) = {}^{\infty}\gamma^{\infty}$ and for any $t < j-1$ and any $i \in \mathbb{Z}$, $F^t(z)_i \notin \{\gamma, \kappa\}$. Consider now the configuration $y = (x_i, z_i)_{i \in \mathbb{Z}}$. By a quick induction on $t < j$, we can see that for any cell $i \in \mathbb{Z}$, only case (3) of the local rule is used, and $\Delta^t_{F,S}(y)_i = (x_i, S^t(z)_i)$. At time j, since $S^{j-1}(y)_i = \gamma$, the second part of the rule is applied and $\Delta^j_{F,S}(y)_i = x_i$. As a result, $x \in \bigcap_{j \in \mathbb{N}_+} \Delta^j_{F,S}(C^{\mathbb{Z}})$. ∎

Lemma 3.6. *Let $x \in \Omega_{\Delta_{F,S}}$ and $i, j \in \mathbb{Z}$ such that $i < j$ and $x_i = (a_i, \gamma), x_j = (a_j, b_j) \in A \times Q$. Then $b_j \in \{\gamma, \kappa\}$.*

Proof. Assume on the contrary that $b_j \notin \{\gamma, \kappa\}$. Let $(x^t)_{t \in \mathbb{Z}}$ be a biorbit of $x = x^0$, *i.e.* a bisequence of configurations such that $\forall t \in \mathbb{Z}, \Delta_{F,S}(x^t) = x^{t+1}$. By an easy recurrence and the fact that $x_i \in A \times Q$ can only be obtained through case (3) of the rule, we can see that for any $t \in \mathbb{N}$, $x^{-t}_{[\![i-rt, i+rt]\!]}$ can be written $(a^{-t}_{i-rt}, b^{-t}_{i-rt}) \dots (a^{-t}_{i+rt}, b^{-t}_{i+rt}) \in (A \times Q)^{1+2rt}$ and $s^t(b^{-t}_{i-r_St} \dots b^{-t}_{i+r_St}) = b_i$; in the same way, $x^{-t}_{[\![j-rt, j+rt]\!]}$ can be written $(a^{-t}_{j-rt}, b^{-t}_{j-rt}) \dots (a_{j+rt}, b_{j+rt}) \in (A \times Q)^{1+2rt}$, and $s^t(b^{-t}_{j-r_St} \dots b^{-t}_{j+r_St}) = b_j$. Then for any $t > \frac{j-i-1}{2r}$, $x^{-t}_{[\![i-2rt, j+2rt]\!]}$ is in $(A \times Q)^{j-i-1+4rt}$ and the image $s^t(x^{-t}_{[\![i-r_St, j+r_St]\!]})$ contains b_i and b_j. In other words, the cylinder $[b_i Q^{j-i-1} b_j]_i$ intersects $S^t(Q^{\mathbb{Z}})$ for any t, and by compactness intersects Ω_S, which contradicts Point 2 of Lemma 3.2. ∎

If $\Sigma \subset A^{\mathbb{Z}}$ is a subshift and $0 \in A$, then we consider the set $0 \bullet \Sigma \bullet 0 = \bigcup_{-\infty \leq l \leq m \leq +\infty} \left\{ x \in A^{\mathbb{Z}} \middle| \, x_{]\!]l,m[\![} \in \mathcal{L}(\Sigma) \text{ and } \forall i \notin]\!]l, m[\![\, , x_i = 0 \right\}$ of configurations or pieces of configurations of Σ surrounded by 0.

Lemma 3.7. $\Omega_{\Delta_{F,S}} \setminus A^{\mathbb{Z}} \subset 0 \bullet (A \times Q)^{\mathbb{Z}} \bullet 0$.

Proof. By shift-invariance, it is sufficient to prove that $\Omega_{\Delta_{F,S}} \cap [A \times Q]_0 \subset 0 \bullet (A \times Q)^{\mathbb{Z}} \bullet 0$. Let us prove by induction on $n \in \mathbb{N}$ that the patterns of $(A \times Q)(A^{2rn} \setminus \{0^{2rn}\})$

are forbidden in $\Omega_{\Delta_{F,S}}$. The base case is trivial (there are no such patterns). Now suppose it is true for $n \in \mathbb{N}$, and suppose there exists a configuration $x \in [(A \times Q)0^{2rn+k}(A\setminus\{0\})]_0 \cap \Omega_{\Delta_{F,S}}$ with $1 \leq k \leq 2r$. Consider a preimage $y \in \Omega_{\Delta_{F,S}}$ of x. On the one hand, in cell 0 of y, we must have applied case (3), so $y_{[\![-r,+r]\!]} \in (A \times Q)^{2r+1}$, and this word does not involve γ. On the other hand, if we have applied case (1) in cell $2nr + k + 1$ of y, then $y_{[\![(2n-1)r+k+1,(2n+1)r+k+1]\!]} \in (A \setminus \{0\})^{2r+1}$, but the space between these two neighborhoods is $(2n - 1)r + k + 1 - r - 1 \leq 2nr - 1$, which contradicts the induction hypothesis. The other possibility was that we have applied case (2) in cell $2nr + k + 1$, which involves a state γ among cells of $y_{[\![(2n-1)r+k+1,(2n+1)r+k+1]\!]}$, which contradicts Lemma 3.6. In the limit, and with a symmetric argument on the left, we obtain that all the configurations of $\Omega_{\Delta_{F,S}} \setminus A^{\mathbb{Z}}$ are in $0 \bullet \Sigma \bullet 0$. ∎

We can abusively denote $A^{\mathbb{Z}} \times X_S = \left\{ (a_i, s_i)_{i\in\mathbb{Z}} \in (A \times Q)^{\mathbb{Z}} \middle| (s_i)_{i\in\mathbb{Z}} \in X_S \right\}$.

Lemma 3.8. $\Omega_{\Delta_{F,S}} = A^{\mathbb{Z}} \cup 0 \bullet (A^{\mathbb{Z}} \times X_S) \bullet 0$.

Proof. Let $x \in C^{\mathbb{Z}}$ and $l, m \in [\![-\infty, +\infty]\!]$ such that $x_{]\!]l,m[\![} \in (A \times Q)^{\mathbb{Z}}$ and for any $i \notin]\!]l, m[\![$, $x_i = 0$. First suppose that $x \in \Omega_{\Delta_{F,S}}$, *i.e.* for any $t \in \mathbb{Z}$, there exists $x^t \in \Delta^t_{F,S}(x)$. By recurrence, we can see that $x_i^{-t} \in A \times Q'$ for all $i \in]\!]l - rt, m + rt[\![$ and $t \geq 1$ since states from $A \times Q$ are only produced by case (3) of the rule. Let $w^{-t} \in Q'^{m-l+2rt+1}$ be the projection of $(x^{-t})_{]\!]l-rt,m+rt[\![}$ on its second component. Clearly, w^0 is in the language of X_S. We deduce that $x = x^0 \in 0 \bullet (A^{\mathbb{Z}} \times X_S) \bullet 0$.

Conversely, suppose that $x \in A^{\mathbb{Z}} \times X_S$, *i.e.* there are $l, m \in [\![-\infty, +\infty]\!]$ and a sequence (y^t) with, for $t \geq 1$, $y^t \in Q'^{\mathbb{Z}}$ and $y^t = S(y^{t+1})$ and, for any $t \in \mathbb{N}$ and any $i \in]\!]l, m[\![$, $x = (a_i, S^t(y^t)_i)$ for some $a_i \in A$. Now take the configuration $\tilde{y}^t \in C^{\mathbb{Z}}$ such that for any $i \notin]\!]l - rt, m + rt[\![$, $\tilde{y}^t_i = 0$, and for any $i \in]\!]l - rt, m + rt[\![$, $\tilde{y}^t_i = (b_i, y^t_i)$ with $b_i \in A$, and $b_i = a_i$ if $i \in]\!]l, m[\![$. By a direct recurrence, for any $j < t$ and any $i \notin]\!]l - rt + rj, m + rt - rj[\![$, we have $\Delta^j_{F,S}(\tilde{y}^j)_i = 0$ and for any $i \in]\!]l - rt + rj, m + rt - rj[\![$, we have $\Delta^j_{F,S}(\tilde{y}^j)_i = (b_i, S^j(y^j)_i)$ (since $y_j \in Q' \in \mathbb{Z}$ case 3 of the definition of $\Delta_{F,S}$ applies at position i of $\tilde{y}^j$). This gives that $\Delta^t_{F,S}(y) = x$. We have proven that $\Omega_{\Delta_{F,S}} \cap 0 \bullet (A \times Q)^{\mathbb{Z}} \bullet 0 = 0 \bullet (A^{\mathbb{Z}} \times X_S) \bullet 0$. We can conclude thanks to Lemmas 3.7 and 3.5. ∎

Corollary 3.9. $\Omega_{\Delta_{F,S}}$ *has an NL-recognizable language.*

Proof. From Lemma 3.8 and Point 3 of Lemma 3.2, the language of the limit set is the finite boolean combination of finite concatenation of NL-recognizable languages. ∎

Proof of Theorem 3.3. Let F be a CA on some alphabet A. We can artificially add some spreading state $0 \notin A$ to build a CA $\tilde{F}$ on alphabet $A \sqcup \{0\}$ which admits F as a sub-automaton. Now we have seen that $\tilde{F}$ is a sub-automaton of $\Delta_{\tilde{F},S}$. From Corollary 3.9, the corresponding limit set has an NL-recognizable language. ∎

4. Analysis of a firing-squad CA

Let S be the firing-squad CA defined in [Kar94]. It has a state set Q of size 16, including a killer state κ, radius 1 and is defined by the transitions appearing

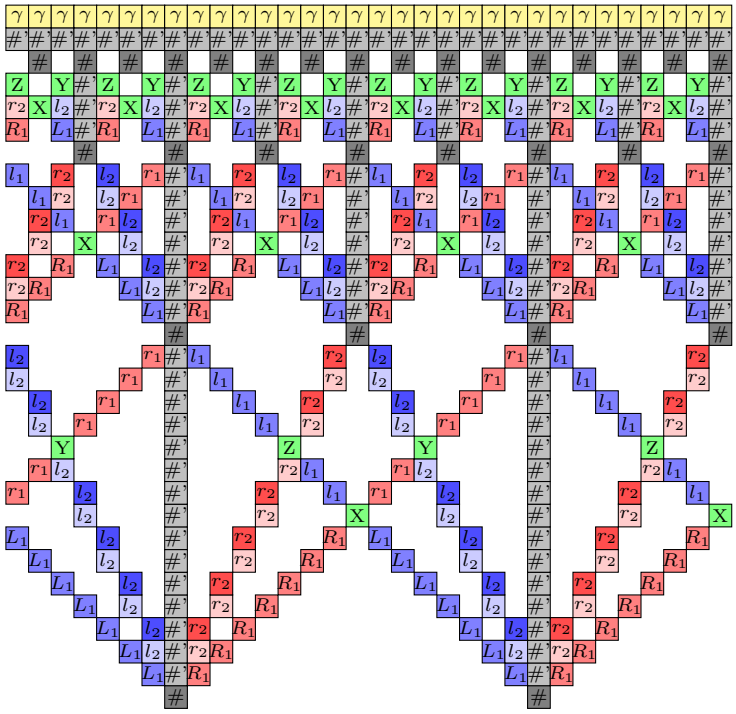


Figure 1: The 16-state firing squad of [Kar94]. Empty spaces represent the blank state.

in Figure 1: precisely, any transition which is not in the space-time diagram of the figure produces the killer state κ. The complete list of transitions is given in [Kar94].

We are interested in history diagrams in X_S, *i.e.* mapping from $\mathbb{Z} \times \mathbb{N}$ to Q of the form: $(z, t) \mapsto x^t(z)$ where (x^t) is a sequence of configuration in X_S such that $S(x^{t+1}) = x^t$. We call them *valid* history diagrams.

When restricted to X_S, the behavior of S is easier to understand via a signal/collision evolving in a quiescent background. More precisely, the background is uniform and made of blank states (denoted B in the sequel) and the signals involved are:

signal	L_1	l_1	l_2	$\#'$	r_2	r_1	R_1
speed	-1	-1	-1/2	0	1/2	1	1

The valid collisions are:

$$L_1 + R_1 \to l_1 + r_1 \qquad l_2 + r_2 \to \#$$
$$l_1 + r_2 \to l_1 + r_2 \qquad l_1 + \#' + r_1 \to \#$$
$$r_1 + l_2 \to r_1 + l_2 \qquad \# \to L_1 + l_2 + \#' + r_2 + R_1$$

Any other intersection of signal is invalid (it raises a κ state). Moreover, the last collision rule (starting from a single $\#$) is valid only if the $\#$ is distant from any other $\#$ by at least 3 cells: if they are 1 cell away, 3 adjacent $\#'$ are generated; if they are 2 cells away, a κ is generated 2 steps later.

To simplify proofs, we will often make reasonings over (portions of) "Euclidean" versions of history diagrams. A Euclidean history diagram is a set of labelled points and labelled (half-)lines or segments in $\mathbb{R}^2$ satisfying the following rules:

- points are only at integer coordinates ($\mathbb{Z}^2$) and labelled by $\#$;

- (half-)lines and segments correspond to signals listed above (label and slope correspond);
- any intersection between lines or segments follow the collision rules above.

Lemma 4.1. *To each history diagram D, we can associate a valid Euclidean history diagram E such that, at any integer coordinate of E containing a point (#) or a single signal, the label gives the state of the corresponding position in D.*

This lemma allows the following proof scheme (used several times below): supposing by sake of contradiction that some word w occurs in a history diagram, we make a reasoning on the corresponding Euclidean diagram, we get a contradiction and finally deduce that no history diagram exists which contain the word w, and therefore that w is not in the language of X_S.

S satisfies points (1) and (2) of Lemma 3.2 as shown in [Kar94, Prop. 4.3]. We give below a complete characterization of the language of X_S which shows that it is NL-recognizable.

Lemma 4.2. *Consider a history diagram containing a word $w \in \#Q^*\#'$ at time t_0. Let z_1 (resp. z_2) be the cell where the first (resp. the last) letter of w occurs. We suppose in addition that the left # of w was created by a l_1 signal. Let t_1 be the first time step in the past when the cell z_1 is in state #, and t_2 be the first time step in the past when the cell z_2 is in state #. Then, both t_1 and t_2 exist.*

We denote by Σ the set $Q' \setminus \{\#, \#'\}$.

Lemma 4.3. *There is no history diagram containing a word w of the form $\#\Sigma^*\#'$, where the left # is created by r_1/l_1 signals.*

Lemma 4.4. *There is no history diagram containing a word w of the form $\#\Sigma^*\#'$, where the left # was created by a l_2/r_2 pair of signals.*

Lemma 4.5. *Any configuration from X_S with at least two # is of the following form, for some value of n: ${}^\omega(\#B^n)^\omega$*

Lemma 4.6. *Let L be the language of configurations from X_S admitting an history diagram where two # occur at some time t in the past. Then $L \in$ NL − recognizable in logarithmic space.*

Lemma 4.7. *The language of configurations from X_S which contain only one state in $\{\#, \#'\}$ is also in* NL.

Lemma 4.8. *Let L be the language of configurations from X_S with two or more $\#'$ and having an history diagram with no #. Then L is regular.*

Lemma 4.9. *The language of the configurations from X_S without any # or $\#'$ is regular.*

Proposition 4.10. *The language of X_S is in* NL.

Proof. There are several cases, and the disjunction on configurations allows to express the language of X_S as a union of 'simple' NL languages.

(1) Configurations with #s or #s. We can descibe the set of these configurations by :

$$^\omega\{L_1, l_1, B\}\{l_2, B\}^*A\{r_2, B\}^*\{R_1, r_1, B\}^\omega$$

where A is one of the following (possibly infinite) configurations:

(a) A has exactly one state in $\{\#, \#'\}$. We conclude in this case with Lemma 4.7.
(b) A has one $\#$, and at least one other $\#$ or $\#'$. Lemmas 4.3, 4.4 and 4.5 show that the configuration satisfy the hypothesis of Lemma 4.6, which allows to conclude.
(c) A has at least two $\#'$ but do not contain any signal. Then Lemma 4.8 conclude.
(d) A has at least two $\#'$, along with some signal(s) between two $\#'$. Denote by c the global configuration in this case. We can simply go back a few steps in the past to find out a configuration c' of case 1b. Then we can apply Lemma 4.6 to c.

(2) Configurations without #s nor $\#'$s. This case is treated in Lemma 4.9.

Thus, since we have described above why each of the possible languages could be recognized in NL, we can just build a non-deterministic machine beginning by making a non-deterministic choice between all of these machines, then doing the computation of the chosen one. ■

Conclusion

We have thus achieved results implying that both limit set and column factors complexities are strongly linked to the factor simulation hierarchy; on the other hand, they are rather orthogonal to the sub-automaton simulation hierarchy.

Many open questions remain.

- We have obtained that universality was not forbidden by some rather strong constraints either on the limit set, or "orthogonally", on the column factors. A natural question is whether we can constrain both at the same time. The two constructions may possibly be composed together, at the price of a (yet) more difficult proof of the NL-recognizability of the limit set.
- Similarly, we believe that our results still hold when alphabets are restricted to $\{0, 1\}$ but at the price of a more technical proof.
- Is there an intrinsically universal CA with an SFT limit set? Following our construction, this raises immediately the following question: is there a firing-squad CA with an SFT limit set?
- What kind of limit system can an intrinsically universal CA have? Can it be injective?
- Is injectivity, expansivity of CA preserved by factor maps?
- Is it enough, for a CA to be a factor of another CA, that the corresponding column factors with some given width be linked by a factor map?
- Is there, for some complexity level λ, an equivalence class for the sub-automaton simulation (with space-time rescalings) of which all the elements have limit sets of complexity λ?

References

[BM97] François Blanchard and Alejandro Maass. Dynamical properties of expansive one-sided cellular automata. *Israel Journal of Mathematics*, 99:149–174, 1997.

[CFG10] Julien Cervelle, Enrico Formenti, and Pierre Guillon. Ultimate traces cellular automata. In Jean-Yves Marion, editor, 27th *International Symposium on Theoretical Aspects of Computer Science (STACS'10)*, Nancy, March 2010.

[ČPY89] Karel Čulik II, Jan K. Pachl, and Sheng Yu. On the limit sets of cellular automata. *SIAM Journal on Computing*, 18(4):831–842, 1989.

[DMOT10a] Marianne Delorme, Jacques Mazoyer, Nicolas Ollinger, and Guillaume Theyssier. Bulking I: an abstract theory of bulking. oai:hal.archives-ouvertes.fr:hal-00451732, January 2010.

[DMOT10b] Marianne Delorme, Jacques Mazoyer, Nicolas Ollinger, and Guillaume Theyssier. Bulking II: Classifications of cellular automata. oai:hal.archives-ouvertes.fr:hal-00451729, January 2010.

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

[GMM93] Eric Goles, Alejandro Maass, and Servet Martínez. On the limit set of some universal cellular automata. *Theoretical Computer Science*, 110:53–78, 1993.

[Hur90] Lyman P. Hurd. Nonrecursive cellular automata invariant sets. *Complex Systems*, 4:131–138, 1990.

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

[Kůr97] Petr Kůrka. Languages, equicontinuity and attractors in cellular automata. *Ergodic Theory & Dynamical Systems*, 17:417–433, 1997.

[Kůr03] Petr Kůrka. *Topological and symbolic dynamics*. Société Mathématique de France, 2003.

[Nas95] Masakazu Nasu. *Textile Systems for Endomorphisms and Automorphisms of the Shift*, volume 114 of *Memoirs of the American Mathematical Society*. American Mathematical Society, Providence, Rhode Island, March 1995.

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

[Oll03] Nicolas Ollinger. The intrinsic universality problem of one-dimensional cellular automata. In Helmut Alt and Michel Habib, editors, 20th *Annual Symposium on Theoretical Aspects of Computer Science (STACS'03)*, volume 2607 of *Lecture Notes in Computer Science*, pages 632–641, Berlin, Germany, February 2003. Springer-Verlag.

[Taa07] Siamak Taati. Cellular automata reversible over limit set. *Journal of Cellular Automata*, 2(2):167–177, 2007.

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

[vN66] John von Neumann. *Theory of Self-Reproducing Automata*. University of Illinois Press, Champaign, IL, USA, 1966.

[Wei73] B. Weiss. Subshifts of finite type and sofic systems. *Monatshefte für Mathematik*, 77:462–474, 1973.

Appendix A. Proof of Lemma 4.1

Proof. There is no rounding problem for signals of slope -1, 0 and 1 (positions of cells of discrete signals correspond exactly to integer position on the continuous signal). Concerning slopes $1/2$ and $-1/2$, S is such that when two # at position p_1 and p_2 are connected by a $1/2$ signal in a discrete space-time diagram, the displacement vector $p_2 - p_1$ is always of the form $(n, 2n)$ (the situation is similar for slope $-1/2$). More precisely, one can check that the set of cells $\{p_1 + (i, 2i) : i \in \mathbb{N}, i \leq n\}$ are in a r_2 signal state, and it corresponds exactly to the integer positions on the continuous segment from p_1 to p_2. ∎

Appendix B. Proof of Lemma 4.2

Proof. First let us prove that t_1 exists. Let us assume that there is no # in the past of cell z_2. The only possible past of this cell, in this case, is necessarily an infinite column of #'s. But we assumed that the left # was created by an l_1 signal, which necesarily either crossed this column, or was generated by it. In both cases, this is a contradiction.

Now that we know this, we prove that t_2 exists. Let us assume that cell z_1 has only #'s in its past. Since the # in the right column was necesarily created by a signal coming from the left, be it r_1 or r_2; this implies that the signal should have "crossed" the left column, which is impossible. Thus, there is at least one # in this column. ∎

Appendix C. Proof of Lemma 4.3

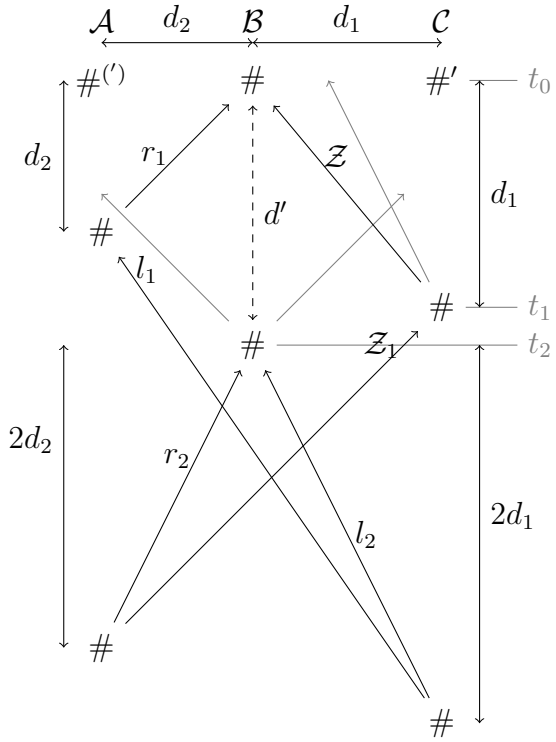
Proof. Let us call $\mathcal{Z}$ the l_1 signal that created the left # of w at t_0. According to Lemma 4.2, there is some time step t_1 in which a # appears in the past of the right #′ (i.e. of column $\mathcal{C}$). Moreover, let t_2 be the most recent step in which a # appears in the past of the # on column $\mathcal{B}$ at t_0. There are two cases:

- This # in column $\mathcal{C}$ was created by an r_1 signal $\mathcal{Z}_1$. In this case, this # is necessarily the same as the one that generated $\mathcal{Z}$. Else, applying the same reasonning as in Lemma 4.2, there would be another # between $\mathcal{Z}$, and the # at t_1, and then $\mathcal{Z}_1$ would have crossed both $\mathcal{Z}$ and the L_1 signal emitted by this intermediate #.

  We now have two subcases:
  - The # at time t_2 in column $\mathcal{B}$ was created by l_2/r_2. We deduce that the # on column $\mathcal{C}$ at time t_1 was created by an l_1/r_1 pair of signals: otherwise we would have a l_2/r_2 intersection between the two columns, which is only possible on a #, but this would create a column of #'s visible at time t_0, which contradicts the hypothesis. Thus we can deduce that:
    (1) Column $\mathcal{C}$ existed before t_1, and we can thus consider the first # on it before t_1. This is the # which emitted the l_2 signal creating the # on column $\mathcal{B}$ at t_2, otherwise a l_1 or l_2 signal emmitted by this # would intersect column $\mathcal{B}$ between t_0 and t_2, which is impossible by minimality of t_2.

(2) the r_2 signal arriving on the # of column $\mathcal{B}$ at time t_2 and the r_1 arriving on the # of column $\mathcal{C}$ at time t_1 meet in the past on some #. This allows to infer the existence of a third column $\mathcal{A}$ on the left of the two previous ones;

We thus have the following situation:



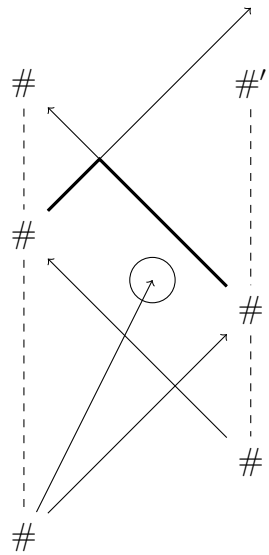
The distances imposed by the signals allow to conclude:

$$\begin{aligned} d_1 + d_2 &= 2d_2 + d' - d_1 \\ 2d_1 &= d_2 + d' \end{aligned} \tag{C.1}$$

$$\begin{aligned} d_2 + d_1 &= 2d_1 + d' - d_2 \\ 2d_2 &= d_1 + d' \end{aligned} \tag{C.2}$$

Equation (C.1) comes from the r_1 signal from the leftmost bottom #, and equation (C.2) comes from the rightmost bottom one. From this we can deduce that $d_1 = d_2 = d'$ and that the R_1/r_1 signal emitted at time t_2 by the central column reaches the right column at time t_0, contradicting the fact that the state at time t_0 in this right column is $\#'$.

– It was created by l_1/r_1, and so was the right # at t_1. In this case, one of them has to have another # in its past, by Lemma 4.2. Thus, this new # sends a r_2 or l_2 signal (depending on which one we consider), that necessarily collides either with the other column of $\#'$s, or to the L_1/R_1 signals represented in thick on the figure below. Since this L_1 signal creates the top left #, this collision should occur necessarily before time t_0 and would create a κ which is forbidden by hypothesis.



- Or it was created by l_2/r_2. In this case, the only possible option for the # at t_2 on the left column is a l_1/r_1, else there would be another $\#/\#'$ column between the # and the #' at time t_0, contrarily to our hypothesis.

  Therefore we are in the exact symmetric of the case studied above, where we supposed that the # at t_2 was created by l_2/r_2, and the left # at t_1 by l_1/r_1. We can conclude to a contradiction by the same reasoning.

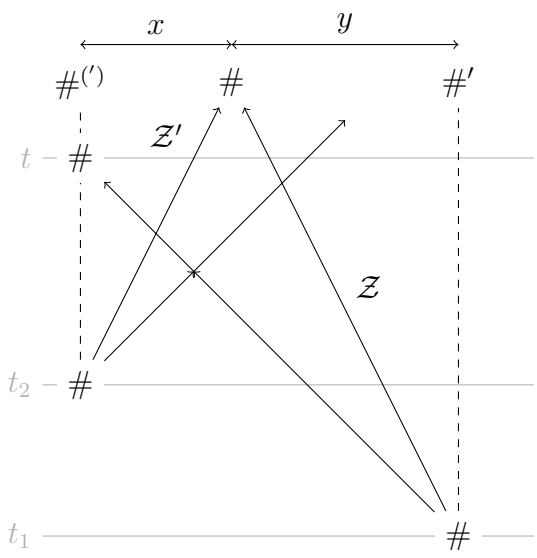
■

Appendix D. Proof of Lemma 4.4

Proof. We want to use the above Lemma 4.3 to prove this one. To do this, we need to prove the existence of a configuration of the form $\#\Sigma^*\#'$ or its symmetric $\#'\Sigma^*\#$, with in any case the # created by l_1/r_1 signals, in the past of our current configuration.

We begin by proving the existence of another column of #, on the left of the left #. Since the right column of #' collides with $\mathcal{Z}$ – the l_2 signal creating the left #, there is necessarily a # in the past of this #', say at time t_1. But then, this # is necessarily the one who emitted $\mathcal{Z}$, for else it would have been formed by l_2/r_2 signals colliding with $\mathcal{Z}$, thus forming another # between the # and the #' of our hypothesis. This shows that $\mathcal{Z}$ was originated by a # in the same column than the right #'.

But then, this # must have emitted also an L_1 signal, that cannot collide with the r_2 signal (let us call it $\mathcal{Z}'$), colliding with $\mathcal{Z}$ to form our left #. Thus, this signal must be converted into an l_1 before meeting the r_2 of the hypothesis; more precisely, it must collide with an R_1 signal. But this one necessarily has a finite origin, since it could not have crossed $\mathcal{Z}'$. This shows the existence of another column of $\#^{(')}$, at t_2.

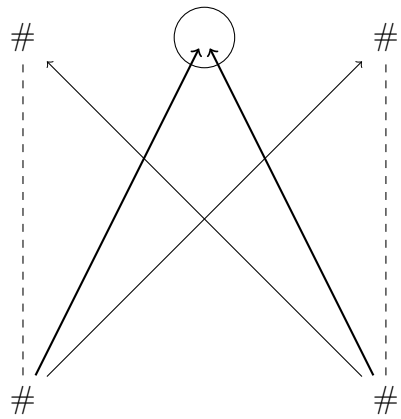


But then, since both #s at t_1 and t_2 emit a L_1 or R_1 signal, and they are desynchronized, one of this signal has to hit the other column before t_0. In fact, to avoid this, we should have at the same time $x + y > 2y$ and $x + y > 2x$, which is impossible. Thus, we can apply Proposition 4.3 or its symmetric to this step (at time t on the figure). ■

Appendix E. Proof of Lemma 4.5

Proof. We first need to show that in a configuration $x \in X_S$ containing #s, there is no history diagram yielding x in which two consecutive #s of x were both created by r_2/l_2 signals, or both by r_1/l_1 signals.

- if two consecutive #s were created by r_2/l_2 signals, these signals would have collided before, thus yielding another # between them.
- if two consecutive #s were created by r_1/l_1, the right r_1 would originate in the left column, and vice-versa:



These signals would then have been created by #s, which also would have emitted r_2s and l_2s intersecting between the two #s in question, thus creating another #.

Now if we have two consecutive #s in a configuration, we can infer the position of a third one from the signals having created them, since they have different speeds

and thus must intersect. Then, we can reproduce this argument at the step where these signals were created.

The only problem with this is that the infered # could be on the same side at each step of the argument, thus yielding a configuration with a rightmost # or a leftmost #. Without loss of generality, let us assume there is rightmost #. Then, the past of this cell would be an infinite column of #s/#′s, with infinitely many #s. But then, some r_2 signal emitted by a # of this column would collide with an R_1 emitted later, which would yield a K state: the configuration we considered would not be in X_S.

Then, it is easy to see that there can be no signal between two consecutive #s, for they would come from a desynchronized configuration in a history diagram of the configuration. ■

Appendix F. Proof of Lemma 4.6

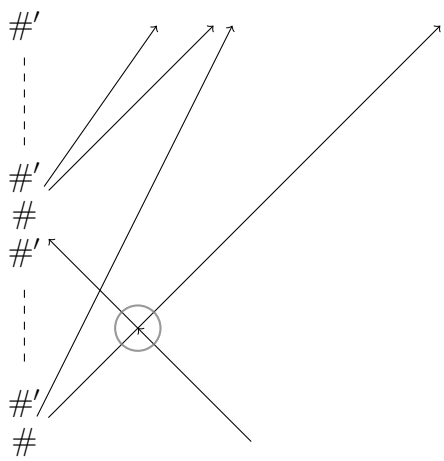
Proof. Applying Lemma 4.5 at time t, we now that the configuration at time t is of the form ${}^{\omega}(\#B^n)^{\omega}$. Moreover, using again the argumentation of Lemma 4.5, we know that their is some time t' before t in history where the configuration is of the form ${}^{\omega}(\#B^{n'})^{\omega}$ with $n' > n$. Hence, L is exactly the language of forward orbits of periodic configurations of the form ${}^{\omega}(\#B^n)^{\omega}$. It is straightforward to check from the definition of S (see Figure 1) that the language of periods of such forward orbits is a finite union of languages of the form:

$$\#' B^{x_1} A_1 \cdots B^{x_i} A_i$$

where $i \leq 4$, $A_i \in \Sigma$ and where the x_i must satisfy some simple linear equations. The lemma follows. ■

Appendix G. Proof of Lemma 4.7

Proof. There are several possibilities for the number of # in the past of the # of #′. There may be at most two, for else, two of them would have been created by l_1 signals crossing the R_1/r_1 emitted by the third #, and that intersection, circled in the following picture, can occur only once.

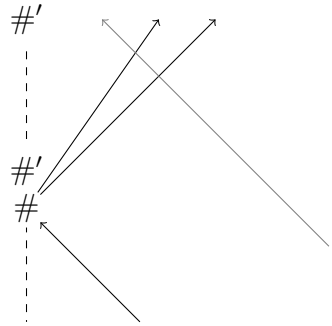


Thus, we are left with:

- Two # in the past: in this case, the r_2 they emitted allows to guess the whole column. After the last r_2 signal, there may be signals in the opposite direction, that cannot collide in the past, i.e. first l_2s, then L_1s and l_1s. Such configurations are described, with $x = 2z + y$, by:

$$\{R_1, r_1, B\}^\omega \{r_2, B\}^* l_1 B^x l_2 B^y L_1 B^z l_2 B^z \{\#', \#\} B^z r_2 B^z R_1 B^y r_2 B^x r_1 \{L_1, l_1, B\}^\omega$$

- With only one # in the past, there may still be some l_1 signal somewhere:



Thus the configurations are of the following form:

$$\{R_1, r_1, B\}^\omega L_1 B^z l_2 B^z \{\#', \#\} B^z r_2 B^z R_1 \{L_1, l_1, B\}^\omega \ ,$$

with possibly a r_1 replacing some state on the left, and an l_1 on the right, like on the figure.

- With no # in the past, we have an infinite column of #′, thus no signal can come from it. These configurations are of the form:

$$^\omega\{R_1, r_1, B\}\{r_2, B\}^* \#' \{l_2, B\}^* \{L_1, l_1, B\}^\omega$$

The lemma easily follows from the characterization of the different possible forms of configuration discussed above. ∎

Appendix H. Proof of Lemma 4.8

Proof. In an history diagram satisfying the hypothesis their can be no collision that would generate a # and even no signal crossing (because this would mean that two signal are going in opposite directions in the same part of a configuration and thus that at least one would meet a #′ in the past, which is forbidden by the hypothesis). Therefore we have:

$$L = \left(B^+\{R_1, r_1\}\right)^* \left(B^+ r_2\right)^* \left(B^+ \#'\right)^* \left(B^+ l_2\right)^* \left(B^+\{L_1, l_1\}\right)^*$$

∎

Appendix I. Proof of Lemma 4.9

Proof. Without any # nor #′, we are left with blank states and signals. The configurations of the limit set are those configurations in which the possible signals do not have to collide in the past, thus fall in one of these three possibilities:

$$^\omega\{l_1, r_2, B\}^\omega$$

$$^\omega\{r_1, l_2, B\}^\omega$$

$$^\omega\{r_1, R_1, B\}\{r_2, B\}^*\{l_2, B\}^*\{l_1, L_1, B\}^\omega$$

■