



**HAL**  
open science

# The Mirage of Universality in Cellular Automata

Guillaume Theyssier

► **To cite this version:**

| Guillaume Theyssier. The Mirage of Universality in Cellular Automata. 2021. hal-03461942

**HAL Id: hal-03461942**

**<https://hal.science/hal-03461942>**

Preprint submitted on 1 Dec 2021

**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.

# The Mirage of Universality in Cellular Automata

Guillaume Theyssier

**Abstract** This note is a survey of examples and results about cellular automata with the purpose of recalling that there is no 'universal' way of being computationally universal. In particular, we show how some cellular automata can embed efficient but bounded computation, while others can embed unbounded computations but not efficiently. We also study two variants of Boolean circuit embedding, transient versus repeatable simulations, and underline their differences. Finally we show how strong forms of universality can be hidden inside some seemingly simple cellular automata according to some classical dynamical parameters.

## 1 Eric, the collector

The present note responds to an invitation to contribute to a book at the occasion of Eric Goles 70th birthday. Before diving into the scientific content, I should say a few words about Eric and the motivation behind this note.

Anyone knowing Eric certainly had the pleasure to listen to some of his colorful anecdotes (I certainly did). He owns a large collection, large enough to adapt to a wide variety of listeners and circumstances. The collection is in fact twice as large, because each anecdote, usually told to an international audience, is doubled with a more confidential Chilean version full of slang words. Eric's pleasure of telling stories is obvious, he has generously shared his collection, but nobody has listened to the same sequence of anecdotes and we all end up with a different global picture, much like the adventurous readers of the antinovel of Cortázar.

The collection of models and systems studied by Eric in its numerous scientific publications is equally striking. It abounds in small examples that are carefully analyzed and shown to capture important phenomena. It connects different points of view and different communities of researchers. It seems to never end up in the

---

Guillaume Theyssier  
CNRS, Université Aix-Marseille, e-mail: guillaume.theyssier@cnrs.fr

exact same theoretical framework and invites us to think about details that make a difference. In short, there is an anti-Bourbakist quality to it.

At the heart of this scientific collection (at least from what I can tell from my collaboration with Eric), there is the question of the computational universality of small dynamical systems, and how it manifests itself in the complexity of various associated decision problems. Computational universality of dynamical systems is a topic that might seem boring to the classical computer scientist (after all Turing showed the existence of a universal machine in the 1930s) and not serious for the dynamical systems community (this is not real maths<sup>1</sup>). Part of the problem is that this kind of research is endangered by what I would call the *mirage of universality*: the illusion that there must be universal consequences to the fact of being “computationally universal” independently of the precise definition used, and that such a statement, even given without technical details, gives information by itself. Pursuing this mirage, one is tempted to put forward vague theorem statements and hide the concrete mathematical result in the proofs (or sketch of). To make an analogy, no paper in computational complexity would use theorem statements like “Problem X is hard” and then, hidden in the proof details, unveil the definition of “hard”. On the contrary, computational complexity theory has been extremely fruitful by putting forward a vast “zoo” of precisely defined complexity classes, often with a corresponding notion of reduction.

Of course, there is a lot to say and a lot has already been said about the mathematical formalization of computational universality in dynamical systems, but my intention here is clearly not to start a comprehensive survey on the topic [7, 21]. Instead, I would like to invite the reader to a quick tour of examples and properties that break this mirage of universality. Most of them were encountered or established during my collaboration with Eric, and I hope this note can give a clue about the richness of Eric’s scientific collection.

**Content of the note:** To simplify exposition, I chose to restrict to (classical) cellular automata and tackle three main topics in three separate sections. Each topic shows examples of “computationally universal” cellular automata that, in some sense, do not behave as expected, or pair of examples that behave differently with respect to some parameter:

- efficient versus unbounded computations: how some cellular automata are able to embed one type of computations but not the other;
- transient versus repeatable circuit simulations: about the existence of (at least) two fundamentally different ways to simulate Boolean circuits in cellular automata, and their consequences;
- hidden universality: how cellular automata might seem ‘simple’ according to some parameter despite being actually universal.

Before starting, some standard definition are given below to set up our framework.

---

<sup>1</sup> It should be noted however that a growing trend in symbolic dynamics has shown the importance of computability considerations. Some of these results were even published in real math journals...

## 2 Standard definitions and notations

For any finite set  $Q$  (the alphabet) and positive integer  $d$  (the dimension), we consider the space of configurations  $Q^{\mathbb{Z}^d}$ , *i.e.* the set of maps giving a state from  $Q$  to each position in the lattice  $\mathbb{Z}^d$ . The state of configuration  $c \in Q^{\mathbb{Z}^d}$  at position  $z \in \mathbb{Z}^d$  will be denoted either  $c(z)$  or  $c_z$ .

For  $n \in \mathbb{N}$ , let  $\mathcal{B}(n)$  be the set of positions of  $\mathbb{Z}^d$  of maximum norm at most  $n$ :

$$\mathcal{B}(n) = \{z \in \mathbb{Z}^d : \|z\|_\infty \leq n\}.$$

Then for any  $u \in Q^{\mathcal{B}(n)}$ , we define the *cylinder set*  $[u]$  centered on cell 0 by:

$$[u] = \{c \in Q^{\mathbb{Z}^d} : \forall z \in \mathcal{B}(n), c_z = u_z\}.$$

These cylinder sets can be chosen as a base of open sets of the space  $Q^{\mathbb{Z}^d}$  endowing it with a compact topology [17]. Equivalently, the same topology can be defined by the Cantor distance:

$$\delta(c, c') = 2^{-\min\{\|z\|_\infty : c_z \neq c'_z\}}.$$

A cellular automaton of dimension  $d$  and state set  $Q$  is a map  $F$  acting continuously on configurations and translation invariant way. Equivalently (Curtis-Lyndon-Heldund theorem [14]), it can be defined locally by a neighborhood  $V$  (a finite subset of  $\mathbb{Z}^d$ ) and a local transition map  $f : Q^V \rightarrow Q$  as follows:

$$\forall z \in \mathbb{Z}^d, \quad F(c)_z = f(c|_{z+V})$$

where  $c|_{z+V}$  denotes the map  $z' \in V \mapsto c_{z+z'}$ .

The *radius* of  $F$  is the smallest integer  $r$  such that  $V \subseteq \mathcal{B}(r)$  where  $V$  is some neighborhood for which there is a local map  $f_V : Q^V \rightarrow Q$  defining  $F$  as above.  $F$  induces an action on finite patterns as follows. For any  $n \in \mathbb{N}$  and any  $u \in Q^{\mathcal{B}(n+r)}$ ,  $F(u)$  is the finite pattern  $v \in Q^{\mathcal{B}(n)}$  obtained by application of  $f$  on  $u$  at each position from  $\mathcal{B}(n)$ , *i.e.* such that

$$\forall c \in [u], F(c) \in [v].$$

We are now going to define a notion of universality for cellular automata. We choose this one for two reasons: first it is one of the strongest form of universality and will serve us as a benchmark in the following, and second, it is intrinsic to the model of cellular automata and make no reference to other models of computation (for more details, see [22, 6, 5]).

This notion, called *intrinsic universality*, is based on a notion of (intrinsic) simulation that is defined through two ingredients [5, 6].

The first ingredient is a notion of cell-wise simulation that works by restriction to a sub alphabet and then projection onto the target alphabet. To be more precise let  $F$  and  $G$  be cellular automata of dimension  $d$ . We denote by  $F \triangleleft G$  the fact that  $F$  is obtained from  $G$  by cell-wise restriction and projection, formally:  $\exists \pi : Q \subseteq Q_G \rightarrow Q_F$  surjective such that for all  $c \in Q^{\mathbb{Z}^d}$

$$\bar{\pi} \circ G(c) = F \circ \bar{\pi}(c)$$

where  $\bar{\pi} : Q^{\mathbb{Z}^d} \rightarrow Q_F^{\mathbb{Z}^d}$  is the cell-wise application of  $\pi$ . In the language of dynamical systems,  $(F, Q_F^{\mathbb{Z}^d})$  is a factor of  $(G, Q_G^{\mathbb{Z}^d})$  which is a sub-system of  $(G, Q_G^{\mathbb{Z}^d})$ .

Now we add the second ingredient, *rescaling*, that allows to turn a cell-wise simulation into a simulation that works by blocks: blocks of cell of the first CA are simulated by blocks of cell of the second CA. Given a rectangular shape  $\mathbf{m} = (m_1, \dots, m_d)$  and some alphabet  $Q$  we define the bloc recoding map  $B_{\mathbf{m}}$  from  $Q^{\mathbb{Z}^d}$  to  $(Q^{m_1 m_2 \dots m_d})^{\mathbb{Z}^d}$  by:

$$\begin{aligned} B_{\mathbf{m}}(x)(z_1, \dots, z_d) = & (x(m_1 z_1, \dots, m_d z_d), \dots, \\ & x(m_1 z_1 + m_1 - 1, m_2 z_2, \dots, m_d z_d), \\ & x(m_1 z_1, m_2 z_2 + 1, m_3 z_3, \dots, m_d z_d), \dots \\ & x(m_1 z_1 + m_1 - 1, \dots, m_d z_d + m_d - 1)) \end{aligned}$$

It is a bijection that recodes any configurations by blocks of shape  $\mathbf{m}$ . Now if  $t$  is a positive integer and  $\mathbf{z} \in \mathbb{Z}^d$ , we define the rescaling of  $F$  of parameters  $\mathbf{m}$  and  $t$  as the CA  $F^{<\mathbf{m}, t>} = B_{\mathbf{m}} \circ F^t \circ B_{\mathbf{m}}^{-1}$ .

We finally say that  $G$  *simulates*  $F$ , denoted  $F \leq G$ , if there are parameters  $\mathbf{m}, t, \mathbf{m}'$  and  $t'$  such that  $F^{<\mathbf{m}, t>} \triangleleft G^{<\mathbf{m}', t'>}$ . We also say that  $G$  *strongly simulates*  $F$  if there are parameters  $\mathbf{m}$  and  $t$  such that  $F \triangleleft G^{<\mathbf{m}, t>}$ . Then, a CA  $G$  is *intrinsically universal* if for any CA  $F$  we have  $F \leq G$ . It can be shown that an intrinsically universal CA can in fact *strongly simulate* any CA [6].

Finally, we assume the reader is familiar with basic notions and results of computability and complexity theory. We will use the following standard classes of decision problems:

- P is the set of problems which can be solve be a deterministic Turing machine in polynomial time;
- NLOGSPACE is the set of problems which can be solve be a non-deterministic Turing machine in logarithmic space;
- $\Sigma_1^0$  is the set of recursively enumerable problems (which contains the halting problem).

Without explicit mention and when speaking about P-completeness we consider LOGSPACE reductions. When speaking about  $\Sigma_1^0$ -completeness we usually consider many-one reductions.

### 3 Efficient vs. unbounded computations

It is well-known that, besides the reference model of Turing machines, there are other ones that fundamentally differ because they either only allow efficient but bounded computation (like Boolean circuits) or unbounded but slow computations

(like Minsky machines) [19]. We would like to illustrate this aspect in the framework of cellular automata in a precise manner. To simplify, we restrict to dimension 1 in this section. We first define two classical problems associated to any CA which will serve as canonical indicators for both aspects mentioned above: efficiency and unboundedness of computations.

The first one is about short-term predictability within a bounded time range and provides a fine-grained complexity measurement within class P.

**Definition 1** Let  $F$  be any CA of radius  $r$  and alphabet  $Q$ . The prediction problem  $\text{PRED}_F$  is defined as follows:

- input:  $t > 0$  and  $u \in Q^{\mathcal{B}(rt)}$
- output:  $F^t(u) \in Q$ .

The second one asks for a prediction about an unbounded future and provides a coarse-grained complexity measure allowed to cross the decidable barrier. It could be refined in many ways as in the definition of universality for dynamical symbolic systems from [7]. We prefer to keep it simple for the clarity of exposition. We say a configuration  $c \in Q^{\mathbb{Z}}$  is eventually bi-periodic if it is eventually periodic to the left and eventually periodic to the right, said differently if it is of the form  ${}^\infty u_L \cdot u \cdot u_R^\infty$  where  $u_L, u$  and  $u_R$  are finite words.

**Definition 2** Let  $F$  be any CA of radius  $r$  and dimension 1. The reachability problem  $\text{UBPRED}_F$  is defined as follows:

- input: an eventually bi-periodic configuration  $c = {}^\infty u_L \cdot u \cdot u_R^\infty$  and a state  $q$ .
- output: decide whether there is  $t \in \mathbb{N}$  such that  $F^t(c)_0 = q$ .

One of the well-know results of computational universality in cellular automata is about elementary rule 110 given by the local rule  $\delta : \{0, 1\}^3 \rightarrow \{0, 1\}$  with

$$\delta(x, y, z) = (1 - xyz) \cdot \max(y, z).$$

It is interesting to note that the first proof of computational universality of this cellular automaton due to M. Cook [3] was enough to prove undecidability of  $\text{UBPRED}_\delta$  but did not give information about problem  $\text{PRED}_\delta$ . It is only later, by a strong improvement in one step the the reduction, that  $\text{PRED}_\delta$  was proven to be P-complete [20]. The purpose of this section is precisely to make clear that there is generally no implication in either direction between the P-hardness of  $\text{PRED}$  and the undecidability of  $\text{UBPRED}$ .

**Definition 3** A CA  $F$  is a *freezing CA* if, for some (partial) order  $\leq$  on states, the state of any cell can only decrease, *i.e.*

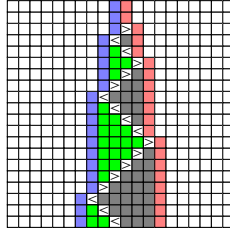
$$F(c)_z \leq c_z$$

for any configuration  $c$  and any cell  $z$ .

The definition above was introduced in [9] in studied more in depth in [23]. Similar cellular automata corresponding to bounded changes or bounded communications were also considered in the literature with the point of view language recognizers [24, 18, 2]. Under the hypothesis that  $\text{NLOGSPACE} \neq \text{P}$ , the following results show examples of cellular automata that can embed arbitrary unbounded computation, but not in an efficient way.

**Theorem 1 (Section 4.3 of [23])** *For any freezing CA  $F$  of dimension 1, the problem  $\text{PRED}_F$  is in  $\text{NLOGSPACE}$ . There exists a 1D freezing CA  $F$  such that  $\text{UBPRED}_F$  is  $\Sigma_1^0$ -complete.*

We are now going to build an example with the opposite computation embedding properties: as hard as it can be in the short term (it can embed efficiently bounded computations), but decidable in the long term (it can not embed unbounded computations). It is inspired from [23, Example 7] and consists in a simulation of some P-complete cellular automaton inside finite zones, with some head controlling the simulation and forced to move back and forth inside the zone and shrink it by one cell at each bounce on a boundary. The simulation is such that one step of the simulated cellular automaton is done at each pass so that there is only a quadratic slowdown (see Figure 1).



**Fig. 1** The shrinking zone trick behind the construction of  $\mathcal{Z}_F$  (time goes from bottom to top).

Let  $F$  be any 1D CA on alphabet  $Q$  with radius 1 and local map  $\delta : Q^3 \rightarrow Q$ . We define  $\mathcal{Z}_F$  on alphabet  $R = \{b, b_+, e\} \cup Q'$  with  $Q' = Q \times Q \times \{\leftarrow, \rightarrow, l, r\}$  and radius 1 as follows:

- $e$ , the *error state*, is a spreading state: any cell with  $e$  in its neighborhood turns into state  $e$ ; a configuration  $c$  is *valid* if  $e$  never appears in its orbit;
- $b$ , the blank state, never changes except in presence of the error state;  $b_+$  becomes  $b$  except in presence of the error state; a maximal connect component of cells in state  $Q'$  is a *working zone*;
- in a working zone, patterns of the form  $(x, y, r)(x', y', l)$ , or  $(x', y', l)(x, y, r)$ , or  $(x, y, z)(x', y', z')$  with  $\{z, z'\} \subseteq \{\leftarrow, \rightarrow\}$ , or  $(x, y, r)(x', y', z)$  or  $(x, y, z)(x', y', l)$  with  $z \in \{\leftarrow, \rightarrow\}$ , are forbidden and generate an  $e$  state when detected; therefore in a valid configuration and in each working zone there is at most one occurrence of a state of the form  $(x, y, \{\leftarrow, \rightarrow\})$  called the *head*;

- a cell without forbidden pattern (from previous item) and without head in its neighborhood doesn't change its state;
- the movements and actions of the heads are as follows:
  - inside a working zone, the head in state  $\leftarrow$  moves left, the head in state  $\rightarrow$  moves right; the local map  $\delta$  is only applied the head moves left to right; precisely we have the following transitions:

$$\begin{aligned} (x, y, l), (x', y', \leftarrow), (x'', y'', r) &\mapsto (x', y', r) \\ (x, y, l), (x', y', l), (x'', y'', \leftarrow) &\mapsto (x', y', \leftarrow) \\ (x, y, l), (x', y', \rightarrow), (x'', y'', r) &\mapsto (x', y', l) \\ (x, y, \rightarrow), (x', y', r), (x'', y'', r) &\mapsto (\delta(y, x', x''), x', \rightarrow) \end{aligned}$$

- when a boundary of the working zone is reached, the head bounces, changes of direction and the working zone get shrunked by one cell; precisely we have the following transitions:

$$\begin{aligned} b, (x, y, l), (x', y', \leftarrow) &\mapsto (x, y, \leftarrow) \\ b, (x, y, \leftarrow), (x', y', r) &\mapsto (x, y, \rightarrow) \\ b, (x, y, \rightarrow), (x', y', r) &\mapsto (y, x, l) \\ b, (x, y, l), (x', y', \rightarrow) &\mapsto b_+ \\ (x, y, \rightarrow), (x', y', r), b &\mapsto (x', y', \rightarrow) \\ (x, y, l), (x', y', \rightarrow), b &\mapsto (x', y', \leftarrow) \\ (x, y, l), (x', y', \leftarrow), b &\mapsto (x', y', r) \\ (x, y, \leftarrow), (x', y', r), b &\mapsto b_+ \end{aligned}$$

(note the swap between  $x$  and  $y$  in the third transition above to initialize the sequential application of  $\delta$ )

- finally the head disappears in a working zone of size 1, precisely:

$$b', (x, y, z), b'' \mapsto (x, y, r)$$

for any  $\{b', b''\} \subseteq \{b, b_+\}$ .

**Theorem 2** For any  $F$ , the problem  $\text{UBPRED}_{Z_F}$  is decidable in polynomial time. If  $F$  is chosen so that  $\text{PRED}_F$  is P-complete, then  $\text{PRED}_{Z_F}$  is P-complete.

*Proof* For the first part of the Theorem, let us consider an eventually bi-periodic configuration  $c = {}^\infty u_L \cdot u \cdot u_R^\infty$ . There are four cases:

- $c$  is not a valid configuration, which means that it contains a working zone with a forbidden pattern. Since the forbidden pattern are locally detectable, such a forbidden pattern must be detected inside the finite word  $u_L u_L u_L u u_r u_R u_R$



(considering the worst case where  $u_L$  or  $u_R$  is of size 1). Therefore, in time  $t$  which is linear in the sizes of  $u_L$ ,  $u$  and  $u_R$  we have  $\mathcal{Z}_F^t(c)_0 = e$ ;

- $c$  is a valid configuration and position 0 belong to a finite working zone in  $c$ . Since the left and right boundary of this zone must belong either to  $u$ , or  $u_L$  or  $u_R$ , the zone is of linear size and it gets completely shrunked in quadratic time, meaning that the state of position 0 will no longer change after a quadratic time;
- $c$  is a valid configuration and position 0 belongs to an infinite zone in  $c$ . In this case, the position of the head must belong to either  $u$ ,  $u_L$  or  $u_R$  and the same for the eventual (unique) boundary of the zone. Therefore, after a linear time in the worst case, cell 0 will never change again (for instance, the head comes from the right, bounces to the left boundary, crosses once more position 0, but never comes back again);
- $c$  is a valid configuration but position 0 does not belong to some working zone, then for any  $t \geq 1$  we have  $\mathcal{Z}_F^t(c)_0 = b$ .

We deduce that that after a quadratic time the state of cell 0 does not change any more, so it is sufficient to simulate  $\mathcal{Z}_F$  on  $c$  for this number of states to solve problem  $\text{UBPRED}_{\mathcal{Z}_F}$ .

For the second part of the Theorem, see [23, Lemma 1 and Proposition 5].  $\square$

## 4 Transient vs. repeatable circuit simulation

In this section we focus on dimension 2 and simulation of Boolean circuit and logical gates by cellular automata. Showing how a cellular automaton can embed Boolean circuits is one of the common methods used to claim its Turing universality (see for instance [8, 1, 21]).

We are now going to describe two modes of simulation of a set of logical gates by a cellular automaton, which were formalized in [11]. The basic simulation mechanism behind both simulation modes uses square blocks concatenated in a grid-like fashion. Each such square block represents a part of a concrete Boolean circuit (either a node or wire). The definition doesn't require any specific way of representing information inside the blocks, just that the family of blocks use coherent representation of information so that the Boolean logic works when assembling them. More concretely, they communicate information with their four neighbors (north, east, south, west) in such a way that each one implements a Boolean function with at most 2 inputs and at most 2 outputs.

In the sequel all considered blocks will compute one of the following maps (we represent them using type  $\{0, 1\}^4 \rightarrow \{0, 1\}^4$  in order to make explicit the position of inputs and outputs among the neighbors in the order north, east, south, west):

$$\begin{aligned}
\text{AND}(x, *, *, y) &= (0, \min(x, y), 0, 0) \\
\text{OR}(x, *, *, y) &= (0, \max(x, y), 0, 0) \\
\text{CROSS}(x, *, *, y) &= (0, y, x, 0) \\
\text{NOP}(*, *, *, *) &= (0, 0, 0, 0) \\
\text{FORK}(*, *, *, x) &= (0, x, x, 0) \\
\text{WIRE}_{i,o}(c \in \{0, 1\}^4) &= k \in \{0, \dots, 3\} \mapsto \begin{cases} c(i) & \text{if } k = o \\ 0 & \text{else} \end{cases}
\end{aligned}$$

for any  $i \neq o \in \{0, \dots, 3\}$ . Note that any function  $f$  above is such that

$$f(0, 0, 0, 0) = (0, 0, 0, 0).$$

We denote by  $\text{Img}(f)$  the set of 4-uple that can be obtained as an image of  $f$ . The  $\text{WIRE}_{i,o}$  functions are just all the possible ways to read a bit on one side and transmit it to another side. Together with the NOP and FORK function they represent the basic planar wiring toolkit denoted  $W$  in the sequel. The NOP gate is special in that one considers it has 4 inputs and 4 outputs.

The two circuit simulation modes share the same block representation of circuit and information, but they differ in their requirement about the dynamical evolution of blocks. In the first mode, called *transient mode*, the gates can be used only once and nothing is granted concerning their evolution afterwards. The second mode, called *repeatable mode*, asks for each gate to go back to some acceptable state each time they are used so that they can be used again. Both modes require the simulation to work in constant time.

Let  $\mathcal{G} \subseteq \{\text{AND}, \text{OR}, \text{CROSS}\}$  be a set of gates. Let  $F$  be a CA with states set  $Q$  and  $N > 0$  be an integer. Consider a set  $V \subseteq Q^{N \times N}$  of patterns, the valid blocks, each of which as a type  $f_u$  where  $f \in \mathcal{G} \cup W$  and  $u \in \text{Img}(f)$  and such that, for any  $f_u$ , there is some block of  $V$  of type  $f_u$ . If a block  $B \in V$  has type  $f_{(a,b,c,d)}$  for some  $f$ , we say it has *north value*  $a$ , *east value*  $b$ , *south value*  $c$  and *west value*  $d$ . Finally let  $\Delta > 0$  be some constant. A configuration is *valid* if it is a concatenation of valid blocks where output sides of a block must face input sides of its corresponding neighbors. Given a block  $B \in V$  of type  $f_u$  in a valid configuration, we say that it *makes the correct transition* if it becomes a block of type  $f_v$  after  $\Delta$  steps where  $v = f(n, e, s, w)$  is the output of  $f$  on the input read from surrounding blocks, precisely: the block at the north of  $B$  has south value  $n$ , the block at the east of  $B$  has west value  $e$ , etc.

**Transient simulation.** We say that  $F$  *simulates the set of gates*  $\mathcal{G}$  *in transient mode* with delay  $\Delta$  and valid blocks  $V$  if for any valid configuration  $c$ , the configuration  $F^\Delta(c)$  is valid and for any  $f \in \mathcal{G} \cup W$ , any block of type  $f_{(0,0,0,0)}$  in  $c$  makes the correct transition.

**Repeatable simulation.** The simulation is *repeatable* if any block in any valid configuration makes the correct transition.

Before stating some theorems, let us define a decision problem associated to any 2D cellular automaton that will serve as a benchmark to separate the two kinds of circuit simulation above.

**Definition 4** Let  $F$  be any 2D CA of radius  $r$  and alphabet  $Q$ , and  $\phi$  a non-decreasing function such that  $1 \leq \phi(n) \leq 2^{O(n)}$ . The prediction problem  $\text{CYCLE}_F^\phi$  is defined as follows:

- input: a periodic configuration  $c$  of period  $n \times n$
- output: is the length of the temporal cycle reached from  $c$  strictly greater than  $\phi(n)$ ?

The first mode of simulation (the repeatable mode) is the strongest one, and is actually equivalent to intrinsic universality even if we use only monotone gates. In this case, although the definition does not explicitly provide crossing gates, it is always possible to realize a dynamical crossing and build arbitrary reusable bloc elements leading to intrinsic universality.

**Theorem 3 ([11])** *A 2D CA  $F$  is intrinsically universal if and only if it can simulate a AND, OR circuitry in a repeatable way. In this case  $\text{PRED}_F$  is P-complete,  $\text{UBPRED}_F$  is  $\Sigma_1^0$ -complete and  $\text{CYCLE}_F^\phi$  is PSPACE-complete for some  $\phi$ .*

To illustrate the difference between the two modes we shall use the symmetric signed majority cellular automaton: it is essentially a majority rule where the state of each neighboring cell can be inverted or not before evaluating majority, this being done according to a local invariant sign vector and in a symmetric way: if cell  $z$  inverts the value of its neighbor  $z'$ , then  $z'$  will also invert the value of  $z$ . We use the von Neumann neighborhood  $V = \{(0, 0), (0, 1), (1, 0), (0, -1), (-1, 0)\}$ . The symmetric signed majority cellular automaton  $F_1$  is defined over state set  $Q = \{-1, 1\}$ <sup>6</sup>. To simplify notation, we will see each state  $q \in Q$  as a pair  $(I(q), S(q))$  where  $I(q) \in \{0, 1\}$  represent the *inner state* and  $S(q) \in \{-1, 1\}^V$  is a *sign vector* associating a sign to each neighbor of the von Neumann neighborhood. For any configuration  $c \in Q^{\mathbb{Z}^2}$ , any cell  $z$  and any cell  $z' \in z + V$  we define the symmetric weight  $w_{zz'} \in \{-1, 1\}$  as  $w_{zz'} = (S(c_z)(z' - z))(S(c_{z'})(z - z'))$ . We note that  $w_{zz'} = w_{z'z}$ , hence the name symmetric.  $F_1$  is then defined as follows.

$$F_1(c)_z = (\alpha, S(c_z))$$

where

$$\alpha = \begin{cases} 1 & \text{if } \sum_{z' \in z+V} w_{zz'} I(c_{z'}) > 0, \\ -1 & \text{otherwise.} \end{cases}$$

The following theorem shows that transient simulations are strictly weaker than repeatable simulations,  $F_1$  being an example capable of the former, but not the latter.

**Theorem 4 ([11])** *If a 2D CA  $F$  can simulate a {AND, OR, CROSS} circuitry in transient mode, then its associated problem  $\text{PRED}_F$  is P-complete.  $F_1$  defined above can simulate a {AND, OR, CROSS} circuitry in transient mode. However,  $F_1$  is not intrinsically universal and such that the problem  $\text{CYCLE}_{F_1}^\phi$  is in P if  $\phi \equiv 1$  and trivial else.*

## 5 Hidden universality

When proving that some cellular automaton is computationally universal, it can be acceptable to avoid a precise definition of universality if the construction makes it clear enough. However, a precise definition seems necessary when ones want to show that some cellular automaton is **not** universal. To avoid formalism, one could be tempted to use a shorter path: prove that, according to some well-chosen parameter, the considered cellular automaton is too simple to be universal. The intuition is that a universal cellular automaton should have roughly the highest complexity for the parameter. This approach can be made precise and yield some proof tools of non-universality in some contexts [12]. The purpose of this section is to recall that things can get counter-intuitive and such a parameter must be chosen carefully.

### 5.1 Hidden Minsky machines simulation

The *limit set* of a cellular automaton  $F$  is the nonempty closed subset

$$\Omega_F = \bigcap_{t \in \mathbb{N}} F^t(X).$$

It represents the set of configurations that may appear arbitrarily far in the evolution and the restriction of  $F$  to  $\Omega_F$  is often considered as the asymptotic dynamics of  $F$ . The limit language is the set of finite patterns that occur in some configuration of  $\Omega_F$ . It is not difficult to see that the limit language is always co-recursively enumerable. However, there are known examples of non-recursive ones [15]. The attentive reader of [4] has probably spotted the affirmation that universal cellular automata have a non-recursive limit set. Depending on the definition of universality, this affirmation can be false. The following theorem shows that arbitrary Minsky machine simulations can be realized while maintaining a simple limit set (see [10] for the precise definition of simulation).

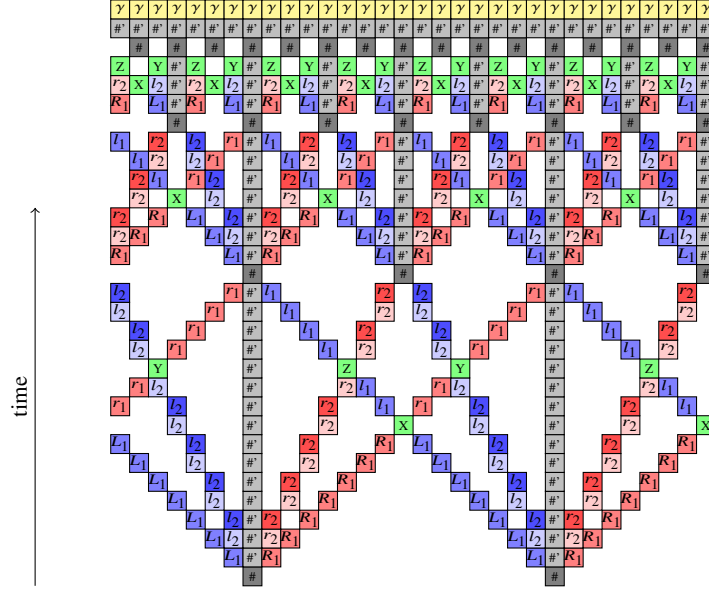
#### **Theorem 5 ([10])**

*For any Minsky machine  $M$  there exists a CA of dimension 1 that simulates  $M$  but whose limit language is regular.*

### 5.2 Hidden intrinsic universality

Following Theorem 5, one can go one step further and hide intrinsic universality behind a simple limit set (at the price of a complexity increase from regular to NLOGSPACE). The main trick of the next theorem is inspired from [16]: adding to a given cellular automaton  $F$  on alphabet  $Q$ , a firing squad component (see Figure 2)

that is able to fill-in the limit set restricted to the  $Q$  component, and therefore make it simple independently of  $F$ .



**Fig. 2** J. Kari's firing squad trick: a synchronous apparition of  $\gamma$  can be triggered arbitrarily far in time, thus allowing to complete the limit set on some component of states to the full-shift.

**Theorem 6 ([13])** *There exists an intrinsically universal CA whose limit language is NLOGSPACE.*

Given  $n \in \mathbb{N}$ , the *column factor* of width  $n$  of  $F$ ,  $\Sigma_n(F)$ , is the set of columns that can appear in space-time diagrams of  $F$ :

$$\Sigma_n(F) = \{(u_t)_{t \in \mathbb{N}} : u_t \in Q^n, u_t = F^t(c)_{[1,n]}, c \in Q^{\mathbb{Z}^d}\}.$$

To  $\Sigma_n(F)$  we associate its language of finite patterns  $L(\Sigma_n(F))$  defined as the set of words  $u_t \cdots u_{t+k}$  for some  $(u_t)_{t \in \mathbb{N}} \in \Sigma_n(F)$  and  $t, k \in \mathbb{N}$ .

The approach of [7] to define universality for general dynamical systems translates into the following in our settings. To  $F$  we associate the model checking problem:

- **input:**  $n$  and a regular language  $L_n$  over alphabet  $Q^n$ ,
- **question:** decide whether  $L_n$  intersects  $L(\Sigma_n(F))$ .

$F$  is *BDK-universal* if its associated model checking problem is r.e.-complete. Like for limit sets, column factors can be filled up and thus simplified by increasing the alphabet starting from an arbitrarily complex cellular automaton.

**Theorem 7 ([13])** *There exists an intrinsically universal CA  $F$  such that  $L(\Sigma_n(F))$  are regular languages computable from  $n$ . In particular, such  $F$  is not BDK-universal.*

## References

1. E. R. Banks. Universality in cellular automata. In *Eleventh Annual Symposium on Switching and Automata Theory*, Santa Monica, California, 1970. IEEE.
2. Olivier Carton, Bruno Guillon, and Fabian Reiter. Counter machines and distributed automata - A story about exchanging space and time. In *Cellular Automata and Discrete Complex Systems - 24th IFIP WG 1.5 International Workshop, AUTOMATA 2018, Ghent, Belgium, June 20-22, 2018, Proceedings*, pages 13–28, 2018.
3. M. Cook. Universality in elementary cellular automata. *Complex Systems*, 15:1–40, 2004.
4. K. Čulik, II, J. Pahl, and S. Yu. On the limit sets of cellular automata. *SIAM Journal on Computing*, 18(4):831–842, August 1989.
5. M. Delorme, J. Mazoyer, N. Ollinger, and G. Theyssier. Bulking II: Classifications of cellular automata. *Theoretical Computer Science*, 412:3881–3905, 2011.
6. 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.
7. Jean-Charles Delvenne, Petr Kurka, and Vincent D. Blondel. Decidability and universality in symbolic dynamical systems. *Fundam. Inform.*, 74(4):463–490, 2006.
8. B. Durand and Z. Róka. *Cellular Automata: a Parallel Model*, volume 460 of *Mathematics and its Applications.*, chapter The game of life:universality revisited., pages 51–74. Kluwer Academic Publishers, 1999.
9. E. Goles, N. Ollinger, and G. Theyssier. Introducing freezing cellular automata. In J. Kari, I. Törmä, and M. Szabados, editors, *Exploratory Papers of Cellular Automata and Discrete Complex Systems (AUTOMATA 2015)*, pages 65–73, 2015.
10. Eric Goles, Alejandro Maass, and Servet Martínez. On the limit set of some universal cellular automata. *Theoretical Computer Science*, 110:53–78, 1993.
11. Eric Goles, Pedro Montealegre, Kévin Perrot, and Guillaume Theyssier. On the complexity of two-dimensional signed majority cellular automata. *J. Comput. Syst. Sci.*, 91:1–32, 2018.
12. Eric Goles Ch., Pierre-Etienne Meunier, Ivan Rapaport, and Guillaume Theyssier. Communication complexity and intrinsic universality in cellular automata. *Theor. Comput. Sci.*, 412(1-2):2–21, 2011.
13. Pierre Guillon, Pierre-Etienne Meunier, and Guillaume Theyssier. Clandestine simulations in cellular automata. In Jarkko Kari, editor, *Second Symposium on Cellular Automata "Journées Automates Cellulaires", JAC 2010, Turku, Finland, December 15-17, 2010. Proceedings*, pages 133–144. Turku Center for Computer Science, 2010.
14. G. A. Hedlund. Endomorphisms and Automorphisms of the Shift Dynamical Systems. *Mathematical Systems Theory*, 3(4):320–375, 1969.
15. Lyman P. Hurd. Nonrecursive cellular automata invariant sets. *Complex Systems*, 4:131–138, 1990.
16. J. Kari. Rice’s theorem for the limit sets of cellular automata. *Theoretical Computer Science*, 127:229–254, 1994.
17. P. Kůrka. *Topological and symbolic dynamics*. Société Mathématique de France, 2003.
18. Martin Kutrib and Andreas Malcher. Cellular automata with sparse communication. *Theor. Comput. Sci.*, 411(38-39):3516–3526, 2010.
19. M. Minsky. *Computation: Finite and Infinite Machines*. Prentice Hall, Englewoods Cliffs, 1967.
20. Turlough Neary and Damien Woods. P-completeness of cellular automaton rule 110. In *In International Colloquium on Automata Languages and Programming (ICALP), volume 4051 of LNCS*, pages 132–143. Springer, 2006.
21. N. Ollinger. Universalities in cellular automata a (short) survey. In B. Durand, editor, *First Symposium on Cellular Automata "Journées Automates Cellulaires" (JAC 2008), Uzès, France, April 21-25, 2008. Proceedings*, pages 102–118. MCCME Publishing House, Moscow, 2008.
22. Nicolas Ollinger. Universalities in cellular automata a (short) survey. In *JAC*, pages 102–118, 2008.

23. Nicolas Ollinger and Guillaume Theyssier. Freezing, bounded-change and convergent cellular automata. *CoRR*, abs/1908.06751, 2019.
24. R. Vollmar. On cellular automata with a finite number of state changes. In Walter Knödel and HansJürgen Schneider, editors, *Parallel Processes and Related Automata / Parallele Prozesse und damit zusammenhängende Automaten*, volume 3 of *Computing Supplementum*, pages 181–191. Springer Vienna, 1981.