A search algorithm for subshift attractors of cellular automata

Enrico Formenti[1], Petr Kůrka[1,2], and Ondřej Zahradník[3]

[1] Laboratoire I3S, Université de Nice Sophia Antipolis,
2000, route des Lucioles, Les Algorithmes - bât Euclide B, BP 121,
06903 Sophia Antipolis - Cedex, France
[2] Center for Theoretical Study,
Academy of Sciences and Charles University in Prague,
Jilská 1, CZ-11000 Praha 1, Czechia
[3] Faculty of Mathematics and Physics, Charles University in Prague,
Malostranské náměstí 25, CZ-11800 Praha 1, Czechia

Abstract. We describe a heuristic algorithm which searches for spreading clopen sets of a cellular automaton. Then the algorithms searches for the corresponding subshift attractors (which are omega-limits of spreading sets found) as forward images of joins of signal subshifts.

Keywords: sofic subshifts, subshift attractors, spreading sets, signal subshifts.

1 Introduction

The concept of attractor and of subshift attractor in particular is essential for understanding the dynamics of cellular automata (see Kůrka [1], Formenti and Kůrka [2]). The special case of maximal attractor, or omega-limit set has been one of the most intensively studied structures of cellular automata theory (see e.g., Culik et al., [3], Maass [4] or Kari [5]). An attractor is the omega-limit of a clopen invariant set. If the attractor is a subshift, the clopen set must be not only invariant but also spreading both to the left and to the right. Clopen sets can be represented by finite sets of words, and the properties of invariance and spreading translate into simple combinatorial properties of these word sets. However, to find all spreading sets of a given CA and to determine the structure of the corresponding subshifts attractors is not always an easy task.

In Formenti and Kůrka [6] we have developped a method which searches for the maximal attractor of a CA in the form of a forward image of join of signal subshifts. A signal subshift (see Kůrka [1]) consists of configurations which after certain time reappear shifted. Signal subshifts are of finite type, can be easily computed, and are contained in the maximal attractor. Having found a certain number of subshift attractors, we construct their join, which is a larger sofic subshift still included in the maximal attractor. Forward images of this join sometimes attain an invariant subshift, which is also sofic. To test whether the obtained subshift is already the maximal attractor, we use a condition of decreasing preimages. A subshift has decreasing preimages, if each preimage of a forbidden word contains, as a factor, a shorter forbidden word. The condition can be tested algorithmically on sofic subshifts. Thus, if the test gives positive result on an invariant sofic subshift, then the subshift is the maximal attractor.

In the present paper we extend this construction method to general subshift attractors and report on its computer implementation. To reduce the more general concept of subshift attractor to the concept of maximal attractor, we use a general concept of cellular automaton as a shift-commuting continuous mappings $F : X \to X$ on a mixing subshift $X \subseteq A^{\mathbb{Z}}$. As argued by Boyle and Kitchens [7] this is the right setting for cellular automata in the context of symbolic dynamics. Then we get a theorem that each subshift attractor is a maximal attractor (of another CA).

As a byproduct of the construction of the algorithm, we show that a CA whose maximal attractor can be constructed as a forward image of joins of signal subshifts has only a finite number of infinite transitive signal subshifts. Dynamics of such CA can be viewed as dynamics of signals which propagate through a neutral medium (of shift-periodic configurations) and transform to simpler signals when they meet.

2 Subshifts

For a finite alphabet A, denote by $A^* := \bigcup_{n\geq 0} A^n$ the set of words over A and by $A^+ := \bigcup_{n>0} A^n$ the set of words of positive length. The length of a word $u = u_0 \dots u_{n-1} \in A^n$ is denoted by $|u| := n$ and the word of zero length is λ. We say that $u \in A^*$ is a subword of $v \in A^*$ ($u \sqsubseteq v$) if there exists k such that $v_{k+i} = u_i$ for all $i < |u|$. We denote by $u_{[i,j)} = u_i \dots u_{j-1}$ and $u_{[i,j]} = u_i \dots u_j$ subwords of u associated to intervals. We denote by $A^{\mathbb{Z}}$ the space of A**-configurations**, or doubly-infinite sequences of letters of A equipped with the metric $d(x, y) := 2^{-n}$, where $n = \min\{i \geq 0 : x_i \neq y_i \text{ or } x_{-i} \neq y_{-i}\}$.

The **shift map** $\sigma : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ is defined by $\sigma(x)_i := x_{i+1}$. For any $u \in A^+$ we have a σ-periodic configuration $u^\infty \in A^{\mathbb{Z}}$ defined by $(u^\infty)_i = u_{|i| \bmod |u|}$. A **subshift** is a nonempty subset $\Sigma \subseteq A^{\mathbb{Z}}$, which is closed and strongly σ-invariant, i.e., $\sigma(\Sigma) = \Sigma$. For a given alphabet A define the bijective k**-block code** $\alpha_k : A^{\mathbb{Z}} \to (A^k)^{\mathbb{Z}}$ by $\alpha_k(x)_i = x_{[i,i+k)}$. For a subshift $\Sigma \subseteq A^{\mathbb{Z}}$ denote by $\Sigma^{[k]} = \alpha_k(\Sigma) \subseteq (A^k)^{\mathbb{Z}}$ its k**-block encoding**. Then $\alpha_k : \Sigma \to \Sigma^{[k]}$ is a conjugacy as it commutes with the shift maps $\alpha_k \sigma = \sigma \alpha_k$.

For a subshift Σ there exists a set $D \subseteq A^*$ of forbidden words such that $\Sigma = \mathcal{S}_D := \{x \in A^{\mathbb{Z}} : \forall u \sqsubseteq x, u \notin D\}$. A subshift is uniquely determined by its **language** $\mathcal{L}(\Sigma) := \{u \in A^* : \exists x \in \Sigma, u \sqsubseteq x\}$. We denote by $\mathcal{L}^n(\Sigma) := \mathcal{L}(\Sigma) \cap A^n$. The language of **first offenders** of Σ is

$$\mathcal{D}(\Sigma) := \{u \in A^+ \setminus \mathcal{L}(\Sigma) : u_{[0,|u|-1)}, u_{[1,|u|)} \in \mathcal{L}(\Sigma)\},$$

so $\Sigma = \mathcal{S}_{\mathcal{D}(\Sigma)}$. If $x \in A^{\mathbb{Z}}$ is a configuration and $I \subseteq \mathbb{Z}$ is an interval, denote by $x_{|I} : I \to A$ the restriction of x to I. The **extended language** of Σ is $\tilde{\mathcal{L}}(\Sigma) = \{x_{|I} : x \in \Sigma, \quad I \subseteq \mathbb{Z} \quad \text{is an interval}\}$.

A subshift $\Sigma \subseteq A^{\mathbb{Z}}$ is **transitive**, if for any words $u, v \in \mathcal{L}(\Sigma)$ there exists $w \in A^*$ such that $uwv \in \mathcal{L}(\Sigma)$. A subshift $\Sigma \subseteq A^{\mathbb{Z}}$ is **mixing**, if for any words $u, v \in \mathcal{L}(\Sigma)$ there exists $n > 0$ such that for all $m > n$ there exists $w \in A^m$ such that $uwv \in \mathcal{L}(\Sigma)$.

Definition 1 *Given an integer $c \geq 0$, the c-**join** $\Sigma_0 \overset{c}{\vee} \Sigma_1$ of subshifts $\Sigma_0, \Sigma_1 \subseteq A^{\mathbb{Z}}$ consists of all configurations $x \in A^{\mathbb{Z}}$ such that either $x \in \Sigma_0 \cup \Sigma_1$, or there exist integers b, a such that $b - a \geq c$, $x_{(-\infty,b)} \in \tilde{\mathcal{L}}(\Sigma_0)$, and $x_{[a,\infty)} \in \tilde{\mathcal{L}}(\Sigma_1)$.*

Examples of joins of subshifts are in Figures 3 and 5.

Proposition 2 *The c-join of two subshifts is a subshift and the operation of c-join is associative. A configuration $x \in A^{\mathbb{Z}}$ belongs to $\Sigma_1 \overset{c}{\vee} \cdots \overset{c}{\vee} \Sigma_n$ iff there exist integers $k > 0$, $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, and intervals $I_1 = (a_1, b_1), I_2 = [a_2, b_2), \ldots, I_k = [a_k, b_k)$ such that $a_1 = -\infty$, $b_k = \infty$, $a_j < a_{j+1}$, $b_j < b_{j+1}$, $b_j - a_{j+1} \geq c$, and $x_{|I_j} \in \tilde{\mathcal{L}}(\Sigma_{i_j})$.*

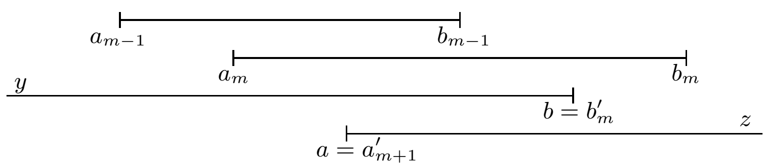

Fig. 1. Associativity of the join.

Proof. It is clear that the c-join of two subshifts is a subshift. We prove by induction the formula for $\Sigma_1 \overset{c}{\vee} \cdots \overset{c}{\vee} \Sigma_n$. By definition, the formula holds for $n = 2$. Let $x \in (\Sigma_1 \overset{c}{\vee} \cdots \overset{c}{\vee} \Sigma_{n-1}) \overset{c}{\vee} \Sigma_n$, so there exists $y \in \Sigma_1 \overset{c}{\vee} \cdots \overset{c}{\vee} \Sigma_{n-1}$, $z \in \Sigma_n$ and integers a, b such that $x_{(-\infty,b)} = y_{(-\infty,b)}$, $x_{[a,\infty)} = z_{[a,\infty)}$ and $b - a \geq c$. By the induction hypothesis there exist integers k and intervals $I_1, \ldots I_k$ such that $y_{|I_j} \in \tilde{\mathcal{L}}(\Sigma_{i_j})$. Let m be the unique index such that $b_{m-1} < b \leq b_m$ (see Figure 1). Set $i'_j := i_j$ for $j \leq m$ and $i'_{m+1} = n$. For $k \leq m+1$ define intervals $I'_j := [a'_j, b'_j)$ by

$$a'_j := \begin{cases} a_j & \text{for} \quad j \leq m \\ \max\{a, a_m + 1\} & \text{for} \quad j = m+1 \end{cases}, \qquad b'_j := \begin{cases} b_j & \text{for} \quad j < m \\ b & \text{for} \quad j = m \\ \infty & \text{for} \quad j = m+1 \end{cases}$$

Then clearly $x_{|I'_j} \in \tilde{\mathcal{L}}(\Sigma_{i'_j})$ and $b'_j - a'_{j+1} = b_j - a_{j+1} \geq c$ for $j < m$. If $a'_{m+1} = a_m + 1$, then $b'_m - a'_{m+1} = b - a_m - 1 \geq b_{m-1} - a_m \geq c$. If $a'_{m+1} = a$, then $b'_m - a'_{m+1} = b - a \geq c$. Similarly it can be shown that the formula holds for $\Sigma_1 \overset{c}{\vee} (\Sigma_2 \overset{c}{\vee} \cdots \overset{c}{\vee} \Sigma_n)$ provided it holds for $\Sigma_2 \overset{c}{\vee} \cdots \overset{c}{\vee} \Sigma_n$. This proves associativity. □

3 Sofic subshifts

A subshift $\Sigma \subseteq A^{\mathbb{Z}}$ is **sofic** if its language $\mathcal{L}(\Sigma)$ is regular. Sofic subshifts are usually described by finite automata or by labelled graphs. A **non-deterministic finite automaton** (NFA) over A (see e.g., Hopcroft and Ullmann [8]) is a system $\mathcal{A} = (Q, \delta, I, R)$, where Q is a finite set of states, $\delta : Q \times A \to \mathcal{P}(Q)$ is the transition function ($\mathcal{P}(Q)$ is the set of subsets of Q), $I \subseteq Q$ is the set of initial states, and $R \subseteq Q$ is the set of rejecting states. A word $u \in A^*$ is accepted by $\mathcal{A}$ if there exists $q \in Q^{|u|+1}$ such that $q_0 \in I$, $q_{i+1} \in \delta(q_i, u_i)$ for $i < |u|$, and $q_{|u|} \notin R$. A NFA $\mathcal{A} = (Q, \delta, I, R)$ is **deterministic finite automaton** (DFA), if I and all $\delta(q, a)$ are singletons. A language $L \subseteq A^*$ is **regular**, if there exists a NFA such that L consists of accepted words. In this case L is also accepted by some DFA.

A **labelled graph** over an alphabet A is a structure $G = (V, E, s, t, l)$, where V is a finite set of vertices, E is a finite set of edges, $s, t : E \to V$ are the **source** and **target maps**, and $l : E \to A$ is a **labelling function**. A finite or infinite word $w \in E^* \cup E^{\mathbb{Z}}$ is a **path** in G if $t(w_i) = s(w_{i+1})$ for all i. The source and target of a finite path $w \in E^n$ are $s(w) := s(w_0)$, $t(w) := t(w_{n-1})$. The **label** of a path is defined by $l(w)_i := l(w_i)$. A subshift Σ is sofic iff there exists a labelled graph G such that $\Sigma = \Sigma_G$ is the set of labels of all doubly infinite paths in G. In this case we say that G is a **presentation** of Σ (see e.g. Lind and Marcus [9], or Kitchens [10]). Given a labelled graph $G = (V, E, s, t, l)$, take a NFA $\mathcal{A} = (V, \delta, V, \emptyset)$, where $\delta(u, a) = \{v \in V : \exists e \in E, s(e) = u, t(e) = v, l(e) = a\}$. Then $\mathcal{A}$ accepts the language $\mathcal{L}(\Sigma_G)$.

A subshift $\Sigma \subseteq A^{\mathbb{Z}}$ is **of finite type (SFT)**, if $\Sigma = \mathcal{S}_D$ for some finite set $D \subseteq A^*$ of forbidden words. The words of D can be assumed to be all of the same length, which is called the order $\mathfrak{o}(\Sigma)$ of Σ. A configuration $x \in A^{\mathbb{Z}}$ belongs to Σ iff $x_{[i,i+\mathfrak{o}(\Sigma))} \in \mathcal{L}(\Sigma)$ for all $i \in \mathbb{Z}$. Any SFT is sofic: if $p = \mathfrak{o}(\Sigma) - 1$, the **canonical graph** $G = (V, E, s, t, l)$ of Σ is given by $V = \mathcal{L}^p(\Sigma)$, $E = \mathcal{L}^{p+1}(\Sigma)$, $s(u) = u_{[0,p)}$, $t(u) = u_{[1,p]}$ and $l(u) = u_p$. If $\Sigma \subset A^{\mathbb{Z}}$ is a finite subshift, then it is of finite type and each its configuration is σ-periodic. The period $\mathfrak{p}(\Sigma)$ of Σ is then the smallest positive integer $\mathfrak{p}(\Sigma)$, such that $\sigma^{\mathfrak{p}(\Sigma)}(x) = x$ for all $x \in \Sigma$.

A labelled graph $G = (V, E, s, t, l)$ is **connected** if for any two vertices $q, q' \in V$ there exists a path $w \in E^*$ from q to q'. A **subgraph** of a graph G is a graph $G' = (V', E', s', t', l')$, such that $V' \subseteq V$, $E' = \{e \in E : s(e) \in V' \;\&\; t(e) \in V'\}$, and s', t', l' coincide respectively with s, t, l on E'. A **connected component** of G is a subgraph of G which is connected and maximal with this property. The subshift of a connected graph is transitive. Conversely, every transitive sofic subshift $\Sigma \subseteq A^{\mathbb{Z}}$ has a connected presentation. A labelled graph $G = (V, E, s, t, l)$ is **aperiodic** if the set of verices cannot be partitioned into disjoint union $V = V_0 \cup \cdots \cup V_{p-1}$ such that $p \geq 2$ and if $s(e) \in V_i$, then $t(e) \in V_{(i+1) \bmod p}$. A sofic subshift is mixing iff it has a connected and aperiodic presentation.

For a labelled graph $G = (V, E, s, t, l)$ and $k > 1$ define the **k-block graph** $G^{[k]} := (V', E', s', t', l')$, where $V' \subseteq E^{k-1}$ is the set of paths of G of length $k - 1$, $E' \subseteq E^k$ is the set of paths of G of length k, s' and t' are the prefix and suffix maps and $l' : E' \to A^k$ is defined by $l'(u)_i = l(u_i)$. Then $G^{[k]}$ is a presentation of the k-block encoding $(\Sigma_G)^{[k]}$ of Σ_G. From $G^{[k]}$ we may in turn get a presentation of Σ_G, if we replace l' by its composition with a projection $\pi_i : A^k \to A$ defined by $\pi(u)_i := u_i$, where $i < k$.

Given two sofic subshifts $\Sigma_0, \Sigma_1 \subseteq A^{\mathbb{Z}}$, their union and intersection (provided non-empty) are sofic subshifts. Moreover there exists an algorithm which constructs a presentation of $\Sigma_0 \cup \Sigma_1$ and $\Sigma_0 \cap \Sigma_1$ from those of Σ_0 and Σ_1. It is also decidable whether $\Sigma_0 \subseteq \Sigma_1$. Given a labelled graph G it is decidable whether Σ_G is a SFT (see Lind and Marcus [9], page 94). It is also decidable, whether a SFT is mixing.

Proposition 3 *Let $\Sigma_0, \Sigma_1 \subseteq A^{\mathbb{Z}}$ be sofic subshifts and $c \geq 0$. Then $\Sigma_0 \overset{c}{\vee} \Sigma_1$ is a sofic subshift. There exists an algorithm which constructs a presentation G of $\Sigma_0 \overset{c}{\vee} \Sigma_1$ from the presentations G_i of Σ_i. Moreover, G has the same connected components as the disjoint union $G_0 \cup G_1$.*

Proof. Let $G_i = (V_i, E_i, s_i, t_i, l_i)$ be presentations of $\Sigma_i^{[c]}$, and assume that $V_0 \cap V_1 = \emptyset$ and $E_0 \cap E_1 = \emptyset$. Construct $G = (V, E, s, t, l)$, where $V = V_0 \cup V_1$,

$$E = E_0 \cup E_1 \cup \{(e_0, e_1) \in E_0 \times E_1 : \ l_0(e_0) = l_1(e_1)\}.$$

The source, target and label maps extend s_i, t_i, l_i. For the new edges we have $s(e_0, e_1) = s_0(e_0)$, $t(e_0, e_1) = t_1(e_1)$, $l(e_0, e_1) = l_0(e_0) = l_1(e_1)$. Then $\Sigma_G = (\Sigma_0 \overset{c}{\vee} \Sigma_1)^{[c]}$, and we get a presentation of $\Sigma_0 \overset{c}{\vee} \Sigma_1$, if we compose l with a projection $\pi_i : A^c \to A$. Clearly, the connected components of $G_0 \cup G_1$ are not changed by the new edges. □

Lemma 4 *If $L \subseteq A^*$ is a regular language and a, b non-negative integers, then the language $L_{[a,b)} := \{v \in A^* : \ |v| \geq a + b, v_{[a,|v|-b)} \in L\}$ is regular.*

(1) If q is the number of states of a NFA which accepts L, then there exists a NFA which accepts $L_{[a,b)}$ with $q + a + b$ states.
(2) If q is the number of states of a DFA which accepts L, then there exists a DFA which accepts $L_{[a,b)}$ with $q + a + 2^b$ states.

The proof is left to the reader.

Proposition 5 *The language $\mathcal{D}(\Sigma)$ of first offenders of a sofic subshift Σ is regular.*

Proof. We have $\mathcal{D}(\Sigma) = (A^+ \setminus \mathcal{L}(\Sigma)) \cap \mathcal{L}(\Sigma)_{[0,1)} \cap \mathcal{L}(\Sigma)_{[1,0)}$. □

4 Cellular automata and subshift attractors

We use a general concept of cellular automaton whose relevance has been argued in Boyle and Kitchens [7].

Definition 6 *A cellular automaton is a pair (X, F), where $X \subseteq A^{\mathbb{Z}}$ is a mixing subshift of finite type, and $F : X \to X$ is a continuous mapping which commutes with the shift, i.e., $F\sigma = \sigma F$.*

For a cellular automaton (X, F) there exists a **local rule** $f : \mathcal{L}^{a-m+1}(X) \to \mathcal{L}^1(X)$ such that $F(x)_i = f(x_{[i+m,i+a]})$. Here $m \leq a$ are integers called **memory** and **anticipation**. The **diameter** of (X, F) is $d := a - m$. If $a = -m \geq 0$, then we say that a is the **radius** of (X, F). The local rule can be extended to a map $f : \mathcal{L}(X) \to \mathcal{L}(X)$ by $f(u)_i := f(u_{[i,i+d]})$ for $0 \leq i < |u| - d$. Thus $|f(u)| = \max\{|u| - d, 0\}$.

The **cylinder set** of a word $u \in \mathcal{L}(X)$ located at $l \in \mathbb{Z}$ is $[u]_l := \{x \in X : x_{[l,l+|u|)} = u\}$. The cylinder set of a finite set $U \subset A^*$ is $[U]_l = \bigcup_{u \in U} [u]_l$. We write also $[u] := [u]_0$ and $[U] := [U]_0$. Every clopen set of X is a cylinder set of a finite set $U \subset A^*$.

Proposition 7 *Let (X, F) be a CA. If $\Sigma \subseteq X$ is a sofic subshift, then $F(\Sigma)$ and $F^{-1}(\Sigma) := \{x \in X : \; F(x) \in \Sigma\}$ are sofic subshifts and there exists an algorithm which constructs their graphs from the local rule f, the graph of Σ and the set of forbidden words of X.*

Proof. We give an explicit (easy) construction which will be used later on. Let $f : \mathcal{L}^{d+1}(X) \to \mathcal{L}^1(X)$ be a local rule of F. We can assume that $d + 1 \geq \mathfrak{o}(X)$, where $\mathfrak{o}(X)$ is the order of X. Let $G = (V, E, s, t, l)$ be a presentation of Σ. Let V_0 be the set of paths of G of length d, and let E_0 be the set of paths of G of length $d+1$. Define the source and target maps by $s_0(u) = u_{[0,d)}$, $t_0(u) = u_{[1,d]}$. Then $G_0 = (V_0, E_0, s_0, t_0, f)$ is a presentation of $F(\Sigma)$.
Set $E_1 = \{(e, u) \in E \times \mathcal{L}^{d+1}(X) : \; l(e) = f(u)\}$, and define s_1, t_1 by $s_1(e, u) = (s(e), u_{[0,d)})$, $t_1(e, u) = (t(e), u_{[1,d]})$. Finally set $V_1 = s_1(E_1) \cup t_1(E_1) \subseteq V \times \mathcal{L}^d(X)$ and define the labelling function by $l_1(e, u) = u_d$. Then $G_1 = (V_1, E_1, s_1, t_1, l_1)$ is a presentation of $F^{-1}(\Sigma)$.

Definition 8 *Let (X, F) be a cellular automaton.*

(1) A set $W \subseteq X$ is **invariant***, if $F(W) \subseteq W$.*
(2) A set $Y \subseteq X$ is **attractor***, if there exists a clopen invariant set $W \subseteq X$ such that $Y = \Omega_F(W) = \bigcap_{n \geq 0} F^n(W)$.*
(3) The maximal attractor of (X, F) is $\Omega_F(X)$.
(4) A clopen F-invariant set $W \subseteq X$ is **spreading to the right** *(or* **left***), if $F^k(W) \subseteq \sigma^{-1}(W)$ (or $F^k(W) \subseteq \sigma(W)$) for some $k > 0$.*
(5) A clopen F-invariant set $W \subseteq X$ is **spreading***, if it is spreading both to the right and to the left.*

Proposition 9 *Let (X, F) be a cellular automaton and $W \subseteq X$ a clopen F-invariant set. Then $\Omega_F(W)$ is a subshift attractor iff W is spreading.*

The proof is the same as in [2].

5 Invariant and spreading word sets

Since every clopen set is a finite union of cylinders of the same length, we can develop the theory of invariant and spreading sets $W \subseteq X$ using their word representations $U \subseteq \mathcal{L}^n(X)$ such that $W = [U]$. Recall that $[U] = [U]_0 = \{x \in X : x_{[0,n)} \in U\}$ and $\mathcal{S}_U$ is the SFT with forbidden set U. If $W = [U]$, then $\widetilde{\Omega}_\sigma(W) := \bigcap_{n \in \mathbb{Z}} \sigma^n(W) = X \cap \mathcal{S}_{A^n \setminus U}$. If W is spreading, then $\Omega_F(W) = \Omega_F(\widetilde{\Omega}_\sigma(W))$.

Definition 10 *Let (X, F) be a CA. We say that a word set $U \subseteq \mathcal{L}^n(X)$ is invariant (spreading), if $[U]$ is invariant (spreading). We say that U is a* **minimal invariant set**, *if U is invariant and there is no proper subset $V \subset U$ such that $\Omega_F([U]) = \Omega_F([V])$.*

The concept of minimality applies only to word sets and not to clopen sets. If $U \subseteq \mathcal{L}^n(X)$ is a minimal invariant word set, then there may exist a proper clopen subset $W \subset [U]$ such that W is spreading and $\Omega_F([U]) = \Omega_F(W)$.

Theorem 11 *Let (X, F) be a cellular automaton, $n \geq \mathfrak{o}(X)$, and $U \subseteq \mathcal{L}^n(X)$ a spreading word set. There exists a set $W \subseteq U$ such that $Y := X \cap \mathcal{S}_{A^n \setminus W}$ is a mixing subshift of finite type and $\Omega_F([U]) = \Omega_F([W]) = \Omega_F(Y)$, so $\Omega_F([U])$ is the maximal attractor of (Y, F).*

Proof. Construct a labelled graph $G = (V, U, s, t, l)$, where $s(u) := u_{[0,n-1)}$, $t(u) := u_{[1,n)}$, $l(u) := u_{n-1}$ and $V := s(U) \cup t(U)$. Then $\widetilde{\Omega}_\sigma([U]) = \Sigma_G$ is F-invariant. If G_0 is a connected component of G, then $F(\Sigma_{G_0})$ is a transitive subshift and there exists a connected component G_1 of G such that $F(\Sigma_{G_0}) \subseteq \Sigma_{G_1}$. In this way we get a dynamical system $\mathcal{F}$ on the finite set $\mathcal{G}$ of connected components of G. This system has some preperiod $m \geq 0$ and period $n > 0$, so $\mathcal{F}^{m+n} = \mathcal{F}^m$. For every $G_0 \in \mathcal{F}^m(\mathcal{G})$ we have $F^n(\Sigma_{G_0}) \subseteq \Sigma_{G_0}$. We show that $\mathcal{F}^m(\mathcal{G})$ consists of a unique connected component. Assume by contradiction that there exist different connected components G_0, G_1 such that $F^n(\Sigma_{G_i}) \subseteq \Sigma_{G_i}$ for $i = 0, 1$. Take some $x \in \Sigma_{G_0}$, $y \in \Sigma_{G_1}$. Since X is mixing, there exists $l > 0$ and a configuration $z \in X$ such that $z_{(-\infty,0)} := x_{(-\infty,0)}$, $z_{[l,\infty)} := y_{[l,\infty)}$. Since U is spreading, there exists $k > 0$ such that $F^k(z) \in \Sigma_G$. If r is the radius of F, then $F^{kn}(z)_{[-q-knr,-knr]} \in \Sigma_{G_0}$, and $F^{kn}(z)_{[l+knr,l+knr+q]} \in \Sigma_{G_1}$. This means that there is a path from G_0 to G_1 in G. Similarly, we prove that there exists a path from G_1 to G_0. Thus $G_0 = \mathcal{F}^n(\mathcal{G})$ is the unique connected component of G. We show that Σ_{G_0} is mixing. If not, then there exists a period $p > 1$ and a disjoint union of closed sets $\Sigma_{G_0} = Y_0 \cup \cdots \cup Y_{p-1}$ such that $\sigma(Y_i) \subseteq Y_{(i+1) \bmod p}$. Take some $x \in Y_0$, $y \in Y_1$ and construct a configuration $z \in A^{\mathbb{Z}}$ such that $z_{(-\infty,0)} := x_{(-\infty,0)}$, $z_{[l,\infty)} := y_{[l,\infty)}$ for some $l > 0$. Since U is spreading, there exists k such that $F^k(z) \in \Sigma_{G_0}$. However, $F^k(z)$ does not belong to any Y_i. This is a contradiction, so Σ_{G_0} is mixing. Let W be the set of edges of G_0. Then $\widetilde{\Omega}_\sigma([U]) = \widetilde{\Omega}_\sigma([W]) = \Sigma_{G_0}$ is mixing, and $\Omega_F([U]) = \Omega_F(\widetilde{\Omega}_\sigma([U])) = \Omega_F(\widetilde{\Omega}_\sigma([W]) = \Omega_F([W])$. Since U is spreading and Σ_{G_0} is F-invariant, $W \subseteq U$ is spreading. $\square$

Definition 12 *Given a CA with memory $m \leq 0$, anticipation $a \geq 0$ and local rule f, define a relation $\xrightarrow{f}$ on words of the same length $|u| = |v|$ as*

$$u \xrightarrow{f} v \iff \exists w \in A^{-m}, \exists z \in A^{a}, (wuz \in \mathcal{L}(X) \;\&\; f(wuz) = v),$$
$$u \xrightarrow{f^+} v \iff u = u^{(0)} \xrightarrow{f} \cdots \xrightarrow{f} u^{(k)} = v \text{ for some } k > 0$$

In particular we have $u \xrightarrow{\sigma} v$ iff there exists $a \in A$ such that $ua \in \mathcal{L}(X)$ and $\sigma(ua) = v$. Here σ denotes (by abuse of notation) the local rule of the shift map defined by $\sigma(u) = u_{[1,|u|)}$. The relation can be applied also to higher iterations of f. The relation $\xrightarrow{f^+}$ is the transitive closure of $\xrightarrow{f}$. If $u \xrightarrow{f^k} v$ for some $k > 0$, then $u \xrightarrow{f^+} v$. The converse implication, however, does not hold.

Proposition 13 *Let (X, F) be a CA, $n \geq \mathfrak{o}(X)$, and $U \subseteq \mathcal{L}^n(X)$.*
(1) U is invariant iff $\forall u, v \in \mathcal{L}^n(X), (u \in U \;\&\; u \xrightarrow{f} v \implies v \in U)$.
(2) If U is a minimal invariant set, then $\forall v \in U, \exists u \in U, u \xrightarrow{f} v$.
(3) If U is a spreading set, then $\forall u \in \mathcal{L}^n(X), \exists v \in U, \exists k > 0, u \xrightarrow{f^k} v$.
(4) If U is a minimal spreading set, then $X \cap \mathcal{S}_{A^n \setminus U}$ is a mixing SFT.
(5) If U is a minimal spreading set, then $(U, \xrightarrow{\sigma})$ is a connected aperiodic graph.

Proof. (1) We have $u \xrightarrow{f} v$ iff $F([u]) \cap [v] \neq \emptyset$.
(2) If the condition were not satisfied for some $v \in U$, then $F([U]) \cap [v] = \emptyset$, so $\Omega_F([U]) = \Omega_F([U \setminus \{v\}])$.
(3) Since X is mixing, there exists $j \in \mathbb{Z}$ and a configuration $x \in [u] \cap \sigma^j([U])$. Since U is spreading, there exists $v \in U$ and $k > 0$ such that $F^k(x) \in [v]$. This means that $u \xrightarrow{f^k} v$.
(4) If $X \cap \Sigma_{A^n \setminus U}$ is not mixing, then by Theorem 11, there exists $V \subseteq U$ such that $\Omega_F([V]) = \Omega_F([U])$ such that $X \cap \mathcal{S}_{A^n \setminus V}$ is mixing, and therefore $V \neq U$. Thus U is not minimal.
(5) is a consequence of (4). □

Corollary 14 *It is decidable whether a given clopen set U is invariant for a given CA (X, F).*

This is an immediate consequence of Proposition 13(1).

Definition 15 *The left and right neighbour sets of a word set $U \subseteq A^n$ are*

$$L_j(U) = \{u \in A^{n+2j+1} : \; u_{[j,j+n)} \notin U \;\&\; u_{[j+1,j+n]} \in U\}$$
$$R_j(U) = \{u \in A^{n+2j+1} : \; u_{[j,j+n)} \in U \;\&\; u_{[j+1,j+n]} \notin U\}.$$

Proposition 16 *Let (X, F) be a CA, $n \geq \mathfrak{o}(X)$, and $U \subseteq A^n$ invariant.*
(1) Assume that $(R_0(U), \xrightarrow{f^k})$ has no connected component for some $k > 0$, i.e., if $u \in R_0(U)$ and $u \xrightarrow{f^k} v$, then $v \notin R_0(U)$. Then U is spreading to the right.

(2) If there exists an F-periodic point $x \in [R_0(U)]$, then U is not spreading to the right.
(3) In particular, if u^∞ is F-periodic for some $u \in R_j(U)$, then U is not spreading to the right.

The proof is obvious.

6 Signal subshifts

Definition 17 *Let (X,F) be a cellular automaton. A configuration $x \in X$ is* **weakly periodic***, if $F^q\sigma^p(x) = x$ for some $q > 0$ and $p \in \mathbb{Z}$. We call (p,q) the* **period** *of x and p/q its speed. Let $\mathbf{S}_{(p,q)}(X,F) := \{x \in X : \; F^q\sigma^p(x) = x\}$ be the set of all weakly periodic configurations with period (p,q). A* **signal subshift** *is any non-empty $\mathbf{S}_{(p,q)}(X,F)$.*

For fixed (X,F) we write $\mathbf{S}_{(p,q)} := \mathbf{S}_{(p,q)}(X,F)$. Note that $\mathbf{S}_{(p,q)}$ is closed and σ-invariant, so it is a subshift provided it is nonempty. Moreover, $\mathbf{S}_{(p,q)}$ is F-invariant and $F : \mathbf{S}_{(p,q)} \to \mathbf{S}_{(p,q)}$ is bijective, so $\mathbf{S}_{(p,q)} \subseteq \Omega_F(X)$. If $\mathbf{S}_{(p,q)}$ is finite, it consists only of σ-periodic configurations. Figures 3 and 5 show some examples of signal subshifts.

Theorem 18 *Let (X,F) be a cellular automaton with memory m and anticipation a, and let $d = a - m$ be the diameter.*
(1) If $\mathbf{S}_{(p,q)}$ is nonempty, then it is a subshift of finite type.
(2) If $\mathbf{S}_{(p,q)}$ is infinite, then $\mathfrak{o}(\mathbf{S}_{(p,q)}) \leq \max\{\mathfrak{o}(X), dq+1\}$ and $-a \leq p/q \leq -m$.
(3) If $p_0/q_0 < p_1/q_1$, then $\mathbf{S}_{(p_0,q_0)} \cap \mathbf{S}_{(p_1,q_1)}$ is finite and $\mathfrak{p}(\mathbf{S}_{(p_0,q_0)} \cap \mathbf{S}_{(p_1,q_1)})$ divides $(\frac{p_1}{q_1} - \frac{p_0}{q_0})q$, where $q := \mathrm{lcm}(q_0,q_1)$ is the least common multiple.

Proof. If $x \in \mathbf{S}_{(p,q)}$, then $x_i = f^q(x_{[i+p+mq,i+p+aq]})$. We consider three cases.
(a) If $p+aq < 0$, then $p+mq < 0$, and we define a function $g : A^{-mq-p} \to A^{-mq-p}$ by

$$g(u)_j := \begin{cases} u_{j+1} & \text{if} \quad j < -mq-p-1 \\ f^q(u_{[0,dq]}) & \text{if} \quad j = -mq-p-1 \end{cases}$$

Then $x_{[i+p+mq+1,i]} = g(x_{[i+p+mq,i-1]})$ for every $i \in \mathbb{Z}$, and this is possible only if x is σ-periodic. Moreover, the period of x is bounded by $|A|^{-mq-p}$, so $\mathbf{S}_{(p,q)}$ is finite and therefore is of finite type.
(b) If $0 < p + mq$, we show similarly as in (a) that $\mathbf{S}_{(p,q)}$ is finite and therefore of finite type.
(c) If $-a \leq p/q \leq -m$, set $u = x_{[i+p+mq,i+p+aq]} \in A^{dq+1}$, so that the condition $x_i = f^q(x_{[i+p+mq,i+p+aq]})$ reads $u_{-p-mq} = f^q(u)$ (note that $0 \leq -p-mq < |u|$). For $D := \{u \in \mathcal{L}^{dq+1}(X) : \; f^q(u) \neq u_{-mq-p}\}$, we get $\mathbf{S}_{(p,q)} = \mathcal{S}_D \cap X$, so $\mathbf{S}_{(p,q)}$ is a SFT of order at most $\max\{\mathfrak{o}(X), dq+1\}$. Cases (a), (b), and (c) together prove (1) and (2).
(3) Set $p_2 = qp_0/q_0$, $p_3 = qp_1/q_1$. If $x \in \mathbf{S}_{(p_0,q_0)} \cap \mathbf{S}_{(p_1,q_1)}$, then

$$\sigma^{-p_2}(x) = \sigma^{-p_2}(F^{q_0}\sigma^{p_0})^{\frac{q}{q_0}}(x) = F^q(x) = \sigma^{-p_3}(F^{q_1}\sigma^{p_1})^{\frac{q}{q_1}}(x) = \sigma^{-p_3}(x),$$

so $\sigma^{p_3-p_2}(x) = x$. □

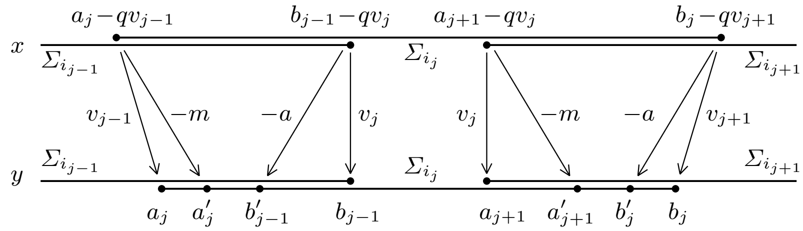

Fig. 2. Preimage of a configuration in the join.

Theorem 19 *Let (X, F) be a cellular automaton with diameter d and assume that $\mathbf{S}_{(p_1,q_1)}, \dots, \mathbf{S}_{(p_n,q_n)}$ are signal subshifts with decreasing speeds, i.e., $p_i/q_i > p_j/q_j$ for $i < j$. Denote by $q := \operatorname{lcm}\{q_1, \dots q_n\}$ the least common multiple. Let $c > 0$ be an integer which satisfies the following inequalities:*

$$\begin{aligned}
&\mathfrak{o}(X) \le c \\
&\forall i \le n, \mathfrak{o}(\mathbf{S}_{(p_i,q_i)}) \le c \\
&\forall i < j \le n, (\mathbf{S}_{(p_i,q_i)} \cap \mathbf{S}_{(p_j,q_j)} \neq \emptyset \implies \mathfrak{p}(\mathbf{S}_{(p_i,q_i)} \cap \mathbf{S}_{(p_j,q_j)}) \le c) \\
&\forall i < j \le n, (\mathbf{S}_{(p_i,q_i)} \cap \mathbf{S}_{(p_j,q_j)} \neq \emptyset \implies (d - \tfrac{p_i}{q_i} + \tfrac{p_j}{q_j})q \le c)
\end{aligned}$$

Then for $\Sigma := \mathbf{S}_{(p_1,q_1)} \overset{c}{\vee} \cdots \overset{c}{\vee} \mathbf{S}_{(p_n,q_n)}$ the following assertions hold:

(1) $\Sigma \subseteq F^q(\Sigma)$ and therefore $\Sigma \subseteq \Omega_F(X)$.

(2) If G is the graph of Σ constructed by Proposition 3 from graphs of $\mathbf{S}_{(p_i,q_i)}$, and if H is a connected component of G, then Σ_H is F^q-invariant.

(3) If G_k is the graph of $F^{kq}(\Sigma)$ constructed to Proposition 7 from G, then for each connected component H_k of G_k there exists a connected component H of G such that $\Sigma_{H_k} = \Sigma_H$.

Proof. (1) Let $y \in \Sigma$, $1 \le i_1 < i_2 \cdots < i_k \le n$ and let $I_1 = (a_1, b_1), \dots, I_k = [a_k, b_k)$ be intervals such that the restrictions $y_{|I_j}$ belong to $\tilde{\mathcal{L}}(\mathbf{S}_{(p_{i_j},q_{i_j})})$ and $b_j - a_{j+1} \ge c$. Let $v_j := p_{i_j}/q_{i_j}$ be the speed of the i_j-th signal. The configurations in $\mathbf{S}_{(p_{i_j},q_{i_j})} \cap \mathbf{S}_{(p_{i_{j+1}},q_{i_{j+1}})}$ are σ-periodic with period $n_j := \mathfrak{p}(\mathbf{S}_{(p_{i_j},q_{i_j})} \cap \mathbf{S}_{(p_{i_{j+1}},q_{i_{j+1}})})$. Let u be the prefix of $y_{|I_j}$ of length n_{j-1} and let v be the suffix of $y_{|I_j}$ of length n_j. Since $c \ge n_{j-1}$ and $c \ge n_j$, we get $u^\infty, v^\infty \in \mathbf{S}_{(p_{i_j},q_{i_j})}$. Since $\mathfrak{o}(\Sigma_{(p_{i_j},q_{i_j})}) \le c$, we get $y^{(j)} = u^\infty(y_{|I_j})v^\infty \in \mathbf{S}_{(p_{i_j},q_{i_j})}$ and $F^q\sigma^{qv_j}(y^{(j)}) = y^{(j)}$. Set $J_j = [a_j - qv_{j-1}, b_j - qv_{j+1})$. For the endpoints of these intervals we get $(b_j - qv_{j+1}) - (a_{j+1} - qv_j) = b_j - a_{j+1} + q(v_j - v_{j+1}) \ge c$. Set $x^{(j)} := \sigma^{qv_j}(y^{(j)})$, so $F^q(x^{(j)}) = y^{(j)}$. For $a_{j+1} - qv_j \le i \le b_j - qv_{j+1}$ we have

$$(x^{(j)})_i = (y^{(j)})_{i+qv_j} = (y^{(j+1)})_{i+qv_j} = (y^{(j+1)})_{i+qv_{j+1}} = (x^{(j+1)})_i.$$

Since $c \ge \mathfrak{o}(X)$, there exists a unique configuration $x \in X$ such that $x_i = (x^{(j)})_i$ for $i \in J_j$. We show $F^q(x) = y$. Set $a'_j := a_j - q(m + v_{j-1})$, $b'_j := b_j - q(a + v_{j+1})$

(see Figure 2). For $a'_j \leq k \leq b'_j$ we have

$$\begin{aligned} F^q(x)_k &= f^q(x_{[k+qm,k+qa]}) = f^q(\sigma^{qv_j}(y^{(j)})_{[k+qm,k+qa]}) \\ &= F^q \sigma^{qv_j}(y^{(j)})_k = (y^{(j)})_k = y_k. \end{aligned}$$

We have $b'_j - a'_{j+1} = b_j - a_{j+1} - q(v_{j+1} - v_j + d) \geq c - c = 0$, so the intervals $[a'_j, b'_j)$ cover whole $\mathbb{Z}$, and $F^q(x) = y$. Thus $\Sigma \subseteq F^q(\Sigma)$. It follows that for any $k > 0$ we have $\Sigma \subseteq F^{kq}(\Sigma) \subseteq F^{kq}(X)$, so $\Sigma \subseteq \Omega_F(X)$.

(2) By Proposition 3, the connected components are not changed by the join construction. If $x \in \Sigma_H$, then $F^q(x) = \sigma^p(x)$ for some $p \in \mathbb{Z}$. Since $\sigma(\Sigma_H) = \Sigma_H$, we get $F^q(x) \in \Sigma_H$.

(3) If H is a connected F^q-invariant component of G, then all paths in G of length kq form a connected component H_k of G_k and they determine the same subshift. Moreover, G_k has no other connected components. □

Corollary 20 *Let (X, F) be a CA. If $\Omega_F(X) = F^k(\mathbf{S}_{(p_1,q_1)} \overset{c}{\vee} \cdots \overset{c}{\vee} \mathbf{S}_{(p_n,q_n)})$ for some $(p_i, q_i) \in \mathbb{Z} \times \mathbb{N}^+$ and $k, c \geq 0$, then (X, F) has only a finite number of infinite transitive signal subshifts.*

Proof. Let $\mathbf{S}_{(p,q)}$ be an infinite transitive signal subshift. Since $\mathbf{S}_{(p,q)}$ is transitive, $\mathbf{S}_{(p,q)} \subseteq \mathbf{S}_{(p_i,q_i)}$ for some i by Theorem 19. Since $\mathbf{S}_{(p,q)}$ is infinite, we get $p/q = p_i/q_i$. If $q > q_i$, then $\mathbf{S}_{(p,q)} = \mathbf{S}(p - p_i, q - q_i)$, so we can assume $q \leq q_i$. Thus there is only a finite number of possibilities for (p, q). □

Example 3 shows that Theorem 20 does not generalize to infinite non-transitive signal subshifts. A CA with infinite number of infinite transitive signal subshifts with different speeds has been constructed in Kůrka [1].

7 Decreasing preimages

Definition 21 *Let $f : \mathcal{L}(X) \to \mathcal{L}(X)$ be a local rule of a cellular automaton (X, F). We say that a subshift $\Sigma \subseteq X$ has r-**decreasing preimages**, if for each $u \in \mathcal{L}(X) \cap \mathcal{D}(\Sigma)$, each $v \in f^{-r}(u)$ contains as a subword a word $w \in \mathcal{D}(\Sigma)$ such that $|w| < |u|$. We say that Σ has* **decreasing preimages**, *if it has r-decreasing preimages for some $r > 0$.*

The condition of Definition 21 is satisfied trivially if $f^{-r}(u) = \emptyset$. Thus for example, each $F^r(X)$ has r-decreasing preimages.

Theorem 22 *If (X, F) is a CA and if a subshift $\Sigma \subseteq X$ has decreasing preimages, then $\Omega_F(X) \subseteq \Sigma$. If moreover $\Sigma \subseteq F^k(\Sigma)$ for some $k > 0$, then $\Omega_F(X) = \Sigma$.*

Proof. If $u \in \mathcal{L}(\Omega_F(X)) \setminus \mathcal{L}(\Sigma) \subseteq \mathcal{L}(X) \setminus \mathcal{L}(\Sigma)$, then each $v \in f^{-r|u|}(u)$ contains as a subword $w \in \mathcal{L}(X) \setminus \mathcal{L}(\Sigma)$ with $|w| = 0$. This is a contradiction, since $w = \lambda \in \mathcal{L}(\Sigma)$. Thus $f^{-r|u|}(u)$ is empty and $u \notin \mathcal{L}(\Omega_F)$. Thus $\mathcal{L}(\Omega_F(X)) \subseteq \mathcal{L}(\Sigma)$ and $\Omega_F(X) \subseteq \Sigma$. If $\Sigma \subseteq F^k(\Sigma)$, then $\Sigma \subseteq F^k(X) \subseteq F^{2k}(X) \cdots$ and therefore $\Sigma \subseteq \Omega_F(X)$. □

Proposition 23 *There exists an algorithm, which decides whether for a given cellular automaton (X,F) and given $r>0$, a given sofic subshift $\Sigma \subseteq X$ has r-decreasing preimages.*

Proof. By the assumption, $\mathcal{L}(X)$, $\mathcal{L}(\Sigma)$, and $f^{-r}(\mathcal{D}(\Sigma))$ are regular languages. It follows that the language

$$\begin{aligned} L_r &:= \{v \in \mathcal{L}(X) : \ f^r(v) \in \mathcal{D}(\Sigma), \ \forall k \leq rd+1, v_{[k,k+|v|-rd-1)} \in \mathcal{L}(\Sigma)\} \\ &= \mathcal{L}(X) \cap f^{-r}(\mathcal{D}(\Sigma)) \cap \bigcap_{k \leq rd+1} \mathcal{L}(\Sigma)_{[k,rd+1-k)} \end{aligned}$$

is regular too (see Lemma 4 for notation). Since L_r is empty iff Σ has r-decreasing f-preimages, we get the deciding procedure. □

Corollary 24
(1) It is decidable whether a given sofic subshift with r-decreasing preimages is the maximal attractor of a given CA.
(2) The set of cellular automata whose maximal attractor is a sofic subshift with decreasing preimages is recursively enumerable.

Proof. (1) This follows from Theorem 22 and Proposition 7.
(2) Generate successively all sofic subshifts, verify whether they are strongly invariant and whether they have decreasing preimages. □

8 Undecidability

The decidability results of Corollary 24 are no longer true when the condition of decreasing preimages is omitted. Moreover, we show that spreading sets cannot be found algorithmically. Recall that a CA $(A^{\mathbb{Z}},F)$ is **nilpotent**, if $\Omega_F(A^{\mathbb{Z}})$ is a singleton. In this case there exists $q \in A$ and $n>0$ such that $\Omega_F(A^{\mathbb{Z}}) = F^n(A^{\mathbb{Z}}) = \{q^\infty\}$. Kari [5] proves that it is undecidable whether a given CA is nilpotent or not. Actually, Kari's proof works even for a smaller class of CA, and this undecidability result can be used to prove undecidability of the spreading property.

Theorem 25 (Kari [5]) *It is undecidable whether for a given CA $(A^{\mathbb{Z}},F)$ with an invariant set $[q]_0$ (where $q \in A$) we have $\Omega_F(A^{\mathbb{Z}}) = \{q^\infty\}$.*

Theorem 26 *It is undecidable, whether for a given CA (X,F) a given clopen invariant set $U \subseteq X$ is spreading or not.*

Proof. Assume by contradiction that such a decision procedure exists. For any CA $(A^{\mathbb{Z}},F)$ with an invariant set $[q]_0$ we construct a CA $(B^{\mathbb{Z}},G)$ and a clopen invariant set $U \subseteq B^{\mathbb{Z}}$ such that U is spreading iff $(A^{\mathbb{Z}},F)$ is nilpotent, and thus obtain a contradiction with Theorem 25. Take $B := A \times \{0,1\}$ and define a function $g : A \times \{0,1\} \to \{0,1\}$ and a CA $G : B^{\mathbb{Z}} \to B^{\mathbb{Z}}$ by

$$g(x,a) := \begin{cases} 1 & \text{if} \quad x \neq q \text{ and } a = 1 \\ 0 & \text{otherwise} \end{cases}, \qquad G(x,y)_i := (F(x)_i, g(x_i,y_i))$$

Clearly the clopen set $U := A^{\mathbb{Z}} \times [0]_0$ is G-invariant. We show that U is spreading iff $\Omega_F(A^{\mathbb{Z}}) = \{q^\infty\}$. Indeed if $(A^{\mathbb{Z}}, F)$ is nilpotent, then $F^n(A^{\mathbb{Z}}) = \{q^\infty\}$ for some $n > 0$. Thus $G^{n+1}(B^{\mathbb{Z}}) = \{(q,0)^\infty\}$, so $(B^{\mathbb{Z}}, G)$ is nilpotent and U is spreading. Conversely, if $(A^{\mathbb{Z}}, F)$ is not nilpotent, then (by compactness) there exists $x \in A^{\mathbb{Z}}$ such that $F^n(x)_1 \neq q$ for all $n \geq 0$. Take $y \in [01]_0$. Then $(x,y) \in U$ but $G^n(x,y) \in A^{\mathbb{Z}} \times [1]_1$ for all $n \geq 0$. This means that $G^n(U) \not\subseteq \sigma(U)$ for all n, so U is not spreading. □

Theorem 27

(1) It is undecidable whether the omega-limit of a given spreading set of a given CA is sofic or not.

(2) It is undecidable whether a given sofic subshift is the maximal attractor of a given CA.

Proof. (1) This is equivalent to the problem, whether the maximal attractor of a CA on a mixing SFT is sofic or not (using the construction of $\widetilde{\Omega}_\sigma(U)$). This is in turn a special case of the problem whether the maximal attractor of CA on $A^{\mathbb{Z}}$ is sofic. This is undecidable by a theorem of Kari [11]), because being a sofic subshift is a non-trivial property.
(2) If there were such a decision procedure, then we could decide whether $\{a^\infty\}$ were the maximal attractor of F. Trying this for all $a \in A$, we would get a deciding procedure for nilpotency. This is impossible by a theorem of Kari [5]. □

9 Search algorithms

We describe now an algorithm which uses our theory to search for subshift attractors of a given cellular automaton. Procedure `Spread` first looks for spreading sets verifying sufficient conditions of Proposition 13. If a spreading set is found, Theorem 11 is used to obtain a mixing SFT, whose maximal attractor is a subshift attractor of the original cellular automaton. Then we use procedure `Omega` to attain the maximal attractor from bellow. Since each signal subshift is included in the maximal attractor, the algorithm successively generates all signal subshifts. Using a certain number of signal subshifts, the algorithm computes their join Σ and its images $F^k(\Sigma)$ which by Theorem 19 are all included in the maximal attractor. Then the condition of decreasing preimages is tested which by Theorem 22 implies that the maximal attractor has been attained.

The input data for the algorithm `Spread` consists of a finite set $D \subset A^*$ of forbidden words of X, the memory m, anticipation a, and a local rule f of F.

`procedure Spread(D,f,m,a,n)`

1. Construct the graphs $(\mathcal{L}^n(\mathcal{S}_D), \xrightarrow{f})$, $(\mathcal{L}^n(\mathcal{S}_D), \xrightarrow{\sigma})$
2. Find all $U \subseteq \mathcal{L}^n(\mathcal{S}_D)$ which satisfy conditions (1), (2), (3), (5) of Proposition 13.
3. For each U from step 2 and for $k := 1, 2, \ldots$ perform steps 4 and 5.

4. If $(L_0(U), \xrightarrow{f^k})$ and $(R_0(U), \xrightarrow{f^k})$ have no connected component, then output U as next spreading set and terminate the cycle of k.
5. If for some $u \in L_k(U) \cup R_k(U)$, u^∞ is F-periodic, then terminate the cycle of k, since U is not spreading.

Step 1 of the procedure `Spread(D,f,m,a,n)` is based on the search for the connected components of a graph. The time complexity of this task is linear in the number of vertices and edges (see Tarjan [12]). Procedure `Spread(D,f,m,a,n)` may fail to terminate if it performs indefinitely steps 4 and 5. If it does terminate for some spreading set U, then the subshift attractor $\Omega_F(U)$ is the maximal attractor of $(X \cap \mathcal{S}_{A^n \setminus U}, F)$. To find the attractor, we use procedure `Omega(D,f,m,a,n)` which searches for the maximal attractor of a CA (X, F) in the form of a forward image of join of signal subshifts whose periods (p, q) satisfy $q \leq n$. The input data for `Omega` are again a set $D \subset A^*$ of forbidden words of X, the memory m, anticipation a, and a local rule f for F.

`procedure Omega(D,f,m,a,n)`

1. Construct all infinite signal subshifts with periods (p_i, q_i) such that $q_i \leq n$, and $-aq_i \leq p_i \leq -mq_i$. Denote by q the least common multiple of q_i.
2. Order the signal subshifts obtained in step 1 by decreasing speeds and construct their c-join Σ, for some c bounded by a value given in Theorem 19.
3. Construct $F^q(\Sigma), F^{2q}(\Sigma), \ldots$ and test whether $F^{kq}(\Sigma) = F^{(k+1)q}(\Sigma)$.
4. If step 3 ends with $F^{kq}(\Sigma) = F^{(k+1)q}(\Sigma)$, verify whether $F^{kq}(\Sigma)$ has r-decreasing preimages for $r := 1, 2, \ldots$. If so, $\Omega_F(\mathcal{S}_D) = F^{kq}(\Sigma)$ has been found.

While the value of c given in Theorem 19 seems to be necessary for the proof of the theorem, procedure `Omega` sometimes works even with smaller values of c for which the constructed data structures are smaller. The procedure `Omega(D,f,m,a,n)` may fail to terminate if it repeats indefinitely step 3, or if step 4 gives negative result for all $r > 0$.

10 Implementation

The algorithm has been implementated in Java and the binary distribution can be downloaded from http://code.google.com/p/cellattractors/. In the implementation, sofic subshifts are represented as NFA because of their simplicity and conciseness. The sizes of these NFA are repeatedly reduced. The algorithms implement standard operations with regular languages like intersection and complementation. To perform tests for equality or complementation, NFA are transformed to DFA. All these NFA or DFA are kept reduced (minimized). A representation of a DFA with n states can be minimized in time $\mathcal{O}(n \log n)$ (see e.g., Knuth [13]). Unfortunately, this is not the case of NFA, for which the state minimization problem is PSPACE-complete (see Jiang and Ravimkur [14]), and the algorithms cannot be used in practice. There exist, however, algorithms which reduce the sizes of NFA without necessarily attaining the minimum. There are at least two feasible NFA reduction techniques. Both techniques are based on merging of states according to some relations $\sim_r$, $\sim_l$. In the first technique $\sim_r$ and $\sim_l$ are constructed as equivalence relations. The equivalence $\sim_r$ is defined as the largest equivalence relation on Q which is right-invariant with respect to $\mathcal{A}$ (see Ilie et al., [15],[16],[17]). The second technique constructs $\sim_r$ and $\sim_l$ as preorder relations smaller than the equivalence relations. (see Champarnaud and Coulon [18], Ilie et al., [17], [16]). The algorithm which combines preorders $\sim_r$ and $\sim_l$ according to simple heuristic (see Champarnaud and Coulon [19]) but not optimally, gives satisfying results and is much easier to implement than the algorithm based on equivalences.

A special operation needed in Proposition 23 is the subword language operation of Lemma 4. The language L_r is the intersection of $rd+3$ regular languages. The time and memory required to construct and represent an automaton representing this language may be prohibitively large. It is therefore worthwhile to test emptiness of intersection of smaller numbers of these languages first. For example, both $f^{-1}(\mathcal{D}(\Omega_F)) \cap \mathcal{L}(\Omega_F)_{[2,1)}$ and $f^{-1}(\mathcal{D}(\Omega_F)) \cap \mathcal{L}(\Omega_F)_{[1,2)}$ are empty for ECA128 (see Example 1). If a subshift does not have r-decreasing preimages, the computation is much more time-consuming.

While the constructions of signal subshifts and their join is rather fast with at most quadratic time complexity, it takes $\mathcal{O}(e^{d+1})$ time to construct a presentation $F(\Sigma)$ of a sofic subshift Σ, where e is number of edges of Σ (see constructive proof of Proposition 7). There are at least q such succesive constructions. We get $\mathcal{O}(e^{(d+1)^q})$ time in the number of edges of given sofic subshift. The construction of DFA from presentation of F-image takes $\mathcal{O}(n)$ time but may require $\mathcal{O}(2^n)$ space. Minimization and normalization of such DFA takes $\mathcal{O}(n \log n + n) = \mathcal{O}(n \log n)$ time, where n is the number of states of DFA. The most time-consuming task is the verification of the decreasing preimages condition according to Proposition 23. Let Σ be sofic subshift with v states and e edges, $L_r = \{v \in A^* : \ f^r(v) \in \mathcal{L}(\Sigma)\}$, and $L'_k = \{v \in A^* : \ v_{[k,k+|v|-rd-1)} \in \mathcal{L}(\Sigma)\}$ be languages from Proposition 23. A recognizing automaton for L_r can be constructed in $\mathcal{O}(ev|\mathcal{L}^{dr}(X)|)$ time and space. A recognizing automaton for L'_k can be constructed in $\mathcal{O}(ev^{rd+2-k})$ time and space. Intersection of two NFA can be done in $\mathcal{O}(v^2)$ time and test on empty language requires $\mathcal{O}(v)$ time.

11 Examples

Example 1 (ECA128) *The product CA* $F(x)_i = x_{i-1}x_ix_{i+1}$ *on* $X = \{0,1\}^{\mathbb{Z}}$.

We have two infinite nontransitive signal subshifts (see Figure 3) $\mathbf{S}_{(1,1)} = \mathcal{S}_{10}$, $\mathbf{S}_{(-1,1)} = \mathcal{S}_{01}$. whose orders are $\mathfrak{o}(\mathbf{S}_{(1,1)}) = \mathfrak{o}(\mathbf{S}_{(-1,1)}) = 2$. Their intersection is $\{0^\infty, 1^\infty\}$, so $\mathfrak{p}(\mathbf{S}_{(1,1)}) \cap \mathbf{S}_{(-1,1)}) = 1$. Finally $(d - \frac{p_1}{q_1} + \frac{p_2}{q_2})q = 0$, so $c = 2$. The maximal attractor is constructed in Figure 3. In the first row, labelled graphs for 2-block encodings of $\mathbf{S}_{(1,1)}$ and $\mathbf{S}_{(-1,1)}$ are given. Their join $\Sigma := \mathbf{S}_{(1,1)} \overset{2}{\vee}$

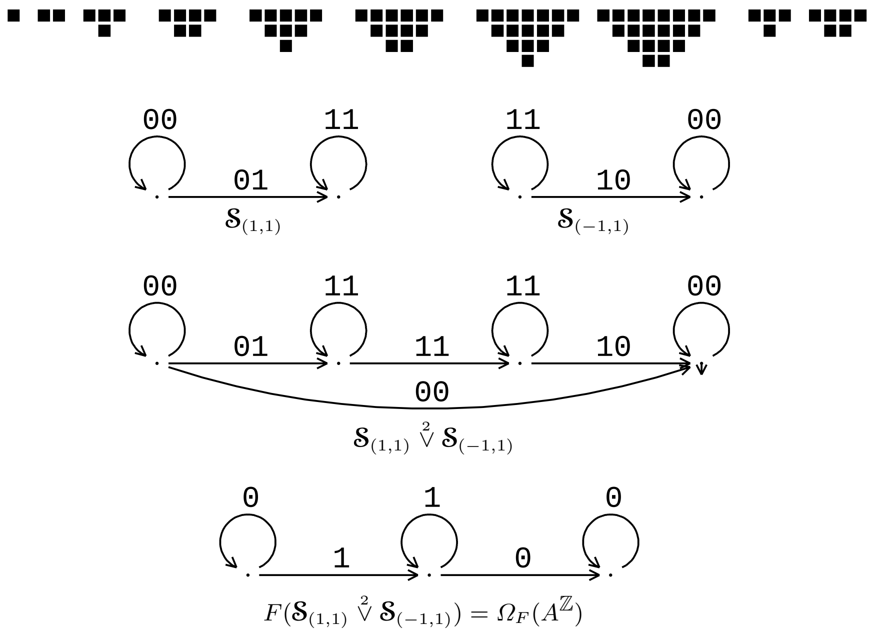

Fig. 3. The product ECA128 and its signal subshifts.

$\mathbb{S}_{(-1,1)} = \mathcal{S}_{\{10^n1:\, n>0\}\cup\{010\}}$ is constructed in the second row. In the third row there is the minimal presentation of (2-block encoding of) $F(\Sigma) = \mathcal{S}_{\{10^n1:\, n>0\}}$. We have $F^2(\Sigma) = F(\Sigma)$ and $F(\Sigma)$ has 1-decreasing preimages: If $u \in \mathcal{D}(F(\Sigma))$, then $u = 10^n1$ for some $n > 0$. If $n \leq 2$, then u has no preimage. If $n > 2$, then the only preimage $1110^{n-2}111$ of u contains a shorter forbidden word $10^{n-2}1$. Thus $f^{-1}(\mathcal{D}(F(\Sigma)) \cap \mathcal{L}(F(\Sigma))_{[2,1)} = \emptyset$ as well as $f^{-1}(\mathcal{D}(F(\Sigma)) \cap \mathcal{L}(F(\Sigma))_{[1,2)} = \emptyset$. Thus $F(\Sigma)$ has 1-decreasing preimages and therefore $F(\Sigma) = \Omega_F(A^{\mathbb{Z}})$.

Example 2 (ECA232) *The majority CA* $F(x)_i = \lfloor \frac{x_{i-1}+x_i+x_{i+1}}{2} \rfloor$.

The CA has a spreading set $U := A^3 \setminus \{010, 101\}$, whose subshift attractor is the signal subshift $\Omega_F(U) = \mathbb{S}_{(0,1)} = \mathcal{S}_{\{010,101\}}$. Graphs $(U, \xrightarrow{f})$ (left) and $(U, \xrightarrow{\sigma})$ (right) are in Figure 4. The left and right neighbours of U are $R_0(U) = \{0010, 1101\}$ and $L_0(U) = \{1011, 0100\}$. There are no arrows in either $R_0(U)$ or $L_0(U)$, so U is a spreading set. There are two other infinite signal subshifts $\mathbb{S}_{(1,1)} = \mathcal{S}_{\{011,100\}}$, $\mathbb{S}_{(-1,1)} = \mathcal{S}_{\{001,110\}}$. The join Σ of all three signal subshifts is constructed in Figure 5. We have $\mathfrak{o}(\mathbb{S}_{(1,1)}) = \mathfrak{o}(\mathbb{S}_{(0,1)}) = \mathfrak{o}(\mathbb{S}_{(-1,1)}) = 3$, $\mathbb{S}_{(1,1)} \cap \mathbb{S}_{(-1,1)} = \{0^\infty, 1^\infty, (01)^\infty, (10)^\infty\}$, $\mathbb{S}_{(1,1)} \cap \mathbb{S}_{(0,1)} = \mathbb{S}_{(0,1)} \cap \mathbb{S}_{(-1,1)} = \{0^\infty, 1^\infty\}$, and $(d - \frac{p_i}{q_i} + \frac{p_j}{q_j})q \leq 2$, so $c = 3$. The arrows from $\mathbb{S}_{(1,1)}$ to $\mathbb{S}_{(0,1)}$ and from $\mathbb{S}_{(0,1)}$ to $\mathbb{S}_{(-1,1)}$ can be omitted without changing the subshift. Thus we get $\Sigma := \mathbb{S}_{(1,1)} \overset{3}{\vee} \mathbb{S}_{(0,1)} \overset{3}{\vee} \mathbb{S}_{(-1,1)} = \mathbb{S}_{(0,1)} \cup (\mathbb{S}_{(1,1)} \overset{3}{\vee} \mathbb{S}_{(-1,1)})$ (see Figure 6). First offenders are $\mathcal{D}(\Sigma) = \{010^n1, 10^n10, 01^n01, 101^n0 :\; n > 1\}$. We have $F(\Sigma) = \Sigma$ and Σ has 1-decreasing preimages. To show it, note that 010, 101 have unique

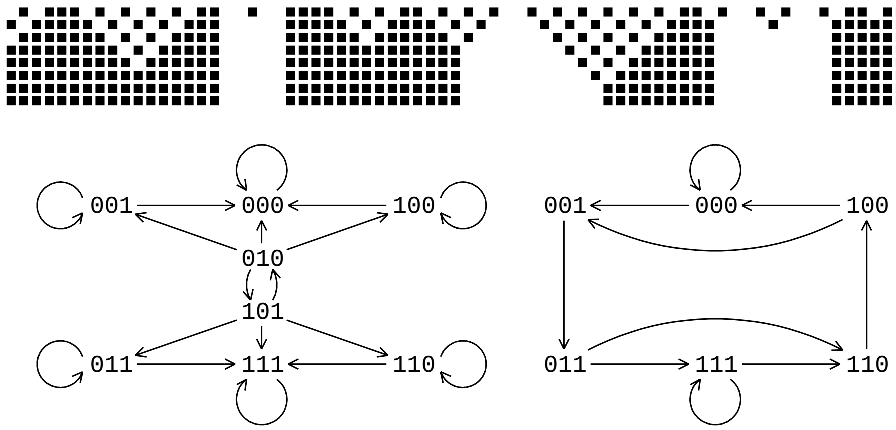

Fig. 4. Graphs $(A^3, \xrightarrow{f})$ (left) and $(U, \xrightarrow{\sigma})$ (right) of the spreading set $U = A^3 \setminus \{010, 101\}$ of the majority ECA232.

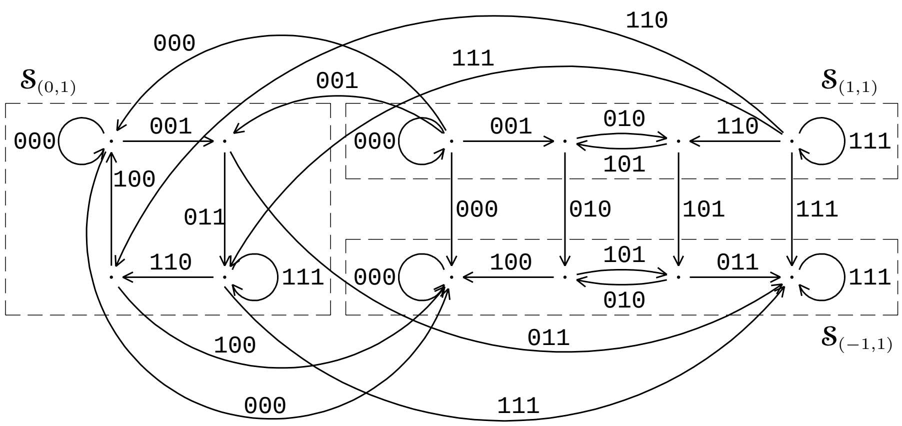

Fig. 5. The join $\mathcal{S}_{(0,1)} \overset{3}{\vee} \mathcal{S}_{(1,1)} \overset{3}{\vee} \mathcal{S}_{(-1,1)}$ of the majority ECA232

preimages $f^{-1}(010) = \{01010\}$, $f^{-1}(101) = \{10101\}$. For the first offenders we get $f^{-1}(01001) = \emptyset$, $f^{-1}(010001) = \{01010011\}$. If $n \geq 4$, then each preimage of 010^n1 has one of the form $01010u0011$ or $01010u0101$, where $|u| = n - 4$ and in each case it contains a shorter forbidden word. The same result we get for other forbidden words. Thus $f^{-1}(\mathcal{D}(\Sigma)) \cap \mathcal{L}(\Sigma)_{[2,1)} \cap \mathcal{L}(\Sigma)_{[1,2)} = \emptyset$, Σ has 1-decreasing preimages, and $\Omega_F(A^{\mathbb{Z}}) = \Sigma$.

Example 3 *There exists a CA with infinite number of infinite (nontransitive) signal subshifts whose maximal attractor is a join of signal subshifts.*

Proof. define a CA with alphabet $A := \{0, 1, 2\}$, memory $m = 0$, anticipation $a = 1$, and local rule $f(ab) = 0$ if $a = 0$ and $f(ab) = b$ otherwise. There

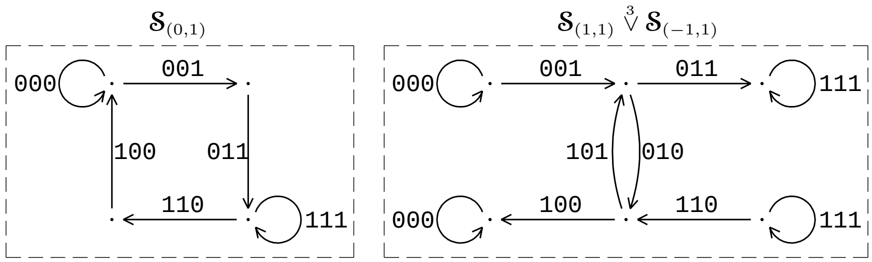

Fig. 6. Maximal attractor $\Omega_F = \boldsymbol{\mathcal{S}}_{(0,1)} \cup (\boldsymbol{\mathcal{S}}_{(1,1)} \overset{3}{\vee} \boldsymbol{\mathcal{S}}_{(-1,1)})$ of the majority ECA232

1121222120021122211212100021111221200
1212221200011222112121000011112212000
2122212000012222112121000001122120000
1222120000022221121210000001221200000

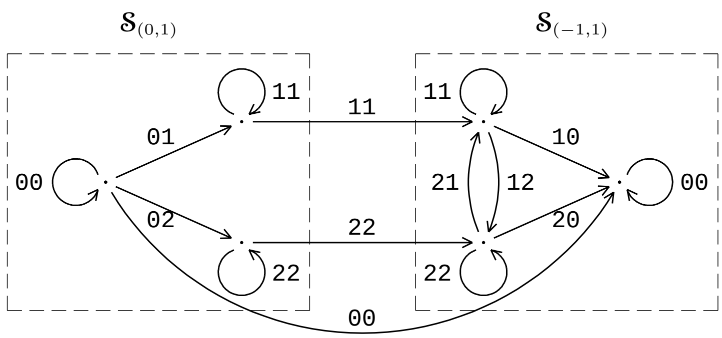

Fig. 7. Infinite number of non-transitive signal subshifts

exist infinite signal subshifts $\boldsymbol{\mathcal{S}}_{(0,1)} = \mathcal{S}_{\{10,12,20,21\}}$, $\boldsymbol{\mathcal{S}}_{(-1,1)} = \mathcal{S}_{\{01,02\}}$ and $\Omega_F = \boldsymbol{\mathcal{S}}_{(0,1)} \overset{2}{\vee} \boldsymbol{\mathcal{S}}_{(-1,1)}$. However there exist an infinite number of different infinite signal subshifts $\boldsymbol{\mathcal{S}}_{(0,q)} = \{0^\infty\} \overset{0}{\vee} \{x \in \{1,2\}^{\mathbb{Z}} :\ \sigma^q(x) = x\}$.

Example 4 *There exists a CA with sofic maximal attractor which has not decreasing preimages.*

Proof. The existence of such a CA follows from Corollary 24. We give an explicit construction. The alphabet is $A = \{0, 1, 2, 3, 4\}$, $r = 1$ and the local rule is

x11:4, x22:4, x33:4, x01:4, x02:4, x03:4, x31:4, x32:4,
x33:4, x14:4, 42x:4, 430:1, 43x:4, 204:4, 400:1, 40x:4,
41x:1, x24:3, x13:3, 13x:0, x23:3, 23x:0, 10x:2, 20x:1.

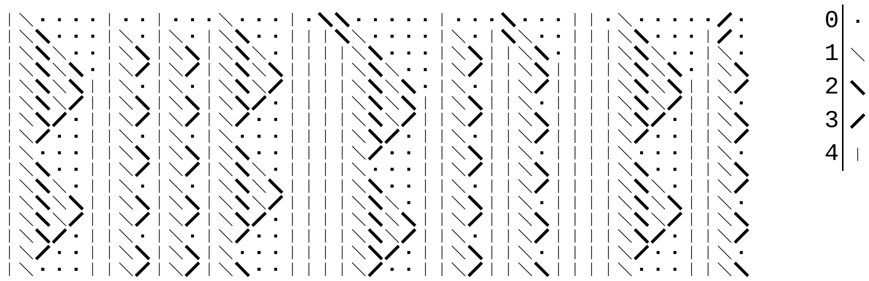

Fig. 8. Sofic subshift attractor without decreasing preimages

Here the first applicable production is used. If there is none, the letter is left unchanged. A simulation can be seen in Figure 8. We have $F^3(A^{\mathbb{Z}}) \subseteq \Sigma_D$, where $D := \{01, 02, 03, 11, 14, 22, 31, 32, 33, 40, 42, 43\}$. The graph of Σ_D is in Figure 9 right. The maximal attractor is a sofic subshift whose graph (Figure 9 left) has two connected components. The right component yields configurations of the form $(12)^\infty 0^{2n+1}4\ldots$ whose some left infinite interval does not contain 4. The word $u = 4(12)^n 04$ does not belong to $L := \mathcal{L}(\Omega_F(X))$, while $(12)^n 04$ does belong to L. The word u has a preimage $v \in f^{-(2n-1)}(u)$ such that all subwords of v of length $|u|-1$ belong to L. In fact v contains as a subword $w = 410^{2n}4 \notin L$. we have $|w| = |u|$ and all proper subwords of w belong to L.

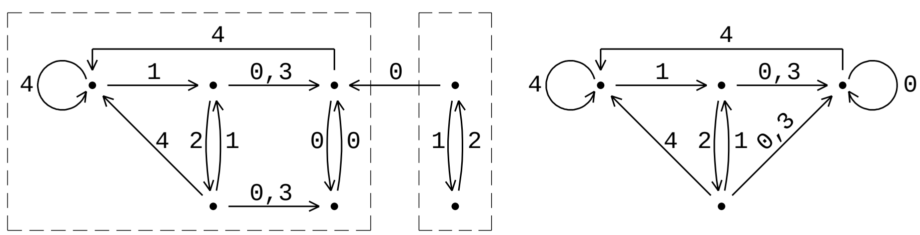

Fig. 9. Maximal attractor and its SFT approximation

12 Conclusion

Our search algorithm works for the simplest CA with small number of infinite signal subshifts. Since many problems connected with the subshift attractors are algorithmically undecidable, this seems to be a maximum what algorithmic procedures can achieve. When the algorithm fails to find any subshift attractor, then this is an indication that the dynamics of the cellular automaton may be complex and interesting. But even in this case, signal subhifts obtained may help to ellucidate the dynamics of a given cellular automaton.

Acknowledgments

The research was partially supported by the Research Program CTS MSM 0021620845.

References

1. Kůrka, P.: On the measure attractor of a cellular automaton. Discrete and Continuous Dynamical Systems **Supplement volume 2005** (2005) 524–535
2. Formenti, E., Kůrka, P.: Subshift attractors of cellular automata. Nonlinearity **20** (2007) 105–117
3. Culik, K., Pachl, J., Yu, S.: On the limit set of cellular automata. SIAM Journal on Computing **18**(4) (1989) 831–842
4. Maass, A.: On the sofic limit set of cellular automata. Ergodic Theory and Dynamical Systems **15** (1995) 663–684
5. Kari, J.: The nilpotency problem of one-dimensional cellular automata. SIAM J. Comput. **21**(3) (1992) 571–586
6. Formenti, E., Kůrka, P.: A search algorithm for the maximal attractor of a cellular automaton. In Thomas, W., Weil, P., eds.: STACS 2007. Volume 4393., Berlin, Springer-Verlag (2007) 356–366
7. Boyle, M., Kitchens, B.: Periodic points for onto cellular automata. Indagationes Math **10**(4) (1999) 483–493
8. Hopcroft, J.E., Ullmann, J.D.: Introduction to Automata Theory, Languages and Computation. Addison-Wesley (1979)
9. Lind, D., Marcus, B.: An introduction to symbolic dynamics and coding. Cambridge University Press, Cambridge (1995)
10. Kitchens, B.P.: Symbolic Dynamics. Springer-Verlag, Berlin (1998)
11. Kari, J.: The Rice's theorem for the limit sets of cellular automata. Theoretical Computer Science **127** (1994) 229–254
12. Tarjan, R.: Depth-first search and linear search algorithms. SIAM Journal for Computing **1** (1972) 146–160
13. Knuutila, T.: Re-describing an algorithm by Hopcroft. Theoretical Computer Science **250** (2001) 333–363
14. Jiang, T., Ravikumar, B.: Minimal NFA problems are hard. SIAM J. Comput. **22**(6) (1993) 1117–1141
15. Ilie, L., Yu, S.: Reducing NFAs by invariant equivalences. Theoretical Computer Science **306** (2003) 373–390
16. Ilie, L., Solis-Oba, R., Yu, S.: Reducing NFAs by equivalences and preorders. In Apostolico, A., Crochemore, M., Park, K., eds.: Proc. of 16th CPM (Jeju island, Korea, 2005). Volume 3537 of Lecture Notes in Computer Science., Berlin, Springer (2005) 310–321
17. Ilie, L., Navarro, G., Yu, S.: On NFA reductions. In Karhumaki, J., ed.: Theory is Forever (Salomaa Festschrift). Volume 3113 of Lecture Notes in Computer Science., Berlin, Springer-Verlag (2004) 112–124
18. Champarnaud, J.M., Coulon, F.: NFA reduction algorithms by means of regular inequalities. Theoretical Computer Science **327**(3) (2004) 241–253
19. Champarnaud, J.M., Coulon, F.: NFA reduction algorithms by means of regular inequalities - correction. Erratum in Theoretical Computer Science (2005)