Contents

Chapter 1

Complex Systems, an introduction

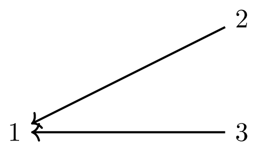
A complex system is a multitude of simple components that, by cooperating, lead to a complex behavior. They can be used as a discrete model for collective intelligence. Thanks to the ease of their implementation (e.g. Netlogo), these models can be used to perform simulations. In these notes, we will present many different models, such as:

- Cellular automata
- Subshifts
- Tiling
- Reaction Systems
- etc.

Complex systems can have multiple properties. We will discuss how the properties of the phenomenon being modeled are reflected in the properties of the model.

Real phenomenon's property

- Reachability

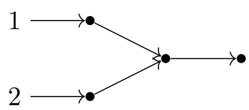


Model's property

- Surjectivity
- Transitivity
- ...

Real phenomenon's property

- Non-collapsing orbits



Model's property

- Injectivity

Real phenomenon's property

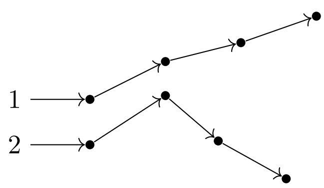
- Reversibility

Model's property

- Injectivity + surjectivity

Real phenomenon's property

- Instability

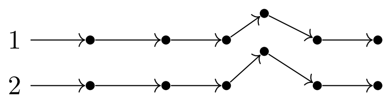


Model's property

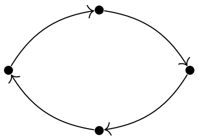
- Sensible dependency on initial conditions
- Caos

Real phenomenon's property

- Stability



- Periodic behavior



Model's property

- Dense set of stability points
- Equicontinuity
- ...

- Eventual periodicity
- Dense set of periodic points

Chapter 2

Cellular automata

In this chapter, we present the first model to be discussed, starting from the objects on which the model operates.

Definition 1. Let A be a finite alphabet, the space of configurations is $A^{\mathbb{Z}} = \{x|x : \mathbb{Z} \to A\}$ i.e., the set of all bi-infinite sequences over A.

By definition, $x \in A^{\mathbb{Z}} \iff x : \mathbb{Z} \to A$ which is equivalent to $i \in \mathbb{Z} \to x_i \in A$.
$x \in A^{\mathbb{Z}}$ can be written as $x = (\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots), \forall i \in Z, x_i \in A$

Example 1. Let $A = \{0, 1\}$ a possible configuration of $\{0, 1\}^{\mathbb{Z}}$ is $x = (\dots, \overset{-4}{0}, \overset{-3}{1}, \overset{-2}{1}, \overset{-1}{0}, \overset{0}{1}, \overset{1}{0}, \overset{2}{1}, \overset{3}{0}, \overset{4}{0}, \dots\}$

Notation 1. $\forall x \in A^{\mathbb{Z}}, \forall h, k \in \mathbb{Z}$ with $h \leq k$, $x_{[h,k]} = x_h x_{h-1} \dots x_k \in A^{k-h+1}$

Example 2. Let $x \in \{0, 1\}^{\mathbb{Z}}$ as of example 1 and $h = -1, k = 1$, then $x_{[-1,1]} = 010 \in A^3$. Suppose now that $h = k = 3$, then $x_{[3,3]} = 0 \in A^1$

2.1 Notion of distance

A core concept required for future discussions is distance over a generic set.

Definition 2. A distance over a generic set X is a function $d : X \times X \to \mathbb{R}_+$ that satisfies the following conditions:

- $\forall x, y \in X, d(x, y) = 0 \iff x = y$
- $\forall x, y \in X, d(x, y) = d(y, x)$
- $\forall x, y, z \in X, d(x, y) \leq d(x, z) + d(z, y)$ (triangle inequality)

We can define many different kinds of distances, such as:

- The trivial distance, which works for every set X, is the function $d_t : X \times X \to \mathbb{R}_+$, where $\forall x, y \in X, d_t(x, y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \neq y \end{cases}$
- The Euclidean distance $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$, where $\forall x, y \in \mathbb{R}, d(x, y) = |x - y|$, note that this can be extended to $\mathbb{R}^n$ by using the norm-2 ($\| \cdot - \cdot \|_2$)

2.1.1 Tychonoff's distance

Here, we discuss a distance notion required to define properties of bi-infinite sequences over an alphabet. The Tychonoff distance is a function $d : A^{\mathbb{Z}} \times A^{\mathbb{Z}} \to \mathbb{R}_+$, such that

$$\forall x, y \in A^{\mathbb{Z}}, d(x, y) = \begin{cases} 0 & \text{if } x = y \\ \frac{1}{2^n} & \text{if } x \neq y \end{cases}$$

where $n = \min\{k \in \mathbb{N} | x_{[-k,k]} \neq y_{[-k,k]}\}$

Example 3. Let $A = \{0,1\}$. Let $x, y \in \{0,1\}^{\mathbb{Z}}$ be

$$x = (\dots, 0, 1, 1, 0, \overset{0}{1}, 0, 1, 0, 0, \dots), \qquad y = (\dots, 0, 1, 1, 0, \overset{0}{1}, 0, 1, 1, 0, \dots).$$

In this case we obtain $n = 3$.

$$\begin{aligned}
&k = 0 \quad x_{[0,0]} = 1, && y_{[0,0]} = 1 \Rightarrow n \geq 0,\\
&k = 1 \quad x_{[-1,1]} = 010, && y_{[-1,1]} = 010 \Rightarrow n \geq 1,\\
&k = 2 \quad x_{[-2,2]} = 10101, && y_{[-2,2]} = 10101 \Rightarrow n \geq 2,\\
&k = 3 \quad x_{[-3,3]} = 1101010, && y_{[-3,3]} = 1101011 \Rightarrow n = 3.
\end{aligned}$$

Therefore

$$d(x,y) = \frac{1}{2^3}.$$

This distance has an important property that will be later used.

Proposition 1. $\forall x, y \in A^{\mathbb{Z}}, \forall n \in \mathbb{N} \quad d(x,y) < \frac{1}{2^n} \iff x_{[-n,n]} = y_{[-n,n]}$

Remark. $(A^{\mathbb{Z}}, d)$ is a compact space, i.e., every sequence of elements of $A^{\mathbb{Z}}$ has a convergent subsequence (in $A^{\mathbb{Z}}$).

2.2 1d cellular automata over a finite alphabet

Definition 3. A one-dimensional cellular automata is a triple $< A, r, f >$, where:

- A is a finite alphabet.
- $r \in \mathbb{N}$ is the radius.
- $f : A^{2r+1} \to A$ is the local rule.

From here on, we will write CA instead of cellular automata. To each CA $< A, r, f >$ is associated a global rule $F : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$, defined as such:

$$\forall x \in A^{\mathbb{Z}}, \forall i \in \mathbb{Z} \quad F(x)_i = f(x_{[i-r,i+r]})$$

$$x = (\dots, \underbrace{x_{i-r}, \dots, x_i, \dots, x_{i+r}}, \dots)$$

$$F(x) = (\dots, \dots, \dots, F(x)_i, \dots, \dots, \dots)$$

Definition 4. A one-dimensional CA with binary alphabet $A = \{0,1\}$ and $r = 1$ is called an elementary CA. Naturally, the local rules is defined ad $f : \{0,1\}^3 \to \{0,1\}$ The total number of possible local rules is $256 = 2^{2^3}$, therefore, there are 256 elementary CA.

Proposition 2. *More in general, for a one-dimentional CA defined over an alphabet which cardinality is $|A|$ and with radius r there are $|A|^{|A|^{2r+1}}$ possible local rules.*

Example 4. Let $< \{0,1\}, 1, f >$ be an elementary CA, where f is defined as the following table:

	$\{0,1\}^3$			f
0	0	0	0	0
1	0	0	1	0
2	0	1	0	0
3	0	1	1	1
4	1	0	0	1
5	1	0	1	1
6	1	1	0	0
7	1	1	1	1

An equivalent way to represent the local rule is using $n_f \in \{0, \dots, 255\}$, where n_f is the decimal number obtained by converting the binary number in the f column (where the bit in row 0 is the least significant bit).

2.2.1 Representation of a cellular automata

Given a CA $< A, r, f >$ there are several way to represent it:

- A table defining f
- A natural number $n_f \in \{0, \ldots, |A|^{|A|^{2r+1}} - 1\}$
- An oriented graph (De Bruijn graph)

Definition 5. The De Bruijn graph associated to the CA $< A, r, f >$ is a graph $< V, E, l >$, where:

- $V = A^{2r}$
- $E = \{(u, v) \in V^2 | u = u_1 \ldots u_{2r}, v = v_1 \ldots v_{2r}, u_1 \ldots u_{2r-1} = v_2 \ldots v_{2r}\}$
- $l : E \to \{0, 1\}$ is a labeling function of the arches where $\forall (u, v) \in E \quad l(u, v) = f(uv_{2r}) = f(u_1 v)$

Example 5. Let $< \{0, 1\}, 1, f >$ be the CA of example 4, i.e. $n_f = 184$. The associated De Bruijn graph is $< V, E, l >$, with $V = \{0, 1\}^2$

$$(00, 00) \in E? \quad \begin{cases} u = 00 \\ v = 00 \end{cases} \quad \text{Yes } \ell(00, 00) = f(000) = 0$$

$$(00, 01) \in E? \quad \begin{cases} u = 00 \\ v = 01 \end{cases} \quad \text{Yes } \ell(00, 01) = f(001) = 0$$

$$(00, 10) \in E? \quad \begin{cases} u = 00 \\ v = 10 \end{cases} \quad \text{No}$$

$$(00, 11) \in E? \quad \begin{cases} u = 00 \\ v = 11 \end{cases} \quad \text{No}$$

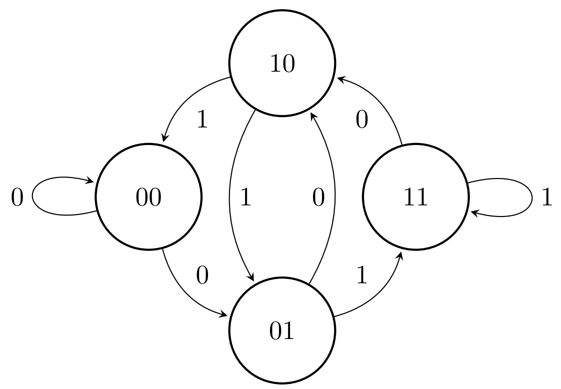
$$(01, 00) \in E? \quad \begin{cases} u = 01 \\ v = 00 \end{cases} \quad \text{No}$$

$$(01, 01) \in E? \quad \begin{cases} u = 01 \\ v = 01 \end{cases} \quad \text{No}$$

$$(01, 10) \in E? \quad \begin{cases} u = 01 \\ v = 10 \end{cases} \quad \text{Yes } \ell(01, 10) = f(010) = 0$$

$$(01, 11) \in E? \quad \begin{cases} u = 01 \\ v = 11 \end{cases} \quad \text{Yes } \ell(01, 11) = f(011) = 1$$

The graph then is:



2.3 Hedlund's theorem

In this section, we will discuss a fundamental theorem regarding CAs. First, we need to define a shift map over bi-infinite words.

Definition 6. A shift map is a function $\sigma : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ where $\forall x \in A^{\mathbb{Z}}, \forall i \in \mathbb{Z}, \sigma(x)_i = x_{i+1}$

Remark. A shift map is the global rule of CA $n_f = 170$.

Theorem 1 (Hedlund)**.** *Let $F : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ be an arbitrary function. F is a global rule of a cellular automata if and only if it holds that:*

- *F is continuous*
- *F is commutative with respect to σ, i.e. $F \circ \sigma = \sigma \circ F$*

Proof. ($\implies$)
Assume F is the global rule of a CA $< A, r, f >$. We now show that F is contiguous, that is

$$\forall x \in A^{\mathbb{Z}}, \forall \varepsilon > 0, \exists \delta > 0 | \forall y \in A^{\mathbb{Z}} \quad d(y, x) < \delta \to d(F(y), F(x)) < \varepsilon$$

To prove this, we choose an arbitrary $x \in A^{\mathbb{Z}}$ and $\varepsilon > 0$. Let $n \in \mathbb{N}$ such that $\frac{1}{2^n} < \varepsilon$. We now show that there exists a $\delta > 0$ such that

$$\forall y \in A^{\mathbb{Z}} \quad d(y, x) < \delta \to d(F(y), F(x)) < \frac{1}{2^n}$$

This always holds true with $\delta = \frac{1}{2^{n+r}}$, by proposition 1 we now that if $d(y, x) < \frac{1}{2^{n+r}}$ then $x_{[-n-r,n+r]} = y_{[-n-r,n+r]}$, naturally then $F(y)_{[-n,n]} = F(x)_{[-n,n]}$ because of how rules are applied.
In order to prove $F \circ \sigma = \sigma \circ F$ we need to show that $\forall x \in A^{\mathbb{Z}}, \sigma(F(x)) = F(\sigma(x))$, in other words

$$\forall x \in A^{\mathbb{Z}}, \forall i \in \mathbb{N} \quad (F(\sigma(x)))_i = (\sigma(F(x)))_i$$

To do so, we expand the left-hand side and right-hand side terms

- $(\sigma(F(x)))_i = (\sigma(f(x_{[i-r,i+r]}))) = f(x_{[i+1-r,i+1+r]})$
- $(F(\sigma(x)))_i = F((\sigma(x))_i) = F(x_{i+1}) = f(x_{[i+1-r,i+1+r]})$

They are the same, which means that applying the global rule first, followed by the shift operator, or applying the shift operator, followed by the global rule, leads to the same result.
($\impliedby$)
Now assume that F is continuous and commutative with respect to σ. Since F is continuous and $A^{\mathbb{Z}}$ is a compact set, then F is uniformly continuous, that is

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x, y \in A^{\mathbb{Z}} \quad d(y, x) < \delta \implies d(F(y), F(x)) < \varepsilon$$

We now choose $\varepsilon = 1$. We are sure there exists a $\delta > 0$ such that

$$\forall x, y \in A^{\mathbb{Z}} \quad d(y, x) < \delta \implies d(F(y), F(x)) < 1$$

We also know that

$$d(F(y), F(x)) < 1 \implies F(y)_0 = F(x)_0$$

Let $r \in \mathbb{N}$ be the smallest natural number such that $\frac{1}{2^r} < \delta$. Then,

$$\forall x, y \in A^{\mathbb{Z}} \quad d(y, x) < \frac{1}{2^r} \implies F(y)_0 = F(x)_0$$

This also means

$$\forall x, y \in A^{\mathbb{Z}} \quad y_{[-r,r]} = x_{[-r,r]} \implies F(y)_0 = F(x)_0$$

We now show that F is a global rule of a CA. To do this, we need to construct a valid CA. We construct the CA $< A, r, f >$, where A is given since we know the function $F : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ and r is the integer described previously.
We now need to define a valid f function such that $\forall x \in A^{\mathbb{Z}}, \forall i \in \mathbb{Z} \quad F(x)_i = f(x_{[i-r,i+r]})$ (we refer to

this condition as $\bigoplus$).
Let $f : A^{2r+1} \to A$ such that $\forall u \in A^{2r+1}, f(u) = F(z)_0$, where z is <u>any</u> configuration of $A^{\mathbb{Z}}$ such that $z_{[-r,r]} = u$.
This definition of the function is valid because $\forall z', z'' \in A^{\mathbb{Z}} \quad z'_{[-r,r]} = z''_{[-r,r]} \implies F(z')_0 = F(z'')_0$. We now prove $\bigoplus$ for $i = 0$, $\forall x \in A^{\mathbb{Z}} \quad F(x)_0 = f(x_{[-r,r]})$, this is obviously true because of the construction of f, picking $z = x$. We now prove it for any $i \neq 0$, let $x \in A^{\mathbb{Z}}$

$$F(x)_i = (\sigma^i(F(x)))_0 = (F(\sigma^i(x)))_0 = f(\sigma^i(x)_{[-r,r]}) = f(x_{[i-r,i+r]})$$

□

2.3.1 Interpretations of the proven theorem

From Hedlund's theorem, we can derive some interesting considerations. First, it gives us some insight into the nature of computation in cellular automata. Cellular automata are a local and uniform computational model. Hedlund's theorem tells us that $F : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ is a global rule of a CA if and only if

1. F is continuous $\to$ Locality

2. $\sigma \circ F = F \circ \sigma \to$ Uniformity

Definition 7. Any element $x \in A^{\mathbb{Z}}$ is said to be a periodic spatial configuration if and only if $\exists K > 0$ such that $\sigma^K(x) = x$.
Equivalently $x \in A^{\mathbb{Z}}$ is said to be a periodic spatial configuration if and only if $\exists u \in A^K$ with $K > 0$ such that

$$x = (\ldots, u, u, u, \ldots) = {}^{\infty}u^{\infty}$$

x can thus be written as the bi-infinite repetition of u.

Corollary 1. *Let $x \in A^{\mathbb{Z}}$ be a periodic spatial configuration and $F : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ some global rule, then $F(x)$ is also a periodic spatial configuration.*

Proof. If $x \in A^{\mathbb{Z}}$ is spatially periodic $\implies \exists k > 0$ such that $\sigma^k(x) = x$. This means that $\exists k > 0$ such that $F(\sigma^k(x)) = F(x)$, but since F is a global rule of a CA this is equivalent to $\exists k > 0$ such that $\sigma^k(F(x)) = F(x)$. It must be noted that if x is a repetition of a finite word of k symbols, then the same holds for $F(x)$, then:

$$\begin{aligned}
x &= {}^{\infty}u^{(0)\infty} \\
&\downarrow F \\
F(x) &= {}^{\infty}u^{(1)\infty} \\
&\downarrow F \\
F^1(x) &= {}^{\infty}u^{(2)\infty} \\
&\downarrow F \\
F^3(x) &= {}^{\infty}u^{(3)\infty} \\
&\vdots \\
&\downarrow F \\
F^t(x) &= {}^{\infty}u^{(t)\infty} = {}^{\infty}u^{(h)\infty} = F^h(x) \text{ with } h < t
\end{aligned}$$

The last step indicates that a cycle exists in the states obtained by applying the global rule to spatially periodic configurations. □

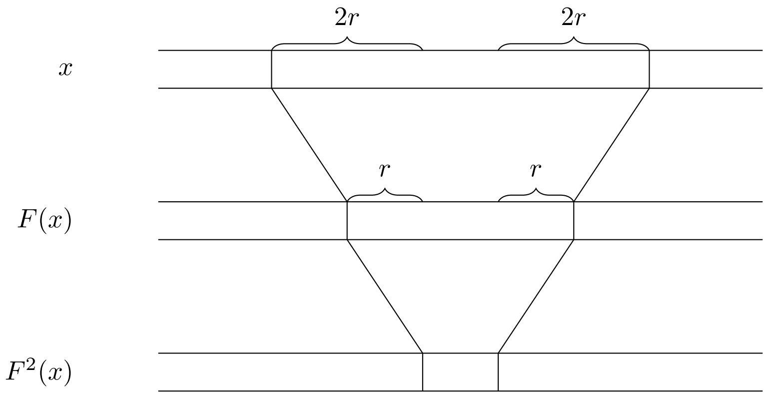


Figure 2.1: Multiple applications of F

Given a CA $< A, r, f >$, it is possible to define a CA $< A, r, f' >$ such that $F' = F^2$ where F is the global rule of the first CA, and F' is that of the last one. We define $f' : A^{4r+1} \to A$, where $\forall u = u_1 \ldots u_{4r+1} \in A^{4r+1} \quad f'(u) = f(f(u_1 \ldots u_{2r+1})f(u_2 \ldots u_{2r+2}) \ldots f(u_{2r+1} \ldots u_{4r+1}))$.
It's easy to prove that $\forall x \in A^{\mathbb{Z}}, \forall i \in \mathbb{Z}, F'(x)_i = f'(x_{[i-2r,i+2r]}$ by using the definition of f' and the fact that $F' = F^2$.

Corollary 2. *If F is a global rule of a CA, then F^2 also is.*

Proof. To show that F^2 is also the global rule of a CA, we can use Hedlund's theorem.

1. F^2 is continuous since F is, and any composition of continuous functions is also continuous.

2. $F^2 \circ \sigma = F \circ (F \circ \sigma) = F \circ (\sigma \circ F) = (F \circ \sigma) \circ F = \sigma \circ F \circ F = \sigma \circ F^2$, then F^2 is commutative with respect to σ.

Then, by Hedlund's theorem, F^2 is the global rule of a CA. □

Corollary 3. *Let F and G be global rules of two different CAs, with the same alphabet A. Then, $F \circ G$ is the global rule of a CA.*

Proof. We can proceed in the same way as the previous proof:

1. $F \circ G$ is obviously continuous as previously stated.

2. The proof of commutativity with respect to σ is the same as the previous corollary, with two generic functions instead of $F \circ F = F^2$.

□

Corollary 4. *Let F be the global rule of a CA. Suppose F is invertible, $\exists F^{-1} : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$. Then F^{-1} is the global rule of a CA.*

Proof. As usual, we use Hedlund's theorem:

1. Since $A^{\mathbb{Z}}$ is a compact set, then F^{-1} is continuous.

2. We show F^{-1} is commutative with respect to σ. We know $F \circ \sigma = \sigma \circ F$, then:
$F^{-1} \circ F \circ \sigma = F^{-1} \circ \sigma F$
$F^{-1} \circ F \circ \sigma \circ F^{-1} = F^{-1} \circ \sigma \circ F \circ F^{-1}$
$Id \circ \sigma \circ F^{-1} = F^{-1} \circ \sigma \circ Id$
$\sigma \circ F^{-1} = F^{-1} \circ \sigma$

□

2.4 Extended rule to finite words

We now extend the definition of local rules to finite words over the alphabet A.

Definition 8. Let $f^* : A^* \to A^*$, where

$$\forall u \in A^* \quad f^*(u) = \begin{cases} \varepsilon & \text{if } |u| < 2r+1 \\ f(u_1 \dots u_{2r+1}) \dots f(u_{n-2r} \dots u_n) & \text{otherwise} \end{cases}$$

From here on, we will only write f instead of f^* when talking about finite words.

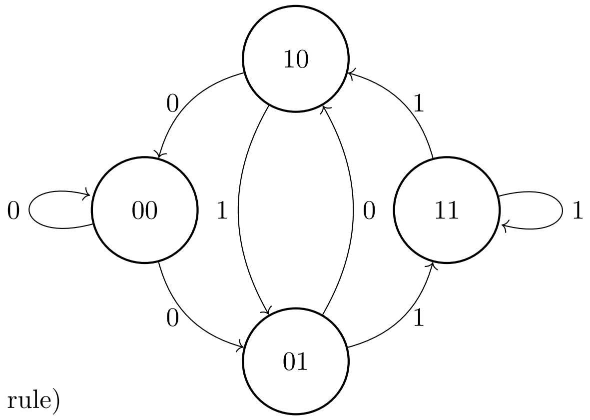
2.4.1 Additional notes on De Bruijn graphs

Let $< A, r, f >$ be a CA, let $\mathcal{G} = (V, E, l)$ be the graph associated with the CA.

Proposition 3. *The bi-infinite paths on the vertices of $\mathcal{G}$ are in a one-to-one correspondence with the elements of $A^{\mathbb{Z}}$. Moreover, $\forall x \in A^{\mathbb{Z}}, F(x)$ is given from the labels of the path on $\mathcal{G}$ corresponding to x. The same holds true $\forall u \in A^* : |u| \geq 2r+1$. The proof is natural given the definition of $\mathcal{G}$*

$\{0,1\}^3$	f
000	0
001	0
010	0
011	1
100	0
101	1
110	1
111	1

$n_f = 232$ (majority rule)



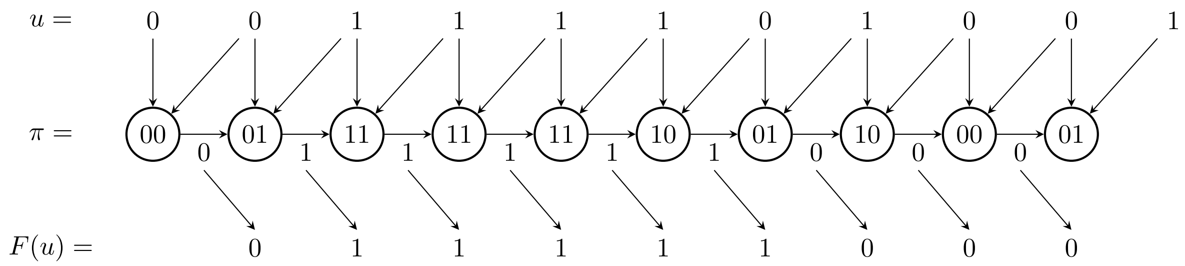


Figure 2.2: Obtaining $F(u)$ from u using $\mathcal{G}$

Remark. A de Bruijn graph of a CA can be considered as a transition graph of a non-deterministic finite state automaton, where every state is both initial and final. The language $\mathcal{L}$ recognized by this automaton is made of all the words obtained by labels of finite paths that start from any state and finish in any other state.

2.5 Surjectivity

Surjectivity can be seen as a weak form of reachability in a system. There are three different ways we can define surjectivity

- $\forall y \in A^{\mathbb{Z}} \exists x \in A^{\mathbb{Z}} : F(x) = y$
- $F(A^{\mathbb{Z}}) = A^{\mathbb{Z}}$

- $\forall y \in A^{\mathbb{Z}}, F^{-1}(y) \neq \emptyset$

Theorem 2. *Let $< A, r, f >$ be a CA, let $F : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ be its global rule. Let $\mathcal{G}$ be its associated De Bruijn graph. Let $\mathcal{L}$ be the language recognized by the FSA whose transition graph is $\mathcal{G}$, and every state is a starting and final state. Then, the following propositions are equivalent:*

1. *F is surjective*
2. *f is balanced, that is $\forall u \in A^+, |f^{-1}(u)| = |A|^{2r}$*
3. *$\mathcal{L} = A^*$*

Proof.
1) $\Rightarrow$ 2)

Here we show that if f is not balanced, then F is not surjective (this is logically equivalent to the original proposition). Suppose that $\exists u \in A^n$ such that $|f^{-1}(u)| \neq |A|^{2r}$. Without loosing generality we can assume that $|f^{-1}(u)| < |A|^{2r}$. We call $s = |f^{-1}(u)|$. Let $k > 1$ be an arbitrary integer. We now build two sets of words, V and W, where a generic word $w \in W$ contains u k times. Between two occurrences of u in w there are $2r$ arbitrary symbols of A. Then, $|W| = |A|^{(k-1)\cdot 2r}$ A generic word $v \in V$ is the juxtaposition of any k pre-images of u, therefore $|V| = s^k$. By construction, we are certain that $\forall v \in V, f(v) \in W$. We now want to find a $w \in W$ such that $f^{-1}(w) = \emptyset$. It's certain that it exists if $|V| < |W|$, in other words $s^k < (|A|^{2r})^{k-1}$, then $\exists k : |V| < |W|$. Then $\exists w \in W : f^{-1}(w) \neq \emptyset$, but then $\exists y \in A^{\mathbb{Z}}$ such that $F^{-1}(y) \neq \emptyset$ (it's sufficient that y contains u), then F is not surjective.

2) $\Rightarrow$ 3)

We now need to show that $\mathcal{L} \subseteq A^* \wedge A^* \subseteq \mathcal{L}$, therefore $A^* = \mathcal{L}$. Choose an arbitrary $v \in A^*$, then we show that $v \in \mathcal{L}$, which means that v is obtained by labels of a path on the vertices of $\mathcal{G}$. Since $|f^{-1}(v)| = |A|^{2r}$, as shown previously, in particular $\exists u \in A^*$ such that $f(u) = v$. By definition of $\mathcal{G}$, since u corresponds to a finite path on the vertices of $\mathcal{G}$, $f(u) = v$ is obtained by the labels of that path.

3) $\Rightarrow$ 1)

Choose an arbitrary $y \in A^{\mathbb{Z}}$, we need to show that $\exists x \in A^{\mathbb{Z}}$ such that $F(x) = y$. $\forall n \in \mathbb{N}$ let $v^{(n)} = y_{[-n,n]} \in A^{2n+1}$. By hypothesis we know that $\exists u^{(n)} \in f^{-1}(v^{(n)}) \subseteq A^{2n+2r+1}$, then $f(u^{(n)}) = v^{(n)}$. Let $x^{(n)} \in A^{\mathbb{Z}}$ be any configuration such that $x^{(n)}_{[-n-r,n+r]} = u^{(n)}$. It's not certain that $F(x^{(n)}) = y$, but it is sure that $\forall n \in \mathbb{N}, F(x^{(n)})_{[-n,n]} = y_{[-n,n]} \implies d(F(x^{(n)}), y) < \frac{1}{2^n}$, which means $\lim_{n\to+\infty} F(x^{(n)}) = y$. Moreover, $\{x^{(n)}\}_{n\in\mathbb{N}}$ has a convergent subsquence $\{x^{(n_h)}\}_{h\in\mathbb{N}}$. Let $x = \lim_{h\to+\infty} x^{(n_h)}$, we have that $F(x) = F(\lim_{h\to+\infty} x^{(n_h)}) = \lim_{h\to+\infty} F(x^{(n_h)}) = y$, then $\exists x \in A^{\mathbb{Z}}$ such that $y = F(x)$. □

Corollary 5. *Surjectivity is a decidable property of 1-d CAs.*

Proof. Property 3) of the previous theorem is easily decidable, it is done by creating a deterministic FSM that recognizes the same language $\mathcal{L}$, minimizing it and checking if $\mathcal{L} = A^*$. □

Remark. In order to check if $\mathcal{L} = A^*$ we need to build a deterministic FSA that recognizes $\mathcal{L}$. This FSM has $\mathcal{B}(V)$ as its set of vertices. Then the complexity of the algorithm is $\geq |\mathcal{B}(V)| = 2^{|A|^{2r}}$.

2.5.1 Finite configurations

Definition 9. Let $q \in A$, then $x \in A^{\mathbb{Z}}$ is a q-finite configuration if $|\{i \in \mathbb{Z} | x_i \neq q\}| < \infty$, in other words x is of the form $x = (\dots qqqqqquqqqqqq \dots)$ for some $u \in A^*$

Let $\mathcal{F}$ be the set of all q-finite configurations. We call $\mathcal{SP}$ the set of spatial periodic configurations, where $\mathcal{SP} = \{x \in A^{\mathbb{Z}} | \exists k \in \mathbb{N}, k > 0, \sigma^k(x) = x\}$

Theorem 3. *Let $< A, r, f >$ be a CA. The following propositions are equivalent:*

1. *F is surjective.*
2. *$F|_{\mathcal{SP}}$ is surjective (F restricted to $\mathcal{SP}$ is surjective).*
3. *$\forall x, y \in \mathcal{F}, x \neq y \implies F(x) \neq F(y)$, F is injective on finite configurations.*

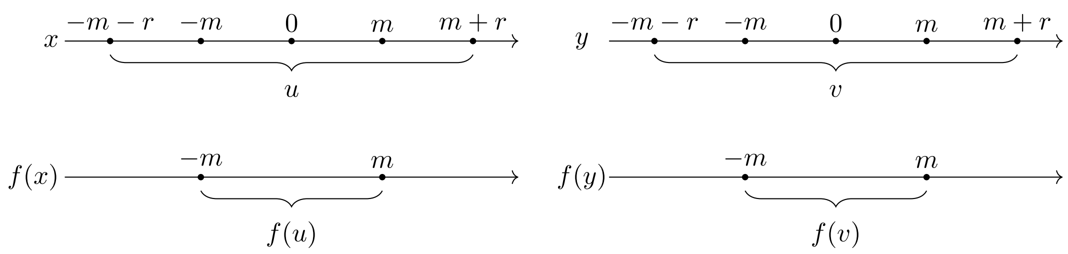
2.6 Injectivity

Theorem 4. *Let $< A, r, f >$ be a CA. The following propositions are equivalent:*

1. *F is injective.*

2. *$\exists m \leq |A|^{4r}$ such that $\forall x, y \in A^{\mathbb{Z}}$ $\quad F(x)_{[-m,m]} = F(y)_{[-m,m]} \implies x_0 = y_0$*

3. *$\exists m \leq |A|^{4r}$ such that $\forall u, v \in A^{2m+2r+1}$ $\quad F(u) = F(v) \implies u_{m+r+1} = v_{m+r+1}$*



Corollary 6. *Injectivity is a decidable property of 1d-CAs.*

Proof. Property 3) is easily decidable since it's a finite combinatorial problem, it is sufficient to check the condition for every possible couple in $A^{2m+2r+1}$ and every $m \leq A^{4r}$. □

Algorithm 1 Decision algorithm for injectivity in 1d-CAs

---

for $m = 0$ to $|A|^{4r}$ **do**
  implic ← true
  for all $u, v \in A^{2m+2r+1} \wedge$ implic = true **do**
    if $(f(u) = f(v) \wedge (u_{m+n+1} \neq v_{m+n+1})$ **then**
      implic ← false
    end if
  end for
  if implic = true **then**
    return true
  end if
end for
return false

Remark. This algorithm has time complexity

$$\gtrsim \frac{1}{2} \sum_{m=0}^{|A|^{4r}} |A|^{2\cdot(2m+2r+1)} = \frac{1}{2}|A|^{4r+2} \cdot \frac{1 - (|A|^4)^{|A|^{4r+1}}}{1 - |A|^4}$$

Theorem 5. *F is injective $\implies$ $F|_{\mathcal{SP}}$ is injective.*

The following properties are equivalent for 1d-CAs:

- Injectivity
- Bijectivity
- Reversibility

Theorem 6. *F is not surjective $\iff$ $\exists$ a diamond, that is $\exists w \in A^{2r}, \exists u, v \in A^{+}$ with $|u| = |v|, u \neq v$, such that $f(wuw) = f(wvw)$.*

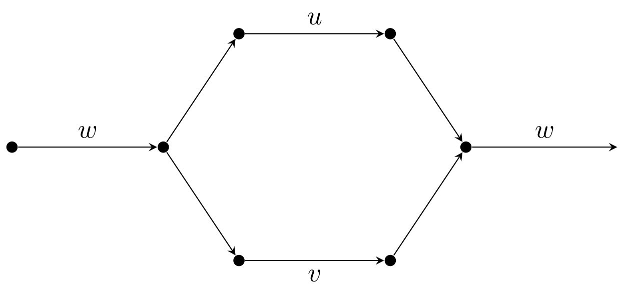


Figure 2.3: Diamond

Definition 10. Let $\mathcal{G} = (V, E, l)$ be a De Bruijn graph of a CA, then its product graph is $(V \times V, E', l')$, where:

$$e' = ((u_1, u_2), (v_1, v_2)) \in E', l(e') = a \iff (u_1, v_1) \in E \wedge (u_2, v_2) \in E \wedge l(u_1, v_1) = l(u_2, v_2) = a$$

A bi-infinite path on the product graph corresponds to a couple $(x_1, x_2) \in A^{\mathbb{Z}} \times A^{\mathbb{Z}}$ such that $F(x_1) = F(x_2)$.

Definition 11. We call a **diagonal** a set $\Delta = \{(u, u)|u \in V\} \subseteq V \times V$

2.6.1 Decidability of injectivity

Here we discuss how to decide whether a global rule of a CA F is injective and surjective. Both questions require knowing which elements have the same image under F. We proceed by constructing a pair graph $\mathcal{G}' = (V', E')$ of the De Brujin graph of the CA $\mathcal{G} = (V, E)$. $V' = V \times V$ where $V = A^{2r}$, for $(u_1, u_2) \in V'$ and $(v_1, v_2) \in V'$ there is an edge with label $a \in A$ if and only if in the De Bruijn graph there are edges with label a from u_1 to v_1 and from u_2 to v_2. Any two-way infinite path p in the pair graph corresponds to two paths in the original De Bruijn graph, obtained by reading only the first or the second components of the pairs. Both paths have the same edge labels, so they correspond to two configurations of the CA with the same image. Recall that Δ is the diagonal set. Notice that the induced subgraph by Δ is isomorphic to $\mathcal{G}$, and in particular Δ is strongly connected. Note that the only cases where two different words result in the same path on $\mathcal{G}'$ are the ones where the path has a vertex outside of Δ.

Proposition 4. *A global rule F of a 1d-CA is:*

1. *Not injective if and only if its pair graph has a cycle that contains a node outside of Δ.*
2. *Not surjective if and only if the pair graph has a cycle that contains a node of Δ and a node outside of Δ.*

Proof.

1. If F is not injective then there are two different spatially periodic configurations c and e with the same image. The path that corresponds to these images in the pair graph is a cycle that contains a node outside of Δ. Conversely, if the pair graph has such a cycle, then the corresponding configuration of the CA has the same image, i.e., the CA is not injective.
2. Let $q \in A$ be abitrary. If F is not surjective then there are two different $q - finite$ configurations c and e with the same image. The corresponding path in the pair graph consists of a loop inside Δ, followed by a cycle that goes outside of Δ and returns to the same cycle inside of Δ. Conversely, if such a cycle exists, then the De Brujin graph has a diamond and the CA is not surjective.

□

2.7 Permutivity (Permutativity)

Here, we define a property that will later be used to give some interesting properties of dynamic systems.

Definition 12. $f : A^{\mathbb{Z}} \to A$ is rightmost-permutive (leftmost-permutive) if and only if $\forall u \in A^{2r}, \forall b \in A, \exists! a \in A$ such that $f(ua) = b$ ($f(au) = b$).

Remark. R-permutivity (L-permutivity) is a decidable property.

Here we give some examples:

$f : \{0,1\}^3 \to \{0,1\}$

x	$f(x)$
000	0
001	1
010	0
011	1
100	0
101	1
110	0
111	1

with $m_f = 170 \quad F = 0$

f is R-perm.

f is not L-perm.

$f : \{0,1\}^3 \to \{0,1\}$

x	$f(x)$
000	0
001	1
010	0
011	1
100	1
101	0
110	1
111	0

with $m_f = 90$

f is R-perm

f is L-perm

$f : \{0,1\}^3 \to \{0,1\}$

x	$f(x)$
000	0
001	1
010	0
011	1
100	0
101	1
110	1
111	0

with $m_f = 106$

f is R-perm

f is not L-perm

Proposition 5. *If f is either R-permutive or L-permutive, then F is surjective.*

Proof. We are going to show that $\forall v \in A^+, |f^{-1}(v)| = |A|^{2r}$. Choose an arbitrary $v \in A^n$ and let $u \in A^{2r}$ be any word. Since f is $R-permutive$, $\exists! a_1 \in A$ such that $f(ua_1) = v_1$, then $\exists! a_2 \in A$ such that $f(u_2 \dots u_n a_1 a_2) = v_2 \dots \exists! a_n \in A$ such that $f(u_n a_1 \dots a_{n-1} a_n) = v_n$. Therefore, $\forall u \in A^{2r} \quad \exists! \alpha = a_1 \dots a_n$ such that $f(u\alpha) = v$, hence v has exactly 1 pre-image $\forall u \in A^{2r}$. ☐

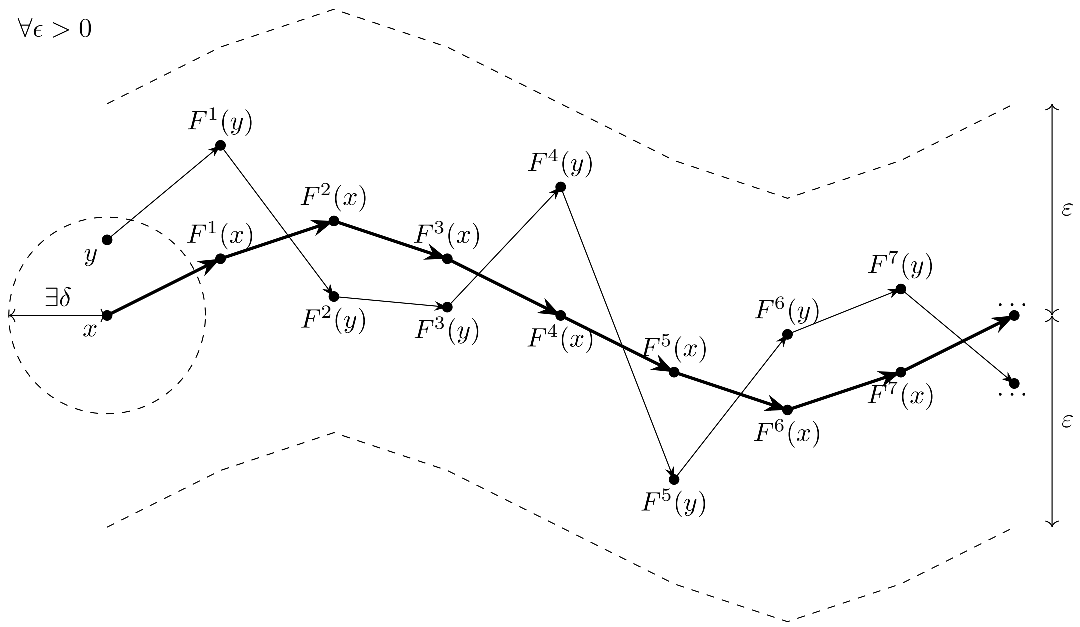
Chapter 3

Stability in discrete-time dynamical systems

The notions given in this chapter (except for when explicitly stated) are valid for any discrete-time dynamical system (**DTDS**), i.e., for any pair (X, F) where (X, d) is a metric space and $F : X \to X$ is a continuous transformation.

3.1 Stability

Definition 13 (Stability point)**.** A point $x \in X$ is a stability (equicontinuity, Lyaponov stable) point if and only if $\forall \varepsilon > 0, \exists \delta > 0, \forall y \in X, d(y, x) < \delta \implies \forall t \in \mathbb{N} \quad d(F^t(y), F^t(x)) < \varepsilon$.



When $X = A^{\mathbb{Z}}$ and d is the Tychonoff distance, one could equivalently write $\forall n \in \mathbb{N}, \exists m \in \mathbb{N}, \forall y \in A^{\mathbb{Z}}, y_{[-m,m]} = x_{[-m.m]} \implies \forall t \in \mathbb{N}, F^t(y)_{[-n,n]} = F^t(x)_{[-n,n]}$.

Definition 14 (Instability point)**.** $x \in X$ is an unstable point if x is not a stability point, that is $\exists \varepsilon > 0, \forall \delta > 0, \forall y \in X, d(y, x) < \delta \wedge \exists t \in \mathbb{N} \quad d(F^t(y), F^t(x)) \geq \varepsilon$

Definition 15 (Stable system (global stability))**.** (X, F) is stable if $\forall x \in X$, x is a stability point, that is:
$\forall x \in X, \forall \varepsilon > 0, \exists \delta > 0, \forall y \in X, d(y, x) < \delta \implies \forall t \in \mathbb{N} \quad d(F^t(y), F^t(x)) < \varepsilon$.

Definition 16 (Equicontinuous systems)**.** $\forall \varepsilon > 0, \exists \delta > 0, \forall x, y \in X, d(y, x) < \delta \implies \forall t \in \mathbb{N} \quad d(F^t(y), F^t(x)) < \varepsilon$.

Remark. Equicontinuity is a stronger condition than global stability.

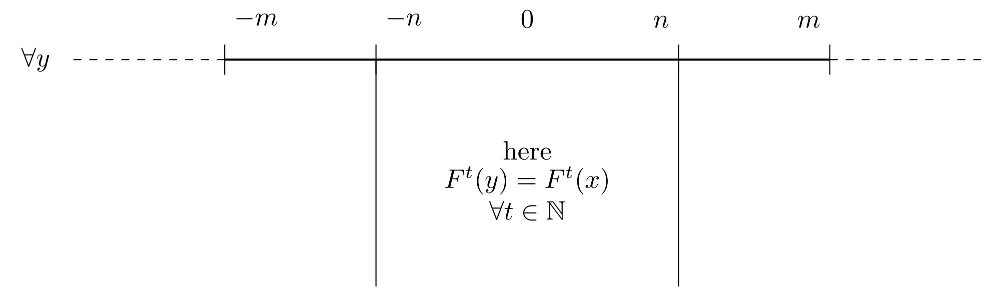
Definition 17 (Unstable system)**.** (X, F) is unstable if $\forall x \in X$, x is an instability point, that is:
$\forall x \in X, \exists \varepsilon > 0, \forall \delta > 0, \forall y \in X, d(y, x) < \delta \wedge \exists t \in \mathbb{N} \quad d(F^t(y), F^t(x)) \geq \varepsilon$

Definition 18 (Sensitivity to initial conditions)**.** (X, F) is sensitive to the initial conditions if
$\exists \varepsilon > 0, \forall \delta > 0, \forall x, y \in X, d(y, x) < \delta \wedge \exists t \in \mathbb{N} \quad d(F^t(y), F^t(x)) \geq \varepsilon$.

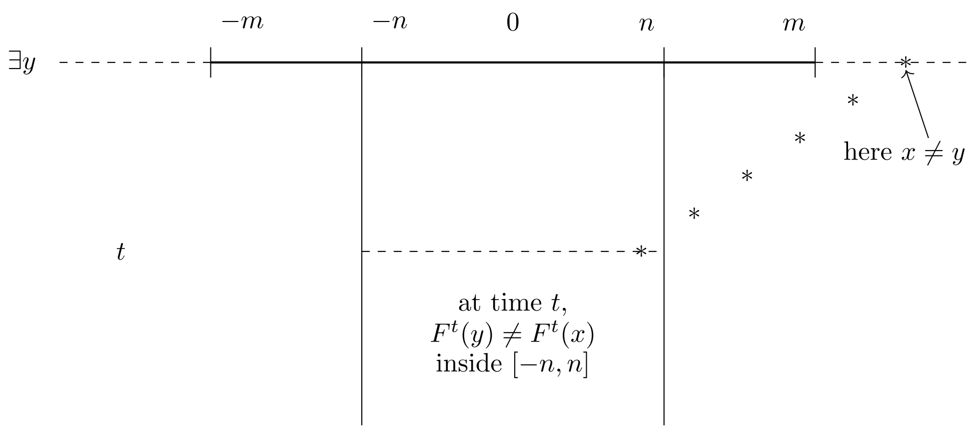
Remark. A sensitive system is an unstable system in which the constant $\varepsilon > 0$ doesn't depend on $x \in X$. Such a constant is called **sensitivity constant**, and it is a specific feature of a system.

Here we show some examples of such definitions on a DTDS $(A^{\mathbb{Z}}, f)$:

- **Equicontinuity point**: a point $x \in A^{\mathbb{Z}}$ such that: $\forall n \in \mathbb{N}, \exists m \in \mathbb{N}, \forall, y \in A^{\mathbb{Z}}, y_{[-m,m]=x_{[-m,m]}} \implies \forall t \in \mathbb{N} \quad F^t(y)_{[-n,n]} = F^t(x))_{[-n,n]}$.



- **Unstable point**: a point $x \in A^{\mathbb{Z}}$ such that $\exists \varepsilon \in \mathbb{N}$, $\exists y \in A^{\mathbb{Z}}, y_{[-m,m]=x_{[-m,m]}} \wedge \exists t \in \mathbb{N} \quad F^t(y)_{[-n,n]} \neq F^t(x))_{[-n,n]}$



We can give a weaker notion of equicontinuity, which will be used later to describe properties

Definition 19 (Almost equicontinuity)**.** A DTDS is almost equicontinuous if the set ξ of its equicontinuity points is residual (ξ is also dense).

What is a residual set? Here we give an example: $I = \bigcap_{\alpha \in \mathbb{Q}} (\mathbb{R} \setminus \{\alpha\})$, then I is residual in $\mathbb{R}$ because:

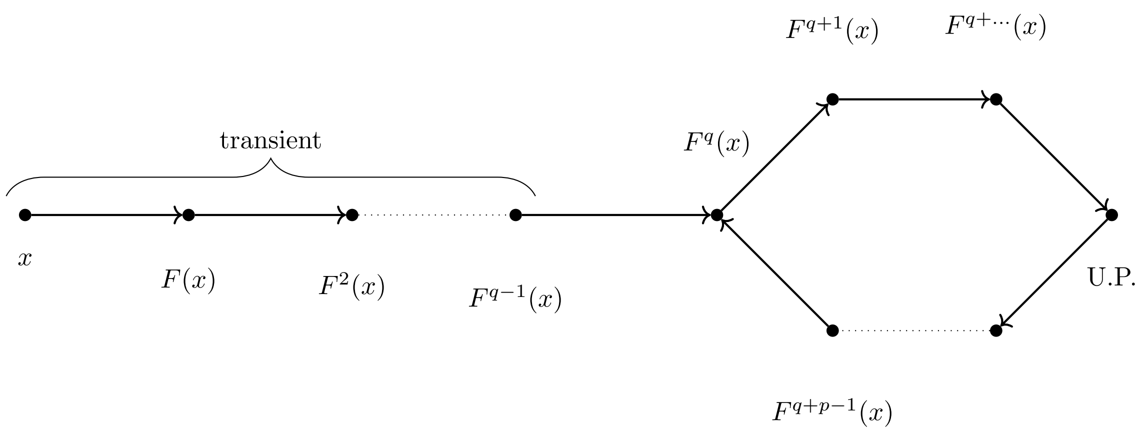
- $\mathbb{R} \setminus \{\alpha\} = (-\infty, \alpha) \cup (\alpha, \infty)$ is open in $\mathbb{R}$.
- $\mathbb{R} \setminus \{\alpha\}$ is dense in $\mathbb{R}$.
- The intersection $\bigcap_{\alpha \in \mathbb{Q}} (\mathbb{R} \setminus \{\alpha\})$ is countable.

I also contains a countable intersection of open and dense subsets of $\mathbb{R}$.

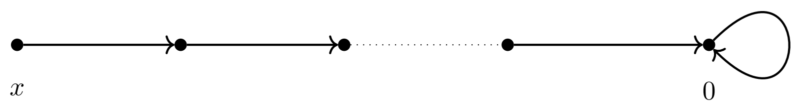
Definition 20. Let (X, d) be a metric space. Let $Y \subseteq X$. Y is residual in X if it contains a countable intersection of open and dense subsets of X.

We can also give a stronger notion of equicontinuity

Definition 21 (Ultimate periodicity (U.P.) and Nilpotency)**.** (X, F) is ultimately periodic if there exist $p, q \in \mathbb{N}, p > 0, q \geq 0$, called respectively period and pre-period, such that $F^{q+p}(x) = F^q(x), \forall x \in X$.



(X, F) is nilpotent if $\exists t_0 \in \mathbb{N} : \forall t > t_0 \forall x \in X, F^t(x) = zero$ where $zero \in X$.



$\{x, F(x), \ldots F^{q-1}(x)\}$ is called **transient**, and its length is equal to the pre-period q.
$\{F^q(x), F^{q+1}(x), \ldots, F^{q+p-1}(x)\}$ is called **cycle**, and its length is equal to the period p.

Remark. Let (X, F) if $|X| < \infty$, then necessarily (X, F) is ultimately periodic, with $q = \max\{q_x | x \in X\}$ and $p = lcm\{p_x | x \in X\}$.

Now we give two interesting properties

Proposition 6. *Let (X, F) be a DTDS, then:*

1. *(X, F) is ultimately periodic and (X, d) is compact $\implies$ (X, F) is equicontinuous*
2. *(X, F) is equicontinuous $\implies$ (X, F) is almost equicontinuous*

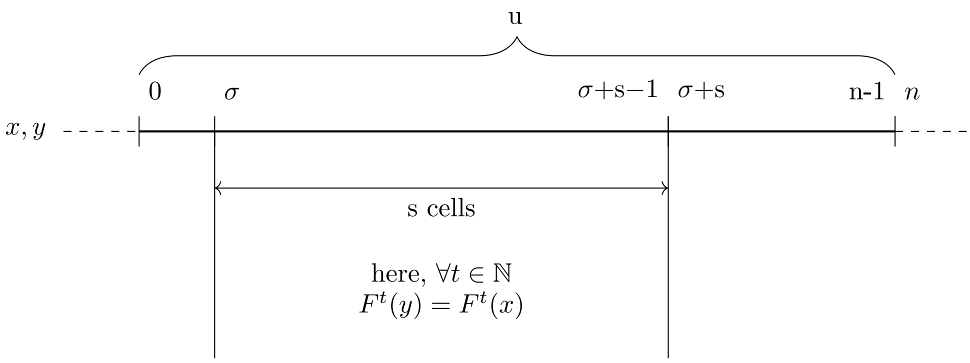
Proof. We first show the second implication since it is easier to prove
2)
If (X, F) is equicontinuous, then $\xi = X$. Trivially, we have $\xi = \bigcap_{n \in \mathbb{N}} X_n$ with $X_n = X$, where X is open and dense. Therefore, (X, F) is also almost equicontinuous.
1)
We show that $\forall \varepsilon > 0, \exists \delta > 0, \forall x, y \in X, d(y, x) < \delta \implies \forall t \in \mathbb{N} d(F^t(y), F^t(x)) < \varepsilon$. Choose an arbitrary $\varepsilon > 0$. Since $F^0 = Id, F^1, \ldots, F^q = F^{q+p}, \ldots, F^{q+p-1}$ are uniformly continuous functions, we have that each of these functions has a certain δ_t that satisfies the inequality. Then $\exists \delta = \min\{\delta_0, \delta_1, \ldots, \delta_q, \ldots, \delta_{q+p-1}\}$, then the implication holds for every F^t since δ is the minimum. ☐

3.2 Stability/instability in cellular automata

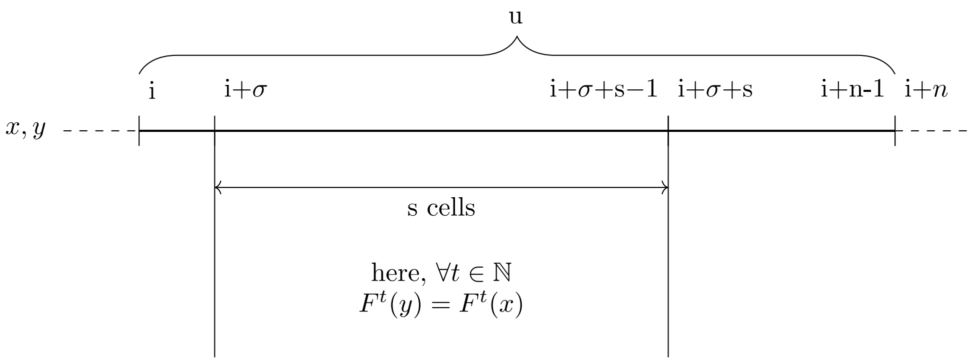
First, we make a remark regarding periodicity in configurations.

Remark. Any spatially periodic configuration is ultimately periodic. Let $x \in A^{\mathbb{Z}}$ be such that $\sigma^n(x) = x$ for some $n \in \mathbb{N}$. Then, $x = ^{\infty} u^{(0)\infty}$ for some $u^{(0)} \in A^n$. Let F be the global rule of a CA. Since the image of a spatially periodic configuration is also spatially periodic, we have that $F(x) = ^{\infty} u^{(1)\infty}$ for some $u^{(1)} \in A^n$, and so on. Since the set A^n is finite, necessarily it holds that $\exists q_x \geq, p_x > 0$ such that $F^{q_x+p_x}(x) = F^{q_x}(x)$, therefore x is ultimately periodic.

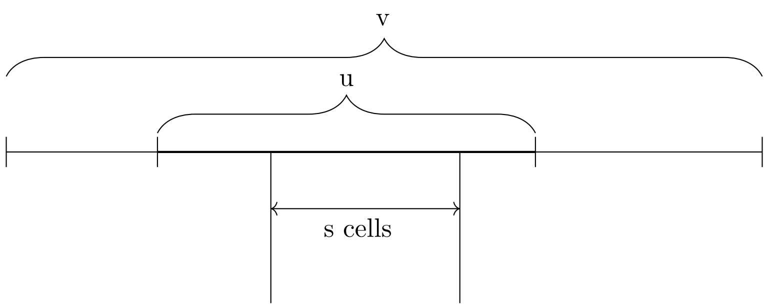
Definition 22 (Blocking word for a CA)**.** Let $u \in A^n$ for some $n \in \mathbb{N}$. Let $s \in \mathbb{N}$ with $s > 0$. u is s-blocking if $\exists \sigma \in [0, n-s]$ (offset) such that $\forall x, y \in A^{\mathbb{Z}} : x_{[0,n)} = y_{[0,n)} = u$ it holds that $\forall t \in \mathbb{N}, F^t(x)_{[\sigma,\sigma+s)} = F^t(y)_{[\sigma,\sigma+s)}$. In other words, u produces a disconnection of the cell space $\mathbb{Z}$. If $s \geq r$ u is a "wall", that means that cells on the left side of u cannot interact with the ones on the right side.



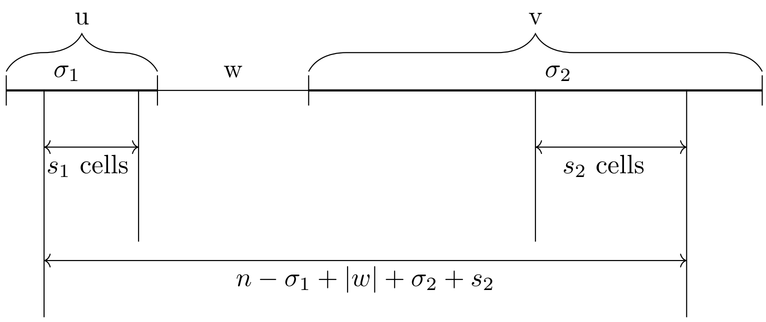
Remark. Since F is the global rule of a CA it is permutative with respect to σ, therefore the previous definiton of blocking word can be extended, to use a generic starting index instead of 0. Let $u \in A^n$ for some $n \in \mathbb{N}$. Let $s \in \mathbb{N}$ with $s > 0$. u is s-blocking if $\exists \sigma \in [0, n-s]$ (offset) such that $\forall i \in \mathbb{Z}, \forall x, y \in A^{\mathbb{Z}}$: $x_{[i,i+n-s]} = y_{[i,i+n-s]} = u$ it holds that $\forall t \in \mathbb{N}, F^t(x)_{[i+\sigma,i+\sigma+s]} = F^t(y)_{[i+\sigma,i+\sigma+s]}$.



Remark. If $u \in A^n$ is an s-blocking word, then every $v \in A^m$ with $m > n$ that contains u as a factor is also an s-blocking word.



Remark. If $u \in A^n$ is an s_1-blocking word with offset σ_1 and $v \in A^m$ is an s_2-blocking word with offset σ_2, then $\forall w \in A^*, uwv$ is a blocking word.



3.2.1 Relationship between blocking words and stability in CAs

Proposition 7. *Let F be a global rule of a CA of radius r. If u is an r-blocking word, then the set ξ of equicontinuity points of F is infinite (and dense).*

sketch of proof. Let $x \in A^{\mathbb{Z}}$ be any configuration of the form $\ldots uSOALuSOALu \ldots$ where $SOAL$ is something of arbitrary length (not necessarily all of the same length). We are going to show that x is an equicontinuity point, i.e., $\forall n \in \mathbb{N}, \exists m \in \mathbb{N}, \forall y \in A^{\mathbb{Z}}, y_{[-m,m]} = x_{[-m,m]} \implies F^t(y)_{[-n,n]} = F^t(x)_{[-n,n]}$. Choose $n \in \mathbb{N}$ and let m be such tat $[-n,n] \subseteq [-m,m]$, and $[-n,n]$ is entirely contained inside the segment starting from the first and the last wall of $x_{[-m,m]}$. Then x is an equicontinuity point □

Theorem 7 (Characterization of CA equicontinuity)**.** *Let $< A, r, f >$ be a cellular automata, let F be its global rule. The following statements are equivalent:*

1. *$(A^{\mathbb{Z}}, F)$ is equicontinuous.*
2. *$\exists k > 0$ such that $\forall u \in A^{2k+1}$, u is $r-blocking$.*
3. *$(A^{\mathbb{Z}}, F)$ is ultimately periodic.*

Proof. 3) $\implies$ 1)
We already proven that in every DTDS U.P. and density implies equicontinuity, since $A^{\mathbb{Z}}$ is dense this always holds true for cellular automata.
1) $\implies$ 2)
We know that $\forall n \in \mathbb{N}, \exists k \in \mathbb{N}, \forall x, y \in A^{\mathbb{Z}}, x_{[-k,k]} = y_{[-k,k]} \implies \forall t \in \mathbb{N} F^t(x)_{[-n,n]} = F^t(y)_{[-n,n]}$. If we pick $n = r$ then we have $\exists k \in \mathbb{N}, \forall x, y \in A^{\mathbb{Z}}, x_{[-k,k]} = y_{[-k,k]} \implies \forall t \in \mathbb{N}, F^t(x)_{[-r,r]} = F^t(y)_{[-r,r]}$. Thus, any $u \in A^{2k+1}$ is $(2r+1)$-blocking, therefore it's also r-blocking.
2) $\implies$ 3)
We are going to show that there exists two integers $p > 0$ and $q \geq 0$ such that $F^{q+p} = F^q$. For each $u \in A^{2k+1}$, consider $x^u =^\infty u^\infty$, which is a spatially periodic configuration. We know that $\{F^t(x^u)\}_{t\in\mathbb{N}}$ is ultimately periodic. Then $\{F^t(x^u)_0\}_t$ is ultimately periodic too. This also means that $\exists p_u > 0$ and $q_u \geq 0$ such that $F^{q_u+p_u}(x^u)_0 = F^{q_u}(x^u)_0$, Set $p = lcm\{p_u | u \in A^{2k+1}\}$ and $q = \max\{q_u | u \in A^{2k+1}\}$. Therefore, $\forall u \in A^{2k+1}, F^{q+p}(x^u)_0 = F^q(x^u)_0$. Let $x \in A^{\mathbb{Z}}$ be any configuration and let $u = x_{[-k,k]}$. Since u is a blocking word $\forall t \in \mathbb{N}$ F^t and $F^t(x)$ are equal inside a window of length r (assume without losing generality that this window contains position 0). Hence, $F^{q+p}(x)_0 = F^q(x)_0$. Now $\forall i \in \mathbb{Z}, F^{q+p}(x)_i = \sigma^i(F^{q+p}(x))_0 = F^{q+p}(\sigma^i(x))_0 = F^q(x)_i = \sigma^i(F^q(x))_0 = F^q(x)_i$. □

Theorem 8 (Characterization of almost equicontinuity)**.** *Let $< A, r, f >$ be a CA, let F be its global rule. The following statements are equivalent:*

1. *$(A^{\mathbb{Z}}, F)$ is not sensitive.*
2. *$(A^{\mathbb{Z}}, F)$ has an r-blocking word.*
3. *$(A^{\mathbb{Z}}, F)$ is almost equicontinuous.*

Proof. 1) $\implies$ 2)
Suppose that F is not sensitive, i.e., $\forall n \exists x \in A^{\mathbb{Z}} \exists m$ such that $\forall y \in A^{\mathbb{Z}}, y_{[-m,m]} = x_{[-m,m]} \implies \forall t \in \mathbb{N}, F^t(y)_{[-n,n]} = F^t(x)_{[-n,n]}$. Choose arbitrarily n such that $2n+1 \geq r$. Then $\exists x \in A^{\mathbb{Z}}$ and $\exists m$ such that $\forall y, y_{[-m,m]} = x_{[-m,m]} \implies \forall t \in \mathbb{N}, F^t(y)_{[-n,n]} = F^t(x)_{[-n,n]}$. Therefore, $u = x_{[-m,m]}$ is $(2n+1)$-blocking. Indeed, $\forall z, z' \in A^{\mathbb{Z}}$ such that $u = x_{[-m,m]} = z_{[-m,m]} = z'_{[-m,m]}$ it holds true

that $\forall t, F^t(z)_{[-n,n]} = F^t(x)_{[-n,n]}$ and $F^t(z')_{[-n,n]} = F^t(x)_{[-n,n]}$, hence $\forall t, F^t(z)_{[-n,n]} = F^t(z')_{[-n,n]} = F^t(x)_{[-n,n]}$. Therefore, u is $(2r+1)$-blocking, and also r-blocking.
2) $\implies$ 3)
We have already proven that 2) $\implies$ there exists an infinite number of equicontinuity points. For the rest of the proof, see "TOPOLOGICAL DYNAMICS OF CELLULAR AUTOMATA" by Petr Kůrka.
3) $\implies$ 1)
Trivially, if $(A^{\mathbb{Z}}, F)$ is almost equicontinuous, then it has one equicontinuity point and therefore, cannot be sensitive. □

3.3 Classification of CA with respect to stability

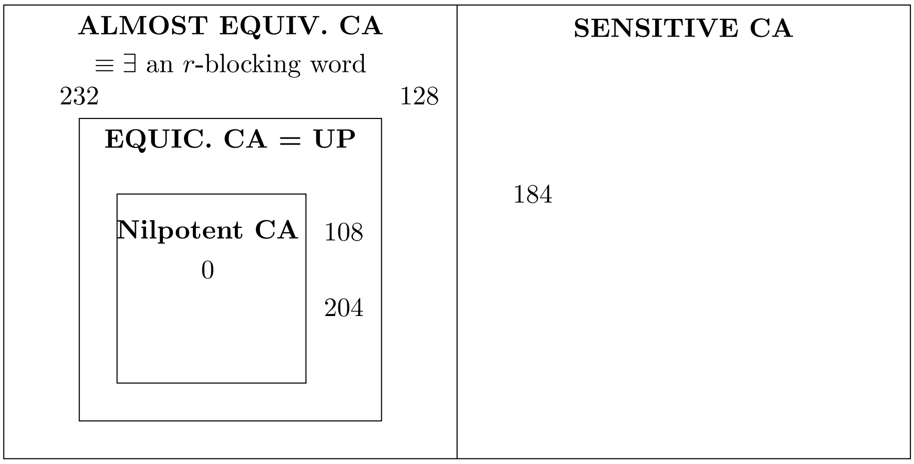


Figure 3.1: Classification of cellular automata previously presented as examples.

Proposition 8. *The problem of establishing whether a CA admits an r-blocking word is semi-decidable (recursively enumerable, not recursive).*

Corollary 7. *It is undecidable to establish whether a CA is almost equicontinuous or sensitive. This is because of the previous proposition. Almost equicontinuity is equivalent to having at least one r-blocking word; it is not decidable. Sensitivity is the negation of almost equicontinuity, and the negation of a semi-decidable problem is undecidable.*

Proposition 9. *It is undecidable to establish whether a CA is nilpotent.*

Corollary 8. *It is undecidable to establish whether a CA is equicontinuous (equiv. UP)*

sketch of proof. Assume ultimately periodicity is decidable. Let $< A, r, f >$ be any CA and let F be its global rule. Since UP is decidable, we can establish whether F falls into one of the following cases:

1. F is not UP, then F is not nilpotent.

2. F is UP, then $\exists p > 0, q \in \mathbb{N}$ such that $F^q = F^{q+p}$. By taking the local rule $f : A^{2r+1} \to A$ and applying it t times we have $f^t : A^{2rt+1} \to A$. Since F is UP, we will have $f^i = f^j$ with $i \neq j$. One of these functions will have a smaller domain; this will give us q.

   (a) $p \neq 1 \implies F$ is not nilpotent.

   (b) $p = 1 \implies$

      i. if $^{\infty}q^{\infty}$ for some $q \in A$ is the unique x such that $F(x) = v \implies F$ is nilpotent.

      ii. otherwise, F is not nilpotent.

However, this is a decision algorithm for nilpotency, which we have stated to be undecidable; therefore, the initial assumption that UP is decidable is false. □

3.3.1 Nilpotency in CA

First, we recall the definition of nilpotency, $\exists zero \in A^{\mathbb{Z}}$ such that $\exists t \in \mathbb{N}, \forall x \in A^{\mathbb{Z}}, F^t(x) = zero$.

Remark. $\sigma(zero) = zero$ and $F(zero) = zero$.

3.3.2 A sufficient condition for CA sensitivity

Let F be the global rule of a CA where its local rule if $f : A^{2r+1} \to A$. For any $t \in \mathbb{N}$, let $f^t : A^{2rt+1} \to A$ be the local rule of the CA with global rule F^t. If f is leftmost/rightmost permutive, then f^t also is.

Proposition 10. *Let F be the global rule of a CA with a leftmost/rightmost permutive local rule f. Then, F is sensitive to the initial conditions.*

Proof. We are going to show that $\exists n = r$ such that $\forall x \in A^{\mathbb{Z}}, \forall m \in \mathbb{N}, \exists y \in A^{\mathbb{Z}}, y \neq x$ with $y_{[-m,m]} = x_{[-m,m]} \wedge \exists t \in \mathbb{N} : F^t(y)_{[-r,r]} \neq F^t(x)_{[-r,r]}$. Choose arbitrarily $x \in A^{\mathbb{Z}}$ and $m \in \mathbb{N}$. Build $y \in A^{\mathbb{Z}}$ as follows: Let $t \in \mathbb{N}$ such that $[-m, m] \subset [r - tr, r + tr]$ and set

1. $y_{[r-tr,r+tr-1]} = x_{[r-tr,r+tr-1]}$

2. $y_{r+tr} \neq x_{r+tr}$

So $y_{[-m,m]} = x_{[-m,m]}$. Since f^t is R-permutive, by 1. and 2. we get:

$$\begin{aligned} F^t(y)_r &= f^t(y_{[r-rt,r+rt]}) \\ &= f^t(y_{[r-rt,r+rt-1]}y_{r+rt}) \\ &= f^t(x_{[r-rt,r+rt-1]}y_{r+rt}) \neq f^t(x_{[r-rt,r+rt-1]}x_{r+rt}) = F^t(x)_r \end{aligned}$$

□

3.4 Expansivity

Here we give additional properties for generic DTDS and cellular automata.

Definition 23 (Positive expansivity)**.** A DTDS (X, F) is said to be positively expansive if $\exists \varepsilon > 0, \forall x, y \in X, x \neq y\ \exists t \in \mathbb{N} : d(F^t(y), F^t(x)) \geq \varepsilon$. In $A^{\mathbb{Z}}$ $\exists n \in \mathbb{N}, \forall x, y \in A^{\mathbb{Z}}, x \neq y\ \exists t \in \mathbb{N} : F^t(y)_{[-n,n]} \neq F^t(x)_{[-n.n]}$

Proposition 11. *Let (X, F) be a DTDS with $|X| = \infty$, where X has no isolated elements. If (X, F) is positively expansive, then it is sensitive.*

This proposition is easy to prove and the proof is left as an exercise to the reader.

Remark. If a DTDS (X, F) has $|X| < \infty$ then:

1. It is ultimately periodic (therefore, equicontinuous).

2. It is expansive.

Proof. Here we only prove 2. since 1. is obvious. Set $\varepsilon = \min\{d(x, y)|x, y \in X, x \neq y\}$. Such an ε exists (since X is finite) and it is strictly positive. Moreover, $\forall x, y \in X, d(x, y) \geq \varepsilon$. □

In general expansivity of (X, F) doesn't imply that F is unstable. Regarding CAs, the following holds true:

Proposition 12. *Let $(A^{\mathbb{Z}}, F)$ be a CA with a local rule which is both leftmost and rightmost permutive. Then, F is positively expansive.*

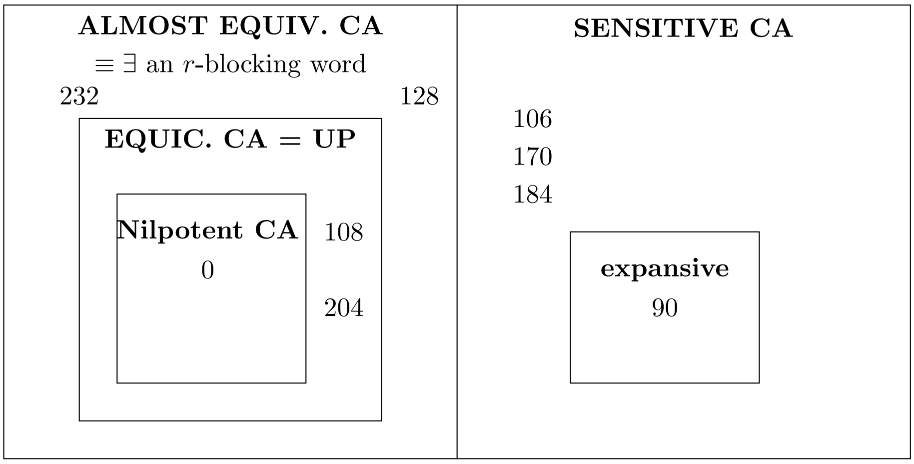


Figure 3.2: Expanded classification of cellular automata previously presented as examples.

3.5 Chaos

In this section we discuss the notions required to define the behavior of a DTDS as chaotic. Sensitivity is not enough to label a DTDS as chaotic, consider (X, F) where:

- $X = \mathbb{R}$
- $\forall x \in X, F(x) = 2 \cdot x$ (F is linear).

Then $\forall t \in \mathbb{N}, F^t(x) = 2^t \cdot x$ for any x. So the dynamical evolution of the initial state x is $(x, 2x, 4x, \ldots, 2^t x, \ldots)$. Thus, $\forall x, y \in \mathbb{R}, d(F^t(x), F^t(y)) = 2^t|x - y|$. It is easy to see that F is expansive, then it is also sensitive. If we interpret chaos as physics usually does, a linear system cannot be chaotic by definition, therefore, we need further notions.

Definition 24 (Transitivity)**.** A DTDS (X, F) is said to be transitive if $\forall x \in X, \forall \delta > 0 \forall y \in X, \forall \varepsilon > 0, \exists z \in X$ with $d(z, x) < \delta$ and $\exists t \in \mathbb{N} : d(F^t(z), y) < \varepsilon$

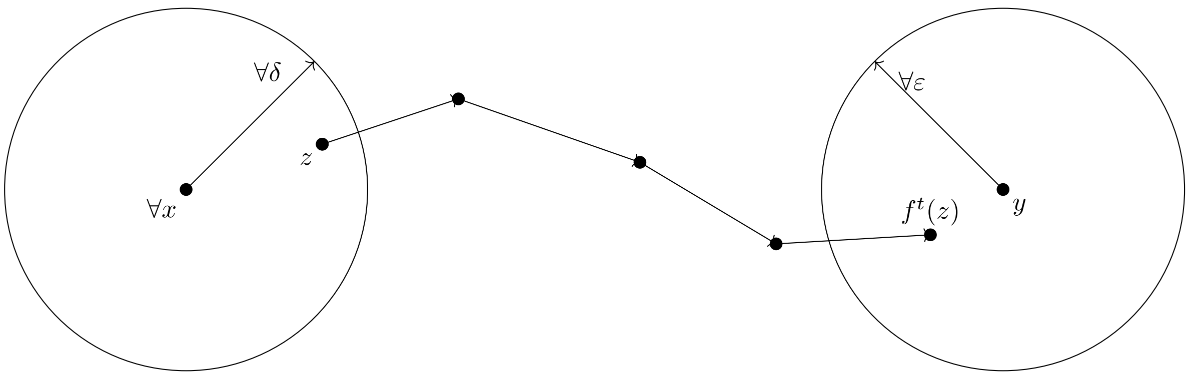


Figure 3.3: Visual representation of transitivity

Definition 25 (Dense periodic orbits (DPO))**.** A DTDS (X, F) has dense periodic orbits if $\forall x \in X, \forall \varepsilon > 0, \exists$ a periodic point p such that $d(x, p) < \varepsilon$. A periodic point is an ultimately periodic point p with preperiod $q = 0$, i.e., with orbit $(p, F(p), \ldots, F^k(p))$ for some $k > 0$.

Definition 26 (Chaotic DTDS (according to Devaney))**.** A DTDS (X, F) is chaotic if all the following conditions are met:

- $|X| = \infty$
- (X, F) is sensitive

- (X, F) is transitive
- (X, F) has DPO

In reality, further research has discovered that sensitivity is redundant because of the following theorem

Theorem 9. *If a DTDS (X, F) with $|X| = \infty$ is transitive and has DPO, then it is sensitive.*

3.5.1 Chaos in cellular automata

Theorem 10. *If a CA $(A^{\mathbb{Z}}, F)$ is transitive, then it is sensitive.*
If a CA $(A^{\mathbb{Z}}, F)$ is expansive, then it is transitive.

Proposition 13. *A CA $(A^{\mathbb{Z}}, F)$ with leftmost (or rightmost) local rule is transitive.*

Proposition 14. *A CA $(A^{\mathbb{Z}}, F)$ with leftmost (or rightmost) local rule has DPO.*

Corollary 9. *From the two previous propositions, then a CA $(A^{\mathbb{Z}}, F)$ with leftmost (or rightmost) local rule is chaotic.*

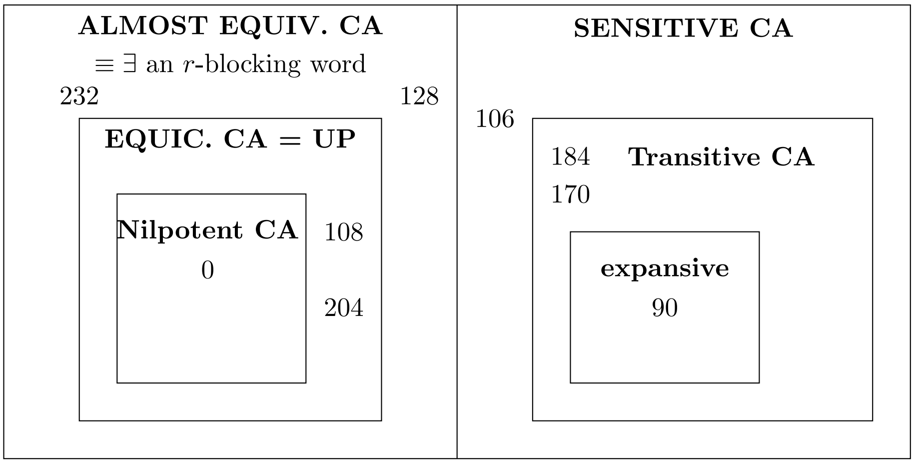


Figure 3.4: Expanded classification of cellular automata previously presented as examples.

Then could we say that transitivity ≡ chaos?

Remark. There exist non-permutative CAs that are transitive or expansive. Take an L-permutive CA with global rule F and an injective CA with global rule G which is neither L nor R-permutive. Then $G^{-1} \circ F \circ G$ is a CA which is not permutive. But it is transitive!

3.6 Linear cellular automata

Let $A = \mathbb{Z}_s = \{[0], [1], \dots, [s-1]\} = \mathbb{Z}/s\mathbb{Z} \quad < \mathbb{Z}_s, +, \cdot >$

Definition 27 (Linear local rule)**.** $f : \mathbb{Z}_s^{2r+1} \to \mathbb{Z}_s$ is linear if $\exists \lambda_{-r} \dots \lambda_0 \dots \lambda_r \in \mathbb{Z}_s$ such that $\forall (x_{-r}, \dots, x_r) \in \mathbb{Z}_s^{2r+1}, f(x_{-r}, \dots, x_r) = (\sum_{i=-r}^{r} \lambda_i \cdot x_i) \mod s$

Definition 28. A CA is linear if it has a linear local rule f.

Theorem 11. *Let $(\mathbb{Z}_s^{\mathbb{Z}}, F)$ be a linear CA defined by a linear local rule with coefficients $\lambda_{-r}, \dots, \lambda_0, \dots, \lambda_r$. Let $\mathcal{P}$ denote the set of prime factors of s, then:*

- *F is surjective* $\iff \gcd(s, \lambda_{-r}, \dots, \lambda_0, \dots, \lambda_r) = 1$
- *F is injective* $\iff \forall p \in \mathcal{P}, \exists! \lambda_i$ *such that* $p \nmid \lambda_i$
- *F is transitive* $\iff \gcd(s, \lambda_{-r}, \dots, \lambda_{-1}, \dots, \lambda_1, \dots, \lambda_r) = 1$
- *F is sensitive* $\iff \exists p \in \mathcal{P} : p \nmid \gcd(\lambda_{-r}, \dots, \lambda_{-1}, \dots, \lambda_1, \dots, \lambda_r)$

- *F is expansive* $\iff$ $\gcd(\lambda_{-r}, \ldots, \lambda_{-1}) = \gcd(\lambda_1, \ldots, \lambda_r) = 1$
- *F is surjective $\iff$ F has DPO.*
- *F is equicontinuous $\iff$ F is almost equicontinuous $\iff$ F is not sensitive.*
- *F is chaotic $\iff$ F is transitive.*

Corollary 10. *All the properties mentioned above are decidable for linear cellular automata.*

Chapter 4

Tiling

In this chapter, we will describe the tiling problem and how it is related to cellular automata. In particular, we will show how properties of the tiling problem let us prove whether properties are decidable for 1d-CAs and higher-dimensional CAs.

Definition 29. Let C be a set of colors.

- A **Wang tile** is a unit square tile with colored edges, with colors from C; one tile is an element of C^4.
- A **tile set** $T \subseteq C^4$ is a finite collection of such tiles where no tile has all four sides of the same color.
- A **valid tiling** is an assignment $\mathbb{Z}^2 \to T$ of tiles on an infinite square lattice, so that the abutting edges of adjacent tiles have the same color. We say that such an assignment is without errors.
- A **finite tiling** is a map from $M \subseteq \mathbb{Z}^2$ to T, with $|M| < \infty$.

4.1 Tiling problem

The tiling problem (or Domino problem) of Wang tiles is the decision problem to determine if a given finite set of Wang tiles admits a valid tiling of the plane.

Theorem 12 (R. Berger 1966)**.** *The Domino problem is undecidable.*

Remark (1)*.* If T admits valid tilings inside squares of arbitrary size, then it admits a valid tiling of the whole infinite plane. There are two equivalent possibilities to formalize this property:

- $\forall n, \exists$ a square $n \times n$ in which T admits a (finite) valid tiling $\implies$ T admits a valid tiling of the whole plane $\mathbb{Z}^2$.
- T does not admit a valid tiling of $\mathbb{Z}^2 \implies \exists n, \forall$ square $n \times n$ T contains a tiling error.

This is a direct consequence of compactness.

Remark (2)*.* There is a semi-algorithm to recursively enumerate tile sets that do not admit valid tilings of the plane. From the previous remark, we know we can try larger squares until one is found.

Remark (3)*.* There is a semi-algorithm to recursively enumerate tile sets that admit a valid periodic tiling. Try all squares until a valid tiling is found where the top and bottom match, and the left and right match.

4.1.1 Correlation to problems in CA

The tiling problem can be reduced to various decision problems concerning two-dimensional cellular automata, so that the undecidability of these problems then follows from Berger's theorem. In a way, one could see Wang tilings as a "static" version of "dynamic" CAs.

Fixed points

Theorem 13. *It is undecidable whether a two-dimensional CA has any fixed point configurations, that is, a configuration c such that $F(c) = c$, where F is the global rule.*

Proof. We show a reduction from the tiling problem. For any given Wang tile set T (with at least two tiles), we effectively construct a two-dimensional CA with state set T, the von Neumann-neighborhood, and a local update rule that keeps a tile unchanged if and only if its colors match with the neighboring tiles. Trivially, $F(c) = c$ if and only if c is a valid tiling. But the domino problem is undecidable; therefore, the existence of fixed points in 2d-CA is also undecidable. □

Remark. The previous theorem doesn't hold for 1d-CA, where fixed points are decidable. Take the De Brujin graph of the CA $\mathcal{G}$, remember how a path in such a graph corresponds to the application of the global rule with respect to the nodes of the path. Then, we construct $\mathcal{G}'$ by removing the arches where the label of the arch does not correspond to the "linking" symbol of the nodes of the arch, since that would mean that the application of the rule changes the symbol. Then, $\exists$ a fixed point for such CA $\iff$ $\mathcal{G}'$ contains a cycle.

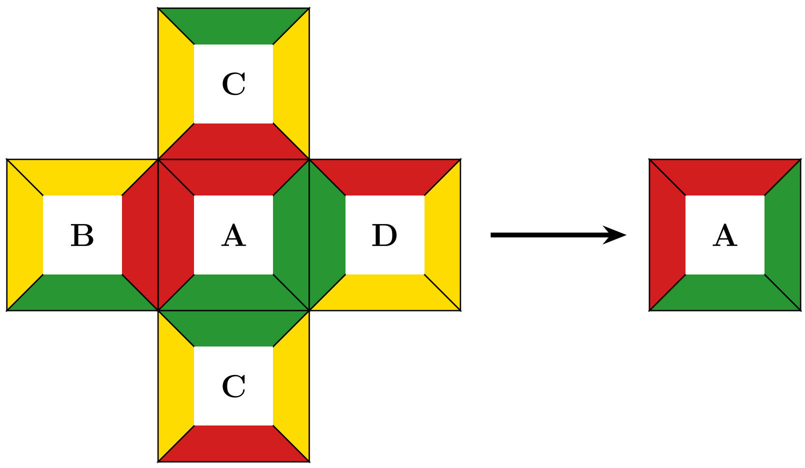
Nilpotency

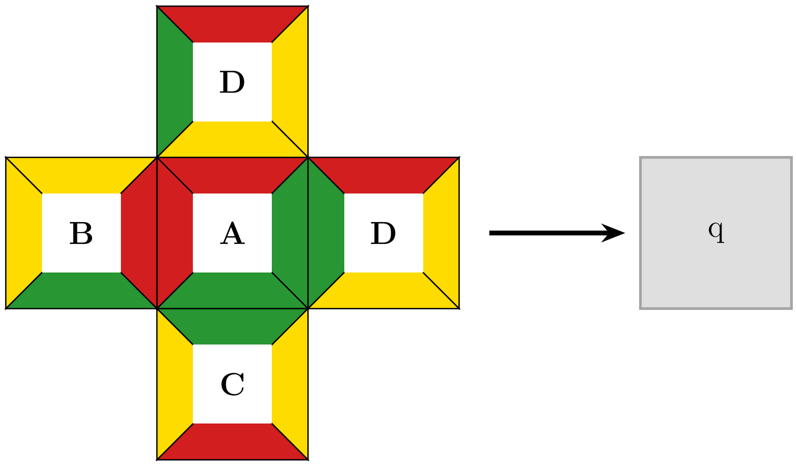
We now discuss a more interesting reduction regarding nilpotency. Remember the definition of nilpotency given in section 3.3.1. Note that the time required for the CA to become quiescent is bounded.

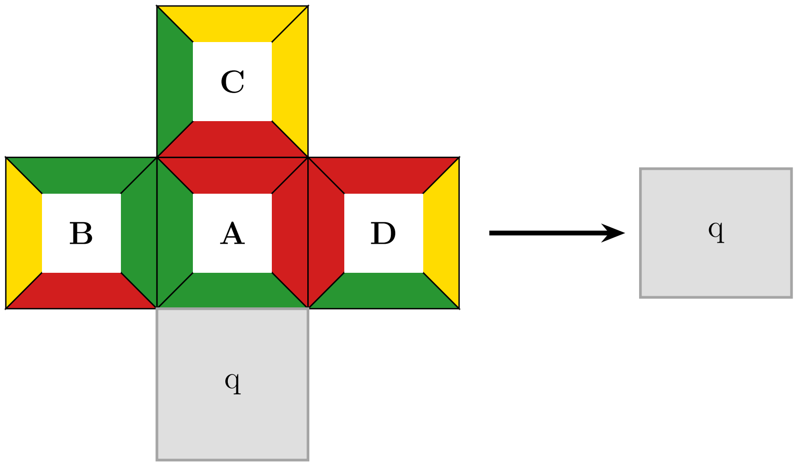
Theorem 14 (Culik, Pachhl, Yu 1889)**.** *It is undecidable whether a given two-dimensional CA is nilpotent.*

Proof. The goal is to show that for any tile set T we can construct a CA that is nilpotent if and only if T does not admit a tiling. For a tile set T we build the following CA:

- State set is $S = T \cup \{q\}$ where q is a new symbol $q \notin T$.
- Von Neumman neighborhood.
- The local rule keeps the state unchanged if all states in the neighborhood are tiles and the tiling constraint is satisfied, in all other cases the new state is q.







$\Longrightarrow$ If T admits a tiling c then c is a non-quiescent fixed point of the CA. So the CA is not nilpotent.
$\Longleftarrow$ If T does not admit a valid tiling then every $n \times n$ square contains a tiling error, for some n. The state q propagates, so in at most $2n$ steps, all cells in the square are q. Then, the CA is nilpotent.
In both cases, the undecidability of the tiling problem implies the undecidability of nilpotency. □

4.1.2 North-West deterministic tiles

While tilings relate naturally to two-dimensional CA, one can strengthen Berger's result so that the nilpotency can be proved undecidable for one-dimensional CA as well. The basic idea is to view space-time diagrams of one-dimensional CA as tilings.

Definition 30. A tile set T is **NW-deterministic** if no two tile have identical colors on their top edges and on their left edges.

In a valid tiling of NW-deterministic tiles, the left and the top neighbor of a tile uniquely determine the tile. Consequently, NE-to-SW diagonals uniquely determine the next diagonal below them.

NW-d tiles as evolution of a 1d-CA

One could interpret a diagonal as a configuration of a one-dimensional CA, therefore a valid tiling represents a space-time diagram. More precisely, for any given NW-deterministic tile set T we construct a 1d-CA as follows:

- State set $S = T \cup \{q\}$, where q is a new symbol $q \notin T$
- Neighborhood is $(0, 1)$, in other words, radius $\frac{1}{2}$, so a cell only sees the one on its right.
- The local rule $f : S^2 \to S$ is defined as follows:

– $f(A, B) = C$ if the right color of A matches the left color of C and the bottom color of B matches the top color of C.

– $F(A, B) = q$ otherwise.

Proposition 15. *The CA we constructed is nilpotent if and only if T does not admit a tiling.*

Proof. $\Longrightarrow$ If T admits a tiling c then diagonals of c are configuration that never evolve into the quiescent configuration. So the CA is not nilpotent.
$\Longleftarrow$ If T does not admit a valid tiling then every $n \times n$ square contains a tiling error, for some n. Hence, state q is created inside every segment of length n. Since q starts spreading once it has been created, the whole configuration eventually becomes quiescent. □

Then, we can strengthen Berger's theorem

Theorem 15. *The tiling problem is undecidable among NW-deterministic tile sets.*

Which implies

Theorem 16. *It is undecidable whether a given one-dimensional CA (with spreading state q) is nilpotent.*

4.2 Snakes

Snakes is a tile set with some interesting (and useful) properties. In addition to the usual colored edges of regular tiles, snake tiles also have an arrow printed on them; these arrows point to an edge of the tile. Given a configuration (valid or invalid tiling) and a starting position, the arrows specify a path on the plane. Each position is followed by the neighboring position indicated by the arrow on the tile. In these paths, we can have loops or the path may be infinite and never return to a tile visited before.

Proposition 16. *In any configuration and on any path that follows the arrows, one of the following conditions is met:*

1. *Either there is a tiling error between two tiles, both of which are on the path;*

2. *Or the path is a plane-filling path, that is, for every positive integer n there is an $n \times n$ square all of whose positions are visited by the path.*

Remark. Note that the tiling may be invalid outside the path, yet the path is forced to "snake" through larger and larger squares.

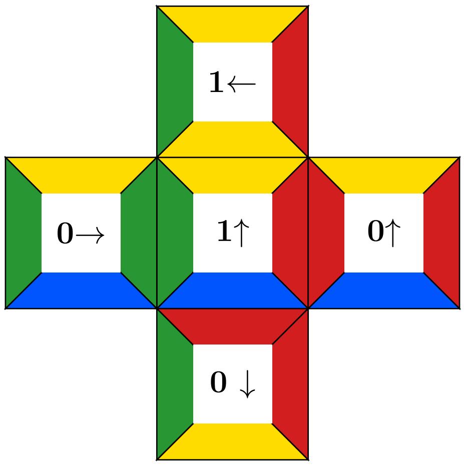
Snakes also has the property that it admits a valid tiling. The paths that Snakes forces when no tiling error is encountered have the shape of the well known plane-filling Hilbert-curve

4.2.1 Applications of snakes

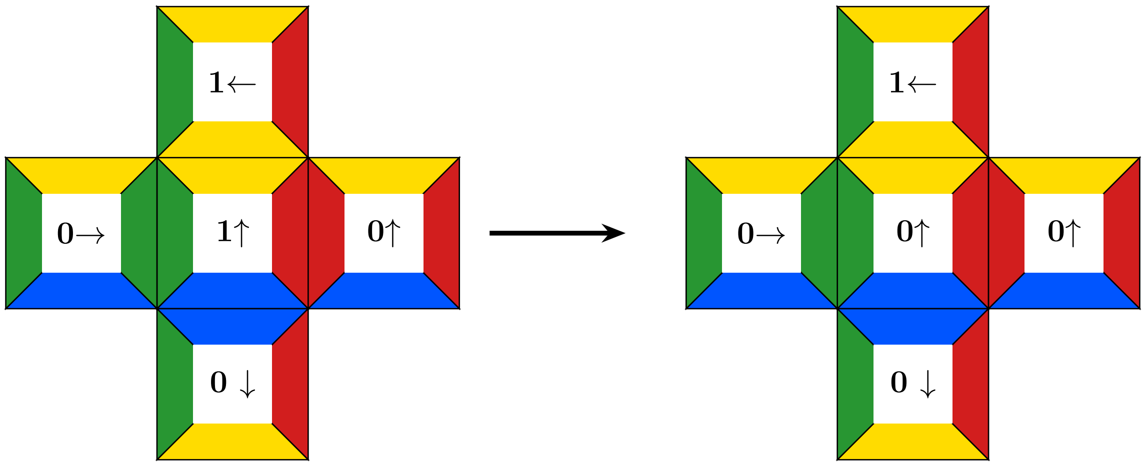
Snake XOR

First, we will see that in 2d-CA injectivity on periodic configurations does not imply injectivity. To prove this, we will show a reduction from snakes. The state set of the CA is $S = Snakes \times \{0, 1\}$, to each snake tile we attach a bit. The local rule checks whether the tiling is valid at the cell level.

- If there is an error, there is no change in the state.



- If the tiling is valid, the cell is **active**, the bits of the neighbor next on the path is XOR'ed to the bit of the cell.



Proposition 17. *Snake XOR is not injective*

Proof. We now show that there are two configurations that have the same image. Since the local rule only modifies the bits on the tiles, in order to have the same image, the two configurations must have the same tiles that form a valid tiling. In one of the configurations all bits are set to 0, and in the other configuration, all bits are set to 1. All cells are active because the tilings are correct. This means that all bits in both configurations become 0. So the two configurations become identical, therefore the CA is not injective. □

Proposition 18. *Snake XOR is injective on periodic configurations*

Proof. Suppose there are different periodic configurations c and d with the same successor. Since only bits may change, c and d must have identical Snakes tiles everywhere. So they must have different bits 0 and 1 in some position $\vec{p_1} \in \mathbb{Z}^2$. Because c and d have identical successors:

- The cell in position $\vec{p_1}$ must be active, that is, the Snakes tiling is valid in position $\vec{p_1}$.
- The bits stored in the next position $\vec{p_2}$. (indicated by the direction) are different in c and d

Hence, we can repeat the reasoning in position $\vec{p_2}$. The same reasoning can be repeated for all positions that form a path that follows the arrows on the tiles. There is no tiling error at any tile on this path. But this contradicts the fact that the plane filling property of snakes guarantees that on periodic configurations every path encounters a tiling error. □

Theorem 17. *In 2d-CA injectivity is undecidable.*

Theorem 18. *In 2d-CA surjectivity is undecidable.*

Theorem 19. *In 2d-CA injectivity is semi-decidable. Enumerate all CA G one by one one and check if G is the inverse of the given CA. Halt once (if ever) the inverse is found.*

Theorem 20. *In 2d-CA non-surjectivity is semi-decidable. Enumerate all finite patterns one by one and halt once (if ever) an orphan is found.*

Undecidability of reversibility

The second application of snakes is to show that it is undecidable to determine if a given two-dimensional CA is reversible. To prove this, we show a reduction from the tiling problem, using snake tiles.

Theorem 21. *It is undecidable to determine if a given 2d-CA is reversible.*

Proof. For any given tile set T we construct a CA with the state set $S = T \times Snakes \times \{0, 1\}$ The local rule is analogous to Snake XOR, with the difference that the correctness of the tiling is checked in both tile layers:

- If there is a tiling error, then the cell is inactive.
- If both tilings are valid, the bit of the neighbor next on the path is XOR'ed to the bit of the cell.

We can reason exactly as with Snake XOR, and show that the CA is reversible if and only if the tile set T does not admit a plane tiling:
$\Longrightarrow$) If a valid tiling of the plane exists, then we can construct two different configurations of the CA that have the same image under G. The Snakes and the T layers of the configurations form the same valid tilings of the plane. In one of the configurations, all bits are 0, and in the other configuration, all bits are 1. All cells are active because the tilings are correct. This means that all bits in both configurations become 0. So the two configurations become identical. The CA is not injective.
$\Longleftarrow$) Conversely, assume that the CA is not injective. Let c and d be two different configurations with the same successor. Hence, T admits tilings of arbitrarily large squares, and consequently a tiling of the infinite plane. Since only bits may change, c and d must have identical Snakes and T layers. So they must have different bits 0 and 1 in some position $\vec{p_1} \in \mathbb{Z}^2$. Because c and d have identical successors:

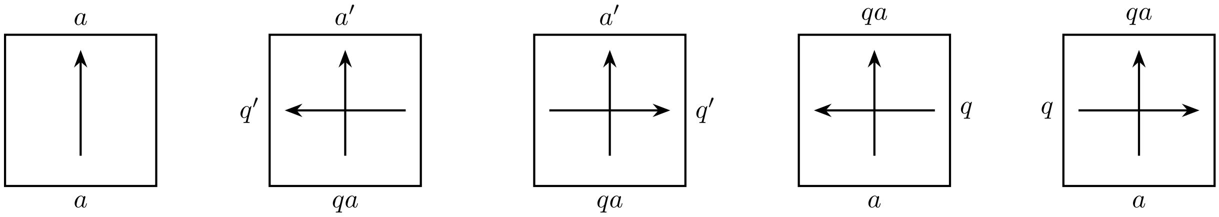
- The cell in position $\vec{p_1}$ must be active, that is, the Snakes and T tilings are both valid in position $\vec{p_1}$.
- The bits stored in the next position $\vec{p_2}$ (indicated by the direction) are different in c and d.

Hence, we can repeat the reasoning in position $\vec{p_2}$. The same reasoning can be repeated over and over again. The positions form a path that follows the arrows on the tiles. There is no tiling error at any tile on this path, so the special property of Snakes forces the path to cover arbitrarily large squares. Hence, T admits tilings of arbitrarily large squares, and consequently a tiling of the infinite plane. □

4.3 Tiling and Turing machines

First, we need to define what a Turing machine is. A Turing machine is a tuple $(Q, q_0, q_a, q_r, \Gamma, \Sigma, b, \delta)$, where

- Q is the set state.
- $q_0, q_a, q_r \in Q$ are respectively the start, accepting, and failure states. In particular, $q_a \neq q_r$.
- Γ is the finite tape alphabet.
- $\Sigma \subset \Gamma$ is the finite input alphabet.
- $b \in \Gamma \setminus \Sigma$
- $\delta : Q \times \Gamma \to Q \times \Gamma \times \{-1, 1\}$. $\forall \gamma \in \Gamma$ we must have $\delta(q_a, \gamma) = (q_a, \gamma, 1), \delta(q_r, \gamma) = (q_r, \gamma, 1)$.



(a) (b) (c) (d)

Figure 4.1: Machine tiles associated with a Turing machine.

To prove undecidability results for cellular automata, it is convenient to take an intermediate step and prove that certain questions concerning Wang tilings are undecidable. The proofs are based on the fact that valid tilings can be forced to contain a complete simulation of a given Turing machine. Given any Turing machine M, we associate to it a tile set shown in figure 4.1, which we call the *machine tiles* of M. Note that in the illustrations, instead of colors, we use labeled arrows on the sides of the tiles. Two adjacent tiles match if and only if an arrowhead meets an arrow tail with the same label. Such an arrow representation can be converted into the usual coloring representation of Wang tiles by assigning to each arrow direction and labeling a unique color. The machine tiles of M contain the following tiles:

- For every tape letter $a \in \Gamma$ a *tape tile* of figure 4.1 (a).
- For every tape letter $a \in \Gamma$ and non-halting state $q \in Q \setminus \{q_a, q_r\}$ and *action tile* of figure 4.1 (b) or (c).
  - Tile (b) is used if $\delta(q, a) = (q', a', -1)$
  - and tile (c) is used if $\delta(q, a) = (q', a', +1)$.
- For every tape letter $a \in \Gamma$ and non-halting state $q \in Q \setminus \{q_a, q_r\}$ two *merging tiles* shown in figure 4.1 (d).

The main idea is that a configuration of the Turing machine M is represented as a row of tiles in such a way that the cell currently scanned by M is represented by an action tile, its neighbor where the machine moves into has a merging tile, and all other tiles on the row are tape tiles. If this is a row of a valid tiling, then it is clear that the rows above must be similar representations of subsequent configurations in the Turing machine computation, until the machine halts.

Chapter 5

Subshifts

In this chapter, we present subshifts as a concept and their applications in coding theory and other fields.

Definition 31. Let A be a finite alphabet, let $x \in A^{\mathbb{Z}}$ and $w \in A^{+}$. We write $w \prec x$ (w is contained in x) if and only if $\exists i, j \in \mathbb{Z}$ such that $x_{[i,j]} = w$.

Let $\mathcal{F} \subseteq A^*$, then we define the set $X_{\mathcal{F}} = \{x \in A^{\mathbb{Z}} | \forall w \in \mathcal{F}, w \not\prec x\}$. Please note that $\mathcal{F}$ stands for *forbidden*, in the sense that $X_{\mathcal{F}}$ is generated from a set of forbidden words.

Definition 32 (Subshift)**.** $X \subseteq A^{\mathbb{Z}}$ is a **subshift** $\iff$ $\exists \mathcal{F} \subseteq A^*$ such that $X = X_{\mathcal{F}}$.

We now show some examples:

1. Let $X = A^{\mathbb{Z}}$, is X a subshift? Yes, because $X = X_{\mathcal{F}}$ with $\mathcal{F} = \emptyset$.

2. Let $A = \{0, 1\}$ and $X \subseteq A^*$ be the set of all bi-infinite binary sequences in which 2 adjacent 1's never appear. X is subshift, because $X = X_{\mathcal{F}}$ with $\mathcal{F} = \{11\}$.

   Remark. We also have that

   - $X = X_{\mathcal{F}'}$ with $\mathcal{F}' = \{11, 111\}$.
   - $X = X_{\mathcal{F}''}$ with $\mathcal{F}'' = \{11, 111, 1111, \dots\}$ which is an infinite set.
   - $\vdots$
   - $X = X_{\mathcal{F}'''}$ with $\mathcal{F}''' = \{110, 111, 011, 111\}$

   We call the subshift $X_{\mathcal{F}}$ the **golden mean shift**.

3. Let $X \subseteq A^{\mathbb{Z}}$ be the set of all bi-infinite binary words in which between two consecutive occurrences of 1, there is always an even number of 0's. X is a subshift, $X = X_{\mathcal{F}}$ with $\mathcal{F} = \{101, 10001, 1000001, \dots\} = \{10^{2n+1}1 | n \in \mathbb{N}\}$

Definition 33 (Language)**.** Let $X \subseteq A^{\mathbb{Z}}$, not necessarily a subshift. Let $\mathcal{B}_n(X) = \{w \in A^n | \exists x \in X, w \prec x\}$. The **language of** X is the set $\mathcal{B}(x) = \cup_{n \in \mathbb{N}} \mathcal{B}_n(X) = \{w \in A^* | \exists x \in X, w \prec x\}$

Definition 34 (Factorial and extendable language)**.** Let $\mathcal{L} \subseteq A^*$ be a language:

- $\mathcal{L}$ is factorial if $\forall u \in \mathcal{L}, \forall w \prec x \implies w \in \mathcal{L}$
- $\mathcal{L}$ is extendable if $\forall u \in \mathcal{L}, \exists v, w \in \mathcal{L}$ such that $vuw \in \mathcal{L}$.

Theorem 22. *Let $X \subseteq A^{\mathbb{Z}}$, X is subshift $\iff$ $\mathcal{B}(x)$ is factorial and extendable.*

5.1 Subshifts of finite type

Definition 35. Let X be a subshift. X is a **subshift of finite type** (SFT) if $\exists \mathcal{F} \subseteq A^*$ with $|\mathcal{F}| < \infty$ such that $X = X_{\mathcal{F}}$.

5.1.1 Memory

Definition 36. X is a SFT of memory M if $\exists \mathcal{F} \subseteq A^{M+1}$ such that $X = X_{\mathcal{F}}$.

Let X be an SFT of memory M. How can we decide if an $x \in A^{\mathcal{Z}}$ is an element of the subshift X?

Proposition 19. $x \in X \iff \forall i, x_{[i,i+M]} \in \mathcal{B}_{M+1}$

Definition 37. The block representation of size $N \in \mathbb{N}, N \geq 1$ is the following way to represent words in $A^{\mathbb{Z}}$. Consider the mab $\beta_N : A^{\mathbb{Z}} \to (A^N)^{\mathbb{Z}}$, then $\forall x \in A^{\mathbb{Z}}, x = (\dots, x_{-1}, x_0, x_1, \dots, x_{N-1}, X_N, \dots)$, we have that $\beta_N(x) = \left(\dots, \begin{pmatrix} x_{-1} \\ x_0 \\ x_1 \\ \vdots \\ x_{N-2} \end{pmatrix}, \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{N-1} \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}, \dots\right)$ $\forall x \in A^{\mathbb{Z}}, \forall i \in \mathbb{Z} \quad \beta_N(x)_i = \begin{pmatrix} x_i \\ \vdots \\ x_{i+N-1} \end{pmatrix}$ We can also use block representations for a subshift X. Let $N \geq 1$, then $X^{[N]} = \beta_N(X)$

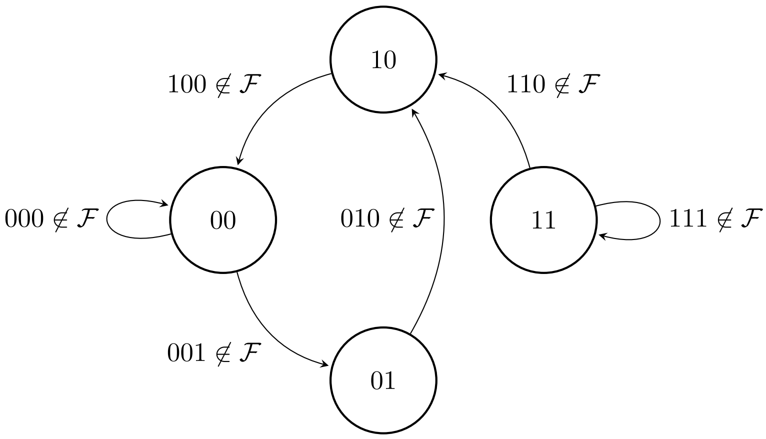
Proposition 20. $X^{[N]}$ *also is a subshift.*

Given an SFT X of memory M we can build a graph that describes it; this is useful because it will be later used to discuss some properties. We build a graph $G = (V, E)$, where

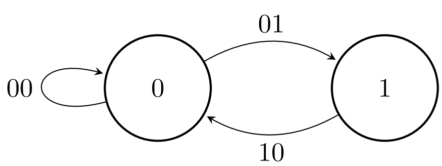
- $V = \mathcal{B}_M(X) \subseteq A^M$
- $\forall u, v \in V, (u, v) \in E$ if and only if :
  - $u = u_1 u_2 \dots u_M$
  - $v = v_1 \dots v_{M-1} v_M$
  - $u_2 \dots u_M = v_1 \dots v_{M-1}$
  - $u v_M = u_1 v \notin \mathcal{F}$

We now show some examples

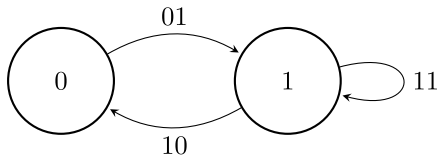
1. $\mathcal{F} = \{101, 011\} \subseteq \{0,1\}^{2+1}$, X is an SFT of memory $M = 2$. Then we build $G = (V, E) \quad V = \mathcal{B}_2(x) = \{00, 01, 10, 11\}$



2. $\mathcal{F} = \{11\} \quad X_{\mathcal{F}}$ is the golden mean shift



3. $\mathcal{F} = \{00\} \quad X_{\mathcal{F}} = X(0,1)$



We can also do the opposite, given a graph $G = (V, E)$ one can build a subshift X_G (which is of finite type) as follows $X_G = \{x \in E^{\mathbb{Z}} | \forall u \in \mathbb{Z}, t(x_i) = i(x_{i+1})\}$. In this case, $A = E$ we use $t(\alpha)$ to denote the terminal vertex of the edge α, and $i(\alpha)$ its initial vertex. With this definition, a word of the subshift is a bi-infinite path on the graph, and it is made by the edges themselves, not the labels of the edges.

5.2 Sofic shifts

Let $G = (V, E)$ be a graph of a subshift X, then we can define $\mathcal{G} = (V, E, l)$, where $l : E \to A$.

Definition 38. $l_\infty : E^{\mathbb{Z}} \to A^{\mathbb{Z}}$, $\forall x \in E^{\mathbb{Z}}, \forall i \in \mathbb{Z} \quad l_\infty(x)_i = l(x_i)$

Definition 39. $X_{\mathcal{G}} = \{y \in A^{\mathbb{Z}} | \exists x \in E^{\mathbb{Z}} : \forall i \in \mathbb{Z} \quad t(x_i) = i(x_{i+1}) \wedge y_i = l(x_i)\}$ is the subshift induced by the graph.

Remark. $y \in X_{\mathcal{G}}$ if and only if it is a bi-infinite path on the labels of the graph $\mathcal{G}$.

Proposition 21. *$X_{\mathcal{G}}$ is a subshift.*

Proof. $X_{\mathcal{G}} = X_{\mathcal{F}}$ with $\mathcal{F} = \mathcal{B}(X_{\mathcal{G}})^C$ is obvious from the previous statement. □

Definition 40. X is a sofic subshift if and only if $\exists \mathcal{G}$ such that $X = X_{\mathcal{G}}$. We call $\mathcal{G}$ the rightmost resolving representation of X if $\mathcal{G}$ is a transition graph of a deterministic finite state automata.

Now we discuss an important property, <u>are SFT sofic?</u> Yes! $X \to G \to (G, L) = \mathcal{G}$, let $e = (u, v)$ be an edge on the graph and $l(e) = u_1$. Then, $x \in X \iff \beta_{M+1}(x) \in X^{M+1} = X_G \iff l_\infty(\beta_{M+1}) \in X_{\mathcal{G}}$.

5.3 Entropy

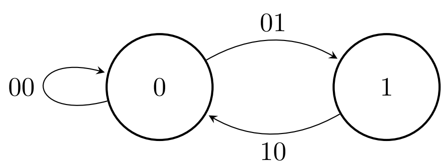
Let X be a subshift, how does $|\mathcal{B}_n(X)|$ grow with n? This measurement gives us a measure of the complexity/richness of X, in other words, a growing rate. Let's start with an example, let $X = A^{\mathbb{Z}}$ be the total shift. Then, $|\mathcal{B}_n(X)| = |A^n| = |A|^n = 2^{\log_2 |A| n}$.
In general, for any subshift X we have $|\mathcal{B}_n(X)| \approx 2^{c \cdot n}$, then $c \approx \frac{\log_2 |\mathcal{B}_n(X)|}{n}$

Definition 41 (Entropy of a subshift X)**.** $h(X) = \lim_{n \to +\infty} \frac{\log_2 |\mathcal{B}_n(X)|}{n}$. Note that we should prove that this limit exists, but we will not provide the proof for the sake of simplicity.

Remark. It is always true that $h(X) \leq \log_2 |A|$ since the full shift is intuitively the shift with maximum entropy.

Here we show an example of how to calculate the entropy. Let $X = X_{\mathcal{F}}$ with $\mathcal{F} = \{11\}$

with G :



which has adjacency matrix $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$

Remember that $M = 1$. We know that $X_G = X^{[2]}$ and $\hat{X}_G = X^{[1]} = X$ where $\hat{X}_G$ is the subshift constructed by paths on the vertices.
Moreover, $\mathcal{B}_n(X^{[M]}) \overset{bijection}{\longleftrightarrow} \mathcal{B}_{n-1}(X^{[M+1]})$. Then $|\mathcal{B}_n(X^{[1]})| = |\mathcal{B}_n(X)| = |\mathcal{B}_{n-1}(X^{[2]})| = |\mathcal{B}_{n-1}(X_G)|$.
But then, the limit (therefore the entropy) is the same. One fundamental problem is calculating $|\mathcal{B}_m(X_G)|$ with any m.

Note that any element of $\mathcal{B}_m(X_G)$ is a word of length m obtained from a path of length m on the edges of G. Then:

$$|\mathcal{B}_m(X_G)| = \text{"the number of paths of length m on any pair of vertices of G"} = \\ = \sum_{i\in V}\sum_{j\in V} a_{i,j}^{(m)}$$

Where $a_{i,j}^{(m)}$ is an element of A^m. Since matrix multiplication is a heavy task, how can we do it efficiently? We will try to diagonalize the adjacency matrix A, so that $\det(A - tId) = 0$ and $t^2 - t - 1 = 0$.

1. Calculate the eigen values of A: $\lambda = \frac{1+\sqrt{5}}{2}$ $\mu = \frac{1-\sqrt{5}}{2}$
2. Compute $A = P \cdot \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \cdot P^{-1}$ where P is a "passage matrix" that has as columns the eigen vectors of λ and μ respectively. $P = \begin{pmatrix} \lambda & \mu \\ 1 & 1 \end{pmatrix}$ and $P^{-1} = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & -\mu \\ -1 & \lambda \end{pmatrix}$.

Therefore, $A^m = P \cdot \begin{pmatrix} \lambda^m & 0 \\ 0 & \mu^m \end{pmatrix} \cdot P^{-1} = \begin{pmatrix} f_{m+1} & f_m \\ f_m & f_{m-1} \end{pmatrix}$ with $f_m = \frac{1}{\sqrt{5}}(\lambda^m - \mu^m)$ and $f_{m+2} = f_{m+1} + f_m$ (Fibonacci!). Finally,

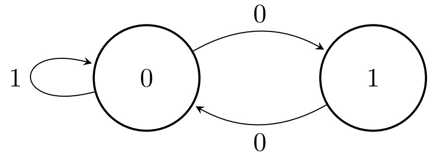
$$\begin{aligned} |\mathcal{B}_m(X_G)| =& \sum_{i\in V}\sum_{j\in V} a_{i,j}^{(m)} = \\ =& f_{m+1} + f_m + f_{m-1} = \\ =& f_{m+2} + f_{m+1} = f_{m+3} \end{aligned}$$

$$|\mathcal{B}_{n-1}(X_G)| = f_{n+2} = \frac{1}{\sqrt{5}}(\lambda_{n+2} - \mu^{n+2}$$

We can now proceed to calculate the entropy of X,

$$\begin{aligned} h(x) &= \lim_{n\to+\infty} \frac{\log_2 \frac{1}{\sqrt{5}}(\lambda^{n+2} - \mu^{n+2}}{n} \\ &= \lim_{n\to+\infty} \frac{\log_2 \frac{1}{\sqrt{5}}\lambda^{n+2}}{n} \\ &= \lim_{n\to+\infty} \frac{\log_2 \frac{1}{\sqrt{5}}\lambda^{n+2}}{n} + \frac{\log_2 \lambda^n}{n} \\ &= 0 + \log_2 \lambda = \log_2 \lambda \quad \text{with } \lambda = \frac{1+\sqrt{5}}{2} \end{aligned}$$

We now show that the same subshift can be obtained from another graph, that is the rightmost representation of X.



This graph $\mathcal{G} = (G, l)$ is isomorphic to G, and we have that $X = X_{\mathcal{G}}$. For any word of $\mathcal{B}(X)$ with a 1 inside, there exists a unique path that identifies it. The same doesn't hold for 2^n, where two different paths exist. Therefore, $|\mathcal{B}_n(X)| = |\mathcal{B}_n(X_G)| - 1$, then $h(X) = h(X_{\mathcal{F}}) = \log_2 \lambda$ with $\mathcal{F} = \{\infty\infty\}$

Proposition 22. *In general, if $\mathcal{G} = (G, l)$ is the right resolving representation of the same subshift, then $h(X_G) = h(X_{\mathcal{G}})$*

5.3.1 Eigenvalues and entropy

Definition 42. A matrix A with all non-negative entries is irreducible if $\forall i, j, \exists n$ such that $(A^n)_{i,j} \geq 0$.

Proposition 23. *Let G be a graph and A its adjacency matrix, then G is strongly connected $\iff$ A is irreducible.*

Theorem 23 (Perron-Frobenius)**.** *Let $A \neq 0$ be an irreducible matrix, then*

- *$\exists v_a$ eigenvector with corresponding eigenvalue $\lambda_a > 0$ that is simple (with corresponding algebraic multiplicity 1).*
- *$\forall \mu$ eigenvalue of A, $|\mu| \leq \lambda_a$ (λ_a is the maximum eigenvalue).*
- *$\forall v' > 0$ eigenvector of A, v' is a positive multiple of v_a.*

Corollary 11. *Let G be a graph and A its adjacency matrix.*

- *G is strongly connected $\implies h(X_G) = \log_2(\lambda_a)$*
- *X SFT with associated graph G which is strongly connected $\implies h(X) = \log_2(\lambda_a)$*
- *X sophic subshift such that $X = X_{\mathcal{G}}$ with $\mathcal{G} = (X, G)$ right resolving $\implies h(X) = \log_2 \lambda_a$*

To compute the entropy of a subshift that is

- *An SFT with associated graph G strongly connected*
- *A sophic X ...*

It is sufficient to calculate the eigenvalue of maximum absolute value $\lambda_a > 0$, then $h(X) = \log_2 \lambda_a$.

5.4 Coding on physical media

In this section we will see how subshifts can be used to define encodings for physical media. This comes from the necessity to not allow certain sequences of bits, since they would render the memorization of data unreliable or more difficult.

Definition 43 (Road coloring)**.** The road coloring of a graph $G = (V, E)$ is $\mathcal{C} : E \to A$ such that $\forall i \in V, |\{(i, j) \in E | j \in V\} = |A|$

Remark. $\forall w \in A^*, \forall i \in V$ there is a single path.

Definition 44. $\mathcal{G} = (V, E, l)$ with $l : E \to A$, l is right-closing with delay D if $\forall \pi, \pi'$ $\pi = e_1 \ldots e_d e_{d+1}, \pi' = e'_1 \ldots e'_d e'_{d+1}$ that start from the same node, $l(\pi) = l(\pi') \implies e_1 = e'_1$

Remark. If $\mathcal{G} = (V, E, l)$ is a graph with $l : E \to A$ right closing, then $\forall I \in V, \forall l, l' \in E^{\mathbb{N}}$ mono-inifnite paths starting from the node I, $l_\infty(l) = l_\infty(l') \implies l = e'$, i.e. for all $I \in V$, l_∞ restricted to the paths that start from I is injective.

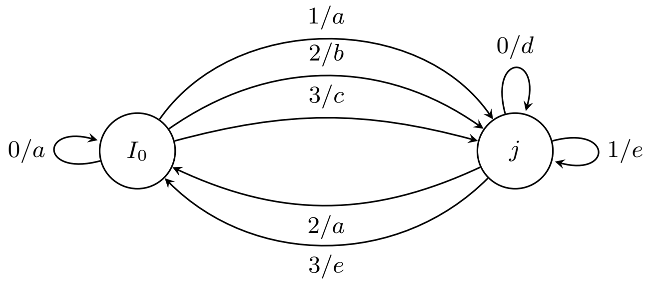
Definition 45 (Finite state code)**.** $(G, \mathcal{I}, \mathcal{O})$, where:

- $\mathcal{I}$ is a road coloring of G (input labeling)
- $\mathcal{O}$ is a right closing labeling of G (output labeling)

We can also give another definition

Definition 46 (Finite state code)**.** (X, n) where X is a subshift is a finite state code in which G has out-degree constant $= n$, and X contains $\mathcal{O}_\infty(X_G) = \{y | \exists x \in X_G : \forall i, y_i = \mathcal{O}(x_i)\}$. Note that $\mathcal{O}_\infty(X_G)$ is sophic by definition.

A finite state code (X, n) can be used to transform sequences of $A^{\mathbb{N}}$ (with $|A| = n$) in sequences of X. Here we show an example, let $n = 4, A = \{1, 2, 3, 4\}$

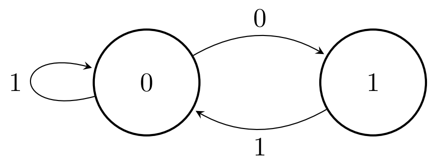


Let $x \in A^{\mathbb{N}}$. There is a unique path $e \in E^{\mathbb{N}}$ that starts from I_0 such that $\mathcal{I}_\infty(e) = x$. For example, let $x = 0231002\ldots$, this sequence is encoded in $\mathcal{O}_\infty^{I_0}(e) = y$, therefore $y = abeaddd\ldots$. Since $\mathcal{O}_\infty^{I_0}$ is injective, given a specific $y \in \mathcal{O}_\infty^{I_0}(E^{\mathbb{N}})$ it is possible to uniquely reconstruct $e \in E^{\mathbb{N}}$, therefore $x_\infty^{I_0}$.

Theorem 24 (Finite state coding)**.** *Let X be a sophic subshift and $n \geq 1$. Then, there exists a finite state coding $(X, n) \iff h(X) \geq \log n$.*

5.4.1 Problems and solutions

What if this condition doesn't hold? Here we show an example that is linked to the problem of writing data on a physical media. We want to transform a sequence $\{0,1\}^{\mathbb{N}}$ in mono-infinite sequences of $X(0,1)$ or $X(1,3)$. This is equivalent to asking if there exists a finite state coding $(X(0,1), 2)$ or $(X(1,3), 2)$. In order to do this we have to compute $h(X(0,1))$



We give the right resolving representation of $X(0,1)$ as sophic. The adjacency matrix is $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ it's maximum eigenvalue is $\lambda_a = \frac{1+\sqrt{5}}{2}$, so $h(X(0,1)) = \log_2 \frac{1+\sqrt{5}}{2} \overset{?}{\geq} \log_2 2$. The inequality doesn't hold. We know this a-priori by knowing that the maximum entropy of $\{0,1\}^{\mathbb{N}}$ is $\log_2 2$.

Remark. We will always have $h(X) < \log_2 n$, therefore it would seem impossible to find a finite state code to encode and decode sequences from $\{0,1\}^{\mathbb{N}}$ into $X(0,1)$ or $X(1,3)$.

The main idea is to divide the input in blocks of length p and the output in blocks of length q.
$x = (\boxed{x_0x_1x_2\ldots x_{p-1}}\boxed{x_px_{p+1}\ldots x_{2p-1}}\ldots)$
$\gamma_p(x) = (\begin{pmatrix} x_0 \\ \vdots \\ x_{p-1} \end{pmatrix}, \begin{pmatrix} x_p \\ \vdots \\ x_{2p-1} \end{pmatrix}, \ldots)$ remark: this is not the block representation of p! In general:

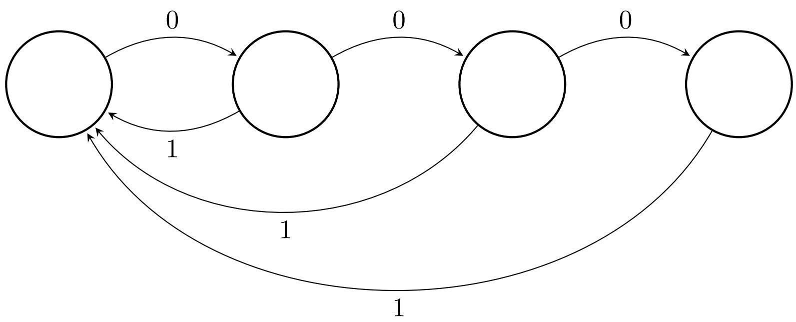
Definition 47 (Power shift)**.** Given $N \geq 1$
$\gamma_N : A^{\mathbb{Z}} \to (A^N)^{\mathbb{Z}}$ (equivalently $A^{\mathbb{N}} \to (A^N)^{\mathbb{N}}$), where $\forall x \in A^{\mathbb{Z}}, \forall i \in \mathbb{Z} (\in \mathbb{N})$ $\gamma_N(x)_i = x_{[iN, iN+i-1]}$
Remember, this is not the block representation of N.

Here we show an example, $N = 4$
$x = (\ldots x_0x_1x_2x_3x_4x_5x_6x_7x_8x_9x_{10}\ldots)$
$\gamma_4(x) = (\ldots \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \begin{pmatrix} x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} \begin{pmatrix} x_8 \\ x_9 \\ x_{10} \\ x_{11} \end{pmatrix} \cdots)$

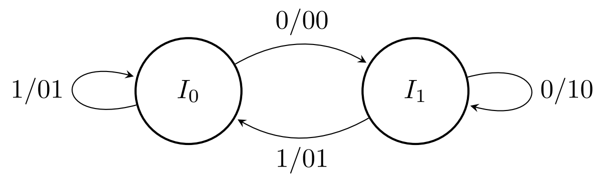
Remark. γ_N can be applied to elements of a subshift. Therefore, given a subshift $X \subseteq A^{\mathbb{Z}}(A^{\mathbb{N}})$ you can consider $\gamma_N(X)$, which we denote with X^N which is called the power shift of X.

Proposition 24. *If* $X \subseteq A^{\mathbb{Z}}(A^{\mathbb{N}})$ *is a subshift, then* $X^N = \gamma_N(X) = \{\gamma_N(x)|x \in X\} \subseteq (A^N)^{\mathbb{Z}} \quad ((A^N)^{\mathbb{N}})$ *is a subshift. Moreover,* $h(X^N) = N \cdot h(X)$.

Getting back to the problem of data memorization, we won't have sequences of $\{0,1\}^{\mathbb{N}}$ to encode into a subshift X (like $X(0,1)$ or $X(1,3)$), but we will have sequences $\gamma_p(\{0,1\}^{\mathbb{Z}})$ to encode in sequences of $\gamma_q(X) = X^q$, with q, p to be chosen. We should choose the value of p and q so that there exists a finite state code $(\gamma_q(X), n)$ where $n = 2^p$, since the sequences to be encoded are in $\gamma_p(\{0,1\}^{\mathbb{N}}) = (\{0,1\}^p)^{\mathbb{N}}$, in other words they are in the full shift of the alphabet $\{0,1\}^p$ instead of $\{0.1\}$. Therefore, $n = |\{0,1\}^p| = 2^p$. In order to exists such finite state code $(\gamma_q(X), n = 2^p)$ that encodes the sequences from $(\{0,1\}^P)^{\mathbb{N}}$ into $X^q = \gamma_q(x)$ we must have $h(X^q) \geq \log_2 n$, that is $q \cdot h(X) \geq \log_2(2^P)$, or $h(x) \geq \frac{p}{q}$.
Here we show an example. Let $X = X(1,3)$



with adjacency matrix $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$ which has maximum eigenvalue $\lambda_a \approx 1,4655$, then $h(X) = \log_2 \lambda_a = 0.55$. We want to have $h(X) = 0.55 \geq \frac{p}{q}$, it is sufficient to pick $p = 1, q = 2$. We have that there exists a finite state code $(X^2, 2)$



Let x be a path that starts from I_0 $x = 1000011$, then $y = 01001010100101$, which is the encoding using $X(1,3)$.