Representing Sturmian words on cellular automata

Francesco Dolce

joint work with Pierre-Adrien Tahay

Journées SDA2

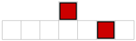
Liège, 14 juin 2022

Cellular automata

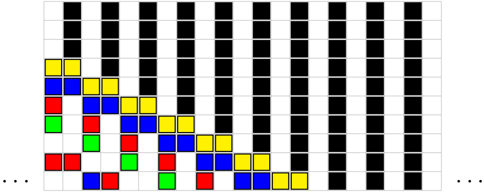
Definition

A (1-dimensional) *cellular automaton* is a dynamical system $(\mathcal{A}, T)$, where $T : \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$, is defined by a *local rule* $\tau : \mathcal{A}^{2r+1} \to \mathcal{A}$ (r is the *radius*).

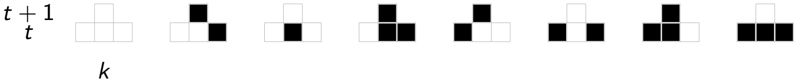
Elements of $\mathcal{A}^{\mathbb{Z}}$ are called *configurations*.
A configuration $\mathbf{x} = (x_k)_k$ is *finite* if $\{k \,:\, x_k \neq 0\}$ is finite.
The *space-time diagram* of a CA is a subset of $\mathcal{A}^{\mathbb{Z} \times \mathbb{N}}$.



$\tau : *****a* \mapsto a$

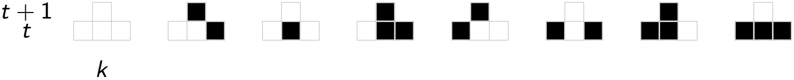


Elementary (1-dimensional) Cellular Automata

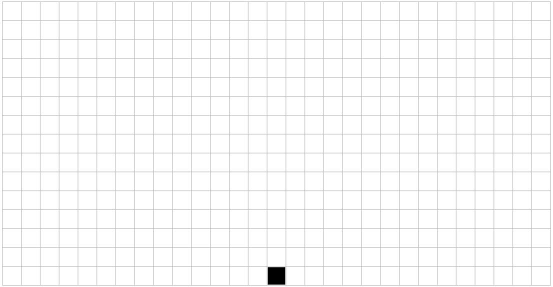


$$\langle k, t+1 \rangle = \langle k-1, t \rangle + \langle k+1, t \rangle \pmod 2$$

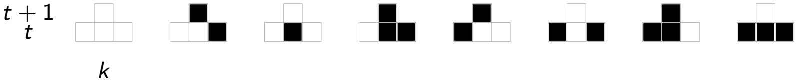
Elementary (1-dimensional) Cellular Automata



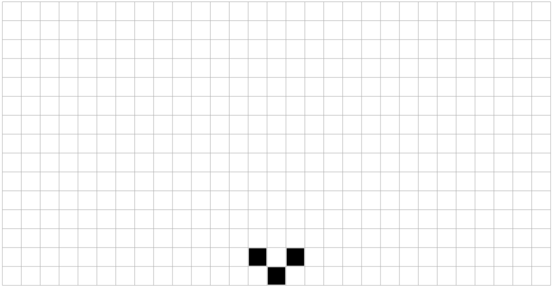
$$\langle k, t+1 \rangle = \langle k-1, t \rangle + \langle k+1, t \rangle \pmod 2$$



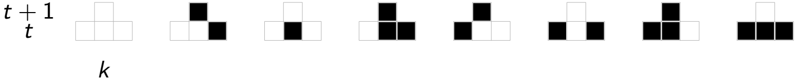
Elementary (1-dimensional) Cellular Automata



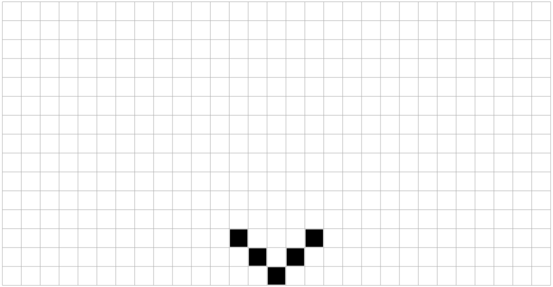
$$\langle k, t+1 \rangle = \langle k-1, t \rangle + \langle k+1, t \rangle \pmod 2$$



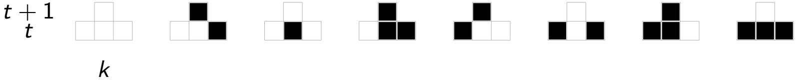
Elementary (1-dimensional) Cellular Automata



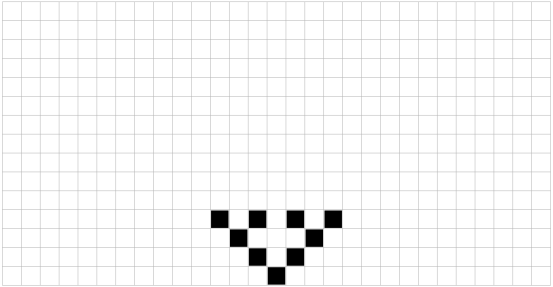
$$\langle k, t+1 \rangle = \langle k-1, t \rangle + \langle k+1, t \rangle \pmod 2$$



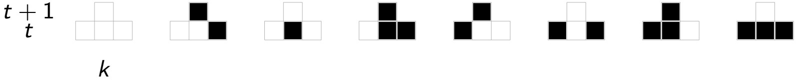
Elementary (1-dimensional) Cellular Automata



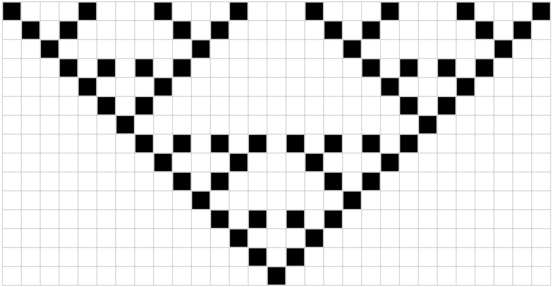
$$\langle k, t+1 \rangle = \langle k-1, t \rangle + \langle k+1, t \rangle \pmod 2$$



Elementary (1-dimensional) Cellular Automata



$$\langle k, t+1 \rangle = \langle k-1, t \rangle + \langle k+1, t \rangle \pmod 2$$



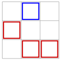
Cellular automata

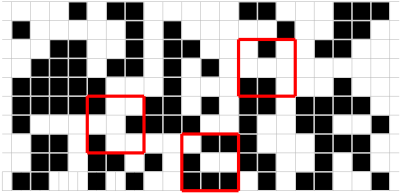
Linear CA with memory

Definition

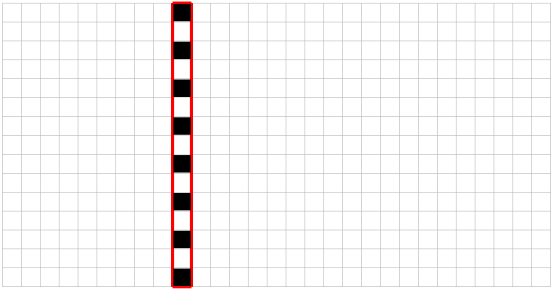
A *cellular automaton with memory* d is defined by $T : (\mathcal{A}^d)^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$.
When $\mathcal{A} = \{0, 1, \ldots, p-1\}$ and T is linear (i.e., the local rule can be written as $\tau((x_i)_{-r \le i \le r}) = \sum_{i=-r}^{r} \alpha_i x_i$) we call the automaton *linear*.

$$\langle t+2, k\rangle = \langle t+1, k-1\rangle + \langle t, k\rangle + \langle t, k+1\rangle \pmod 2$$

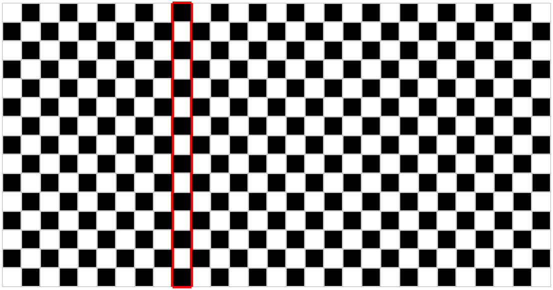




Column representation



Column representation



$\mathbf{x}_0$ 

τ :  ,  .

Column representation

Question :
Which infinite sequences can be represented *using a finite number of states* on a column of a cellular automaton ?

More formally : given

$$\mathcal{S} = \left\{ (T^n(\mathbf{x})_0)_{n\geq 0} \in \mathcal{A}^{\mathbb{N}} \ : \ T \text{ is a } 0\text{-quiescent CA on } \mathcal{A}^{\mathbb{Z}} \text{ and } \mathbf{x} \text{ is finite} \right\}$$

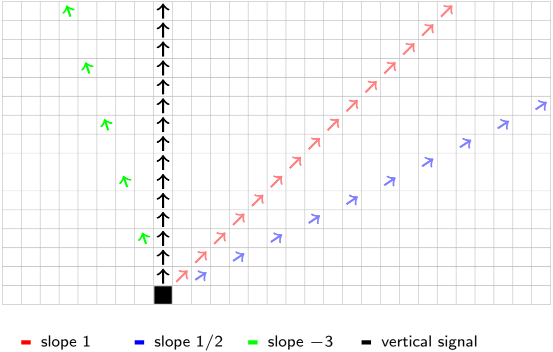
determine whether $\mathbf{w} \in \mathcal{S}$ (and if yes construct the CA).

A Cellular Automaton T is 0-*quiescent* if $T(0^{\mathbb{Z}}) = 0^{\mathbb{Z}} = \cdots 000 \cdots$.

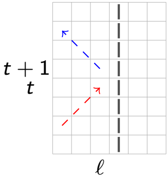
Signals

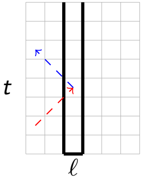
communication space-time diagram

The *slope* of a signal is the ratio $\frac{t}{\ell' - \ell}$.



Walls





Column representation

Theorem

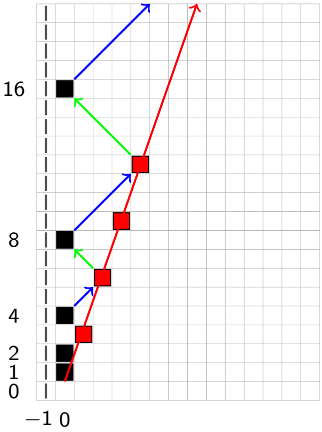
The characteristic function $\mathbb{1}_P$ of the set of prime numbers is in $\mathcal{S}$.

- P.C. Fisher (1965), 30.000+ states ;
- I. Korec (1997), 11 states.

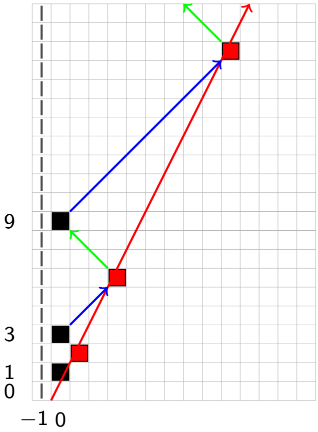
$$\mathbb{1}_P(x) = \begin{cases} 1 & \text{if } x \text{ is a prime,} \\ 0 & \text{otherwise.} \end{cases}$$

Representation of automatic sequences

Powers : 2^n and 3^n [Mazoyer, Terrier (1999)]



slope 3 = $\frac{2+1}{2-1}$ slope 1 slope −1



slope 2 = $\frac{3+1}{3-1}$ slope 1 slope −1

Representation of automatic sequences

Theorem [Litow, Dumas (1993)]

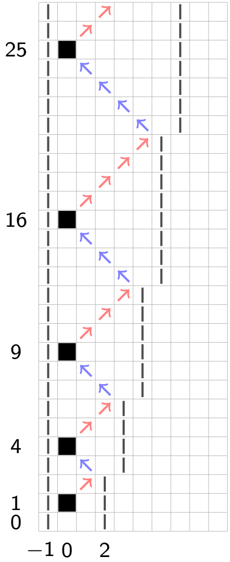
Let p be a prime number. Every column of a linear CA over $\{0, \dots, p-1\}$ (with finite initial configuration) is p-automatic.

Theorem [Rowland, Yassawi (2015)]

Let p be a prime number and $q = p^m$.
A sequence of elements in $\{0, \dots, q-1\}$ is p-automatic if and only if it is a column of a spacetime diagram of a linear cellular automaton with memory over $\{0, \dots, q-1\}$ whose initial conditions are eventually periodic in both directions.

Representation of non-automatic sequences

Squares [Poupet, Sablik, Theyssier (2011)]

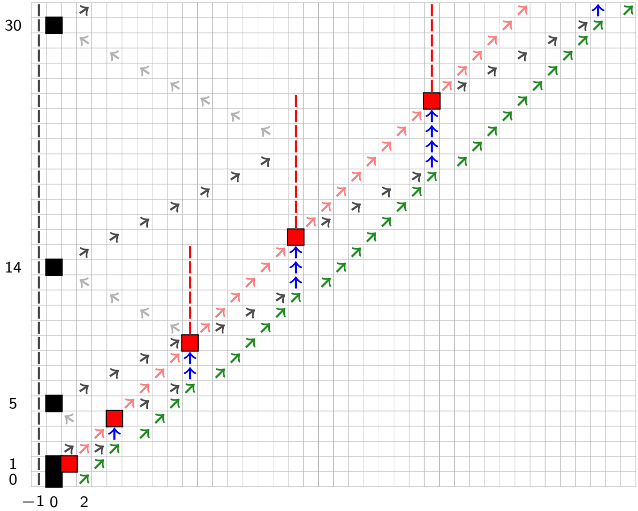


slope 1

slope −1

Representation of non-automatic sequences

Sum of squares [Marcovici, Stoll, Tahay (2018)]



Sturmian sequences

Definition

An infinite word $\mathbf{w}$ is *Sturmian* if it has exactly $n+1$ distinct factors of length n for every $n \geq 0$.
If both sequences $\mathtt{a}\mathbf{w}$ and $\mathtt{b}\mathbf{w}$ are Sturmian, we call $\mathbf{w}$ a *characteristic Sturmian word*.

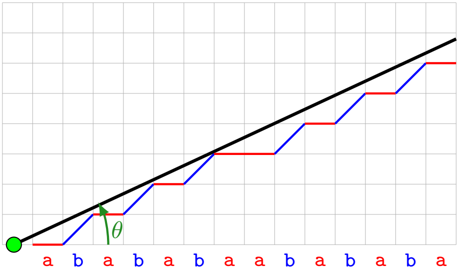
Example (Fibonacci)

$$\mathbf{f} = \mathtt{abaababaabaababa}\cdots$$

$$\mathcal{L}(\mathbf{f}) = \{\underbrace{\varepsilon}_{1}, \underbrace{\mathtt{a}, \mathtt{b}}_{2}, \underbrace{\mathtt{aa}, \mathtt{ab}, \mathtt{ba}}_{3}, \underbrace{\mathtt{aab}, \mathtt{aba}, \mathtt{baa}, \mathtt{bab}}_{4}, \underbrace{\mathtt{aaba}, \mathtt{abaa}, \mathtt{abab}, \mathtt{baab}, \mathtt{baba}}_{5}, \ldots\}$$

Sturmian sequences

Mechanical words



Remark : Characteristic Sturmian words correspond to *balanced mechanical words* with irrational intercept equal to the slope.

Let $\mathbf{c}_\alpha$ denote the unique characteristic Sturmian with slope (and intercept) α.

Continued fraction expansion

The *continued fraction expansion* of $\theta \in \mathbb{R}$ is defined as $[c_0, c_1, c_2, \ldots]$ whenever

$$\theta = c_0 + \cfrac{1}{c_1 + \cfrac{1}{c_2 + \ddots}}$$

with $c_0 \in \mathbb{Z}$ and $c_i \in \mathbb{Z}^+$ for every $i > 0$.

Theorem

If θ is a positive irrational, its continued fraction expansion is unique.

$$e = [2, 1, 2, 1, 1, 4, 1, 1, 6, \ldots], \qquad \pi = [3, 7, 15, 1, 292, 1, 1, 1, 2, \ldots]$$

Continued fraction expansion

The *continued fraction expansion* of $\theta \in \mathbb{R}$ is defined as $[c_0, c_1, c_2, \ldots]$ whenever

$$\theta = c_0 + \cfrac{1}{c_1 + \cfrac{1}{c_2 + \ddots}}$$

with $c_0 \in \mathbb{Z}$ and $c_i \in \mathbb{Z}^+$ for every $i > 0$.

Theorem

If θ is a positive irrational, its continued fraction expansion is unique.
If θ is a quadratic irrational, its continued fraction expansion is eventually periodic.

$$e = [2, 1, 2, 1, 1, 4, 1, 1, 6, \ldots], \qquad \pi = [3, 7, 15, 1, 292, 1, 1, 1, 2, \ldots]$$

$$\varphi = \frac{1+\sqrt{5}}{2} = [\,\overline{1}\,] = [1, 1, 1, \ldots].$$

Continued fraction expansion

and Sturmian sequences

Theorem

Let α be an irrational number, $0 < \alpha < 1$ having continued fraction expansion $\alpha = [0, d_1 + 1, d_2, d_3, \ldots]$. Then $\mathbf{c}_\alpha = \lim_{n \to \infty} w_n$, where

$$w_{-1} = \mathtt{b}, \qquad w_0 = \mathtt{a}, \qquad \text{and} \qquad w_n = w_{n-1}^{d_n} w_{n-2} \quad \text{for every } n \geq 1.$$

Example (Fibonacci, $\theta = [0, 2, \overline{1}]$)

$$\mathbf{f} = \mathbf{c}_\theta = \mathtt{abaababaabaababa} \cdots$$

Indeed $w_{-1} = \mathtt{b}$, $w_0 = \mathtt{a}$, and

- $w_1 = \mathtt{ab}$
- $w_2 = \mathtt{aba}$
- $w_3 = \mathtt{abaab}$
- ...

Representation of Sturmian sequences

Theorem [D., Tahay (2022)]

A Sturmian word with quadratic slope can be represented as a column in the space-time diagram of a one-dimensional cellular automaton.

Representation of Sturmian sequences

Step 1 : Prefix lengths

Proposition

Let $(d_n)_{n\geq 1}$ be an eventually periodic integer sequence, $d_1 \geq 0$, $d_i > 0$ for every $i \geq 2$. Let $(S_n)_{n\geq 0}$ be the integer sequence defined by

$$S_{n+1} = d_{n+1} S_n + S_{n-1}$$

for every $n \geq 0$, with $S_{-1}, S_0 > 0$. Then $\mathbb{1}_{\{S_n\}} \in \mathcal{S}$.

Representation of Sturmian sequences

Step 2 : Prefixes

Proposition

Let $(w_n)_{n \geq -1}$ be defined by

$$w_{-1} = \mathtt{b}, \qquad w_0 = \mathtt{a}, \qquad \text{and} \qquad w_n = w_{n-1}^{d_n} w_{n-2} \quad \text{for every } n \geq 1$$

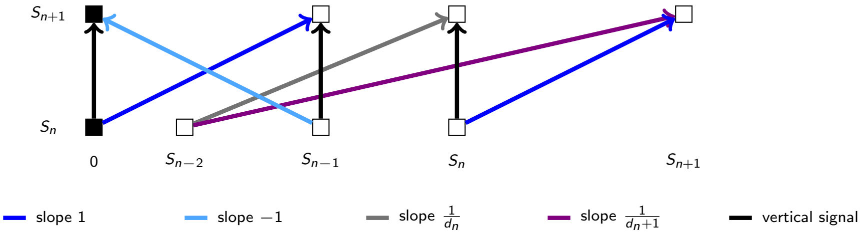
where $(d_n)_n$ is eventually periodic, with $d_1 \geq 0$ and $d_i > 0$ for every $i > 0$.

Then $\mathbf{w} = \lim_{n \to \infty} w_n \in \mathcal{S}$.

Representation of Sturmian sequences

Step 1 : Prefix lengths - Proof 1/2 ($d_{n+1} = 1$)

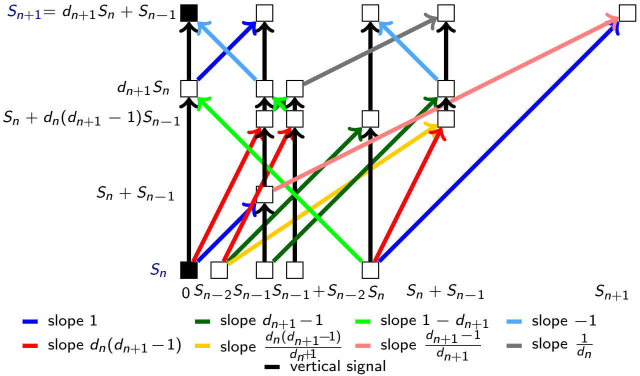
$$S_{n+1} = d_{n+1} S_n + S_{n-1}$$



Representation of Sturmian sequences

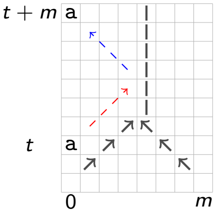
Step 1 : Prefix lengths - Proof 2/2 ($d_{n+1} \neq 1$)

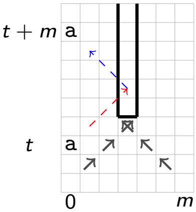
$$\begin{aligned} S_{n+1} &= d_{n+1}S_n + S_{n-1} \\ &= S_n + (d_{n+1}-1)S_n + S_{n-1} \\ &= S_n + (d_{n+1}-1)(d_nS_{n-1} + S_{n-2}) + S_{n-1} \\ &= S_n + d_n(d_{n+1}-1)S_{n-1} + (d_{n+1}-1)S_{n-2} + S_{n-1} \end{aligned}$$



Representation of Sturmian sequences

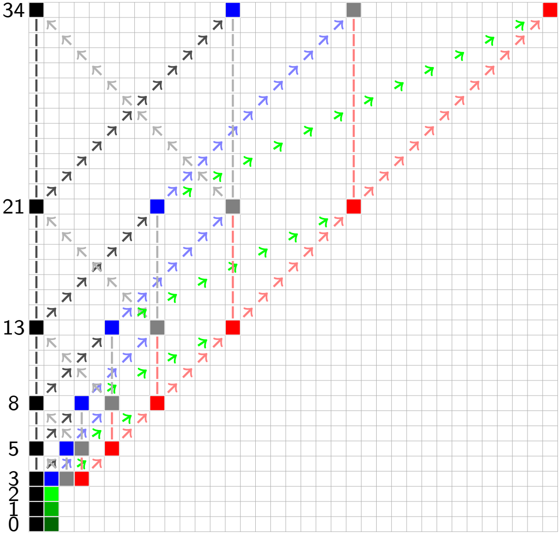
Step 2 : Prefixes - main idea of the proof





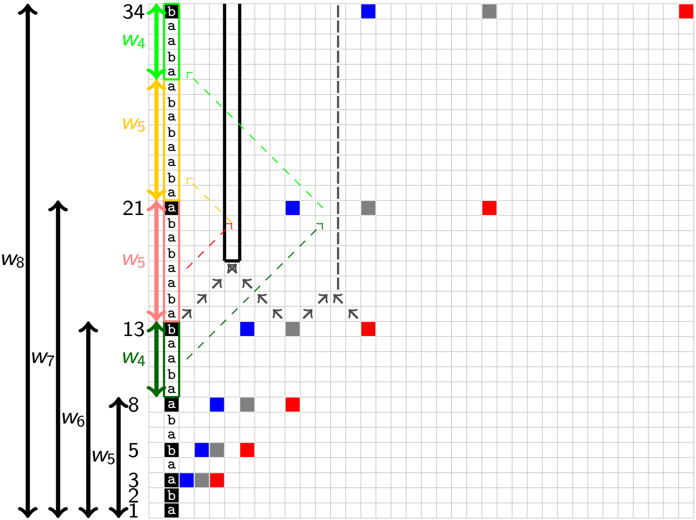
Representation of Sturmian sequences

Example : Fibonacci 1/2



Representation of Sturmian sequences

Example : Fibonacci 2/2



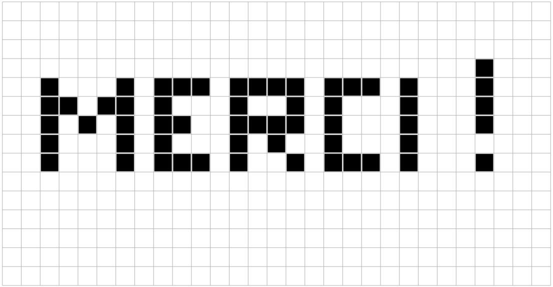
Question :
Is it possible to generalize to larger alphabets (Arnoux-Rauzy, dendric, etc.) ?

Idea :
Every Sturmian word $\mathbf{w}$ can be written as

$$\mathbf{w} = \lim_{n \to \infty} \psi_0 \psi_1 \cdots \psi_n \left(w^{(n)} \right)$$

where $w^{(i)}$ is a word and $(\psi_i)_i \in \{G, D\}^{\mathbb{N}}$ (infinitely many of each), with

$$G = \begin{cases} \mathtt{a} \mapsto \mathtt{a} \\ \mathtt{b} \mapsto \mathtt{ab} \end{cases} \quad \text{and} \quad D = \begin{cases} \mathtt{a} \mapsto \mathtt{ba} \\ \mathtt{b} \mapsto \mathtt{b} \end{cases} .$$


MERCI !