A new characterization of p-automatic sequences

Eric Rowland, joint work with Reem Yassawi

2013 April 5

Outline

1. Old characterizations of automatic sequences
2. Cellular automata
3. Automatic sequences and cellular automata

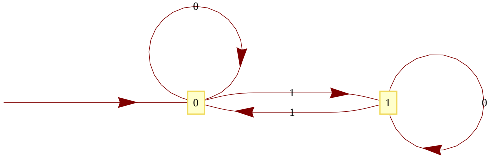
The Thue–Morse sequence

$$T(n) = \begin{cases} 0 & \text{if the binary representation of } n \text{ has an even number of 1s} \\ 1 & \text{if the binary representation of } n \text{ has an odd number of 1s.} \end{cases}$$

The Thue–Morse sequence $T(n)_{n \geq 0}$ is

$$0\,1\,1\,0\,1\,0\,0\,1\,1\,0\,0\,1\,0\,1\,1\,0\,1\,0\,0\,1\,0\,1\,1\,0\,0\,1\,1\,0\,1\,0\,0\,1 \cdots .$$

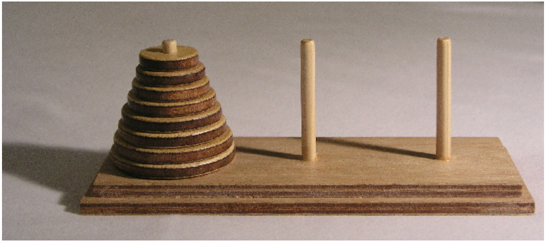
The Thue–Morse sequence is 2-automatic.
It is computed by a deterministic finite automaton with output in base 2:



A sequence $s(n)_{n \geq 0}$ is k-automatic if there is DFAO whose output is $s(n)$ when fed the base-k digits of n, from least to most significant.

Tower of Hanoi

The sequence's alphabet is not necessarily the digits $\{0, 1, \ldots, k\}$.



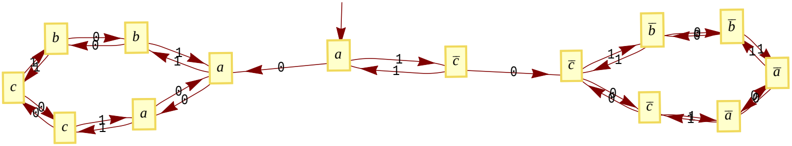
The optimal solution

$$a\overline{c}bac\overline{b}a\overline{c}b\overline{a}cba\overline{c}bac\overline{b}acb\overline{a}c\overline{b}a\overline{c}bac\overline{b}a\overline{c}\cdots$$

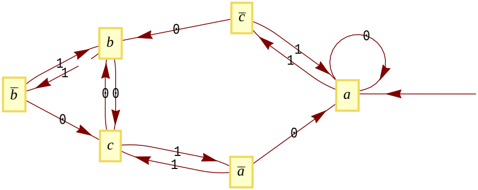
to the "infinite" tower of Hanoi puzzle is 2-automatic.

Tower of Hanoi automata

An automaton:



An automaton that reads the most significant digit first:



Thue–Morse arises in ...

- Combinatorics on words:
  An overlap is a word *awawa* where *w* is a word and *a* is a letter.
  The Thue–Morse word $0110100110010110\cdots$ is overlap-free.

- Multigrades:

$$0^0 + 3^0 + 5^0 + 6^0 = 1^0 + 2^0 + 4^0 + 7^0 = 4$$
$$0^1 + 3^1 + 5^1 + 6^1 = 1^1 + 2^1 + 4^1 + 7^1 = 14$$
$$0^2 + 3^2 + 5^2 + 6^2 = 1^2 + 2^2 + 4^2 + 7^2 = 70$$

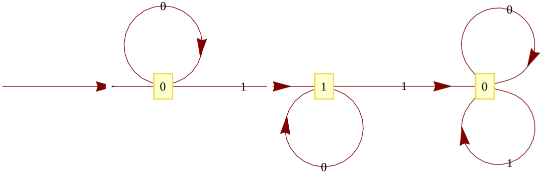
  In general, $\displaystyle\sum_{n=0}^{2^\ell - 1} (-1)^{T(n)} n^m = 0$ for $0 \leq m \leq \ell - 1$.

- Interesting products:

$$\prod_{n \geq 0} (n+1)^{(-1)^{T(n)}} = \frac{1 \cdot 4 \cdot 6 \cdot 7 \cdot 10 \cdot 11 \cdot 13 \cdot 16 \cdots}{2 \cdot 3 \cdot 5 \cdot 8 \cdot 9 \cdot 12 \cdot 14 \cdot 15 \cdots} = \frac{1}{\sqrt{2}}$$

More 2-automatic sequences

- The characteristic sequence of powers of 2 is 2-automatic:

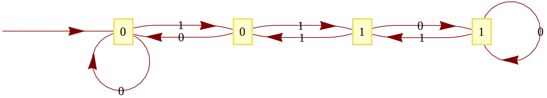


- The Rudin–Shapiro sequence

$$s(n) = \begin{cases} 0 & \text{if the binary representation of } n \text{ has an even number of 11s} \\ 1 & \text{if the binary representation of } n \text{ has an odd number of 11s} \end{cases}$$

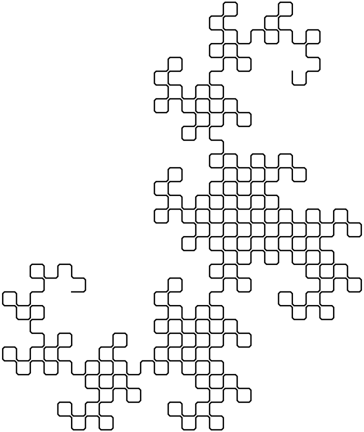
is 2-automatic:

$$0\,0\,0\,1\,0\,0\,1\,0\,0\,0\,0\,1\,1\,1\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,1\,1\,1\,0\,0\,0\,1\,0 \cdots.$$



Dragon curve

The sequence of turns in the paperfolding curve is 2-automatic.



Characterization as a fixed point

Several characterizations of automatic sequences are known.

Theorem (Cobham 1972)

A sequence is k-automatic if and only if it is the image, under a coding, of a fixed point of a k-uniform morphism.

Thue–Morse: $\varphi(0) = 01, \varphi(1) = 10$; then $\varphi^\omega(0) = 01101001 \cdots$.

Rudin–Shapiro: $\varphi(a) = ab, \varphi(b) = ac, \varphi(c) = db, \varphi(d) = dc$;
$\tau(a) = \tau(b) = 0, \tau(c) = \tau(d) = 1$.

$$\varphi^\omega(a) = abacabdbabacdcacabacabdbdcdbabdb \cdots$$
$$\tau(\varphi^\omega(a)) = 00010010000111010001001011100010 \cdots$$

Subsequences characterization

Theorem (Eilenberg 1974)

A sequence $s(n)_{n\geq 0}$ is k-automatic if and only if the set of subsequences $\{s(k^e n + r)_{n\geq 0} : e \geq 0, 0 \leq r \leq k^e - 1\}$ is finite.

Thue–Morse: $T(2n) = T(n)$, $T(2n+1) = 1 - T(n)$

Rudin–Shapiro:

$$\begin{aligned} s(4n+1) &= s(2n) = s(n) \\ s(8n+7) &= s(2n+1) \\ s(16n+3) &= s(8n+3) \\ s(16n+11) &= s(4n+3) \end{aligned}$$

Algebraic characterization

Let p be a prime.
Let $\mathbb{F}_q$ be a finite field of characteristic p.

Theorem (Christol–Kamae–Mendès France–Rauzy 1980)

A sequence $s(n)_{n\geq 0}$ of elements in $\mathbb{F}_q$ is p-automatic if and only if the formal power series $\sum_{n\geq 0} s(n)t^n$ is algebraic over $\mathbb{F}_q(t)$.

For Thue–Morse, $G(t) = \sum_{n\geq 0} T(n)t^n$ over $\mathbb{F}_2(t)$ satisfies

$$tG(t) + (1+t)G(t)^2 + (1+t^4)G(t)^4 = 0.$$

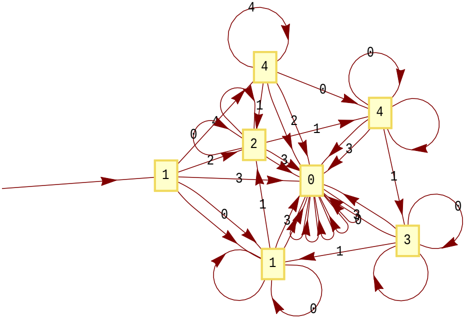
The assumption $s(n) \in \mathbb{F}_q$ is not restrictive:
For a sequence on Σ, any injection $\Sigma \hookrightarrow \mathbb{F}_q$ gives an algebraic series.

Algebraic sequences modulo p

Let $C(n)$ be the nth Catalan number and $G(t) = \sum_{n \geq 0} C(n) t^n$. Reducing

$$1 - G(t) + tG(t)^2 = 0$$

modulo p implies that $C(n) \bmod p$ is p-automatic. For $p = 5$:

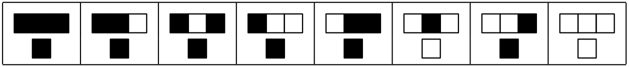


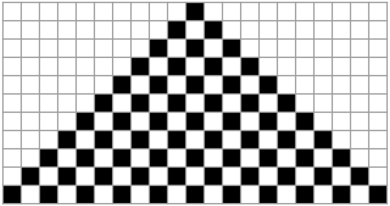
Outline

1. Old characterizations of automatic sequences
2. Cellular automata
3. Automatic sequences and cellular automata

One-dimensional cellular automata

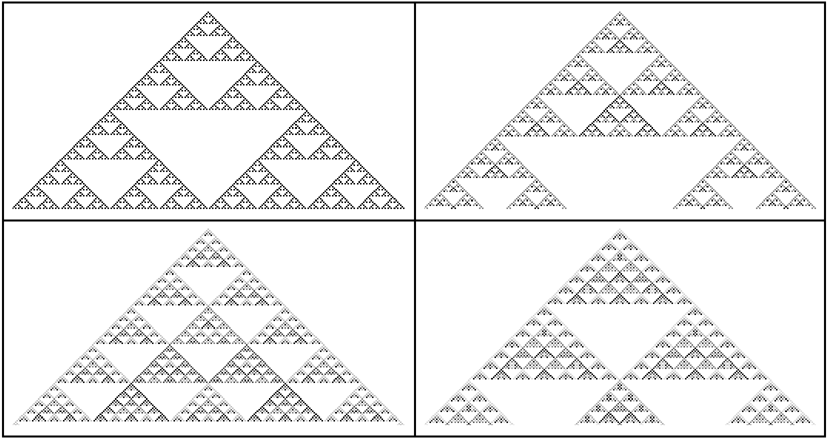
- finite alphabet Σ (for example $\{\square, \blacksquare\}$)
- function $i : \mathbb{Z} \to \Sigma$ (the initial condition)
- integer $\ell \geq 0$
- function $f : \Sigma^\ell \to \Sigma$ (the local update rule)





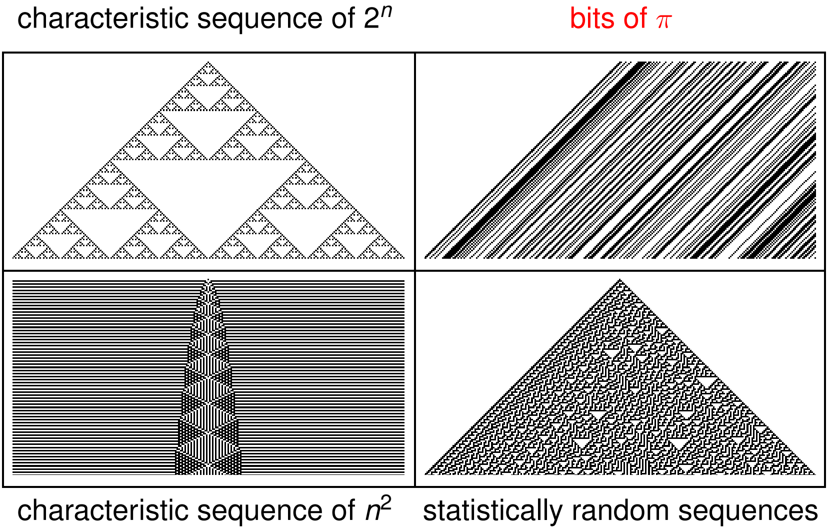
Binomial coefficients

Binomial coefficients modulo k are produced by cellular automata.



The local rule is $f(u, v, w) = u + w$ modulo k.

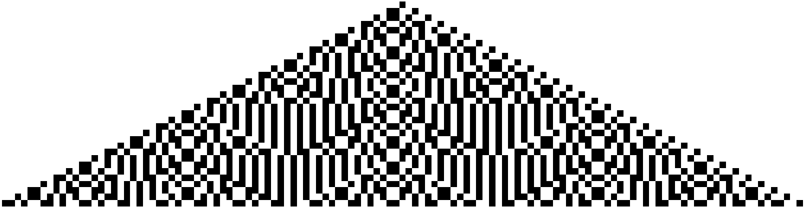
Column sequences



Finiteness condition

Finiteness condition:
The initial condition is eventually periodic in both directions.

The Thue–Morse sequence occurs in this $\ell = 5$ cellular automaton:



$$0\,1\,1\,0\,1\,0\,0\,1\,1\,0\,0\,1\,0\,1\,1\,0\cdots$$

Outline

1. Old characterizations of automatic sequences
2. Cellular automata
3. Automatic sequences and cellular automata

Linear cellular automata

A cellular automaton is linear if the local rule $f : \mathbb{F}_q^\ell \to \mathbb{F}_q$ is $\mathbb{F}_q$-linear.

For example, $f(u, v, w) = u + w$ for binomial coefficients modulo p.

Theorem (Litow–Dumas 1993)

Every column of a linear cellular automaton over $\mathbb{F}_p$ is p-automatic.

The proof uses two theorems about formal power series — Christol's theorem and Furstenberg's theorem.

Furstenberg's theorem

The diagonal of a bivariate series $\sum_{n\geq 0}\sum_{m\geq 0} a(n,m)t^n x^m$ is

$$\sum_{n\geq 0} a(n,n)t^n.$$

Theorem (Furstenberg 1967)

A formal power series $G(t)$ is algebraic over $\mathbb{F}_q(t)$ if and only if $G(t)$ is the diagonal of a bivariate rational series $F(t,x)$.

Sketch of Litow–Dumas proof

Represent the nth row $\cdots\ a(n,-1)\,a(n,0)\,a(n,1)\ \cdots$ by

$$R_n(x) = \cdots + a(n,-1)x^{-1} + a(n,0)x^0 + a(n,1)x^1 + \cdots\,,$$

which is rational since the initial condition is eventually periodic.

Linearity of the rule means $R_{n+1}(x) = C(x)R_n(x)$ for some $C(x)$.
For binomial coefficients, $C(x) = x + \frac{1}{x}$.

Then the bivariate series $F(t,x) = \sum_{n\geq 0}\sum_{m\in\mathbb{Z}} a(n,m)t^n x^m = \sum_{n\geq 0} R_n(x)t^n = \sum_{n\geq 0}(C(x)t)^n R_0(x)$ is rational.

Column m of $F(t,x)$ is the diagonal of $x^{-m}F(tx,x)$, hence it is algebraic (Furstenberg) and hence p-automatic (Christol et al.).

The converse

Given a p-automatic sequence, can we compute a cellular automaton?

We can reverse the proof, using the other directions of Christol's and Furstenberg's theorems.

Issue 1: We may not get a recurrence for $R_n(x)$ of order 1.
In general, $C_0(x)R_n(x) = \sum_{i=1}^{d} C_i(x)R_{n-i}(x)$.

To deal with this, we introduce memory into the cellular automaton.

Issue 2: We need $C_0(x)$ to be a (nonzero) monomial so that each $\frac{C_i(x)}{C_0(x)}$ is a Laurent polynomial, so that the update rule is local.

Constructing a Thue–Morse cellular automaton

Christol's theorem gives that $x = \sum_{n \geq 0} T(n) t^n$ satisfies

$$tx + (1+t)x^2 + (1+t^4)x^4 = 0.$$

Replace $x \mapsto 0 + 1t + 1t^2 + t^2 x$, and divide by t^3.
Then $G(t) := \sum_{n \geq 0} T(n+3) t^n$ satisfies $P(t, G(t)) = 0$, where

$$P(t,x) = x + x^2 t + (1+x^2)t^2 + x^4 t^5 + (1+x^4)t^9.$$

By Furstenberg's theorem, $T(n+2)$ is the coefficient of x^{-2} in $R_n(x)$:

$$\frac{\frac{\partial P}{\partial x}(t,x)}{P(t,x)} = \frac{1}{x} + t + \left(\frac{1}{x^2} + 1 + x\right) t^2 + \cdots = \sum_{n \geq 0} R_n(x) t^n.$$

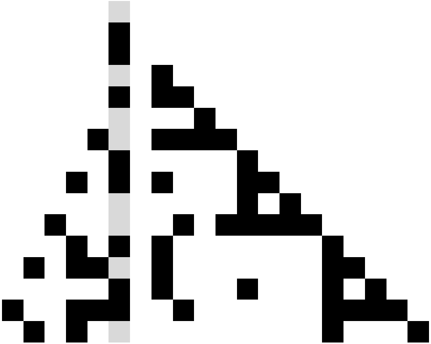
$R_n(x)$ satisfies the recurrence

$$R_n(x) = xR_{n-1}(x) + \left(\frac{1}{x} + x\right) R_{n-2}(x) + x^3 R_{n-5}(x) + \left(\frac{1}{x} + x^3\right) R_{n-9}(x)$$

for all $n \geq 10$, which determines a linear cellular automaton rule with memory 9.

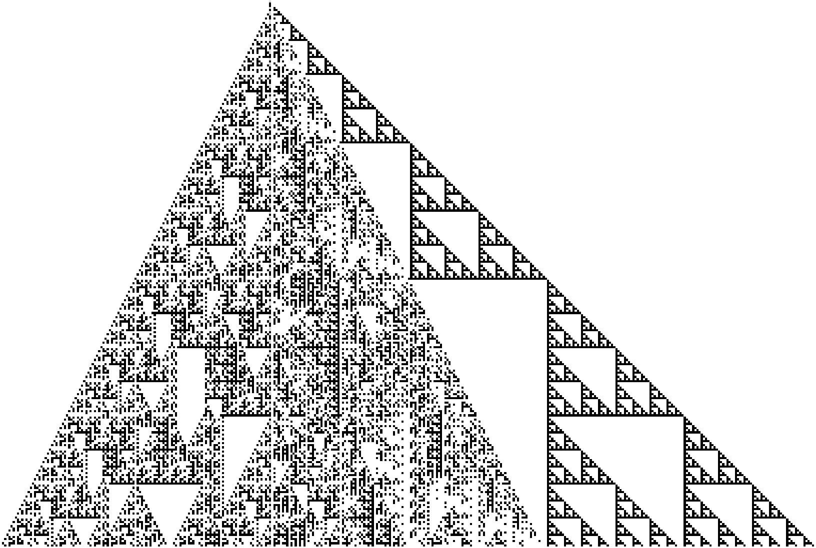
Restoring initial terms

Extend the memory to $9 + 3 = 12$ without introducing dependence on the earliest 3 rows.
Then $T(n)_{n \geq 0}$ occurs in Column -2 from initial conditions $R_{-2}, \ldots, R_9$.



0 1 1 0 1 0 0 1 1 0 0 1 0 1 1 0 $\cdots$

Thue–Morse cellular automaton with memory 12



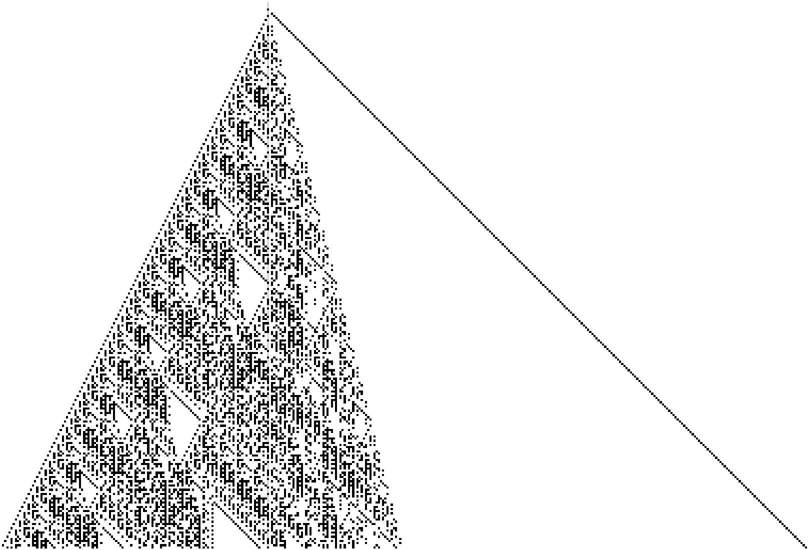
Result

Theorem (Rowland–Yassawi)

Every p-automatic sequence of elements in $\mathbb{F}_q$ occurs as a column of a linear cellular automaton over $\mathbb{F}_q$ with memory whose initial conditions are eventually periodic in both directions.

Combined with the Litow–Dumas result, we have a new characterization of *p*-automatic sequences (for prime *p*).

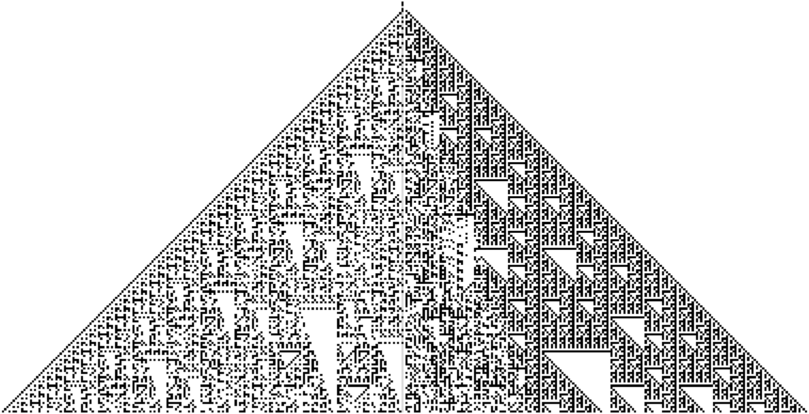
Rudin–Shapiro cellular automaton with memory 20



Baum–Sweet cellular automaton with memory 27

The Baum–Sweet sequence 1 1 0 1 1 0 0 1 0 1 0 0 $\cdots$ is defined by

$$s(n) = \begin{cases} 0 & \text{if the binary representation of } n \text{ contains a block of 0s of odd length} \\ 1 & \text{if not.} \end{cases}$$



Removing memory

If we give up linearity, we can get a cellular automaton without memory.

Corollary

Every p-automatic sequence occurs as a column of a cellular automaton whose initial condition is eventually periodic in both directions.

Invertible cellular automata

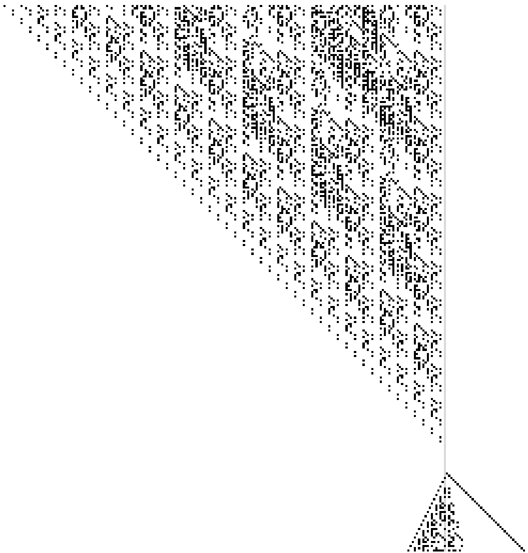
A cellular automaton rule is invertible if it has an inverse rule.

In other words, it can be evolved backward in time as well as forward.

Corollary

If $s(n)_{n\geq 0}$ is a p-automatic sequence, then for some $r \geq 0$ the sequence $s(n)_{n\geq r}$ occurs as a column of an invertible cellular automaton with memory.

Invertible Rudin–Shapiro cellular automaton



Open questions

- Given a p-automatic sequence on the alphabet $\Sigma \subset \mathbb{F}_q$, one can find a cellular automaton (without memory) with at most $q^{d+r+1} + |\Sigma|$ states containing the sequence as a column. Can this bound be improved?
- Does there exist a 3-automatic sequence $s(n)_{n \geq 0}$ on a binary alphabet such that $s(n)$ is not eventually periodic and $s(n)$ occurs as a column of a (nonlinear) 2-state cellular automaton?
- Exhibit a sequence that does not occur as the column of a cellular automaton with eventually periodic initial conditions.