Algorithmical unsolvability of the ergodicity problem for binary cellular automata [1]

André Toom

Federal University of Pernambuco, department of statistics

Recife, PE, 50740-540, Brazil

E-mail toom@de.ufpe.br, toom@member.ams.org

Abstract

We prove algorithmical unsolvability of the ergodicity problem for a class of one-dimensional translation-invariant cellular automata with two states, all transition probabilities being in $\{0,\ 1/2,\ 1\}$, and a similar statement for finite space.

Key words: random processes, local interaction, cellular automata, ergodicity, unsolvability, Turing machine.

AMS Subject Classification: Primary 60K35, 68Q80, 03D10, 03D35 Secondary 82C22, 03B25

[1]Supported by FAPESP, grant #98/15994-0

Our configuration space is $\{0,1\}^Z$, whose elements are written as $v = (v_i) = (\ldots, v_{-1},\ v_0,\ v_1, \ldots)$, where all $v_i \in \{0,1\}$. We denote $\mathcal{M}$ the set of normed measures on $\{0,1\}^Z$ (i.e. on the σ-algebra generated by cylinder sets). For any $v \in \{0,1\}^Z$ we denote $\delta_v \in \mathcal{M}$ the measure concentrated in v. For any natural number r and non-negative transition probabilities $\theta(x|y_1, \ldots, y_r)$ for $x, y_1, \ldots, y_r \in \{0,1\}$, where $\sum_x \theta(x|y_1, \ldots, y_r) \equiv 1$, we define a linear operator $P : \mathcal{M} \to \mathcal{M}$ as follows: for any $w = (w_i) \in \{0,1\}^Z$ the result $P\,\delta_w$ of application of P to δ_w is a product measure, for which

$$\forall\ i \in Z,\ \ x \in \{0,1\}\ :\ (P\,\delta_w)(v_i = x) = \theta(x|w_{i+1}, \ldots, w_{i+r}).$$

As usual, we call a measure $\mu \in \mathcal{M}$ invariant if $P\mu = \mu$ and we call an operator ergodic if for any initial measure μ the sequence $P^t\,\mu$ tends to one and the same limit. When speaking about our set of operators and other sets of objects, we assume whenever necessary that they are enumerated in some constructive way.

Theorem. Consider only those of operators described above, for which all the transition probabilities $\theta(x|y_1, \ldots, y_r)$ equal 0 or $1/2$ or 1. There is no algorithm to decide which of these operators are ergodic.

Our method is essentially that of Kurdyumov [1, 2, 3], who proved algorithmic unsolvability of the ergodicity problem for a class of cellular automata, where the set of states of every site was arbitrary finite, but the transition probabililies depended only on three neighbors. To prove our theorem, we first have to consider a case of this sort, but we cannot simply refer to Kurdyumov's theorem; we need to reformulate his construction. As a by-product, we make his result stronger, because our construction involves only two rather than three neighbors. Let us take any non-empty finite set S and denote $\mathcal{M}_S$ the set of normed measures on the configuration space

S^Z . For any non-negative transition probabilities $\theta(x|y,z)$ for all $x,y,z \in S$, where $\sum_x \theta(x|y,z) \equiv 1$, we define a linear operator $P_S : \mathcal{M}_S \to \mathcal{M}_S$ as follows: for any configuration $w = (w_i) \in S^Z$, the result $P_S\,\delta_w$ of application of P_S to δ_w is a product measure, for which

$$\forall\ i \in Z,\quad x \in S\ :\ (P_S\,\delta_w)(v_i = x) = \theta(x|w_i,\ w_{i+1}).$$

Lemma 1. Consider only those of operators P_S described above, for which all the transition probabilities $\theta(x|y,z)$ equal 0 or $1/2$ or 1 . There is no algorithm to decide which of them are ergodic.

In proving lemma 1, we shall use the following set of Turing machines with one head and one bi-infinite tape. To describe a Turing machine of our class, we choose two non-empty finite sets G and H , where G is the set of tape symbols and $H \cup \{stop\}$ is the set of head states. Also we choose three functions:

$$F_{\text{tape}}\ :\ G \times H \to G,$$
$$F_{\text{head}}\ :\ G \times H \to H \cup \{stop\},$$
$$F_{\text{move}}\ :\ G \times H \to \{-1, 0, 1\}.$$

When the machine starts, all cells of the tape are filled with the initial symbol $g_1 \in G$, the head is in the initial state $h_1 \in H$ and the head observes the 0-th cell of the tape. At every step the head simultaneously writes into that cell of the tape, which it observes, a new symbol according to the function F_{tape} , goes to a new state according to the function F_{head} , and moves one cell left or does not move or moves one cell right along the tape according to the values $-1, 0, 1$ of the function F_{move} respectively, the arguments of all the three functions being the symbol in the presently observed cell of the tape and the present state of the head. The machine stops when and if the head reaches the state *stop* . (That is why we don't need to define our functions when the head is in the state *stop* .) It is well-known that the problem of deciding, which of these machines ever stop, is algorithmically unsolvable. Now, for any Turing machine M we shall construct an operator P_S . We set

$$S = S_{left} \times S_{right} \times S_{par} \times S_{tape} \times S_{head}, \tag{1}$$

where

$$S_{left} = S_{right} = S_{par} = \{0, 1\}, \qquad S_{tape} = G, \qquad S_{head} = H \cup \{0\}.$$

Accordingly, we write a generic element of S as

$$x = (left(x),\ right(x),\ par(x),\ tape(x),\ head(x)). \tag{2}$$

We say that a state x has a *left bracket* if $left(x) = 1$ and that it has a *right bracket* if $right(x) = 1$. We call x *even* if $par(x) = 0$ and *odd* otherwise. We call x a *no-head* if $head(x) = 0$ and a *head* otherwise. We call x a *stop-head* if $head(x) = stop$. The state $(0,\ 0,\ 0,\ g_1,\ 0)$ is called $empty$, the state $(1,\ 1,\ 1,\ g_1,\ h_1)$ is called $newborn$ and the state $(0,\ 0,\ 0,\ g_1,\ stop)$ is called $final$. For brevity we shall write $F_*(x) = F_*(tape(x),\ head(x))$, where $*$ means 'tape', 'head' or 'move'. We say that a head x wants to move left, to stay or to move right when $F_{\text{move}}(x)$ equals -1 , 0 or 1 respectively. Using i.i.d. random variables $b(i)$, which equal 0 or 1 with probabilities 1/2 and 1/2, we define our operator P_S as follows: For any $w = (w_i) \in S^Z$ the measure $P_S\,\delta_w$ is induced by this distribution with the following map: the i-th component equals $d_{b(i)}(w_i,\ w_{i+1})$, where the functions $d_0,\ d_1\ :\ S^2 \to S$ are defined as follows:

$$d_0(y, z) = \begin{cases} final & \text{if } y \text{ or } z \text{ is a stop-head,} \\ newborn & \text{otherwise.} \end{cases} \tag{3}$$

The definition of $d_1(\cdot)$ consists of several rules.

Rule 0. If y or z is a stop-head, then $d_1(y,z) = final$.

Formulating all the other rules, we assume that neither y nor z is a stop-head. We call a pair $(y,z) \in S^2$ *normal* if $par(y) = par(z)$ and at least one of y, z is a no-head. We call a normal pair (y,z) *even* if $par(y) = 0$ and *odd* otherwise.

Rule 1. Whenever the pair (y,z) is not normal, $d_1(y,z) = empty$.

It remains to define $d_1(y,z)$ when the pair (y,z) is normal. All the rules pertaining to this case come in symmetric pairs, each pair consisting of one even rule applied to even pairs and one odd rule applied to odd pairs. First we define $d_1(y,z)$ when the pair (y,z) is even. For any $x \in S$ let us denote $\overline{x}$ that element of S , which has all the same components as x in the representation (2) except $par(\overline{x}) = 1 - par(x)$. Here are the even rules:

Rule 2-even. If both y, z are no-heads, then $d_1(y,z) = \overline{y}$.

Rule 3-even. If y is a head which wants to move left, then

$$d_1(y,z) = (0,\ right(y),\ 1,\ F_{\text{tape}}(y),\ 0).$$

Rule 4-even. If y is a head which wants to stay, then

$$d_1(y,z) = (left(y),\ right(y),\ 1,\ F_{\text{tape}}(y),\ F_{\text{head}}(y)).$$

Rule 5-even. If y is a head which wants to move right, then

$$d_1(y,z) = (left(y),\ 0,\ 1,\ F_{\text{tape}}(y),\ 0).$$

Rule 6-even. If z is a head, which wants to move left and has a left bracket, then

$$d_1(y,z) = (1,\ 0,\ 1,\ g_1,\ F_{\text{head}}(z)).$$

Rule 7-even. If z is a head, which wants to move left and has no left bracket, then

$$d_1(y,z) = (left(y),\ 0,\ 1,\ tape(y),\ F_{\text{head}}(z)).$$

Rule 8-even. If z is a head, which wants to stay or move right, then $d_1(y,z) = \overline{y}$.

Thus all the even rules are defined. When the normal pair (y,z) is odd, the definition of $d_1(\cdot)$ is obtained from the definition for the even case by permuting the two values of $par(\cdot)$. permuting y and z and permuting left and right. Our operator P_S is defined.

Lemma 2. Thus constructed operator P_S is ergodic if and only if the Turing machine M stops.

Proof of lemma 2. Denote δ_{final} the measure concentrated in the configuration "all components are in the state $final$". Due to the upper line in (3) and rule 0, this measure is invariant for P_S. Therefore P_S is ergodic if and only if $P_S^t\,\delta_w$ tends to δ_{final} for any initial configuration w. Any $(s,t) \in Z \times Z_+$ will be called a *point* and any map $\rho : Z \times Z_+ \to S$ will be called a *realization*. We shall represent the measures $P_S^t\,\delta_w$ for all $t \geq 0$ as "time-slices" of one distribution on the set of realizations. With all points (s,t), where $t > 0$, we associate i.i.d. variables $b(s,t)$, which equal 0 or 1 with probabilities 1/2 and 1/2. Using the inductive rule

$$\rho(s,t) = d_{b(s,t)}(\rho(s,t-1),\ \rho(s+1,\ t-1))$$

for all $s \in Z$ and $t > 0$, and the initial condition $\rho(s,0) = w_s$ for all $s \in Z$, we define a distribution on the set of realizations, whose restrictions when t is fixed coincide with $P_S^t\,\delta_w$. We say that a *birth* occurs at a point (s,t) if $b(s,t) = 1$. Now we argue in the following two directions.

One direction: Let us suppose that M stops after T steps and prove that $P_S^t\,\delta_w$ tends to δ_{final} for any initial w. Let us consider a region $[s_0 - N,\ s_0 + N] \subset Z$, where $N = 9T$. If a stop-head is present there, it turns into $final$, which expands left due to the first line of (3) and rule 0. If there is no stop-head there, then the following scenario has a positive probability: First, at some time t_0 births may occur in all sites in the range $[s_0 - N,\ s_0 + N]$, which always may happen with a positive probability due to the second line in (3). At the next time step all of these sites except the last one become empty a.s. due to rule 1. At the next time step birth may occurs at the s_0-th site and this may be the only birth that occurs in the space-time region

$$\{(s,t) \mid s_0 - N \le s < s_0 + N - (t - t_0),\ 0 < t - t_0 \le N\},$$

the probability of which is also positive. Under these conditions, the restrictions of measures $P_S^t\,\delta_w$ to this region are concentrated in configurations imitating the functioning of M during time long enough for M to stop. As soon as the head stops, it turns into $final$, which expands left due to rule 0. This scenario happens somewhere on the right side of any given place almost sure, whence $P_S^t\,\delta_w \to \delta_{\text{final}}$.

The other direction: Let us assume that M never stops, i.e. continues to function forever. According to its functioning, for all $s \in Z$ and $t \in Z_+$ we denote: $tape_M(s,t)$ - the symbol in the s-th site of the tape at the moment t and $head_M(s,t)$ - the head state of M at time t if the head is at s at this time and 0 if the head is not there. Let us take the initial measure δ_{empty} concentrated in the configuration "all components are in the empty state". If P_S is ergodic, then $P_S^t\,\delta_{\text{empty}}$ must tend to δ_{final}, whence the event

$$\text{"the } 0\text{-th particle is a stop-head at time } T_{stop}\text{"} \qquad (4)$$

must have a positive probability for some value of T_{stop}. We shall prove, however, that this probability is zero. Given a realization ρ and a point (s,t) such that $\rho(s,t)$ is a head, we define two integer numbers $L_\rho(s,t)$ and $R_\rho(s,t)$ as follows: $L_\rho(s,t)$ is the smallest integer number L such that

$$L < i \le s \Rightarrow left(i,t) = 0 \quad \text{and} \quad L \le i < s \Rightarrow right(i,t) = 0$$

if such a number exists; otherwise $L_\rho(s,t)$ is undefined. $R_\rho(s,t)$ is the greatest integer number R such that

$$s \le i < R \Rightarrow right(i,t) = 0 \quad \text{and} \quad s < i \le R \Rightarrow left(i,t) = 0$$

if such a number exists; otherwise $R_\rho(s,t)$ is undefined. For our operator both $L_\rho(s,t)$ and $R_\rho(s,t)$ are defined a.s for any (s,t), where $\rho(s,t)$ is a head. Now for any realization ρ and any point (s,t), where $\rho(s,t)$ is a head, we call the sequence $\rho(L_\rho(s,t),t),\ldots,\rho(R_\rho(s,t),t)$ the *domain* of the point (s,t), given ρ. It is easy to prove by induction that if a sequence $x_0,\ldots,x_n$ has a positive probability to occur in a realization of $P^t\,\delta_{\text{empty}}$ as a domain, then:

$$\begin{cases} \bullet\,)\ \text{All } x_0,\ldots,x_n \text{ have one and the same parity.} \\ \bullet\,)\ \text{Exactly one of } x_0,\ldots,x_n \text{ is a head.} \\ \bullet\,)\ \text{There are integer numbers } \Delta s \text{ and } \Delta t \text{ such that for all } s \in [0,n] \\ tape(x_s) = tape_M(s+\Delta s,\ \Delta t) \text{ and } head(x_s) = head_M(s+\Delta s,\ \Delta t)\,. \end{cases}$$

Now assume that the event (4) has a positive probability. We cover the event (4) by a countable set of events, in everyone of which the domain of the point $(0,\ T_{stop})$ is specified. Take one of these events with a certain domain

$$\rho(L_\rho(0,\ T_{stop}),\ T_{stop}),\ldots,\rho(R_\rho(0,\ T_{stop}),\ T_{stop}).$$

Then there are integer numbers Δs and Δt such that

$$head(\rho(0,\ T_{stop})) = head_M(\Delta s,\ \Delta t).$$

Hence from (4) M stops, which contradicts our assumption. Thus lemma 2 is proved, whence lemma 1 immediately follows.

Proof of the theorem. We call a *bit* a variable, which may be equal only 0 or 1. We call a k-*code* or just a *code* any sequence of k bits. An infinite code is an infinite sequence of bits. Given a finite code C, we call its *length* and denote $length(C)$ the number of bits in it. There is an empty code or 0-code, whose length is 0. Given a finite or infinite code C, we denote $C[i]$ the i-th bit in C. Also for any positive $i \leq j \leq length(C)$ we denote $C[i, j]$ and call a sub-code of C the sequence of its bits from the i-th to j-th one. Given finite codes $C_1, \ldots, C_n$, their concatenation, that is the code obtained by writing them one after another, is denoted $concat(C_1, \ldots, C_n)$. We shall also use $concat(\ldots, C_{-1},\ C_0,\ C_1, \ldots)$, concatenation of a bi-infinite sequence of finite codes.

Suppose that a Turing machine M is given and the corresponding operator P_S is already constructed as described above. Now we construct an operator P. To do this, first let us enumerate elements of S: $S = \{e_1, \ldots, e_{|S|}\}$. For every $e_i \in S$ we denote $bin(e_i)$ the $|S|$-code, in which the i-th bit is 1 and all the others are zeros. For any $\alpha \in S$ and $\beta \in \{0, 1\}$ we denote:

$$frame(\alpha,\ \beta) = concat(bin(\alpha),\ 0,\ \beta,\ 0110). \tag{5}$$

Any $(|S| + 6)$-code, which has the form (5), is called a *frame.* Given a frame C of the form (5), $\alpha(C)$ and $\beta(C)$ denote α and β respectively. Let us take $r = 3(|S| + 6) - 1$ and call a r-code C *regular* if it can be represented as

$$C = concat(A,\ C_0,\ C_1,\ B),$$

where $C_0,\ C_1$ are frames. Notice that this representation of C, if it exists, is unique. Now we can define our transition probabilities. If the code $y_1, \ldots, y_r$ is not regular, then

$$\theta(x|y_1, \ldots, y_r) = \begin{cases} 1 & \text{if } x = 1, \\ 0 & \text{if } x = 0. \end{cases}$$

When $y_1, \ldots, y_r$ is regular, let is first define a frame C_{new} as follows:

$$C_{new} = frame(d_\beta(y_0,\ y_1),\ 0,\ 0),$$

where $\beta = \beta(C_0)$ and $y_i = \alpha(C_i)$ for $i = 0,\ 1$.

Now we define $\theta(x|y_1, \ldots, y_r)$ when the code $y_1, \ldots, y_r$ is regular as follows: If $length(A) = 4$, then $\theta(x|y_1, \ldots, y_r) = 1/2$. Otherwise

$$\theta(x|y_1, \ldots, y_r) = \begin{cases} 1 & \text{if } x = C_{new}[\,(|S| + 6) - length(A)\,], \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 3. Thus constructed operator P is ergodic if and only if the Turing machine M stops.

Proof. Denote δ_1 the measure concentrated in the configuration "all the sites are in the state 1". Since the r-code consisting of ones is not regular, δ_1 is invariant. Therefore our operator P is ergodic if and only if $P^t\,\delta_w$ tends to δ_1 for any initial $w \in \{0,1\}^Z$. Thus it is sufficient to argue in the following two directions.

One direction: Let us call a *triple* the code 111. Given any $v \in \{0,1\}^Z$, it is evident that if for any $m \in Z$ there is $n > m$ such that $v[n,\ n+2]$ is a triple, then each triple expands left, whence $P^t\,\delta_v \to \delta_1$. Now suppose that M stops, take any initial $w \in \{0,1\}^Z$ and consider two cases:

Either for any $m \in Z$ there is $n > m$ such that $w[n+1,\ n+(|S|+6)]$ is not regular. Then at $t=1$ we obtain a triple near that area.

Or there is $m \in Z$ such that for all $n > m$ the sub-code $w[n+1,\ n+(|S|+6)]$ is regular. Then w on the right side of some number is a concatenation of frames. Therefore at $t=1$ we a.s. obtain infinitely many sub-codes, which are frames, whose α is *newborn* . Each of them, in the absence of other births in a large enough region, imitates the functioning of M , which leads to a stop-head. This will happen a.s. on the right side of any number, after which there will be a triple there.

The other direction. Let us call $v \in \{0,1\}^Z$ a *frame-configuration* if it can be written as $v = concat(\ldots, C_{-1},\ C_0,\ C_1,\ \ldots)$, where all C_i are frames. Substituting every frame C_i in this formula by $\alpha(C_i)$, we obtain a sequence of elements of S , which we denote $\sigma(v)$. Let us assume that M never stops, but $P^t\,\mu$ tends to δ_1 for any initial μ and come to a contradiction. It is sufficient to take μ , which is a product-measure, in which every v_i equals:

•) 1 and 0 with probabilities 1/2 and 1/2 if $i \equiv |S|+2 \pmod{(|S|+6)}$;

•) $frame(empty,\ 0,\ 0)[j]$, where $1 \le j \le |S|+6$ and $j \equiv i \pmod{(|S|+6)}$ otherwise.

It is easy to prove that for every t the measure $P^t\,\delta_w$ is concentrated in frame-configurations and imitates $P_S^t\,\delta_{\text{empty}}$ in the following sense: the map σ turns $P^t\,\delta_w$ into $P_S^t\,\delta_{\text{empty}}$. The last part of the proof of lemma 2 shows that $(P_S^t\,\delta_{\text{empty}})(v_i = x)$ can be positive only if x is non-stop. Therefore a.s. no stop-head will ever appear and this imitation will continue to infinity, whence the measure $P^t\delta_w$ will not tend to δ_1. Thus our theorem is proved.

Note. A similar statement can be proved for finite analogs of processes described above. Let us say that we have a pattern if a natural r and transition probabilities $\theta(x|y_1,\ldots,y_r)$ for $x,y_1,\ldots,y_r \in \{0,1\}$, where $\sum_x \theta(x|y_1,\ldots,y_r) \equiv 1$, are chosen for all $x,y_1,\ldots,y_r \in \{0,1\}$. For every pattern and every natural m we can define a Markov chain with a finite set of states $\{0,1\}^m$, the space being the set Z_m of residues modulo m. Then there is no algorithm to decide for which of these patterns, even if all $\theta(\cdot)$ equal 0 or 1/2 or 1, the Markov chain is ergodic for all m large enough. To prove this, it is sufficient to use the same set of Turing machines and the same definition of the pattern corresponding to any Turing machine as described above.

References

[1] G. L. Kurdyumov. An algorithm-theoretic method in studying homogeneous random networks, In: R. L. Dobrushin, V. I. Kryukov, and A. L. Toom (editors). Locally Interacting Systems and Their Application in Biloogy. Lecture Notes in Maqthematics, 653, Springer, pp. 37-55.

[2] G. L. Kurdyumov. An algorithm-theoretic method for the study of homogeneous random networks. Adv. Probab., 6, pp. 471-504.

[3] A. L. Toom, N. B. Vasilyev, O. N. Stavskaya, L. G. Mityushin, G. L. Kurdyumov and S. A. Pirogov. Discrete Local Markov Systems. *Stochastic Cellular Systems : ergodicity, memory, morphogenesis.* Ed. by R. Dobrushin, V. Kryukov and A. Toom. Nonlinear Science: theory and applications, Manchester University Press, 1990, pp. 1-182.