Rev. Mat. Apl. **12**:93-102

Revista de
Matemáticas Aplicadas

Departamento de Ingeniería Matemática

ERGODICITY FOR A CLASS OF PROBABILISTIC CELLULAR AUTOMATA

P.A. FERRARI
Instituto de Matemática e Estatística
Universidade de São Paulo
Cx. Postal 20570
01498 São Paulo, Brasil
pferrari@brusp.bitnet

Abstract

We study a class of probabilistic cellular automata (PCA) which includes majority vote models, discrete time Glauber dynamics and combinations of these processes with mixing dynamics. We give sufficient condition for the ergodicity of the processes. The method is based on a graphical representation and the construction of a "generalized dual process".

AMS Classifications: 60K35.

Key Words: Probabilistic cellular automata, ergodic systems, generalized duality, graphical representation.

1.- Introduction

A probabilistic cellular automata (PCA) is a discrete time Markov process σ_n in the state space $\mathbf{X} := \{-1,1\}^{\mathbb{Z}^d}$. At each time n the spin at site $x, x \in \mathbb{Z}^d$ is updated according to a random rule that takes account of the configuration at time $n-1$ in a finite neighborhood of x. The rule can be time dependent, for instance when spins belonging to different sublattices are updated at different times. A measure μ on $\mathbf{X}$ is called invariant if σ_n has distribution μ implies that also σ_{n+1} has distribution μ.

A PCA is called ergodic if there exists only one invariant measure μ and starting from any measure the system converges to μ. The study of ergodicity is one of the most relevants problems in systems with infinitely many components. The first sufficient condition for ergodicity of spin flip processes was given by Dobrushin [3]. There are many papers dealing with systems that accept as invariant and reversible the Gibbs states of statistical mechanics. There is a close connection between the lack of ergodicity of the system and the existence of more than one Gibbs state for a given potential (phase transition). This is reviewed in [13] for interacting particle systems, the continuous time version of PCA. We refer to [6] and [14] for reviews and examples of PCA. The sovietic literature is being reviewed in [18].

The aim of this paper is to illustrate the use of what we call generalized duality to prove ergodic properties for PCA. It extends the notion of duality, a technique introduced in interacting particle systems. In words duality consists in expressing functions of the configuration at time t depending on coordinates in a finite set A by local functions of the configuration at time zero depending of a finite set A_t. The process admits a dual if A_t is Markov and the set of functions for which this is possible generates the continuous functions on X. Using duality one transforms a problem that originally has infinitely many components in a problem depending only on a finite number of coordinates. This has been largely developed for proving ergodic properties for the simple exclusion process, voter models, cancellative systems and Glauber-type dynamics. The standard references are Griffeath [11] Liggett [13] Gray [8] and Durrett [4].

Nevertheless not all the processes accept a dual. We show that it is possible to give sufficient conditions for ergodicity of any system by constructing a generalized dual process. This process is not Markovian in general. The construction that we present here appeared first for continuous time in [2] for the study on the hydrodynamics of Glauber-Kawasaky dynamics and in [5] for the study of ergodicity of those processes.

2.- Definitions and Results

A probabilistic automata is a discrete time stochastic process on the state space $\mathbf{X} := \{-1, 1\}^{\mathbb{Z}^d}$. We study two types of dynamics: spin-flip and mixing.

The spin-flip dynamics is represented as follows: let $\sigma \in X$ be a given configuration. Define $G\sigma$ as the random configuration with distribution given by

$$\begin{aligned} &P(G\sigma(x) = -\sigma(x)) = h(x, \sigma) \\ &P(G\sigma(x) = \sigma(x)) = 1 - h(x, \sigma) \end{aligned} \tag{2.1}$$

for all $x \in \mathbb{Z}^d$ where the function h depends on a finite set of coordinates R_x and satisfies $0 \le h(x, \sigma) \le 1$. Furthermore, for fixed σ, the random variables $\{G\sigma(x)\}_x$ are mutually independent.

In order to represent the mixing dynamics, let $\{V_i\}_{i \in I}$ be a partition of $\mathbb{Z}^d$, where all the V_i's have a uniformly bounded number of elements:

$|V_i| \leq v \in I\!N$, for all i. For a given configuration σ, we define $M\sigma$, the configuration obtained when the spins in each of the V_i are permuted in some fixed way. More precisely, if for each $i, q_i : V_i \to V_i$ is a bijection, we define, for $x \in V_i$,

$$M\sigma(x) = \sigma(q_i(x)). \tag{2.2}$$

where $1\{\cdot\}$ is the characteristic function of $\{\cdot\}$. This is a deterministic (no random) dynamics.

The mixing dynamics can also be random. To define that, consider, for each V_i, a family of permutations (bijections) $q_{i,j}, j \in J$, and define the random configuration $M\sigma$ with distribution given by:

$$P[M\sigma(x) = \sigma(q_{i,j}(x)), \text{for all } x \in V_i] = p_{i,j}, \tag{2.3}$$

where, for each $i, p_i, .$ is a probability over J, and the random vectors $\{M\sigma(x), x \in V_i\}_i$ are mutually independent.

We also consider cases where more than one rule can be applied successively. For instance, if G_1 and G_2 are as in (2.1), we may consider $G = G_1 G_2$. The same is valid for the mixing rules.

We define the spin flip mixing automata σ_n, as the Markov process on X with the following probability transition function:

$$E(f(\sigma_n)|\sigma_{n-1} = \sigma) = E(f(GM\sigma)). \tag{2.4}$$

2.1.- Examples

Pure noise. This is the simplest dynamics. All spins flip independently. The flip probability is given by

$$h(x, \sigma) = \epsilon$$

for all x, σ.

Nearest neighbors voter model. For this model

$$h(x, \sigma) = \frac{1}{2d} \sum_{e:|e|=1} \frac{|\sigma(x) - \sigma(x+e)|}{2}$$

in such a way that the probability of flipping the spin at x is proportional to the number of nearest neighbor sites with opposite spin. This system is not ergodic. It has at least 3 invariant measures: δ_1, δ_{-1}, and $\frac{1}{2}\delta_{\eta_1} + \frac{1}{2}\delta_{\eta_2}$, where $\eta_1(x_1, ..., x_d) = 1$ (respectively -1) if $\sum x_i$ is even (odd) and $\eta_2(x_1, ..., x_d) = -1$ (respectively 1) if $\sum x_i$ is even (odd). The last measure reflects the fact that in this model there are two independent space-time sublattices.

Deterministic majority vote model. In this case

$$h(x, \sigma) = 1\{\sum_{e:|e|=1} \frac{|\sigma(x) - \sigma(x+e)|}{2} > d\}$$

in such a way that the spin flips if the majority of its nearest neighbor sites have the opposite spin. This process has lots of configurations that are invariant.

Majority vote model. This is the model resulting when the deterministic majority vote model rule and the pure noise with parameter ϵ are applied successively. This model was widely studied by Gray [9], [10]. The majority vote model is expected to be ergodic in one dimension and to be not ergodic in two or more dimensions.

Glauber dynamics. We call Glauber any dynamics that has as reversible measure a Gibbs measure. There is a parametrized family of Glauber dynamics for any Gibbs measure. As an example consider the nearest neighbor ferromagnetic Gibbs measure. It is defined using the conditional probabilities

$$\mu(\sigma(x) = \xi(x), x \in F | \sigma(x) = \xi(x), x \notin F) = Z^{-1}(\xi) \exp\{\beta \sum_{x,y} \xi(x)\xi(y)\},$$

where β is a parameter (the "inverse temperature") and the sum runs over the set $\{(x,y) \in \mathbb{Z}^d : x \in F \subset \mathbb{Z}, |x-y| = 1\}, F$ is any finite set. The normalizing constant $Z(\xi)$ makes μ a probability. Define $H(x,\sigma) := \sum_{|e|=1} \sigma(x)\sigma(x+e)$. Let $h(x,\sigma)$ be a function satisfying

$$\frac{h(x,\sigma)}{h(x,\sigma^x)} = \frac{\exp\{\beta H(x,\sigma^x)\}}{\exp\{\beta H(x,\sigma)\}} \tag{2.5}$$

where

$$\sigma^x(z) = \begin{cases} \sigma(z) & \text{if } z \neq x \\ -\sigma(z) & \text{if } z = x \end{cases}$$

Now, consider the rules G_1 and G_2 (odd and even) defined by

$$P(G_1\sigma(x) = -\sigma(x)) = h(x,\sigma) 1\{x \text{ is odd}\}$$
$$P(G_2\sigma(x) = -\sigma(x)) = h(x,\sigma) 1\{x \text{ is even}\}$$

where $x = (x_1, ..., x_d) \in \mathbb{Z}^d$ is odd (even) if Σx_i is odd (even). Then the process σ_n defined by $\sigma_{n+1} = G_1 G_2 \sigma_n$ is a Glauber dynamics. The reader can also check that the simultaneous updating of all the spins does not give rise to a Glauber Dynamics. The function h can be chosen in many ways. The most famous are Gibbs sampler:

$$h(x,\sigma) = \frac{\exp\{\beta H(x,\sigma^x)\}}{\exp\{\beta H(x,\sigma)\} + \exp\{\beta H(x,\sigma^x)\}} \tag{2.6}$$

and Metropolis:

$$h(x,\sigma) = \begin{cases} 1 & \text{if } H(x,\sigma) < H(x,\sigma^x) \\ \exp \beta\{H(x,\sigma^x) - H(x,\sigma)\} & \text{otherwise} \end{cases}$$

Nearest neighbors deterministic mixing. We define this dynamics in one dimension. Extensions to more dimensions are immediate. Define $V_i^1 = \{2i, 2i-1\}$, $i \in \mathbb{Z}$,

$$q_i^1(x) = \begin{cases} x+1 & \text{if } x = 2i-1 \\ x-1 & \text{if } x = 2i \end{cases} \tag{2.7}$$

and $V_i^2 = \{2i, 2i+1\}$,

$$q_i^2(x) = \begin{cases} x+1 & \text{if } x = 2i \\ x-1 & \text{if } x = 2i+1 \end{cases} \tag{2.8}$$

and let M_1 and M_2 be the corresponding rules as in (2.2). Then define $\sigma_{2n} = M_1\sigma_{2n-1}$, $\sigma_{2n+1} = M_2\sigma_{2n}$. This dynamics is the discrete analogous of the free gas with two velocities. Spins starting at even sites move at speed one while those starting at odd sites move at speed -1.

There exists a simple program that simulates the above models and others for pc compatible micro computers [16].

Let

$$r = \sup_x |R_x|, \tag{2.9}$$

where R_x is the set of coordinates determining $h(x,\sigma)$, defined in (2.1). Let

$$k = \inf_x \inf_\sigma \min\{h(x,\sigma), 1-h(x,\sigma)\}. \tag{2.10}$$

Intuitively k is the minimun between the probability of flipping and the probability of staying. Since those probabilities sum one, $k \le \frac{1}{2}$.

We say that a process is exponentially ergodic if for any cylindric f there exist positive constants $\alpha_1 = \alpha_1(f)$ and α_2 such that $|Ef(\sigma_n) - \mu f| \le \alpha_1 e^{-\alpha_2 n}$. In the next Theorem α_1 is proportional to the number of sites determining f times the infinite norm of f.

Theorem 2.1. *Let σ_n be a spin flip mixing automata as defined in (2.4). If*

$$r(1-2k) < 1, \tag{2.11}$$

then the process is exponentially ergodic for any mixing dynamics M. If equality holds in (2.11), then the process is ergodic.

Notice that in (2.11) the maximun probability of flipping is controlled as well as the minimun. Otherwise a process with big probability of flipping could be not ergodic. An extreme example in this direction is the process in which all spins flip simultaneously with probability one.

The Theorem is proved in the next section.

3.- Graphic Representation and Duality

In order to represent graphically the process, we rewrite (2.1) partitioning $\mathbf{X}$ in subsets A_i with constant $h(x,\sigma)$:

$$h(x,\sigma) = \sum_{i=1}^{p} a_i(x) 1\{\sigma \in A_i(x)\}$$

where the a_i's are nonnegative ordered constants such that if $\sigma \in A_i(x)$ then $h(x,\sigma) = a_i$. The $A_i(x)$'s are cylindric sets depending on σ only through R_x and for each x, $\{A_i(x)\}_i$ is a partition of $\mathbf{X}$. The number of terms of the sum, $p = p(h(x,.))$, is at most the number of parts of R_x; $\sup_x |R_x| \le r < \infty$. Assume that the cylindric set $A_i(x) \subset \mathbf{X}$ depends on the set $F_j(x) \subset R_x$. The process will be constructed in the semi-space $\mathbb{Z}^d \times \mathbb{Z}^+$, where the last coordinate represents time. Now denote $a_0 := k$, defined in (2.11) and agrupe terms to get

$$\begin{aligned} h(x,\sigma) &= k\sum_{i=1}^{p} 1\{\sigma \in A_i(x)\} + \sum_{i=1}^{p}(a_i(x) - k) 1\{\sigma \in A_i(x)\} \\ &= k + \sum_{i=1}^{p}(a_i(x) - a_{i-1}(x)) \sum_{j \ge i} 1\{\sigma \in A_j(x)\}. \end{aligned} \tag{3.1}$$

In other words, with probability k the spin at x flips without regarding at the configuration σ. It follows from the definition of k that with at least probability k the spin at x stays the same.

Graphic construction of the process. To each point $(x,n), x \in \mathbb{Z}^d$, $n \in \mathbb{Z}^+$ associate a random variable I assuming values $i = \delta^+, \delta^-, 1, \dots, p$, defined by

$$\begin{aligned} &P(I(x,n) = \delta^+) = P(I(x,n) = \delta^-) = a_0 \\ &P(I(x,n) = i) = a_i(x) - a_{i-1}(x), \quad 1 \le i \le p. \end{aligned} \tag{3.2}$$

Assume that $\{I(x,n)\}$ are mutually independent and let $(\Omega, \mathbf{F}, P)$ be the probability space induzed by this family. Let $\{i(x,n) : x \in \mathbb{Z}^d, n \ge 1\}$ be a realization of these random variables.

We construct the process by induction. Suppose known σ_n, the configuration at time n. Then, the (random) configuration at time $n+1$ is given by

$$\sigma_{n+1}(x) = \begin{cases} +1 & \text{if } i(x,n+1) = \delta^+; \\ -1 & \text{if } i(x,n+1) = \delta^-; \\ -\sigma_n(x) & \text{if } \sigma_n \in A_j(x), \text{ for some } j \ge i(x,n+1); \\ \sigma_n(x), & \text{otherwise.} \end{cases} \tag{3.3}$$

It is not difficult to see that the process defined this way has distribution given by (3.1).

Construction of the dual process. The generalized dual of the process is a marked branching structure constructed in the space

$$\mathbb{Z}^d \times \{\delta, 1, \ldots, p\} \times \mathbb{Z}^+ \tag{3.4}$$

Assume that in the time interval $[0,t]$ we have a realization $i(x,n)$ of the random variables $I(x,n)$, $0 \leq n \leq t$. Reverse the time, calling $m = t - n$. Let D be a subset of $\mathbb{Z}^d$. This set is the "base" of our branching structure. It is the projection of the structure at time $n = 0$. Suppose that we know the structure in the time interval $[0,m]$. Call D_m its projection at time m. Then, at time $m+1$, the projection of the structure is given by:

$$D_{m+1} = \bigcup_{x \in D_m} \quad \bigcup_{j \geq i(x,t-m)} F_j(x)$$

where we use the convention that, if $i(x,t-m) = \delta^{\pm}$, then $\cup_{j \geq i(x,t-m)} F_j(x) = \theta$.

The knowledge of the structure D_m^D, $0 \leq m \leq t$, *plus* the marks $i(x,t-m)$, $x \in D_m$, $0 \leq m \leq t$ allows us to recover the configuration σ_t in the set D, as far as the initial configuration σ be known in the set D_t. Notice that D_t can be the empty set. In this case the configuration σ_t in D does not depend on $\sigma_0 = \sigma$. The structure can be described formally. Let $i(x,n)$ be a realization of the random variable $I(x,n)$. Then the generalized dual process is

$$\hat{D}^D_{[0,t]} = \{(m, x, i(x,t-m), D_m^D) : x \in D_m^D, 0 \leq m \leq t\} \tag{3.5}$$

where D_m is the projection of $\hat{D}_{[0,t]}$ at time m.

The duality formula is just formal:

$$1\{\sigma_t(x) = 1 : x \in D\} = \Phi(\hat{D}^D_{[0,t]}, \sigma(y), y \in D_t^D), \tag{3.6}$$

where Φ is some function that can be easily computed for each realization, but it is in general imposible to have a formula for it. When the function Φ depends just on $\sigma(y)$, $y \in D_t^D$, we have the usual notion of duality. The point in this construction is that, if $D_t^D = \phi$, then σ_t in D does not depend on σ_0. A sufficient condition for

$$P(D_t^D = \phi) < \alpha_1 e^{-\alpha_2 t} \tag{3.7}$$

can be found by dominating $|D_t^D|$ with a usual discrete time branching process $R_t^{|D|}$, for which at each step, each branch dies and creates $r =$

$\sup_x |R_x|$ offsprings with probability $1-2a_0$ and no offspring with probability $2a_0$; $R_0^{|D|} = |D|$. Such a branching process dies exponentially fast, i.e. $P(R_t^{|D|} \neq \phi) \leq \alpha_1 e^{-t\alpha_2}$, $\alpha_1, \alpha_2 > 0$, if the average number of offsprings

$$r(1 - 2a_0) < 1. \tag{3.8}$$

See for instance [12]. Equation (3.6) implies (3.5) which implies exponential ergodicity in Theorem 2.1 when there is no mixing process.

Graphical representation of the mixing marks. We consider that the mixing rule is applied at times such that the spin flip rule is not applied. This is not a loss of generality as it stands only for a change of time variables. The times of application of mixing rules can be random or deterministic. In the first case the realization of an appropriate random variable indicates which are the mixing times. Now, if t is a mixing time, connect each site x at time t with the corresponding site x' at $t+1$, according to the mixing rule applied at this time. To realize the process apply the Glauber rules as described before at the Glauber times and if t is a mixing time, simply decide that at time $t+1$ the spin x' is the same than the spin at x at time t.

The proof of the Theorem when mixing marks are present follows the lines of the proof above. The same branching process still dominates the branching structure constructed with the Glauber and mixing marks. The reason is that the presence of mixing marks does not change the number of branches of the branching structure.

4.- Concluding Remarks

To conclude we comment some open problems. The positive rates conjecture is one of the most important open problems ([13] and [7]). Is it true that any one dimensional finite range PCA with positive k is ergodic? This is proven for Glauber dynamics, by using absense of phase transition for the corresponding Gibbs measure [13]. One can start trying to show ergodicity for the nearest neighbor majority vote model described above. In this model at each given time each spin takes the value of the majority of the set composed by itself and the two nearest neighbors with probability $1-\epsilon$ and the opposite value with probability ϵ. The analogous model in continuous time has as invariant measure a Gibbs measure. Also discrete time majority vote models are ergodic if the updating is done by sublattices is such a way that never two neighbors are updated at the same time and all sites are updated infinitely often. For simultaneous updating, it follows from Theorem 2.1 that any (finite range) majority vote model is ergodic if ϵ is close enough to $1/2$. This has been proven in [15] by coupling methods. Gray [10] proved this result when ϵ is small enough. One could try to improve the condition for ergodicity (2.11) by dominating the branching structure by an oriented percolation type of process [4].

An important step in proving ergodicity would be to show monotonicity properties for functions of the noise ϵ. For instance, for the majority vote model, one would like to show that when the initial condition is the configuration "all ones" the average magnetization (i.e. the average value of the spin) at any given time is a decreasing function of the noise parameter ϵ. The difficulty here is that when ϵ increases the spin correlations will decrease.

Approach to a product measure when the probability of mixing is big compared with the probability of flipping. Assume that the PCA has distribution

$$E(f(\sigma_n)|\sigma_{n-1} = \sigma) = \lambda E(f(M\sigma)) + (1 - \lambda)E(f(G\sigma))$$

where λ is a parameter between 0 and 1. That is, the mixing and Glauber rules are applied randomly with probability λ and $1 - \lambda$ respectively. It is presumably possible to prove that in the exponentially ergodic regime of G the only invariant measure μ_λ converges as $\lambda \to 1$ to a product measure. This was proven by [5] in the continuous case.

References

[1] M. Aizenmann, R. Holley, "Rapid covergence to equilibrium of stochastic Ising models in the Dobrushin-Shlosman regime", in Percolation Theory and Ergodic Theory of Infinite Particle Systems, Springer, Berlin, H. Kesten (ed), IMA Vol. Math. Appl. 8, 1-11 (1987).

[2] A. De Masi, P.A. Ferrari, J.L. Lebowitz, "Reaction-diffusion equations for interacting particle systems", J. Stat. Phys. 44, 589-644 (1986).

[3] R.L. Dobrushin, "Markov with a large number of locally interacting components: existence of a limit process and its ergodicity", Problems Inform. Transmission 7, 149-164 (1971).

[4] R. Durrett, "Lecture notes on particle systems and percolation", Wordworth & Brooks/Cole. Advanced Books and Software, Pacific Grove, California (1988).

[5] P.A. Ferrari, "Ergodicity for spin systems with stirrings", Ann. Probab. 18(4), 1523-1538 (1990).

[6] A. Georges, P. Le Doussal, "From equilibrium to probabilistic cellular automata", Preprint (1988).

[7] L. Gray, "The positive rates problem for attractive nearest neighbor spin systems on $\mathbb{Z}$", Z. Wahrsch. Verw. Gebiete. 51, 171-184 (1982).

[8] L. Gray, "Duality for general attractive spin systems with applications in one dimension", Ann. Probab. 14, 371-396 (1986).

[9] L. Gray, "The behavior of processes with statistical mechanical properties", In Percolation Theory and Ergodic Theory of Infinite Particle Systems, Kesten (ed), The IMA volumes in mathematics and its applications 8, Springer Verlag, New York (1987).

[10] L. Gray, "Finite range majority vote models in one dimension", Preprint (1986).

[11] D. Griffeath, "Aditive and cancellative interacting particle systems", Springer, Lecture Notes in Mathematics 724 (1979).

[12] T.E. Harris, "The theory of braching processes", Springer, Berlin (1963).

[13] T.M. Liggett, "Interacting particle systems", Springer Verlag (1985).

[14] Ch. Maes, "Some aspects of the stochastic evolution of interacting spin systems", in Leuven Notes in Mathematical and Theoretical Physics, M. Fanner and A. Verbeure (ed), Leuven Univ. Press, Leuven, Belgium, 1A (1989).

[15] A.L. Toom, N.B. Vasilyev, O.N. Stavakaya, L.G. Mitjushin, G.L. Kardunov, S.A. Pigorov, "Discrete local Markov systems", Preprint (1989).

Software

[16] P.A. Ferrari, E. De Mauro, "Automat, a simulator of stochastic automata", Sao Paulo, IME-USP (1989).

Acknowledgments

Partially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq, Brasil.

I thank the Centro Internacional de Física for very warm hospitality in Bogotá.