

Random Networks of Automata: A Simple Annealed Approximation.

B. DERRIDA and Y. POMEAU

*Service de Physique Théorique, CEN Saclay
91191 Gif-sur-Yvette Cedex, France*

(received 8 October 1985; accepted 31 October 1985)

PACS. 05.40. – Fluctuation phenomena, random process, and Brownian motion.

Abstract. – Kauffman's model is a random complex automata where nodes are randomly assembled. Each node σ_i receives K inputs from K randomly chosen nodes and the values of σ_i at time $t + 1$ is a random Boolean function of the K inputs at time t . Numerical simulations have shown that the behaviour of this model is very different for $K > 2$ and $K \leq 2$. It is the purpose of this work to give a simple annealed approximation which predicts $K = 2$ as the critical value of K . This approximation gives also quantitative predictions for distances between iterated configurations. These predictions agree rather well with the numerical simulations. A possible way of improving this annealed approximation is proposed.

The study of automata has become recently more and more popular among physicists [1]. Problems of cellular automata [2] are very related to classical models of statistical mechanics (percolation, Ising models) [3]. Also the ideas developed in the pattern recognition like the Hopfield model [4, 5].

Networks of Boolean automata to study the behaviour of generic regulatory systems were introduced by KAUFFMAN [6] in 1969. The subsequent studies have revealed surprisingly ordered structures in randomly constructed networks. In particular the most highly organized behaviour (small attractors, smaller number of attractors, stable attractors, etc.) appeared to occur in networks where each node receives inputs from two other nodes.

Two approaches have been taken to attempt to understand this emergence of ordered behaviour. One focused on the percolation of a particular kind of subgraphs (forcing structures [7-10]). A second approach is based on arguments concerning one-step convergence in phase space [8, 11, 12]. While these approaches have been, to some extent, successful, they have not yet explained some of the most important numerical observations. In this note, we shall suggest a novel and independent approach to these questions.

Let us first briefly describe Kauffman's model. The model depends on a parameter K . The system consists of N spins σ_i which can take two possible values ($\sigma_i = 0$ or 1). The time

evolution of this system is given by N Boolean functions of K variables each

$$\sigma_i^{(t+1)} = f_i(\sigma_{i_1}^{(t)}, \sigma_{i_2}^{(t)}, \dots, \sigma_{i_K}^{(t)}). \quad (1)$$

For each i , the spins $\sigma_{i_1}, \sigma_{i_2}, \dots, \sigma_{i_K}$ are randomly chosen among the N spins. They need not be different (*i.e.* i_1 can be equal to i_2 for example). So the system is defined once a function f_i and a set i_1, i_2, \dots, i_K have been chosen for each site i . In fig. 1, we can see an example for $N = 4, K = 2$.

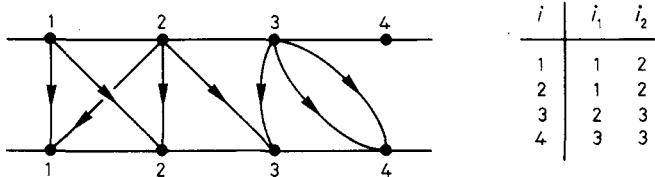


Fig. 1. – An example for $N = 4$ and $K = 2$.

There exist 2^{2^K} possible Boolean functions of K variables. In Kauffman's model, each function f_i is randomly chosen among these 2^{2^K} possible functions.

The system is *random* because, for each i , the set i_1, \dots, i_K is randomly chosen and because the function f_i is random. This randomness is *quenched* because the functions f_i and the sets i_1, \dots, i_K are quenched (they do not change with time).

The most interesting questions one can ask about this system concern its time evolution. For finite N , the long-time behaviour must be periodic because the system has 2^N possible configurations \mathcal{C} of the spins σ_i . So after a long time the system comes back to a configuration that it had already visited and then the evolution becomes periodic.

There are many questions one can ask about this model:

What is the length of the limit cycles?

What is the number of different limit cycles?

If one considers 2 different spin configurations at time $t = 0$, what is the probability that they become the same at time t ?

If one starts with two randomly chosen configurations \mathcal{C}_1 and \mathcal{C}_2 of the system at time $t = 0$, what will be the distance between these configurations after time t and in the limit $t \rightarrow \infty$?

There have been several numerical studies of this model [8, 10]. Let us summarize briefly what has been observed. For $K \leq 2$, the length of cycles seems to be small and of order $\beta(K)N^t$, whereas for $K \geq 3$, it increases like $\exp[\alpha(K)N]$. Therefore, there must be a critical K_c with $2 \leq K_c \leq 3$, where $\alpha(K)$ and $\beta(K)$ should have a singular behaviour. The number of different cycles has been measured. Also the distance between a configuration and its successor has been measured for several cases ($K = 2, 3, 5, 7$).

The purpose of this letter is to present a simple approximation to this model which explains why there is a critical value K_c of K and which allows one to calculate many properties of this model. This approximation neglects the fact that the functions f_i and the sets i_1, \dots, i_K are quenched (*i.e.* constant over time) and instead randomly reassigns inputs and functions to all spins at each time step.

Thus this is a kind of *annealed approximation*. It turns out that its qualitative and quantitative agreement with what has been observed in numerical simulations is surprisingly good.

Let us consider 2 spin configurations \mathcal{C}_1 and \mathcal{C}_2 which are at distance n . By definition the distance $d(\mathcal{C}_1, \mathcal{C}_2)$ between two spin configurations is n , if the number of spins which are different in the two configurations is n . If one considers two randomly chosen configuration \mathcal{C}_1 and \mathcal{C}_2 such that $d(\mathcal{C}_1, \mathcal{C}_2) = n$ at time $t = 0$, one can calculate the probability $P_1(m, n)$ that the distance $d(\mathcal{C}_1', \mathcal{C}_2')$ between their images \mathcal{C}_1' and \mathcal{C}_2' at time $t = 1$ is $d(\mathcal{C}_1', \mathcal{C}_2') = m$.

To calculate $P_1(m, n)$, let us call A the set of spins which are identical in \mathcal{C}_1 and \mathcal{C}_2 and B the set of spins which are different. Set A contains $N - n$ spins, whereas B contains n spins. Let us call $Q(N_0)$ the probability that N_0 spins have all their K arrows coming from set A . One has

$$Q(N_0) = \binom{N}{N_0} \left[\left(\frac{N-n}{N} \right)^K \right]^{N_0} \left[1 - \left(\frac{N-n}{N} \right)^K \right]^{N-N_0}. \quad (2)$$

These N_0 spins will be of course identical in \mathcal{C}_1' and \mathcal{C}_2' . For the remaining $N - N_0$ spins, since at least one of their inputs is different in \mathcal{C}_1 and \mathcal{C}_2 , there is a probability $\frac{1}{2}$ that they are the same in \mathcal{C}_1' and \mathcal{C}_2' and $\frac{1}{2}$ that they differ. Therefore, the probability $P_1(m, n)$ that $d(\mathcal{C}_1', \mathcal{C}_2') = m$ is

$$P_1(m, n) = \sum_{N_0=0}^N Q(N_0) \left(\frac{1}{2} \right)^{N-N_0} \binom{N-N_0}{m}. \quad (3)$$

One can easily compute the sum over N_0 and one finds

$$P_1(m, n) = \frac{N!}{m!(N-m)!} \frac{1}{2^N} \left[1 + \left(1 - \frac{n}{N} \right)^K \right]^{N-m} \left[1 - \left(1 - \frac{n}{N} \right)^K \right]^m. \quad (4)$$

This formula is the exact probability $P_1(m, n)$ for Kauffman's quenched model if \mathcal{C}_1 and \mathcal{C}_2 have been randomly chosen at time $t = 0$.

If one wants to calculate the probability $P_t(m, n)$ that the configurations $\mathcal{C}_1^{(t)}$ and $\mathcal{C}_2^{(t)}$ obtained at time t are at distance m , given that $d(\mathcal{C}_1, \mathcal{C}_2) = n$ at $t = 0$, for Kauffman's model one must take into account that the configurations $\mathcal{C}_i^{(t)}$ for $t \geq 1$ are correlated to the functions f_i and the sets i_1, \dots, i_K . So

$$P_2(m, n) \neq P_2^{\text{annealed}}(m, n) = \sum_{q=0}^N P_1(m, q) P_1(q, n). \quad (5)$$

However, if one neglects these correlations, *i.e.* one studies the model where at each time the functions f_i and the sets i_1, \dots, i_K are changed (so that the functions f_i and the sets i_1, \dots, i_K are no longer quenched), then $P_t(m, n)$ in this *annealed approximation* is given by

$$P_t^{\text{annealed}}(m, n) = \sum_{q_1=0}^N \dots \sum_{q_{t-1}=0}^N P_1(m, q_{t-1}) P_1(q_{t-1}, q_{t-2}) \dots P_1(q_1, n). \quad (6)$$

For large N , one can introduce continuous variables

$$\frac{n}{N} = x, \quad \frac{m}{N} = y \quad (7)$$

and one can see easily from (4) and (6) that all the $P_t^{\text{annealed}}(m, n)$ become very peaked around a well-defined value of y .

From formula (4), one sees that $P_1(m, n)$ is very peaked around a value y_1 given by

$$y_1 = \frac{1 - (1-x)^K}{2}. \quad (8)$$

For similar reasons $P_t^{\text{annealed}}(m, n)$ is peaked around a value of $m = Ny_t$, where y_t is given by

$$y_t = \frac{1 - (1 - y_{t-1})^K}{2}. \quad (9)$$

Y_1 being given by (8).

So we see that once the distance Nx between \mathcal{C}_1 and \mathcal{C}_2 is given, then with probability 1 in the limit $N \rightarrow \infty$, the distance between their images at time t is Ny_t . Therefore, the problem reduces to the study of the map (9).

For $K \leq 2$, the fixed point $x = 0$ is the only fixed point of the map and it is attractive. For any starting value x , $y_t \rightarrow 0$ as $t \rightarrow \infty$. So

$$\lim_{t \rightarrow \infty} \lim_{N \rightarrow \infty} \frac{d(\mathcal{C}_1^{(t)}, \mathcal{C}_2^{(t)})}{N} = 0 \quad \text{for } K \leq 2. \quad (10)$$

For $K > 2$, the fixed point $x = 0$ is unstable and there appears a new fixed point y^* of (9) which is attractive. Therefore, in the limit $t \rightarrow \infty$,

$$\lim_{t \rightarrow \infty} \lim_{N \rightarrow \infty} \frac{d(\mathcal{C}_1^{(t)}, \mathcal{C}_2^{(t)})}{N} = y^*, \quad \text{for } K > 2. \quad (11)$$

So we see that this simple annealed approximation gives a critical value $K_c = 2$.

One can compare the value y^* with what has been measured in numerical simulations of Kauffman's model (see table I). One sees that the agreement is surprisingly good.

TABLE I. – Comparison of the value of y^* with measured distances in simulations of Kauffman's model [10].

K	y^* fixed point of (9)	Masured distance between 24-th and 25-th iterates for $N = 10^8$
2	0	0.113 ± 0.097
3	0.38197	0.37 ± 0.078
5	0.48121	0.485 ± 0.047

It is also possible to calculate within this annealed approximation the mean fluctuations of y around y^* in the limit $t \rightarrow \infty$. This can be done easily by relating the moments of m to those of n using formula (4), assuming that one has a probability distribution of the distances n

$$\begin{cases} \langle m \rangle = \frac{N}{2} \left\langle 1 - \left(1 - \frac{n}{N}\right)^K \right\rangle, \\ \langle m^2 \rangle = \frac{N}{2} + \frac{N(N-1)}{4} \left\langle \left[1 - \left(1 - \frac{n}{N}\right)^K\right]^2 \right\rangle. \end{cases} \quad (12)$$

This gives in the limits $N \rightarrow \infty$ and $t \rightarrow \infty$

$$\langle y^2 \rangle - \langle y \rangle^2 = \frac{\langle m^2 \rangle - \langle m \rangle^2}{N^2} = \frac{\langle n^2 \rangle - \langle n \rangle^2}{N^2} = \frac{1}{N} \left[\frac{1}{2} - (y^*)^2 \right] / \left[1 - \frac{K^2}{4} (1 - y^*)^{2K-2} \right]. \quad (13)$$

So we see that this annealed approximation predicts a critical value $K_c = 2$ and allows calculation of the K -dependence of several quantities.

This annealed approximation can be extended to generalizations of Kauffman's model. For example if instead of having all spins with K inputs, one considers a model where the

number of inputs is not the same for all the spins. Then, if there are N_{ρ_K} spins having K inputs, one can show easily that formula (9) becomes

$$y_{t+1} = \sum_K \rho_K \frac{1 - (1 - y_t)^K}{2}. \quad (14)$$

Another generalization of Kauffman's model consists in giving different weights to the functions f_i . A function f_i takes 2^K possible values 0 or 1. Formula (9) was obtained by giving the same weight to all possible Boolean functions. If instead of that, one gives a weight which depends on the number of times that the function f_i takes the value 1 (namely the probability that a function f_i is proportional to $p^\alpha(1-p)^{2^K-\alpha}$, where α is the number of times that function f_i takes the value 1), then formula (9) becomes

$$y_{t+1} = [1 - (1 - y_t)^K] 2p(1 - p). \quad (15)$$

One big difference between the annealed model studied here and Kauffman's model is that in the annealed model there is no limit cycle because the rule is changed at each time step. So *a priori* the annealed approximation cannot give any prediction for the length of the cycles. However, if one takes a slightly different point of view, one can relate the lengths of the cycles in Kauffman's model to an analogous quantity in the annealed model. Suppose that we start at time $t = 0$ with all the 2^N spins configurations. The system has, at $t = 0$, 2^N different configurations. If one iterates, one can pose the question: how many different configurations will remain after time t ? In the limit of very long time, in Kauffman's model, one obtains the number of points which belong to limit cycles. In the annealed model, one can ask the same question for $K > 2$. At the moment we have not found a way of calculating this quantity and, therefore, we do not know if it would give a good estimate for the total number of periodic points in Kauffman's model. But we think that this quantity would be worth studying.

Lastly, it seems that one can try to improve the annealed approximation described in the present work. In the annealed approximation, one calculates exactly $P_1(m, n)$ for Kauffman's model and then one uses formula (6) to approximate $P_t(m, n)$. One could try to calculate exactly $P_2(m, n)$ for Kauffman's model because this would take into account partly the fact that the f_i and the sets (i_1, \dots, i_K) are quenched and then make an annealed approximation similar to formula (6) based on $P_2(m, n)$.

* * *

This work has benifited a lot from discussions with S. KAUFFMAN. We would like to thank him for his encouragement.

REFERENCES

- [1] *Cellular Automata*, *Physica D*, **10**, 1 (1984).
- [2] S. WOLFRAM: *Rev. Mod. Phys.*, **55**, 601 (1983).
- [3] W. KINZEL: *Z. Phys. B*, **58**, 229 (1985) and references therein; E. DOMANY and W. KINZEL: *Phys. Rev. Lett.*, **53**, 311 (1984); G. GRINSTEIN, C. JAYAPRAKASH and YU HE: preprint (1985).
- [4] J. J. HOPFIELD: *Proc. Nat. Acad. Sci.*, **79**, 2554 (1982).
- [5] D. J. AMIT, H. GUTFREUND and H. SOMPOLINSKY: preprint (1985); W. KINZEL: preprints (1985).
- [6] S. A. KAUFFMAN: *J. Theor. Biol.*, **22**, 437 (1969).
- [7] S. A. KAUFFMAN: in *Current Topics in Developmental Biology*, Vol. **6**, Edited by A. MOSCANA and A. MONROY (New York, N.Y., 1971).
- [8] S. A. KAUFFMAN: *J. Theor. Biol.*, **44**, (1974); *Physica D*, **10**, 145 (1984).
- [9] F. FOGELMAN SOULIE: *Discrete Appl. Math.*, **9**, 139 (1984).
- [10] A. E. GEL'FAND and C. C. WALKER: *Ensemble Modeling* (Marcel Dekker Inc., New York., N.Y. and Basel, 1984).
- [11] R. A. SHERLOCK: *Bull. Math. Biol.*, **41**, 687, 707 ((1979).
- [12] S. A. KAUFFMAN: preprint (1985).