Phase transitions in random networks: Simple analytic determination of critical points

Bartolo Luque [1,2] and Ricard V. Solé [1,3]
[1] *Complex Systems Research Group, Departament de Fisica i Enginyeria Nuclear, Universitat Politécnica de Catalunya, Sor Eulalia d'Anzizu s/n, Campus Nord, Mòdul B5, 08034 Barcelona, Spain*
[2] *Facultad de Matemáticas, Universidad Nacional Autónoma de Mexico, Mexico Distrito Federal, Mexico*
[3] *Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501*
(Received 9 July 1996)

The standard method for determining the critical points in random networks (such as Boolean nets) is based on the Derrida approximation. This method leads to the critical points based on the so-called *annealed* approximation. In this paper, we present a very simple annealed approximation based on the study of damage spreading when single elements are modified. Several examples are analyzed. [S1063-651X(97)08401-8]

PACS number(s): 05.40.+j

Complexity seem to appear, very often, close to critical points [1,2]. At such points, several characteristic features such as fractal structures, $1/f$ noise, and optimum information transfer spontaneously emerge. In this context, the existence of universal behavior close to critical points justifies the use of very simple, generic models. At the critical boundary, important features of large-scale phenomena are roughly insensitive to the particular details of the models and are shared by very (apparently) different systems [3].

A theoretical approach to a wide class of complex systems was provided by the introduction of random Boolean networks (RBN's), also called Kauffman nets [4–8]. First introduced by Kauffman, a set of N binary elements $\mathbf{S}(t)=(S_1(t),\ldots,S_N(t))$, with

$$S_i(t)\in\Sigma\equiv\{0,1\}$$

$(i=1,\ldots,N)$, is updated by means of the the dynamic equations

$$S_i(t+1)=\Lambda_i(S_{i_1}(t),S_{i_2}(t),\ldots,S_{i_K}(t)). \tag{1}$$

Such dynamical systems share several properties with the so-called cellular automata models [9], but here randomness is introduced at several levels. Each automaton is randomly connected with exactly K others that send inputs to it. Here Λ_i is a Boolean function also randomly choosen from a set $\mathcal{F}_K$ of all the Boolean functions with connectivity K. An additional source of randomness is introduced through the random choice of the initial condition $\mathbf{S}(0)\equiv\{S_i(0)\}$, taken from the set $\mathcal{C}(N)$ of Boolean N strings. In spite of these random choices, RBNs exhibit a critical transition at a given connectivity $K_c=2$. For $K<K_c$, a *frozen phase* is observed and for $K>K_c$ a *chaotic phase* appears [7,8]. At the critical point, a small number of attractors $[\approx O(\sqrt{N})]$ is observed, which show high homeostatic stability (i.e., high stability against minimal perturbations in single elements or Boolean functions) and low reachability among different attractors [7,8]. These properties are clearly observed in some complex systems as the genome [5,6,10].

This critical point was first estimated through numerical simulations [5,6] and later analytically obtained [11,12] by means of the so-called Derrida annealed approximation the (DAA). Roughly, DAA considers two random initial configurations $\{S_i^{(1)}(t)\},\{S_i^{(2)}(t)\}\in\mathcal{C}(N)$, which have a given initial (normalized) overlap

$$a_{12}(t)=\frac{1}{N}\sum_{i=1}^{N}\Theta(S_i^{(1)}(t)-S_i^{(2)}(t)) \tag{2}$$

[with $\Theta(z)=1$ if $z=0$ and zero otherwise]. Then, using Eq. (1), we update the system once. The new overlap $a_{12}(t+1)$ is then computed. Then a new set of connections and Boolean functions are again chosen at random (i.e., the dynamics is annealed) and we follow the dynamics of the overlap in time. In such a way, each time step each unit receives exactly the same number of inputs, though the neighborhoods and specific interactions are modified.

It can be shown [10,11] that it evolves following the one-dimensional nonlinear map

$$a_{12}(t+1)=\tfrac{1}{2}[1+a_{12}^K(t)]. \tag{3}$$

If the normalized Hamming distance $D_t\equiv 1-a_{12}(t)$ is used, the fixed point $D^*=0$ becomes unstable at $K_c=2$. This method has been successfuly used in generalizations of RBNs [10]. It can be extended to more complex situations (see below) and, after some tedious algebraic manipulations, the critical points are obtained.

In this paper, a much more simple analytic procedure for calculating critical points in random networks will be introduced. It is based on a different view of how pertubations propagate, starting from a change in the state of a given element. We study the set of outputs starting from such a unit and, using an annealed approximation, we follow the propagation of changes through the network.

To be more specific, let us consider the following case. The standard RBN is used, but involving a distribution of connections $f(K_i)$ $(K_i=1,2,\ldots,K_m)$, i.e.,

$$\sum_{K_i=1}^{K_m} f(K_i)=1,$$

and so a mean connectivity will be

$$\langle K\rangle = \sum_{K_i=1}^{K_m} K_i f(K_i).$$

Additionally, a bias p in the sampling of Boolean functions will be used, that is to say the probability

$$p \equiv P[\Lambda_i(S_{i_1}(t), S_{i_2}(t), \ldots, S_{i_K}(t)) = 1].$$

Now the underlying dynamical system has to be generalized to

$$S_i(t+1) = \Lambda_i(S_{i_1}(t), S_{i_2}(t), \ldots, S_{i_{K_i}}(t)) \tag{4}$$

(i.e., each S_i receives $K_i \in \{1,2,\ldots,K_m\}$ inputs) and the Boolean functions Λ_i are randomly chosen from the set

$$\mathcal{S}(K_m, p, \Sigma) = \bigcup_{K_i=1}^{K_m} \mathcal{F}_{K_i,p}(N, \Sigma).$$

Here $K_m \in \mathbf{N}$ is the maximum connectivity. Following the DAA, it can be shown [10] that the equation for the overlap evolution is now a more complicated one

$$a_{12}(t+1) = 2p(1-p)\left[1 + \sum_{K_i=1}^{K_m} f(K_i) a_{12}^{K_i}(t)\right]. \tag{5}$$

Now the critical curve on the parameter space $(p, \langle K\rangle)$ is

$$\langle K\rangle = G(p) = \frac{1}{2p(1-p)}. \tag{6}$$

For $\langle K\rangle > G(p)$ a chaotic phase is reached and a frozen one otherwise. For $p = 1/2$ it reduces to the standard RBN problem.

This result can be obtained in a different way. We consider an annealed network with input connectivity $\langle K\rangle_i$. Obviously, the output connectivity $\langle K\rangle_o$ will be the same (so we have $\langle K\rangle_o = \langle K\rangle_i = \langle K\rangle$). Now let us build up a tree representation of our annealed dynamics. We start from a single unit $S_i(t) \in \mathbf{S}$, which sends $\langle K\rangle$ *outputs* to other units, which in turn send $\langle K\rangle$ outputs to other units. Here each tree level will correspond to the network states at different time steps (Fig. 1). This tree give us a picture of the possible paths followed by a change in S_i. These paths are shown by arrows reaching a set of units $\{S_i^1, S_i^2, \ldots, S_i^{K_i}\}$. From the point of view of graph theory, we are constructing a Cayley tree (i.e., a Bethe lattice) with coordination number $z = \langle K\rangle$.

Intuitively, we know that in the frozen phase such a change does not propagate through the network (in other words, damage does not spread through the system). At the chaotic phase, a change in a single unit can generate an avalanche of changes through all the net. Let us consider a given tree and now let us consider exactly the same tree, but with a single change in S_i at t (i.e., a nonnormalized $D = 1$ Hamming distance).

Two possible changes can be introduced: $S_i = 1 \to 0$ or $S_i = 0 \to 1$. This change (damage) can modify the state of one or more units. Let us consider the unit S_i^j After one update, this unit will be in state $S_i^j = 1$ with probability p (as defined previously) and a change in S_i will modify S_i^j towards $S_i^j = 1$ with probability $1-p$. The complementary case is trivially obtained. So, taking into account both cases, the probability of change in S_i^j is $P(S_i^j = 0 \to 1 \vee S_i^j = 1 \to 0) = 2p(1-p)$. If all outputs are considered (we have, on average, $\langle K\rangle$ outputs), at least one change will occur if $\langle K\rangle 2p(1-p) \geq 1$, which leads to the critical condition

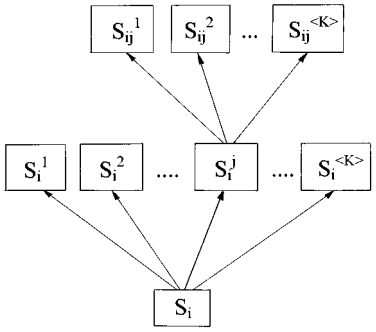

FIG. 1. Tree structure used in our annealed approximation. The unit S_i is modified and this change can (or cannot) propagate through some of the $\langle K\rangle$ units that receive inputs from S_i. This process is repeated using the annealed approach described in the text. Eventually, a single change can percolate through the whole tree.

$$\langle K\rangle = \frac{1}{2p(1-p)}, \tag{7}$$

which is the same as before, though now just a simple probabilistic argument has been used.

Our second example involves a network where now the units can take a wider set of S integer values, i.e., $S_i \in \Sigma \equiv \{0,1,\ldots,S-1\}$. For this case, we consider a mean connectivity $\langle K\rangle$ and a given distribution of probability for our states.

For simplicity, let us assume that $p = P[\Lambda_i(\mathbf{S}) = 0]$ and the other states have the same probability, i.e., $p(S \neq 0) = \{1 - P[\Lambda_i(\mathbf{S}) = 1]\}/(S-1)$. For a mean connectivity $\langle K\rangle$, it can be shown that the Derrida approximation leads, after some algebra, to the overlap equation

$$a_{12}(t+1) = a_{12}^{\langle K\rangle}(t) + \left[p^2 + \frac{(1-p)^2}{S-1}\right][1 - a_{12}^{\langle K\rangle}(t)] \tag{8}$$

and the marginal stability condition gives the critical point relation

$$p_c = \frac{1}{S}\left[1 - \left(1 - \frac{S}{K}[(2-S)K + S - 1]\right)^{1/2}\right], \tag{9}$$

which defines a critical surface in the (S,K) parameter space (Fig. 2). For $S = 2$ (Boolean net), Eq. (9) leads to

$$p = \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{2}{K}} \tag{10}$$

and is the corresponding critical condition given by Eq. (7).

Now let us consider our approach. Let us take a unit $S_i \in \Sigma$ and let us switch its state. Remember that

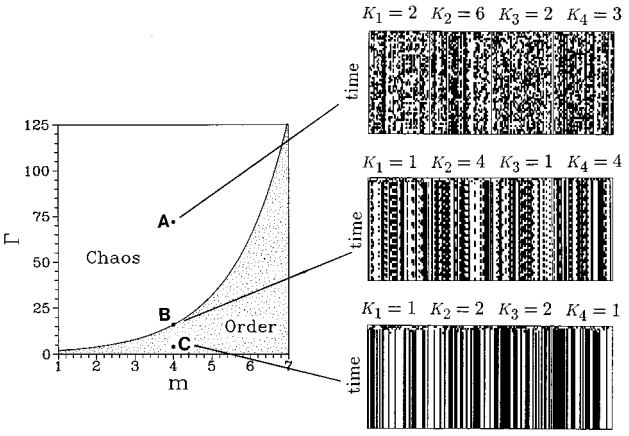

FIG. 2. Phase space of CRBNs. The two phases are observed, separated by the critical condition given by Eq. (19). Three numerical examples are also shown corresponding to three different situations. Here there are $m=4$ coupled nets, with $N=40$ units. We have three typical behaviors: (a) chaotic, (b) critical, and (c) frozen. The connectivities of each net are indicated.

$P(S_j=0)=p$ and $P(S_j\in\Sigma-\{0\})=(1-p)/(S-1)$. A change in S_i will modify a given S_i^j with probability $(1-p)[1-(1-p)]/(S-1)$ if $S_i^j=0$ and with probability $(1-p)^2(S-2)/(S-1)$ if $S_i^j\in\Sigma-\{0\}$. So no changes will occur if

$$\Gamma^*\equiv\langle K\rangle\left[2p(1-p)+\frac{S-2}{S-1}(1-p)^2\right]<1 \tag{11}$$

and the critical point will be given by $\Gamma^*=1$. It can be easily shown that this result is the same as the one given by Eq. (9).

Now let us consider a third example: coupled RBNs (CRBN). Though this is a more sophisticated system, some types of CRBNs have been used as models of cooperative games [1]. Here a simpler dynamics is considered. Let us take a couple of RBNs with connectivities K_1 and K_2, respectively. Let $\mathbf{S}_1=(S_1^1(t),\ldots,S_N^1(t))$ and $\mathbf{S}_2=(S_1^2(t),\ldots,S_N^2(t))$ be the states of such nets at t. Then the dynamics is defined as

$$S_i^1(t+1)=\Lambda_i^1(S_{i_1}(t),\ldots,S_{i_{K_1}}(t)) \tag{12}$$

(i.e., we update the first net once). Then we use $\mathbf{S}_1(t+1)$ as input for $\mathbf{S}_2$, i.e., $\mathbf{S}_2(t+1)=\mathbf{S}_1(t+1)$, and then the second net is updated

$$S_i^2(t+2)=\Lambda_i^2(S_{i_1}(t+1),\ldots,S_{i_{K_1}}(t+1)) \tag{13}$$

and we use this state as input for $\mathbf{S}_1$. Then this rule is applied again.

The DAA leads, for this system, to the following coupled equations for the normalized Hamming distance:

$$D_{t+1}=\tfrac{1}{2}[1-(1-D_t)^{K_1}], \tag{14a}$$

$$D_{t+2}=\tfrac{1}{2}[1-(1-D_{t+1})^{K_2}]. \tag{14b}$$

Using the composition $D_{t+2}=D_{t+1}(D_t)$, where again $D^*=0$ is the fixed point, we get the marginal stability condition

$$\left(\frac{\partial D_{t+2}}{\partial D_t}\right)_{D^*=0}=\frac{1}{4}K_1K_2=1, \tag{15}$$

which leads to the critical point $K_1K_2=4$.

This development is easily extended to a set of m coupled nets, with connectivities $K_1,\ldots,K_m$ (again $D^*=0$ is a fixed point). Let us assume that, for $m-1$ nets, we have

$$\left(\frac{\partial D_{t+m-1}}{\partial D_t}\right)_{D^*=0}=\frac{1}{2^{m-1}}\prod_{i=1}^{m-1}K_i^{m-1}, \tag{16}$$

then, for m nets we get

$$\left(\frac{\partial D_{t+m}}{\partial D_t}\right)_{D^*=0}=\left(\frac{\partial D_{t+m}}{\partial D_{t+m-1}}\frac{\partial D_{t+m-1}}{\partial D_t}\right)_{D^*=0}$$

$$=\frac{1}{K_m}\frac{1}{2^{m-1}}\prod_{i=1}^{m}K_i^m, \tag{17}$$

i.e., we have a general result

$$\left(\frac{\partial D_{t+m}}{\partial D_t}\right)_{D^*=0}=\frac{1}{2^m}\prod_{i=1}^{m}K_i^m \tag{18}$$

so the critical point for m coupled nets is simply

$$\prod_{i=1}^{m}K_i=2^m. \tag{19}$$

This result can be extended to nets with a distribution of biases $\{p_j\}$. The critical point is now given by

$$\prod_{i=1}^{m}K_ip_i(1-p_i)=\frac{1}{2^m}. \tag{20}$$

Using our method, now we have m trees and each time step we move from one tree to the next one. A percolating tree in the coupled system implies percolation in each tree. This is a set of independent events (in the annealed approximation) and we have stability (a frozen regime) if:

$$\Gamma^*\equiv 2p(1-p)K_1\cdots 2p(1-p)K_m<1, \tag{21}$$

which is, at $\Gamma^*=1$, the result of Eq. (20). The phase space for these coupled nets is shown in Fig. 2. We also show, for $m=4$, three examples of the dynamics, using $N=40$ networks.

Finally, let us consider a different type of random network: a discrete neural network with asymmetric connections [13]. The dynamic equation is given by a standard neural network equation

$$S_i(t+1)=\mathrm{sgn}\left(\sum_j C_{ij}S_j(t)+h\right), \tag{22}$$

with $i=1,\ldots,N$ neurons and $S_i\in\{0,1\}$. Here h is a given threshold and each neuron receives exactly K inputs

$(1 \leqslant K \leqslant N)$. The weigth of the connections C_{ij} are randomly choosen from a given distribution $\rho(C_{ij})$. If neuron j sends no input to neuron i, C_{ij} is set equal to zero. Kurten [13] has shown that such a neural network has two different phases (in the thermodynamic limit): the frozen one and the chaotic one, as in our previous RBN. This network was shown to share behavioral similarities with the Kauffman model, pointing to the existence of formal relationships.

In the analytic treatment by Kurten, he uses the formal expression for the Hamming distance in terms of a polynomial spline function of K order:

$$D_{t+1}(K) = \sum_{\nu=1}^{K} (-1)^{\nu+1} \binom{K}{\nu} a_\nu D_t^\nu(K), \tag{23}$$

with

$$a_\nu = 1 + \sum_{m=1}^{\nu} (-1)^m \binom{m}{\nu} I_{Km}, \tag{24}$$

with $\nu = 1,2,\ldots,K$ and I_{km} is given by the integral

$$I_{Km}(\rho,h) = \int \cdots \int dx_1 \cdots dx_K \rho(x_1) \cdots \rho(x_K) \times \Theta((x_{m+1} + \ldots + x_K + h)^2 - (x_1 + \ldots + x_m)^2),$$

with $\Theta(x)$ the Heaviside function. The phase transition point for $m = 1$ is determined by the nonlinear map [13]

$$S_K(h) = \left(\frac{\partial D_{t+1}(K)}{\partial D_t(K)} \right)_{D(K)=0} = K[1 - I_{K1}(h)] = 1.$$

Using our method, let us consider the change of a single unit $S_j(t) \to -S_j(t)$ and let us see if such a change percolates through the tree. $I_{K1}(h)$ is the probability that a change in a single unit does not modify the sign of $\Sigma_j C_{ij} S_j(t) + h$. We want to obtain the critical condition, which, in our approach, is directly given by the condition

$$K[1 - I_{K1}(h)] = 1$$

where I_{k1} is defined by the previous integral.

To sum up, we have used an analytic method of determination of critical points in random networks. It has been shown that the method exactly reproduces the classical results obtained from the DAA. Several examples, involving S-state random nets, coupled random nets and asymmetric neural networks have been presented.

Our method is, in fact, equivalent to the Derrida approximation. The DAA starts from two annealed nets, with identical dynamics and two initial conditions with a given Hamming distance D_t. The dynamical equation for such a distance is obtained, and the existence of two qualitative dynamical regimes is demonstrated by depending of the connectivity K. If D_t goes to zero assymptotically, we are in the frozen regime. Otherwise, we are in the chaotic domain. In our method, we also start from two annealed nets with identical dynamics. Now, however, the Hamming distance among them is the minimum one: only of a unit. In the ordered regime, such a perturbation will disappear. In the chaotic one, it will be amplified. By considering the output neighbor trees, starting from the perturbed unit, we follow the propagation of this change. At each tree level we check if the distance is finite or zero. In this way, our approach can be understood in terms of percolation in Bethe's lattice.

We would like to thank Stuart Kauffman, Andy Wuensche, Jordi Delgado, Susanna C. Manrubia, Jordi Bascompte, Pedro Miramonte, Octavio Miramonte, and Brian Goodwin for useful discussions. This work has been partially supported by DGYCIT through Grant No. PB-94 1195. R.S. thanks the Santa Fe Institute for financial support and B.L. was supported by an Intercampus Grant.

[1] S. A. Kauffman, *The Origins of Order* (Oxford University Press, Oxford, 1993).

[2] R. V. Solé, S. C. Manrubia, B. Luque, J. Delgado, and J. Bascompte, Complexity **1** 13 (1996). A possible scenario for reaching critical points was suggested by P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. A **38**, 364 (1988).

[3] J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. E. J. Newman, *The Theory of Critical Phenomena: An Introduction to the Renormalization Group* (Clarendon, Oxford, 1995).

[4] S. A. Kauffman, *At Home in the Universe* (Oxford University Press, Oxford, 1995).

[5] S. A. Kauffman, J. Theor. Biol. **22**, 437 (1969).

[6] S. A. Kauffman, J. Theor. Biol. **44**, 167 (1974).

[7] S. A. Kauffman, Physica D **10**, 145 (1984).

[8] S. A. Kauffman, Physica D **42**, 135 (1990).

[9] S. Wolfram, Physica D **10**, 1 (1984).

[10] R. V. Solé and B. Luque, Phys. Lett. A **196**, 331 (1995). For a recent review of the application of RBNs to the genome, see R. Somogyi and C. A. Sniegoski, Complexity **1**, 45 (1996). Though the genome regulation involves, in fact, continuous concentrations of RNA and proteins, the high nonlinearities implicit in the gene-gene interactions leads to a basically on-off response. Even if a continuous model of gene interaction is considered, a nearly binary response is reached. This is, for example, what we get in the two-gene model of the λ-phage behavior, where two proteins interact in such a way that they show mutual inhibition. A continuous model can be easily derived [see D. Kaplan and L. Glass, *Understanding Nonlinear Dynamical Systems* (Springer, New York, 1995)] leading to two attractors where one of the concentrations is high and the other is very low. This happens through a symmetry breaking and can be easily modeled by a $N = 2$ Boolean net. However, as the size of the system grows, continuous models give us little useful information (in particular, no general macroscopic properties are shared with the real genome). Actually, RBNs are the only successful model leading to predictable properties.

[11] B. Derrida and Y. Pomeau, Europhys. Lett. **1**, 45 (1986).

[12] B. Derrida and Y. Pomeau, Europhys. Lett. **2**, 739 (1986).

[13] K. E. Kurten, Phys. Lett. A **129**, 157 (1992).