# Dynamics of Complex Systems: Scaling Laws for the Period of Boolean Networks

Réka Albert and Albert-László Barabási*

*Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556*

Boolean networks serve as models for complex systems, such as social or genetic networks, where each vertex, based on inputs received from selected vertices, makes its own decision about its state. Despite their simplicity, little is known about the dynamical properties of these systems. Here we propose a method to calculate the period of a finite Boolean system, by identifying the mechanisms determining its value. The proposed method can be applied to systems of arbitrary topology, and can serve as a roadmap for understanding the dynamics of large interacting systems in general.

PACS numbers: 84.35.+i, 05.40.–a, 64.60.Ak, 87.23.Ge

Microscopic and spatial order, as well as compositional homogeneity, being intrinsic properties of numerous materials, have been the source of spectacular advances in many branches of contemporary physics. However, we often encounter systems that are composed of *nonidentical* units with *different functions*, between which the interaction is often *long range* and *random*. For example, living systems are often viewed as genetic networks [1], whose vertices are macromolecules with different functions, connected through links of chemical origin. Similarly, the society is a web of individuals interacting with a selected group of other individuals [2]. The economy can be viewed as a complex web of companies with different interests that are linked by diverse business relationships [3]. Despite the rather diverse functions the vertices have in these examples, each shows a surprising degree of self-organization that allows their continued functioning. These examples, far from being exhaustive, offer a paradigm different from that so successfully addressed by physical sciences: many complex systems are composed of units (vertices) with diverse and distinct functions that are connected in a random fashion to other vertices in the system. Understanding the emergence of order in these complex systems is a formidable challenge to statistical mechanics.

A first step in this direction has been the introduction of Boolean networks [4] that allow the "individuals" (vertices) to follow different agendas. The Boolean model (Fig. 1) consists of $N$ vertices that are characterized by Boolean variables (spins) that can take up the values $\sigma_i = 0$ or 1. Each vertex receives input from $K_i$ selected vertices. The value of the Boolean variable at every time step is updated according to the rule $\sigma_i(t + 1) = \mathcal{B}_i(\sigma_1(t), \sigma_2(t), \ldots, \sigma_{K_i}(t))$ where $\mathcal{B}_i$ is a randomly chosen Boolean function, predetermined and fixed for each vertex. While in general $K_i$ can vary, to decrease the complexity of the problem, Kauffman [5] proposed a model in which each vertex is connected to exactly $K$ randomly selected vertices. The dynamical properties of the model are determined both by $K$ [5] and the bias $Q$ of the Boolean law [6–8] (see Fig. 1b). For fixed $K(>2)$ the model displays a phase transition from an ordered or "frozen" state

for $Q > Q_c$, in which the system breaks into isolated islands, to a "chaotic" state for $Q < Q_c$, where the system behaves as a highly connected single random net, and a perturbation in the state of a single vertex propagates over the whole system [6,7].

An important quantity characterizing the dynamics of a system of $N$ vertices is the total period $T(N)$. The system has $2^N$ possible states, thus the dynamics follows a periodic orbit with a maximum period $2^N$. Despite the numerous attempts to determine $T(N)$ as a function of $N$ [5,7,9,10], success has been achieved only for the fully connected case ($K = N$) [10]. Numerical investigations have indicated that in the ordered phase the period depends on a power of $N$, while in the chaotic phase this dependence is exponential [5,7]. However, due to the large fluctuations, the conclusive nature of these numerical results was limited.
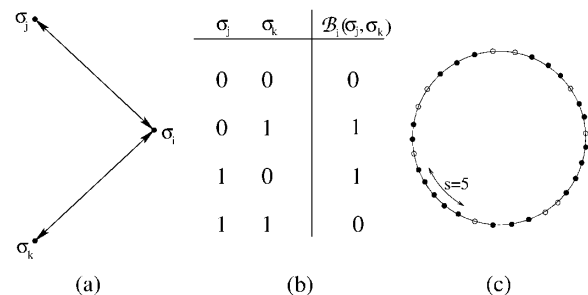


FIG. 1. In a Boolean system each vertex receives inputs from $K$ vertices. The spins of these neighbors can take up $2^K$ different configurations. The Boolean function $\mathcal{B}_i$ assigns each of these $2^K$ inputs an output value of 0 or 1. The outputs are selected randomly, choosing with probability $Q$ the value 0 and with probability $(1 - Q)$ the value 1. The Boolean table determines at any time the spin of vertex $i$, based on the spins of its $K$ neighbors. In the figure we show an example for $K = 2$, where vertex $i$ receives four different inputs from two neighbors, $j$ and $k$ (a). Table (b) defines $\mathcal{B}_i$, by assigning an output [right column, $\mathcal{B}_i(\sigma_j, \sigma_k)$] to each input $\sigma_j$ and $\sigma_k$ (two left columns). The 0 and 1 values in the right column have been selected randomly. (c) Illustration of clustering in the one-dimensional system. The active vertices (filled circles), whose spin fluctuates in time, form clusters of different sizes $s$, separated by frozen spins (open circles), whose value is independent of time.

Thus, despite numerous numerical and analytical attempts, a quantitative understanding of the dynamical properties of the Boolean systems is still missing. Here we take an important step towards filling the gap: we identify the factors determining the period $T(N)$ of a Boolean system, allowing us to explicitly predict $T(N)$ for the first time. We find that $T(N)$ can have a rather complex functional form, determined by the interplay between the system's complex microscopic properties and its topology. Our predictions for $T(N)$ are in very good agreement with the numerical results, and help to explain the uncertainties in earlier numerical studies. Furthermore, we propose a method that helps uncovering the period $T(N)$ for systems of arbitrary topology, and can serve as a roadmap for the investigation of similar complex systems.

*The one-dimensional system (1D).*—To uncover the main factors determining the period, we first study a 1D system that has the simplest topology which nevertheless gives nontrivial dynamics. In 1D the vertices form a circle, so every vertex is connected to two neighbors ($K = 2$). For a given $Q$ the system is broken into one-dimensional active clusters that are separated by inactive sites [5,9] (see Fig. 1c). The origin of this clustering lies in the fact that a fraction of the possible Boolean laws $\mathcal{B}_i$ result in the freezing of the spin $\sigma_i$, such that its value does not change with time. Obvious examples are the tautology ($\sigma_i = 1$ for any input) and the contradiction ($\sigma_i = 0$ for any input), but spins can freeze for other $\mathcal{B}_i$ as well, depending on the dynamics of their neighbors. After a transient period only a fraction $P$ of the vertices is active. Each active cluster of size $s$ fluctuates *independently* with a period $T^i(s) \leq 2^s$. Consequently, the total period $T(N)$ is the least common multiplier of the periods of the individual clusters. Thus, in order to determine the expectation value of $T(N)$, we need to know the size distribution of the active clusters, $F(s)$, and the expected dependence of the individual cluster periods on their size, i.e., the function $T(s)$. Whether a vertex $i$ is active or frozen depends on the updating rule $\mathcal{B}_i$, and on the dynamics of the neighboring spins $\sigma_{i-1}$ and $\sigma_{i+1}$. If the vertices were to be independent, given the probability $P$ of a site to be active, percolation theory [11] predicts that in 1D the fraction of active clusters of size $s$ is $F(s) = [(1 - P)/P^2] \exp(-cs)$, where $c = -\ln(P)$. As Fig. 2a indicates, the numerical results are in qualitative agreement with this prediction, having the form

$$F_{\text{num}}(s) \sim \exp[-c(Q)s], \qquad (1)$$

with $c(Q)$ diverging as $Q$ increases. Since $F(s)$ is exponential, the average cluster size $\langle s \rangle$ does not depend on the system size, thus the average number of clusters is $n = n(Q) \propto N$. A special role is played by the largest cluster in the system, with size $s_{\max}$. The probability that $s_{\max}$ is smaller than a given value $x$ is $\text{Prob}(s_{\max} < x) = [\text{Prob}(s < x)]^n = [1 - \sum_x^N F(s)]^n$. For large $n$, $s_{\max}$ is the solution of the equation
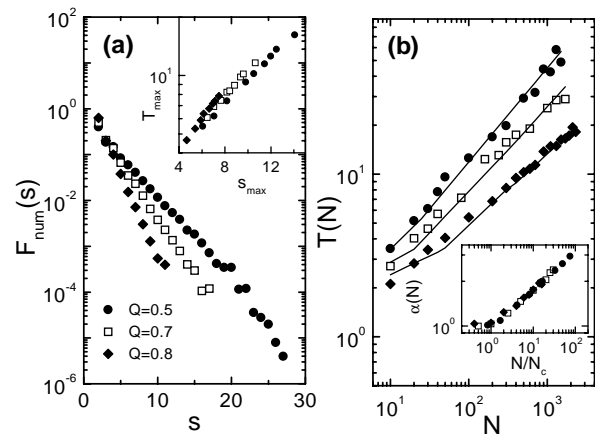


FIG. 2.   (a) Size distribution of the active clusters in the 1D system for three values of $Q$. Inset: The dependence between the size $s_{\max}$ and the period $T_{\max}$ of the largest active cluster. (b) Comparison between the prediction of Eq. (6) (solid line) and the numerically obtained period. Inset: $N$ dependence of the proportionality factor between $T(N)$ and $T_{\max}$. In all simulations we average over 200 realizations of Boolean functions, and 20 initial spin configurations. After a transient of 5000 time steps, we identify the active vertices by studying their spin values for another 5000 steps.

$$\sum_{s_{\max}}^{N} F(s) = \frac{1}{n}. \qquad (2)$$

Substituting (1) in (2), we obtain

$$s_{\max} = 2 + \frac{\ln n(Q)}{c(Q)}, \qquad (3)$$

and, since the average cluster number $n \propto N$, $s_{\max}$ depends logarithmically on $N$.

The period $T(N)$ is the least common multiplier of the periods of the individual clusters, $T^i(s)$. The most dominant contribution to $T(N)$ is given by the largest individual period, $T_{\max}$. We expect this to be the period of the largest dynamic cluster, $s_{\max}$ (we checked this numerically). We have determined the scaling of $T_{\max}$ with $s_{\max}$ numerically, obtaining

$$T_{\max} \sim \exp[\gamma(Q)s_{\max}], \qquad (4)$$

where $\gamma(Q)$ increases with $Q$ (see the inset of Fig. 2a).

For small systems there are only a few clusters and $T_{\max}$ is identical to the least common multiplier of the individual periods, $T(N)$. But the number of clusters increases with $N$, and the probability that clusters different from $s_{\max}$ contribute to $T(N)$ also increases. Thus $T(N)$ is a product of $T_{\max}$ and a factor $\alpha(N)$ that accounts for the appearance of smaller clusters with a period different from the divisors of $T_{\max}$, i.e., $T(N) = \alpha(N)T_{\max}$. As the inset of Fig. 2b illustrates, we find that $\alpha(N)$ follows

$$\alpha(N) \sim \begin{cases} 1 & \text{if } N \ll N_c(Q), \\ N^\beta & \text{if } N \gg N_c(Q), \end{cases} \qquad (5)$$

where $\beta = 0.25$ and $N_c(Q)$ increases with $Q$.

Finally, combining Eqs. (3)–(5) we obtain

$$T(N) = \begin{cases} C_1 N^{[\gamma(Q)/c(Q)]} & \text{if } N \ll N_c(Q), \\ C_2 N^{\beta + [\gamma(Q)/c(Q)]} & \text{if } N \gg N_c(Q), \end{cases} \quad (6)$$

where $C_1$ and $C_2$ are the products of the constants appearing in Eqs. (3)–(5). Figure 2b compares Eq. (6) and the numerically obtained $T(N)$. The fit, which contains no free parameters, is rather good for all $Q$ values, underlying the applicability of our method. Since in the 1D system percolation can be achieved only for $P = 1$, the system is in the ordered regime for any $Q$. Thus Eq. (6) agrees with numerical findings that in the ordered regime the period has a power-law $N$ dependence [5].

The agreement demonstrated in Fig. 3 indicates that $T(N)$ is determined by three factors: (a) the clustering of the vertices into islands of active vertices isolated by frozen vertices, quantified by $F(s)$, which determines $s_{max}$; (b) the relationship between the size and the period for a


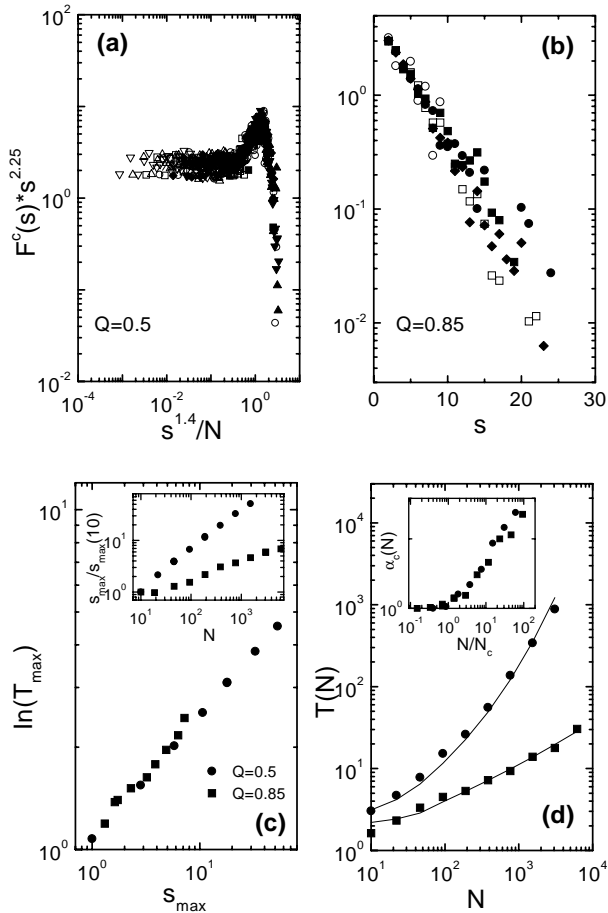
FIG. 3.   Size distribution of the active clusters for the Cayley tree in the (a) chaotic and the (b) ordered phases for $N = 10-3070$. (c) Dependence of $T_{max}$ on $\eta \, s_{max}$ for the Cayley tree, where $\eta$ is defined in [14]. The functional form is independent of $Q$. Inset: Dependence of $s_{max}$ on $N$, following a power law. (d) Comparison between the prediction (11) (solid line) and the numerical results for the Cayley tree. The behavior of $T(N)$ is very different for the two $Q$ values, indicating that they correspond to different phases. Inset: $N$ dependence of $\alpha^c(N)$.

cluster, i.e., $T_{max}(s_{max})$; (c) the relationship between the periods of individual clusters and the total period of the system, i.e., $\alpha(N)$. Since each of these can be determined *independently*, we can formulate the following conjecture:

$$T(N) = \alpha(N) T_{max}[s_{max}(N)]. \quad (7)$$

At this point it is not clear whether (7) applies to systems with more complex topology. In random networks, at the percolation transition, the clusters form a loopless tree with average connectivity $K_c = 2$ [12]. Thus a good model of the cluster topology at connectivity $K_c$ is the Cayley tree, to which we apply Eq. (7) next.

*The Cayley tree.*—We consider a Cayley tree with coordination number 3, thus the sites inside the tree are connected to three neighbors, while those at the perimeter have only one connection, giving the average connectivity as $\langle K \rangle = 2 - 2/N$. Changing $Q$ induces a phase transition from the chaotic to the ordered phase [8], thus to check the validity of (7) in both of these phases, we investigated separately the system for $Q = 0.5$ (chaotic phase) and $Q = 0.85$ (ordered phase).

Following the agenda set by Eq. (7), we first determine the size distribution of active clusters, $F(s)$. Because of the rather different topology, the functional form of $F(s)$ is different from (1). Assuming that the vertices are independent, percolation theory gives [11]

$$F^c(s) = \begin{cases} s^{-\tau} f(s^\sigma/\xi) & \text{if } \xi \ll N, \\ s^{-\tau} g(s^\sigma/N) & \text{if } \xi \sim N, \end{cases} \quad (8)$$

where $\xi$ is the correlation length of the active clusters, $f(x)$ is an exponentially decaying function, and $g(x) = $ const for $x \ll 1$ and decays rapidly for $x \gg 1$.

Figures 3a and 3b present the data collapse for the distribution of active clusters, indicating that the two $Q$ values correspond to different percolation regimes: for $Q = 0.85$ the active clusters are small, following the first alternative of Eq. (8), while for $Q = 0.5$ they percolate through the system, following the second alternative. The data collapse was obtained using $\tau = 2.25$ and $\sigma = 1.4$, different from $\tau_{perc} = 5/2$, and $\sigma_{perc} = 1/4$. The origin of this difference lies in the nature of the Boolean law: the activity of a vertex depends on the dynamics of the neighboring vertices, generating correlations in the system [13]. Thus, while the functional form of $F(s)$ is correctly predicted by percolation, both in 1D and for the Cayley tree the exponents are different due to correlation between spins.

Substituting the distribution function $F^c(s)$ in Eq. (2), we obtain that $s_{max}$ has an asymptotic solution of the form

$$s_{max} \sim n^{\delta(Q)}, \quad (9)$$

where $\delta(Q)$ decreases with $Q$. The scaling of the size of the largest cluster, $s_{max}$, with $N$ follows a power law (see Fig. 3c), in contrast to the logarithmic dependence (3) obtained for the one-dimensional case.

As Fig. 3c indicates, the dependence of $T_{max}$ on $s_{max}$ follows a stretch exponential, different from the exponential

(4) in 1D,

$$T_{\max} \sim \exp[(\eta s_{\max})^\theta], \qquad (10)$$

where we find that $\eta$ and $\theta$ are independent of $Q$ [14].

Finally, $T(N)$ is proportional to $T_{\max}$, the proportionality factor $\alpha^c(N)$ having the same $N$ dependence as in the one-dimensional case (5), with $N_c(0.5) = 25$, $N_c(0.85) = 66$, and $\beta = 0.25$ (see Fig. 3d). Combining these results and using Eq. (7) we obtain

$$T(N) = \begin{cases} C_3 \exp[A(Q)N^{\delta(Q)}]^\theta & \text{if } N \ll N_c, \\ C_4 N^\beta \exp[A(Q)N^{\delta(Q)}]^\theta & \text{if } N \gg N_c, \end{cases} \qquad (11)$$

where $A(Q) = 2\eta(\frac{n(Q)}{N})^{\delta(Q)}$, and since $n(Q) \propto N$, $A(Q)$ is independent of $N$. Figure 3d compares the conjecture (11) and the numerically obtained $T(N)$. Although the dependence of $T$ on $N$ is quite involved, the agreement is remarkable. While for both $Q$ Eq. (11) has an exponential contribution, the overall behavior of the period is quite different: for $Q = 0.5$ (chaotic phase) the exponential component dominates, while for $Q = 0.85$ (ordered phase), the power-law term $N^\beta$ has the stronger influence. This explains the earlier numerical results that $T(N)$ can be approximated by a power law in the ordered regime, while it is exponential in the chaotic regime.

In conclusion, we have identified the factors determining the period of a Boolean system, showing that they can be combined into a single equation, whose components can be determined independently by a combination of analytical and numerical tools. We find that depending on the system's topology, Eq. (7) can lead to rather complex and unexpected $T(N)$ functions. Without identifying all constituents in (7), numerical investigation can easily approximate $T(N)$ with simpler functions. Indeed, this is the source of the inconclusive nature of numerical simulations that focused only on $T(N)$ [7,15]. The strength of Eq. (7) lies in the fact that it focuses our attention on the *factors determining the period*, thus offering a *roadmap* for the investigation of the dynamics of the Boolean networks. We hope that this roadmap will induce further research towards understanding the universality and the origin of the components that, combined together, determine the dynamics of the Boolean networks. It also indicates that the period, due to its convoluted character, is not the best quantity to investigate if one wishes to understand the phase transition in Boolean systems. Rather, investigating the changes in the components of $T$, as given by Eq. (7), might give better measures of changes in the system as the transition point is approached. Furthermore, Eq. (7) will be helpful in addressing the dynamics of randomly connected Boolean nets, a more realistic model of complex systems. Indeed, we have recently performed simulations where spins were placed on a random network [12], finding that following the decomposition outlined in this paper we could obtain excellent agreement between the total period and Eq. (7) [16]. Finally, it would be rather important to investigate how will the period of the Boolean systems change if one uses as support more realistic networks, like scale-free [17] or directed networks.

*Email address: alb@nd.edu

[1] U. Alon, M. G. Surette, N. Barkai, and S. Leibler, Nature (London) **397**, 168 (1999); U. S. Bhalla and R. Iyengar, Science **283**, 381 (1999).

[2] S. Wassermann and K. Faust, *Social Network Analysis* (Cambridge University Press, Cambridge, England, 1994).

[3] C. C. Walker and A. E. Gelfand, Behav. Sci. **24**, 112 (1979); A. E. Gelfand and C. C. Walker, *Ensemble Modeling* (Marcel Dekker, New York, 1984); T. Lux and M. Marchesi, Nature (London) **397**, 498 (1999).

[4] C. C. Walker and W. R. Ashby, Kybernetics **3**, 100 (1965).

[5] S. A. Kauffman, J. Theor. Biol. **22**, 437 (1969); S. A. Kauffman, *The Origins of Order* (Oxford University Press, New York, 1993).

[6] B. Derrida and D. Stauffer, Europhys. Lett. **2**, 739 (1986).

[7] G. Weisbuch and D. Stauffer, J. Phys. (Paris) **48**, 11 (1987); L. de Arcangelis and D. Stauffer, J. Phys. (Paris) **48**, 1881 (1987).

[8] B. Luque and R. V. Solé, Phys. Rev. E **55**, 257 (1997).

[9] H. Atlan, F. Fogelman-Soulie, J. Salomon, and G. Weisbuch, Cybern. Syst. **12**, 103 (1981); F. Fogelman-Soulie, Discrete Appl. Math. **9**, 139 (1984); Theor. Comput. Sci. **40**, 275 (1985).

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

[11] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, London, 1991); A. Bunde and S. Havlin, *Fractals in Science* (Springer-Verlag, Berlin, 1994).

[12] P. Erdős and A. Rényi, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960); B. Bollobás, *Random Graphs* (Academic Press, New York, 1985); J. E. Cohen, Discrete Appl. Math. **19**, 113 (1988).

[13] S. Prakash, S. Havlin, M. Schwartz, and H. E. Stanley, Phys. Rev. A **46**, R1724 (1992).

[14] Sites in direct contact with frozen vertices oscillate in phase with their active neighbors, so a better characteristic size of an active cluster is the size of the bulk, $s^b$, which in 1D is $s^b = s - 2$. For a Cayley tree the bulk $s^b$ is proportional to $s$. Numerical results indicate $s^b = \eta s$, with $\eta = 0.66$, independent of $Q$.

[15] D. Stauffer, Philos. Mag. B **56**, 901 (1987).

[16] R. Albert and A.-L. Barabási (unpublished).

[17] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999); A.-L. Barabási and R. Albert, Science **286**, 509 (1999); A.-L. Barabási, R. Albert, and H. Jeong, Physica (Amsterdam) **272A**, 173 (1999).