Basin Entropy in Boolean Network Ensembles

Peter Krawitz[1,2] and Ilya Shmulevich[1]

[1]*Institute for Systems Biology, Seattle, Washington 98103, USA*

[2]*Fakultät für Physik, Ludwig Maximilians Universität, 80799 München, Germany*

(Received 22 December 2006; published 9 April 2007)

The information processing capacity of a complex dynamical system is reflected in the partitioning of its state space into disjoint basins of attraction, with state trajectories in each basin flowing towards their corresponding attractor. We introduce a novel network parameter, the basin entropy, as a measure of the complexity of information that such a system is capable of storing. By studying ensembles of random Boolean networks, we find that the basin entropy scales with system size only in critical regimes, suggesting that the informationally optimal partition of the state space is achieved when the system is operating at the critical boundary between the ordered and disordered phases.

DOI: 10.1103/PhysRevLett.98.158701 PACS numbers: 89.75.Hc, 05.45.−a, 89.75.Da, 89.75.Fb

Introduction—Complex dynamical systems generically possess state spaces that are partitioned into disjoint basins of attraction. Within each basin, state trajectories flow to a single attractor, which, in general, may be a steady-state fixed point, an oscillation—including limit cycles—or a strange (chaotic) attractor. Such classification of states via attractors and their basins of attraction constitutes the information processing ability of a complex system and represents a type of associative memory: All states in a basin are in the same class in that they are all associated with the same attractor. The complexity of information that such a system is capable of processing depends on the manner in which its state space is partitioned, which, in turn, is largely determined by the dynamical regime of the system. By using a novel network parameter, the basin entropy, and the model class of random Boolean networks, for which there exist well-defined and extensively studied notions of ordered, critical, and chaotic dynamics [1–4], we show here that the informationally optimal partition of the state space is achieved when the system is operating at the critical boundary between the ordered and disordered regimes.

Random Boolean networks (RBNs) are commonly studied as generic models of dynamically interacting entities, such as gene regulatory networks. A RBN, originally introduced by Kauffman [5], consists of n nodes, each of which can have two possible values (0 and 1). Each node receives input from k randomly chosen nodes that determine its value at the next time step via a randomly chosen Boolean function assigned to that node. The output of the function is chosen to be 1 with probability p, known as the bias [6]. A state of the network is the collective activity of the nodes. All nodes are updated synchronously, and the network transitions from one state to another, thus tracing out a state trajectory that eventually flows into a series of periodically recurring states called an attractor. The transient states that flow into an attractor constitute the basin of attraction for that attractor.

In the limit of large n, RBNs exhibit a phase transition between a dynamically ordered and a chaotic regime. Depending on the parameters k and p, small perturbations die out over time in the ordered regime and increase exponentially in the chaotic regime (see, e.g., [4]). Networks that operate at the boundary between the ordered and the chaotic phase have been of particular interest as models for gene regulatory networks, as they exhibit complex dynamics combined with stability under perturbations [7–10]. The average lengths of attractors have been extensively studied numerically and analytically for a wide range of Kauffman networks. In the chaotic phase, the average length of attractors increases exponentially with system size [6]. In the highly chaotic case, where the state space of a RBN can be approximated by a random map, the expected number of attractors increases linearly with system size [11]. In the ordered phase, where the fraction of nodes that freeze to a constant value approaches one, the average number and length of attractors is bounded [2,12]. In contrast to former assumptions, it has been recently shown that the average number and length of attractors increases superpolynomially in critical Kauffman networks [13].

In this Letter, we introduce a new network parameter, the basin entropy h of a Boolean network, and show that this quantity increases with system size in critical ensembles, whereas it reaches a fixed value in the ordered *and* in the highly chaotic phase. The partition of the state space into basins of attraction induces a probability mass function over the state space, with the weight of each basin defined by its size relative to the other basins. The basin entropy (hereafter, simply entropy) of a network is then calculated from this probability distribution. The partition that a given network imposes on its state space is not unique and, thus, neither is its entropy. Certain network instances in the ordered phase may have the same entropy as some networks in the chaotic phase, and thus it becomes necessary to study the average entropies of network ensembles.

 © 2007 The American Physical Society

In the informational sense, the entropy of a Boolean network is a measure of the uncertainty about its dynamical behavior. In a network with higher entropy, more information is needed to describe its future behavior, and its effective complexity is higher [14]. Thus, if the network is viewed as a classifier, then the entropy is a measure for the complexity of this classification process. Herein, we show that an increase in such classification complexity as a function of system size occurs only when the system is poised at the boundary between an ordered and a disordered phase.

A Boolean network $B = (\mathbf{N}, \mathbf{F})$ is defined by the set $\mathbf{N}$ of its nodes, $\mathbf{N} = \{1, \ldots, n\}$, and the set $\mathbf{F}$ of their corresponding Boolean updating functions, $\mathbf{F} = \{f_1, \ldots, f_n\}$, with f_i: $\{0, 1\}^k \rightarrow \{0, 1\}$. The value x_i of node i at time $t + 1$ is determined by the values of its k controlling elements as $x_i(t + 1) = f_i(x_{i_1}(t), \ldots, x_{i_k}(t))$.

The sensitivity s of a network, defined as the average sensitivity of the Boolean functions used in the network [15], is an order parameter that specifies the average number of nodes that are affected by a perturbation of a random node. Thus, an average sensitivity of $s = 1$ indicates that a perturbation of a random node is on average propagated to one other node. This defines the point of phase transition between the ordered regime ($s < 1$), where perturbations die out over time, and the chaotic regime ($s > 1$), where even small perturbations increase over time [4]. In classical Kauffman networks, the relationship between the network parameters k and p and the average sensitivity s is $s = 2kp(1 - p)$ [15]. The logarithm of this parameter may also be interpreted as the Lyapunov exponent $\lambda = \log s$ [3].

In the following, we study ensembles $\mathbf{B}(n, k, p)$ of Boolean networks with n nodes parametrized by k and p. For $p = \frac{1}{2}$, all networks of the ensemble have equal probability. In the general case, where certain Boolean functions might be chosen with different probabilities, we indicate by υ_i the probability of a certain network instance i in the ensemble. We will also find it convenient to refer to the ensemble of networks $\mathbf{B}(n, s)$ with average sensitivity s.

Average basin entropy.—Any Boolean network $B \in \mathbf{B}(n, s)$ partitions its state space into attractors and corresponding basin states. The weight w_ρ of an attractor ρ is the length of its attractor plus the number of basin states draining into that attractor, normalized by the size of the state space (2^n), so that $\sum_\rho w_\rho = 1$. The basin entropy h of a network B is defined as

$$h(B) = -\sum_\rho w_\rho \ln w_\rho. \tag{1}$$

The average entropy $\langle h \rangle$ of an ensemble $\mathbf{B}(n, s)$ of networks is defined as

$$\langle h \rangle[\mathbf{B}(n, s)] = -\sum_{i \in \mathbf{B}(n,s)} \upsilon_i \sum_{\rho_i} w_{\rho_i} \ln w_{\rho_i}, \tag{2}$$

where υ_i is the probability of a network instance in the ensemble. To determine the state space partition of a network exactly, one has to link each state to its attractor. We performed an exhaustive computation of the state space partition to estimate the average entropy of the ensembles $\mathbf{B}(n, k, p = 1/2)$, with $n = 10, \ldots, 20$ and $k = 1, \ldots, 10$ (for every ensemble, the average was taken over more than 1000 networks). As shown in Fig. 1, the average entropies of the critical ensembles $k = 2$ grow with n, whereas the average entropies of the chaotic ensembles $k > 2$ approach a finite value of approximately 0.61, independent of n. Before considering the scaling behavior of the average entropy in critical network ensembles $\mathbf{B}(n, s = 1)$, we will discuss the average entropy in the chaotic regime in the limit of large n.

When k reaches n, the ensemble $\mathbf{B}(n, k = n, p = 1/2)$ can be identified with the random map model, which is a simple disordered system with deterministic dynamics. For each state in the state space, another state (not necessarily a different one) is randomly chosen as its successor. Thus, a random map can be regarded as an unbiased Boolean network where each node depends on all n variables. The average sensitivity of the random map is thus $s = n/2$, and its average entropy can be expressed as

$$\langle h \rangle[\mathbf{B}(n, s = n/2)] = -\sum_{j=1}^{2^n} g_j w_j \ln w_j, \tag{3}$$

where the sum is taken over all possible weights $w_j \in \{1/2^n, 2/2^n, \ldots, 1\}$, multiplied by the normalized frequency g_j of their occurrence in the ensemble $\mathbf{B}(n, s = n/2)$.

As $n \rightarrow \infty$, we may replace the sum by an integral, where $g(w)\, dw$ indicates the average number of attractors with a weight between w and $w + dw$:

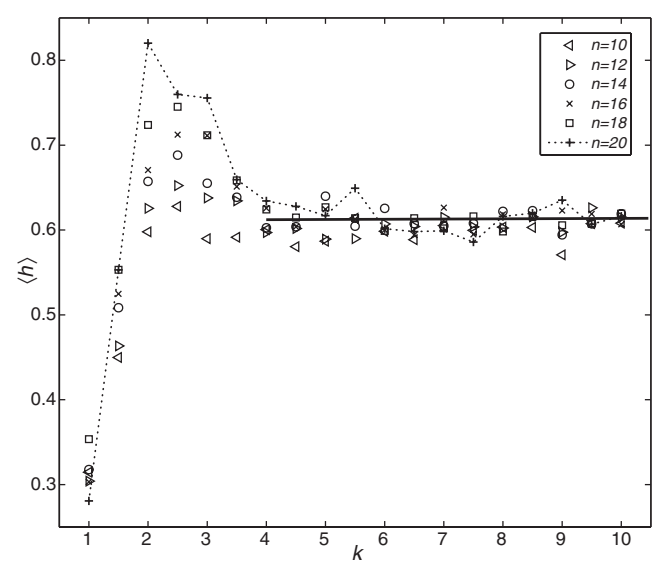

FIG. 1. Average entropy in ensembles $\mathbf{B}(n, k, p = 1/2)$ with $n = 10, \ldots, 20$ and $k = 1, \ldots, 10$. The average entropy grows with n for the critical ensembles $\mathbf{B}(n, k = 2, p = 1/2)$ and approaches a finite value in the chaotic regime independent of n.

$$\lim_{n\to\infty}\langle h\rangle[\mathbf{B}(n, s = n/2)] = -\int_0^1 g(w)w\log(w)dw. \quad (4)$$

Derrida derived the following analytical expression for the distribution of g [16]: $g(w) = \frac{1}{2}w^{-1}(1-w)^{-1/2}$. Substituting this expression into Eq. (4) yields $\lim_{n\to\infty}\langle h\rangle \times [\mathbf{B}(n, s = n/2)] = 2(1 - \ln 2) = 0.613\,71$, which is the average entropy of the random map in the limit of large n. It is remarkable that ensembles of small network sizes already approach this value soon after entering the chaotic regime (see the horizontal black line in Fig. 1).

Before discussing the scaling behavior in critical ensembles, let us briefly recapitulate the concept of irrelevant and relevant nodes in a network [2,6] and their meaning in the context of entropy. A node that is updated by a constant Boolean function is a frozen node. All nodes that eventually take the same constant value on every attractor are called clamped nodes [12]. These nodes build the frozen core of the network. A relevant node is one that eventually influences its own state. The number and the length of the attractors in a network can be determined from the set of relevant nodes.

Nonfrozen nodes that do not influence their own state are irrelevant for the attractor dynamics and are, therefore, called nonfrozen irrelevant nodes. As soon as the network reaches an attractor, the future behavior of such a nonfrozen irrelevant node can be determined as a function of only the relevant nodes. When we add a nonfrozen irrelevant node to a network, the number of states flowing into each of the attractors doubles. Thus, the attractor weights do not change by adding or removing nonfrozen irrelevant nodes. Further, the addition of a frozen node does not, on average, change the attractor weight distribution. Thus, the average entropy of an ensemble depends only on the organization of its relevant nodes. A set including all relevant nodes can be obtained by iteratively removing nodes that freeze and become part of the frozen core and nonfrozen nodes that influence only irrelevant nodes. The remaining set is also referred to as the computational core of a network and exists only in the critical and chaotic regime with high probability [17,18]. The organization of the relevant components is crucial for the understanding of the entropy of the entire network, since it can be calculated by adding the entropies of the independent nonconnected relevant components.

For networks in the ordered regime, the number of relevant nodes approaches a finite limit for large system sizes [19], and the proportion of frozen nodes in the network approaches one [2]. Therefore, the mean number of attractors, and thus the average entropy, are bounded for large systems. These results indicate that increasing the system size does not, on average, increase the entropy of networks in the ordered or chaotic regimes, implying that complexity of classification cannot be increased in these regimes.

Critical ensembles.—We first treat the special case of a critical network with connectivity $k = 1$, as in this case the scaling behavior can be discussed analytically. In these networks, all nodes are updated by either the Boolean "copy" or the "invert" function. The topology of such networks consists of loops and trees rooted in loops. Only nodes on loops are relevant. Frozen nodes do not exist in these networks. Consequently, attractors of the same length have the same weight. The entropy of a critical $k = 1$ network is the sum of the entropies of its loops. In the limit of large n, the number of relevant nodes scales as $n_r \sim (\frac{\pi n}{2})^{1/2}$ (a critical $k = 1$ network is a random mapping digraph, and the number of relevant nodes is equivalent to the number of vertices on cycles; see, e.g., [20]). The probabilities of having n_l loops of size l are independent and Poisson distributed with mean $\lambda = l^{-1}$ [21]. If we approximate the entropy of a loop of size l by $h_l \approx l\ln 2$ and take the sum over the expected number of such loops for $l \le n_r$, we get the following scaling behavior for the entropy in critical $k = 1$ networks [22]:

$$\langle h\rangle \sim \sum_{l=1}^{n_r}\frac{h_l}{l} = \ln 2\left(\frac{\pi n}{2}\right)^{1/2}. \quad (5)$$

In networks with a connectivity higher than one, relevant nodes may depend on more than just one relevant node. In addition to simple loops, such networks may contain complex relevant components with cross-links and attractors of the same length do not necessarily have the same weight any more. The scaling behavior of relevant nodes in a general class of critical random Boolean networks and structural properties of the complex relevant components were recently discussed in an extraordinary series of publications by Drossel, Kaufman, and Mihaljev [19,23,24]. The key result of their work is that the number of relevant nodes n_r scales as $n_r \sim n^{1/3}$ with the system size in all critical ensembles with $k > 1$. Furthermore, the proportion of relevant nodes that depend on more than one relevant input approaches zero with growing n. From Theorem 1.3 in Ref. [25], we find that the expected number of nodes on simple loops increases if the probability of relevant nodes depending on one relevant input increases. As we already discussed the scaling behavior of the average entropy of simple loops, it follows that the average entropy in all critical ensembles increases with system size. Figure 2 shows that the average entropy grows with the expected number of relevant nodes in the critical ensembles $k = 1, 2, 3$ for $n = 20, \ldots, 500$ [26]. The lower average entropy in the $k = 3$ ensemble compared to the $k = 2$ ensemble might be explained by a higher proportion of frozen nodes.

Remarks.—Our findings are particularly relevant to the study of biological networks, as the information processing in living systems is based on massively parallel dynamics in complex molecular networks that underlie ontogeny, immune responses, and cognitive behavior. Kauffman hypothesized that attractors of genetic regulatory networks can be viewed as cell types [7]. If so, then cell types, like classes in parallel processing networks, drain basins of attraction. Recent experimental evidence supports this hypothesis [27].

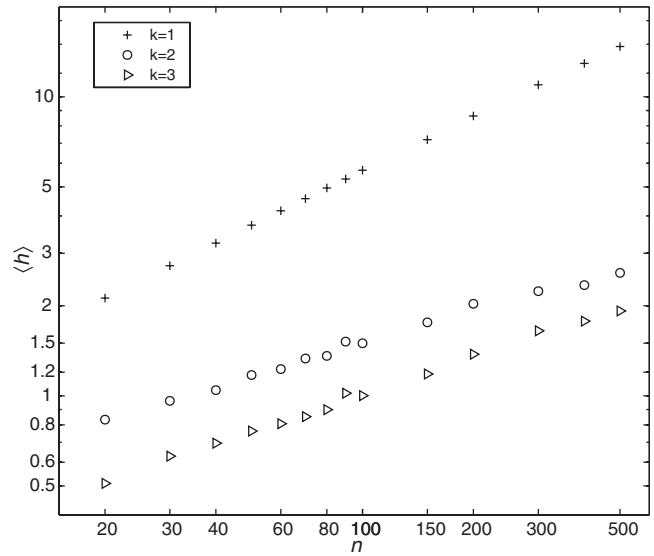

FIG. 2. The average entropy $\langle h \rangle$ of critical network ensembles scales with the number of relevant nodes as $\sim n^{1/2}$ for $k = 1$ and as $\sim n^{1/3}$ for $k = 2, 3$.

The cell is also able to respond to changes in its environment by changing its own behavior. Examples of such decisions include initiation of cell division (cell cycle), execution of specific functions (differentiation), programmed cell death (apoptosis), and other cellular functional states. Such decision-making motivates the view that a living system "classifies" its environment according to its steady states. When viewed in terms of information processing, a cell "reads" the information in its environment, propagates that information through its intracellular networks of interacting biomolecules, and responds by exhibiting a corresponding steady state. The information in the steady state reflects the information in the environment. If maximal classification complexity confers a selective advantage, our results lend support to the longstanding hypothesis that living systems are critical.

The classical Kauffman network ensembles studied herein have characteristics that do not reflect certain aspects of biological networks. First, the synchronous updating scheme does not reflect the fact that elements in biomolecular networks exhibit different time scales. Second, the organization of biological networks is highly modular, containing many loosely coupled modules of similar sizes. In contrast, most relevant nodes in random Boolean networks are, with high probability, organized into one giant component. Since the total entropy is the sum of the entropies of the components, a growing number of such regulatory modules will certainly increase entropy, independent of the updating scheme. Thus, understanding the entropy of complex relevant components under different updating rules is an important topic for future studies.

This work was supported by NIH/NIGMS No. GM070600, No. GM072855, and No. P50-GM076547 and by the Max Weber-Programm Bayern.

[1] B. Derrida and Y. Pomeau, Europhys. Lett. **1**, 45 (1986).
[2] H. Flyvbjerg, J. Phys. A **21**, L955 (1988).
[3] B. Luque and R. V. Solé, Physica (Amsterdam) **284A**, 33 (2000).
[4] M. Aldana, S. Coppersmith, and L. P. Kadanoff, in *Perspectives and Problems in Nonlinear Science*, edited by E. Kaplan, J. E. Marsden, and K. R. Sreenivasan (Springer, New York, 2002), pp. 23–89.
[5] S. A. Kauffman, J. Theor. Biol. **22**, 437 (1969).
[6] U. Bastolla and G. Parisi, Physica (Amsterdam) **115D**, 203 (1998).
[7] S. A. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution* (Oxford University Press, New York, 1993).
[8] P. Rämö, J. Kesseli, and O. Yli-Harja, J. Theor. Biol. **242**, 164 (2006).
[9] I. Shmulevich, S. Kauffman, and M. Aldana, Proc. Natl. Acad. Sci. U.S.A. **102**, 13 439 (2005).
[10] R. Serra, M. Villani, and A. Semeria, J. Theor. Biol. **227**, 149 (2004).
[11] M. Kruskal, Am. Math. Mon. **61**, 392 (1954).
[12] J. Socolar and S. Kauffman, Phys. Rev. Lett. **90**, 068702 (2003).
[13] B. Samuelsson and C. Troein, Phys. Rev. Lett. **90**, 098701 (2003).
[14] M. Gell-Mann and S. Lloyd, Complexity **2**, 44 (1996).
[15] I. Shmulevich and S. Kauffman, Phys. Rev. Lett. **93**, 048701 (2004).
[16] B. Derrida, J. Phys. (Paris) **48**, 971 (1988).
[17] L. Correale, M. Leone, A. Pagnani, M. Weigt, and R. Zecchina, Phys. Rev. Lett. **96**, 018101 (2006).
[18] L. Correale, M. Leone, A. Pagnani, M. Weigt, and R. Zecchina, J. Stat. Mech. (2006) P03002.
[19] T. Mihaljev and B. Drossel, Phys. Rev. E **74**, 046101 (2006).
[20] B. Harris, Ann. Math. Stat. **31**, 1045 (1960).
[21] B. Drossel, T. Mihaljev, and F. Greil, Phys. Rev. Lett. **94**, 088701 (2005).
[22] In a loop, the length l refers to the number of relevant nodes. A loop with an even (odd) number of "invert" functions is called an even (odd) loop. In an even loop of prime length l, there are $(2^l - 2)/l$ attractors of length l and $(2^l - 1)/2l$ attractors of length $2l$ in the odd case. If l is not prime, then there are additional attractors of shorter length [see N. de Bruijn, Verh. K. Ned. Akad. Wet. **49**, 578 (1946)]. We approximated the entropy of a loop by considering only the "longest" attractors and considering even and odd loops as equally likely: $h(l) \approx \frac{1}{2}\ln[(2^l/l) - 2] + \frac{1}{2}\ln[(2^l/2l) - 1] \approx l\ln 2 - \ln\sqrt{2}l$.
[23] V. Kaufman, T. Mihaljev, and B. Drossel, Phys. Rev. E **72**, 046124 (2005).
[24] V. Kaufman and B. Drossel, New J. Phys. **8**, 228 (2006).
[25] C. Cooper and A. Frieze, Comb. Probab. Comput. **13**, 319 (2004).
[26] For $k = 1$, the weight distribution was calculated exactly, whereas for $k = 2, 3$, it was estimated by linking over 2000 random states to their attractors, and the entropy was then averaged over more than 2000 random networks from each ensemble.
[27] S. Huang, G. Eichler, Y. Bar-Yam, and D. E. Ingber, Phys. Rev. Lett. **94**, 128701 (2005).