Portland State University

PDXScholar

Electrical and Computer Engineering Faculty Publications and Presentations

Electrical and Computer Engineering

3-1-2012

Emergent Criticality through Adaptive Information Processing in Boolean Networks

Alireza Goudarzi
Portland State University

Christof Teuscher
Portland State University

Natali Gulbahce

Thimo Rohlf

Follow this and additional works at: https://pdxscholar.library.pdx.edu/ece_fac

 Part of the Electrical and Computer Engineering Commons

Let us know how access to this document benefits you.

Citation Details

Goudarzi, A., Teuscher, C., Gulbahce, N., & Rohlf, T. (2012). Emergent Criticality through Adaptive Information Processing in Boolean Networks. Physical Review Letters, 108(12), 128702.

This Article is brought to you for free and open access. It has been accepted for inclusion in Electrical and Computer Engineering Faculty Publications and Presentations by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Emergent Criticality through Adaptive Information Processing in Boolean Networks

Alireza Goudarzi,[1] Christof Teuscher,[1] Natali Gulbahce,[2] and Thimo Rohlf[3,4]
[1]Portland State University, 1900 4th SW Avenue, Portland, Oregon 97206, USA
[2]University of California, San Francisco, 1700 4th, San Francisco, California 94158, USA
[3]Interdisciplinary Center for Bioinformatics, University Leipzig, Haertelstrasse 16-18, D-04107 Leipzig, Germany
[4]Max-Planck-Institute for Mathematics in the Sciences, Inselstrasse 22, D-04103 Leipzig, Germany
(Received 20 April 2011; published 23 March 2012)

We study information processing in populations of Boolean networks with evolving connectivity and systematically explore the interplay between the learning capability, robustness, the network topology, and the task complexity. We solve a long-standing open question and find computationally that, for large system sizes N, adaptive information processing drives the networks to a critical connectivity $K_c = 2$. For finite size networks, the connectivity approaches the critical value with a power law of the system size N. We show that network learning and generalization are optimized near criticality, given that the task complexity and the amount of information provided surpass threshold values. Both random and evolved networks exhibit maximal topological diversity near K_c. We hypothesize that this diversity supports efficient exploration and robustness of solutions. Also reflected in our observation is that the variance of the fitness values is maximal in critical network populations. Finally, we discuss implications of our results for determining the optimal topology of adaptive dynamical networks that solve computational tasks.

DOI: 10.1103/PhysRevLett.108.128702 PACS numbers: 89.75.Hc, 05.45.−a, 05.65.+b

In 1948, Turing proposed several unorganized machines made up from randomly interconnected two-input NAND logic gates [1] as a biologically plausible model for computing. He also proposed to train such networks by means of a ''genetical or evolutionary search.'' Much later, random Boolean networks (RBNs) were introduced as simplified models of gene regulation [2,3], focusing on a system-wide perspective rather than on the often unknown details of regulatory interactions [4]. In the thermodynamic limit, these disordered dynamical systems exhibit a dynamical order-disorder transition at a sparse critical connectivity K_c [5]. For a finite system size N, the dynamics of RBNs converge to periodic attractors after a finite number of updates. At K_c, the phase space structure in terms of attractor periods [6], the number of different attractors [7], and the distribution of basins of attraction [8] is complex, showing many properties reminiscent of biological networks [3]. In cellular automata, complex computation has been hypothesized to occur where the rules show complex dynamics at ''the edge of chaos'' [9,10]. This claim was refuted in Ref. [11]. However, the argument in Ref. [11] rests on symmetric spaces in the cellular automata lattice and rule space. These results therefore do not apply to RBNs. Phase transition in information dynamics was studied in Ref. [12]. State-topology coevolution in RBNs was studied in Refs. [13–15], and it was shown that networks evolved toward a critical connectivity $K_c = 2$. This Letter presents the first study to link complex dynamics, topology, and task solving in an open RBN.

In Refs. [16–19], simulated annealing and genetic algorithms (GAs) were used to train feedforward RBNs and to study the thermodynamics of learning. For a given task with predefined input-output mappings, only a fraction of the input space is required to train networks that generalize perfectly on all input patterns. This fraction depends on the network size and the task complexity. Moreover, the more inputs a task has, the smaller the training set needs to be to obtain full generalization. In this context, *learning* refers to correctly solving the task for the training samples, while *generalization* refers to correctly solving the task for novel inputs. We use *adaptation* to refer to the phase where networks have to adapt to ongoing mutations (i.e., noise and fluctuations) but have already learned the input-output mapping. In this Letter, we study adaptive information processing in populations of Boolean networks with an evolving topology. Rewiring of connections and mutations of the functions occur at random, without bias toward particular topologies (e.g., feedforward). We systematically explore the interplay between the learning capability, the network topology, the system size N, the training sample T, and the complexity of the computational task.

First, let us define the dynamics of RBNs. A RBN is a discrete dynamical system composed of N automata. Each automaton is a Boolean variable with two possible states: $\{0, 1\}$, and the dynamics is such that $\mathbf{F}$:$\{0, 1\}^N \mapsto \{0, 1\}^N$, where $\mathbf{F} = (f_1, \ldots, f_i, \ldots, f_N)$, and each f_i is represented by a look-up table of K_i inputs randomly chosen from the set of N automata. Initially, K_i neighbors and a look-up table are assigned to each automaton at random. For practical reasons, we restrict the maximum K_i to 8. An automaton state $\sigma_i^t \in \{0, 1\}$ is updated by using its corresponding Boolean function $\sigma_i^{t+1} = f_i(\sigma_{i_1}^t, \sigma_{i_2}^t, \ldots, \sigma_{i_{K_i}}^t)$.

The automata are updated synchronously by using their corresponding Boolean functions. For the purpose of

 © 2012 American Physical Society

solving computational tasks, we define I inputs and O outputs. The inputs of the computational task are randomly connected to an arbitrary number of automata. The connections from the inputs to the automata are subject to rewiring and are counted to determine the average network connectivity $\langle K \rangle$. The outputs are read from a randomly chosen but fixed set of O automata. All automata states are initialized to "0" for each input pattern before simulating the network.

Methodology.—We evolve the networks by means of a traditional GA to solve three computational tasks of varying difficulty, each of which defined on a 3-bit input: full-adder (FA), even-odd (EO), and the cellular automata rule 85 (R85) [20]. The FA task receives two binary inputs A and B and an input carry bit C_{in} and outputs the binary sum of the three inputs $S = A + B + C_{\text{in}}$ on the 2-bit output and the carry bit C_{out}. The EO task outputs a 1 if there is an odd number of 1's in the input (independent of the order) and a 0 otherwise. R85 is defined for three binary inputs A, B, and C and outputs the negation of C. The output for R85 task therefore depends only on one input bit. The EO task represents the most difficult task, followed by the FA and R85 tasks. Task difficulty is the complexity of information integration needed in the input to determine the output. This can be measured through information-theoretical decomposability of a task. We can represent the task itself as the contingency table of its inputs and outputs. Different decomposition models of the task are the different ways that we can calculate the marginal probabilities from the original contingency table [21]. We calculate a weighted sum of the vector of the information content of all possible decomposition models of a task F. This can be summarized in $\text{decomposition}_F = \sum_{m \in \text{Models}_F} w_m \ln f_m$, where Models_F is the set of all decomposition models of F and the weight w_m of a model is proportional to its degrees of freedom. The information content of a model is calculated by using $\ln f_m = 1 - \frac{H_m - H_F}{H_{\text{ind}} - H_F}$. Here, H_m is the entropy of the model, H_F is the entropy of F, and H_{ind} is the entropy of the independence model (all input and output variables are assumed independent). Higher values for decomposition_F mean that the task is more decomposable and therefore less difficult.

The genetic algorithm we use is mutation-based only; i.e., no crossover operation is applied. For all experiments we ran a population of 30 networks with initial connectivity $\langle K_{\text{in}} \rangle = 1$ and a mutation rate of 0.8. Each mutation is decomposed into $1 + \alpha$ steps repeated with probability $p(\alpha) = 0.5^{\alpha+1}$, where $\alpha \geq 0$. Each step involves flipping a random location of the look-up table of a random automaton combined with adding or deleting one link. Each population is run for 30 000 generations. We repeat each evolutionary run 30 times and average the results. In each generation and for each tested input configuration, the RBN is run for a convergence time $t \propto N$ updates. Afterward, we run the network for an additional $t \propto N$ updates to record the activity of the output nodes. If the activity of an output node is "1" for at least half of the t time steps, we interpret the output as a 1 and as 0 otherwise. For an evolutionary run of training size T, the training sample set M is randomly chosen at each generation without replacement from the 2^3 possible input patterns. During each generation, the fitness of each individual is determined by $f = 1 - E_M$, where E_M is the normalized average error over the T random training samples: $E_M = \frac{1}{T} \sum_{i \in M} \sum_{j \in O} (a_{ij} - o_{ij})^2$. a_{ij} is the value of the output automata j for the input pattern i, and o_{ij} is the correct value of the same bit for the corresponding task. The generalization score is calculated by using the same equation with M including all 2^3 inputs rather than a random sample. Finally, selection is applied to the population as a deterministic tournament. Two individuals are picked randomly from the old population, and their fitness values are compared. The better individual is mutated and inserted into the new population; the worse individual is discarded. We repeat the process until we have 30 new individuals in the new population.

Results.—We observe a convergence of $\langle K \rangle$ close to the critical value $K_c = 2$ for large system sizes N and training sample sizes larger than or equal to $T = 4$. For $T = 8$, populations always evolve close to criticality for moderate N. For smaller T, the average over all evolutionary runs is found at slightly higher values of $\langle K \rangle$ (Fig. 1). If the average is taken only over the *best individuals*, however, $\langle K \rangle$ values close to K_c are recovered. This observation can be explained from the fact that for $T < 8$, due to the limited information provided for learning, some populations cannot escape local optima and, hence, do not reach maximum fitness. Suboptimal network populations show a large scatter in $\langle K \rangle$ values in the evolutionary steady state, while those with high fitness scores cluster around $K_c = 2$ (Fig. 1, inset). For the simple R85 task, we do not observe

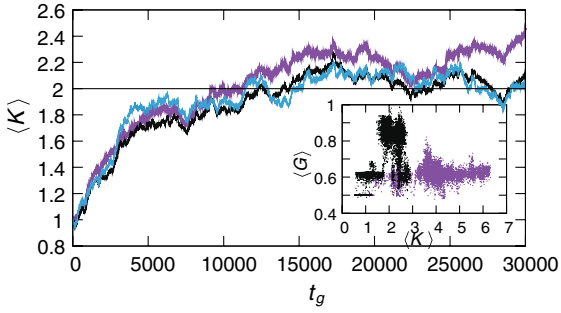

FIG. 1 (color online). Convergence of the average network connectivity as a function of the GA generations t_g. The FA task with $T = 4$ and $N = 100$. The curves are averaged over 30 evolutionary runs (red), only the 22 best (green), and the 15 best (light blue) populations, respectively. Inset: Scatter plot correlating average $\langle K \rangle(t_g)$ and average generalization $\langle G \rangle(t_g)$ of a successful population (black) and a suboptimal population (purple).

any convergence to $K_c = 2$, independent of the training samples. For the other tasks, the finite size scaling of $\langle K \rangle$ (Fig. 2) exhibits convergence towards K_c with a power law as a function of the system size N. For $T = 8$, the exponent b of the power law for the three tasks EO, FA, and R85 is -1.63, -1.11, and -0.30, respectively (Fig. 2). Altogether, these results suggest that the amount of information provided by the input training sample helps to drive the network to a critical connectivity.

Interestingly, the population dynamics in our model follow Fisher's fundamental theorem of natural selection, which attributes the rate of increase in the mean fitness to the increased fitness variance in the population [22]. It has been shown in GAs that the diversity maximization [23] makes more configurations of the search space accessible to the genetic search to find optimal solutions [16–18].

Indeed, we find that the standard deviation of the fitness values in the populations has a local maximum near K_c (Fig. 3, inset), with a sharp decay toward larger $\langle K \rangle$, indicative of maximum diversity near criticality. Evidently, this diversity helps to maintain a high fitness population in the face of continuous mutations with a fairly high rate (0.8 in our study). While the average fitness can be lower (and often is), compared to less diverse populations, the probability to find and maintain high fitness solutions is strongly increased. Indeed, we find that populations where the best mutant has maximum fitness ($f = 1$) sharply peak near K_c (Fig. 3), as well as populations where the best mutant reaches perfect generalization. To find a possible source of fitness diversity, we determined several topological measures of the networks [24]: the eccentricity (maximum shortest path between a vertex v and any other vertex in a graph), the betweenness centrality (the average fraction of shortest paths between all vertices in a graph that passes through a vertex v), the participation, and the characteristic path length. These measures were calculated for Erdös-Rényi, exponential random graphs, as well as for the evolved networks (Fig. 4). In fact, we find that the graph-theoretical measures have maximal variance near $K_c = 2$. Similarly, other authors have shown that dynamical diversity is maximized near K_c, too [25]. Our results suggest that evolving RBNs can indeed exploit this diversity to optimize learning.

In addition, we find that, during the learning process of the networks, the in-degree distribution changes from a Poissonian to an exponential distribution. In particular, we observe that the topological properties of the networks reach a compromise between Erdös-Rényi graphs and the exponential random graphs. The same observation was made in input- and outputless RBNs that were driven to criticality by using a local rewiring rule [14]. This significant topology change is related to diversity (entropy) maximization during the learning phase [26]. However, this is beyond the scope of this Letter and will be discussed in a separate publication.

Finally, we measured the damage spreading in the evolved RBNs [5] to determine their dynamical regime. The damage spreading d_{t+1} is measured by changing the state of a randomly selected node in two identical networks. The two networks are simulated for a single time step, and the damage spreading $\bar{d}$ is then calculated by averaging the ratio $\frac{d_{t+1}}{d_t}$ over many trails with random initial network configurations. One observes that for critical networks $\bar{d} = 1$, for supercritical networks $\bar{d} > 1$, and for subcritical networks $\bar{d} < 1$. We see that, for networks with a high fitness, $\bar{d}$ peaks around 1 for all N.

Discussion.—We investigated the learning and generalization capabilities in RBNs and showed that they evolve toward a critical connectivity of $K_c \approx 2$ for large networks and large input sample sizes. For finite size networks, the connectivity approaches the critical value with a power law

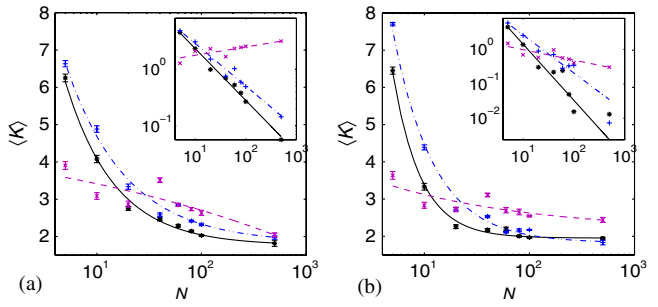

FIG. 2 (color online). Finite size scaling of $\langle K \rangle$ as a function of N for the three tasks, EO (black), FA (blue), R85 (magenta), and the training sample size $T = 4$ (a) and $T = 8$ (b). Points represent the data of the evolved networks; lines represent the fits. The finite size scaling for $\langle K \rangle$ shows that it scales with a power law as a function of the system size N. The dashed lines represent the power-law fit of the form $a * x^b + c$. We favor the data for larger N by weighting the data according to $N/N_{\max}$, where $N_{\max} = 500$. The insets show $K_c - c$ as a function of N on a log-log scale.

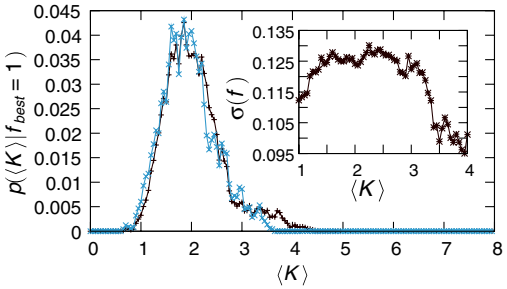

FIG. 3 (color online). The conditional probability that evolving populations, where the best mutant reaches maximum fitness (i.e., $f_{\text{best}} = f_{\max} = 1$), have average connectivity $\langle K \rangle$ shows a sharp peak near K_c (black curve); the same is found for maximum generalization (light blue). Inset: The diversity of evolving populations, quantified in terms of the standard deviation $\sigma(f)$ of fitness distributions, has a maximum near $K_c = 2$. All data were sampled over the best 22 out of 30 populations for the full-adder task with $T = 4$ and $N = 100$.

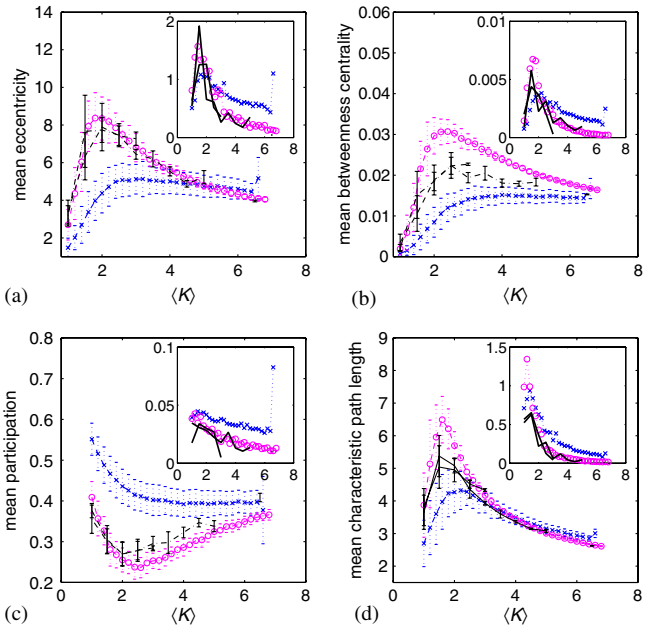

FIG. 4 (color online). Near K_c, the topology of the network shows maximal variance. The insets show the standard deviation of the topological measures for initial Erdös-Rényi networks (magenta), the evolved networks (black), and the exponential random graphs (blue). The solid lines represent the topological measures in random networks. The dotted line represents the same measure in RBNs.

of the system size N. We showed that network learning and generalization are optimized near criticality, given that the task complexity and the amount of information provided surpass threshold values. Furthermore, critical RBN populations exhibit the largest diversity (variance) in fitness values, which supports learning and robustness of solutions under continuous mutations. By considering graph-theoretical measures, we determined that K_c corresponds to a region in network ensemble space where the topological diversity is maximized, which may explain the observed diversity in critical populations.

Interestingly, we observe that RBN populations that are optimal with respect to learning and generalization tend to show average connectivity values slightly below $\langle K \rangle = 2$. This may be related to previous results indicating that $K_c < 2$ in finite size RBNs [27].

Examination of the attractors of the final population confirms that the computation happens as partitioning of the state space into disjoint attractors [28]. During the evolution, the attractor landscape changes so that there are enough attractors to properly process the inputs. The entire task is encoded as a hypercycle (i.e., a set of mutually reachable attractors) in the network dynamics. The input combinations play the role of a multivalued switch that pushes the dynamics out of one attractor into the next along the hypercycle. The emergence of the large attractor basins makes the computation highly robust to perturbations in the node state while maintaining sensitivity to input signals. All networks in our final population converge to fixed-point or cyclic attractors.

In summary, we solved a long-standing question and showed that learning of classification tasks and adaptation can drive RBNs to the edge of chaos [3], where high-diversity populations are maintained and ongoing adaptation and robustness are optimized. Our study may have important implications for determining the optimal topology of a much larger class of complex dynamical networks where adaptive information processing needs to be achieved efficiently, robustly, and with limited connectivity (i.e., resources). This has applications, e.g., in the area of neural networks and complex networks and, more specifically, in the area of emerging molecular and nanoscale networks and computing devices, which are expected to be built in a bottom-up way from vast numbers of simple, densely arranged components that exhibit high failure rates, are relatively slow, and are connected in an unstructured way.

This work was partly funded by NSF Grant No. 1028120. The first author is grateful for the fruitful discussions with Guy Feldman and Lukas Svec.

[1] A. M. Turing, in *Machine Intelligence*, edited by B. Meltzer and D. Michie (Edinburgh University, Edinburgh, United Kingdom, 1969), p. 3–23.
[2] S. A. Kauffman, J. Theor. Biol. **22**, 437 (1969).
[3] S. A. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution* (Oxford University, New York, 1993).
[4] S. Bornholdt, Biol. Chem. **382**, 1289 (2001).
[5] B. Derrida and Y. Pomeau, Europhys. Lett. **1**, 45 (1986).
[6] R. Albert and A. L. Barabási, Phys. Rev. Lett. **84**, 5660 (2000).
[7] B. Samuelsson and C. Troein, Phys. Rev. Lett. **90**, 098701 (2003).
[8] U. Bastola and G. Parisi, Physica (Amsterdam) **115D**, 203 (1998).
[9] C. G. Langton, Physica (Amsterdam) **42D**, 12 (1990).
[10] N. H. Packard, in *Dynamic Patterns in Complex Systems*, edited by J. A. S. Kelso, A. J. Mandell, and M. F. Shlesinger (World Scientific, Singapore, 1988), pp. 293–301.
[11] M. Mitchell, P. T. Hraber, and J. P. Crutchfield, Complex Syst. **7**, 89 (1993).
[12] J. Lizier, M. Prokopenko, and A. Zomaya, in *Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems (Alife XI)* (MIT, Cambridge, MA, 2008), pp. 374–381.
[13] S. A. Kauffman, J. Theor. Biol. **149**, 467 (1991).
[14] S. Bornholdt and T. Rohlf, Phys. Rev. Lett. **84**, 6114 (2000).
[15] M. Liu and K. E. Bassler, Phys. Rev. E **74**, 041910 (2006).
[16] A. Patarnello and P. Carnevali, Europhys. Lett. **4**, 503 (1987).
[17] P. Carnevali and S. Patarnello, Europhys. Lett. **4**, 1199 (1987).

[18] S. Patarnello and P. Carnevali, in *Neural Computing Architectures: The Design of Brain-Like Machines*, edited by I. Aleksander (North Oxford Academic, London, 1989), p. 117.
[19] C. Van den Broeck and R. Kawai, Phys. Rev. A **42**, 6210 (1990).
[20] S. Wolfram, Rev. Mod. Phys. **55**, 601 (1983).
[21] M. Zwick, Kybernetes **33**, 877 (2004).
[22] A. W. F. Edwards, Biol. Rev. Camb. Philos. Soc. **69**, 443 (1994).
[23] M. A. Bedau and A. Bahm, in *Artificial Life IV: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*, edited by R. A. Brooks and P. Maes (MIT, Bradford, MA, 1994), pp. 258–268.
[24] M. Rubinov and O. Sporns, NeuroImage **52**, 1059 (2010).
[25] M. Nykter, N. D. Price, A. Larjo, T. Aho, S. A. Kauffman, O. Yli-Harja, and I. Shmulevich, Phys. Rev. Lett. **100**, 058702 (2008).
[26] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Rev. Mod. Phys. **80**, 1275 (2008).
[27] T. Rohlf, N. Gulbahce, and C. Teuscher, Phys. Rev. Lett. **99**, 248701 (2007).
[28] P. Krawitz and I. Shmulevich, Phys. Rev. Lett. **98**, 158701 (2007).