

# Commentary on “Emergence in Biomolecular Networks?”

James F. Lynch  
Department of Mathematics  
Box 5815  
Clarkson University  
Potsdam, NY 13699-5815  
Telephone: 315-268-2374  
Fax: 315-268-2371  
email: jlynch@clarkson.edu

## Abstract

We examine several versions of the notion of “emergence” in the context of random Boolean networks. Certain properties of the behavior of these networks have often been claimed to be emergent. We discuss the strengths of these claims under different interpretations of emergence and different standards of mathematical proof.

## 1 Introduction

“Emergence” is a term from dynamical systems theory that, like the related term “complexity,” seems to have a variety of meanings. We will discuss some of these meanings, first in the general context of dynamical systems theory, and then for a specific class of dynamical systems known as agent-based models, where the term may have some utility.

A dynamical system is described by a pair  $\langle \mathcal{S}, R \rangle$ , where  $\mathcal{S}$  is its set of states, and  $R$  is the rule defining the state transitions. The rule can be deterministic, nondeterministic, or probabilistic. For example, the state of a Boolean network is specified by the values of its nodes, the functions assigned to the nodes, and the connections between the nodes. In Kauffman’s random Boolean networks [10], all of these features are chosen randomly at the start, but once chosen, the functions and connections stay fixed. Therefore only the values of the nodes change during a transition, and since the next value of each node is uniquely specified by the present values of its input nodes, Kauffman’s Boolean networks are deterministic. On the other hand, the Boolean networks of Derrida and Pomeau [4] are probabilistic because the functions and connections are randomly chosen at each step.

We propose the following as a strict definition of emergence.

**Definition 1.** *Let  $\mathcal{P}$  be a proof system. An emergent property of a dynamical system (with respect to  $\mathcal{P}$ ) is a property that is not a logical consequence of the definition of the system.*

Proofs of emergence thus belong to metamathematics, since they are proofs of unprovability, and as Professor Malaterre points out, such proofs are very rare.

To give an example of an emergent property that has a perhaps superficial relation to questions about stability and order, take any class of dynamical systems that has the capability of universal computation, such as Turing machines or cellular automata and any consistent proof system in which the transition rules can be defined, and consider the property that the system halts on every starting state. Halting may be regarded as a very strong form of stability, where every component in the system eventually stops changing. It follows from standard arguments about Turing-recognizability that there are systems for which this property is emergent. Thus we have an emergent property involving stability. This argument could not work for the Boolean networks of Kauffman and Derrida and Pomeau because nodes are never created once a network begins operation, and therefore it can only reach a finite number of states from any given starting state. But if the transition rules included the possibility of adding nodes, then Boolean networks with the capability for simulating universal computation could be defined, and certain forms of ordered behavior would be emergent. Of course, these examples of emergence would be very artificial, and it seems unlikely that any property that is provably emergent in the strong sense would have biological relevance.

Even though examples of strict emergence are very uncommon and the known examples are not interesting from a scientific viewpoint, biologists still use the term frequently, albeit casually, and some even claim that it is a useful notion. We will describe one such version of emergence that is often applied to agent-based models. An agent-based model is a type of dynamical system where states consist of populations of agents or individuals, with structural relations imposed on the agents. The structural relations can pertain to the individuals themselves or to pairs of individuals, triples, and so on. Boolean networks are agent-based models, where the nodes are the agents, and the structural relations are the Boolean functions assigned to the nodes and the connections between the nodes. Grimm and Railsback [8] propose the following criteria for emergent properties of agent-based models:

- Emergent properties are not just the sum of the properties of the individuals.
- Emergent properties are of a different type than the properties of the individuals.
- Emergent properties cannot be easily predicted from the properties of the individuals.

Admittedly, these criteria are not precise, but Grimm and Railsback claim that emergence is a meaningful concept that can lead to better understanding of behavior.

When scientists design a model to explain some form of behavior, they can either impose the behavior explicitly on the individuals, or they can try to model the true actions of the individuals and hope that the desired behavior emerges. In general, the latter approach results in more complex models, and poses the problem of explaining the emergent behavior.

Grimm and Railsback give numerous examples of both approaches in ecology, but the same principles apply to any domain where agent-based modeling is used.

According to the criteria above, the existence of a critical connectivity  $K_c$  for random Boolean networks is emergent. It is certainly difficult to derive from the properties of the nodes. Kauffman hypothesized  $K_c = 2$  and made several other conjectures about ordered and chaotic behavior of random Boolean networks in the late 1960's [9], but it was not until 1991 when Luczak and Cohen [11] published the first mathematically rigorous proofs of any of Kauffman's claims, and strictly speaking, the question of existence of a critical connectivity is still not completely resolved.

In addition to being unexpected, the existence of a critical connectivity illustrates significant differences in outlook between two scientific communities. Derrida and Pomeau [4] proved  $K_c = 2$  for the annealed version of Boolean networks, and then extended it by analogy to Kauffman's quenched model. The analogy is that in the thermodynamic limit the two models show similar behavior. More specifically, for any time  $t$ , as the number of nodes  $N \rightarrow \infty$ , the behavior of a given node at time  $t$  is asymptotically the same in both models. Although this is true for fixed  $t$ , what really needs to be shown is that, as  $N \rightarrow \infty$ , for all  $t$ , the behavior of a given node at time  $t$  is asymptotically the same in both models.

There is now a substantial body of analytic work, mostly by physicists such as Derrida and Pomeau, on random Boolean networks. There is a much smaller body of work by mathematicians. Luczak and Cohen (who cited Derrida and Pomeau) proved that Kauffman's  $K = 2$  networks exhibited certain forms of stable behavior. They left open some questions about other forms of ordered behavior, in particular the conjectures that the average limit cycle size was on the order of  $\sqrt{N}$  when  $K = 2$ , and that  $K > 2$  was the onset of chaotic behavior. The first conjecture was later shown to be false: the average limit cycle size when  $K = 2$  grows faster than any polynomial in  $N$  [12]. There are some results indicating that networks with  $K > 2$  are not as stable as networks with  $K \leq 2$  [13], but they fall short of proving long-term chaotic behavior. Thus the picture is more complicated than originally supposed, and it is not yet complete.

A fundamental question is whether the results about Boolean networks extrapolate to real genomic networks. Here the picture is more uncertain. The review by de Jong [3] and the book by Bower and Bolouri [2] give an idea of the variety of more complex models of genomic and metabolic networks. It is not at

all clear when and if more complex models are better than Boolean networks for capturing essential features of biomolecular systems. There are some arguments that, at least in some cases, the Boolean formalism is adequate [1, 7], provided nodes with different numbers of connections are allowed.

Analysis of more complex models is more difficult, of course. Again, the physicists have led the way. Professor Malaterre mentioned the work by Solé, Luque, and Kauffman [16] on networks constructed from multivalued logic nodes. Other authors have investigated how different probability distributions on the Boolean functions assigned to nodes and the connectivity between nodes affect the critical connectivity [6, 14, 15]. Most of their arguments rely on mathematically unwarranted assumptions like the thermodynamic limit analogy and the mean-field assumption, which essentially says that a population of diverse individuals behaves like a homogeneous population of average individuals. Other than the citations already listed and several related papers, there seem to be very few mathematically rigorous papers analyzing extensions of the basic random Boolean network model.

Finally there are the meta-theoretic questions of whether the informal notion of emergence as used by biologists and physicists can be defined more precisely, and whether that would be useful. Perhaps the best view of emergence is that it is simply a way of identifying challenges to scientists, and progress occurs when a property formerly regarded as emergent is no longer regarded as such. As stated by di Paolo, Noble, and Bullock [5], “For a simulation model to be of any use, both obvious and nonobvious patterns must be explained and not brushed under the carpet of emergence as this amounts to an admission of failure.”

## References

- [1] R. Albert and H. G. Othmer. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*. *J. Theoret. Biol.*, 223:1–18, 2003.
- [2] J. M. Bower and H. Bolouri. *Computational Modeling of Genetic and Biochemical Networks*. The MIT Press, Cambridge, MA, 2001.
- [3] H. de Jong. Modeling and simulation of genetic regulatory systems: A literature review. *J. Computational Biol.*, 9:67–103, 2002.
- [4] B. Derrida and Y. Pomeau. Random networks of automata: a simple annealed approximation. *Europhys. Lett.*, 1:45–49, 1986.
- [5] E. A. di Paolo, J. Noble, and S. Bullock. Simulation models as opaque thought experiments. In M. A. Bedau, J. S. McCaskill, N. H. Packard, and S. Rassmussen, editors, *Artificial Life VII: The Seventh International Conference on the Simulation and Synthesis of Living Systems*, pages 497–506, Reed College, Portland, OR, 2000. MIT Press/Bradford Books, Cambridge, MA.
- [6] J. J. Fox and C. C. Hill. From topology to dynamics in biochemical networks. *Chaos*, 11:809–815, 2001.
- [7] L. Glass and S. A. Kauffman. The logical analysis of continuous, non-linear biochemical control networks. *J. Theoret. Biol.*, 39:103–129, 1973.
- [8] V. Grimm and S. F. Railsback. *Individual-Based Modeling Ecology*. Princeton University Press, 2005.

- [9] S. A. Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. *J. Theoret. Biol.*, 22:437–467, 1969.
- [10] S. A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, New York, 1993.
- [11] T. Luczak and J. E. Cohen. Stability of vertices in random Boolean cellular automata. *Random Structures and Algorithms*, 2:327–334, 1991.
- [12] J. F. Lynch. On the threshold of chaos in random Boolean cellular automata. *Random Structures and Algorithms*, 6:239–260, 1995.
- [13] J. F. Lynch. Dynamics of random Boolean networks. In K. Mahdavi, R. Culshaw, and J. Boucher, editors, *Current Developments in Mathematical Biology, Proceedings of the Conference on Mathematical Biology and Dynamical Systems*, pages 15–38. World Scientific Publishing Co., 2007.
- [14] C. Oosawa and M. A. Savageau. Effects of alternative connectivity on behavior of randomly constructed Boolean networks. *Physica D*, 170:143–161, 2002.
- [15] M. Skarja, B. Remic, and I. Jerman. Boolean networks with variable number of inputs ( $k$ ). *Chaos*, 14:205–216, 2004.
- [16] R. V. Solé, B. Luque, and S. A. Kauffman. Phase transition in random networks with multiple states. *Phys. Rev. A*, submitted.