

Hip Bone is connected to... II

Leo P. Kadanoff

e-mail: LeoP@UChicago.edu

In a previous column (Physics Today, September 2005, page 10), I described how computer models might be used to mimic and describe a few of the small networks that control and drive biological systems. I focused upon the understanding that could be obtained from moderately accurate descriptions of relatively simple biological systems. Here, I wish to look at conceptualizations of much larger networks.

Forty years ago when Stuart Kauffman took on the immensely challenging task of understanding something about the interlinked chemical activity within a living cell. He despaired of the task of doing this problem in anything like its full detail but instead decided to describe this problem by a vastly simplified and generalized model. He started with N variables or “nodes”, each representing a chemical species that might be present in a cell¹. The possible states of the cell was specified by snapshots which described each species as “present” or “absent” at a given time. The cell dynamics was a vastly simplified, stepwise process. Each compound’s presence at time step $t+1$ depends on the presence or absence of a few other compounds at time t . Thus for example compound A would be present at time $t+1$, if B and C were present at time t , but D was absent. The entire biological cell was described by listing rules like this one for all the N compounds in the cell. Two parameters would describe the average properties of the biological cell, the number of compounds, N , and the average number of precursors, K , whose presence or absence would determine the formation of a given compound.

Of course, one could make up a huge number of networks which would fit the description just given. Different networks would be specified by giving different connections and different rules for the formation of each particular chemical compound. The job of finding the rules for a given biological system could be expected to be immense, providing a job more than one generation of biochemists and biologists.

Kauffman was unwilling to wait. Instead he built upon work of Paul Erdos and

Alfred Renyi,² in which they studied ensembles constructed from all possible networks of a given type. Kauffman¹ picked at random from all possible networks with given N and K , asking what could be learned by saying that biological systems are constructed from a “typical” set of rules. This approach is somewhat similar to the Boltzmann-Gibbs strategy for statistical mechanics, which says that a configuration for a set of gas molecules might be described as picked at random from all possible configurations of N molecules with a given energy. A biological cell is represented by a network with fixed connections and fixed rules. Each chemical compound is a node in the network, and is linked by connecting lines to those compounds which directly affected its presence.

In recent years people have in fact learned the detailed structure of such networks for a few biological systems by painstaking experimental study of all their reactions and interactions. Lots can be learned from examining such specific examples. (Figure 1 shows a network of this kind.) But it is also interesting to look back and see what has been learned from the general features of Kauffman’s analysis. Because each compound may be either present or absent, the number of configurations is 2^N , so that eventually the cell must return to a configuration it has visited earlier. Because the time-development rules are fixed, thereafter the cell can only retrace its earlier steps. Thus, the system will eventually settle down into a repetitive, cyclical behavior. A given system can support several different repetitive cycles so that, depending upon initial data, the cell will fall into one of several cycles. Kauffman took the different cycles to each represents a different kind of cell. Our bodies contain many different cell types.... skin cell, brain cell, a muscle cell... each with an identical genetic makeup. Just as different cycles of dynamical systems might be generated by giving the systems different initial conditions, so different cell types emerge from identical dynamics but different starting conditions.

Now one might begin to ask about the generic properties of such networks. For an overview of the work done with models of this kind, see the review paper of Max Aldana, Susan Coppersmith and myself³. At least three separate kinds of behavior can be discerned in a specified network. If the connections among nodes are sparse, different parts of the system each have their own dynamics independent of the other parts. For such systems, which arise

when the average K is less than two, we get a kind of frozen behavior in which there is too little linkage among the different nodes for the system to exhibit the kinds of complexity characteristic of actual biological situations. Conversely, for $K > 2$, the network motion is totally chaotic in that a change in the rule at almost any node can change the subsequent behavior in a finite fraction of all the nodes of the system. This kind of structure is much too noise-sensitive to represent the noise-tolerant response characteristic of real biological systems. For K very close to 2, the system displays a kind of critical behavior intermediate between frozen and chaotic states. Here, if you change an initial value of a node or change the rule for updating a certain node a small subset of the nodes will change their behavior, but most of the nodes will continue to follow the same pattern as before. This kind of partially flexible, but mostly **unchanging**, behavior is characteristic of most biological systems. Biologists describe such behavior by using the word “robust”.

Kauffman thus argues that biological systems might well show a kind of critical dynamics, akin to the dynamics seen near the critical point of a phase transition. This argument is rather widely accepted as giving a very rough but reasonable result.

This early work was necessarily related to the gross structure of biological networks. In the meantime, however, biologists have been able to see the actual dynamics of a few much-studied networks. Each of these networks had previously been analyzed piece by piece in experimental work by biochemists and biologists. The nodes each contain the concentration of the different chemical compounds, each linked to the compounds which determine their production rate. These studies do not support the initial presupposition that biological networks look like they have been picked from an randomly constructed ensemble. Indeed, many of the networks have been understood as being structured from small pieces which provide rather elementary functions, like the AND or OR gates found on a computer chip. For example, Rene Thomas⁴ has done an extensive analysis of the component pieces of biological networks.

Furthermore, recent work has shown that these networks most often are constructed from a few preferred small structures. Uri Alon⁵ and coworkers has analyzed small pieces-- containing only a few nodes and their

interconnections--- carved from large biological networks and counted the frequency of occurrence of the different possible structures. Alon then compared these frequencies with ones drawn from randomly connected networks, a la Kauffman. The biological networks showed a detailed structure very different from the purely random systems. In particular, in each network a very few of these structures, called motifs, appeared far more frequently than one would expect from randomness.

For example, according to Alon, in contrast to a random network there are many more negative feedback loops with the structure of figure 2a than one might expect at random for a network characterized by an average connectivity, K . These negative feedback loops act as “thermostats” which can control in a precise and reliable manner the concentration of important chemical species. In contrast, the number of positive feedback loops is far smaller because biological systems have little use for runaway quantities.

In different biological networks have different motifs because different kinds of small structures might serve useful purposes in each context. Evolution will then especially select those structures that are robust, i.e. maintain their functionality even when changed slightly. Alon further argues that, like a computer programmer, evolution tends to duplicate and reuse structures already proven to be useful. That is why he expect that each kind of network will show its own list of oft-repeated motifs.

Thus recent studies move far from Kauffman’s random networks. However, the randomness is not the important take-home message from the earlier work. Rather the enduring point is the three states of the system: frozen, critical, and chaotic and the contrast between the way information flows in each of these states. This message is far broader than the model which originally supported it and is in fact used in many different areas of physics, mathematics, and computer science. The change between the ordered and chaotic modes is called the percolation transition and is a ubiquitous descriptor of the slow transfer of information.

1. S. Kauffman, "Metabolic stability and epigenesis in randomly constructed genetic nets," in: *Journal of Theoretical Biology*, 22:437-467, 1969.

2. Paul Erdos and Alfred Renyi, On random graphs, *Publ. Math. (Debrecen)* 6:290 (1959)

3. M. Aldana, S. Coppersmith and L.P. Kadanoff. “Boolean Dynamics with Random

Couplings”, In Perspectives and Problems in Nonlinear Science. A celebratory volume in honor of Lawrence Sirovich. Springer Applied Mathematical Sciences Series. Ehud Kaplan, Jerrold E. Marsden, and Katepalli R. Sreenivasan Eds. May 2003. ISBN: 0-387-00312-6.

4. <http://www.ulb.ac.be/cenoliw3/PERSO-PAGES/rthomas.html>

5. Uri Alon, An Introduction to Systems Biology, Design Principles of Biological Circuits, Chapman and Hall, 2007

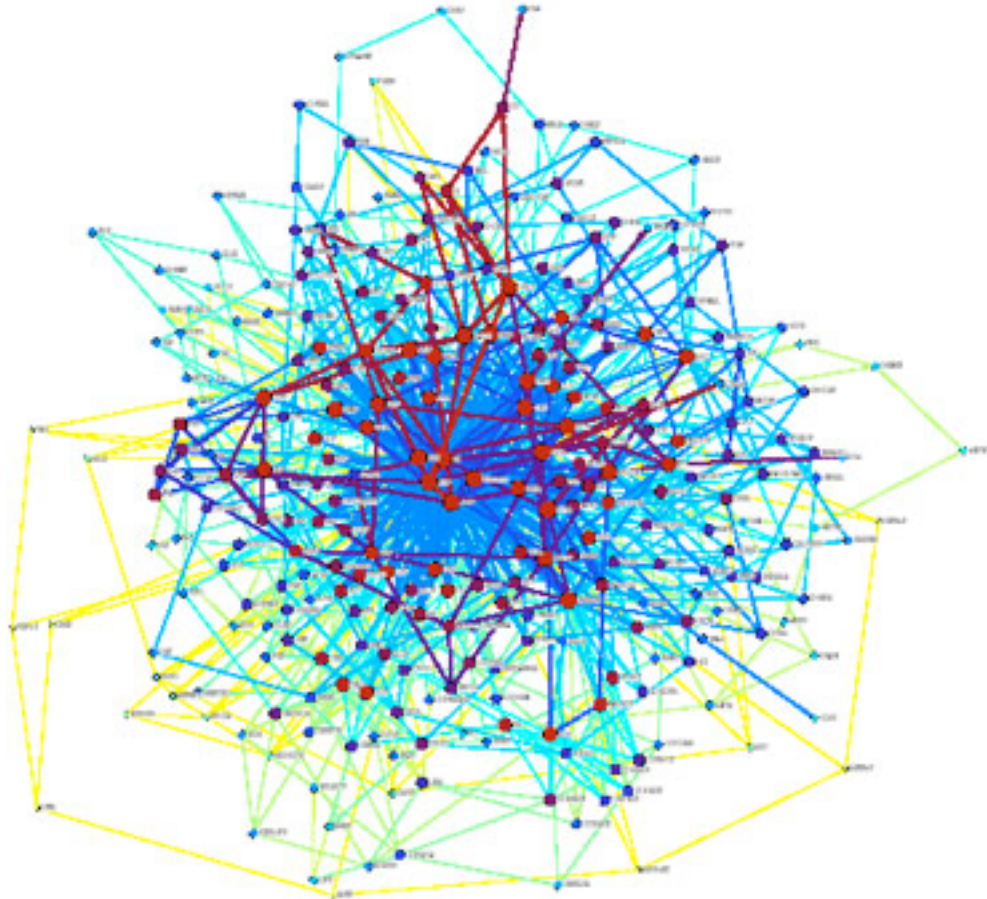


Figure 1. The metabolic network of the bacteria, E. coli. Alon analyzed the structure of all the small pieces in this network. Here the network is visualized using Himmeli software package by Ville Mäkinen.

www.lce.hut.fi/publications/pdf/LCE_Annual_Report_2005.pdf

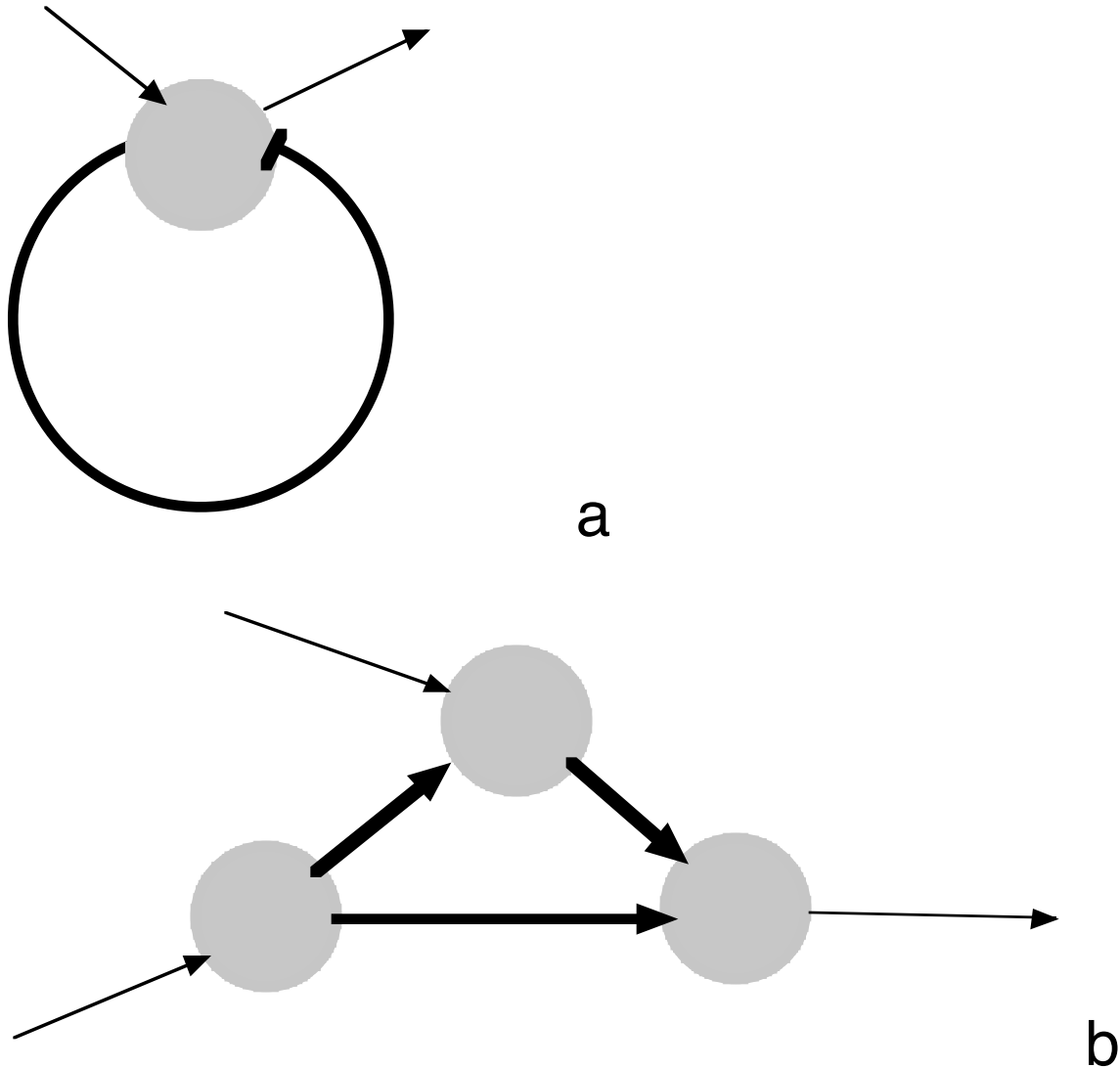


Figure 2. Two motifs from E. Coli gene regularitory network. The gray circles are nodes. The heavy lines represent connections with the motifs, while the light lines connect them to the rest of the network. Part a is a negative feedback circuit, with the bar at the end oif the arc indicating that this particular signal is inhibitory. Part b might represent and “and” gate in which a signal is produced on the right if both inputs are present.

copyright Leo Kadanoff (2008, 2009)