Philosophy of Systems Biology: 5 Questions
Sara Green (ed.)

William Bechtel

Professor
Department of Philosophy and Center for Circadian Biology, University of California, San Diego

Systems Biology: Negotiating between Holism and Reductionism

1. How and why were you initially drawn to systems biology?

The historical conflict between holists (vitalists) and mechanists has long fascinated me. Often holists pointed out real limitations of mechanistic explanations of their day. The problem, however, is that the holists generally did not have a productive research strategy—a way to advance and defend accounts of how systems as wholes functioned. So the history has largely been one in which mechanists developed more sophisticated accounts of mechanisms that addressed some of the shortcomings of previous mechanistic accounts. In an early, but illuminating example, Bernard responded to the vitalist objections posed by Bichat regarding the indeterminacy of physiological responses and the apparent drive of living organisms to resist death by positing that living systems are organized to maintain the constancy of their internal environment. Subsequently, this provided part of the grounding for Cannon's conception of homeostasis. With the articulation of negative feedback processes as tools for control, mechanisms could be envisaged that would exhibit phenomena that Bichat thought were beyond the capacity of mechanisms.

The growing interest in systems by the Cyberneticists and the proponents of General Systems Theory in the mid 20th Century began to provide conceptual tools that could address whole systems, but these remained, for the most part, small steps that often lost contact with actual biology. The emergence of systems biology in the 21st Century reflects the further development of research tools that can give holists a research agenda that can truly complement the mechanist's highly successful enterprise of identifying component parts and characterizing their operations. To my mind, there are two distinct types of research methodologies that systems biology has brought forward. The first is the ability to collect and analyze very large corpora of data so that one can gain information about, for example, the pattern of expression of large numbers of genes or activities of large numbers of molecules in cells. The second, and the one that has attracted my interest, is the development of mathematical tools that enable researchers to represent the organization and behavior of systems of large numbers of components that interact non-linearly and are organized non-sequentially. These include the tools of graph theory, computational modeling, and dynamical systems theory.

In many ways, the rhetoric associated with systems biology has exceeded the results that have been secured to date. This has aroused skepticism among critics. In many respects, this early history of systems biology resembles that of artificial intelligence. With new tools (production systems and list programming) proponents of early AI often presented themselves as on the road to accomplishments that proved ultimately to be far more difficult than initially envisaged. But while AI systems that match human cognitive performance overall are still far in the future, AI has made very substantial progress, and in fact many of its tools are being put to use in systems biology. I expect much the same path for systems biology. It is not offering instant solutions. But what is exciting is that it has provided new research tools whose potential will require time to realize.

2. How do you view the relation between philosophy and systems biology, and (how) can these fields inform each other?

As a philosopher of biology, I am most interested in how systems biology can inform philosophy of science by, for example, revealing scientific practices that have not been adequately characterized in philosophical accounts. I do hold out the hope that philosophy of science may make contributions to actual science, but given the sordid track record of applying philosophical analyses within science, I think philosophers should proceed with great humility. One problem that afflicts any cross disciplinary endeavor is that people in one discipline adopt what is put forward in another discipline as decisive when in fact it is one of many views and often represents a stage in the development of that discipline. The remedy is not to rely on textbooks or even review articles, but to engage with the original literature and even the researchers themselves. This may be easier for philosophers than for scientists since we do not have laboratories to run, and investigating the scientific results and processes is the object of our research. But we can also do more to make our work relevant to scientists by presenting our philosophical analyses in the context relevant to the scientists' research.

I turn, though, to my particular interest as a naturalistically oriented philosopher of science (i.e., a philosopher who takes as his mission to characterize science as it is actually done) in systems biology. Much of my work has focused on explanation and what systems biology provides are examples of different strategies for explaining phenomena than those I have examined previously. This may lead to a replay of what happened when philosophers seriously engaged with cell and molecular biology, which resulted in rejecting the traditional empiricist framework in which the explanatory strategy was to discover laws from which descriptions of phenomena could be derived. Biologists seldom appeal to laws except for those they borrow from physics and chemistry, but instead appeal to mechanisms. Beginning in the late 20th century several philosophers of biology, myself included, began to try to understand what biologists took mechanisms to be, how they represented them and reasoned about them, and especially the strategies they invoked in discovering them. Systems biologists often differentiate their endeavors from those of mechanistic inquiry and even if ultimately their results can be reconciled with mechanistic accounts, what this makes clear is that systems biologists are approaching biological phenomena with a different theoretical perspective and set of research tools. This provides philosophers a rich resource both to examine the challenges systems biologists face in deploying these tools and the types of explanatory accounts they develop using them.

3. What do you consider the most neglected topics and/or contributions in late 20th Century (philosophy of) biology?

The most neglected topic, in large part because of the epistemic challenges in understanding it, is the role of organization of systems in giving rise to their dynamic behavior. In the last years of the 20th Century philosophers of biology began to catch up with domains of biology that had been pursuing mechanistic explanations by decomposing mechanisms into their component parts and operations. In fields such as cell and molecular biology, researchers in the 20th Century developed powerful techniques for identifying component structures of biological systems and determining the operations they perform. Although all accounts of mechanistic explanation included reference to the organization of parts and operations as crucial to the ability of mechanisms to generate the phenomenon investigators are trying to explain, there was little discussion of modes of organization and the consequences different modes of organization might have for the dynamic functioning of the mechanism. The epistemic challenge is to identify patterns of organization and determine their effects.

Largely as a consequence of the fact that humans develop new thoughts sequentially, when we approach the interactions of multiple components we think of them carrying out their operations sequentially. When biologists represent a biological mechanism, they often rely on box and arrow diagrams. Inferring the behavior of the mechanism involves following a path through the diagram and mentally simulating the effects of each operation. This strategy works reasonably well when the organization is sequential and the interactions are linear. It can even be extended to slightly less sequential cases such as those in which an operation feeds back on an operation envisioned as occurring earlier. Bernard, Cannon, and the cyberneticists were able to recognize how negative feedback could maintain a system in a stable configuration. By the early 20th Century engineers recognized that negative feedback could also give rise to oscillations, but it is far more challenging to determine through mental rehearsal whether a negative feedback system will generate sustained oscillations, as this depends on the non-

linearities in the interactions. Doing so requires mathematical modeling, which depends on not just an appropriate mathematical description of the system but appropriate tools for performing the computations required. Inspired by the feedback mechanism in the operon described by Jacob and Monod, Goodwin developed a mathematical representation, but using an analog computer, he seriously underestimated the value needed for a critical parameter to sustain oscillations. When circadian researchers identified a negative feedback loop as a candidate mechanism for circadian rhythms, Goldbeter revived Goodwin's model, but incorporated delays to enable sustained oscillations with a more realistic parameter value than Goodwin's model required.

Negative and even positive feedback were invoked in the analysis of biological systems in the later years of the 20th Century, but discovering other modes of organization lagged behind. It appeared that there might not be any general principles that could characterize complex biological mechanisms—each might have to be analyzed using its own computational model. If so, the understanding of organization would at best be derivative of mechanistic research directed at identifying parts and operations—after these were identified, mathematical models could be used to determine if the mechanism could generate the phenomenon in question. One could not hope to anticipate the organization, however, from the type of behavior exhibited by the phenomena.

4. What have been the most significant advances in systems biology?

To my mind, the most significant advance brought by systems biology is new tools for analyzing patterns of organization of biological systems that enable biologists to begin to reverse engineer biological systems. These tools involve applying and developing resources from graph theory, computational modeling, and dynamical systems theory to understand biological systems. I focus first on graph theory. Graph theorists in the mid-20th Century analyzed the most mathematically tractable graphs—random networks and regular lattices. Both of these had important applications to biology in accounting for synchronization of components or repeating sequences of behavior. But at the end of the century Watts and Strogatz focused attention on an intermediate mode of organization in which nodes that are highly connected to their neighbors also have a few connections to more distant nodes. They identified these as small-world networks and, besides noting their widespread occurrence in naturally occurring systems, pointed to their power in processing information. Soon after Barabási and his colleagues recognized that many real-world networks violate another simplifying assumption made in 20th century graph theory—that the number of edges originating in nodes is distributed normally. Instead, the distribution often approximates a power law, with highly connected nodes constituting hubs linking other nodes into modules or providing connections between modules.

Although recognizing that many biological networks, including gene regulatory networks and protein interaction networks, exhibit small-world organization with approximately power-law distribution of connections has proven useful in accounting for properties such as the robustness of biological systems, such analyses of global network structure are still in early stages of development. My hope is that as research progresses, sub-categories within the small-world region might be identified and their dynamical properties analyzed. Just such progress has occurred with respect to local configurations within network structures. By identifying frequently occurring patterns of connections between small numbers of units, commonly referred to as *motifs*, and then developing models to determine the types of behaviors to which such sub-graphs would give rise, Alon, his colleagues, and other researchers have provided potent resources for understanding how systems will behave. An instructive example, analyzed by Tyson and Novak, involves two units each feeding back negatively on the other. Without outside inputs, such a motif will settle into a state in which whatever unit had the highest activation becomes more active. With inputs to one or the other unit, the motif can switch between these two states, and with appropriate parameter values, it realizes a bistable switch in which a much greater increase or decrease in input is required to cause the switch to reverse than was required to drive it into a state to begin with. Given its functionality, it is not surprising that Tyson and Novak found several instances of bistable switches in the eukaryotic cell cycle at points regulating advances from one stage to another at which it is important for the system not to revert to an earlier stage.

Motif analysis has most frequently been applied at the level of local circuits, but an alternative and potentially extremely valuable application is its use to understand the underlying principles that explain behavior in a complex mechanism. I will exhibit this strategy in Ueda's research on the circadian clock mechanism in animals. Following the discovery of the first clock gene, *per*, in *Drosophila*, and the determination that concentrations of both *per* mRNA and the protein Per oscillate in cells, with the protein lagging several hours behind the mRNA, Hardin et al. proposed a negative feedback mechanism. In their proposal, when Per is in high concentration, it inhibits the transcription of its own gene, thereby reducing transcription until it degrades; only once it is degraded are transcription and translation able to resume. After the initial transcription-translation feedback model was developed, circadian researchers discovered a host of additional genes and proteins involved and identified multiple feedback loops through which they affect each other. Figure 1 provides a representative view of the mechanism. Starting with Goldbeter's 1995 model, the mechanism has been the focus of numerous computational models, many advanced with the goal of trying to identify what are the critical components that enable the mechanism to generate circadian oscillations. Ueda pursued a different approach, which involved finding a way to abstract from the particular genes and proteins to develop a framework that revealed what he proposes is the core organizational principle.

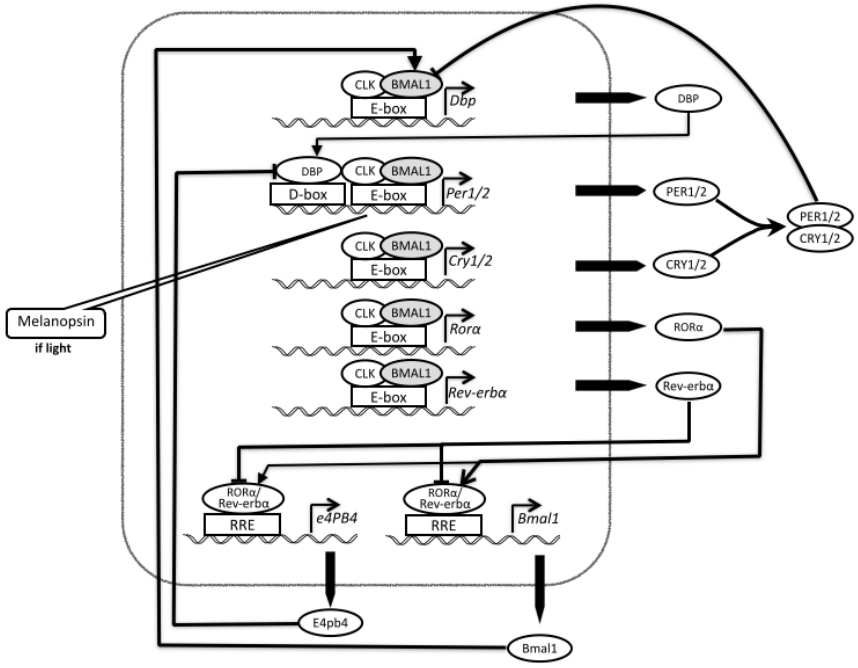

Figure 1. Diagrammatic representation of the mammalian circadian clock mechanism showing feedback loops through which proteins expressed from clock genes feed back to bind or affecting the binding at promoters on these clock genes.

Ueda focused on the fact that expression of clock genes is regulated by transcription factors that bind to sequences of DNA in the promoter region known as E-boxes, D-boxes, and Ror-elements (RREs). These are commonly portrayed in representations of the circadian clock, such as Figure 1, but what distinguishes Ueda's approach is that he makes them central. Since one or more are found not only in the promoter regions of all genes viewed as part of the clock mechanism but also many genes that are regulated by the clock, Ueda calls them clock-controlled elements (CCEs). By inserting destabilized luciferase genes into the region regulated by the promoters in a cell culture system and recording the timing of maximum bioluminescence, Ueda was able to determine the time when each CCE was most active. Although the precise time varies by tissue, in the suprachiasmatic nucleus, thought to be the locus of the central clock, E-boxes are most active in the morning, D-boxes about five hours later (evening) and RREs about eight hours later (night). As a result of this investigation, Ueda has introduced an alternative representation of the circadian

mechanism (Figure 2a) in which the three CCEs are central and the genes/proteins that are activators or inhibitors of the CCEs are shown as green or purple ovals with arrows or flat-ended lines connecting them to CCEs to which they bind. Having downplayed the genes and proteins to simply being the vehicles by which the promoter boxes regulate each other, he collapsed multiple linkages that achieve the same effect into single arrows, arriving at Figure 2b. He then recognized that this network is composed of the two motifs shown in Figure 2c (a repressilator and a negative feedback loop), both of which are known to be capable of generating sustained interactions.

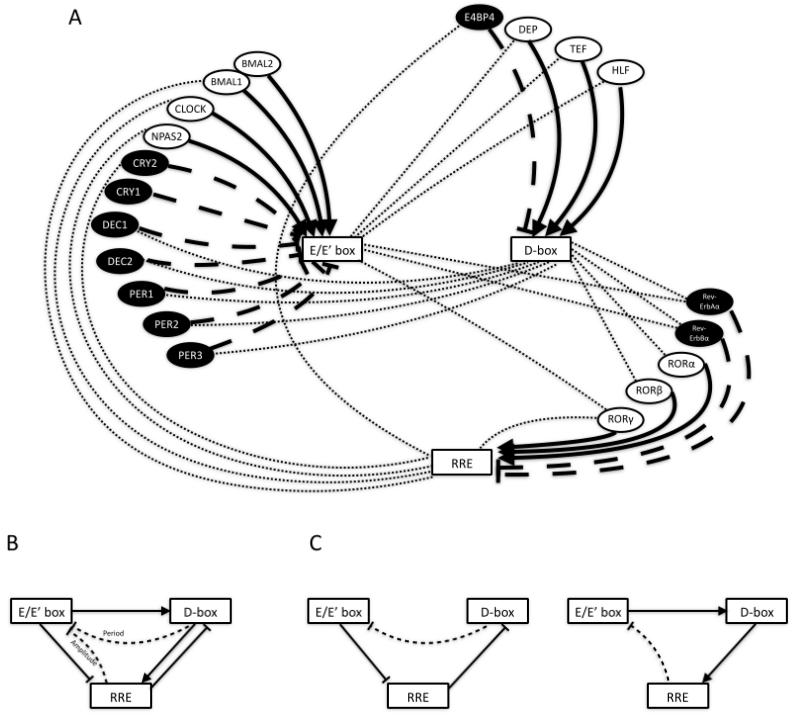

Figure 2A. The format Ueda developed for representing the mammalian circadian clock mechanism that makes the promoter boxes central. Genes and the proteins into which they are transcribed are not differentiated. Light dotted lines indicated genes/proteins regulated by a given box, dark arrows indicate that the protein activates the promoter and dark dashed end-edged lines indicate that the protein inhibits the promoter. Adapted from Ukai-Tadenuma, M., Kasukawa, T., & Ueda, H. R. (2008). Proof-by-synthesis of the transcriptional logic of mammalian circadian clocks. *Nat Cell Biol, 10*, 1154-1163. **2B-2C.** Ueda's strategy for reducing the multiple connections between promoter boxes to single arrows. C decomposes B into two motifs, a repressilator and a negative feedback loop. B and C are adapted from Ukai-Tadenuma, M., Yamada, R. G., Xu, H., Ripperger, J. A., Liu, A. C., & Ueda, H. R. (2011). Delay in feedback repression by cryptochrome 1 is required for circadian clock function. *Cell, 144*, 268-281.

The reason that graph structures such as small-worlds and motifs are of interest in systems biology is that they can provide a foundation for understanding the dynamical behavior of biological systems. To understand the behavior of motifs under different conditions (represented in different parameter values), Alon and Tyson relied on computational modeling. But if researchers want not just to know the results of a computational model, but also to understand how it generates the behavior, they need a means to represent the operation of the model. Dynamical systems theory has provided a number of graphical tools for both mathematically characterizing and graphically representing how a complex system changes its state over time. Graphically, one can represent the state of a system at any time as a point in state space defined by the different variables used to describe the system. Change over time will then involve a trajectory through state space. As a result of the organization of the system, the set of trajectories in state space will be limited, and one can characterize the geometry of the space. Points or regions of the space in which

trajectories terminate constitute attractors (these regions might be points (representing a system that settles), continuous cycles (systems that sustain oscillations), or regions in which the system never visits the same exact location twice (systems that are in a chaotic regime). By using one dimension to represent the probability of the system ending up in the location defined by the other dimensions, one can present a landscape in which valleys represent attractors, mountain tops repellers (unstable configurations that can evolve in multiple directs), and ridges separatrices differentiating valleys. The challenge with landscape diagrams is that we can only actually construct such diagrams using two dimensions to represent variables and one for the probability of the defined state. This works well for motifs in which there are only two states whose values are represented as variables. Although we cannot visualize higher-dimensional spaces, the same procedures for constructing a landscape can be applied, allowing researchers to characterize attractors (which may themselves be complex multidimensional structures), repellers, etc., and invoke these to describe the evolution in time of a complex network.

5. What do you consider the most important problems in (philosophy of) systems biology and what are the prospects for progress in this respect?

New tools such as graph theory, mathematical modeling, and dynamical systems theory provide holists with a research program that they previously lacked. Beyond merely criticizing extant mechanistic accounts for failing to take into account the context of the system in which mechanisms operate, they can represent and analyze the behavior of complex systems. The new challenge, though, is to integrate systems-level theorizing with accounts of the underlying mechanism. The need for this stems from the fact that for systems theorists to go beyond describing possible systems by demonstrating that the system structure they identify is actually realized in the biological system in question. The ability of a dynamical model to characterize nuances of the phenomena that are detected empirically provides some evidence that the proffered dynamical account actually characterizes the system. But by far the strongest evidence that a graph represents the actual mechanism is that the nodes can be related to identified parts and edges to operations detected in the mechanism. Likewise, the strongest evidence that a given dynamical account characterizes an actual mechanism is that the variables correspond to properties of parts of the system that can be measured and the equations describe operations that can be empirically investigated.

Relating dynamical and mechanistic accounts is a huge challenge. The tools and reasoning of mechanistic researchers and dynamical systems theorists are very different. In their attempts to decompose mechanisms into parts and operations, mechanistic biologists set up experimental contexts in which many variables are controlled, creating what Cartwright refers to as a nomological machine. If these are well designed, researchers can characterize the structure of the parts and measure precisely the operations in which they engage. But these are often very different from those found naturally in which the components are enmeshed in large-scale systems in which a variety of other components can affect the behavior of any given component. Systems theorists interested in the complex dynamics, on the other hand, turn to mathematical models which may include variables and relations that cannot be experimentally verified but which are needed to generate the desired phenomena. Especially troubling for many mechanists is that dynamical models not only abstract from the detail of actual mechanisms but often idealize by introducing components not believed to occur in the actual system. As Nersessian has observed, when mechanistic investigators and dynamical modelers try to engage each other, they often talk past each other. The modelers ask for data that cannot be procured with current research tools and experimentalists expect answers that modelers cannot provide.

I finish with what I take to be one promising strategy for bridging the localist approach of mechanists and the global approach of holists that invokes graph theory representations of systems and the analysis of motifs. In pursuing their work on motifs, both Alon and Tyson have applied the analysis of the motifs to well-studied biological mechanisms in which the motifs arise. Here it is possible to relate the values of variables in the mathematical models to those of specific molecules and the relations between variables to reactions that can be demonstrated in experimental conditions. The challenge, however, is that the components characterized in the motif do not occur in isolation, but are embedded in networks. Activity elsewhere in the network can alter the states of components represented as

variables in the models or even the parameters used in characterizing relations between variables. This might seem to simply undermine the approach of analyzing motifs, but the hope is that one can add in additional components in a step-by-step manner, preserving key ideas from the motif analysis while incorporating additional features of the systemic context. One reason to hope that such an approach might be successful, at least in many domains, is the recognition that many biological systems have been shown to exhibit small-world organization with the number of edges per node distributed according to a power law. This tends to result in modules with highly interconnected components but with some connections to components of other modules. When motifs occur within modules, then they may show their typical effects being only moderately affected by activity elsewhere in the system.

In discussions of systems biology there is often reference to top-down versus bottom-up causation. In previous work I have argued for making sense of the phenomena for which these terms are used while restricting causation to intra-level contexts (accommodating the downward and upward relations in terms of the constitution relation between parts and wholes). But graph theory provides a potentially more informative way to visualize what have been thought of as intra- and inter-level causal processes while jettisoning the not well-articulated sense of levels. What is needed is a means of detecting within a network cases when a set of nodes constitutes a module that exhibits endogenous dynamical activity (e.g., sustained endogenous oscillation) while still open to influences from elsewhere. Such a module is what might have been characterized as a higher-level entity, but in the graph it is simply an organization of nodes. If one wants, one could draw a circle around the module to indicate that it exhibits an endogenous dynamic in which each component is responding to other nodes in the module and the response to any input from outside will be modulated by the dynamics within the module. Using a circle for the components in the module as well is appropriate since in fact the entities corresponding to most nodes within a network can themselves be decomposed into a set of organized components. Whether one does so will depend on whether the details of the internal processes in the entities corresponding to the nodes are taken to be important for the explanation sought. Although a causal process will typically impinge on some components of a module more than others, which will then affect other components of the module over time, it may suffice for one's explanatory purposes to simply view the process as affecting the module. Within this network perspective, bottom-up causation arises when local causal processes propagate within a module to generate a system whose responses depend on those local processes. These very causal processes, though, are also what mediate top-down effects as they result in the components of the network behaving differently depending on the state of the network as a whole. This account applies iteratively as one moves (a) in to more local parts of the network and identifies modules with their own systemic patterns of organization wherein the behavior of parts is largely determined by the behavior of other parts or (b) out to more global modules that are formed by realizing such organization between local modules that the activity within local modules is modulated by activity in other modules of the larger structure.