EVELYN FOX KELLER*

Organisms, Machines, and Thunderstorms: A History of Self-Organization, Part One

ABSTRACT

Over the last quarter century, the term "self-organization" has acquired a currency that, notwithstanding its long history, has been taken to signal a paradigm shift, and perhaps even a scientific revolution, introducing a new *Weltanschauung* in fields as diverse as mathematics, physics, biology, ecology, cybernetics, economics, sociology, and engineering. But there is a prehistory to this revolution, as to the term itself, with at least two earlier episodes in which the same term was used to signal two other, quite different revolutions. In this paper, I review the pre-history of "self-organization," starting with Immanuel Kant, who first introduced the term, and then turn to the dramatic reframing of the concept by mid-twentieth century engineers. In a subsequent paper, I will review the more recent history of this concept when the term was once again reframed, this time by physicists. My aim will be to situate this latest incarnation of "self-organization" against the backdrop of earlier discussions.

KEY WORDS: Self-regulation, stability, homeostasis, cybernetics, self-organization

INTRODUCTION

This history of self-organization is, in part, a history of the complex and shifting relations between the life sciences, the physical sciences, and engineering in their ongoing struggle for methodological and epistemological supremacy.

*E51-171, Massachusetts Institute of Technology, Cambridge, MA 02139; efkeller@mit.edu.

This essay had its origin in a Rothschild Lecture at Harvard (2005) and grew into the Ryle lecture series at Trent University (Nov 2005). In its present form, I have borrowed liberally both from an earlier work (E. F. Keller, "Marrying the Pre-Modern to the Post-Modern: Computers and Organisms after WWII," in *Growing Explanations: Historical Perspectives on Recent Science*, ed. M. Norton Wise [Durham, NC: Duke University Press, 2004], 181–200) and a more recent overview in E. F. Keller, "The Disappearance of Function from 'Self-Organizing Systems,'" in *Systems Biology: Philosophical Foundations*, ed. Fred C. Boogerd, Frank J. Bruggeman, Jan-Hendrik S. Hofmeyr, and Hans V. Westerhoff (Amsterdam: Elsevier Science, 2007), 303–17.

Historical Studies in the Natural Sciences, Vol. 38, Number 1, pps. 45–75. ISSN 1939-1811, electronic ISSN 1939-182X. DOI: 10.1525/hsns.2008.38.1.45.

It begins in the late eighteenth century, on the eve of the designation of biology as a science of its own. The subject of the new science, the *organism*, is specified by the definition of biology—it is an entity assumed both to be given by nature and, even though visibly temporal, unchanging in its basic character (at least over such a period of time as that with which this history is concerned). By contrast, because it is a product of human industry, the *machine*, the subject of the mechanical arts (or engineering), may reasonably be expected to change over the course of history, and certainly, what qualifies as a machine has changed dramatically over the last two centuries or more. The subject of the physical sciences is of yet a third kind, for although it too is presumed to be given by nature, and hence unchanging, it is designated both by convention and by the state of its science, and for that reason it too has changed. In my title I chose the thunderstorm to signal the subject of the physical sciences, but in fact that choice only makes sense in the context of the preoccupations of today's physical sciences. At the beginning of this history, a better choice might have been the planetary system, or even, as it was for William Paley, the stone on the heath. As to the machine, I do not specify the kind at issue, but in the early days, the clock would have been a good stand-in; today a better choice would be the computer. In other words, not only do the sciences themselves have a history, but so do their subjects. Given this double historicity, the relations between and among the subjects of these various sciences inevitably, and necessarily, also have a history. My aim in this essay is to explore that history by following the particular concept of "self-organization," chosen both because it is a term that was introduced in order to clearly—and, it might be said, newly—demarcate these three subjects, and because of the complex and varied role it has since played in the changing configurations of organisms, machines, and inanimate, and naturally occurring, physical systems over this period.

I begin with Immanuel Kant, for it was he who originally introduced the term, and he did so as a way of characterizing what it was that so conspicuously singled out organisms from other subjects. For the next 150 years, the distinctiveness of organisms from machines held firm. Each might serve as analogy for the other, but only as analogy; organisms cannot be confused with machines. But their distinctiveness from inanimate, physical systems was less stable. With the emergence of thermodynamics in the mid-nineteenth century, and Helmholtz's success in demonstrating the equivalence of animal heat and energy, the relations between animate and inanimate systems became uncertain. By the turn of the century, however, worries began to accumulate, including those about the compatibility of life with the second law of thermodynamics. New confidence in the clarity of the distinction between the two kinds of systems came as well with a renewed focus in the early twentieth century on the active role of organisms in maintaining their properties, even those that seemed to resemble properties of physical systems. And at this time, it once again became scientifically respectable to maintain the singularity of organisms, to defend a fundamental divide between the living and the non-living.

But the most dramatic mutation in this tradition came with the radical, and surprisingly rapid, breakdown of one of the founding divisions, namely the divide between organisms and machines, and this came with the rise of cybernetics in the immediate aftermath of World War II. The basic claim of cybernetics was that the relation between organisms and machines was not merely analogous, but homologous: organisms *were* machines, and at least some machines could be organisms. Accordingly, it ought to be possible to build machines with the same self-organizing capacities as organisms. The architects of cybernetics effectively bypassed the confusions and disagreements about the relation between life and the second law of thermodynamics by turning directly to engineering, and more specifically to the design and construction of machines with just those properties that had previously been thought to make them so distinctive. Ironically, the term employed to refer to the well-regarded, and well-funded, program that emerged to build such machines, and which persisted for the next twenty years, was "self-organization." I end Part One of this essay with the effective collapse of this program, leaving the next chapter of this history, in which still another vision of self-organization resurfaces, for Part Two. In this most recent vision, it is physics that (once again) takes the lead: organisms are assimilated not to machines but to non-linear dynamical physical systems—like thunderstorms, for example. But the story does not end here, for this vision of self-organization too falls short of Kant's high standard. I end the essay with a brief look at some signs of things to come, hints of yet another vision of self-organization in the wings. My intent, however, is to view all of these variants against the backdrop of the history that preceded them.

KANT AND HIS FOLLOWERS

The term *self-organization* first appears in the second part of Immanuel Kant's *Critique of Judgment* (1790), where he took upon himself the exceedingly difficult question of how far science can help us understand those peculiar entities of nature that we call living beings. To address this question, he needed first to ask: What is an organism? What is the special property, or feature, that distinguishes a living system from a collection of inanimate matter? At the time, in the late eighteenth century, even before biology was defined as a separate and distinctive science, at least the form of an answer was already in the air. For what led to the common grouping of plants and animals in the first place—what makes "the two genres of organized beings" (as Buffon referred to them) *organisms*—was a new focus on just that feature, that is, on their conspicuous property not simply of being organized, but of being organized in a particular way. A century earlier, John Locke had already laid great emphasis on the importance of organization in establishing the identity of plants and animals:

> For this organization, being at any one instant in any one collection of matter, is in that particular concrete distinguished from all other, and is that individual life, which existing constantly from that moment both forwards and backwards, in the same continuity of insensibly succeeding parts united to the living body of the plant, it has that identity which makes the same plant, and all the parts of it, parts of the same plant, during all the time that they exist united in that continued organization, which is fit to convey that common life to all the parts so united.[1]

But Locke's concern lay primarily with the identity of the individual plant and animal, whereas the end of the eighteenth century brought a particular concern with what distinguished plants and animals, now grouped together under the category of the living from that altogether distinct category of the non-living. Here the emphasis shifts to the particular ways in which living beings were organized.

What made it possible to distinguish an organism from its Greek root, *organon*, or tool, was that special arrangement and interaction of parts that brings the wellsprings of form and behavior of an organism *inside* itself. A tool, of necessity, requires a tool-user, whereas an organism is a system of organs (or tools) that is self-steering and self-governing: it behaves as if it has a mind of its own. Indeed, Kant gave one of the first modern definitions of an *organism*—not as a definition per se, but rather as a principle or "maxim" which, he wrote, "serves to define what is meant as an organism"—namely, "one in which every part is reciprocally both end and means. In such a product nothing is in vain, without an end, or to be ascribed to a blind mechanism of nature."[2] Organisms, he wrote, are the beings that

1. John Locke, *An Essay Concerning Human Understanding* (http://enlightenment.supersaturated.com/johnlocke/BOOKIIChapterXXVII.html, 1690), bk. 2, ch. 27, §4.

2. Immanuel Kant, *Critique of Judgement,* vol. 39 of *Great Books* (Chicago: Encyclopedia Britannica, 1993), 558, §66.

> first afford objective reality to the conception of an *end* that is an end of *nature* and not a practical end. They supply natural science with the basis for a teleology . . . that would otherwise be absolutely unjustifiable to introduce into that science—seeing that we are quite unable to perceive *a priori* the possibility of such a kind of causality.[3]

Elaborating on this kind of causality, he wrote:

> In such a natural product as this every part is thought as *owing* its presence to the agency of all the remaining parts, and also as existing *for the sake of the others* and of the whole, that is as an instrument, or organ. . . . [T]he part must be an organ *producing* the other parts—each, consequently, reciprocally producing the others. . . . Only under these conditions and upon these terms can such a product be an *organized* and *self-organized being*, and, as such, be called a *physical end*.[4]

Here is where the term "self-organized" first makes its appearance. An organism is not merely self-steering, self-governing, and self-maintaining; it is also self-organizing. More, it is self-generating. As Kant put it, an organism is both "*cause and effect of itself*."[5] The conspicuous organization of living beings and their consequent appearance of design are not due to any external agent, but only to their own nature. Organisms are products of nature. No external force, no divine architect, is responsible for this peculiar entity, only the internal dynamics of the being itself.

To say what an organism *is* would thus require description and delineation of the particular character of the organization that defined its inner purposiveness, that gave it a mind of its own, that enabled it to organize itself. What is an organism? It is a bounded body capable not only of self-regulation, self-steering, but also, and perhaps most important, of self-formation and self-generation; it is both an organized and a self-organizing being. An organism is a body which, by virtue of its peculiar and particular organization, is constituted as a "self"—an entity that, even though not hermetically sealed (or perhaps because it is not hermetically sealed), achieves both autonomy and the capacity for self-generation.

At the close of the eighteenth century and the dawn of the nineteenth, it was evident to Kant, as well as to his contemporaries, that neither blind chance

3. Ibid., 558, §66.
4. Ibid., 557, §65 (italics in original).
5. Ibid., 555, §64 (italics in original).

nor mere mechanism, and certainly no machine that was then available, could suffice to account for this special kind of organization. "For a machine," Kant wrote, "has solely *motive power*, whereas an organism possesses *formative* power."[6] "Strictly speaking," he elaborated, "the organization of nature has nothing analogous to any causality known to us."[7] The science of such an entity would thus have to be a new kind of science, one that Kant assumed to be irreducible to physics and chemistry. As he famously wrote, there "will never be a Newton for a blade of grass." Not only can there be no laws of motion generating the organization of living organisms, but also there could be no such laws accounting for their origin. It is "contrary to reason," he wrote, that "raw material could have originally formed itself according to mechanical laws, that life could have originated from the nature of the lifeless, and that matter could have arranged itself in the form of a self-maintaining purposiveness."[8] Life's distinctiveness from non-life is thus two-fold, effectively corresponding on the one hand to an organism's ontogeny and its phylogeny, and on the other hand to its taxonomic relation to man-made machines and to structures arising spontaneously from dead, inert matter.[9]

The two terms—organism and self-organization—remained tightly linked until well into the twentieth century, with the conjunction frequently invoked to mark the taxonomic divide (the distinctiveness of organisms both from the realm of the inanimate objects of nature [like Paley's rocks], and that of machines [e.g., clocks]). Inanimate matter lacks both the organization and self-organization of organisms, and machines, though organized, and even designed, are organized and designed from without. The organization and design of organisms is, by contrast, internally generated. The burden of the concept of self-organization thus fell on the term *self*, for it was the self as source of the organization that prevents an organism from ever being confused with a machine. We might analogize organisms to machines, or to artifacts, but their self-production prevents us from ever actually confusing the two. It might be useful, or even necessary, for us to regard them *as if* designed, but because they are natural products, the notion of design must remain metaphorical. Indeed, Kant's invocation of final causation seems itself to have been analogical: the organic must be explained "as if" it were constituted as teleological. Teleological judgment is

6. Ibid., 557, §65.

7. Ibid.

8. Ibid., 581, §81. Hannah Ginsborg, "Kant on Understanding Organisms as Natural Purposes," in *Kant and the Sciences*, ed. Eric Watkins (Oxford: Oxford University Press, 2001), 243.

9. Alix A. Cohen, "A Kantian Stance on Teleology in Biology," *South African Journal of Philosophy* 26 (2007): 109–21.

appropriate to the investigation of nature "only with a view to bringing it under principles of observation and research by analogy to the causality that looks to ends, while not pretending to explain it by this means."[10] Kant does not in fact exclude the possibility that such products are intrinsically mechanical, but only the possibility that we, given the limits to our powers of judgment, can ever fully explain them in terms of blind mechanism. Which is not to say that we should not try, but only that we must accept our own cognitive limits. His argument is that it is precisely because of our limitations that we must use the concept of teleology in our investigations of biological systems.

For the next fifty years, Kant's formulation inspired a rich vein of research in German biology that Tim Lenoir has described as "teleo-mechanist." Scientists like Ernst von Baer, Johannes Müller, Justus von Liebig, Rudolph Leukart, and Theodor Ludwig Wilhelm Bischoff all sought formulations of biological phenomena that were as faithful to mechanistic accounts as was possible while still respecting the centrality of purposive organization. As Bischoff put it, however much physical and chemical analyses can help us in understanding the form and composition of an organ, and even the material changes associated with its functioning, they can only go so far. For understanding the generation and maintenance of the organ, other forces, beyond the reach of our analyses, are required.[11]

LIFE AND THE FIRST LAW OF THERMODYNAMICS

The middle of the nineteenth century brought with it a revolution in the physical sciences that, to many, made the teleo-mechanist position come to seem untenable. Especially important were, first, Hermann Helmholtz's tour de force demonstration in 1845 that only material processes are involved in the generation of animal heat, and, second, in 1847, his formulation of the principle of conservation of energy (or the first law of thermodynamics). These provided powerful inspiration for a wholesale attack on the need for any notion of purposive organization in biology. Helmholtz himself saw the issue in these terms, introducing his findings as follows:

> One of the most important questions of physiology, one immediately concerning the nature of the *Lebenskraft*, is whether the life of organic bodies is the effect of a special self-generating, purposive force or whether it is the result of forces

10. Kant, *Critique of Judgement* (ref. 2), 550, §61.
11. Timothy Lenoir, *The Strategy of Life* (Chicago: University of Chicago Press, 1982), 229–30.

> active in inorganic nature which are specially modified through the manner of their interaction.[12]

Clearly, he believed he had demonstrated the latter. And so did many others. Three years later, Emil du Bois-Reymond wrote:

> It can therefore no longer remain doubtful, what is to be made of the question, whether there exists a single recognizable difference between the processes of inorganic and organic nature. No difference exists. No new forces are attributable to organic forms, no forces which are not also active outside them. There are no forces therefore, which deserve the title *Lebenskräfte*. . . . There are no *Lebenskräfte* in this sense because the effects ascribed to them are to be reduced to those which originate from the central forces of material particles. There are no such *Lebenskräfte.*[13]

Lenoir claimed that "after 1842 no German physiologist worth his salt really believed that on physiological grounds it was necessary to invoke a special vital force to explain the functioning of the animal machine."[14] But the teleomechanists had introduced a crucial distinction between *Lebenskräfte* and *Zweckmässigkeit* (purposive organization). Yet no such distinctions survived in the heat of the debate that ensued. Lenoir continued, "However much they sought to avoid the appellation, as Helmhotz saw it, they were vitalists."[15]

Yet, for all the success of Helmholtz, du Bois-Reymond, and their followers, some resistance—some insistence on the special properties of organisms—did survive. It survived not in the defense of either *Lebenskräfte* or *Zweckmässigkeit*, but by recourse to a seemingly more physicalist language of stability, equilibrium, and fixity (and later, of steady states) in describing features of biological regulation, a language having recourse to the preoccupations of both the new physics and new developments in chemistry. Perhaps especially important was the attention to temporal dynamics awakened by studies of chemical reactions. In this context, Paley's rock could no longer serve adequately to signal the gap between inanimate and animate, as avatar of inanimate stability, for now the temporality of chemical activity became crucial to the organization of both living and (at least some) non-living systems. Where then, if anywhere, was the difference between them to be located?

12. Hermann Helmholtz, "Über den Stoffverbrauch bei der Muskelaktion," *Archiv für Anatomie und Physiologie* (1845), 72–83, on 72, quoted in Lenoir, *The Strategy of Life* (ref. 11), 200.

13. Emil du Bois-Reymond, *Untersuchungen über thierische Elektricität* (Berlin: G. Reimer, 1848), xliii–xliv, quoted in Lenoir, *The Strategy of Life* (ref. 11), 216.

14. Lenoir, *The Strategy of Life* (ref. 11), 232.

15. Ibid., 236.

LIFE, DEATH, AND THE AMBIGUITIES OF STABILITY AND EQUILIBRIUM

In the second half of the nineteenth century, stability, equilibrium, and fixity were crucial concepts for a number of philosophers and scientists as they formulated general principles for all systems, and at the same time characterized the distinctive features of living systems. Yet only rarely does one find a serious effort to distinguish the kinds of stability or equilibria discussed by physicists and chemists from those seen in biological systems. Herbert Spencer came closest in his *First Principles*, published in 1862. Arguing for a general principle of evolution from the inherently unstable homogeneous to a more stable heterogeneity, in order to distinguish the evolution of organic systems from that of inorganic systems, he introduced a distinction between mechanical instability (exemplified by a stick standing on its end) and "dependent moving equilibrium," a state best illustrated by the living organism in which the state of equilibrium is maintained by the organism's distinctively active responses. Citing the analogy of the steam engine, he wrote, "the organic functions constitute a dependent moving equilibrium—a moving equilibrium, of which the motive power is ever being dissipated through the special equilibrations just exemplified, and is ever being renewed by the taking in of additional motive power."[16] Spencer was clearly conversant with the physics of his time, and he struggled valiantly to reconcile an organic principle that maintains an organism in a "balanced state" with evolution's inevitable drift toward equilibrium. But he made no explicit reference to the second law of thermodynamics in this work, and indeed, his arguments straddle that law with telling ambiguity.

Eleven years later, Gustav Fechner (1873) argued for the universality of another principle, one he called the principle of a "tendency toward stability." In his early years, Fechner had been a professor of physics in Leipzig, but by the time he formulated this principle, his interests had broadened dramatically, reaching a breadth comparable to Spencer's. Like Spencer, Fechner sought a principle that would apply to all systems, both organic and inorganic, yet at the same time would be capable of distinguishing between the two domains. Also like Spencer (and despite its growth in prominence over the intervening years), Fechner made no mention of the second law of thermodynamics, and the relation of his principle to that law is similarly ambiguous. Even if

16. Herbert Spencer, *First Principles* (http://etext.lib.virginia.edu/toc/modeng/public/SpeFirs.html, 1862), §173.

Fechner was not familiar with Spencer's work (he made no mention of Spencer in this text), resemblance between the two principles is striking.

Michael Heidelberger has urged us to see Fechner's arguments as a forerunner to modern theories of self-organization, but more noteworthy is his attribution to Fechner of a distinction between open and closed systems. According to Heidelberger, the crucial difference between organic and inorganic is whether the system is open or closed. A closed system (i.e., a system subject to constant forces), merely as a consequence of "its inner forces," "gradually approximates a so-called stable state, without relapsing, meaning that it approximates a state in which its parts periodically . . . return to the same positions and the same relations of movement regarding one another."[17] By contrast, the principle of the tendency to stability has only approximate validity in open systems (like living beings). As Heidelberger put it, such a system "struggles to maintain its approximate stability vis à vis the external world to the greatest possible extent."[18] Only when the system collapses into an inorganic state (i.e., in death) does the system achieve absolute stability. In this way, Fechner's principle simultaneously unites organic beings with inorganic systems and differentiates them. Where closed, inorganic systems achieve stability naturally, "by their inner forces," organic systems, by contrast, must struggle to achieve what both is and is not the same end; that is, their end is crucially only an approximate version of the stability of closed systems. Both the unattainability of full stability and the active striving for approximate stability are thus defining features of organic systems.

Finally, in an effort to integrate the psychic realm with the material, Fechner proposed that the physical tendency to stability provides the basis of a psychical tendency to pleasure, and indeed sometimes referred to the latter as the "principle of pleasure." "The physical tendency toward stability," he argued, is "the carrier of a psychical tendency to create and maintain just those states, from which the physical originates."[19]

The connection to Freud's "pleasure principle" is hardly incidental. Freud was greatly influenced by Fechner, and he duly acknowledged his debt to his predecessor. He was especially impressed by Fechner's hypothesis that "every psychophysical motion rising above the threshold of consciousness is attended

17. Michael Heidelberger, *Nature from Within: Gustav Theodor Fechner and His Psychophysical Worldview* (Pittsburgh: University of Pittsburgh Press, 2004), 250.

18. Ibid., 251.

19. Ibid., 254.

by pleasure in proportion as, beyond a certain limit, it approximates to complete stability, and is attended by unpleasure in proportion as, beyond a certain limit, it deviates from complete stability."[20] But in his appropriation of Fechner's arguments, all the uncertainties in the meaning of stability, and especially its ambiguous relation to life on the one hand and death on the other, came home to roost. Rewriting *stability* as *constancy*, Freud changed Fechner's striving for stability (e.g., in the form of rhythmic cycles of excitation) to a striving for the reduction (if not cessation) of excitation, thus paving the way for a reformulation that changes Fechner's principle of living organization to the death instinct. Interestingly enough, the net effect (at least from today's perspective) was to leave Fechner and Freud on opposite sides of the question about the relation between living systems and the second law of thermodynamics.

All of these considerations were speculative in the extreme, and the obvious question arises as to what if any influence they had on the experimental study of living systems that had been built so successfully in the nineteenth century. The answer is probably very little. For many, the rise of experimental biology brought with it a major curtailing of such grandiose intellectual ambitions, confining attention to questions that experimental analysis might have something to say about. But as we can see, large-scale speculation did not come to a halt; rather, it led a parallel life under the name of "philosophy," running side by side and only occasionally interacting with scientific practice.

Claude Bernard provides a case in point, illustrating the precarious relation between philosophical speculation and experimental science. Bernard perhaps did more than any other to establish the tradition of experimental physiology in the nineteenth century, but he was also a philosopher. His work on self-regulation and his theory of the fixity of the *milieu intérieur* (a principle he regarded as the most important of his career) straddled his dual identity in interesting ways. In his *Introduction à l'étude de la médecine expérimentale*, written in the early 1860s, and then more fully in a series of lectures delivered towards the end of his life, Bernard elaborated on his principle of a constant internal environment. Although there is no evidence that he was directly influenced by either of his predecessors, Bernard's principle bears an obvious if unremarked resemblance both to Spencer's "dependent moving equilibrium" and to Fechner's "approximate stability." The physiologist, however, elevates his principle to something approaching a defining feature of life, at least for higher organisms. "La fixité du milieu intérieur est

20. Sigmund Freud, *Complete Psychological Works of S. Freud,* ed. J. Strachey, vol. 18 (London: Hogarth, 1955), 8–9.

la condition de la vie libre," he wrote. But the principle was hardly restricted to higher organisms: "All the vital mechanisms, however varied they may be, have only one object, that of preserving constant the conditions of life in the internal environment."[21]

Bernard's concern was with self-regulation rather than "self-organization," yet in many respects his arguments echoed Kant's own hopes for a mechanistic but not reductionist biology, and on this point, Bernard tended to agree. As William Coleman has observed, "The crucial matter was to accept the purposiveness of life as a principle of analysis, but also to insist upon the legitimacy and necessity of approaching the living organism in a deterministic manner."[22] For all his emphasis on the importance of physical and chemical analysis, Bernard, like Kant, was convinced of the ultimate inadequacy of physical and chemical analysis of the component parts, and correspondingly ambivalent about the necessity of some sort of teleological principle for explaining both the formation and the endurance of "living machines."[23] It is possible that the narrowing of his focus to self-regulation followed directly from his disciplinary orientation as a scientist; but for readers of his texts, it also had the effect of drawing him closer to the concerns of writers like Fechner (with whom he seems to have had no interaction), and (more importantly for this essay) to the concerns of those twentieth century authors to whom I will next turn.

But before leaving the nineteenth century, some further comment is in order regarding the language of fixity, stability, steady states, and equilibria, and the absence of clear definitions of these terms. It is true that both Fechner and Spencer made some efforts in this direction, but such subtleties were widely ignored. As a consequence, distinctions that loom so large in hindsight were given little attention, and for some authors of that time (e.g., Freud) they failed to register altogether. Indeed, the use of an undemarcated family of terms for a wide range of different phenomena may have been essential to the generalizing impulse behind these different projects.

21. Claude Bernard, *Leçons sur les phénomènes de la vie, communs aux animaux et aux végétaux*, publié par A. Dastre, 2 vols. (Paris: Baillière, 1878–1879), 113, 121.

22. William Coleman, "The Cognitive Basis of the Discipline: Claude Bernard on Physiology," *Isis* 76 (1985): 53.

23. "[W]hat distinguishes a living machine is not the nature of its physico-chemical properties, complex as they may be, but rather the creation of the machine which develops under our eyes in conditions proper to itself and according to a definite idea which expresses the living being's nature and the very essence of life" (Claude Bernard, *An Introduction to the Study of Experimental Medicine* (New York: Henry Schuman, Inc., 1927), 93.

More specifically, I suggest that the absence of distinctions might also have been useful, perhaps especially to Bernard and his readers, in allowing assimilation with the physical and chemical sciences even while asserting the distinctiveness of biology. Indeed, once a "law" of chemical equilibrium was formulated by the French chemist Henri Louis Le Chatelier in 1884, it seemed to more than a few that Bernard's principle had found a chemical foundation. Le Chatelier's principle stated that

> Every system in chemical equilibrium, under the influence of a change of any single one of the factors of equilibrium, undergoes a transformation in such direction that, if this transformation took place alone, it would produce a change in the opposite direction of the factor in question.[24]

Taken alone, without the qualifications that followed, one would have been hard pressed to distinguish this principle from at least some of the formulations of Bernard's principle (for example, Léon Fredericq's version that "the living being is an agency of such sort that each disturbing influence induces by itself the calling forth of compensatory activity to neutralize or repair the disturbance").[25] It is small wonder, then, that later biologists sometimes referred to Le Chatelier's principle as providing a direct connection between biological stability and chemical stability.[26]

Yet despite the confusion to which this sometimes gave rise, I hazard to suggest that the failure to clearly mark the difference between biological and chemical "equilibria" was not indicative of a lapse, but rather that it was a positive expression of a profound ambivalence about the relation between organisms and physico-chemical systems that persisted through the second half of the nineteenth century.

Of the two great laws to emerge from nineteenth-century physics, the relation of biological systems to the first, the conservation of energy (the first law of thermodynamics), was settled relatively quickly. Energy may not be (indeed, is not) conserved within the organism per se, but it is within in a closed system in which the organism is embedded. Accordingly, conservation of energy requires that the organism itself be characterized (either implicitly or explicitly)

24. Alfred J. Lotka, *Elements of Physical Biology* (Baltimore: Williams and Wilkins Company, 1925); reprinted as *Elements of Mathematical Biology* (New York: Dover Publications, 1956), 281.

25. Ibid., 284.

26. Ibid., 283–85.

as an open rather than a closed system. But would it then still qualify as a "self"? To many, the answer was apparently yes, for organisms do more than convert fuel into sustenance; they also self-regulate and self-organize, and these properties could be seen (and were seen by many) as continuing to distinguish living beings from physical systems.[27] These issues arose not in relation to the first law of thermodynamics, but to the second law. Entropy, organization, stability, equilibrium—all of these concepts raised immensely difficult problems for reconciling biological processes with the new science of thermodynamics. Indeed, in the absence of clear definitions, from either biologists or physicists, the questions could not even be well formulated. Yet at the time, a certain amount of ambiguity was in fact probably useful—if nothing else, it made it possible to avoid insoluble problems.

ONCE MORE, A WEDGE BETWEEN PHYSICS AND BIOLOGY

It was not until the late 1920s and early 1930s that a clear distinction between equilibrium (as understood in the physical sciences) and steady states (as used in the biological sciences) was articulated; and for the most part, it was a distinction drawn not by biologists, but by chemists and biophysicists. The inadequacy of the strict assumption of equilibrium for living systems had clearly begun to register already by the mid-1920s, but it was still hoped by many that such an assumption would not be so far off as to be useless. Only in the late 1920s, apparently forced by a set of empirical studies of systems maintaining constant values despite being far from equilibrium, did an explicit marking of distinctions between equilibrium and steady state, and between equilibrium and non-equilibrium, come to seem necessary. The physical chemist Frederick George Donnan offered a particularly clear statement of the difference in 1928 in his address to the British Association for the Advancement of Science: Living beings

> do not live and act in an environment which is in perfect physical and chemical equilibrium. It is the non-equilibrium, the free and available energy of the

27. As an aside, I might point out that even now, the juxtaposition of the requisite openness of the biological system with the property of being "self-organizing" highlights a fundamental ambiguity in the notion of "self"—an ambiguity suggesting an essential relationality to the processes of "self-regulation" and "self-organization." This ambiguity will resurface time and again in the second half of the twentieth century; but in the nineteenth century, it does not seem to have yet been made explicit.

> environment, which is the sole source of their life and activity. As Bayliss so finely put it, equilibrium is death.[28]

Based on the work of the biophysicist A. V. Hill (who had shared the Nobel prize with Otto Meyerhoff in 1922 for their work on the energetics of carbohydrate catabolism in skeletal muscle), Donnan concluded that the living cell "requires constant oxidation to preserve the peculiar organization or organized molecular structure of the cell. The life machine is therefore totally unlike our ordinary mechanical machines. Its structure and organization are not static. They are in reality dynamic equilibria, which depend on oxidation for their very existence."[29] When Donnan's address was published in *Nature* later that year, the accompanying editorial drew the obvious implications about the essential relationality of the living organism, sharply distinguishing the picture Donnan drew from the emphasis on autonomy implied in Bernard's conception of the fixity of the internal environment as maintained *in spite of* external variations: "The fundamental distinction between the living and the non-living is that whilst it is possible to isolate the phenomena of the inorganic world, it is impossible to consider a living organism apart from its environment; it is, in fact, its reactions and adaptations to changes in its surroundings which distinguish the living from the inanimate and form the basis of the science of biology."[30] In fact, one might paraphrase the claim being made here by saying that the autonomy of the living organism, the very property that invites its characterization as a "self," is itself a product of the organism's relation to its environment.

In a volume published three years later, A. V. Hill provided a fuller elaboration of his "conception of the steady state."[31] Reflecting, for example, on studies of osmotic pressure in shore crabs (performed in 1929), Hill found himself "forced to admit that the tissues and membranes are standing ready for immediate change,

28. F. G. Donnan, "The Mystery of Life," *Nature* 122 (1928): 512–14, on 512.

29. Ibid., 514.

30. "Life and death," *Nature* 122 (1928): 501–03, on 501. The contrast is especially sharp in L. J. Henderson's reading of Bernard's principle. As J. S. Haldane wrote, "Henderson treats the constancy of reaction in the living body as if it depended on the physico-chemical properties of blood. In actual fact this constancy depends during health on the coordinated activity of the kidneys and respiratory organs. . . . Not all the buffering in the world would keep the reaction constant otherwise, though the buffering greatly smooths the regulation." J. S. Haldane, "Claude Bernard's Conception of the Internal Environment," *Science* 69 (1929): 453–54.

31. A. V. Hill, *Adventures in Biophysics* (Philadelphia: University of Pennsylvania Press, 1931), 55.

in response to an alteration of the environment: and yet, the steady state in which they are is very far from equilibrium. It must be that a control of some kind is being exerted on the passage of dissolved substances . . . in and out of the animal; somehow, by some positive active process, the concentration of the blood and of the environment are being held apart." The question is, "How is this strange and interesting inequality of concentration maintained?" It is not clear what physics can tell us here—indeed, he observes in a footnote, "It is not always safe to argue from physics to physiology."[32] There is a crucial difference between the steady states that are the result of passive diffusion (with which physics can deal) and those resulting from active transport (with which it cannot). At the end of his extensive review of the different kinds of steady states that had been observed (not all of which are far from equilibrium), Hill concludes:

> The problem is, in a sense, a single one in all these cases—we are not dealing with a thermodynamic equilibrium at one end of the scale, with an organized factory at the other. Throughout we are involved, not with genuine equilibrium, but with conditions maintained constant by delicate governors and by a continual expenditure of energy. How that energy is supplied, how it is utilised to maintain the structure and the organisation, is, I think, the major problem of biophysics.[33]

Ludwig von Bertalanffy is often credited with the argument that living systems should be thought of as open, non-equilibrium systems, in contrast to the more familiar closed, equilibrium systems of physics and chemistry that percolated through the literature; but by the time he was writing, the argument was already established in mainstream scientific literature.[34] From the 1930s on, a shift in notions of stability that were seen as relevant to biological systems, from those of equilibrium systems to those characteristic of steady state,

32. Ibid., 68, 69, 79.

33. Ibid., 79.

34. Ibid.; Ludwig von Bertalanffy, "Der Organismus als physikalisches System betrachtet," *Naturwissenschaften* 28 (1940): 521–31. In his later years, von Bertalanffy would probably have preferred being introduced in the next section, for, with the rise of cybernetics, he often attempted to claim priority for many of the arguments for the possibility of "self-organizing" machines. He appears here because it is clear that, in his early work, he vigorously resisted any such assimilation between organisms and machines. For example, in Ludwig von Bertalanffy, *Modern Theories of Development*, trans. J. H. Woodger (Oxford: Oxford University Press, 1934), he argued that it is evident that "every machine is where and what it is for a definite purpose, and that it presupposes the engineer who has conceived and constructed it" (pp. 37–38). Here, self-regulation and self-organization are, as they were for Kant, defining features of living systems.

non-equilibrium systems, became increasingly well established in the biochemical and biophysical literature.[35] Yet even with this distinction, an ambiguity still persisted, for, as Hill remarked, there are many kinds of steady states; and even though he proposed that they should all be thought in terms of control and regulation, for others more committed to the adequacy of conventional physics the term steady state could also be used (and was used) to refer primarily to those states resulting from passive (rather than from active) transport. In fact, as I will discuss in Part Two of this essay, much the same ambiguity returns to seriously haunt a later generation of thinkers on these issues.

But for now, it is Walter B. Cannon, a bridge figure par excellence, who must claim our attention. Cannon was a physiologist in the mold of Bernard, and strongly influenced by his nineteenth century precursor, but his particular intervention provided a crucial stepping-stone for the radical shift in terms that was to come in the mid-twentieth century. His contribution to this discussion was in much the same spirit as Hill's, even though coming from a different tradition. Cannon introduced the term "homeostasis" in 1926, precisely to distinguish the kind of stability maintained by biological systems from simple physico-chemical equilibria (and perhaps from simple states of passive diffusion as well). Three years later, he published a fuller account of what he intended the term to cover.[36] There, he acknowledged Bernard's priority for "the central idea" and explained the need for a new term as follows: [37]

> The highly developed living being is an open system having many relations to its surroundings. . . . Changes in the surroundings excite reactions in this system, or affect it directly, so that internal disturbances are produced. Such disturbances are normally kept within narrow limits, because automatic adjustments within the system are brought into action, . . . and the internal conditions are

35. See A. C. Burton, "The Properties of the Steady State Compared to Those of Equilibrium as Shown in Characteristic Biological Behavior," *Journal of Cellular and Comparative Physiology* 14 (1939): 327–49; Rudolf Schoenheimer, *The Dynamic State of Body Constituents* (Cambridge, MA: Harvard University Press, 1942); J. M. Reiner and S. Spiegelman, "The Energetics of Transient and Steady States, With Special Reference to Biological Systems," *Journal of Physical Chemistry* 49 (1945): 81–92.

36. W. B. Cannon, "Organization for Physiological Homeostasis," *Physiology Review* 9 (1929), 2.

37. Cannon paid somewhat fuller tribute to Bernard in the French translation of *The Wisdom of the Body*. There he wrote, "The central idea of this book, 'the stability of the inner medium of the organism in higher vertebrates,' is directly inspired by the precise views and deep understanding of the eminent French physiologist Claude Bernard. This book can even be considered a tribute to his memory." L. L. Langley, *Homeostasis* (New York: Van Nostrand Reinhold Co., 1973), 2.

> held fairly constant. The term "equilibrium" might be used to designate these constant conditions. That term, however, has come to have exact meaning as applied to relatively simple physico-chemical states in closed systems where known forces are balanced. . . . . The present discussion is concerned with the physiological rather than the physical arrangements for attaining constancy. The coordinated physiological reactions which maintain most of the steady states in the body are so complex, and so peculiar to the living organism, that it has been suggested that a specific designation for these states be employed—*homeostasis.*[38]

Hill's emphasis on the "strange and interesting" properties of physiological steady states, on the "positive active process" by which they were maintained, and Cannon's introduction of a new term to mark the distinctiveness of biological dynamics, might be said to mark the end of an era—an era that had begun with Kant's coining of *self-organization* to signal the distinctiveness of organisms, and ends with Cannon's use of *homeostasis* to much the same end. Kant had no doubt about the exceptionality of living systems, nor did Bernard, nor, I believe, did Cannon. Despite all the successes of Helmholtz and his colleagues, despite the fierce campaigns mounted against anti-reductionism that ensued, and despite frequent references to Darwin as the "Newton of biology," there remained considerable question about the extent to which biological systems could be assimilated with the kinds of systems with which physics and chemistry deals. To many physiologists in the early part of the twentieth century (and certainly to most embryologists), the prospects of a "Newton for a blade of grass" would have seemed as remote as it had for Kant. Darwin may have provided a mechanism for the evident order among the kinds of living things in the world, but the question of how to account for the order (or organization) of an individual organism remained opaque.

As to the divide between organisms and man-made products, that still held; organisms were not yet to be confused with machines. Certainly, attributes of machines were employed in descriptions of organisms, but direct comparisons were few and far between. For example, the idea that the maintenance of life required compensation (or regulation) was well accepted by the end of the nineteeth century, and indeed, it had been around for more than a century. In 1790, Lavoisier had clearly described the "compensations" that enable an organism to maintain equilibrium—for example, how the rate of respiration is regulated in response to changes in temperature and levels of exertion: "The animal machine," he had written, "is governed principally by three main regulators: respiration

38. Cannon, "Homeostasis" (ref. 36), 400.

which consumes hydrogen and carbon and provides the caloric; perspiration, which increases or decreases, depending on whether more or less caloric is needed; and finally, digestion, which restores to the blood what it has lost through respiration and perspiration."[39] But even though Watt's steam engine governor had been introduced two years before Lavoisier's, to my knowledge neither he nor his nineteenth century colleagues seem to have considered the governor of the steam engine as a model (either analogical or homological) for physiological regulation. Helmholtz invoked the steam engine to illustrate "the manner in which [the animal body] obtains heat and force";[40] Spencer used it to illustrate his notion of "dependent moving equilibrium";[41] and in 1858 Alfred Wallace employed it to relate it to the evolution of species.[42] But the first attempt to explicitly integrate thinking about mechanical control and regulation with experimental studies of physiological control and regulation that I have been able to find appears only in the twentieth century. The reference appears in Hill's book of 1931, but in a form that can still be read as analogical. The positing of the kind of homological relation that has become so ubiquitous in today's literature—a relation strong enough to entirely negate the divide between organisms and machines that was taken for granted by Kant and his followers, and even by Cannon and (the early) von Bertalanffy—that, I will argue, comes only in the late 1940s, with the rise of cybernetics.

A NEW WORLD ORDER: MECHANIZING SELF-ORGANIZATION

The steam engine is often credited with spawning the science of thermodynamics, but the science (or technology) of the regulation of that steam engine (as distinct from its conservation of energy) had little impact on nineteenth-century physics and chemistry. There is even less evidence of its impact on the biological sciences. For all the uncertainty about the relation between organisms and physical-chemical systems that pervaded nineteenth century thinking in the biological sciences, no comparable uncertainty seemed to surround the divide between organisms and machines.

39. Lavoisier, *Oeuvres*, vol. II, 700, http://historyofscience.free.fr/Lavoisier-Friends/a_ chap4_lavoisier.html

40. Hermann von Hemholtz, *Popular Scientific Lectures* (New York: D. Appleton, 1854), 183.

41. Spencer, *First Principles* (ref. 16), §173.

42. Alfred Wallace, "On the Tendency of Varieties to Depart Indefinitely From the Original Type (S43: 1858)," http://www.wku.edu/~smithch/wallace/S043.htm: 62.

The idea that mechanical systems could be designed to be self-regulating was already well established by the dawn of that century. Devices to regulate the direction of the windmill and the flow of grain had been around for a hundred and fifty years; temperature regulators were even older. Float regulators were of course far older still—although long forgotten, they had been used in ancient times for water clocks and other applications. But these too were revived in the eighteenth century. And James Watt's centrifugal fly-ball governor stimulated a veritable industry of such control devices in the nineteenth century. In other words, mechanical regulators were plentiful when Bernard was writing, and had even begun to appear at the time Lavoisier was writing almost a century before. Yet if there is evidence of a link having been drawn between biological and mechanical regulation during that entire period, it has eluded me. I think it is fair to say that the link was simply not made.

Of course, other kinds of connection *were* being made; in particular, the analogy between Watt's governor (and before Watt's, to Plato's metaphor of a steersman) and social regulation was frequently drawn in the nineteenth century, and indeed, efforts to mechanize social organization and regiment the workplace in the nineteenth century led to widespread protest (not to mention a rich body of critical literature that has accumulated since). It was in this context that the term *cybernetics* first made its appearance. In 1834, the physicist André Ampère referred to his vision of the art of governing nations as *cybernétique*, and in 1843, the Polish philosopher Bronislaw Trentowski published a volume on the art of governance, calling his treatise *Cybernetyka.* But it would appear that Bernard never imagined linking his work on biological regulation to the increasing presence of control mechanisms, and it is even doubtful that Cannon, writing half a century later, did either (although Cannon did leave the door open for the presence of homeostasis in social and industrial systems).[43] On the engineering side, David Mindell argues that in the first half of the twentieth century, "engineers building feedback devices had little interaction with biologists and their ideas of homeostasis."[44] So the question arises, when did

43. "It seems not impossible that the means employed by the more highly evolved animals for preserving uniform and stable their internal economy (i.e., for preserving homeostasis) may present some general principles for the establishment, regulation and control of steady states, that would be suggestive for other kinds of organization—even social and industrial—which suffer from distressing perturbations." (W. B. Cannon, *The Wisdom of the Body* (London: Kegan Paul, 1932).

44. David Mindell, *Between Human and Machine: Feedback, Control, and Computing before Cybernetics* (Baltimore: Johns Hopkins University Press, 2002), 13.

the connecting lines between control theory and homeostasis begin to be drawn? And why then but not before?

Hints of a vital realignment between organisms, machines, and physico-chemical systems began to appear in the 1920s and 30s. But even though some connections had already begun to be drawn in the early 1930s (see, e.g., A. V. Hill's reference to "delicate governors," quoted above), it was the intense concentration of technical efforts in World War II that cemented that realignment, and it was Norbert Wiener who gave the connection both an explicit formulation and wide currency. "Out of the wickedness of war," as Warren Weaver put it, emerged not only a new machine, but a new vision of a science of the inanimate, a science based on principles of feedback and circular causality, and aimed at the mechanical implementation of exactly the kind of purposive organization of which Kant had written and that was so vividly exemplified by biological organisms; in other words, a science that would repudiate the very distinction between organism and machine on which the concept of *self-organization* was originally predicated. Following on the one hand the writings of Claude Bernard, Walter Cannon, and numerous others, and on the other hand the use of feedback in the development of servo-mechanisms, Norbert Wiener re-coined the term *cybernetics* in 1948 to name this new science of feedback, control, and communication. Wiener's vision of not simply the analogy between animals and machines, but of the close homology (if not identity) between the two, became the focal point for the interests of many of the mathematicians, physicists, and engineers who had been actively working on feedback, control, and communication throughout the war years, as well as drawing in a number of newcomers. It also attracted a large and enthusiastic public following.

In his book of 1948, *Cybernetics: or, Control and Communication in the Animal and the Machine*, Wiener wrote, "[O]ur inner economy must contain an assembly of thermostats, automatic hydrogen-ion-concentration controls, governors, and the like, which would be adequate for a great chemical plant. These are what we know collectively as our homeostatic mechanism."[45] A few years later, he elaborated the connection: "Walter Cannon, going back to Claude Bernard, emphasized that the health and even the very existence of the body depends on homeostatic processes . . . that is, the apparent equilibrium of life is an active equilibrium, in which each deviation from the norm brings on a

45. Norbert Wiener, *Cybernetics: or, Control and Communication in the Animal and the Machine* (Cambridge, MA: MIT Press, 1948), 113.

reaction in the opposite direction, which is of the nature of what we call negative feedback."[46]

Yet while Wiener worked to dissolve the boundary between organisms and machines, he also reasserted the distinction between biological processes and the rest of the physical and chemical universe, emphasizing that "the apparent equilibrium of life is an active equilibrium." Elsewhere in the same text he said,

> Life is an island here and now in a dying world. The process by which we living beings resist the general stream of corruption and decay is known as *homeostasis*.
>
> We can continue to live in the very special environment which we carry forward with us until we begin to decay more quickly than we reconstitute ourselves. Then we die.[47]

In other words, Wiener accepted the consensus that had come to prevail in the 1920s and 30s about the paradox that living systems seemed to pose in relation to the second law of thermodynamics: While there appeared no reason to assume that organisms, taken together with their environmental context, violate the second law, it seemed clear to virtually all commentators that they embody a form of organization that locally resists the action of the second law. Just how this is achieved, as Schrödinger had so eloquently put it just a few years earlier, to "feed upon" the order in its environment, remained the dominant puzzle.[48] But unlike Schrödinger, Wiener's approach was to look not to quantum mechanics, but to control theory in engineering—to a tradition that seems, virtually by definition, to bypass the ambiguities that had plagued earlier discussions between stability, constancy, and equilibrium. Indeed, defined either as the process by which the stability of a system is ensured through time, or as the means by which system identity is sustained in an uncertain environment, the concept of control has come to our ears to seem ideally suited to the description of biological processes.

Wiener did not, however, invent the field of activity that he aptly named and so articulately promoted; many others had been working along closely parallel lines throughout the 1940s. The work of W. Ross Ashby, an English psychiatrist who was a major influence in the development of postwar cybernetics, was especially critical, particularly for his tackling of the notion of self-organization. In 1947 he published a paper entitled, "Principles of the Self-organizing Dynamic System," which began with the observation, "It has been

46. Ibid., 251.

47. Ibid., 95–96.

48. Erwin Schrödinger, *What is Life? The Physical Aspect of the Living Cell & Mind and Matter* (Cambridge: Cambridge University Press, 1967), 75–76.

widely denied that a machine can be 'self-organizing,' i.e., that it can be determinate and yet able to undergo spontaneous changes of internal organization."[49] Ashby's primary concern was with the nervous system, which, he argued, had both these properties—it was a determinate physico-chemical system *and* it could undergo "self-induced" reorganizations that result in a change in behavior. His aim was to put to rest the widespread suspicion that it could not be both, and the remainder of the paper was devoted to an argument purporting to show "that a machine can be at the same time (a) strictly determinate in its actions, and (b) yet demonstrate a self-induced change of organization."[50] One year later, another paper reported his success in actually building a machine that could satisfy these dual criteria, and that therefore (at least in Ashby's mind) could serve as a primitive model for a brain. He called his machine a *homeostat*.[51]

The inspiration for Ashby's homeostat came not so much from studies of physiological regulation as from the study of behavioral adaptation, and especially from the link that biologists had frequently drawn between adaptation, survival, and "equilibrium." Ivan Pavlov, for example, had written:

> The animal must respond to changes in the environment in such a manner that its responsive activity is directed towards the preservation of its existence. This conclusion holds also if we consider the living organism in terms of physical and chemical science. Every material system can exist as an entity only so long as its internal forces, attraction, cohesion, etc., balance the external forces acting upon it. . . . Being a definite circumscribed material system, it can only continue to exist so long as it is in continuous equilibrium with the forces external to it.[52]

In other words, animals adapt to changes in their environment with behavior that works to restore the system to normality. Strictly speaking, Ashby's use of the term *homeostat* stretches the meaning of homeostasis beyond that intended by Cannon; and insofar as it would be used as a model for "self-organization," it recouped at least part of the larger meaning of the term as Kant had used it.

Ashby's aim was to automate such adaptive behavior. In this effort he translated this principle of adaptation into the assumption that the organization of a system would be modified according to the results obtained. In the simplest

49. W. R. Ashby, "Principles of the Self-organizing Dynamic System," *Journal of General Psychology* 37 (1947): 125–28.

50. Ibid., 125.

51. W. R. Ashby, "The Homeostat," *Electron* 20 (1948): 380.

52. Ivan Pavlov, "Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex, Lecture One" (1927), http://www.ivanpavlov.com/lectures/ivan_pavlov-lecture_001.htm

form, a negative result would trigger a random restructuring, while a positive result would signal no change. Or, to put it differently, "Heads I win, tails we start again." Of course, not everyone was so impressed by Ashby's homeostat. After all, what did it do? This was a machine for doing nothing, or, as Grey Walter put it, a *machina sopora*, a sleep machine. But to Ashby this was to miss the central point of the device.[53] And that was that a reorganization (rewiring, in this case) would be automatically triggered every time the system departed from its range of stability, thereby guaranteeing that it would *always* reach a stable state no matter how serious the perturbation of the inputs. It was not merely stable, it was "ultrastable."

This notion of *ultrastability* was of great importance to Ashby. It marked a crucial distinction between the capacity of his device and that of others being offered up as candidates for self-organizing machines. For example, this was not simply a case of a mechanical self-assembly in which an unorganized system becomes organized. Rather, here was a machine that could automatically change its organization in a positive direction—it could spontaneously shift from a "bad" organization, or way of behaving, to a "good" one. It could adapt. He later wrote:

> A well-known example is the child that starts with a brain organization that makes it fire-seeking; then a change occurs, and a new brain organization appears that makes the child fire-avoiding. Another example would occur if an automatic pilot and a plane were so coupled, by mistake, that positive feedback made the whole error-aggravating rather than error-correcting. Here the organization is bad. The system would be "self-organizing" if a change were automatically made to the feedback, changing it from positive to negative; then the whole would have changed from a bad organization to a good.[54]

This, he claimed, was precisely what his homeostat did. It showed that a properly designed machine could exhibit autonomous, self-organizing behavior of just the kind that animals displayed.[55]

But, he cautioned, there is a problem, for strictly speaking, "no machine can be self-organizing in this sense."[56] The appearance of self-organization is in fact a product of the relationship a machine has with another machine, or, more

53. Grey Walter, *The Living Brain* (New York: W. W. Norton, 1953), 123.

54. W. R. Ashby, "Principles of the Self-organizing System," in *Principles of Self-Organization*, ed. H. von Foerster and G. W. Zopf (New York: Pergamon Press, 1962), 255–278.

55. See Andrew Pickering, "Cybernetics and the Mangle: Ashby, Beer and Pask," *Social Studies of Science* 32 (2002): 413–37 for further discussion of Ashby.

56. Ashby, "Principles" (ref. 54), 267.

familiarly, the environment from which it receives the necessary input. More precisely, only if the "self" is enlarged to include the system to which the machine is coupled can one speak of self-organization.

What then about the origin of life? Is this not an example of self-organization, perhaps the most obvious instance of self-organization; and did it not occur on an isolated planet? Well yes, but the living systems that emerged did so as sub-systems of a larger system, namely the planet, and it would be more accurate to speak of the self-organization of the planet that resulted in the appearance of such structures. Ashby was convinced that such analyses had crucial light to shed on the problem of the origin of life—indeed, that they turned conventional wisdom on its head:

> In the past, when a writer discussed the topic, he usually assumed that the generation of life was rare and peculiar, and he then tried to display some way that would enable this rare and peculiar event to occur. . . . The truth is the opposite—every dynamic system generates its own form of intelligent life, is self-organizing in this sense.[57]

Just what the basis of this claim was remains unclear (perhaps we should call it an article of faith), but it is certainly striking, especially in view of the fact that the same article of faith will resurface repeatedly at the end of the century (see Part Two).

In the meantime, others of Ashby's time (notably, Frank Rosenblatt and Heinz von Foerster) were taking somewhat different approaches to self-organization: Supported by the U.S. Office of Naval Research, at Cornell's Astronautical Laboratory, Rosenblatt was inspired by Donald Hebb's claim that an ensemble of neurons could learn if the connection between neurons were strengthened by excitation, and he attempted to build a machine that could learn as Hebb had suggested.[58] The Perceptron, a simple neuron-like learning device, is now seen as the first connectionist machine. Rosenblatt argued that it should be able to "perceive, recognize and identify its surroundings without human training or control." Inspired as well by the work of von Neumann in the late 1940s on self-reproducing automata, Rosenblatt went on to suggest that "[i]n principle, it would be possible to build Perceptrons that could reproduce themselves on an assembly line and which would be 'conscious' of their existence."[59]

57. Ibid., 270.
58. D. O. Hebb, *The Organization of Behavior* (New York: Wiley, 1949).
59. Frank Rosenblatt, "Electronic 'Brain' Teaches Itself," *New York Times*, 13 Jul 1958.

The commonalities among these different efforts were sufficiently strong to lead Marshall C. Yovits, director of a new branch of the Office of Naval Research on "Information Systems" in the late 1950s, to conclude that "self-organization" was a subject that had come of age and that it was of sufficient importance to warrant concentrating his ONR resources on it. Together with Scott Cameron, he hosted the first conference on self-organizing systems in 1959, thereby initiating a series of conferences that ran through the early 1960s. A major impetus behind these conferences was the Navy's hope of building a new kind of machine, more adept and more versatile—indeed, more autonomous—than the digital computer developed in the war years.[60] If one wanted to build a machine that could handle truly complex problems, that could learn from experience, the trick, they argued, would be to study the natural systems (i.e., living organisms) that did so with such obvious success.[61] As Yovits explained in his introductory remarks to the first conference,

> certain types of problems . . . can be solved efficiently only with the use of machines exhibiting a high degree of learning or self-organizing capability. . . . On the one hand the psychologist, the embryologist, the neurophysiologist and others involved in the life sciences were attempting to understand the self-organizing properties of biological systems, while mathematicians, engineers, and physical scientists were attempting to design artificial systems which could exhibit self-organizing properties.
>
> Accordingly, [we] decided to sponsor a conference to enable workers in the many disciplines to meet together to discuss [these problems].[62]

Yet, notwithstanding their multidisciplinarity, all of these discussions and investigations were conducted with an explicitly engineering aim, namely, the

60. The organizer of one of the early conferences on "self-organizing systems" suggested a slightly different aim, one he described as "man's dominating aim," namely "the replication of himself by himself by technological means" (C. A. Muses, ed., *Aspects of the Theory of Artificial Intelligence* [New York: Plenum, 1962], 114). He introduced the published volume with the suggestion that "Man Build Thyself?" might be a better title (v).

61. The U.S. military's interest in bionics had a similar basis. In 1960, Harvey Saley of the Air Force addressed a Bionics Symposium as follows: "The Air Force, along with other military services, has recently shown an increasing interest in biology as a source of principles applicable to engineering. The reason clearly is that our technology is faced with problems of increasing complexity. In living things, problems of organized complexity have been solved with a success that invites our wonder and admiration." Quoted in Geoffrey C. Bowker, "How to be Universal: Some Cybernetic Strategies, 1943–1970," *Social Studies of Science* 2 (1993): 107–27, on 118.

62. Marshall Yovits and Scott Cameron, ed., *Self-Organizing Systems* (New York: Pergamon Press, 1960), v–vi.

design and construction of systems that could organize themselves, grow themselves, and perhaps even reproduce themselves.

A second commonality of importance can be found in the general agreement that the crucial property for the realization of these goals (especially as Ashby, Pask, and von Foerster formulated it) lay in the relationality between components. It is out of such relations, the claim goes, that function and purpose emerge, presumably spontaneously. Von Foerster opened the first conference on self-organization with the provocative claim, "There are no such things as self-organizing systems."[63] Systems achieved their apparent self-organization by virtue of their interactions with other systems, with an environment. In a similar vein, Robert Rosen noted the "logical paradox implicit in the notion of a self-reproducing automaton,"[64] and in 1962, Ashby reiterated the main point: "the appearance of being 'self-organizing' can be given only by the machine *S* being coupled to another machine [α]. . . . Only in this partial and strictly qualified sense can we understand that a system is '*self*-organizing' without being self-contradictory." Indeed, "Since no system can correctly be said to be self-organizing, and since use of the phrase 'self-organizing' tends to perpetuate a fundamentally confused and inconsistent way of looking at the subject, the phrase is probably better allowed to die out."[65] The goal, for all of these participants, was to understand the logic of the coupling (or "conversation") between sub-systems that would generate the capacity of such a system to have its own objectives, that is, to be a natural system. In hindsight, von Foerster introduced the term "second-order cybernetics" to distinguish their efforts from the concerns of the early cybernetics, to denote the explicit preoccupation of the new cybernetics with the nature of relational and self-reflexive systems, and to cement an alliance with the work of Humberto Maturana and his student Francisco Varela. Thirty years later, Pask summarized the differences between the old and new cybernetics as a shift in emphasis, from information to coupling, from transmission of data to conversation, from stability to "organizational closure," from external to participant observation.[66] In short, the shift went from

63. Ibid., 31.

64. Robert Rosen, " On a Logical Paradox Implicit in the Notion of a Self–Reproducing Automata," *Bulletin of Mathematical Biophysics* 21 (1959): 387–94.

65. Ashby, "Principles" (ref. 54), 267, 269.

66. Gordon Pask, "Introduction: Different Kinds of Cybernetics," in *New Perspectives on Cybernetics: Self-Organization, Autonomy, and Connectionism*, ed. Gertrudis van de Vijver (Dordrecht: Kluwer, 1992), 24–25.

Wiener's "command, control, and communication" to an approach that could be assimilated to Maturana and Varela's concept of "autopoiesis."[67]

What came of their efforts? Not a great deal, at least not in the laboratories of mainstream science and engineering, and especially not in the short run. Throughout the 1960s some developmental biologists attempted to make use of this new alliance, but by the end of the decade that effort collapsed, as (at least in the U.S.) did cybernetics. The 1960s were the years in which molecular biology established its hegemony in biology, and by the end of the decade the successes of that effort lent its own agenda an authority with which it was not possible to contend. Even its appropriation of cybernetic language worked effectively to eclipse the inherently global questions of developmental biology, especially the question of what it is that makes an organism "self-organizing." The cybernetic vision also collapsed in computer science, where it gave way to competing efforts from what Daniel Dennett has called "High-Church computationalism."[68] The starting assumption of the latter was that, just like digital computers, minds are nothing more than physical symbol systems, and intelligent behavior could therefore be generated from a formal representation of the world. A fierce funding war raged between the two schools in the early 1960s, championed respectively by the ONR and the Air Force. By the mid-60s, with the award of more than ten million dollars from ARPA to Marvin Minsky (through Project MAC, created in 1963) for the development of Artificial Intelligence at MIT, it was effectively over. Funding for the Perceptron and its requisite hardware (especially for parallel processors) dried up (Yovits says "it was starved to death"[69]), and the ONR- supported conferences on Self-Organizing

67. See Humberto Maturana and Francisco Varela, *Autopoiesis and Cognition: The Realization of the Living* (Boston: D. Reidel, 1979). The main arguments for autopoiesis were developed in Chile in the late 1960s, but Maturana was a frequent visitor at von Foerster's "Biological Computing Laboratory" in Urbana during that period, and after 1968, so was Varela; indeed, Varela thanks von Foerster for his "role in the gestation and early days of the notion of autopoiesis." Francisco Varela, "The Early Days of Autopoiesis: Heinz and Chile," *Systems Research* 13 (1996): 407–416, on 407. Both Maturana and Varela remained close personal friends of von Foerster until his death in 2002, but in my reading of their work, there is a considerable distance between the theory of autopoisesis and the work of von Foerster and Pask, especially as originally represented under the banner of "self-organization."

68. D. C. Dennet, "The Logical Geography of Computational Approaches: A View from the East Pole," in *The Representation of Knowledge and Belief*, ed. M. Brand and R. M. Harnish (Tucson: University of Arizona Press, 1986), 59–79.

69. Personal communication, 1997.

Systems ceased. In 1969, Marvin Minsky and Seymour Papert published a devastating attack on the entire project, explaining their intent:

> Both of the present authors (first independently and later together) became involved with a somewhat therapeutic compulsion: to dispel what we feared to be the first shadows of a "holistic" or "Gestalt" misconception that would threaten to haunt the fields of engineering and artificial intelligence as it had earlier haunted biology and psychology.[70]

Shortly after (July 11, 1971), Frank Rosenblatt, that project's guiding inspiration, died in a boating accident that some say was a suicide.

But in fact, within the Artificial Intelligence community, early connectionism had already met its demise.[71] In the absence of funding, and of the kinds of concrete successes in embodying biological principles in non-biological systems that might have generated new alliances and new sources of funding, the union promised between organisms and computers failed to materialize. Neural net models, which had been the heart of Rosenblatt's effort, had to wait another fifteen years to be revived. Interactionist models of development had to wait even longer. Some have argued that the soil of American science in the 1960s was not quite right for this largely European transplant to take root. Even so, as Hubert and Stuart Dreyfus wrote, "blaming the rout of the connectionists on an antiholistic bias is too simple."[72] In a rather expansive sweep, they fault the influence of an atomistic tradition of Western philosophy dating from Plato to Kant. Von Foerster, by contrast, blamed both the funding policies of American science (the short funding cycle, the emphasis on targeted research) and the excessively narrow training of American students.[73] But there were surely other factors as well. There is no question of the intellectual vitality of these early meetings. Ideas, supported by a motley range of ingenious if ad hoc arguments, were generated in abundance, some of them touching on issues of

70. Marvin Minsky and Seymour Papert, *Perceptrons* (Cambridge, MA: MIT Press, 1969), 19.

71. See Hubert Dreyfus and Stuart Dreyfus, "Making a Mind Versus Modeling the Brain: Artificial Intelligence Back at a Branchpoint," *Daedalus* 117 (1988): 15–44, on 24. Von Foerster, however, was able to keep his center, the Biological Computing Laboratory, going for another five years, largely by forging new alliances with the National Institutes of Health and the Department of Education.

72. Dreyfus and Dreyfus, "Making a Mind" (ref. 71), 24.

73. Heinz von Foerster, interview with Stefano Franchi, Güven Güzeldere, and Eric Minch, SEHR 4:2: *Constructions of the Mind, Updated 26 June 1995,* http://www.stanford.edu/group/SHR/4-2/text/interviewvonf.html.

great profundity. But American science in the 1960s was less interested in clever, possibly even profound, quasi-technical, quasi-philosophical arguments than in methods of analysis that could be employed in a variety of settings, that could, in short, generate a research program. And these were clearly lacking. Looking back on these papers today, it is difficult to find a single example of a procedure that could be generalized.

To the extent that the work of Maturana and Varela on autopoiesis is to be associated with "second-order cybernetics" (although it is debatable how meaningful that category is, and how close the relation between concepts of autopoiesis and "self-organization" actually were), the project of the Chilean school must almost certainly be counted as the most philosophically ambitious, and in many ways the most successful, venture to come out of that association. Yet it must be acknowledged that, with the exception of Maturana's laboratory in Chile and Varela's in Paris, the primary domains in which their ideas found a positive reception lay far from the mainstream of biology, engineering, or even mainstream philosophy.

I would venture to say that, today, the main interest in notions of "self-organization" associated with "second-order cybernetics" is retrospective, arising from subsequent developments (like connectionism, robotics, and control theory) that had grown along largely independent trajectories. At the time, and for all the enthusiasm and inventiveness of Ashby, Rosenblatt, Yovits, von Foerster, and their colleagues, the efforts of this group to establish "self-organization" as a general concept, capable of unifying the animate and inanimate worlds, failed to capture the imagination of either the public or the larger scientific community. And certainly, no machine was yet built with anything like the capacities expected of the Perceptron; indeed, no one from that period succeeded in actually building a machine that could vindicate Wiener's confidence in a strict organizational homology between organisms and machines.

CONCLUSION

The story is clearly not over; indeed, even this part of the story is not over. In the U.S., the place of cybernetics was quickly filled by its close cousin, control theory, where yet another revolution was in its infancy. The launching of *Sputnik* in 1957 had awakened a generation of engineers and mathematicians to the advances that Russian mathematicians had been making in the analysis of nonlinear systems. The work of Popov and Lyapunov, among others, began to be made available to English-speaking readers, introducing an entirely new set of tools into control theory (at least as practiced in the West), stimulating the explosion of new mathematical work in non-linear dynamics, and transforming the analytic methods of mathematical physics as well. Little if any of this work bore on the nature of biological organization; but it prepared the ground for the emergence in the late 1970s of a new vision of "self-organization" in the biological realm, and yet another reconfiguration of the relations between organisms, machines, and spontaneously arising structures in the inanimate realm. Once again, the singularity of living systems was challenged; only this time around, it was not so much machines with which organisms were to be wed, but the kinds of systems that had long been familiar, if largely intractable, to physical scientists. Like thunderstorms, biological processes were likened to "far-from-equilibrium dissipative structures," or, more straightforwardly, to the stable solutions of nonlinear differential equations. And nineteenth-century notions of stability could be seen as finding a new home in the basins of attraction of chaos theory. I will argue, however, that the new paradigm of "self-organization" has thus far done little to satisfy the demands of Kant's original concept and, if for that reason alone, constitutes not a concluding chapter, but rather an episode, in an ongoing history.

ACKNOWLEDGMENTS

I thank Abigail Lustig for her sophisticated editorial help in preparing this version.