10 :: CYBERNETICS

N. KATHERINE HAYLES

Cybernetics flourished for about thirty years, from about 1940 to 1970, and then all but vanished from the academy as an identifiable discipline. At present there are only a handful of cybernetic departments in the United States and Europe (notably at UCLA and the University of Reading). Yet cybernetics did not disappear altogether; rather, it flowed over a broad alluvial plain of intellectual inquiry, at once everywhere and nowhere. In a sense it is more important than ever, although more for the inspiration it provides and the general framework it pioneered than for contributions as a discrete field of inquiry. For media studies, cybernetics remains a central orientation, represented by approaches that focus on information flows within and between humans and intelligent machines. The connection between media studies and cybernetics is prefigured by Gordon Pask's definition of cybernetics as the field concerned with information flows *in all media*, including biological, mechanical, and even cosmological systems. Now, as information is reimagined within the context of simulations and the computational universe, cybernetics is again surfacing as an area of active debate.

As is well known, the term *cybernetics*, adapted from the Greek word for "steersman," was coined by Norbert Wiener. Feedback mechanisms have been known since antiquity (one such device being Ktesibios's water clock) and are seen also in the eighteenth century (the governor on James Watt's steam engine) and the nineteenth and early twentieth centuries (homeostatic systems within animal physiology). However, cybernetics arguably did not come into existence as such until the traditional notion of the feedback loop was joined with the modern concept of information in the early twentieth century. It took shape as an interdisciplinary field of study during the mid-1930s to 1940s, coming of age in two key papers published in 1943, "Behavior, Purpose and Teleology" by Arturo Rosenblueth, Norbert Wiener, and Julian Bigelow, and "A Logical Calculus of the Ideas Immanent in Nervous Activity" by Warren McCulloch and Warren Pitts. From the beginning, cybernetics was conceived as a field that would create a framework encompassing both biological and mechanical systems; the ambition is clear in Norbert Wiener's 1948 book *Cybernetics, or Control and Communication in the Animal and the Machine*. During the famous Macy Conferences on Cybernetics, held from 1943 to 1952, much of the discussion revolved around the legitimacy of machine-animal-human comparisons. Analogies were drawn between the behavior of relatively simple mechanisms, such as Claude Shannon's electric "rat," William Grey Walter's electric "tortoise," Wiener's "Moth" and "Bedbug," and Ross Ashby's homeostat and that of much more complex animal and human systems.

The model of the neuron presented in McCulloch and Pitts's paper was centrally important in justifying such comparisons. Pitts proved a theorem demonstrating that the neuron model was capable of formulating any proposition that could be proved by a universal Turing machine. That animal and human neurons, acting singly and as a group in neural nets, were capable of performing computational acts was one of the strong justifications for considering both machines and biological organisms as cybernetic entities. Another lynchpin of early cybernetic theory was Claude Shannon's information theory, along with the similar theory developed by Norbert Wiener. Defining information as a function of message probabilities, the Shannon-Wiener approach detached information from context and consequently from meaning. The decontextualization of information was a crucial move in conceptualizing it as a disembodied flow that can move between different substrates and different kinds of embodiment. So powerful was Shannon's theory that it was quickly adopted in fields far removed from electrical engineering, including communication theory, linguistics, and psychology, leading eventually to Hans Moravec's claim that the human brain is nothing but an informational pattern that can be represented in any medium. Hence, Moravec argues, it will possible within the next few decades to upload the human brain into a computer without losing anything essential. Such a view carries into the contemporary period the cybernetic tendency to equate complex biological processes with relatively simple mechanistic ones, now envisioned as ushering in a postbiological era in which humans will shuffle off the mortal coil with which evolution has burdened them and become free to inhabit any kind of computational device they please, as long as it has sufficient storage and processing power.

All along, however, many researchers cautioned that the story is not so simple—that embodiment and context are crucially important factors that cannot be ignored. Donald MacKay, a British neurophysiologist, boldly argued for a theory that defined information relative to the changes it brought about in an embodied and historically situated receiver. The Shannon-Wiener model won out, however, because it produced information as a consistently quantifiable entity, whereas the MacKay model, precisely because it conveyed a richer, more complex sense of information, was intractable to exact quantification.

Indeed, the story of cybernetics as a whole can be seen as a struggle between the simplifications necessary to yield reliable quantitative results and more complex views that yield richer models but thwart robust quantification (see chapter 14, "Technology"). Elsewhere I have suggested that the history of cybernetics can be understood as evolving through three distinct phases; since then, it has become evident that we are well into a fourth phase, and the chronology needs to be updated accordingly. But I am getting ahead of my story; first let us consider the phases as they appeared in 1996, which I formulated as a three-part progression.

In the period from 1943 to 1960, homeostasis, disembodied information, and self-regulation were the central foci of research, and simple mechanisms stood in for much more complex biological organisms. This period is often called first-order cybernetics, because the organism/mechanism was theorized as an entity distinct from the environment in which it was embedded.

In the period from 1960 to 1985, the balance swung away from the simplifications of the first period to encompass the complexity introduced by considering the observer as part of the system. Reflexivity, understood as a shift of focus that brings the framing mechanism into view, was discussed in terms of bringing the observer *inside* the system, rather than assuming an external (and largely unnoticed) observer. This period, concerned both with cybernetic systems and with observers of cybernetic systems, is often referred to as second-order cybernetics. Its characteristic reflexive move is aptly signified by the punning title of Heinz von Foerster's book, *Observing Systems*, which can be taken to mean either that one is observing a system or that one is oneself a system that observes and that can be observed in turn by an observer who can also be observed, ad infinitum.

In autopoietic theory, a theory of the living pioneered by Humberto Maturana and Francisco Varela, the observer is taken into account in a negative way, by making a strong distinction between what an observer perceives and what the system can be understood as producing in its own terms. While an observer may posit causal links between events in the environment and an animal's behavior, autopoietic theory argues that within the living system as such, everything takes place in terms of the system's own organization, which always operates so as continually to produce and reproduce itself. Events outside the system can trigger events within, but no information passes from the environment to the autopoietic system. This premise allows information and meaning to be reconnected, but only as reflexive loops circulating within the system boundaries.

The third phase, as I suggested in 1996, can be understood as virtuality. From the beginning, as we have seen, cybernetics has been concerned with understanding how information, messages, and signals work within systemic boundaries, and with an orientation that implies humans, animals, and machines can all be understood within such a framework. Given this theoretical perspective, media are important primarily for their differential capacities to store, transmit, and process information. The cybernetic perspective implies that human and animal bodies, no less than cybernetic mechanisms, are media because they too have the capacity for storing, transmitting, and processing information. The construction of the body as an informational medium, implicit in the McCulloch-Pitts neuron, in the contemporary period takes the form of bodies tightly coupled with other media, especially computational media such as the Internet and Web. Thus the much-heralded virtuality of cyberspace merits the implicit connection with the history of cybernetics that the prefix *cyber-* invokes.

Today that kind of virtuality is apt to appear as a transition toward a social condition that has once again entangled the frame with the picture, although in a very different way from the second-order reflexive paradigm of autopoiesis. A decade or two ago there was much talk of virtual realms as "cyber" locations distinct from the real world, typified in the 1980s by the Polhemus helmet, which constrained the user within a tangle of wires defining the limits of physical movement. Nowdays that constraining frame is apt to be constituted not by a VR helmet but the Graphical User Interface (GUI) of microcomputers, and it is increasingly giving way to the pervasiveness, flexibility, and robustness of ubiquitous media. Instead of constructing virtual reality as a sphere separate from the real world, today's media have tended to move out of the box and overlay virtual information and functionalities onto physical locations and actual objects. Mobile phones, GPS technology, and RFID (radio frequency identification) tags, along with embedded sensors and actuators, have created environments in which physical and virtual realms merge in fluid and seamless ways. This fourth phase is characterized by an integration of virtuality and actuality that may appropriately be called mixed reality.

Bruce Sterling, in his book *Shaping Things*, proposes the neologism *spime* to denote this condition. Spimes are virtual/actual entities whose trajectories can be tracked through space and time, a capacity that RFID technology makes easily possible. As Sterling conceives the term, however, it implies more than the devices by themselves, connoting the transition from thinking of the object as the primary reality to perceiving it as data in computational environments, through which it is designed, accessed, managed, and recycled into other objects. The object is simply the hard copy output for these integrated processes. The spime is "a set of relationships first and always, and an object now and then" (Sterling 2005, 77); it is "not about the material object, but where it came from, where it is, how long it stays there, when it goes away, and what comes next" (109). Here the information flows *in all media* that Gordon Pask proposed as a definition of cybernetics comes to fruition as data streams processed in pervasive computational environments.

Complementing the move of computational devices out of the box, some researchers have suggested the possibility of a third-order cybernetics. Whereas first-order cybernetics was concerned with the flow of information in a system, and second-order cybernetics with interactions between the observer and the system, third-order cybernetics is concerned with how the observer is constructed within social and linguistic environments. Arnulf Hauan and Jon-Arild Johannessen, for example, see third-order cybernetics as interrogating the construction of the observer located within social networks; Vincent Kenny and Philip Boxer invoke Lacan along with Maturana to explore the linguistic construction of the observer in language communities. Other theorists have associated third-order cybernetics with complex adaptive systems, arguing that the autopoietic model of second-order cybernetics is not sufficiently attentive to the potential of complex systems to evolve and adapt in multiagent, multicausal environments.

From the beginning, the social, cultural, and theoretical impact of cybernetics has been associated with its tendency to reconfigure boundaries. First-order cybernetics, subverting the boundary separating biological organisms and machines, nevertheless implicitly drew a boundary around the system that left the observer outside the frame. Second-order cybernetics redrew the boundary to include the observer as well as the system (or, in the terms that Maturana and Varela develop, the autopoietic, informationally closed system plus the observer looking at the system). Third-order cybernetics redraws the boundary once again to locate both the observer and the system within complex, networked, adaptive, and coevolving environments through which information and data are pervasively flowing, a move catalyzed by the rapid development of ubiquitous technologies and mixed reality systems.

An equally strong tendency within the cybernetic tradition has been the impulse to construct a framework within which animals, humans, and machines can all be located. With the growing importance and increasing power of computational media, this framework has tended to be seen not only as a flow of information but specifically as computational processes. The ultimate boundary-enlarging move is the claim put forward by some scientists, including Stephen Wolfram and Edward Fredkin, that the universe is a giant computer, ceaselessly generating physical reality by means of computational processes that it both embodies and is. The claim has important implications for the nature of reality. In Fredkin's view, for example, it implies that space and time, and indeed all of physical reality, are discrete rather than continuous. The claim also shifts the focus from information flows to computational processes, and this opens the way for a reconceptualization of information.

In a recent presentation, Fredkin suggested that "the meaning of information is given by the processes that interpret it." Although Fredkin did not develop the idea (aside from giving the example of an MP3 player, which interprets digital files to create music), the formulation is solidly within the cybernetic tradition in that it crafts a framework within which the behavior of organisms and machines can be understood in similar terms. At the same time, the formulation goes significantly beyond first- and second-order cybernetics in giving a more enactive and embodied sense of information. Specifically, it breaks new ground by changing the meaning of "interpretation" and of "meaning." Information in this view is inherently processual and contextual, with the context specified by the mechanisms of interpretation. These processes take place not only within consciousness but within subcognitive and noncognitive contexts, both biological and mechanical. A computer, for example, gives information one kind of meaning when voltages are correlated with binary code. Another kind of meaning, much easier for humans to understand than strings of ones and zeros, emerges with high-level programs such as C++, and still another kind when C++ commands are used to generate even more accessible screen displays and behaviors. Human cognition, for its part, arises from contexts that include sensory processing, which interprets information from the environment and gives it meaning within this context. The meaning that emerges from these processes undergoes further interpretation and transformation when it reaches the central nervous system; these meanings are transformed yet again as the CNS interacts with the neocortex, resulting in conscious thoughts.

Donald MacKay, cited earlier as a pioneer in formulating an embodied theory of information, had already envisioned a series of hierarchical and interrelated contexts that included subcognitive processes, anticipating in this respect Fredkin's formulation. MacKay insisted, for example, that the meaning of a message "can be fully represented only in terms of the full basic-symbol complex defined by all the elementary responses evoked. These may include visceral responses and hormonal secretions and what have you" (1969, 42). Fredkin adds to this vision a way of understanding meaning that extends it to mechanical nonhuman processes. Indexed to local subcognitive and noncognitive contexts, "interpretation" ceases to be solely a high-level process that occurs only in consciousness. Rather, it becomes a multilayered distributed activity in which the "aboutness" of intentionality (traditionally used by philosophers as the touchstone of cognition) consists of establishing a relation between some form of input and a transformed output through context-specific local processes. By breaking the overall context of reception into many local contexts, Fredkin's formulation makes the processes at least partially amenable to reliable quantification. Many of these local contexts already have metrics that work: voltages, processing speeds, and bits per second in computers; neural responses, fatigue rates, and the like in humans. The important point is a shift of vision that enables us to see these subcognitive and noncognitive processes not just as contributing to conscious thought but as *themselves* acts of interpretation and meaning.

Fredkin's proposal provides a convenient instance to explore what cybernetics has to contribute to media studies (and vice versa). Although it may seem strange in a book devoted to critical terms for media studies, I want to ask what "media" are. Although references to media are ubiquitous, when I searched a couple of years ago for definitions, I was astonished at the dearth of proposals. It seems everyone knows what media are, but few specify what the term denotes. Ideally I want a definition less broad than McLuhan's ("any extension of ourselves," which includes roads, for example, as well as hammers and language) and not as narrow as most offered by dictionaries (typically along the lines of "anything that stores, transmits, or processes data"). Following a scheme developed by Annagret Heitman, I suggest that media can be understood through four principal levels of analysis, which function, in effect, as a definition through specification: materiality, technology, semiotics, and social contexts. The materiality level specifies what a medium is in its physical composition; for a desktop computer, this would include silicon chips, circuits, display screen, and the like. The technology level specifies how the material object works and functions; in the case of the computer, this includes analyses of hardware and software, their interactions and interconnections. The semiotic level addresses the basic function of media, which I take to be the facilitation of communication; for the computer, this includes binary code, scripting languages, compiled and interpreted languages, and Internet protocols. The social context includes not only the way people use computers but also the corporations that produce hardware and software, the market mechanisms through which they are disseminated into a population, repair and maintenance industries, and so on.

Clearly the computer qualifies as a medium, for even as a stand-alone device with no network capabilities, it facilitates communication for an individual user; moreover, it is never simply a solitary device, for it depends on dense, distributed, and tightly integrated networks of programmers who write the code, software engineers who design the machine, organizations that specify standards for these operations, and many more social, cultural, and technical processes. Add to this network capabilities and the convergence of media into integrated digital devices, and one could plausibly argue that networked and programmable machines (or, as John Cayley prefers to call them, "programmatons") are the most important and pervasive media of the contemporary period.

Returning to the question of cybernetics in relation to media studies, I will focus on the three general contributions: the idea of the feedback loop joined with a quantitative definition of information; the idea of a theoretical framework capable of analyzing control and communication in animals, humans, and machines; and artifacts instantiating these ideas, from the McCulloch-Pitts neuron model to virtual reality platforms such as a CAVE and, more recently, the Internet and Web. The material and technological levels of computational media correspond to cybernetic artifacts, while the semiotic level is routinely understood as information flows, and the social contexts involve myriad feedback loops between humans and computers that continue to reconfigure social, economic, and technological conditions for people throughout the world. In this sense, then, not only do computational media continue the cybernetic tradition; arguably, computational media are the *principal* arenas in which cybernetics and media co-construct each other.

In a move reminiscent of the recursive feedback loops important to cybernetics, we can now cycle back to consider how the computational universe both produces and is produced by cybernetic dynamics. As I have argued elsewhere, the idea of the computational universe is catalyzed by the power and pervasiveness of contemporary programmatons; it is precisely because networked and programmable media have been able to do so much that the analogy between their operations and the natural world becomes compelling. Moreover, the feedback loop circles back into contemporary technologies of computation to authorize computation as the language of nature, displacing the traditional claim of mathematical equations to this role. As Harold Morowitz, among others, has argued, the twenty-first century is witnessing a move away from the study of relatively simple systems that can be modeled by equations with explicit solutions to complex adaptive systems that resist mathematical formalization and can be studied only through computational simulations. What calculus was to the eighteenth and nineteenth centuries, Morowitz argues, simulations are to the twentieth and twenty-first centuries. Feedback loops recursively cycle between and among humans and computers, between autonomous agents in simulations and the environments in which they operate, and between contemporary technologies and theories about the Universal Computer. These feedback loops are simultaneously topics central to media studies and the dynamics through which the most important and central theories of media studies have been constituted.

There is fitting irony in the fact that just as cybernetics disappeared as a discrete field only to reappear in many different areas, so computation, expanded to cosmic scale in the Universal Computer, disappears into the fabric of the universe. Can something still be called a medium when it is the foundation for physical reality, continuously generating the world and everything we know through its computations? Here we come upon another legacy from cybernetics that harbors a subtle paradox, if not a contradiction: artifactual systems, because they are simple enough to be understood in their entirety, reinforce the interpretation of much more complex biological entities as cybernetic systems, while at the same time subverting the distinction that separates the artificial from the natural. If the natural world is itself a giant computer, are not computers the most "natural" objects imaginable? As Bruno Latour has shown, nature/culture hybrids are far from unusual. Cybernetics, in its ambition to create frameworks that apply equally to machines and to bodies, has been one of the forces driving twentieth and twenty-first century thought to interpret the mind/body in computational terms and to think about computers as cognizers capable of evolving in ways that parallel the emergence of humans as thinking beings.

Since computers are also media, considering the body as a medium is a thoroughly cybernetic move. Traditionally, "media" have been understood primarily as artifacts and artifactual systems, but as we have seen, the cybernetic impulse is to erode this distinction. Moreover, the boundary work that would put bodies on one side of a nature/culture divide and media on the other has already been rendered unstable by technologies that penetrate the body's interior and make it available as images through such media devices as ultrasound, MRI, and CT and PET scans. The highly publicized endeavors of the online Visible Human Project and Gunther von Hagen's *Korperwelten* ("Body Worlds") exhibitions have further disseminated the sense of bodies as media into the cultural imaginary. Bernadette Wegenstein's proposal to "reconfigure the discipline of body criticism into one of new media criticism" is a predictable strategy. She argues that "the medium and questions around mediation have literally taken over the space and place of the individual body . . . [and] the body . . . has emerged in place of . . . the very mediation that once represented it for us. The medium, in other words, has become the body" (2006, 121). I differ from her assessment only in cautioning that the feedback loops connecting bodies and computational media should not be allowed to disappear from sight by too easily collapsing bodies and media together, for they form a crucial constellation that reveals yet another way in which the cybernetic project continues to inform contemporary thought.

Understood in evolutionary terms, these recursive feedback loops between culture and computation create a coevolutionary dynamics in which computational media and humans mutually modify, influence, and help to constitute one another in the phenomenon known as technogenesis. What these coevolutionary cycles imply, and where they might be headed, are matters of intense debate. At the extreme end of the spectrum are transhumanist advocates such as Hans Moravec and Ray Kurzweil, who predict that within a few decades, human consciousness will be uploaded into computers, achieving effective immortality. More moderate views, articulated by Rodney Brooks among others, caution that the process will take much longer and may not be possible at all, considering the enormous gap that separates the complexity of the human organism from even the most sophisticated computer. Still others, notably Francis Fukuyama, want the boundary between bodies and cybernetic technologies to be vigorously policed to ensure that human beings will not lose their biological and evolutionary heritages. Mark Hansen stakes out yet another position in acknowledging the importance of technogenesis but arguing that the embodied observer must remain at the center of our understanding of our interactions with digital media.

My own view is that the predictions of Moravec and others that the mid-twenty-first century will witness a "singularity"—a point of transformation so dramatic that the nature of human being will be forever changed—downplays the enormous differences between biological organisms and computers and dehistoricizes what has been a very long process that was already underway when *Homo sapiens* became a distinct species. The emergence of bipedalism, the advent of speech, the escalating complexity and speed of tool invention and use have all irrevocably changed the biology, culture, and cognitions of humans. At the other end of the spectrum, Fukuyama essentializes and indeed fetishizes "human nature" in an equally unhistorical and untenable way. Technogenesis is one of the major forces shaping humans now, as it has been for eons. Important changes have already taken place within the twentieth century, and we can expect the trend to continue with increasing speed and momentum in the twenty-first (assuming, of course, that environmental disasters and cataclysmic events such as nuclear war do not qualitatively change the nature of the game). On my view, responsible theorizing about these changes requires close attention to the materiality of bodies and computational media, a clear understanding of the recursive feedback loops cycling between them, and contextualizations of bodies and machines that reveal how meaning is created through the cascading processes that interpret information.

If we return now to the three major contributions of cybernetics—joining information with feedback, creating a framework in which humans and machines can be understood in similar terms, and creating artifacts that make these ideas materially tangible—we may conclude that contemporary media studies would scarcely be conceivable without the contributions of cybernetics. Friedrich Kittler's aphorism that "media determine our situation," the media archaeology he pioneered and that has been further developed by Bernhard Siegert and others, the technogenesis explored by Mark Hansen and Bernard Stiegler, the hypothesis of the Universal Computer advocated by Stephan Wofram and Edward Fredkin, and the contributions of innumerable others are indebted in both obvious and subtle ways to the ideas that found pioneering expression in the four waves of cybernetics. If cybernetics seems today to have been reduced to a prefix that echoes pervasively across the culture, that is because its essential concepts have been absorbed deeply into the fabric of contemporary thought. Remembering the history of cybernetics reminds us not only of the debt we owe to it but also the necessity to keep its premises in sight so that, in the spirit of the remarkable innovators who launched the cybernetic movement, we can continue to interrogate our assumptions as we search for better ways to envision and create a life-enhancing, life-affirming future.

References and Suggested Readings

Dupuy, Jean-Pierre. 2009. *On the Origins of Cognitive Science: The Mechanization of Mind*. Cambridge, MA: MIT Press.

Haraway, Donna J. 1990. "A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century." In *Simians, Cyborgs, and Women: The Reinvention of Nature*, 149–82. New York: Routledge.

Hayles, N. Katherine. 1999. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: University of Chicago Press.

Heims, Steve J. 1991. *The Cybernetics Group*. Cambridge, MA: MIT Press.

MacKay, Donald. 1969. *Information, Mechanism, and Meaning*. Cambridge, MA: MIT Press.

Maturana, Humberto R., and Francisco J. Varela. 1991. *Autopoiesis and Cognition: The Realization of the Living.* New York: Springer.

Shannon, Claude E., and Warren Weaver. [1948] 1998. *The Mathematical Theory of Communication*. Champaign: University of Illinois Press.

Sterling, Bruce. 2005. *Shaping Things*. Cambridge, MA: MIT Press.

Varela, Francisco J., Evan Thompson, and Eleanor Rosche. 1992. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

Von Foerester, Heinz. 2002. *Understanding Understanding: Essays on Cybernetics and Cognition*. New York: Springer.

Wegenstein, Bernadette. 2006. *Getting under the Skin: Body and Media Theory*. Cambridge, MA: MIT Press.

Wiener, Norbert. 1965. *Cybernetics, or Control and Communication in the Animal and the Machine*. 2nd ed. Cambridge, MA: MIT Press.

Wolfram, Stephen. 2002. *A New Kind of Science*. New York: Wolfram Media.