

Autopoiesis and Computation

Stephen Ramsay

Willard McCarty, in an essay intended to describe fairly and succinctly the project of humanities computing, proposes that we “ask in the context of computing what can (and must) be known of our artefacts, how we know what we know about them and how new knowledge is made” (McCarty 1). In one sense, this is the same question that runs through the whole of hermeneutical philosophy from Schleiermacher to Heidegger and beyond. But when such questions are put forth “in the context of computing,” they take on a newness—and a new urgency—borne of methodological uncertainty and doubt. We who bring technology to bear on humanistic problems would like to know “what can be known,” not for its own sake, but because computers seem to require a model or pattern of that knowing. The machine is waiting to be told what we know and how we know it, and while we may skip ahead to epistemology in order that methodology might reveal itself more clearly, methodology presses upon us with a new insistence.

This is something like the problem that confronted computer science when artificial intelligence went from being thought experiment to being actual experiment in the latter half of the twentieth century. Since a computer is a mechanistic entity, biological models of understanding and

interpretation seemed unavoidable, and so AI sought to encode something like our methodology—something like our own process of understanding and interpreting—into another physical instantiation. Yet it did so with an acute awareness that we are woefully ignorant of the process by which we ourselves negotiate the extraordinarily complex set of transactions that make up the thinking, interpreting self. The result was a land shift in our approach to epistemology. For the first time, we began to consider meaning, knowing, understanding, and interpretation, not as problems in philosophy, but as problems in biology, cognitive psychology, and systems theory.

But when we turn to biological explanations, we are immediately confronted with a deep conundrum. The physical properties of the brain—the properties and behavior of axons and dendrites—bear no obvious relationship to the broader act of understanding. The American engineer and cognitive scientist Heinz von Foerster, one of the original cyberneticists and an attendant at the Macy conferences during the 1940s, put it thus:

The response of a nerve cell does *not* encode the physical nature of the agents that caused its response. Encoded is only “how much” at this point on my body, but not “what.”

[...]

[F]or indeed “out there” there is no light and no color, there are only electro-mechanical waves; “out there” there is no sound and no music, there are only periodic variations of the air pressure; “out there” there is no heat and no cold, there are only moving molecules with more or less kinetic energy, and so on. Finally, for sure, “out there” there is no pain.

[...]

Since the physical nature of the stimulus—its *quality*—is not encoded into nervous activity, the fundamental question arises as to how does our brain conjure up the tremendous variety

of this colorful world as we experience it any moment while awake, and sometimes in dreams while asleep. (Foerster 293–4)

It is a paradox at least as old as Leibniz's *Monadology*. The baffling implications of a brain so unlike the perceptions with which it is apparently involved provide both the wellspring of Cartesian dualism and the essential research project of contemporary neuroscience. Yet it also complicates the question McCarty poses.

McCarty's question takes on a different tenor when we consider it under the peculiar conditions of the Leibnizian Gap (as the disconnect between our brains and the world is sometimes called). In the case of literary and artistic works, we end up with a set of propositions that are ironically conformable to many of the insights of post-structuralist thought. There is no meaning to the poem "out there"—only the graphic orderings of luminous similitude and contrast impressing themselves upon our nervous system. There is no interpretation of the stageplay "out there," but only sound waves and kinetic energy interacting with the physical interface of what we obliquely refer to as the self.

I believe the most radical answer ever proposed to the how and why of this disconnect may be found in the notion of *autopoiesis* as put forth by the Chilean neuroscientists Humberto Maturana and Francisco Varela. Autopoiesis proceeds from the premise that the Leibnizian Gap is a permanent fact of biological instantiation. In observing such systems, we find that we simply cannot account for them by enumerating their parts. We can speak about the "transfer functions" of cells, of the role that serotonin plays in neurotransmission, and of the various ways in which axons and dendrites communicate across charged synapses in living systems, but none of this seems to bring us any closer to understanding what it means to think, to know, and to interpret (Maturana 5-6). We can discern rules and patterns (even such suggestive ones as generative grammar), but as Maturana and Varela remind us again and again, "these are regularities

... that lie exclusively in the realm of the observer. ... [T]ransitions from state to state as internal processes in any system are unrelated to the nature of the interactions to which they give rise” (Maturana 34). We can construct models of mind (as well as hermeneutical recommendations) that depend—fundamentally—on memory, and yet, “. . . memory as a storage of representations of the environment to be used on different occasions in real life *does not exist* as a neurophysiological function” (Maturana 37, emphasis mine). In the end, Maturana conclude (or rather, begin) by putting forth the idea that the study of living systems cannot be “the study of a mapping of a colorful world on the nervous system, but rather that it [has] to be the understanding of the participation of the [nervous system] in the generation of the color space of the observer” (Maturana xii).

One may discern in this weird reversal the seeds of the common post-structuralist notion that the reader creates the text. But here we are casting a much more extreme form of skepticism on a much more fundamental series of epistemological claims: that, for example, human beings process the world, that their behaviors are influenced by their environment, and that objective descriptions are descriptions of what is known about the world. Where Kant had consigned identity to the realm of the transcendental, Maturana and Varela propose that biological organisms (including human interpreters) are characterized by the autonomic maintenance of identity. Again and again Maturana and Varela define living systems as recursive generators of themselves. Maturana:

The circular organization [of living systems] constitutes a homeostatic system whose function is to produce and maintain this very same circular organization by determining that the *components* that specify it be those whose synthesis or maintenance it secures. (Maturana 9)

Varela:

An autopoietic machine is a machine organized . . . as a network of processes of production (transformation and destruction) of components that produces the components which: (i) through their interactions and transformation continuously regenerate and realize the network of processes (relations) that produced them. (Maturana 79)

We are not determined by our interactions. The circular and autonomous nature of living systems determines the interactions into which they can enter, and they maintain their identity only as long as they maintain the basic circularity of that organization.

The notion of a fully autonomic existence—of entities that do not respond to the environment so much as define it through the maintenance of their own solipsistic identity—would seem to render linguistic activity problematic, if not impossible. Linguistic communication manifestly occurs among autonomous, autopoietic systems, and yet it manifestly cannot be the transmission of information, “as if something were transmitted from organism to organism, in a manner such that the domain of uncertainties of the ‘receiver’ were reduced according to the specifications of the ‘sender.’ ” The listener, acting in a fully autonomous fashion, makes a choice. Maturana:

The choice is caused by the “message,” but the orientation thus produced is independent of what the message represents for the orienter. In a strict sense then, there is no transfer of thought from the speaker to the interlocutor; the listener creates information by reducing his uncertainty through his interactions in his cognitive domain. Consensus arises only through cooperative interactions in which the resulting behavior of each organism becomes subservient to the maintenance of both (Maturana 32)¹

¹The persistence of the “conduit metaphor” of information transmission in human lin-

Let us assume that the neuroscientists are right: the quality of the phenomenal world is not encoded in the nervous system. It remains to say whether Maturana and Varela have used this insight successfully to characterize living systems as something distinct from non-living ones. Certainly, most nonliving systems are not autopoietic (though Maturana and Varela have left open and are at odds with one another concerning the possibility that social organizations, economies, and some machines are autopoietic, and therefore may be coherently described as living systems). What does this mean for hermeneutics in general, and humanities computing in particular?

It means, first of all, that interpretation consists not merely of ideations put forth by the interpreter, but of a series (we should say, a process) of alternative stimuli that enable those ideations to come forth. Such alternative formations stand toward the interpreting subject as the rest of the phenomenal world does; that is, as further means of reorientation. Interpretation is, as Von Foerster said of perception, much more a matter of correlation than confirmation—the reality we perceive is, in purely biological terms, “a” rather than “the” reality (Foerster 295). We do not create the physical text (the collection of stimuli), but neither do we interpret a fixed entity. Instead, we recursively generate descriptions (and sometimes descriptions of ourselves describing ourselves, as is the case here) that exist solely within our own cognitive space as observers. We do this because, as von Foerster puts it, “The nervous system is organized (or organizes itself) so that it computes a stable reality” (Foerster 306).

Computationally assisted literary criticism—what I have elsewhere called algorithmic criticism²—may be conceived as the attempt to insert additional computations into the recursivity of cognition. By subjecting texts to algorithmic transformations, we create alternative arrangements that

guistic behavior is well documented in Joe Grady’s “The Conduit Metaphor Revisited: A Reassessment of Metaphors for Communication” (1998). See Grady, *op cit.*

²“Toward an Algorithmic Criticism,” *Literary and Linguistic Computing*, 2003, *op cit.*

necessarily alter our sense of the object. We do so not as a way of apprehending the reality that the work puts forth, but as a way of adding to the possible realities we may describe in light of the work. The possibility of mere solipsism is avoided, because while the radical constructivism that governs the subject presupposes autonomy, these constructions are embedded within a social network which, like the thinking self, orders itself toward the computation of a stable reality. The outputs of algorithmic transformations and interpretations of them do not stand apart from the larger linguistic order in which acts are evaluated and truth claims put forth. Such are the necessary and sufficient conditions that allow this stable, collective reality to emerge.

The computer is ideally suited to the task of dilating possible worlds—not because its insistent logic of truth and falsehood cracks through the shell of truth and falsehood, but because that insistence is so alien to the autopoietic negotiation that governs our own perception. We adapt our environment in order to maintain our identity; our cognitive biology defines in advance the regime of our interpretive interactions. The computer disrupts this order with the unexpected pattern—the previously unknowable (because unviewable) pattern. The computer orients us to the cognitive domain of an observer wholly unlike ourselves.

This is perhaps what McCarty refers to when he notes that:

Research in humanities computing begins ...in the breakdown, when tools become models... By definition a model cannot be true; it is, as Max Black says, a heuristic fiction ([70]). Hence, although better knowledge of the modeled object and of the analytical method results, modeling is essentially a quest for meaningful failure. [...] A computational model ...does this in a rigorous and particular way by demanding absolute consistency and complete explicitness. It also raises the question in an environment of indefinite flexibility, inviting indefinite heuristic play. (McCarty)

And:

Since the evolution of computing seems increasingly to facilitate interactive play, a place is needed for it in an epistemology of computing: an apparently blind (and therefore black-box) “feeling toward” new knowledge rather than a targeting of it, in which what matters most is, as Hacking says, not observation but being observant.

What is needed, then, is an ethic of play, an aesthetic of computation, and a hermeneutics that takes the Leibnizian Gap as the basis of its underlying epistemology—not because we cannot behave without such a theory, or because we are incapable of making hermeneutical recommendations in the absence of such a theory, but because such a theory might allow us to speak of humanistic engagement (and in particular, its more playful aspects) as something other than the hermeneutics of “anything goes.” Von Foerster suggests the following imperatives, and they are a good place to start:

The Ethical Imperative: Act always so as to increase the number of choices.

The Aesthetic Imperative: If you desire to see, learn how to act. (Foerster 308)

Works Cited

- Foerster, Heinz von. *Observing Systems*. Seaside, CA: Intersystems, 1984.
- Grady, Joe. “Discourse and Cognition: Bridging the Gap.” (1998).

- Maturana, Humberto R. and Francisco J. Varela. *Autopoiesis and Cognition: The Realization of the Living*. Boston Studies in the Philosophy of Science 42. Dordrecht: Reidel, 1972.
- McCarty, Willard. "Humanities Computing." *Encyclopedia of Library and Information Science*. New York: Dekker. 2003.
- Ramsay, Stephen. "Toward an Algorithmic Criticism." *Literary and Linguistic Computing* 18 (2003): 167–74.