From Chapter Three: Emerson's Environments

Autopoietic Systems

Humberto Maturana and Francisco Varela begin their essay, "Autopoiesis: The Organization of the Living" by noting, "A universe comes into being when a space is severed[1] into two. A unity is defined. The description, invention and manipulation of unities is at the base of all scientific inquiry" (Autopoiesis 73). The "cut" creates two entities, and separating out one instantly identifies the other. In making such distinctions, scientific inquiry--any inquiry, really--finds purchase on the world. Emerson might have agreed. He says in the introduction to Nature, "All science has one aim, namely, to find a theory of nature," and goes on to claim that "Philosophically considered, the universe is composed of Nature and the Soul" (CW 1:8). Emerson makes his opening move by finding what he perceives as a crack in the smooth eggshell of the world, which he uses to establish two identities: his own soul/mind/self, and everything else--"Nature" or, as he will also say, the "NOT ME" (what is often called elsewhere the mind/body or the subject/object dualism). While Emerson's goal in the rest of Nature is to repair this crack (and he fails, of course, to do any more than displace it), his founding gesture replicates the strategy which Maturana and Varela view as critical for autonomous, self-organizing systems: establishing an impermeable, organizational boundary between the system and its environment.

Like many other theorists in the life and cognitive sciences, Maturana and Varela wondered how complex biological systems can maintain their coherence in the face of even more complex environments, retaining their formal integrity even as their internal components adapt and change over time. How, too, they asked, can such systems develop greater complexity than their constituent components and structures taken separately might suggest? In other words, as the question was once framed, why is the sum of an organism so much more than its parts? The term "autopoiesis" was coined by Maturana to draw attention to what he decided were the defining features of living systems: their circular, self-referential organization, or autonomy; and their continual creation (poiesis) of the components which structure and sustain that autonomy (Autopoiesis xvii). Maturana and Varela imagine a living system as an "autopoietic machine," organized "as a network of processes of production, transformation and destruction of components that produces the components which: (i) through their interactions and transformations regenerate and realize the network or processes (relations) that produced them; and (ii) constitute it

as a concrete unity in the space in which they exist by specifying the topological domain of its realization as such a network" (135). Autopoietic systems thus maintain their organization (a defining relationship among the various components and processes of a unity) even as their structure (the actual assemblage of components and processes that make up the unity) may change as a result of perturbations in the environment. The example of a toilet can demonstrate the difference between organization and structure: the toilet mechanism is organized as "an apparatus capable of detecting the water level and another apparatus capable of stopping the inflow of water. The toilet unit embodies a mixed system of plastic and metal comprising a float and a by-pass valve. This specific structure, however, could be modified by replacing the plastic with wood, without changing the fact that there would still be a toilet organization" (Tree 47).[1]

To say that a "toilet" is organized on the basis of its water-regulating capability may not strike all of us as obvious. We tend to think that a toilet is organized on the basis of its function in human waste disposal, which is certainly its primary function both as it is used and as it was conceived, and as far as the toilet bears on these human goals, the water-regulating function appears as a necessary but certainly not sufficient feature in the toilet's definition. However, as Maturana and Varela explain, that we can create a "purpose" for an autopoietic machine, or simply place it within a larger context of organization, are notions "intrinsic to the domain of observation, and cannot be used to characterize any particular type of machine organization" (Autopoiesis 78). It is thus a matter of convenience, of nomenclature, to ascribe any purpose at all to an autopoietic entity. Varela puts it this way:

---

[1] As Katherine Hayles notes, the example of the toilet seems somewhat inapt, since in The Tree of Knowledge Maturana and Varela are less interested in machines than they are living systems, and for living organisms the key point is that material (i.e., protein) change is less important than "the way the material is organized" ("Making the Cut" 47). As well--and I think this point may be even more germane—toilet organization seems to smack of an allopoietic system rather than an autopoietic one. Allopoietic systems can be compared to the cybernetic system described by Bateson, in which homeostasis is the goal, but homeostasis based on balancing inputs and inputs with the environment. The toilet is not immediately identifiable as consisting only in its water regulating structures; there are other systems involved as well: connecting plumbing, flush handle, septic system. However, as Maturana and Varela suggest, even autopoietic systems can be part of larger autopoietic systems (e.g., cell organelles within cells, nervous systems within the body). Thus, the toilet example could be an autopoietic subsystem of a larger sewage system, and their point still valid.

> It is important to realize that we are not using the term organization in the definition of an autopoietic machine in a transcendental sense, pretending that it has an explanatory value of its own. We are using it only to refer to the specific relations that define an autopoietic system. Thus, autopoietic organization simply means processes concatenated in a specific form: a form that constitutes and specifies the system as a unity. It is for this reason that we can say that if any time this organization is actually realized as a concrete system in a given space, then the domain of the deformations that the system can withstand without loss of identity (that is, maintain its organization) is the domain of changes in which it exists as a unity. (Principles 13)

If this passage sounds like a tautology, it is: a system is organized as it is organized, and if its organization changes then it is no longer organized in the way that it was originally. But we must remember that "autopoiesis" is conceived precisely to provide a definition of system based only on its organizational unity and the reproduction of that unity. Recalling the founding role of "making a distinction," we can say that the system is what it is, and is not what it is not: there is the toilet itself, and there is the rest of the world. The reference of the system only to its own immanent self-regulating, self-organizing autonomous structural processes is what simultaneously allows it to distinguish itself from its environment and constitute its unity. If a system can continue to meet the demands of its own organization, despite structural changes, then it maintains its identity (the toilet stays a toilet); if changes in structure force a change in organization, the system loses its identity, either by disintegrating, or by creating a new identity with a new organization (the toilet becomes, say, a water pump). When an observer ascribes purpose or use to an autopoietic machine he places it outside of its own autopoietic space, within which the only "purpose" is always simply to preserve autopoiesis. A corollary is that self-conscious systems (i.e., humans) may observe their own behavior and place it in contexts the consciousness-system does not properly have at the fundamental level of its organization. We stand outside ourselves and observe our actions and behaviors; only rarely do we attribute them to the simple fact of pursuing the goalless "goal" of preserving organizational coherence. In general, according to the theory of autopoiesis, though we consciously grant ourselves and others all sorts of motivations and purposes, at the level of our autopoietic functions there is no teleonomy at all. Marjorie Levinson says, "this amounts to a question of representation as survival," such that "The ghost in the rationalist and materialist machine is shown to be

nothing but the survival of a form of organization in the real world as perceived by another system in that world, a system perforce specified by that organization and thus included in it" (122).

{....}

If an autopoietic system works exclusively to reproduce its autopoiesis, how does it interact with its environment? Would it not be blind to that environment, and likely to stumble inevitably in self-defeating directions, perhaps to the point of destroying its environment? In fact, this is always possible, so that just as with the evolution of species, only the successfully adapted systems persist.[2] By definition, an environment is always more complex than the system; the system's challenge is to continuously respond to those environmental factors that have a bearing on it--and screen out the rest. It does this through structural coupling, which Maturana and Varela define as "the continued interactions of a structurally plastic system in an environment with recurrent perturbations [which] will produce a continual selection of the system's structure" (Principles 33). In effect, the environment may trigger changes in the system because the system is structurally open, but because the

---

[2] The contrast between an autopoietic view of systems and other paradigmatic ecological theories is well illustrated by comparing the work of Maturana and Varela to Edward Goldsmith's The Way: An World View. In this ecosophical text, the former editor of The Ecologist suggests that "For natural systems to achieve their goal of maintaining their own and Gaian stability, they must be able to predict environmental changes to which they must adapt, as well as the environmental effects of such adaptations. There is every reason to suppose that they are well capable of doing so, providing such changes occur within their tolerance ranges" (153). Furthermore, "A living thing apprehends its environment by detecting data that appear relevant to its behavior pattern and interpreting them in light of its mental model of its relationship with it. This means that it seeks to establish their meaning and thereby to understand them" (158). Under autopoiesis there is an element of predictability to systems' behavior, in that recurrent "successes" of a system in its environment may give it the appearance of correct adaptation to environmental factors. But this appearance is illusory. Autopoietic systems cannot "predict" or "represent" internally their environment, and they do not preview the effects on the environment of their own processes. Instead, they attempt to maintain their autonomy by modifying their internal structures to conserve their organization, regardless of what is happening in the environment: what looks to an observer like the system's responsiveness to environment is simply the system's specification of its own internal state. Goldsmith posits an open system, and as with the Batesonian cybernetic loop, optimizing the inputs and outputs between organism and environment leads to the desired homeostasis (even to the point of a global, "Gaian" balance). With Maturana and Varela, however, it is not a question of inputs and outputs, but of closure--and whether or not the system can maintain it in the face of environmental irritation.

system is also operationally closed what forms those changes take are determined solely by the system itself as it strives to maintain its organization. Only the system can determine its structural changes, and only if further autopoiesis is secured can the changes to environmental pressures be considered appropriate.[3] The environment, in other words, can initiate changes but cannot determine how they play out in the system. For example, a nervous system will react to various sensory stimuli (touch, image, etc.) by reconjugating the order and arrangement of neural firings, which in turn may lead to particular motor responses. But considered at the level of its organization, the nervous system is simply attempting to maintain the already existing organization of its components which have been perturbed by the environmental stimuli. Richard Lewontin, a geneticist whose work has been informed by dialectical materialism, confirms that "organisms actually change the basic physical nature of signals that come to them from the external world. As the temperature in a room rises, my liver detects that change, not as a rise in temperature, but as a change in the concentration of sugar in my blood and the concentration of certain hormones" (90). In a passage that should resonate well with the readers of Maturana and Varela, Lewontin says "The last rule of the relation between organism and environment is that the very physical nature of the environment as it is relevant to organisms is determined by the organisms themselves" (91).

Perhaps what is most interesting here is that the operational closure of the nervous system suggests some far-reaching epistemological consequences for systems in general and cognitive ones in particular. When we define the nervous system as closed to its environment, we seem to be saying that it is essentially solipsistic, in that its operations are self-referential and directed solely towards the maintenance of specific relations between internal structures. Yet at the same time, autopoietic systems do have the ability to structurally couple with environmental elements, and thus in some sense must be able to account for the environment internally. To put a finer point on the ramifications of their work, Maturana and Varela claim that the brain

> is not solipsistic, because . . .it participates in the interactions of the nervous system in its environment. These interactions continuously

[3] Structural changes can be induced by the environment (e.g., an injury or a disease), but they are of a different order and clearly have nothing to do with the system's normal structural adjustments based on autopoietic factors.

> trigger in it the structural changes that modulate its dynamics of states. . . This is so despite the fact that, for the operation of the nervous system, there is no inside or outside, but only the maintenance of correlations that continuously change. . . . Nor is it representational, for in each interaction it is the nervous system's structural state that specifies what perturbations are possible and hat changes trigger them. . .it brings forth a world by specifying what patterns of the environment are perturbations and what changes trigger them in the organism. (Tree 169)

Ultimately, "Anything said is said by an observer" (Autopoiesis 8), a crucial point made over and over again by Maturana and Varela to emphasize that communications are always made from within an observational domain and so are distinct from system-specific operations. We can dispense with debates about representationalism and solipsism if we simply remember that while systems do adjust themselves as they attempt to maintain their organization in a complex environment, it is observers who attribute those adjustments to specific system/environment interactions (i.e., behavior).

We have arrived now at the point of the traditional clash between realism and idealism: is the real world grasped by the nervous system according to some sort of empirical mapping procedure, or is the real word produced by a priori categories of perception hard-wired into the brain? In what is Maturana and Varela's crucial insight, this question can be set aside when we realize that the operational closure/structural openness of systems under the autopoietic paradigm allows us "to maintain a clear logical accounting" (Tree 135) between the domains under consideration. The authors offer the analogy of a person who has (improbably) lived all his life in a submarine and been trained to keep dials and gauges all within certain parameters by correctly manipulating levers and knobs to compensate for fluctuations. This "pilot" is a metaphor for the autopoietic process itself. When the submarine arrives near a shore, an observer congratulates him for his ability to handle the craft, for successfully negotiating the reefs and surfacing the submarine at its destination. The pilot replies, "'What's this about reefs and surfacing? All I did was push some levers and turn knobs and make certain relationships between indicators as I operated the levers and knobs. It was all done in a prescribed sequence which I'm used to. I didn't do any special maneuvering, and on top of that, you talk to me about a submarine. You must be kidding!'" (Tree 136-137). The outside observer assumes the pilot is representing the situation in the same way he is, i.e., as if the

pilot is trying to guide the submarine through the sea and towards a destination. The outside observer thinks the pilot sees what the observer sees because the pilot is assumed to reconstruct through his instruments the same picture of the world the observer constructs through visual sampling. But the pilot is simply observing the internal dynamics of system-specific operations; he believes he is carrying out certain procedures and corrections in order to maintain the correct organization of the system. In other words, the pilot does not even know he is a pilot. But "logical accounting" tells us that the two perspectives are both correct within their particular domains. The observer is correct because from a system-external viewpoint the submarine has displayed purposive action in an environment; he sees the submarine's movements as the pilot's way of dealing with particular environmental stimuli. The pilot is correct too, because he is speaking from within the domain of the system itself, in which the only pertinent goal is the maintenance of the system's organization, and in which environmental stimuli appear not as representations of an outside world but as perturbations of structures (i.e., dials and gauges) which, by reference to prior states of stability, the pilot attempts to correct.

Social Systems and their Environments

Here it might be pointed out that autopoietic theory seems to flirt with the relativistic paradox familiar to anyone with even a passing knowledge of postmodernism and contemporary debates on epistemology. If Maturana and Varela are suggesting that different observational perspectives produce different observations, and that observations could always be otherwise, then we appear to be traveling down a well-worn path. Furthermore, if they are heard to say that autopoietic theory provides a meta-observational, "observation of observation," whereby we take a perspective that includes all other partial ones, in what way do they depart from Bateson's concept of a "pattern that connects?" In fact, however, they reject tout court the possibility of an extramundane perspective from which diverse observations could be bound together, nor do they obey the traditional pragmatic or Russelian injunction to simply ignore nettlesome logical paradoxes by retreating to the next level of logic. They refuse as well to make the deconstructive gesture of fixating on the contingency of observation--the fact that it is "observation all the way down"--and so do not reinscribe "reflexivity" itself as kind of privileged signifier. In fact, they are not interested in probing how we may evade epistemological relativism at all, but would rather ask, to paraphrase Richard Rorty, relative to what do we think we are being relative? The crucial move here is a

reinterpretation of what epistemology is supposed to ground in the first place, namely, the subject's knowledge of the world. Varela proposes that "the proper units of knowledge are primarily concrete, embodied, incorporated, lived. This unique, concrete knowledge, its historicity and context, is not 'noise' that occludes the brighter pattern to be captured in its true essence, an abstraction, nor is it a step toward something else." Rather, it is "how we arrive and where we stay" ("Reenchantment" 320). The issue of relativism becomes a moot point because knowledge is always embodied knowledge, situated by the observational "cut" that defines the system/observer. Relativism is not something we need to "solve;" it is rather just the condition of possibility of observing the world. When I want to look at what is in front of me, I do not lament that I can no longer see what is in back of me. The sacrifice I make to see where I am going is that I do not see what I have left behind.

"Knowing" under this regime is not to be understood as a representation of a pre-given reality but as a kind of mental coping activity--cognition that is better described as Darwinian than Cartesian. We move from philosophy to biology, simply leaving behind the dilemmas that cognition (through philosophy) has presented to itself. Cognition becomes a function of the "fast dynamics" (rapid emergence and selection) of cognitive structures or "micro-states," the myriad neural ensembles which give rise to thoughts and actions by drawing from the history of structural coupling between organism and environment as embodied in previous cognitive micro-states. We learn by doing and by doing we learn: "perception and action are embodied in self-organizing sensorimotor processes; . . cognitive structures emerge from recurrent patterns of sensorimotor activity" (335). Purposive thoughts and actions are quickly assembled from available cognitive structures, often jerry-built to meet the demands of new and unfamiliar situations from "ready-to-use," habitual patterns of behavior. Varela summarizes as follows:

> First, knowledge appears more and more as built from small domains, that is, microworlds and microidentities. . . What all living beings seem to have in common. . . is knowledge that is always constituted on the basis of the concrete; what we call the "general" and the "abstract" are aggregates of readiness-for-action. Second, such microworlds are not coherent or integrated into some enormous totality regulating the veracity of the smaller parts. It is more like an unruly conversational interaction: the very presence of this unruliness allows a cognitive moment to come into being according to the system's constitution and

> history. The very heart of this autonomy, the rapidity of the agent's behavior selection, is forever lost to the cognitive system itself. Thus, what we traditionally call the "irrational" and the "nonconscious" does not contradict what appears as rational and purposeful: it is its very underpinning. (336) [4]

At the social systems level, the jostling of observational domains mimics the action of the brain itself, resulting in an "ungrounded" epistemology that allows the acquisition and validation of knowledge to proceed not by referring statements to some transcendent check-list but rather via continuous self-observations combined with the constant reference of one domain to another. In fact, as we will see in a moment, any particular observational domain is actually blind to its own categories of selection, the distinctions it uses to navigate the world; it simply cannot see what it uses to see. What counts as "common knowledge" (though a better term might be "common ignorance") is the sum of many observational domains pointing out the blindspots of one another. Each social domain privileges its own view and slights those of its peers, and out of all the crisscrossing something like integration emerges. Knowledge looks over its shoulder to avoid tripping over its feet. When we try to ground knowledge in one domain to the exclusion of others, then it can be said that knowledge as such ceases to exist. Just as a brain restricted to one channel of interaction is autistic, a society focused on one mode of observation, and therefore one kind of knowledge, is likely dysfunctional.

[4] For a complete discussion of these ideas, see The Embodied Mind by Francisco Varela, Eleanor Rosch, and William Thompson.

from Chapter Four: Thoreau's Moral Vision

Conscious Systems: From Hybridity to Embodiment

{...}

But what sort of "subjectivity" would be associated with conscious, autopoietic systems? We can start to address that question by reviewing some recent work on the structure and function of the brain itself. The same cybernetics revolution that informs the social systems approach of Luhmann also gave rise to a new branch of inquiry, namely, cognitive science. Cognitive science's dominant paradigm--the so-called "cognitivist" approach, much indebted to figures like Warren McCulloch, Noam Chomsky, and Marvin Minsky--has contributed mightily to a great many areas outside of brain research itself, most sensationally the attempt to design Artificial Intelligence, computers capable of understanding human language and responding appropriately to any input (i.e., to generate its own programs spontaneously). Such a "thinking machine" would by definition pass the "Turing Test," which is to say its responses could not be distinguished logically from a human's in a blind examination.

The assumption has been that our own cognitive flexibility originates in the brain's ability to create--somewhere within its 10^{15} neurons--symbolic representations of the contents of the world (and by "representation" I refer to the strong, epistemological sense of the word as a correspondence between symbol-in-the-mind and pre-given thing, not in the weak, semantic sense as simply the way we use one thing to stand for something else in everyday communication). The AI would be equipped with a similarly comprehensive symbolic catalog; it would be able to turn inputs into appropriate symbols, rearrange them in new patterns based on complicated algorithms, and produce the most effective response to a given stimulus. But while current rule-based programming languages have a remarkable capacity to process routinized information, the cognitivist approach is seen increasingly as fatally flawed: the generative grammar necessary to respond to an infinite number of input contingencies may be impossible to design. In fact, as early as the famous Macy conferences in the mid-fifties, the notion of cognition as a centralized, rule-governed process of information coding, decoding, and routing was being questioned. The "connectionist" paradigm has been offered as a possible alternative. In this model we find the concept of emergence--a view of the brain as a networked system of connections that produces global consequences arising spontaneously from many localized neural firings--beginning to take center stage. Emergence depends not

simply on real-time effective response to external stimuli but also on the history of previous neural activity. This means that if particular neurons have in the past participated in local activity, the bond between them and their partners is strengthened and, conversely, diminished when there is less coordinated activity (Hebb's Rule). As Varela puts it, "the system's connectivity becomes inseparable from its history of transformation, and related to the kind of task defined for the system" ("Whence" 245). In other words, an intelligent computer designed along these lines would "grow" its pattern-recognizing capacities just as a human does: by learning and shoring up its cognitive capacities over time. A significant implication of the theory of emergence is that the notion of the brain as primarily a symbol-processing organ must be discarded, for cognition arises from patterns of neural activity that occur at a sub-symbolic level, meaning that particular symbols are not encoded point-to-point with the environment but instead are relative to the overall state of the neural system.

But despite this important shift, connectionism as an approach to AI once again awaits the development of some kind of generative grammar, for no matter how "fine-grained" the grammar may seem, "At the heart of the most volatile pattern-recognition system ('connectionist' or not) lies a von Neumann engine, chugging along, computing a computable function," as Daniel Dennett playfully puts it (269). The connectionist model is in the final analysis still a representational one, because the brain "works" only insofar as it can replicate within its distributed structures some sort of accurate picture of the world within which it must operate.

To fully leave behind the representationalist approach Varela has proposed a modified connectionist approach. This "enactive" model looks very much like the logical extension of his and Maturana's general theory of autopoiesis to the study of cognitive systems. According to Varela, enactive cognition can be delineated by answering three basic questions:

> Question #1: What is cognition?
> Answer: Effective action: History of structural coupling which enacts (brings forth) a world.
> Question # 2: How does it work?
> Answer: Through a network of interconnected elements capable of structural changes undergoing an uninterrupted history.
> Question # 3: How do I know when a cognitive system is functioning adequately?

> Answer: When it becomes part of an existing on-going world of meaning (in ontogeny), or shapes a new one (in phylogeny). ("Whence" 256)

Taking these three points together, we understand that the crucial departure from the connectionist view is that now we can dispense completely with the idea of the brain reconstructing internally an external, pre-given world, and instead say simply that cognition and world arise simultaneously. The brain does not merely produce behaviors that accurately register the contours of an outside world but, based on biological, psychological, and social histories, determines for itself what counts as its world. In no way can this process be thought of as "representing," because the world brought forth can quite literally vary from brain to brain depending on the history of structural coupling produced during phylogenetic and even ontogenetic development. For example, colors, shapes, and shadings are not "out there" to be recovered (nor, to anticipate the charge of idealism, "in here" to be projected) but instead result from specific selections of information, selections that emerge during a history of interaction. The world as seen though the compound eyes of an insect as opposed to the binocular vision of a human is not simply a different view of the same fundamental reality: what is taken as "world" actually co-evolves with the sensorimotor apparatuses that have coupled with it. Depending on how those apparatuses have developed, the "embodied actions" of the organisms produce particular cognitive states which in turn reproduce the effective actions that guide perception in that world. In other words, the world appears as it does neither because we project our internal states upon it nor because we retrieve from it invariant properties, but because we have evolved to see and act on it in specific ways--and, to a very substantial degree, it has evolved with us.[5] In The Embodied Mind, Varela and co-authors Evan Thompson and Eleanor Rosch write:

---

[5] For example, it is a relatively well-established point in evolutionary biology that the colors of flowers and fruit have co-evolved with the vision of the insects and animals that pollinate them and disperse their seeds. Both plants and animals benefit from the relationship (the plants increase their reproductive capacity and the animals get food) so that selective pressures favored striking colors and the color vision required to notice them. The larger question, of course, is "Did not colors exist before there were eyes to see them?" The answer is that the reflectance properties of elements and chemical compounds did exist but that without the evolution of vision to distinguish the different wavelengths of light our world would seem a drastically different place. Without color vision, would gold or precious gems have any value to us? More importantly, would these substances--now various shades of gray--actually "be" gold, red, green, etc.? The "reality" of colors is nothing more than a function of

> The key point, then, is that the species brings forth and specifies its own domain of problems to be solved by satisficing [solving problems by using whatever works, not what is optimal]; this domain does not exist "out there" in an environment that acts as a landing pad for organisms that somehow drop or parachute into the world. Instead living beings and their environments stand in relation to each other through mutual specification or codetermination. Thus what we describe as environmental regularities are not external features that have been internalized, as representation and adaptationism both assume. Environmental regularities are the result of a conjoint history, a congruence that unfolds form a long history of codetermination. (198)

The point to bear in mind is that the brain does not receive information from the environment as input; it creates its own information on the basis of physiologically opening itself selectively to perturbations in the environment. In reductive terms, we are describing a kind of "adequation as survival," a jerry-rigging of perception, cognition, and action to what surrounds us: we see what we see because it works, and until it stops working we will continue to see it that way. "Thus," say Varela, Thompson, and Rosch, "the overall concern of an enactive approach to perception is not to determine how some perceiver-independent world is recovered; it is, rather, to determine the common principles or lawful linkages between sensory and motor systems that explain how action can be perceptually guided in a perceiver-dependent world" (173).[6]

---

our ability to see them (as opposed to other possible wavelengths), or as Luhmann might say, the blindspot of particular distinguishing operations that select some realities at the exclusion of others.

[6] The obvious connection to certain strains in phenomenology is encapsulated in a statement by Merleau-Ponty that Varela is fond of quoting: "When I begin to reflect, my reflection bears upon an unreflective experience, moreover my reflection cannot be unaware of itself as an event, and so it appears to itself in the light of a truly creative act, of a changed structure of consciousness, and yet it has to recognize, as having priority over its own operations, the world which is given to the subject because the subject is given to himself. . . . Perception is not a science of the world, it is not even an act, a deliberate taking up of a position; it is the background from which all acts stand out, and is presupposed by them: The world is not an object such that I have in my possession the law of its making; it is the natural setting of, and field for, all my thoughts and all my explicit perceptions" (Phenomenology of Perception x-xi)

Varela has theorized that conscious human systems (i.e., what we call "minds") operate through this same combination of organizational closure and structural openness that we have discussed many times previously. Once again, a main concern here is to steer a course between, as Maturana and Varela say, the Scylla of realism and the Charybdis of idealism. Thinking machines designed along these lines would, through many receptors and their own associated effectors, respond to various environmental stimuli appropriately--without the presence of a central processor with a universal grammar coordinating the total activity through a master representation of the environment. What looks to an observer like a system using an inner map to navigate a territory is rather an ensemble of micro-responses, each sensorimotor unit pursuing its own "goals" with the only proviso being that the various units (structures) must be compatible (i.e., to have evolved in parallel) so that their independent activities do not lead to the interruption or destruction of the totality of activities (the organization).

An apt analogy here is the coordinated movement of a school of fish or a flock of birds: although it is tempting to imagine that each animal is aware of the overall direction and purpose of the group as it moves through its medium, the observed coordination of the group depends solely on the sensorimotor activity of each animal as it responds to stimulus in its own immediate environment. Leading and flanking animals, for example, may respond to food or prey by changing their direction toward or away from the stimulus, while interior or tailing animals may simply react to minute changes in the spatial disposition of the animal in front of or beside them. As long as the members of the group continue to react appropriately (i.e., nothing occurs to compel each individual to follow other, more deeply necessitous behaviors) the school or flock remains organized as such. In essence, what is taken by an observer to be the astonishingly well-orchestrated behavior of the school or flock (the "group mind," according to an earlier interpretation) is not orchestrated at all: it is nothing more than the total effect of many small course corrections (structural changes) in compliance with a modest set of organizational rules. The larger point here--with respect to the proposed AI or the human brain--is that what we like to think of as mind, ego, self, subjectivity, etc., is simply our attribution of a governing core to a distributed, organizationally conservative process such as this, what Daniel Dennett calls a "Center of Narrative Gravity" (410). There is no "seat" of consciousness, a cogito in the Cartesian sense (or "Me" in the Emersonian one), but only many local sites where the brain responds to stimuli, the

global outcome of which is the coordinated behavior we believe (wrongly) is being masterminded by a central processor.

Admittedly, the prospect of an imaginary mind seems to violate common sense. The self surely must be a thing apart from the physio-chemical interactions that compose the brain and body, a hard kernel of ego that keeps us who we are across space and time. Even those who have given up on the "soul" find it hard to relinquish their "mind," for how could there not be some sort of ghost in the machine, an entelechy that animates and gives focus to these three pounds of skull-bound meat? But "common sense" (as in the quotidian thoughts and feelings that shape our inner lives) is precisely what is at issue here, for in the enactive view common sense now arises from the successful history of coupling between many sensorimotor structures and their specified environments, which together compose the ongoing project of individual growth (ontogeny) or species evolution (phylogeny). "Common sense" would then be defined as just this sort of embodied knowledge, knowledge that allows us to perceive, predict, create, and act effectively within our surround. The "mind" or "self" is perhaps no more than the sum-total of this knowledge, a repertoire of habituated neural firings, flexible patterns of sensorimotor activity, and ready-to-hand experience. In other words, what we think of as the self may simply be the constant pulling-together of a variety of neural operations, a self-generating program designed only by evolution, cultural history, and individual psycho-physiology. Our perennial grasping after the certainty of ego-identity is thus a surface effect of a great many subsystems and micro-structures working in tandem to preserve the brain's autopoiesis.[7]

---

[7] An interesting theory as to the origin of self-consciousness from a systems perspective is outlined by Eric Rosseel: "Self-observation occurs when the structural coupling of a human system to a specified ambience gets blocked by an environmental perturbation that is compensated by a 'breakdown' of the action of the system. As the system does not dispose of an adequate action that would have continued his drift in the ambience, 'mindless' automated action is interrupted and self-observation helps explore the plasticity of the systems and to develop a new 'creative' action that removes the blockage. . . Only in those moments , we do see what we see, we look at the world and talk about the images we have formed. . . In other words we act with Intelligence" (236). What I take this to mean is that self-consciousness emerges as a response by the neural system to problems that cannot be solved more economically by autonomic or instinctual operations. At moments when habitual actions and responses are inadequate, we find ourselves thrust into the creative, reflective, self-aware states of mind we take for granted as the normal state of affairs; rather, according to this view, such states are, strictly speaking, the abnormal condition.

The autopoietic view of mind as an "answer" rather than a "question" is not so very different from the Lacanian notion of the subject as a retroactively constructed entity that exists only as an explanation for various displacements in a field--in other words, an effect in search of a cause. The reason the subject so often feels alienated, ungrounded, "at loose ends" is because she really is: that is what it feels like to be a subject, this void at the center of a whirlpool. Likewise in deconstruction it is the rupture between signifier and signified that impels subjects to traverse endless chains of substitutions towards a ground, an identity, a "transcendental signified," that never actually materializes. Such materialization, to paraphrase Derrida, could only be marked as the arrival of death or God, for in a manner of speaking the end of signification is either a pure absence or a pure plenitude--and for all intents and purposes the result would be the same: an end to the signifying process and therefore subjectivity itself. In both versions of poststructuralism, self-consciousness is finally just this "desire" for self, for the certitude of wholeness in the face of radical contingency and difference.

In systems parlance, a self-reflexive self--a mind that can consider its own existence--is the propitious issue of a highly-differentiated, multi-layered, closed form of neural organization, at the heart of which are a host of difference-producing operations and structures that permit the organization to exist at the price of its continual separation from an insuperable Other (its environment). According to this interpretation, we can dispense with the mind/body duality that has vexed philosophy for so long, and conclude simply that mind and body are two parts of the same autopoietic process. Cognitive organization (mind) arises from a particular concatenation of bodily senses and neural networks which in turn continue to concatenate only so long as their coordinated organization can be maintained: Mind equals body equals mind.

If it seems that I have only added yet another overlay of arcane technical language to an experience of self that most of us can safely take for granted, I think the epistemological lesson is perhaps a little more useful: If subjectivity is the self-description of a closed unity structurally open to the environment through a variety of sense organs and effectors, then subjects' feelings of both disconnection and contact with that environment come to be seen in a different light. Firstly, we now consider "knowledge" to be just another of the many adaptations we have evolved to allow us to "cope" (i.e., structurally couple) with environments; it is storable,

transmittable, action-in-the-world.[8] Secondly, because this knowledge of the world is embodied, not objective, there is no particular path the subject is forced to take in the realization of his autopoiesis; any path is permitted as long as it is viable, that is, it does not destroy his cognitive and/or bodily organization. Crudely speaking, if the subject decides to walk off a cliff he will shortly discover the non-viability of this manner of structural coupling, not because his ability to properly "represent" the features of the world has let him down but because, despite the environmental cues that should have warned him away, he failed to secure the further organization his cognitive (not to mention physical) system.

Ultimately, then, we are simultaneously connected to this world through our sensorium yet disconnected to it, in that information selections can only contribute to the maintenance of our cognitive organization. The world is just the world we enact through living in it, and there is no ground but the ground we walk on. Thus our feelings of rootlessness are fully justified--but so too are our feelings of connection. The difference in affect may simply depend on the mental state we find ourselves in. When we engage in reflective, self-conscious thinking we feel disembodied, inasmuch as we are dealing with problems of "organization" (i.e., how we can adjust our minds to conform to what our sense are telling us or, said another way, how we can "rationalize" the world so as to preserve our cognitive integrity). But when we are not engaged in this problem-oriented manner of thinking, perhaps during well-rehearsed activities or autonomic functions, or simply when we allow our

---

[8] In his fascinating discussion of evolutionary epistemology Darwin Machines and the Nature of Knowledge, Henry Plotkin argues that the term "knowledge" can be applied to any sort of adaptation, that, for example, "the fleshy water-conserving cactus stem constitutes a form of knowledge of the scarcity of water in the world of the cactus" (228). Adaptations are in effect "forms of 'incorporation' of the world into the structure and organization of living things" (xv). He contends that "knowledge is a complex set of relationships between genes and past selection pressures, between genetically guided developmental pathways and the conditions under which development occurs, and between a past of the consequent phenotypic organization and specific features of environmental order" (228). The more commonplace sense of the word "knowledge" can conversely be thought of as an epiphenomenal consequence of the more profound adaptive cognitive structures which gave an advantage to their possessors. A trivial piece of knowledge, such as the fact that "Germany won the World Cup, is really only the visible, or potentially visible, part of a complex multiple-layered and historically ordered hierarchical structure involving the genes which code for the brain structures that enable me to gain knowledge, development which led to the establishment of the required brain mechanisms, brain and cognitive states that are the present embodiment of that knowledge, and culture and its artefacts that allow me to learn rapidly and accurately what is occurring in a distant part of Europe" (229).

minds to wander, that is when we feel most embodied and grounded, most open to the "a-rational" possibilities of the world we have enacted through living.[9]

As Varela, Thompson, and Rosch have argued, this latter state can be cultivated. Drawing on the Eastern tradition of philosophy, they find a model for a state of heightened openness to the environment in "mindfulness/awareness meditation." They suggest that in our Cartesian heritage we have developed too much of a penchant for disembodied modes of reflection, those styles of thinking which abstract all experience down to the closed circuit of our consciousness and leave is under the impression the world is real only insofar as it can be represented confidently in the theater of our minds. In effect, the Cartesian tradition exacerbates the much-ballyhooed discontinuity between self and world, mind and body (and by extension, society and nature) by focusing, as it were, on the organizational closure of subjectivity at the expense of its structural openness. Varela et al propose instead that once we learn to reflect on "reflection" itself as just another embodied experience, we can escape the closed loop of disembodied self-consciousness, and regain that sense of connection we seem to long for so desperately. "When reflection is done that way," they claim, "it can cut the chain of habitual thought patterns and preconceptions such that it can be an open-ended reflection, open to possibilities other than those contained in one's current representations of the life-space" (27). Practically, they suggest that a deeper understanding of the "ego grasping" that is a response to the apprehension of groundlessness has potential importance for the individual (who can learn to moderate such ego-grasping and attain a greater degree of psychological comfort), as well as for society as a whole, the institutions of which (particularly those of science) are far too engaged in processes that separate, to our detriment, experience from facts. (This discussion, by the way, is not meant to introduce yet another "bad" dualism, that between reason and intuitive thinking, for all knowledge is finally embodied knowledge, dependent on a history of bodily action and response, and all thinking is a form of self-observation in which current states of mind are compared against previous ones, and projections of future states are either confirmed or denied. So what is meant here, rather, is that just as there are different sorts of ambulatory motion--walking, running, jumping--there are different ways of thinking, and that what I have called the "disembodied" style has been cultivated over the "embodied" one, even though both are finally examples of a common cognitive activity.)

---

[9] For a shorter, non-technical discussion of these ideas, see Varela's Ethical Know-How.

As others have remarked, part of Thoreau's own intellectual achievement was to have aspired to a vision of nature, life, and literature in which the remoteness of the Cartesian subject is replaced by an observer who is very much a part of the world. Both as amateur scientist and as artist, Thoreau was committed to proving the facts of the world through his body before putting pen to ink; in this sense, "embodiment" appears similar to what Laura Walls calls his "epistemology of contact," by which she means

> a scientia that would be relational rather than objective. This "relational knowing" extended and applied the possibilities opened up by the disintegration of subject/object dualism, which encouraged the subject to "know" by seeing correspondence in the world's objects, as if they were the mirror of the self, or by "reading" the book of nature as if it were a text ready-made for decoding. By contrast, knowing as an active process in Thoreau's sense becomes no less than what H. Daniel Peck calls "worlding," the making of a world "by the interaction--the dance--of the creative self and the world." (147)

Quoting from the Journal, Walls supports her view that against the objective fallacy and the problems associated with representation, Thoreau proposes just such a relational stance: "The important fact is its [the object's] effect on me. . . With regard to such objects, I find it is not they themselves (with which the men of science deal) that concern me; the point of interest is somewhere between me and them (i.e. the objects)" (10:164-65; quoted in Walls 206). Walls tells us that in "Thoreau's alternative science authority comes from individual involvement and experience," and that "we are all similarly involved, implicated. As our designs tangle with those of willows and squirrels and oaks and beggar-ticks, we all become co-producers of Concord, and by extension, the 'environment' around us, wherever we are" (207). Walls' point is that Thoreau provides us with an example of how we too can move toward an "epistemology of contact" and away from our current model of knowledge (which we might wish to call the "epistemology of disembodiment"). As neither a Hoon-like solipsist making the world in which he walks nor a detached objectivist recording "just the facts," Thoreau treads a middle way that brings the two perspectives into accord.

Sharon Cameron also finds in the Journal evidence for this "dance" between subjectivity and objectivity, whereby observation and reflection on nature constitute not merely steps toward an increase of knowledge but instead are methods to actually insert nature into the mind: "The internalization [of nature] is not the result

of analogic correspondence conventionally understood. . . analogies cannot effect comparisons between nature and human nature (the two are incomparable); they must rather effect transfers, in the express sense of moving natural phenomena. . . into the mind and onto the page, where the mind can testify from close up to the fundamental difference between itself and what it contemplates" (15). Thus, Thoreau's practice in the Journal of attending to and describing the minute details of nature is understood by Cameron to mean that "to see nature--whether the flooded Sudbury meadows, or the swamp, or winter, or the mosquitos' humming--is to take it into the mind while all the time recognizing it is not of the mind" (153). His descriptions are therefore not sentimental representations, intended to evoke memories of favored natural objects by means of analogy or metaphor (although they do do that, in the weak sense that all writing is representative) but, as Cameron puts it, "they rather exist to enact displacements," (150)--or in my terms, they act as a kind of cognitive catalyst.

What exactly do they catalyze? Not a rapid search of a mental portrait gallery of nature scenes but rather the same state of neural activity that the described objects themselves originally excited. What seems like an image pulled out from an internal Rolodex of memories is instead a particular synaptic cascade that, once triggered, reproduces the mental conditions at the point of the original coupling. In terms of autopoietic cognition, "writing nature" would then be Thoreau's version of embodied knowledge: inscriptions of nature represent histories of couplings (visual, tactile, etc.) between his cognitive system and the world. It may seem like a small point--whether memories are pictures in the brain or neural ensembles that allow us to constantly recreate the pictures--but it means that, in a manner of speaking, the world is its own representation, and that we keep bringing it forth again and again as we interact with it through coupling. The closed organization of the brain suggests that we do not mirror the contents of the world in our heads but instead at every moment use those contents as guides for coordinating our cognitive operations.[10]

---

[10] The brain is not a mirror of nature, but neither is it a record needle that plays nature's tune; in effect it plays its own tune so long as nothing forces it to change the melody. Again Rorty is instructive: on the same subject his analogy between representation and world is to imagine mind as an encyclopedia, which "can get changed by things outside itself, but [which] can only be checked by having bits of itself compared to other bits. You cannot check a sentence against an object, although an object can cause you to stop asserting a sentence. You can only check a sentence against other sentences, sentences to which it is connected by various labyrinthine relationships" ("Pragmatist's Progress" 100). In a sense, the brain is such

Without the ongoing history of structural coupling between the brain and an environment we would become autistic, for we need the world to act as the template on which the brain can anchor the overall coherence and disposition of its sensorimotor structures and processes, i.e., its organization. Absent those structural regularities the organization would be stressed; and without the time to develop new structures (either through phylogenetic or individual change, depending on the speed and manner of environmental perturbation) the organization would eventually break down, as its structures found themselves in the grave situation of being unselective and unresponsive to the new environment.[11]

The balance between intimacy and otherness of nature both Walls and Cameron attribute to Thoreau in his inscription of nature is very much in accord with the cognitive stance proposed by Varela et al in The Embodied Mind, their "middle way" by which to "get in touch" with self and world by "meta-reflecting" on our embodied interaction with it. Thoreau demonstrates the middle way as he oscillates between assembling the hard facts of nature so as to develop accurate, scientific representations of it and reflecting that such representations depend on the subjective experiences which color those facts in the first place. Nevertheless, my point will be that the cognitive dislocation one expects to see produced in trying to straddle both styles of thinking--and which we do observe in Thoreau--is never the consequence of this balancing act but is in fact its prerequisite: cognitive dislocation comes first, as a result of the closed/open design of embodied cognition. So we must bear in mind that the reaffirmation of embodied knowing is not actually a "solution" to the Cartesian predilection for disembodied reflection: in the end both sorts of knowledge are always options for self-observing cognitive systems, whose fundamental mode of operation is drawing distinctions between one thing and

---

a closed encyclopedia, which can reconjugate its contents in response to new facts so that it remains consistent and coherent across its organization. The relation between what the encyclopedia says and the contents of the world is relevant only in that the world can force the encyclopedia to reconjugate its contents: the world thus acts as stimulus, not as a point-to-point guarantee of "truth," for "truth" only consists in the integrity of this organization, that is, there is no contradiction between the contents.

[11] Note that this is not quite the same as Bateson's idea that, because our mind and environment are cybernetically linked, when we destroy the environment we automatically destroy the mind. For, in fact, the mind is only concerned with maintaining its own organization, and organization may be preservable even though much in its environment is destroyed--as we know from bitter experience. It is the case, however, that one outcome might very well be the degradation of the cognitive organization, i.e, "madness," assuming the right kind and degree of environmental perturbation.

something else. To be able to consider one's autonomy, to be able to reflect on one's reflections, to be able to consider the difference between what one takes to be one's self and what is taken to be the Other--all of these modes of thinking depend on the openness of structure and the closure of organization, i.e., the idea that although we can sense an environment that is not the same as us, this difference cannot tell us how to organize our cognitive processes but only provide, by its continual reentry, the means to do so. The various dualisms we have come to know so well are therefore different ways of putting a name to the fundamental cognitive dissonance that comes about as conscious systems perceive the insuperable difference between themselves and their environments--and a fortiori the various attempts to eradicate those dualisms are fated only to produce new dualisms. In other words, the "split"--no longer understood as a subject/object, reason/intuition, or mind/body split but rather distilled down to the difference between the cognitive system and its environment—is not something we are ever likely to heal, expunge, or otherwise resolve, because it is the very precondition of self-observation in the first place.

Ironically, then, what Varela, Walls, Cameron, Peck, and presumably Thoreau himself hold up as an exemplary way of knowing ("relational knowing," "worlding," or "mindfulness/awareness") stands revealed as just such an attempt to enfold mind into environment, to give oneself over to the most "embodied" of all the embodied knowledges provided from the history of structural coupling, as if the contemplation of the bodily roots of thinking and perceiving could tell us how to think and act. But the most the attendance to such knowledge can really do is allow us to ignore, however briefly, the wound on which is built our subjectivity. If attention to structural "contact" gains any purchase at all--if, literally, we "forget the pain" of being ourselves for a moment--it can only mean that the fundamental rift between conscious system and environment is still in play, still presenting itself as a separation to be overcome. In other words, the system persists according to the principles of self-organization and autonomy, open structurally to the vagaries of its environment but creating its own information about how those environmental perturbations shall be taken. The environment cannot tell us what to think about it, how to relate to it, how to navigate through it; all it can do is cause us to "try" something else or continue to do what we have been doing successfully. So the final outcome of Thoreau's effort to hybridize distance with intimacy, objectivity with subjectivity is that it will fail to close the cognitive rift that allows the generation of those dualisms, will in fact only allow him to forget about them for a while--although clearly, as Varela et al argue, that goal alone is certainly compelling. But once the

observation of observation enters the evolutionary scene and offers a second-order, self-reflexive awareness of the difference between self (system) and Other (environment), the possibility of eradicating this rift has been closed off forever, for between death and God there exists only the autonomous consciousness constantly comparing itself against its previous states, never able to transcend the difference-generating operations that distinguish it from not only the environment but from itself from one moment to the next. To be an embodied conscious system means, at the most fundamental level, that cognitive displacement or dissonance will never cease.