[This draft is a pre-publication typescript of my recently completed book. Responses are welcome at jaogilvy@gmail.com.]

Coming Together

How the Emergence of Life, Evolution, and Language Shed Light on the Emergence of Consciousness and Love, Wealth and Artistic Creativity

James Ogilvy

2024

Contents

Introduction

It's easy to underestimate just how much we *don't* know. Mountains of knowledge obscure whole ranges of ignorance. Ever since the pioneering work of Galen in the 2nd Century and William Harvey much later, we've known how the heart pumps blood through our circulatory system. But who among us knows the nature of love? Ever since Watson and Crick unlocked the double helix we've known how DNA works to replicate strings of amino acids. But who among us knows the origin, or even the nature, of life? Linguists have mapped, compared, and contrasted the grammars and vocabularies of hundreds of languages. But no one claims to know with any certainty the origin of language, or even how small children learn language before our very eyes.

And consciousness! What is it? How does it relate to the brain, the mind, the self? Its nature is so obscure that researchers now speak of "the race for consciousness" the way scientists in the 1960s talked about the race for the moon.

There is so much we don't know.

It's worth reminding ourselves of our ignorance as a way of clearing a space for knowledge of a fundamentally new kind. It's worth remembering from time to time how little we know about life and language. We are as savages, primitives compared to those who will come after us. Vis-à-vis life we are like cave men who use fire, wonder at its warmth, but don't know how to start it. Those who follow us will look back at us and marvel at what we did not know just as we look back at those before Galen who knew so little about the circulatory system.

There is so much we don't know.

Just as we can look back and see how medieval ways of knowing made it difficult for those at the dawn of experimental science to convince their colleagues of new ways of knowing, so they will look back at us and see how our ways of knowing made it difficult for us to understand the things that elude our science. And there are many such things, e.g. life, language, love, consciousness, money and creativity.

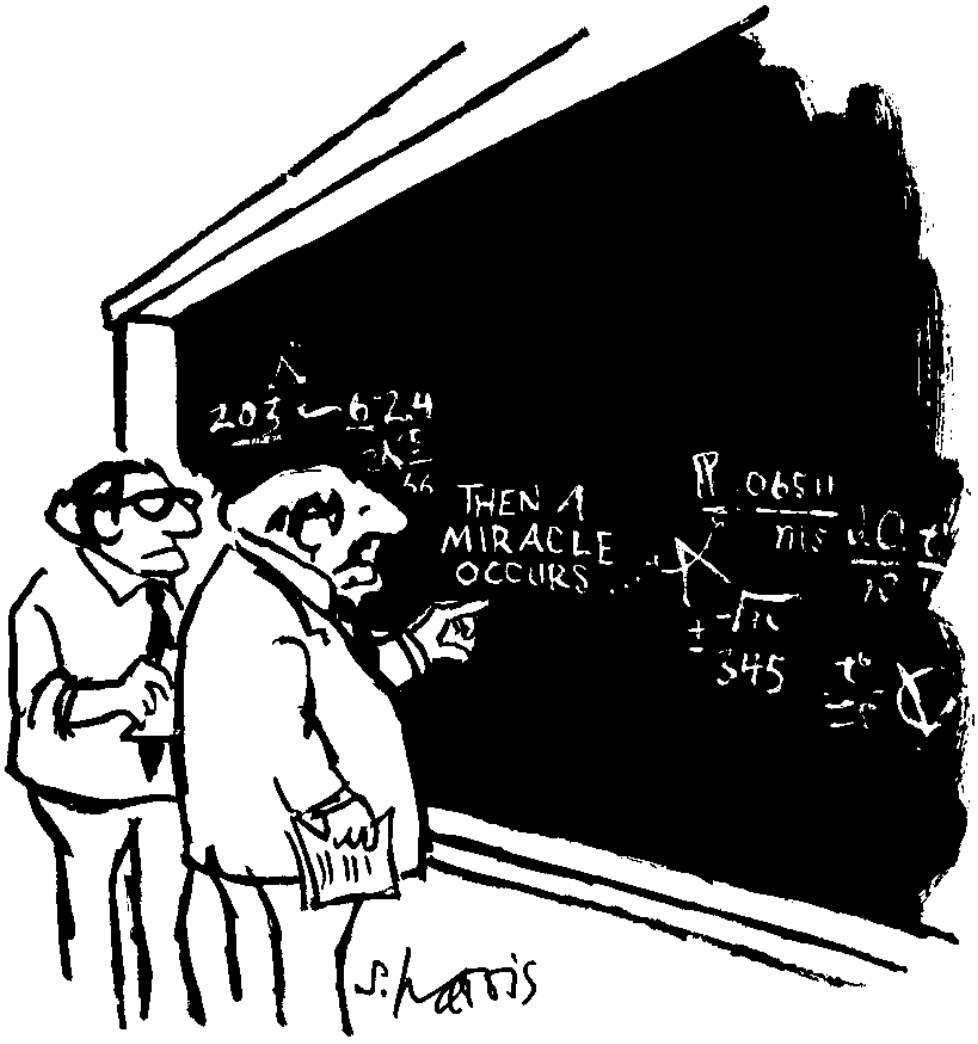
Despite the magnificent achievements of science—the discoveries, the inventions, the well-earned Nobel Prizes—the proud recounting of our successes should not obscure a

humble recognition of our failures. We don't know how to foster economic growth in so-called 'developing economies.' We don't know how to prevent or cure depression. We don't even know that much about knowledge itself: what it is to know, how knowledge happens. If we did, we might do a better job of helping students climb the mountains of knowledge we have. But we don't, as our manifest failures at educational reform attest.

This book is built on a premise of ignorance, and the hypothesis that a new way of knowing may be necessary if we are to come to know processes like life and language. If so, then we might get a better purchase on other emergent phenomena, including consciousness. Once having become fluent in the language of several more familiar emergent systems, then the mysteries of money, love, and creativity might open up. Three other emergent phenomena--ant colonies, cities, and software, don't need our attention here because they have already been so insightfully discussed in Steven Johnson's book, *Emergence* (Scribner, 2001), over two decades later still the most accessible introduction to the concept of emergence, which can be somewhat clumsily defined as the coalescence or coming together of many parts into wholes whose properties are neither predictable from, nor reducible to the properties of those prior parts. Desire, consciousness, and emergence: these are among the three themes that come together in this book. We will treat life and language in some detail; move on to see what light they cast on love and consciousness; then conclude with more cursory treatments of money, and creativity.

Desire and consciousness are familiar aspects of everyday experience, if obscure to theoretical understanding. Emergence is less familiar: something of a riddle. In the literature where emergence is a familiar term, too often it appears as a miracle. If writers are having trouble giving a causal account of the origins of life, they reach for emergence by 'emergence'. as if it provided an explanation. If they are looking for a physical or materialistic account of the nature of consciousness, they invoke emergence from the neurophysiology of the brain. But the invocation of emergence in the context of the usual scientific explanation explains nothing at all, because we cannot give a rigorous account

of just what we mean by ‘emergence.’



The word ‘emergence’ plays the role of a rhetorical skyhook, an expression often used by philosopher, Dan Dennett that nicely expresses a lifting by something that has no purchase form which to lift. It stands in for that little note in the cartoon: “Then a miracle occurs.” Emergence used in this way amounts to a surrender of reason rather than an extension of reason.

This book offers an account of emergence and emergent systems that goes beyond the invocation of a miracle. It can be shown not only that emergence happens, but also what it is. It's slippery, because so many of our habits, so many of our criteria for what counts as rational, have been honed by a science aimed at explaining non-emergent phenomena—levers and pulleys, forces and impacts. In order to *explain* emergence, we'll have to let a new kind of explanation emerge. No small order.

This book is peculiar in a number of ways. It is not a book of science even though it talks quite a bit about science. Most of my academic training is in philosophy, but this book bears little resemblance to philosophy of science. There was a time when people spoke of "natural philosophy." Perhaps this book is best placed in that ancient tradition. It's about how the world works. Again, no small order.

Be forewarned that this book is radical. You will be invited to think thoughts and entertain questions in ways that are fundamentally at odds with much of what common sense has inherited from modern science. Specifically: we've learned the first law of thermodynamics that asserts the constant conservation of mass and energy; that you can't get more out of less. But the concept of emergence is defined by one of its foremost proponents as, precisely, "getting more out of less."

We've learned that good explanations explain complex phenomena by reducing them to their simpler components and then applying so-called covering laws to state descriptions of those simple components. 'Reductionism' it's called. But explanation by reduction runs contrary to emergence, downwards toward simple components as opposed to upwards toward emergent properties. Both science and common sense are used to explaining the "higher" by the "lower," the idea of emergence turns that world upside down.

We've learned that there are no "final causes," or, as Aristotle called them, *teloi.* Causality runs only forwards, from past to present, not backwards, from future to present. Teleological explanations were banished from the universe by the rationalists of the Enlightenment—Bacon, Newton, Galileo, and Spinoza. While Aristotle may have understood change on the model of an acorn becoming an oak as it actualized its indwelling potential, we've known for more than 300 years that Newton's falling apple was not *seeking its proper place* as Aristotle claimed. It is simply obeying the law of gravity. Teleological explanations that talk about "seeking" and "proper place" may be alright for explaining the actions of people who can articulate plans and purposes; but inanimate things cannot *seek.* They have no indwelling purposes, no *teloi.* Modern science respects pushes from the past but no pulls from the future; efficient causes but no final causes.

It's not altogether clear how thoroughly the rejection of final causes has migrated from science to common sense. Teleological reasoning persists in everyday discourse as when people cope with hardship with phrases like, "Everything happens for a reason." But this is wrong. This is wishful thinking. People seem to need purposes, a sense of *meaning* in their lives. So teleological reasoning persists in wrong-headed ways, whether scientists like it or not. So scientists and philosophers of science turn around and try to purge the universe of final causes . . . even where they are real and unavoidable.

Even in the sciences themselves, especially biology, teleological language persists in statements like, "The function of the heart is to pump blood." Final causes may be officially dead in modern philosophy of science, but the language of *function* lives on even though it has lost its metaphysical license.

The constant conservation of matter and energy, reductionism, and the elimination of teleological explanation by final causes are fundamental pillars of modern materialism—a picture of ultimate reality as nothing but matter and motion in space and time. According to this materialistic metaphysics, as Nobel Physicist Steven Weinberg puts it, "All the explanatory arrows point down." And further, "The better we understand the world, the more we see that it lacks purpose."

Modern materialism may be vastly superior to Aristotelian metaphysics when it comes to putting a man on the moon, curing polio, or splitting the atom. But when it comes to leading a meaningful life, reductionistic materialism comes up short. The rise of religious fundamentalism and the debate over so-called Intelligent Design may be seen as refusals to accept the metaphysics of materialism. People are understandably reluctant to accept the idea that their lives are meaningless; that, ultimately speaking, it's all nothing but matter in motion in space and time, a massive 3-D pool hall of molecules bumping off of one another in a galactic dance headed for ultimate heat death. This way lies nihilism: No meaning, no value, no glory, nothing . . .People are understandably reluctant to accept this idea. Nihilism can make it hard to get out of bed in the morning. Why bother?

Of course, there's always religion to fall back on . . . unless your intellectual honesty won't let you. The new atheists—Richard Dawkins, Sam Harris, Dan Dennett and Christopher Hitchens—have made it harder for thoughtful people to hold on to their religious beliefs. There is simply too much cognitive dissonance involved in embracing modern materialism on the one hand and believing in the sacraments on the other. If you believe in the ultimate truth of modern science, then you can't seriously believe that wine turns into blood, or bread into the flesh of Jesus. You can't seriously believe in an afterlife, or a personal god who intervenes in earthly life to work anti-scientific miracles.

Though Nietzsche announced the death of God over a century ago, people have been slow to get the news. But the new atheists are now making the choice more stark than

ever: *Either* you grow up and accept nihilistic materialism *or* you persist in childish beliefs according to which Santa Claus, the tooth fairy, and a personal God are available to give meaning to life, guarantee Christmas presents, quarters for teeth, and an afterlife in Heaven.

This book is about a third way between the poles of this very stark, very high stakes dilemma. The science of emergence promises a naturalistic, up-from-the-bottom account of the way purpose and meaning come to be in a universe that once lacked them. Purpose and meaning are neither absent from the world as scientific materialism would suggest; nor are they hard-wired into the nature of reality by a designer god as religion would suggest. Purposes and meanings *emerge*.

In order to understand how purpose and meaning can emerge, however, it will be necessary to challenge each of the pillars of scientific materialism—that you can't get more out of less, that there are no final causes, and that reductionism is the only form of satisfactory explanation. Radical!

You, dear reader, will have to ask yourself which direction you generally tilt when faced with the stark dilemma. Do you tilt toward materialism, damn the consequences? Are you *that* rigorous, *that* tough-minded, *that* strong? Or do you fall back on religion to give meaning to your life? Do you find goodness in a world order designed by a benign deity? Or do you lean this way or that depending on the day of the week or the situation at hand? Do you live with the cognitive dissonance of embracing both science and religion, materialism and God?

For my part, just to come clean, I've spent most of my life as a Nietzsche-carrying atheist. But reading the new atheists has made me question my atheism. I'm with them as long as they are poking holes in some of the more preposterous claims of religion . . . but then I find myself uncomfortable with their materialistic reductionism. *There's more to life,* I hear myself saying, *even if it isn't handed down by God.*

So, dear readers—and now I must invoke the plural as I turn alternately toward both horns of the stark dilemma—to those of you who take pride in your tough-minded willingness to stare into the nihilistic abyss and get out of bed anyway, I say: Perk up; you can have meaning and value without falling back on religion. And to those of faith I say: You don't *need* the fairy-tales told by most religions in order to justify your values.

These are, then, glad tidings . . . but fundamentally at odds with what just about everybody believes, whether secular materialists on the one hand or religious believers on the other. Radical! This third way is a road less traveled, *far* less traveled. Whether you come at it from the side of the secular or from the side of the sacred, you are almost

bound to confuse it with one or the other . . . because we lack a science of emergent systems. It simply doesn't exist.

Bits and pieces of a science of emergent systems are to be found in a number of places, and we'll go hunting for many of them: in the literature on the origins of life, in evolutionary theory, in the cognitive sciences, in complexity theory. Human experience is populated by several emergent phenomena that are famously difficult to get our minds around using materialistic reductionism. We have a difficult time with nation building because a nation is an emergent system and we're not very good at talking about emergent systems. Likewise happiness. What makes for happiness? What is a good teacher? Can you identify the necessary and sufficient conditions for good education? Each of these great big, very familiar, but famously elusive subjects continue to elude our understanding because we're not very good at talking about emergence and how it works.

Additionally the problem: *consciousness*. This is the brass ring after which countless thinkers are grasping today. They're coming from all directions: neuroscientists are studying the brain to find consciousness. Computer programmers are pursuing consciousness by way of artificial intelligence. Psychologists and philosophers bring conceptual analyses. There are many differing opinions and differing approaches. Almost all of these searchers and researchers can agree with the following proposition: Consciousness is an emergent property of the brain. They can agree so easily because none of them really knows what he or she means by the word, 'emergent.'

The strategy of this book is therefore as follows: Rather than mount a headlong assault on consciousness, money, love, or any number of emergent systems, we'll take a more circuitous route. What can we learn about emergence by first beating about the neighboring bushes in the literature on the origins of life, the literature on evolution, the literature on the origins of language. My hypothesis: If we can become fluent in the language of emergent systems by examining several of them and exploring the common traits of emergent systems in each of them, then consciousness, love, etc. will yield up some of their storied elusiveness.

The argument will not proceed, as Bacon and Descartes suggested we should in their discourses on method: from the clear and distinct, the simple and certain, then working step by step toward the conquest of the complex and unfamiliar. Instead we'll proceed from the abstract and obscure toward a re-appropriation of the concrete and familiar:

> Part One presents eight traits of emergent systems at a level of abstraction that might be thought of as *syntactical.* What is the grammar of emergence? What are its abstract rules, and how do those rules relate to one another?

Part Two turns from syntax to semantics. Part Two introduces content to the formal syntax of Part One. Part Two shows how the eight traits apply across several different domains: E.g., the origins of life, the story of evolution, and the origins of language.

Part Three is devoted to the nature of consciousness.

Part Four, playfully, is devoted to The Phallusy of Misplaced Physics.

Let me now offer brief statements of the eight traits of emergent systems and a preview of their applications.

Part One: The Eight Traits of Emergent Systems

First Trait: The impossibility of first instances

There could not have been a first word, for what made a barely articulate grunt into a word was the presence of other words constituting a rudimentary language. There had to be a language before there could be a word. But in that case, no single word could have been *first.* Likewise, there cannot have been a first thought to provide the origin of consciousness. The image capture of a camera is not a thought. Nor is the registration of light by the eye of a newt. At what point can we say that an information processing event counts as *a thought?* Not until that event occurs within the context of consciousness.

Second Trait: Emergent systems pop.

Because they come, all of a piece, emergent systems come about all of a sudden. In retrospect, their emergence has the appearance of discontinuity, however gradual and continuous the processes leading up to emergence may have been. The continuous emergence of discontinuity sounds paradoxical. Indeed, part of the reason we need a science or discourse of emergent systems is precisely because, without it, phenomena that exhibit emergence are likely to be dismissed as too weird, too paradoxical, to be real.

Third Trait: Holism—The Whole influences the nature of the part.

What was purportedly the first word could not be a word if it lacked the context of a language. A language had to be already there for a sound to be a word. Because the emergence of an emergent system requires the coming together of many pieces to create a system, it will seem as if that system came from nowhere. At one time there was no language. Then there was language. In the beginning, there was not one word, but many. Otherwise there could be no language. Prior to language, the sounds that became words were not words, but sounds. Only after many different sounds became correlated with many different objects and behaviors could it be said that those sounds acquired the meaningfulness associated with words and language.

The appearance of synchrony makes sense in light of this all or nothing presence or absence of language. Because there is no first instance, no gradual and incremental diachronic or sequential building of second word upon first, third upon second, fourth upon third . . . the emergence of language is synchronic.

On this reading, the phenomena of synchrony are not to be understood as some sort of miraculous causation at a distance—so-called synchronicity—as if event A in one place were causing event B to happen at the very same moment at some other place. Synchronicity, as discussed by C. G. Jung and others is an instance of what will be described as an instance of The Phallusy of Misplaced Physics. The concept of Synchronicity involves the attempt to read meaning into the chance coincidence of simultaneous events. Unlike synchronicity, synchrony should not be taken as some sort of miraculous violation of causality. Synchrony is instead a systemic, holistic feature of emergence that follows from the coming together of a set of complex interrelationships.

Fourth Trait: Emergent Systems are recursive.

Every emergent system achieves some form of closure, sometimes by way of a relatively simple feedback loop of the kind found between a thermostat and a furnace, sometimes by way of more complex forms of self-reference. A self-enclosed wholeness may be achieved by auto-poeisis according to Francisco Varela, (Francisco Varela, *Principles of Biological Autonomy,* North Holland Press, New York, 1979.) or auto-catalytic closure according to Stuart Kauffman of the Santa Fe Institute. (Kauffman, *At Home in the Universe,* Oxford University Press, New York, 1995; and *Investigations,* Oxford, 2000.) Doug Engelbart, the inventor of several of the components of the personal computer, talks about "boot-strapping." For philosopher Hans Georg Gadamer, the relationship between objective knowledge and subjective interpretation loops back in what he calls a "hermeneutic circle." Reflexivity is also central to the work of John Holland . (John Holland, *Hidden Order: How Adaptation Builds Complexity,* Addison-Wesley, Reading, Mass., 1995.) On both Kauffman and Holland and the work of their colleagues at the Santa Fe Institute, see Mitchell Waldrop, *Complexity: The emerging science at the edge of order and chaos,* Simon & Schuster, New York, 1993.

As this list of references is intended to show, recursion occurs at every level of emergent complexity: at the origin of life, language, consciousness, and cultural paradigms—what Michel Foucault called epistemes.

Fifth Trait: Emergent systems are unpredictable from the properties of their *component parts.*

Sixth Trait: Emergent systems are irreducible to the properties of their component parts.

These two linked and symmetrical traits of emergent systems are the most familiar features of emergence. In the past they have sometimes served to provide the primary definition of emergence. (See many sources, from the works of C. Lloyd Morgan to P. E. Meehl and Wilfrid Sellars, "The Concept of Emergence, in *Minnesota Studies in the*

Philosophy of Science, Vol 1, edited by Herbert Feigl and Michael Scriven, Minneapolis, 1956, pp. 239-252, to many essays in the journal, *Emergence.*) Apart from the context of the other traits of emergent systems, however, these frequently defining features of emergent systems have done little more than point to mystery. In the context of a science that seeks to explain things by placing them under covering laws that generate testable predictions, the invocation of unpredictability simply exiles those unpredictables from the domain of what counts as science. When confronted with the unpredictability of emergent systems, generations of scientists and philosophers of science have been left scratching their heads, or looking for "hidden variables" that would allow prediction after all.

Seventh Trait: Desire

Most systems operate most of the time under conditions that can be accurately described as seeking equilibrium, e.g., the normal satisfaction of demand by supply at a given price, or the satisfaction of ordinary hunger by nutrition. For such systems, contentment at equilibrium is the rule. For a new or higher or more complex system to emerge, desire must, first, upset the balance of contentment and, second, work to bind a new whole at a higher level.

The workings of desire are to be seen at every level, from that of a uni-cellular organism swimming upstream in a glucose gradient, to the works of love that defy the logic of mere utility. The hierarchy of desire includes Maslow's famous hierarchy of needs, but stretches further down into biology and further up into society and culture than Maslow's hierarchy, which applies only to the graduated hierarchy of needs in individual human beings.

Explication of the role of desire in consciousness, in particular, will show that consciousness does not result from the successive accretion of sufficient levels or feats of purely cognitive computation. Countless attempts in this direction have yielded only confusion at "the world knot," or frustration at "the last remaining mystery," (Cf. Dan Dennett, *Consciousness Explained,* Little Brown & Co., Boston, 1991, pp. 21-22.) or abject defeat at a problem deemed insoluble. (Cf. Colin McGinn, "Can we solve the mind-body problem?," *Mind,* XCVIII:891, 349-366.) Eugene O. Mills, "Giving Up on the Hard Problem of Consciousness," in *Explaining Consciousness—The Hard Problem,* edited by Jonathan Shear, MIT Press, Cambridge, Mass, 1998, pp. 109-115. Also David Chalmers, "Facing Up to the Problem of Consciousness," in Shear, *op. cit.,* pp. 9-30. We need a Copernican revolution that replaces cognition at the center of the solar system of consciousness with desire. If you follow both Hume and Spinoza in seeing cognition in the service of desire, rather than desire as a satellite of cognition, the hard problem of

consciousness, the world-knot, can be untangled. As neuro-scientist Antonio Damasio puts it, “We are feeling beings that think, not thinking beings that feel.”

Eighth Trait: Coming Apart

Whatsoever cometh together can also come apart. A critical feature of emergent systems is their fragility, their evanescence. As mysterious as life may appear to a science focused on parts rather than on wholes, death is equally mysterious: at one moment all the molecules making up Aunt Sally were parts of a living whole, then . . . next moment she was dead. Same molecules, but now they are no longer parts of a living whole.

Emergent systems can de-cohere. Love can end in divorce. Consciousness falls asleep. Wealth evaporates when a bull market goes bust. Whole societies collapse. See the record from Gibbon’s *Decline and Fall of the Roman Empire,* to Jared Diamond’s *Collapse,* to the sudden fall of the Soviet Union. Bubbles burst.

These eight traits of emergent systems come in pairs. The first two—no first instance and popping—are about beginnings; the last two are about endings, as desired outcome or dissolution. The third and fourth—holism and recursion—are about the nature of closure or self-containment that determines the *wholeness* of an emergent system. The fifth and sixth—unpredictability and irreducibility—relate to the context of *explanation,* answers to the question, *Why?*

Once we pin down the meaning of emergence by analyzing its traits in the emergence of life, evolution, and language, then we can use that understanding to unravel the mysteries of consciousness and love, money and creativity. There’s no end to the range of phenomena that light up when illuminated by the traits of emergent systems. With their application to each new realm, the traits of emergent systems are then further confirmed: From Michael Porter’s descriptions of the emergence of industry clusters as motors of economic growth, to the nature of humor, jokes, and the neuro-anatomy of laughter; from the magic of music to the gift of morality.

To anticipate the way elusive ideas can be caught in the net of the eight traits of emergent systems, consider as an anticipatory example *happiness,* a state of being that is highly valued but famously difficult to define, much less achieve. Note how neatly the eight traits fit some of the intuitively familiar features of happiness:

1. **No First Instance:** Happiness has no recognizable onset, no first instance. You’ll have to have been happy for some time before you can be called or consider yourself happy.

2. Emergent systems pop: Despite the fact that happiness pervades a whole life, the experience of happiness erupts. It seizes you when you least expect it. It comes over you all of a sudden. It suffuses a moment.

3. The Whole influences the nature of the parts: One's whole life must be happy, not just part of it. You can't be happy at home and unhappy at work, or vice versa. Happiness is something that qualifies the whole of life, not in the sense of beginning-to-end, but in the sense of all-of-the-parts-of-life-together-in-a-whole. This is why Aristotle said that you cannot really say of someone that he is happy until after he's dead.

4. Emergent Systems are recursive: Happiness declares itself. You have to *know that* you're happy in order to really *be* happy.

5. Unpredictability: Happiness is famously unpredictable. It cannot be predicted reliably on the basis of income. Happiness can't be predicted on the basis of location. Nor is happiness predictable as lying on the other side of some hurdle or other, e.g., gradualtion from school, or marriage, or the attainment of tenure. Somewhere it is written, "Give up the goal of happiness. That sweet bird lights only when least expected." (James Ogilvy, *Living Without A Goal,* Doubleday Currency, New York, 1995.)

6. Irreducibility: Happiness is not reducible to any one of its components, not *love alone,* not *work alone,* nor any other component that might be proposed, quite rightly, as essential to happiness. This is why Freud, when asked what was necessary for happiness, replied, "*Lieben und Arbeiten"*—love *and* work.

7. Desire: Happiness has a lot to do with the fulfillment of desire, not, *pace* Buddhists, with its cessation, desirelessness. There is a kind of desirelessness that is consistent with what is being said here, namely that if you desire happiness too directly, with a fervor and focus of intent, then it will surely elude you. But if you put aside your lust for happiness, if you relax your longing, then she is much more likely to come 'round. But never guaranteed. Happiness emerges only when it all comes together.

8. Coming Apart: Happiness can end abruptly in tragedy, as with the loss of a loved one. Just as happiness is an emergent system, so too is its opposite, depression. No one thing causes happiness; no one thing causes depression. Just as coming apart is the flip side of emergence, so depression is the flip side of happiness.

Both the explication and confirmation of the traits of emergent systems are necessarily systemic. The whole is greater than the sum of the parts. Taken one by one as puzzles or paradoxes in the context of ordinary science as axiomatized and understood by codifiers like Hempel and Oppenheim, Ernest Nagel, Rudolf Carnap, or Wilfrid Sellars, each of these traits is not a trait at all, but a puzzle, a paradox, an anomaly or a mystery. Consider several traits in turn if taken out of the context of a science of emergent systems:

Take the first trait, the paradox of the impossibility of first instances. Surely, one can assert in the context of ordinary science and an ordinary understanding of ordinal systems, that if a more complex system evolves or emerges from a less complex system, whether we're talking about the origin of life or the origin of language or the origin of consciousness or the origin of a socio-cultural *episteme* (as Foucault calls a paradigm) . . . then in each case, given what we take for granted about the linearity of time, for each instance of emergence, *there must have been a first instance.* At one time there were no languages. Then there was language. There must have been a first word! How can you get a language of 500 words without a first word? If we can add a new word to a language, say, the 501st, then doesn't it make sense to imagine that the 500th word was added to a language of 499 words? And if we can work backwards from 500 to 499, as well as forwards to 501, doesn't it make sense to work back all the way to the first word? *There must have been a first* according to the linear logic of common sense and its "long arm," ordinary science. (For the idea that science is the long arm of common sense, see Gustav Bergmann, *Meaning and Existence,* The University of Wisconsin Press, Madison, 1960.)

Taken alone, the First Trait of Emergent Systems, the impossibility of first instances, makes no sense.

Take the fourth trait, the recursivity of emergent systems. This feature of emergent systems runs afoul of Bertrand Russell's theory of types according to which, "no function can have among its values anything which presupposes the function." (Bertrand Russell and A. N. Whitehead, *Principia Mathematica,* Cambridge: Cambridge University Press, paperback edition, 1962 pp. 39 and 56. See also Charles S. Chihara, "Russell's Theory of Types, in *Bertrand Russell: A Collection of Critical Essays,"* ed. D. F. Pears, Garden City, N.Y., Doubleday Anchor, 1972, pp. 245 ff.) Russell invented his theory of types in order to avoid paradoxes of the sort that ensue when you think through the truth or falsity of the proposition, "All Cretins are liars and I am a Cretin." The theory of types sets out a hierarchy of types such that propositions on a higher level of the hierarchy can refer to classes of propositions on lower levels, but no proposition can refer, self-reflexively, to a class on the same level, of the same "type," of which it is a member. All well and good when one is trying to construct an internally consistent logical hierarchy for the

foundations of mathematics, but Russell's injunction against self-reference as "vicious circularity" goes too far if taken as an injunction against all recursive self-reference. (Cf Douglas Hofstedter's marvelous discussion of Russell's unreasonable attempt to ban self-reference in the name of reason in *I Am a Strange Loop,* Basic Books, 2007.)

Take the fifth and sixth traits, the impossibility of prediction and reduction. But prediction and reduction are the very engines of explanation. If those engines go idle—as the principles of unpredictability and irreducibility would seem to entail—then explanations must cease. It would seem that the elimination of these traits reduce science to a dumb stupor.

Take the seventh trait, desire. Taken by itself, at any level of complexity, the invocation of desire looks like blatant anthropomorphism. People *want* things, but things can't *want.* Things just respond to causes, efficient causes in Aristotle's lexicon: pushes, not pulls. To think otherwise is to be guilty of teleological reasoning, putting the effect before the cause, the end before the means, the cart before the horse. As all rational theorists have known since Spinoza and the eighteenth century Enlightenment, this sort of teleological reasoning is simply mistaken. This is not how the world works. Causes precede effects. There is no final cause, no purpose to the universe. Things do not happen "for the best."

Eighth Trait: Happiness can end very quickly. Imagine a happy person who suddenly suffers the loss of a loved one. Misery is always hovering to threaten happiness. But to consider happiness out of the context of emergence is to imagine happiness as eternal.

So to summarize the points made thus far: each of the eight traits of emergent systems, taken singly and by itself, makes little sense in the context of the science we have so laboriously built since the Eighteenth Century. In what follows we'll see how each of these traits makes a great deal of sense when all eight are taken together as mutually coherent parts of a science of emergent systems. The argument of this book is emergent. It takes all eight traits to constitute this revolutionary account of emergence. This book is radical, as a friend has put it on reading the manuscript, "downright revolutionary."

A Note on method

In keeping with the *content* to be described, the *form* of the description suffers under the constraint of the third trait, that the whole must be present before any part makes sense. The science of emergent systems cannot be described as an axiomatic unfolding from first principles. It emerges, instead, from the coherence of all of its pieces taken together. Given the inevitably linear form of reading and writing, the best way to start is to grasp as

much of the whole as possible in the form of a single (if incomplete) table that will provide a kind of architecture of the whole.

After addressing the whole in tabular outline, we'll return to a more detailed exposition of each of the parts by traversing first the columns (the traits in their generality) and then the rows (each of the emergent levels in their specificity). Such an exposition will dispel, little by little, any impression that emergent systems rest on magic. Those who traverse the following argument, little by little, step by step, will eventually return to the following table and will then see, all of a piece, the coherence of a system that explains in greater detail why each of the pieces, taken apart from the completed system, should have justly appeared mysterious and inexplicable from the blinkered perspective of Enlightenment science and rationality.

Nota bene — note well—no one is claiming that science or rationality as we have come to know it is crudely *wrong*. Current science is simply incomplete. The phrase 'current science' covers a lot of ground. Does it make sense to make such a sweeping claim as that *current science is incomplete?*

Monological Science

In order to be clear about what is meant by 'current science,' or 'Enlightenment science and rationality,' it will be helpful to establish a clear foil for the science of emergent systems. Let the phrase, 'monological science,' stand for that kind of scientific reasoning that is not so much wrong as incomplete. The term 'monological' has been chosen to suggest a scientific rationality that over-emphasizes singleness, unity or unification in three different respects:

> **Formal unity**: the view that to understand or explain something is to subsume it under some singular, Platonic ideal; the subsumption of many particulars under one universal; the subsumption of several different explanatory systems under one covering law, e.g. the explanation of rusting, respiration and combustion by their subsumption under the single concept of oxidation; the further extrapolation from such limited unifications toward the demand for a unified field theory for the explanation of everything, as if, for the universe to be rational, there *must* be some single peak to the hierarchical pyramid of monological abstractions; a Platonic Form of forms as deserving of rational respect as the Lord of lords is deserving of monotheistic reverence.
>
> **Substantive unity:** a metaphysics that privileges things over structures, differences, fields, or relationships; the long tradition since Aristotle who affirmed

that to be is to be an individual; atomistic materialism; the preference for interpreting Bohr's complementarity principle in terms of particles rather than waves. In cases of emergence, we rrevert from an atomistic metaphysics to a *relational* metaphysics.

Explanatory monism: the view that to explain something is to find a single cause, e.g., the mosquito as the cause of the spread of malaria; the plot structure of the whodunit; the quest for the "smoking gun," or "silver bullet;" the germ theory of disease; an understanding of causality that takes the combination shot on the pool table as a model for mono-linear sequences of causes and effects—cue stick strikes cue ball which then drives the 3-ball into the 7-ball in order to nudge the 9-ball into the side pocket, all in a single string of mono-linear causes and effects. Explanatory monism imagines that every effect can be explained by such a mono-linear sequence.

To appreciate the mindset of monological science, it's important to see all three ways that unity is privileged over multiplicity. The valorization of unity has its epistemological (formal unity), ontological (substantive unity), and nomological (explanatory) correlates. The world of monological science is a world of atoms in motion according to laws arranged in a hypothetico-deductive hierarchy under a unified field theory. This worldview has many variants running from the Greek atomists, Lucretius and Democritus, through the science of Laplace to the logical positivism of the early Wittgenstein, A.J.Ayer, and Hans Reichenbach. (Cf. Hans Reichenbach, *The Rise of Scientific Philosophy*, Berkeley, University of California Press, 1959, for a good statement of the philosophy of monological science.)

It is worth emphasizing once again that however many revisions may need to be made to particular parts of this worldview, e.g., Einstein's advances over Newton, the worldview of monological science retains tremendous explanatory power. Monological science cured polio and put men on the moon. It's not simply wrong.

But monological science is inadequate to the task of explaining, on its own terms, emergent systems like life, language, consciousness or creativity. Each of these *systems* is unpredictable from its evolutionary precursors, and is irreducible to them.

Consider as just one illustrative example of such unaccountability the following passage from the autobiography of philosopher of science, Paul Feyerabend:

> . . . a moral character cannot be created by argument, "education," or an act of will. It cannot be created by any kind of planned action, whether scientific, political, moral, or religious. Like true love, it is a gift, not an achievement. It depends on accidents such as parental affection, some kind of stability, friendship,

> and—following therefrom—on a delicate balance between self-confidence and a concern for others. We can create the conditions that favor the balance; we cannot create the balance itself. Guilt, responsibility, obligation—these ideas make sense when the balance is given. They are empty words, even obstacles, when it is lacking. (P. K. Feyerabend, *Killing Time,* The University of Chicago Press, Chicago, 1995, p. 174.)

Unlike Feyerabend who boldly invokes the gift of love, some early readers of this book have advised against dwelling on love: Too warm and fuzzy; too unscientific; too close to the kind of folk psychology that a rigorous philosophy of mind or a careful cognitive science should avoid. But that is precisely the point: Precisely to the extent that love escapes easy definition, precisely to the extent that it is *not* one simple thing but a congeries of culturally mediated romance and biological urge, for that very reason it provides both an illustration and a test-bed for a discourse on emergent systems.

Love is not simple. Neither is it complex in the way a computer is complex. Love comes to be and passes away. Much must come together for love to emerge—the physical, the mental, the emotional and, though sometimes in retreat, the rational. Love is an open system that is also dynamic, and therefore amenable to systems dynamics of the sort that R. D. Laing picked up from Gregory Bateson and the existentialists. Love has its feedback loops, its own particular cybernetics. Love must declare itself in poems and songs, in oaths and troths.

There are those who would claim that love is an illusion. They would reduce it to evolutionary biology's trick to propagate the species. Others might grant its reality only to reduce its *real* nature to some underlying plumbing of libido and sublimation. Still others will grant love's undeniable prominence in the course of human affairs, but will advise prudence and caution before tackling so vast a subject. Yet again, that is precisely the point. Love serves the argument of this book not because it appeals to the romantics among us; love serves the argument of this book because, by its very bulk, it demonstrates the coming together of many component parts.

Love calls out for a language of emergent systems because other approaches, with the possible exceptions of songs and verses that count the ways, famously fail to get their arms around love's elusiveness. And if a fluency in the language of emergent systems *can* encompass the otherwise imponderable vastness of love, then perhaps such a language can encompass other imponderables as well, like life, language, wealth, beauty . . . and depression. The last is important as a further illustration of the strategy for approaching love: not so much because it is devoutly to be wished, but rather because it is so hard to understand using the ordinary tools of analysis and monological explanation. If the language of emergent systems can shed light on love, then it may just be able to illuminate mysteries of life, consciousness, and language as well. A language that can

illuminate love's dark corners may also be able to shed light on money and artistic creativity.

At the end of the day, we can have a science that accounts for both the ordinary and the extraordinary, the normal workings of systems seeking equilibrium, and the eruptions of new order we call emergent systems. We can maintain both monological science and a science of emergent systems in a kind of stereoscopic vision of our external and internal environments.

Monological science alone is blinkered. Like the single eye of Cyclops, it will not give us the depth perception we need to see a world in which consciousness is more than a behavioral epiphenomenon. In order to see a world that has the depth that consciousness introduces, we need a "second eye," a second epistemology, a way of seeing that allows us to account for the emergent systems we call life, language, and wealth. Sticking to the finality or fundamentality of monological thinking is the epistemological equivalent of insisting on monaural sound—like the Luddites who resisted stereo saying, "It's the *same music.* It's *Beethoven's Ninth* whether you send it through two speakers or one.*"* But listen to it! Use your senses! Yes, both experiences are based on the same score. But what a thin abstraction the monaural version represents when compared with the higher resolution and added dimensionality of the stereo experience!

So, likewise, the *identity theorists* in the philosophy of mind—the computationalists, the eliminative materialists—would have us believe that the score is all that matters; that human consciousness can be represented as the running of software on the wetware of the brain. The what-it-is-to-be-conscious, the *qualia,* the *experience of* consciousness is regarded as an epiphenomenon that is mis-described by "folk psychology" with words that will eventually be dropped from our language, like 'phlogiston' and 'entelechy'. But they are wrong.

They are as wrong as the Luddites who wanted to hold on to monaural by *refusing* to buy that second speaker, by *not* buying a new record changer with a more sensitive arm and a completely different—and fairly expensive—stylus, to say nothing of a new amplifier that could handle not one but two channels of sound.

Literal stereoscopic vision—with two eyes—is not about doubling the vision of one eye any more than stereophonic sound simply reproduced mono in a different part of the room. Nor, however, does stereoscopic vision show you some kind of *other world* with properties altogether different from the first, e.g. "spiritual" as opposed to "material." Like stereophonic sound, stereoscopic vision adds a depth and dimensionality to experience by "splitting" the visuals in such a way that it would be a mistake to say you're getting *the same* visuals from different perspectives. You aren't. Even though you're looking at *the same* objects.

Monological science is not wrong. It is simply limited. It gives a relatively thin representation compared to the richness and depth of reality as represented by stereoscopic vision. This relatively thin representation can seem incredibly rich, as did Toscanini's monaural recordings of Beethoven with the Philadelphia Symphony Orchestra. But wait till you hear Beethoven in stereo! You had to hear it to believe it. And now you have to "see" it to believe it. But now, to the extent that it is not precisely sight that we are talking about, but only using sight vs. sound as an analogy for classical-plus-emergentist science vs. monological epistemology . . . it's equally true that you have to believe it in order to "see" it. You have to *get* what I mean by stereoscopic "vision" before you can grasp the added dimensionality that it offers.

The strategy of exposition is therefore to examine the application of emergence in a number of different realms, e.g. the nature and origins of life; the nature and origins of language. These are not trivial subjects. They have their difficulties. But their difficulties are as nothing compared to the difficulties of understanding consciousness or love.

If we can achieve stereoscopic vision in these other realms, then when we return to the mysteries of wealth and creativity, we stand a better chance of unlocking their mysteries. If we keep looking at them through the Cyclopean eye of monological science, we will get nowhere. The concept of emergence is the key to unlocking that large and imposing door to subjectivity, the experience of awareness, the very nature of what it is to be conscious being, a *me*, an *I,* a *human self* .

As the following chart is meant to show, each trait applies in slightly different ways with different language at each level of emergence. The power of the chart, the power of the whole account showing the application of all traits on all levels, lies in the way different levels illuminate one another. Where there has been careful and rigorous thinking related to one trait on one level, what we learn there can suggest directions for the clarification of what appear to be mysterious on other levels.

	No First Instance	Emergent systems pop	Holism	Recursion	Unpredict-ability	Irreduci-bility	Desire	Coming Apart
Origins of Life	No first Germ cell	Birth	Life of whole influences life of part	Autopoiesis; Auto-catalytic sets	Organic life unpredictable from inorganic elements	Life not reducible to presence of *élan vital*	Tropisms; Hunger	Death
Evolution	One mutant does not a species make	Cambrian explosion; punctuated equilibria	Niche as important as genes; Evo-devo	Re-production true to species	Sporting; Mutation; Randomness of variation	Phenotypic expression not reducible to genotype	Evolution from mitosis to sexual re-production	Extinction
Origins of Language	No first word; no simple language	Symbolic reference does not accrue from indexical	Lattice work of language; synchronic structure	Symbols as words about words	Intention can't be reduced to reference or denotation	De Saussure's arbitrarines s of the sign	Deacon's deduction of language from com-mitment	Devolution from signal to noise
Love	Love at first sight as exceptional	"Falling" in love; orgasm	Love colors everything	Love must declare itself	Why arranged marriages don't work	Can't say, "I love her because…"	Role of Eros/libido in love	Falling out of love; divorce
Conscious-ness	No first idea	Awakening	"Arche-writing" – Derrida; Grammar, Syntax	Re-entry – Edelman; Self-Cs. as conditioning Cs. - Sartre	Creativity; subjective intention; free will	Failure of eliminative materialism	Hegel's deduction of desire	Falling asleep; Death
States, Politics	No first citizen	Independ-ence; Inaugura-tion	Subsidiar-ity	Constitution; "We the people . . . "	Role of freedom, intention and volition	Critique of Realism and Neo-realism	Patriotism	Failed states; collapse of Soviet Union
Wealth, Markets	Takes two to trade	Take-off; How Hits Happen	Porter's industrial clusters	Brian Arthur's positive returns	If it's predictable, it's not a market.	Austrian subjectivist economics	Consumer demand; "consump-tivity"	Bursting bubbles; bear markets; depression

The Periodic Table of Elements of Emergent Systems Each of the emergent systems in the left-hand column exhibits all eight traits. Other emergent systems—humor, morality, happiness, depression, beauty, leadership, a good education, governance—all cry out for similar analysis.

This first part follows the columns in Table 1. The meaning of each of the eight traits will become clear from the application of each trait to several different levels of emergent systems. Equipped with a better understanding of each of the eight traits, Part Two takes a closer look at the specifics of several levels by showing how the traits interrelate to one another at each level.

Trait One: The impossibility of first instances

The intuition driving both the first and third traits is the notion that, for emergent systems, no part of the system can be any part of the system— first, last or otherwise— until the whole system exists as a system. And given the availability of a system in all of its complexity, no single part can be described as first. How is this possible? How can such systems come to be if there is no first part of that coming to be?

Rather than starting with the simplest cases, as we might be inclined to do if we were operating within the paradigm of Enlightenment rationality, here it will be both easier and more informative to begin with the most complex cases and then reason backwards. So let us begin with the example introduced in the quotation from Feyerabend—morality. Here there are noble precedents for making the case we want to make, all the way from Plato's doctrine of the unity of the virtues to Alasdair MacIntyre's treatment of the virtues in *After Virtue.* Both Plato and MacIntyre make a case for understanding virtue as embedded in a context of moral practice. A single act of bravery in battle may not count as bravery unless it occurs in the context of other such acts and other virtues like temperance and wisdom. Apart from a measure of wisdom and prudence, a soldier's bold actions may be better described as foolish bravado or the wild act of a deranged man. (See Plato's dialogue, *Laches.* Cf. MacIntyre, *After Virtue,* University of Notre Dame Press, Notre Dame, Indiana, 1981.)

Granted, it may be difficult to imagine how virtue can come to be if there can be no first act of virtue. But the problem lies in the lenses we are using to look at this phenomenon, not with the phenomenon itself. We *expect* to be able to find a first instance, because that is the way non-emergent phenomena come to be. Think of the building of a house that begins with a foundation, then framing, etc. When our expectations aren't met, we become skeptical, as does Socrates in many of Plato's dialogues. After taking the virtues one by one and trying to nail down tight definitions for them, most of the early dialogues end skeptically, at least on the surface or explicit level.

In the *Meno,* for example, Socrates begins by asking whether virtue can be taught. After quizzing Meno fairly rigorously, the interlocutors conclude, much like Feyerabend, that virtue is a gift, and that it cannot be taught or learned. Yet close readers of Plato's Socratically ironic texts find grounds for a less skeptical reading. If we examine a series of fairly subtle hints dropped on the dramatic level of the dialogue that runs alongside and beneath its logical argument, we notice that Meno, who began the dialogue as a fairly brash and arrogant young man, is gradually, little by little, step by step, tamed by Socrates. By the end of the dialogue he gives the impression of being more virtuous than when the dialogue began. He has been taught virtue. (I owe this reading to Robert

Brumbaugh. But see also Jacob Klein, *A Commentary on Plato's Meno,* The University of North Carolina Press, Chapel Hill, 1965.)

How are we meant to respond to Plato's subtlety? Precisely as Feyerabend does, namely, by granting that virtue and morality *are* indeed available to us, but that facile and glib claims on the part of sophists or religious leaders that they can teach or bestow morality in three quick steps or ten commandments deserve our profound skepticism. Aristotle's teachings are perfectly in accord with this conclusion. In his *Nicomachean Ethics* (Book II, Chap. 1) we are told that "moral virtue comes about as a result of habit, whence also its name *ethike* is one that is formed by a slight variation from the word *ethos* (habit)." (1103a16-18)

At the level of complex human behavior, it's easy enough to solve the apparent paradox of the impossibility of first instances: We're all familiar with activities that can be learned by, first, going through the motions in order to acquire certain habits, and only later can we claim mastery. Fake it till you make it. First we walk the walk, then we talk the talk, and only after engaging in a sustained practice of virtuous behavior can others rightfully ascribe virtue to us (for we would be guilty of the vice of pride or arrogance to claim on our own behalf to have mastered any virtue).

Having familiar examples like the preceding can help us to achieve the imaginative reach we need to make sense of other domains of emergence, like the emergence of language, which will be forever shrouded from our curious gaze by the mists of pre-history. Though every parent gets to witness the re-birth of language in infants and toddlers, it's too easy to claim that each new speaker has the benefit of listening to others who already know the language. What about *the first* language? Was there ever a *first word?*

A logico-analytic approach would lead us to expect that there was, as if some unsung *homo sapiens* must have somehow invented the word for *rock* or *fire* or *wooly mammoth*, and from such humble beginnings, built his or her vocabulary word by word, then syntactical rule by syntactical rule, until finally there was something that counted as a language. But this is absurd. No sounds are words, strictly speaking, until they function in the context of linguistic communication. Not only could there not have been a first word, there could not have been a first speaker. It takes at least two for communication to occur.

Reasoning backwards from the complex case of morality (rather than forwards from some putative first word or first speaker), we have to imagine some analog to habit, something like what Derrida calls an "arche-writing" that precedes speech. We have to imagine a time when people simply played with different sounds, and gradually built up certain associations between certain sounds and certain things or acts. Only when there was a sufficiently rich variety of sounds, a sufficiently large stock of *differences* among different sounds, was it possible to have something like a language, and therefore words.

Second Trait: Emergent systems pop

Consider first as a list, and then, one by one, the following candidates for description in terms of emergence. The point of working our way through this list of seemingly unrelated phenomena is to *show* (not just say) how many different phenomena pop. In showing each item on the list, and the length of the list, we achieve an existence proof for the reality of emergence. Apart from some such existence proof, the die-hard rationalist is likely to doubt that any single item on the list is a genuine instance of emergence. Any particular instance, taken by itself, will be subject to explanation in terms of monological science. If each particular instance is so reducible to monological explanation, then who needs a science of emergent systems? But if we take the list as a *whole,* then that divide and conquer strategy starts to look like the addition of epicycles to the old Ptolemaic model of the solar system. As Kuhn showed, you *could* hold on to the Ptolemaic model of the solar system by constructing elaborate epicycles for each of the planets. But if you simply replaced the earth with the sun at the center of the planetary system, the whole picture suddenly became a lot simpler and more elegant. So likewise with the following list: If we see each of them as an emergent system popping rather than reducing each item to some elaborate monological explanation, the whole system becomes a lot simpler and more elegant:

(a) Paradigm shifts – T. S. Kuhn
(b) Incommensurability – Feyerabend and Scriven
(c) Epistemes – Foucault
(d) Tipping points – Gladwell
(e) Quantum mechanics, Fission/Fusion (critical mass)
(f) Punctuated equilibrium – Eldredge and Gould
(g) Substantial vs. accidental change – Aristotle
(h) Catastrophic Morphogenesis – Rene Thom
(i) Enantiodromia – C. G. Jung
(j) *Aufhebung* (dialectical synthesis) – Hegel
(k) Synchrony
(l) Singularity – V. Vinge, Kurzweil
(m) Peak experience – Maslow
(n) Beauty
(o) Orgasm

These phenomena occur. Granted, many of the items on this list derive from contexts that already contain a theoretical commitment to something like emergence. E.g., Rene Thom's idea of catastrophic morphogenesis is part of a richly mathematized and theory-

laden description of bifurcations in biology. He describes the genesis of new forms by processes of branching; he offers a mathematical model for the rapid transition from fight to flight. While the phenomena he illuminates through his theorizing are very real and very observable, we can hardly claim that 'catastrophic morphogenesis' is itself part of an observation language. (The putative distinction between an observation language and a theory language was one of the cornerstones of logical positivism. Simple and indubitable sense data were to be recorded in an observation language that described atomic facts of location, temperature, etc., while a theory language was built with more complex inferences about relationships and generalizations. (For a classic and very accessible statement of this approach, see Henry Margenau, *The Nature of Physical Reality,* New York, McGraw-Hill, 1950, esp. Chapters 3-5, pp. 33-101.) For a critique of that classic view that muddies the distinction between observation language and theory language by showing how "theory-loaded" or theory-laden observation terms inevitably are, see Norwood Russell Hanson, *Patterns of Discovery,* Cambridge, Cambridge University Press, 1965, esp. pp. 54 ff.) It is part of a theory language that presupposes a theory that already countenances the kinds of discontinuities that the monological rationalist is inclined to reject. Basing an argument for the existence of emergence on any one of these theory-laden phenomena would, thus, beg the question. It would assume what is to be proven.

'Nature abhors a vacuum.' 'Nature makes no leaps.' 'The principle of plenitude asserts that the Great Chain of Being has no missing links.' With such axioms and postulates the die-hard rationalist will defend the so-called *principle of sufficient reason.* In the face of a farrago of such mutually supportive postulates, any one of the items on the above list is vulnerable to refutation and the existence proof will fail. In the face of *all* of the items on the above list, however, even the die-hard rationalist might have to pause and consider: *Something* is escaping the reach of rationalistic prediction and reduction. *What could it be?* Any one of these phenomena on the above list might be swept aside as a singular anomaly. Any one alone is not enough to disconfirm the universal reach of monological reason. But when a whole pack of anomalies crop up . . . then we have reason to suspect that we're looking at a paradigm in trouble and a paradigm shift in the making, and with that reflection, we're off and running toward an inductive inquiry toward new generalizations, not a deductive explanation from old covering laws.

The following discussions of different items on the above list will *not* try to show that most of the items *are really nothing but* instances of any one item on the list. We needn't reduce the many to the one. We needn't *explain* all of the items on the list by showing how they can be understood as particular instances of one general principle. Instead, we can approach each item phenomenologically —with a minimum of presuppositions, not (impossibly) no presuppositions at all—and then generalize toward a second trait of emergent systems. *Emergent systems pop!*

Let's begin this inductive inquiry precisely where we are: With the observation that the list of candidates for emergence is long enough to qualify for the kind of treatment that T.S.Kuhn gives to the epicycles in the Ptolemaic model of the heavens. (T.S.Kuhn, *The Structure of Scientific Revolutions,* University of Chicago Press, 1962.) If there had been one or only two heavenly bodies whose paths across the heavens resisted mapping in earth-centric orbits, then the Ptolemaic model might have prevailed in the face of the Copernican revolution. We might have maintained the conceit of living at the center of the universe; we might have continued to believe the grand illusion that the world revolved around us. But the anomalies were too many. The old Ptolemaic paradigm had to fall, and not just one planet at a time, little by little, but catastrophically, all at once.

(a) Paradigm shifts – T. S. Kuhn

What Kuhn so famously achieved was a revolution in our thinking about scientific revolutions. Prior to Kuhn, the generally accepted view held that science proceeded step by step, incrementally, such that each new theory would mark an incremental advance over the old in giving us a progressively more accurate description of our environment. According to this older view, Einstein's relativity theory does not prove Newtonian mechanics wrong. Rather, Newtonian mechanics remains true as a special case, under specified boundary conditions, e.g. the order of magnitude of a baseball diamond. Newtonian mechanics are just fine for mapping the trajectory of a well hit baseball. Relativity theory will not help even the best outfielder shag any more fly balls. Indeed, with the proper specification of boundary conditions, Newtonian mechanics could be deduced from relativity theory—according to pre-Kuhnian, incremental history and positivistic philosophy of science.

Kuhn—with a little help from Norwood Russell Hanson, Paul Feyerabend, Michael Scriven, and Wilfrid Sellars—changed all that. Kuhn distinguished between *ordinary science* as incremental puzzle solving, and *revolutionary science* as a discontinuous breaking of the mold such that the new theory does not just re-describe the same old things in slightly different ways; the new theory, if it really *is* a piece of revolutionary rather than ordinary science, shows us a world that changes our understanding about the very things to be explained. It doesn't just re-describe the same old things; it shows us a world that contains new and different things. E.g., the world is no longer divided between sub-lunary (under-the-moon=earth) things and super-lunary (over-the-moon=all the stars, the planets and the sun) things. The new theory gives us a fundamentally new understanding of what a planet *is,* and now the earth joins the other planets as one more satellite of the sun.

When Lavoisier discovered oxygen, not only did he subsume under the new theory of oxygen chemistry the disparate phenomena of rusting, combustion and respiration; he eliminated from the scientific lexicon and from our inventory of the universe an old thing that *had* served as an explainer: phlogiston.

Now there's a particular aspect of paradigms and paradigm shifts that calls for attention, and serves as an excuse for why some readers will find this book hard going. Paradigms are *trans-disciplinary.* A single discovery in a single discipline does not a paradigm shift make. Paradigms cover all the disciplines at once. Because the involve our ways of seeing and knowing, paradigm shifts change our ways of seeing and knowing everything, so we have to explore a lot of disciplines to describe paradigm shifts.

Aristotelian science with its final causes and *entelechies* affected his physics, his biology, his cosmology, everything. The the paradigm shift to modern reductionist materialism affected the way we see *everything.* Now, in order to explore the paradigm shift to emergent systems, we have to delve into many disciplines: biology, evolution, physics, economics, romance literature, and the emergence of artistic creativity. This is no small order, and many a reader is likely to feel lost in certain unfamiliar realms. But be not afraid. In the end, *it will all come together.*

I cannot remember the number of times I found myself pacing around at our company, Global Business Network, as a complex project was working its way to a conclusion. "*It's all coming together,"* I would say, smacking one fist into the other palm. And I don't mind the sexual connotation of 'coming together.' Love and orgasm are two of our best examples. But the summons to traverse many different disciplines is daunting. But that's how paradigms work: if they're not trans-disciplinary, they're not paradigms. When Peter Schwartz and I wrote the SRI (formerly Stanford Research Institute) report, *The Emergent Paradigm: Changing Patterns of thought and Belief*, we covered no less than thirteen different disciplines. He took the hard sciences, I took the humanities. But you, dear reader, whether you are a scientist or a humanist, are going to occasionally find yourself in strange waters in what follows. Hold on. Persevere. *It will all come together.*

(b) Incommensurability – Feyerabend and Scriven

It may be true that new theories help us to explain why old theories lasted as long as they did, and in this sense it was tempting to believe that older theories could be deduced, under suitable boundary conditions, from newer theories. If so, then a certain incremental continuity could be preserved. The same old stuff was just being re-described. But Feyerabend and Scriven show quite convincingly that the observation terms of the old theory are "incommensurable" with the observation terms of the new theory. 'Space'

does not mean the same thing in Newtonian science and Einsteinian science. 'Atom' does not mean the same thing to Democritus and to Neils Bohr. Even 'temperature' has a different meaning before and after Boltzmann's reduction of thermodynamics to statistical mechanics. As Norwood Russell Hanson was fond of putting it, our so-called primitive observation terms are always "theory-laden." Change the theory, not by incremental extension but by revolutionary paradigm shift, and you change the meaning of the observation terms. You're not really talking about *the same things* any more. After Lavoisier, we don't re-describe phlogiston. You can't measure how much phlogiston is contained in a cubic foot of oxygen at 70 degrees Celsius. Phlogiston and oxygen are *incommensurable.*

True, you can use a new theory to explain why the old theory was as accurate as it was under certain boundary conditions, but you cannot, strictly speaking, *deduce* the old theory from the new theory, or *reduce* the terms in the object language of the new theory to descriptions in terms of the old theory. Nor can you predict the behaviors described in the new theory by simply adding some new initial conditions or new boundary conditions to the old theory. Instead, to quote Wilfrid Sellars:

> Theories about observable things *do not 'explain' empirical traits in the manner described* [namely by deducing the earlier theories from the later theories], *they explain empirical traits by explaining why observable things obey to the extent that they do, these empirical traits.* (Wilfrid Sellars, "The Language of Theories," in *Science, Perception and Reality,* New York, The Humanities Press, p. 121.)

All of this talk of non-deducibility and incommensurability was offensive to the old school of Enlightenment rationalists and logical positivists. There seemed no logical path from here to there, from old theory to new theory, predictively, or from new to old deductively or retrodictively. And this is why Kuhn concluded that, for a new paradigm to really take hold, the holders of the old paradigm would simply have to die off—the ultimate discontinuity. And why economists say that the discipline of economics advances one funeral at a time.

While there's much more that could be gleaned from Kuhn, Feyerabend, Scriven and Hanson, let's move on to other items on our list of anomalies. Kuhn makes for a good beginning for several reasons: First, because he provides us with some of the language we need to do the rest of the work, e.g., the very idea of "anomalies," and the need for a "paradigm shift" from reductionistic explanation to a paradigm that includes emergent systems that are explicable even though they are not reducible to their precursors. Second, Kuhn's clarifiers (Feyerabend, Scriven, *et. al*.) offer a degree of rigor and precision that is lacking in some of the accounts of other items on the list of anomalies. It

pays to start with examples that are, if not certain and irrefutable, at least clear and distinct.

Kuhn *et. al.* have not spoken the last word on the *current* scientific revolution from a hypothetico-deductive science of the sort described by Hempel and Oppenheim to a stereoscopic science that *also* includes emergent systems. Indeed, as solid and admirable as Kuhnian history and philosophy of science may be, stunningly missing in Kuhn and his heirs is a recognition of the degree to which their achievements are *not the first instance* of a new paradigm, but part and parcel of the whole range of anomalies listed above.

(c) Epistemes – Foucault

Perhaps it is just a matter of Anglo-American provincialism, the degree to which most English and American philosophy departments have taught philosophy for the last half century as if *real* philosophy began with Ludwig Wittgenstein and Bertrand Russell and everything before could be swept away and forgotten . . . but it is remarkable how few contemporaries have noticed the "family resemblances" between Kuhnian thinking and the tradition from Kant and Hegel to Michel Foucault. (See Wittgenstein, *Philosophical Investigations,* trans. G. E. M. Anscombe, Oxford, Basil Blackwell, 1963, ## 65-67 *,* pp. 31-32, for an account of the way "family resemblances" can stand in for the Platonic "one thing in common" that supposedly bound together the particular instances of a universal.) There are exceptions like Richard Rorty and Richard Bernstein. But for the most part, the immense popularity and impact of Kuhn's work in the English speaking world can be taken as a measure of an ignorance and/or incomprehension of continental philosophy since Kant.

Almost two centuries prior to Kuhn, Kant showed how much our knowledge is constructed by our categories of understanding and forms of intuition (read 'paradigm'). Hegel then showed how our forms of consciousness *(Gestalten des Bewusstseins)* change from one historical era to another (read 'paradigm *shifts')*. Kant, it is true, played into the enlightenment rationalist paradigm by maintaining that the form of consciousness he was describing was *universal*—that the categories of understanding and forms of intuition he had identified constituted *the* structure of rationality for everyone everywhere for all time. And Hegel, it is true, played into a similar faith in the universality of the *science* of spirit (*Geisteswissenshaft)* by claiming that there was a scientifically *necessary* progression from one form of consciousness to another, and that his *Phenomenology of Spirit* and, even more pretentiously, his *Encyclopedia,* represented a science of the necessity of the unfolding of spirit according to a rigorous dialectic. Not until Michel Foucault does the continental tradition attain the courage of its anti-rationalist convictions

enough to maintain that the succession of forms of consciousness (what Foucault calls *epistemes)* are as incommensurable and non-deducible from one another as Kuhnian paradigms.

Plumb Foucault's *Archeology of Knowledge,* plow your way through the dense details of his history of epistemes in *The Order of Things,* or his three volume *History of Sexuality,* and nowhere will you find any suggestion of an orderly or rational progression from one episteme to another. As nostalgic rationalists are inclined to put it in criticizing Foucault, according to Foucault it's just one damn thing after another, no rhyme or reason, no science of spirit, no orderly progression of a predictable or retrodictable history. Hence the hasty classification of Foucault as one more of those postmodernists who don't believe in progress.

For Foucault, as well as for Kuhn and his heirs, there is no logical path of prediction or reduction that will get you from one paradigm or episteme to another. With some exceptions, like Wittgenstein himself whose early and late philosophies span the distance from positivist to post-positivist paradigms, most holders of an older episteme will simply have to die off before a new episteme can take hold.

(d) Tipping Points – Gladwell

In his elegant little book, *Tipping Points,* Malcom Gladwell takes epidemics as one of his prime examples of phenomena that tip from one regime to another in ways that are difficult to predict or describe. So he sets about to describe the unpredictable by looking, as we are doing here, not for some theory from which he could deduce each instance of tipping, but rather, as a journalist is better at doing than an academic, by investigating the details, the phenomenology, of what we see before our everyday faces. And those details, once we look at them in a fairly unbiased, non-theoretical, journalistic, commonsensical way, are not all that magical or mysterious.

Epidemics happen. We all know this. Black plague in the fourteenth century, AIDS and Covid in the 20th and 21st centuries, influenza during the First World War, etc. And it's not that hard to see how and why epidemics happen: A disease reaches epidemic proportions if and only if the acceleration of its spread outpaces the deceleration of our ability to cure it. The common cold does not become a flu epidemic if enough people can get over their colds before passing them on to others. But if a cold is bad enough that each person keeps sniffling and sneezing long enough to pass it on to three others before getting over it him or herself, then what *was* the common cold *becomes a flu epidemic*. Just do the math and you see an exponential curve: one to three, three to nine, nine to 27.

When does the common cold tip over into a flu epidemic? Can you isolate the first instance of an epidemic? In Randy Shilts' magnificent book on AIDS, *And the Band Played on: Politics, People, and the AIDS Epidemic,* he's able to isolate the first carrier. But at that point, AIDS was not an epidemic; it was an anomaly. Not until much later, not until many people had contracted HIV and the disease was spreading rapidly through the bath-houses of San Francisco, did anyone talk of an "epidemic." But by that time, you could not have plausibly pointed to the first instance of the epidemic, even if you could isolate the first carrier of AIDS.

Gladwell uses the logic of epidemics to talk about other phenomena from the reversal of a crime wave in a city to the spread of fads and fashions. In each case he's able to identify signs of tipping. For turning around crime, it was the repair of broken windows and the elimination of graffiti from the New York subways. For the spread of fads and fashions, he identifies particular individuals who have a unique ability to receive a *meme* and pass it on to many others. But in no case does Gladwell give us a theory or a covering law such that if certain initial conditions are satisfied, then a hit *must* happen, predictably and on cue. Wouldn't marketing departments like such a theory! But that's not going to happen because that's not the way emergent systems work. They are not predictable! Yes, you can talk about a given fashion reaching "critical mass" among taste-makers and trend-setters. But "critical mass," in the world of fashion, is only a metaphor, not a logical argument.

(e) Nuclear Physics, Quantum Mechanics, Fission/Fusion (critical mass)

'Critical mass' comes from nuclear physics. Within the domain of nuclear physics it has a perfectly clear and rigorous definition, about which we don't need to get too technical. The main point to be made is that, despite the availability of a scientifically rigorous definition of critical mass in nuclear physics, no one would claim that nuclear physics explains how the attainment of "critical mass" makes hits happen in the fashion world. The concept of critical mass functions as a metaphor, not as an explanatory law, for what happens when an album goes platinum. (Cf. Winslow Farrell, *How Hits Happen,* Harper Business, New York, 1998.)

Quantum mechanics presents us with still further examples of phenomena that exhibit features of emergent systems that are not magic or mysterious, yet pose challenges to the principle of sufficient reason and classical scientific rationalism. A phenomenology of the reception of quantum theory shows just how weird it looked when viewed through the lenses of classical mechanics. That weirdness, not the full body of quantum mechanical

theory, serves as further evidence for the fact that we are in the presence of one more paradigm breaking anomaly, one more chink in the armor of deterministic rationalism.

In a book entitled *The Quantum Brain*, Jeffrey Satinover gives an excellent description of quantum theory's offensiveness to classical scientific rationality:

> At the subcellular level, matter itself actually looks and behaves (in the words of one physicist) "more like a thought" than like the cogs of a machine. *Nothing in the world that causes the particle to jump,* discovered the first quantum mechanists. But the first premise of science is that *everything happens solely as a result of causes in the world.* "If we are going to stick to this damned quantum-jumping," complained one of its founders, "then I regret that I ever had anything to do with quantum theory." (Jeffrey Satinover, *The Quantum Brain,* John Wiley & Sons, New York, 2001, p. 7. In support of his quotations, Satinover cites Abraham Pais, *Niels Bohr's Times in Physics, Philosophy, and Polity,* Clarendon Press, Oxford, 1991, p. 299; and W. Pauli, L. Rosenfeld, and V. Weisskopf, eds., *Niels Bohr and the Development of Physics,* McGraw-Hill, New York, 1966, p. 349.)

The point here is *not* (as Satinover and others want to argue), that quantum theory *explains* the emergent system we call consciousness. The point, whose subtlety approaches that of the complex quotation from Sellars a few pages back, is that quantum theory, like the epicycle describing Saturn's course through the Ptolemaic heavens, is just weird enough so that when you put quantum theory together with all of the other anomalies, then just as the epicycle of Saturn together with the epicycles of all of the other planets helped to tumble the Ptolemaic model, so, the combined force of all these weirdnesses should begin to suggest that the adequacy of the kind of simple old causality that Niels Bohr pined for so poignantly can no more stand up to this full range of anomalies than Ptolemaic astronomy could stand up to the Copernican alternative. The force of this argument is thus very different from an attempt to explain the emergence of consciousness by reference to quantum theory. This latter attempt is an instance of what I'll later describe as *the phallusy of misplaced physics.* We'll come back to this fallacy later, and give further examples. For now, speed is of the essence so that *all* of the items on the list of anomalies can be seen closely enough together so that their initial status as anomalies can gain a critical mass sufficient to tumble the old paradigm.

(f) Punctuated equilibrium – Eldredge and Gould

Punctuated equilibrium is the name given by Niles Eldredge and Stephen Jay Gould to their theory of the evolution of species according to which evolution proceeds by fits and starts, not as gradually and continuously as had been thought. Notice a similarity, a family resemblance, to Kuhn's view of the history of science? Or to those damnable

jumps abhored by Niels Bohr? Or to the sequence of epistemes described by Foucault? According to Eldredge and Gould, tensions in an ecological regime build up until—snap, or *tip,* or *pop*—a whole group of species co-evolve toward a different ecological regime fairly quickly.

In order to describe this tipping, this punctuation of equilibrium, Eldredge and Gould introduce a marvelous image: Galton's polyhedron. Think of continuous change as like the rolling of a sphere. Now replace that sphere with an almost spherical polyhedron with, say, 40-50 different facets. That polyhedron does not roll as smoothly as a sphere. Instead, if nudged by a small force, it sits flat on one facet. If nudged a little harder, it might lift off its facet and settle back to the same facet. Only if nudged by enough force will it *tip* to settle on another facet of its surface. *That's* how evolution works, according to Eldredge and Gould, and they can point to alternations among layers in the paleontological fossil record to prove it. Some layers show relatively little change in the balance of different species over long periods of time, like those periods where a single paradigm reigns and ordinary science proceeds with its puzzle solving. Then other layers show the rapid introduction of large numbers of new species, like those periods in the history of science when there is a scientific revolution and a shift to a new ecological paradigm, a new ecosystem, a new biological regime.

(g) Substantial vs. accidental change – Aristotle

The next item on the list, Aristotle's distinction between accidental and essential change, is not exactly an anomaly. It is not a recent theory or discovery that challenges the paradigm of classical rationalism. Quite the contrary, it pre-dates the Enlightenment by many centuries. But it's worth considering, like the Greek concept of *enantiodromia,* as yet another way of describing a distinction between different kinds of change.

Aristotle distinguished what he called accidental change, like the changing of the color of a leaf, from essential change, as when one substance turns into a different kind of substance. By 'accidental', Aristotle does not mean 'by chance,' or without cause or reason. The term 'accident' gains its meaning from its opposition with 'essential'. 'Essence' refers to a thing's substance, its fundamental nature; 'accidents' are the properties of that substance. A substance can suffer the alteration of some of its properties and still remain the same substance. A plant can grow larger and still remain the same plant. A man can become rich or poor, but still he remains a man. While Aristotle offers many examples of accidental change, he offers very few for substantial change. Like the other items on the above list, essential change is relatively rare, extraordinary, out of the ordinary.

Change of substance —trans-substantiation—seems almost miraculous, and for that very reason Aristotle's analysis of the different kinds of change was eagerly adopted by Christian theologians as a way of talking about the trans-substantiation of wine into the blood of Christ and bread into his body in the holy sacraments. Cows do not turn into horses. Water does not turn into wine. So it's not easy to find examples of essential or substantial change, or to explain how such a thing could happen. But Aristotle felt it necessary to talk about such transformations. Why? One answer would have it that the possibility of essential change follows as a result of other features of his metaphysics; like it or not, it simply has to be countenanced as a consequence of other parts of his system. (Aristotle's metaphysics includes a distinction between what he called prime matter and proximate matter. As wood, proximate matter, is to a bed, so prime matter is to wood. While wood has form—its hardness, its grain, its woodiness—which can then take on further properties that make it into a bed—the shape and structure given to it by a carpenter—prime matter is featureless. It takes on the features of wood or of stone. In this three layered system, the possibility of essential change, changes in the essences inhering in prime matter, follows as a consequence of the analogy with changes among the properties in the proximate matter of substances.) Another answer would say that Aristotle's system is as it is because he sensed that the way the world works, there really are these different kinds of changes—the usual, normal types of changes that can be described as accidental or incremental, and other, rarer, more extraordinary kinds of change. Interestingly enough, one of the few examples of substantial change that he offers spans one of the very transitions we've already mentioned as calling for a science of emergent systems: The difference between life and death. Aristotle uses the transition from man to corpse as an example of essential or substantial change.

Aristotle was a biologist. His worldview lacked a sense of time that was historical. He believed that the number of species was fixed. Nonetheless he was intensely interested in species, in natural kinds, and in the processes of coming to be and passing away. One of his books is titled *On Generation and Corruption.* As contemporary philosophy of science shifts its focus from physics to biology, as has been happening over recent decades, (For support of this very general claim, see W. W. Bartley, "Philosophy of Biology versus Philosophy of Physics," Chapter 1 of *Evolutionary Epistemology, Rationality, and the Sociology of Knowledge,* ed. Gerard Radnitzky and W. W. Bartley, Open Court, LaSalle, Illinois, 1987, pp. 7-46; Freeman Dyson, *Infinite in All Directions,* Harper & Row, New York, 1988, pp. 6-8; and *Imagined Worlds,* Harvard University Press, Cambridge, 1997, Chapter 2; Hans Jonas, *The Phenomenon of Life,* University of Chicago Press, Chicago, 1966; Elliott Sober, *The Nature of Selection: Evolutionary Theory in Philosophical Focus,* MIT Press, Cambridge, 1987.) we find ourselves re-encountering some of the explanatory challenges that Aristotle dealt with, but which slipped off the radar during the early 20th Century when physics was the focus of the philosophy of science. During the rise of logical positivism, the questions that seemed

worth asking and answering in the shadow of Einstein and Heisenberg had to do with the nature of space and time and causality. No one was talking about natural kinds or species. That kind of talk was thought to have been left behind by modern science. Newtonian cause and effect, impact and reaction, replaced all talk of natural kinds and their essences, or explanation of change by reference to potentialities and actualities. The Aristotelian language and taxonomy was not modified by incremental changes. The Aristotelians died. Their paradigm was not modified; it was superceded, replaced.

(h) Catastrophic Morphogenesis – Rene Thom

Now, however, we find ourselves, in the shadow of Darwin, confronting some of the same questions that so preoccupied Aristotle. Consider the title of Rene Thom's *Structural Stability and Morphogenesis.* 'Structural stability'—how is it, why is it, that things remain true to type, that cows don't turn into horses? 'Morphogenesis'—*morphe* = form, *genesis* = coming to be, so 'morphogenesis' is about the coming to be of forms. How, Rene Thom wants to know, do structures come to be out of other structures, and then retain their form in the face of forces for further change?

When modern biologists return to these very Aristotelian questions, but without Aristotle's metaphysics, and from within a paradigm largely shaped by modern physics, they find themselves baffled and tongue-tied. Without suggesting that we should go back to Aristotle for answers, it is nonetheless intriguing to watch modern biologists struggling to fit the round peg of biology into the square hole of physics. How difficult it is to do biology from within a paradigm shaped by physics! Questions that seemed relatively easy to solve (albeit *wrongly* from the perspective of modern science) using Aristotle's biologically derived metaphysics are now much harder to solve using a paradigm derived from physics.

Not to suggest that we go back to Aristotle for *answers*; but as we give up the positivistic attempt to *reduce* biology to physics, we will find ourselves struggling with some of the same *questions* that plagued Aristotle. As we approach biology—life—as an emergent system that is irreducible to the traits of physics, we will be better able to describe the coming-to-be of some radically new and different kind of thing: the emergence of life. And as we come back to some of Aristotle's questions, we may find that some of his ideas have some cogency. We won't attempt to deduce Aristotelian science from our new science. But we might find that the stereoscopic vision of a modern science supplemented by a science of emergent systems allows us to show how and why things behave in ways that conform to Aristotle's descriptions as well as they do. We may find ourselves re-describing what Aristotle called *substantial change* or *potentiality* or *final cause (telos).*

Certainly Rene Thom found himself compelled to entertain the kinds of questions that Aristotle asked, e.g., about potentiality and teleology. Thom was suspicious of monological explanation. He questioned a reliance on singular causes or point-sources of change.

> I am reluctant to subscribe to the current belief that a point mutation, affecting just one nucleotide, is sufficient to inhibit the activity of a gene; this seems to me to repeat on another plane the error of the morphologists who believed that the destruction of one neuron in the brain would stop the process of thinking . . . Even if life is only a tissue of catastrophes, as is often said, we must take into account that these catastrophes are constrained by the global stability of the process and are not the more-or-less hazardous game of a mad molecular combination. Even adopting the anthropomorphic point of view that there is a mechanism for reading the DNA that is perturbed by errors, might we not push this anthropomorphism to its full extent and admit that the errors are oriented, as in Freudian psychology, by the "unconscious" needs and desires of the ambient metabolism? (Rene Thom, *Structural Stability and Morphogenesis,* trans. D. H. Fowlder, Reading, Mass., Benjamin/Cummings Publishing Co., 1975, p. 282.)

Or, to put it in language we'll return to below, could it be that the environment, the "ambient metabolism," the outside world, *wants* something new? Thom argues that "Metazoa have, located in the kinetic configuration of the metabolism of their gametocytes, a model of their actual conditions of existence," or, in plainer English, animals carry inside themselves a map of what is outside themselves.

> The topology of this universal model will reflect less the phylogenetic relationships than the functional interaction between species, so that the distance between bee and snapdragon will be less than that between bee and butterfly. The big evolutionary advances of history will be described by global deformations of this universal model; the metabolism of gametocytes will appear as a kind of research laboratory, a device simulating conditions of existence close to the actual ones, and an evolutionary advance in this device will provoke unstable virtual catastrophes that will correspond to a new functional morphology, expressing the organic and physiological adaptations made necessary by this evolutionary change. Only when this new morphology is sufficiently stabilized, can evolutionary advance begin, and it will manifest itself as a mutation, a rearrangement of chromosome stock, and the appearance of new organic forms. According to these ideas, the fish already "knew", before they became amphibious, that a life on land would be possible for them, and what new organs they would need. (*Ibid.,* p. 294.)

In this passage, Thom is anticipating Eldredge and Gould's theory of punctuated equilibrium. He is saying that the fish contains within itself a kind of map, not only of its existing environment—water—but also, by virtue of the fact that the fish is part of a universe, an "ambient metabolism," containing more than water, the fish can model in its

"research laboratory" another possible environment—land—and on the basis of that modeling capacity, can mutate in a way that is "not the more-or-less hazardous game of a mad molecular combination," but rather a more directed, more Lamarckian "aiming" toward that new global order. Just as much as Thom resists relying on a monological point-source in a single gene, so likewise he will resist any preconceived point-destination in a pre-destined telos.

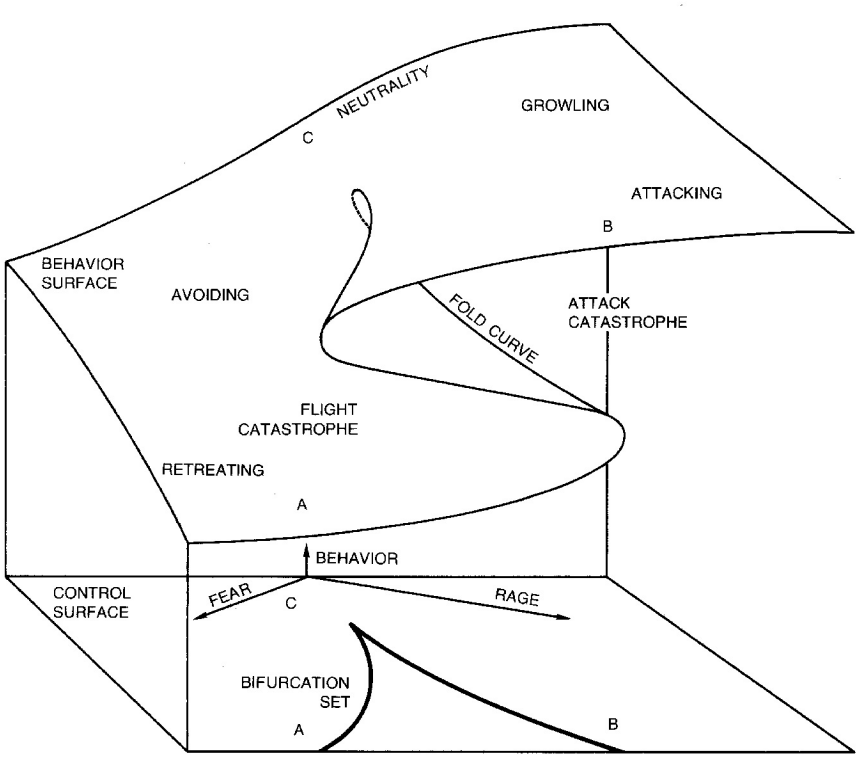
> Global finalists (like Theilhard de Chardin) push the analogy between evolution and epigenesis of the egg to the extent that they believe that, just as the embryo develops according to an established and unchanging plan, so must the wave of evolution . . . but this neglects the important difference that the development of the embryo is reproducible and thus an object of science, whereas the wave of evolution is not. To assert that a unique and unrepeatable phenomenon occurs according to a plan is gratuitous and otiose. (*Ibid.,* p. 291.)

Thom embraces teleonomy without teleology. That is, he acknowledges purposive behavior in evolutionary change, but without pinning his account on the existence of some single point-destination or telos, which he regards as "gratuitous and otiose."

How does Thom mathematize the discontinuities described in the phrase, "unstable virtual catastrophes that will correspond to a new functional morphology"? Thom developed a series of mathematical models he called "catastrophes," not because they are bad or disastrous, but rather because he wanted to call attention to their mapping of *discontinuities*. Thom's "catastrophe theory" has a lot in common with what Nobel Prize winner, Ilya Prigogine, called "bifurcation theory." Both men were interested in the ways that nature *branches* from time to time, and often in ways that a calculus of continuity has difficulty describing.

Thom theorized seven different types of "catastrophe." The force and power of his analysis can be grasped from a brief description of just one of his seven different types, the "fold" catastrophe, and its application to modeling the "bifurcation" between fight and flight in the behavior of dogs.

Figure 3 Model of Aggression in Dogs



Imagine the behavior of a dog as following a path on the behavior surface beginning at point C and moving, on a path of increasing rage, toward B. While launching an attack, the dog sees a man with a large stick. The dog's behavior then moves from point B towards point A. The beauty of this mathematical model, the beauty of Thom's "catastrophe theory," lies in the way it models, on a continuous surface, the discontinuity in behavior we observe when we see the dog rapidly shift from fight to flight. Moving on the behavior surface from B to A, the dog will enter the intrinsically unstable area between the dark lines on the control surface labeled "bifurcation set." Mapping this shift on the behavior surface, it's easy to see where the path from B to A will "fall" discontinuously from the "attacking" plane of the behavior surface to the "retreating" plane. (This application of Thom's catastrophe theory is drawn from E.C. Zeeman's article, "Catastrophe Theory," *Scientific American,* April 1976.)

(i) Enantiodromia – C. G. Jung

Thom was not the first to notice such non-linear phase shifts or transitions. He simply (or not so simply) found a way to mathematize and thus lend rigor to what had earlier been described with terms like *enantiodromia.* This term dates from Greek philosophy. It is best translated as "turning over into its opposite." Examples might include the young liberal turning into an arch-conservative, or, as Plato describes in the *Phaedo*, the starry-eyed lover of humanity who, after some bitter disappointments, turns into a misanthrope. (Cf. Plato, *Phaedo* 89d.) Or, to get even closer to home, the person who puts so much stock in reason that, when confronted with some anomalies, turns into a radical skeptic.

(j) *Aufhebung* – Hegel

Another thinker preoccupied with things "turning over into their opposites" was Hegel. His dialectic is largely a process of showing how forms of consciousness, when pushed to their extremes, have a way of taking on the characteristics of their opposites through the very process of opposition. In the master/slave dialectic, it is the slave who gains the greater mastery of self-consciousness. In the battle of enlightenment with superstition, rational enlightenment begins to take on some of the characteristics of a religion. As a result of this dance of mutual enantiodromias, an *Aufhebung* occurs. This almost untranslatable term (literally, lifting up; *heben*=lift, *auf*=up) has been translated as 'sublation,' which sounds very close to 'sublimation.' The idea, strange and weird as it might sound to enlightenment rationalism, is that sublation "takes up but transcends" the dialectical contraries that gave rise to it.

Coming to *Aufhebung* as late on the list as we have, we now have the vocabulary to make this idea sound a little less strange. Of course sublation transcends its precursors and is unpredictable from them. Yet it also "takes them up" in a way that makes *Aufhebung* irreducible to its precursors. And of course a flailing about from one extreme to another is likely to generate the conditions for *Aufhebung*. Such a pushing of extremes is just what is required to push a system far from equilibrium, to punctuate equilibrium.

The appearance of extremes is just the sort of thing that results from a bifurcation. And the resolution of tension created by extremes, according to Hegelian dialectics, has nothing to do with mere compromise or a seeking of some commensurable mean. An *Aufhebung* marks the appearance of something radically new under the sun, an emergence that transcends the old even as it sublimates its dialectical extremes in a way that is incommensurable with them. The transition may be catastrophic, as was the French Revolution, a period of history fraught with violent conflict. It yielded a new era of *liberté, egalité, fraternité* . . . along with a reign of terror, and the transition from monarchy to democracy.

It would be a mistake to claim that everything else on our list of anomalies is really nothing but a case of *Aufhebung.* Nor is Hegel's concept of *Aufhebung* reducible to some other item on the list. Instead, the point is to note the family resemblance between what Hegel was talking about and what some of the other theorists represented on the list we're traversing. If so many of these concepts and theories seem to be knocking on the same door in roughly the same rhythm, shouldn't we begin to suspect some sort of club on the other side of that door?

(k) Synchrony -- Strogatz

It's all coming together. Is there a word for this kind of convergence? Yes there is: *Synchrony*. (Steven Strogatz, *Sync: The Emerging Science of Spontaneous Order,* Hyperion, New York, 2003.) It happens.

Why does it happen? Apart from a science or theory of emergent systems, synchrony will be transformed by the phallusy of misplaced physics (See later part with that title) into "synchronicity," an idea that is weird, almost uncanny. As developed by theorists from C. G. Jung (C. G. Jung, *Syncronicity: An Acausal Connecting Principle,* trans. R. F. C. Hull, Princeton University Press, Princeton, 1973.) to Joseph Jaworski, (Joseph Jaworski, *Synchronicity: The Inner Path of Leadership,* Berrett-Koehler, San Francisco, 1996.) 'synchronicity' is used to describe extraordinary coincidences that seem to suggest some hidden order behind the manifest phenomena.

Jung describes a day in which an extraordinarily large number of references to fishes pile up. In another example, he recounts a conversation with a patient who reports a dream in which she was given a golden scarab.

> While she was telling me this dream I sat with my back to the closed window. Suddenly I heard a noise behind me, like a gentle tapping. I turned round and saw a flying insect knocking against the windowpane from outside. I opened the window and caught the creature in the air as it flew in. It was the nearest analogy to a golden scarab that one finds in our latitudes, a scarabaeid beetle, the common rose-chafer *(Cetonia aurata)* which contrary to its usual habits had evidently felt an urge to get into a dark room at this particular moment. (C. G. Jung, *Synchronicity: An Acausal Connecting Principle,* Berrett-Koehler, San Francisco, 1996.)

How can such extraordinary coincidences be explained? Etymologically, the term 'synchronic' derives its meaning from its (dialectical) opposition to 'diachronic,' which means taking place according to the linear sequence of time. 'Synchronic' means something close to simultaneous, as when it all comes together at once. As if by magic? Certainly not predictably.

Interest in synchronicity is widespread. (See, for example, F. David Peat, *Synchronicity: The Bridge Between Matter and Mind,* Bantam Books, New York, 1987.) Like emergence itself, synchronicity is sometimes invoked in New Age circles in a way that smacks of the miraculous. Taken as such, instances of synchronicity hardly help as existence proofs for naturalistic (rather than super-natural) evidence of emergence. Syncrhonicity looks like synchrony *by design*. But if a naturalistic account of emergence can be given, then instances of synchrony, but *not synchronicity,* will surely be part of that account.

(l) Singularity – V. Vinge, Kurzweil

The coming together of several components—*now, tip, pop, all of a sudden, catastrophically*—has yet another name in addition to synchrony: 'singularity'. Both Verner Vinge and Ray Kurzweil use this term to describe the moment when the pace of accelerating technological change catches up with itself and a series of mutually accelerating dynamics go hyper-exponential. The singularity is the asymptote towards which exponential curves rise ever closer to touching. Moore's Law—that the speed of microchips doubles every eighteen months even as the cost of computation is cut in half—represents one of the best known examples of an exponentially accelerating curve. Where will it end? Just how fast, how cheap, can computation become? At some point a singularity must occur; the curve goes so close to vertical that the speed is *almost* indistinguishable from simultaneity, synchrony. At the asymptote, time would seem to cease.

Time, as one wit put it, is Mother Nature's little way of keeping everything from happening at once. But at the singularity, everything *does* happen at once. In a blinding epiphany—*now, pop*—time seems to stand still. Everything, all eternity, is present at an instant.

Will it happen? Could it happen? Who knows? But the term has entered our vocabulary and, if the concept has any legitimate place, it is in the context of things that go *pop!*

(m) Peak experience – Maslow

Mystics speak about such moments, when they are not standing silent before the peace that passeth all understanding. (Cf. Evelyn Underhill, *Mysticism,* Meridian, New York, 1955.) Scientists are now approaching such mysteries with the mathematics of asymptotic curves of accelerating change. And psychologists, mainly in the wake of Abraham Maslow, are attempting to describe the experience of such moments. Maslow called them "peak experiences." (Abraham Maslow, *The Psychology of Being,* Wiley, 1998) Modern

descriptions are pre-figured in the writings of the mystics, and in the writings of Soren Kierkegaard, who puzzled over the paradox of eternity occurring in time. (Soren Kierkegaard, *Philosophical Fragments,* trans. David Swenson, Princeton University Press, Princeton, N. J., 1962.) In the context of enlightenment rationalism, and an understanding of the world that sought explanations by way of causes preceding effects, there could be no understanding of ecstatic singularities. But so many people report these experiences of time standing still and everything coming together in the moment. Granted, they are often tongue-tied, and understandably reluctant to eff the ineffable. But they find ways to point to the experience, if not to reduce it to nothing but a hallucination induced by some odd chemical imbalance. Apart from the ruminations of the mystics, and the writings of transpersonal psychologists, there is a controversial literature surrounding the use of hallucinogens as a route of access to peak experiences. For some, the use of drugs stands as a clear demonstration that such experiences have no transcendent religious referent—God, or The One—but are instead nothing but epiphenomena induced by chemical means. For others, from Aldous Huxley to Huston Smith, and Michal Pollan the route of access is less significant than the self-validating evidence of the experience itself. (Cf. Huston Smith, *Cleansing the Doors of Perception,* Tarcher, 2000.)

(n) Beauty

We needn't go so far afield for experiences that reveal the architecture of epiphany. Art and the experience of beauty give us intimations of the sublime. Like a number of the other items on the list, beauty has the earmarks of an emergent phenomenon. Beauty is not reducible to its components which, apart from the whole of a work of art or natural beauty, give no indication of a potential for beauty, e.g., a patch of brown that is part of the Mona Lisa. Beauty is not reducible to its components, nor is beauty predictable from any one component. Try to fill in the blank: "She is beautiful because _____." Whatever you say will sound ridiculous.

Beauty is not *caused* by any one component; it is the result of their combination, their coherence, their proportions . . . and very soon all attempts at the *explanation* of beauty will draw you into the hazy and inexact realm of aesthetics. Once there, it is precisely the haze that is worth appreciating more than the products or conclusions of any particular aesthetic theory. The haze tells us that we're in the presence of phenomena for which monological science is inadequate. Hence such hackneyed expressions as, "I don't know much about art, but I know what I like," or "*De gustibus, non disputandum est."* (There's no disputing taste.)

Beauty emerges, whether from nature, or from the creativity and genius of the artist and/or performer. We know it when we experience it. We are drawn to beauty. Beauty induces desire, and desire is gratified by the experience of beauty. The structure of gratification is prefigured in the way a piece of music plays with dissonances that "want" to resolve in a cadence. The final resolution on the tonic—pop!—brings the melody to a completion that satisfies and gratifies, which is not to say that there won't be other delights along the way.

Unlike monological science whose truths are universal or they are not truths, the experience of beauty is almost always rooted in some particular culture or genre or tradition. Art has a history. Mondrian had his time and place, and if an art student today paints like Mondrian, she will be criticized as derivative. Art is always situated in a context. Consequently, any examples used to prove a point about the experience of beauty will be open to criticism from those who don't share the same context, taste, or tradition.

This much said, so that no readers will feel left out or disputatious, it becomes possible to point to *one example* of the emergence of beauty without claiming anything like universal reach or knock-down-drag-out validity for this particular example: Live performances by the late Jerry Garcia and the Grateful Dead. If this example doesn't work for you—if you are more drawn to Mozart, or to Rembrandt, or to the plays of Eugene O'Neil—then read what follows with your own example in mind.

The Grateful Dead were the leading live performance band in the world for the decades they played, from the late sixties until Jerry Garcia's death in 1997. No other band sold as many tickets, though many other bands sold more records. To their millions of fans they offered unique experiences. As all Dead-heads know and have intoned many times, "There is nothing like a Grateful Dead concert." Part of the magic lay precisely in its unpredictability. Not every show was special. Therein lay part of the suspense, part of the mystery. But in some shows there would come a moment—pop!—when it all came together: the music, the performance, the audience, the melody, the words, the rhythm . . . and the experience was like nothing else: ineffable, unrepeatable, evanescent, but moving and, for some, transformative— life changing.

The Grateful Dead attracted followers, almost a cult; but there was no doctrine, no teaching. If you "got it," if you shared the experience, you became part of a loosely structured, non-hierarchical community. Tens of thousands of cars and trucks bear the red, white and blue death's head decal suggesting that the occupants "got it." Getting it is not reducible to any one element. The magic is a matter of the way it all comes together in particularly ecstatic moments.

"Getting it" is itself one of those all-at-once, all-of-a-piece, tippings. Think about what it's like to get a joke. It's sudden. Laughter is the body's reaction to the suddenness of the transformation of one meaning of a *double entendre* flipping into its other meaning. You can't explain a joke. That would be tantamount to offering the first meaning as a first instance, then to be followed, in rational sequence, by the second meaning. But separating the first instance as first from the second meaning as second then separates the two in a way that destroys the tension of their unity in the joke. Laughter is the body's expression of the mind's spontaneous delight at *getting it.*

(o) Orgasm

'O' is for orgasm, unpredictable for many, irreducible for all; like nothing else, least of all a sneeze. The onset of orgasm, so intimately or secretly familiar to every post-adolescent, must be a complete mystery to every pre-adolescent. In this virtually universal, oh so familiar experience, we find an everyday or every night example of singular synchrony where it all comes together in "a glowing tingle which *now* had reached that state of absolute security, confidence and reliance not found elsewhere in conscious life," or so writes Vladimir Nabokov in his stirring novel, *Lolita.*

Orgasm—what can one say? It occurs. It is the last, and perhaps most obvious and irrefutable of our existence proofs for the reality of emergence. It all comes together: the caresses, the feelings, the primitive urges and their higher sublimation in love and eroticism. Orgasm, right before our very noses, but hidden by concerns for propriety. Orgasm: the climax of foreplay, and the climax of this section on anomalies that won't yield to reduction or prediction. Orgasm: a perfect transition from the first six traits of emergent systems to the seventh, desire. For orgasm is, of course, the climax of desire, the singularity, the asymptote toward which excitement rises to punctuate the equilibrium of mere contentment.

But first we have four more traits before getting to the seventh . . .

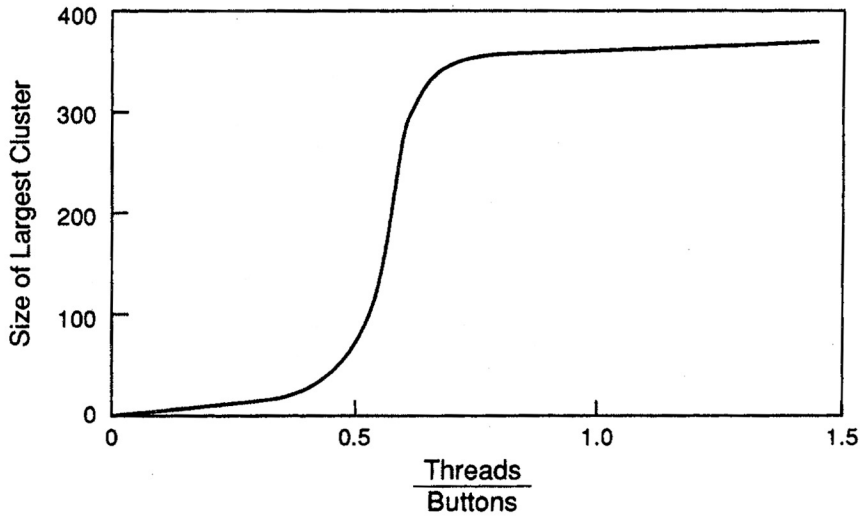
Third Trait: Holism -- The whole influences the nature of the parts or *The importance of requisite complexity*

In considering the First Trait of the impossibility of first instances in several different domains, we are driven fairly quickly to a consideration of its corollary, the requirement of a complex whole for the influence of any part. My friend, Alfonso Montouri caught me out on an earlier formulation—that the whole *determines* the nature of the part.

Holism is not deterministic. It *influences* but does not monologically *cause* the nature of the part. Quite the opposite of the logico-analytic quest for the simplest or first elements of a system, in the case of emergent systems, *complexity* is a prerequisite. You can't have a word without a language. You can't have a language without a lot of words.

You can't have an idea without a mind. Locke, Hume, and their empiricist heirs were wrong to imagine that you could build up a mind out of simple ideas taken in through the senses. An appreciation for emergent systems will tell us that minds are complex, thoughts presuppose minds, and there is therefore no such thing as a simple idea. *All* ideas are complex, however much the objects of some ideas—say, redness—may appear simpler than the objects of other ideas—say, tragedy. Complexity is a prerequisite of mind, and mind is a prerequisite of any idea at all.

At the level of molecular biology and the origins of life, we now have the benefit of the research performed at the Santa Fe Institute (and elsewhere) to lend rigor to our understanding of requisite complexity. Stuart Kauffman asks us to imagine a scattering of 500 buttons on a table. Connect three pairs of buttons with three lengths of thread, then close your eyes and choose any of the 500 buttons at random while blindfolded, pick it up, take off the blindfold and see how many buttons you have picked up. Plot the answer on a graph with the number of buttons picked up on the y-axis, and the ratio of connecting threads to the total number of buttons on the x-axis. Now connect several more pairs of buttons together, and try again to pick up some buttons by choosing one button at random. Plot the results. As you continue adding connective threads, picking up buttons at random, and plotting the results, you will end up with a curve like this:



What is lovely about this curve (Stuart Kauffman, *At Home in the Universe,* Oxford University Press, New York, 1995.) is the way it illustrates how what will seem like a

discontinuous phase transition—a step jump—can result from a continuous accretion of threads. When the ratio of threads to buttons reaches approximately 0.5, *something happens!* Vast clusters of buttons emerge! The core of Kauffman's argument in *At Home in the Universe* is to be found right here. His argument for the origin of life draws on an analogy between those buttons and a stew of molecules sufficiently complex to virtually guarantee the cooking of something like life.

How like life? What is life? What a simple question. You would think that science would have given us a relatively simple answer. But just because life is an emergent system, the science we have inherited from the Enlightenment has been hard pressed to come up with any simple definition of life at all. James Lovelock, author of *Gaia: A New Look at Life on Earth,* found himself in search of a definition of life that he could use in order to answer the question of whether the earth and its ecosystem could be regarded as a living system.

> I expected to discover somewhere in the scientific literature a comprehensive definition of life as a physical process, on which one could base the design of life-detection experiments, but I was surprised to find how little had been written about the nature of life itself. . . Data galore had been accumulated on every conceivable aspect of living species, from their outermost to their innermost parts, but in the whole vast encyclopedia of facts, the crux of the matter, life itself, was almost totally ignored. (James Lovelock, *Gaia,* Oxford University Press, New York, 1979, p. 3.)

More than a decade after Lovelock published these lines, his co-worker, Lynn Margulis and Dorion Sagan published their book, *What is Life?,* in which they write:

> In physics, Werner Heisenberg's uncertainty principle limits what is measurable. In mathematics, Kurt Gödel's incompleteness theorem warns that every mathematical system, if complete, cannot be consistent and, if consistent, cannot be complete, since to define it axioms are needed from outside the system. Such scientific uncertainty also impedes any search to define life. On the one hand, a final definition of life by life may be like kissing your elbow or rolling your eyes to see your own optic nerve: impossible. (Lynn Margulis and Dorion Sagan, *What is Life?,* Simon & Schuster, New York, 1995, p. 40.)

Researchers in the relatively recent field known as *artificial life,* or *A-life,* agree that the question, *What is Life?,* calls for a deeper understanding. "What drives men and women engaged in the quest for a-life is a desire to decipher the vast tangle of obscurities that nature has laid before us, particularly in regard to the deepest question of all, What is life?" (Steven Levy, *Artificial Life: The Quest for a New Creation,* Pantheon Books, New York, 1992, p. 10.)

So, what *is* life? As with consciousness, these and other writers often describe it as one of the great mysteries. Why? Not because we've yet to find some hidden variable that would explain life the way the mosquito was discovered to explain the spread of malaria. Instead, life is described as mysterious because, like consciousness, it is an emergent phenomenon and, as such, defies the kind of definition or explanation that are adequate for non-emergent phenomena. There is no analogue to the mosquito that can explain life under a covering law, no *élan vital* to serve in a monological *cause*. As an emergent system, life will exhibit its own traits that cannot be predicted from the traits of the precursor systems. Nor can the traits of the precursor systems, say classical chemistry and physics, be deduced from the traits of a successor system that would satisfy us as an adequate account of life. Nor can life be explained by reduction to monological chemistry and physics.

Kauffman's graph of the connected buttons has the shape of a phase shift. First there was a table full of disconnected buttons and a few pairs that were connected. Pick up any button when there are only a few connecting threads and chances are you didn't pick up many. But then as the ratio of threads to buttons approached 0.5, *shazzam, pop,* you get a great hairball of buttons every time you go to pick up just one. *Shazzam* has the sound of magic. Certainly it suggests a discontinuity, a certain suddenness.

Once again, the intention here is to replace the aura of magic with the satisfaction of scientific explanation. . . but a particular kind of scientific explanation. These initial approaches to the origins of life, language, and consciousness might strike the reader as fundamentally unsatisfying. This is as it should be. A somewhat more satisfying account will come later when we travel horizontally across the rows of the chart showing the inter-relationships among the eight traits at each level of emergence. Not until then will the pieces of the puzzle about consciousness, or the pieces of the puzzle about life really come together. For now, we are still at the stage of laying out the pieces before putting them together in a new way.

Still, it might be worth previewing part of what comes later, namely the nature of explanation. *The funny thing that happens when you go to explain the evolution of emergent systems is that explaining the evolution of emergent systems makes you think differently about the nature of explanation.*

Rather than explaining something by stating a covering law, $p \rightarrow q$, then explaining q by demonstrating the existence of initial condition, p, when it comes to the evolution of emergent systems, we'll need a very different kind of explanation. We'll end up explaining q by seeing the shape of the hole left by what is *not q*— the entire environment around q that *fits q* as glove to hand but is nonetheless *not q.* When you have seen and

appreciated *not-q,* then you come to appreciate why *q* has to be just the way it is, why it could not be otherwise.

Think of the relationship between the roles of the genotype and the phenotype in the evolution of species. Focus on the genotype, and the traditional covering law model for explanation seems adequate. If *q* represents the genotype of an offspring, you can make some predictions about that genotype by virtue of a description of the genotype of the parents, represented by *p.* If the parents exhibit a certain genotype (say, a propensity toward breast cancer), then the female children are at greater risk than normal. $p \rightarrow q$ says, If the parents have the genes, the kids will have the propensity. Supply information that demonstrates that the parents *have* the genes, then conclude that the (female) children of those parents are at risk.

All very well and good. But when you want to attend to the evolution of species, and to the survivability of a certain gene pool, attention to the genes alone, using a covering law as above, will not be adequate. You must also look at the survivability of the phenotype *in a given niche.* And at that point, the shape of the niche—does it include mamograms, etc.—is essential. The shape of the niche, the shape of the "hole" left by *not-q,* is as important as the genotype, *p,* in determining the evolution of the species. This evolutionary style of explanation, this testimonial of the survivor as to its adaptability to its environment, is not like explanation under a covering law. It is a different kind of explanation. It is a kind of explanation that grows out of the account of the world we get with a stereoscopic vision involving the eye of monological science *and* the eye of a science of emergent systems. (For a discussion of the difference between what he calls "cybernetic explanation" and causal explanation, see Gregory Bateson, "Cybernetic Explanation," in Gregory Bateson, *Steps to an Ecology of Mind,* Ballantine Books, New York, 1972, pp. 399-410.)

Fourth Trait: Recursivity

A self-enclosed wholeness is achieved by some form of closure. He're I'm going to give a list of references to *diffparerent* approaches to the theme of recursivity, much as I gave a list of things that go *pop,* only without letteters or bullet points, but you'll get part of the point from the simple multiplicity of this list. Again, in order to prove a paradigm shift, your proof has to necessarily be trans-disciplinary. Otherwise we're not talking about a paradigm shift, just a single discovery. So here goes: *feedback,* according to Norbert Wiener's cybernetics; *auto-poeisis* according to Varela, (Francisco Varela, *Principles of Biological Autonomy,* North Holland Press, New York, 1979.) or "hyper-cycles" according to Manfred Eigen (Eigen and Schuster, *The Hyper-cycle: A principle of natural self-organization*, Springer-Verlag, Heidelberg, 1979), or *auto-catalytic closure* according to Stuart Kauffman, (*At Home in the Universe,* Oxford University Press, New York, 1995; and *Investigations,* Oxford, 2000.) or John Holland (John Holland, *Hidden*

Order: How Adaptation Builds Complexity, Addison-Wesley, Reading, Mass., 1995.) On both Kauffman and Holland and the work of their colleagues at the Santa Fe Institute, see also Mitchell Waldrop, (*Complexity: The emerging science at the edge of order and chaos,* Simon & Schuster, New York, 1993.*)* or self-referentiality as discussed by Gregory Bateson, (Gregory Bateson, *Steps to an Ecology of Mind,* Ballantine Books, New York, 1972*)* or "boot-strapping" according to Doug Engelbart, or "synergy" according to Peter Corning, (Peter Corning, *The Synergism Hypothesis,* McGraw Hill, 1983.) or "reentry in the thalamocortical system" according to Gerald Edelman, (Gerald M. Edelman and Giulio Tononi, *A Universe of Consciousness: How matter becomes imagination,* Basic Books, New York, 2000.) or the claim that "there is nothing outside the text," according to Derrida, or the hermeneutic circle according to Gadamer. (Hans Georg Gadamer, *Wahrheit und Methode: Grundzuge einer philosophischen Hermeneutik,* J.C.B.Mohr, Tubingen, 1965.) and Foucault. (Michel Foucault, *The Order of Things* (Routledge, 1970). OK, that's enough to prove that we've got something significant to talk about, *recursivity.*

Emergent systems are loopy. They double back on themselves. They are self-referential, reflexive, reflective, re-entrant, cyclical, *recursive.* The kind of biological reproduction characteristic of life requires not just repetition but rebuilding out to reproduce itself. The egg and the seed contribute their DNA to the DNA of the embryo.

One of John von Neumann's great insights in his work on self-reproducing automata was that it's not enough for a machine to make another copy of itself; it has to make another copy of itself that can make yet another copy of itself. (John von Neumann, *Theory of Self-reproducing Automata,* A. W. Burks (ed.), University of Illinois Press, Urbana, 1966.)

Because, after all, the first machine was able to make another one of *it*self. So if it's going to produce another one of *itself,* the one it produces has to be just as capable of producing another one of *it*self. Otherwise the reproduction wasn't an *exact* reproduction. If reproduction does *not* achieve this double loop recursion, it's as if a fertile horse gave birth to an infertile donkey—far from exact reproduction. So the loops of self-representation and self-reproduction must continue by way of a passing along of the blueprint. Each new unit has to be able to produce the next in kind, and do so in a way that the next can produce the next, which will produce the next . . . *ad infinitum.*

Sometimes the loopiness of emergent systems takes the form of feedback—a relatively straightforward concept very familiar to information theorists and sound engineers. Think of the shriek created when a microphone gets too close to a speaker and feedback gets amplified. That's positive, or runaway, feedback. A thermostat, or a governor on an engine, gives negative feedback, limiting an extreme and bringing the system back inside

a set of boundaries, hot or cold, acid or base, salty or fresh. Negative feedback curbs extremes and keeps systems inside a set of boundaries.

Some loops don't involve circular tracks of energy or matter. Some are purely logical or semiotic: self-referential in a way that uses a map of the self to re-present the self--referring first of all, but referring to self rather than referring to some other referent. The word 'I' and indexicals like 'this' and 'here' have this structure. Mirroring redoubles the image of its source, but with a left/right difference. Narcissus is the patron saint of this particular form of loopiness.

Reflexive sociology, as practiced by Alvin Gouldner and Ernest Gellner and others, takes up the existential challenge of changing the conditions of the society under study. Rather than practicing a value-free inquiry that leaves everything as it was, reflexive sociology is a critical sociology that tries to use its reflection to change the society it is reflecting. It chases its own tail. To the extent that it is *critical,* it would be happy to put itself out of business by issuing self-refuting prophesies. Example: The clarion calls of ecological doom issued by the Club of Rome in the book, *Limits to Growth.* Some saw that book as simply wrong because it ignored the role of human ingenuity. Others saw its self-referential strategy, and its success at provoking the very environmental sensitivities that would help to forestall its most alarming scenarios.

Existentialism is a philosophy that puts into question the existence of the questioner: Who are you? What are you? Whom would you become? And what is your being such that it is able to put its being so in question? Such thin ice the existentialist skates on! No wonder there is anxiety. No wonder there is dread at the nearness of the abyss! A whiff of nothingness.

Self-referentiality can seem like lifting oneself by one's own bootstraps. Lacking any firm foundation outside the self, lacking any impossible skyhooks from which to suspend itself, the self *must* lift itself by its own bootstraps. Lacking any designer or creator outside itself, the self *must* accomplish self-organization. The discovery of self-organizing systems was a crucial antidote to the death of the great Organizer, the creator God of Genesis. Organization happens. It doesn't need an outside organizer. But such self-organization is therefore inevitably self-referential, recursive.

Understanding self-organizing systems has been the holy grail of late 20th Century thinking, from von Neumann's self-reproducing automata to Maturana and Varela's work on autopoiesis, to Chris Langton's preoccupation with Artificial Life, to Manfred Eigen's hypercycles, to Kauffman's work on autocatalytic sets, to Brian Arthur's work on positive sum, self-reproducing wealth through increasing returns in economics.

You *can* get more out of less. This is a large part of the meaning of emergence according to John Holland, author of *Emergence* (Oxford University Press, 2000). But emergence is always very loopy, to the point of counting the same thing twice from time to time—dubious accounting: double dipping in the same stream, contrary to what Heraclitus said was even possible.

How recursive can you get? The Kantian transcendental turn is pretty deep: What is necessary in order that X be possible—put in whatever you want for X, e.g., knowledge, science, peace, the state, money, love, you name it. The transcendental turn, the quest for the conditions for the possibility of some phenomenon-X, is a powerful trope. It gets you to thinking.

A useful way to *start* thinking about recursion and emergence is to ask: What is the *simplest* system in which a recursive topology works to bootstrap that system into existence? Biological anthropologist Terrence Deacon, in collaboration with colleagues from several different disciplines—Ty Cashman, Jeremy Sherman, Julie Hui, and myself on occasion—has been developing and testing this hypothesis. Synthesizing established principles in biomechanics, complexity theory and evolutionary biology, Terry has developed what appears to a growing number of scientists to be a plausible and sufficiently detailed theoretical model for the transition from matter to life, and from matter to mattering. In what follows, we're going to down a scientific rat-hole that some humanists will find uncomfortable. Buckle up, read slowly, and you might just learn something

The dauntingly simple challenge is to conceive of a plausible model for the emergence of life in a universe governed only by the established laws of physics and chemistry—and the Second Law of Thermodynamics that says that things tend to become less ordered more than more ordered. It's called entropy. How in the world can we get the comples order we call life in a world that is tending toward disorder?

We start with a universe of atoms bonding differentially to form molecules of diverse shapes. These molecules interact randomly with each other in water. Kinetic energy is sufficient to cause interactions, most of which would appear simply as molecules bouncing into and off of one another. Some interactions, however, would appear as a stickiness forming the molecules in contact with each other. The strength of these bonds is a function of the quantity of surfaces that come in contact with one another. The greater the surface contact at the interface between two molecules, the more likely, and the longer, the surfaces will stick together. We thus see differential bonding—some molecules sticking together so well that they would appear to have formed new stable molecules, others sticking together so negligibly as to appear to bounce off of each other.

Let's consider "soup" as a precursor to life. "Soup" is a molecular solution with a discernible flavor caused by the cultivated presence of certain molecules and the cultivated absence of other molecules, as opposed to a solution that contains any and all species of molecules. (I owe to Jeremy Sherman my embrace of this idea of "soup.") A God could certainly concoct a soup. An intentional agent could as well. The question here is whether chance interactions strictly within the sphere of physics and chemistry could concoct something like soup spontaneously. The first symptoms of pre-life to look for would be directional changes in the pattern and population distribution of certain molecular species.

A rare but non-negligible configuration of catalysts might arise. Imagine, for example, that the molecule formed when A + B combined (=AB) happened also to be a catalyst itself, that increased the likelihood of molecules C+D combining (=CD). Imagine also that CD also happened to be a catalyst that itself brought together molecules (E+F) and that the molecule thus formed (EF) happened to be the original catalyst in the chain, capable of catalyzing A+B. This kind of reciprocal catalysis would indeed be rare, but it would have an unusual effect on directionality. When such an "auto-catalytic cycle" formed, it would not only transform many molecules, it would do so at an accelerating rate. Each cycle would produce one more of each of the catalytic members. Since catalysts themselves are unaffected by their catalytic action, the catalysts that produced the additional catalytic members would remain available to catalyze still more molecules, transforming them into more functioning catalysts as well. The effect would be a self-amplification akin to compound interest, whereby that which is generated increases the rate at which still more can be generated.

Is this "soup"? Getting there perhaps, though realistically life is still a long way off. For one thing, such autocatalytic sets would be rare; for another, should the rare one occur there is nothing about it that would capture its formula to set itself in motion again should it be disrupted. There is nothing about the set by itself that could perpetuate it beyond its local and temporary directionality. The destiny for such a set would tend strongly toward depletion and degradation. Once any of the catalyze-able molecules (A through E) were depleted, the compounding effect would slow. At some point all available catalyze-able molecules would be transformed and the cycle would end, degrade and regress to an entropic state, not the least like soup. Once dissolved, the probability of this autocatalytic set reforming itself would be no better than chance. Unlike life, the growing catalytic set bequeaths no legacy, has no capacity to replicate itself again.

Nothing described up to this point exceeds the commonplace in simple chemistry, the sort taught in a freshman course. Stretching out from the inanimate world toward the animate, this much of the bridge from physics and chemistry over to biology has been well established. How to build the remaining bridge? That is what we'll take up next.

Auto-catalytic sets need not be as text-book simple as the one just described. They could arise with a combination of anabolic (headed toward greater togetherness) and catabolic (headed toward destruction) catalysts. Imagine that instead of an anabolic catalyst combining C+D to produce CD, a catabolic catalyst breaks apart a molecule (call it CD) producing C, which happens to be the catalyst that increases the likelihood of E and F uniting.

What then of D, the other product of this catabolic break-up? D might simply be a byproduct of the auto-catalytic cycle, a molecule that is neither affected by nor affecting any of the other molecules in the set. For the duration of the cycle's growth, molecules D would be generated with no effect upon the cycles' function. Imagine that one such byproduct, while not a member of the autocatalytic set itself, were nonetheless conducive to the growth of the set by means of increasing the local availability of participating molecules A through E. There is a candidate for this: Lipids.

Lipids are common elongated inorganic molecules that happen to be polarized in their affinity to water. One end of the lipid is hydrophilic, or highly water-bonding. The remaining length of the molecule (its 'tail') is hydrophobic (or non water-bonding which is to say it bonds more strongly to other molecules than it does to water molecules). Since the hydrophobic tail is more likely to stick to itself than to any of the surrounding water, a single lipid molecule collapses in on itself. But in quantity, lipid molecules align in relation to each other. The hydrophobic 'tails' straighten out and bond to each other along their lengths. This would leave the tail-tips—the points of greatest hydrophobia—exposed to water. They tend strongly therefore to align with each other, hydrophobic-tip to hydrophobic-tip. The resulting cluster of lipids forms what is known as a bi-layer. Two plies of lipids aligned to each other with their hydrophobic ends touching where the plies connect, and the hydrophilic ends facing into the water. The bi-layer forms in sheets that grow as they accumulate more lipid molecules, become unstable, and eventually curl in on themselves, spontaneously forming bubbles, or sacs, similar to those formed by oil in a well-shaken bottle of vinegar and oil salad dressing.

Imagine then the extremely rare autocatalytic set that happens to form lipid molecules as a byproduct. These lipids would tend to form into sacs enclosing in close proximity any available elements of the catalytic set. Would this increase the likelihood of each catalytic encounter necessary to the auto-catalytic cycling? In many cases, no; in some cases, yes. We can imagine lipid membranes capturing only some of the molecules necessary to the cycle, forming a barrier preventing access by the catalysts to their catalyze-able molecules. We can also imagine cases in which the lipid sac's formation with its random tearing and sealing, opening and closing, forcing proximity and interaction among the elements in the autocatalytic set which would also produce more lipid byproducts: A

growing constellation of interdependent molecules enshrouded within a chance protective sac.

Is it soup *yet*?

Again, local accretion in certain molecular species would result. Again, the greatest likelihood would be eventual depletion and dissipation, as the cycle churned through all available catalyze-able molecules. Still, it is possible that the lipid sac would preserve the integrity of the catalytic set, inadvertently buffering it against the vicissitudes of kinetic energy, entropy, and disruptive bonding.

Is it life yet? It grows. It has a sort of selfhood we would recognize as a function of its membrane. It has a potential for fittedness. That is to say, an outside observer could monitor the interactions among the molecules and identify which of these interactions increased the proto-self's chances of persisting and growing. But all increases in fitness would be strictly fortuitous; there would be nothing about this combination of molecules that has even the tiniest intent to preserve itself. As we have noted, the membrane itself could either increase or decrease the chances of the catalytic set persisting, either by containing within or secluding outside of its boundaries the molecules that would "run" the catalytic set.

Life evolves by variation and selection. Do these lipid enshrouded autocatalytic sets evolve? We can imagine a non-zero sum "competition" among several of them, each by happenstance consuming resources that might otherwise "cycle" each other. We can imagine those with the most fortuitous combination of autocatalytic elements and lipid sac structures being the "fittest" and thereby out-competing others. We can imagine others running out of raw ingredients and dissolving—the non-survival of the non-fit.

And the fittest—what spoils to the victor? None to speak of, and this because eventually it too would run through all available catalyze-able molecules. It would slow, degrade, dissipate and the probability of reforming would remain no better than chance. Though it is subject to competition, it is not subject to evolution, because it has no way to replicate its formulae for success, no way to pass on its pattern to subsequent "generations."

No, this is still not life. Not even sustainable soup.

Still, let's take this plausible but unlikely scenario one step further toward the improbable. Suppose that a few members of the auto-catalytic set had the tendency to bond to the surface of the lipid membrane. When the set depleted all available resources, it would break apart, and scraps of the lipid membrane would float away. A scrap that happened to have a few sufficiently representative molecules from the auto-catalytic set would become something of an auto-catalytic starter kit, in effect a seed that could restart

the set if that scrap should drift into a quantity of catalyze-able molecules. At this point—the point at which the interaction of the fortuitous molecular constellation is captured and represented by the starter kit—we have crossed a significant threshold in the steps toward life. For the first time, a combination of molecules that happened to be fortuitously sustaining would also have the markedly fortuitous capacity to record its formula for success. For the first time there would be a better-than-chance probability of the catalytic set's reformulating in a new environment. This is the advent of intergenerational legacy, something upon which evolution, not merely competition for resources, can act. And suddenly life seems not nearly as distant and unlikely a prospect.

Let us reflect on what has—and what has not—been achieved in this allegorical story of a path toward proto-life. From the random collisions and adhesions of an assortment of molecules—matter—the rudiments of dynamical form have emerged: an enclosure, a lipid sac that contains molecules A through E, thus maintaining their proximity, and increasing the likelihood of their auto-catalytic reaction cycle. Further, if the first containment is broken, the availability of a starter kit that has recorded the formula for success increases the likelihood of replicating those conditions for success. Here we have a system that can be called morphodynamic (*morphe=f*orm*; dynamic=*interacting through time*).* It can produce more of itself by virtue of the fact that the memory of the system is not so much of the *matter* but of the *form* of the autocatalytic cycle. What's carried forward is the habit, the pattern of the process, not the stuff. A topology of interactions gets reproduced. The auto-catalytic cycle binding molecules A through E will itself be catalyzed *again.* We have a double-loop system, an auto-catalysis of auto-catalysis—a reciprocal auto-catalysis among auto-catalysts, as in Manfred Eigen's cycles of cycles, what he called *hypercycles*. One is about space—the containment in a lipid sac; the other is about materials, the molecules A through E. When one cycle is weakened, as by the breakdown of a lipid sac, the other, in the form of the starter kit, will help to reconstitute the initial auto-catalytic cycle. So here we have a higher order morphodynamic cycle. We are in the presence of the origin of disposition, a directionality toward more of the same. This constitutes a second order morphodynamic process, a process that links morphodynamic processes. In this second order morphodynamic cycle we have a very interesting part/whole relationship: the starter kit is a part that helps to reproduce the whole, and the whole—the lipid sac and its contents—tends to reproduce the parts. This part/whole relationship is important because it prefigures the relationships between organism and environment and is thus one of the minimal conditions for the emergence of selection of some parts over others as a result of the affordances granted by an environment.

At this stage of the story we've derived form—approximating Aristotle's formal cause—from material and efficient causes. The random distribution of molecular substrates supply the material cause; the combination of thermodynamic jostling and catalytic

bonding supply the efficient causes. The resultant enclosure has a discrete form that sets it off from its environment as inside from outside, proto-self from other. But we do not yet have final cause. There may be a directionality toward the reproduction of more of the same, but there is no tendency or intention toward anything different. For that step to be accomplished, there must be conditions for variation and selective retention. There must be some way for this topology of mutually autocatalytic processes to achieve variation by substituting different substrates for the original set of molecules, A through E.

The topology itself can be embodied or realized differently. The same topology can be achieved by variance of the molecules drawn from the surrounding solution. The substitutions that persist are not just doing self-generation in the sense of replicating exact copies of the original. Instead they are variants that preserve the original topology, but with different substrates drawn from the environment, like so many different fonts (different substrates) spelling the very same words (topologies). The makeup of the surrounding solution will selectively favor some substitution sets over others. And now what you have is a new kind of part/whole relationship. You have the original auto-catalytic cycle and its general topology instantiated or realized in several different sets of substrates; second, over time, you have a differential favoring of one set of substrates over others; so now, third, you have a larger whole—part and environment—that are in a dynamical relationship with one another. The more successful parts exhibit their differential fitness to that environment.

Now, in addition to the original topology of the auto-catalytic cycle of molecules A through E, you have a larger cycle produced by recursive sampling. Recursion means that you bring the results of your last sampling to your next sampling, as in compound interest. Output becomes input. The consequences of earlier samplings are being sampled by later samplings.

In a simple cybernetic feedback loop like a thermostat, *we* determine a set-point—say 70 degrees Fahrenheit for the living room—from the outside. Here the topology is determined from the inside out through multiple levels of the part/whole relationships. With each disruption of containment, closure is broken. With various openings, there will be various samplings of different substrates. You now have reciprocal morphodynamic relationships that can create and be nested in higher order relationships. This process has the wherewithal to point to what it is not. Rather than simply reproducing identical tokens of a type, the process as described thus far—which is still only proto-life, not life—nonetheless has the rudiments of intentionality or aboutness in this primitive "pointing" from one assemblage of molecules with an autocatalytic topology to other assemblages with the same topology but different substrates. What is the "direction"? Very simply, towards *more.* But the reason we can talk of directionality rather than simple replication of exactly the same is that the variation of substrate molecules introduces a degree of

difference such that the system is moving *from here* to *there* rather than from here to here to here to here again. Differential selection from among slight variants exhibiting identical topologies allows the system to capture the amplifications of self organizing form and matter.

What brings function into being—the story of the origin of functionality—is something that the functioning entity *is not yet:* a state of things that does not yet exist. The beginnings of ends—the origins of purposiveness— lies in the unlikely, but non-magical, accretion of causal topologies, starting with the thermodynamic shuffling of molecules, adding the morphodynamic layer of autocatalytic chemistry and physical containment, and then a further process of recursive sampling to produce a representation of what is *not yet* in what already *is.*

By nesting the autocatalysis of morphodynamic autocatalyses inside the temporal process of a system that can recursively sample and select for biases, this account shows how a relatively simple system of proto-life can project itself toward successively more successful instantiations of its initial topology.

At this simplest level of autocatalytic closure, recursive dynamics are crucial. A + B produced AB, which catalyzed C + D to produce CD, which catalyzed E + F to produce EF, which catalyzes the original reaction, A + B. Meanwhile the byproduct, D, turns out to assist in the creation of a lipid sac that holds all of the ingredients in proximity with one another, thereby increasing the likelihood of their autocatalytic reaction cycle. There is a double closure: not only is the reaction cycle a closed loop; there is a closed sac to increase the speed and sustainability of the reaction cycle. And no one designed it. It is a self-organizing system.

We'll come back to this allegory of self-organization in later sections—in particular, the sections on *desire* later in Part One, and the section on *language* in Part Two. This story has many parts, beginning with Kauffman's buttons, which first appeared in relation to Trait Two—*popping.* We're now discussing Trait Four, reflexivity. Reflexivity—recursion—is a basic trait of emergent systems, as this allegory shows. Autocatalysis is part of the story, and with the help of recursion, assisted by the enclosure of a lipid sac, we get autocatalysis of autocatalysis—a double recursion. And there's more.

In his naturalistic account of emergent systems, Terrence Deacon has shown how purposiveness can accrue from non-purposive components. This is an immense achievement that has vast import for a science of emergence. In several different papers and books, Deacon has given a careful account of three levels of emergence that are most briefly described as follows:

First-Order Emergence: Properties emerge as a consequence of shape interactions. Example: The interaction of water molecules (nothing but) generates a new property, surface tension (something more).

Second-Order Emergence: Properties emerge as a consequence of shape interactions played out over time, where what happens next is highly influenced by what has happened before. Example: The formation of a snowflake, where initial and boundary conditions become amplified in effect over time. In general, complex or "self-organizing" systems display second-order emergence.

Third-Order Emergence: Properties emerge as a consequence of shape, time, and "remembering how to do it." Example: biology, where genetic and epigenetic instructions place constraints on second-order systems and thereby specify particular outcomes called biological traits. These traits then become substrates for natural selection by virtue of the fact that 1) their instructions are encoded and 2) they endow organisms with adaptive properties. (Terrence Deacon, "Memes as Signs in the Dynamic Logic of Semiosis: Molecular Science meets Computation Theory," Note 2. See also Deacon, "The Hierarchical Logic of Emergence: Untangling the interdependence of evolution and self-organization," in B. Weber and D. Depew (ed.), *Evolution and Learning: The Baldwin Effect Reconsidered,* MIT Press, 2003.)

Deacon has established the conditions for the possibility of the beginnings of ends—the origins of teleology—in a naturalistic, up-from-the-bottom, no-homunculi-needed account. But this account is an allegory, a story, and not yet science. The precise mechanisms by which molecules of DNA refer to different reactions among proteins have not been established. Nor have the precise mechanisms by which symbols refer to complex constellations of things like justice, youth, transportation, or disjunction. But this three-stage logic of emergence—also called (1) thermodynamic, (2) morphodynamic, and (3) teleodynamic—in a non-magical way, the pre-conditions for life. To foreshadow later parts of the account, this allegory of recursion lays a foundation for the later accounts of desire, intentionality, reference, and language. Deacon's powerful analysis will recur later in this text, but for now we'll leave him to cite further varieties of recursive experience.

Now there, that wasn't too hard.

To further illustrate the workings of reflexivity at this stage of the overall argument, let's move up many layers of biological complexity to Gerald Edelman's research on the neuro-anatomy of the brain, and on his development of what he calls *re-entry.* In order to clarify this single part, re-entry, we need to see its place in the context of the whole: How it functions as part of an account of consciousness. As suggested earlier, consciousness,

like life, strikes many as a mystery. As with confronting the mysteriousness of life, we should take these reports of mysteriousness not as signs of danger to inquiring minds, but as evidence we're looking in the right place—down the block in the dark where the drunk actually lost his keys, to invoke an old joke, not under the street lamp at the corner where the light of ordinary science lets us think we can see better. The fact that consciousness studies should yield up a group of theorists whom Owen Flanagan could call "the New Mysterians" (Owen Flanagan, *The Science of the Mind,* MIT Press, Cambridge, 1984 and 1991.**)** tells us that we are in the presence of an emergent system. If it's an emergent system, it *should* look mysterious to the first eye of ordinary science. It *should* require the second eye of a science of emergent systems to provide a satisfactory account.

Why is it that some events in the brain are conscious and others, like the management of blood pressure, are not? Further, what is it about the neuro-physiology of the brain that makes it capable of supporting not only those unconscious processes like the management of blood pressure, but also the subjective *experience* of consciousness as well?

Edelman's answer lies with what he calls *re-entry*. The appeal of demonstrating the fourth trait of emergent systems with Edelman's research lies in the fact that, like Deacon's description of recursive topology at the bottom of the ladder at the level of proto-life, Edelman can give a clear and scientifically verifiable description of re-entry that does not rely only on images of kissing elbows or rolling one's eyes back to see the optic nerve. Edelman is able to isolate specific bundles of neurons that link some brain functions to other brain functions in loops that connect some functional clusters of neurons to others in patterns and structures that support the *experience* of consciousness.

Experiments using electrodes to record directly from neural cells in animals have demonstrated that short-term temporal correlations can be found within single areas of the brain, as well as between different areas. In some cases, it has even been demonstrated that short-term temporal correlations between the two cerebral hemispheres are due to direct reentrant interactions. If the millions of reentrant fibers connecting the hemispheres are cut, these short-term correlations vanish. We take these findings as direct evidence that integration and rapid functional clustering occur in the thalamocortical system *and that reentry is the major mechanism by which integration is achieved."*
(Gerald Edelman and Giulio Tononi, *A Universe of Consciousness, op. cit.,* p. 124. Edelman and Tononi are also very clear about the need to account for the subjective *experience* of consciousness. See, for example, pp. 2, 10-19, and p. 157: "In no case can a theory or description substitute for an individual's experience of a quale, no matter how correct such a theory is in describing its underlying mechanisms."*)*

As we have emphasized, if there is one central structural principle that underlies the appearance of consciousness, it is the emergence during evolution of new anatomically based reentrant systems. (*Ibid.,* p. 109.)

Edelman and others have uncovered the neuro-anatomical correlates, the platforms, for the subjective experience of consciousness. What is becoming increasingly clear is that the old afferent-efferent architecture that we learned in school is a misleading model for thinking about thinking. It's a mistake to think of consciousness in terms of a stimulus-response arc that begins with input (afferent) to a central processor (homunculus) and then back out to output (efferent). That model leaves us in a state of mystification regarding what goes on inside the black box of the so-called central processor.

Once we zoom in on the so-called central processor, what we discover is that it is not central at all. However apparently singular the *experience* of consciousness may be—that sense of first-person singularity, that sense of being an *I* and not a *we*—the neuro-anatomy of consciousness is in fact a fairly distributed process involving millions upon millions of neurons. As Edelman and Tononi repeat several times over in slightly different ways:

> Our analysis leads to several conclusions. First, conscious experience appears to be associated with neural activity that is distributed simultaneously across neuronal groups in many different regions of the brain. Consciousness is therefore not the prerogative of any one brain area; instead, its neural substrates are widely dispersed throughout the so-called thalamocortical system and associated regions. Second, to support conscious experience, a large number of groups of neurons must interact rapidly and reciprocally through the process called reentry. (Edelman and Tononi, *op. cit.,* p. 36. Cf. P. 44 for reference to, "the lack of a central coordinative area." p. 55: "Every conscious task involves the activation or deactivation of widely distributed areas of the brain." And p. 115: "No superordinate area coordinates the responses of the model.")

What makes us conscious, then, is not some specifiable, locatable seat of consciousness, but instead a process that Edelman calls reentry, a process that qualifies as recursive or self-reflexive.

What is becoming increasingly clear from the writings of Edelman, Damasio and others is that so-called consciousness should not be considered as a simple on/off, digital system. Instead, there are degrees of consciousness. While some minimal degree of self-reflexive reentry is necessary to support what both Edelman and Damasio call "primary consciousness"—as in the case of an awareness of the color red, for example—further loops of self-reflexivity are required for the kind of higher consciousness associated with a sense of self or a narrative sense of self-identity.

> As in the case of primary consciousness, a key step in the evolution of higher-order consciousness was the development of a specific kind of reentrant connectivity, this time between the systems for language and the existing conceptual regions of the brain. . . When narrative capabilities emerged and affected linguistic and conceptual memory, higher-order consciousness could foster the development of concepts of the past and future related to that self and to others. (*Ibid.,* p. 195. For a graphic illustration of the addition of further loops, compare Figure 15.1 on p. 194 to Figure 9.1 on p. 108.)

This much said by way of re-introducing some degree of complexity, now consider the contrast between emergent systems and non-emergent systems. In the world of non-emergent systems, we are fond of saying things like, "You cannot lift yourself by your own bootstraps." For emergent systems, on the contrary, if they don't reach down and self-reflexively lift themselves by their own bootstraps, then they are not emergent systems at all, but the predictable consequences of some prior cause.

Edelman and Tononi are explicit in the use of this image of bootstrapping:

> The short-term memory that is fundamental to primary consciousness reflects previous categorical and conceptual experiences. The interaction of the memory system with current perception occurs over periods of fractions of a second in a kind of bootstrapping: What is new perceptually can be incorporated in short order into memory that arose from previous categorizations. The ability to construct a conscious scene is the ability to construct, within fractions of seconds, a remembered present. (Edelman and Tononi, *op. cit.,* p. 109. This process, and its description, bear comparison to what Dennett describes using his metaphor of "drafting" and "editing." Cf. *Consciousness Explained, op. cit.*)

Doug Engelbart, the genius who invented the computer mouse and much else that went into today's PCs, spent the final years of his astoundingly productive career on what he calls his Bootstrap Institute, the mission of which is to improve our processes of improvement—an appropriately self-reflexive dynamic.

Zooming out, now, from the single cell in Table 1 where the level of consciousness intersects with the column under the fourth trait that says, "All Emergent Systems are Self-reflexive," let's see whether the rigor that Edelman and Tononi give to self-reflexivity in their study of the neuro-anatomy of consciousness can illuminate similar dynamics on other levels. We'll want to list and describe all the forms of reflexivity from self-referentiality to self-contradiction; from feedback to reentry; from "strange loops" as described by Douglas Hofstadter in *Gödel, Escher, Bach,* to self-consciousness; from apperception as described by Immanuel Kant to introspection as described by William James; from being "self-contained" to auto-catalytic closure; from autopeoeisis as described by Umberto Maturana and Francisco Varela to Manfred Eigen's hypercycle.

Consciousness cannot be explained absent several kinds and several layers of recursivity If you combine Antonio Damasio's models of different levels of consciousness and awareness together with Terrence Deacon's multi-level model of mind, you already find four or five distinct layers, below which lie Edelman's description of reentry at the neuro-physiological level and above which lie further forms of reflexivity, from Sartre's analysis of the For-itself at the level of the individual to Anthony Giddens' descriptions of "reflexive sociology" at the level of society. Trying to grapple with the concept of consciousness with anything less than a full toolkit of 'reflexors' would be as hard as trying to disassemble and reassemble a motorcycle without a full set of wrenches.

Dan Dennett is moving in the right direction in his description of consciousness when he works the analogy with *editing.* Editing is a fairly loopy activity, e.g. in the medium of photography. First you take the picture. Then you develop the picture. Then you look at the picture. Then you arrange different parts of the picture, or different pictures, into "the edited version" —as distinct from "the first draft."

Consider the medium of writing. When you edit your own writing you make corrections. You change things. You delete some words, you add others. You strike out whole sentences, whole paragraphs. The final draft may be quite different from the first draft. All of that said, it remains the case that the so-called first draft is not altogether *raw* or *un*edited. The words are in a language that was learned. That language is part of a culture that has evolved. Both language and culture operate as editors of the ideas and impressions in even the "first" draft. There is, in a sense, *no first instance* of editing. There are many edits and none is a purely rough cut. Every draft is, to use Derrida's oft-used phrase, *always already* edited.

This is why 'consciousness,' the word, has so many meanings, from waking as opposed to sleeping, in one very frequent usage—" 'Ah, she's conscious,' said Dr. Tim as he leaned over little Sally's bed and noticed as she finally showed signs of waking from her coma"—to 'consciousness' as the Marxists used to talk about it in such contexts as: "Only praxis will break through the false consciousness of the Bourgeoisie." *False consciousness* isn't a feature describable in neuro-physiological or even solely psychological terms. It is only graspable on a cultural, political, ideological level. It is a kind of involuntary blindness to injustice induced by a subtle blend of privilege and indoctrination, imposed under the guise of a high quality and often very expensive education.

Consciousness occurs in different forms, on different levels, all up and down a hierarchy of levels of emergence. It may not be possible—or even necessary—to locate and identify a lowest level of consciousness, a first instance. Indeed the attempt to do so would

indicate either (a) that we aren't talking about an emergent system, or (b) we do not understand the traits of emergent systems, the very first of which states: No first instance.

Francisco Varela's *Principles of Biological Autonomy* stands as a pioneering effort to break away from attempts to explain the origins of life by some prior cause. His concept of *auto-poiesis* is clearly aimed at showing how a chain of molecules can 'reenter' itself to create a membrane, and how a chain of cells can 'reenter' itself to create an autonomous organism.

Likewise, the work of John Holland and Stuart Kauffman at the Santa Fe Institute has been devoted to unpacking the bootstrapping potential of what they call 'auto-catalytic sets.'

> In short, among the possibilities already demonstrated are modestly complex reaction networks of autocatalytic and cross-catalytic peptides. For example, A might catalyze its own formation as well as the formation of B, while B catalyzes its own formation and the formation of A as well, in a catalytic structure named a 'hypercycle' by Nobel laureate Manfred Eigen and Peter Schuster in 1977. . .
>
> So, in the American vernacular, self-reproducing molecular systems are a done deal.
>
> Now what?
>
> Now much. . .
>
> Now the first hints of a new technology based on self-reproducing, evolvable molecular systems.
>
> Now the hard, hard push to explore a terra nova . . .

Namely, the science of emergent systems and its several traits, without which any one of the traits seems paradoxical at best, nonsensical at worst. After all, *things just can't raise themselves by their own bootstraps! Levitation is not possible.*

But life *has happened.* Consciousness exists, even if there was never a first idea. Languages exist, even if there was never a first word, nor a cause of the first word on the basis of which the first word was perfectly predictable. And, if we launch our imaginations on the wings of analogy and homology, we see instances of economic behavior that defy the logic of a simple equilibrium between supply and demand. We see cases of "increasing returns," to use Brian Arthur's phrase, examples like the contemporary art of Basquiat or Jeff Koons where the higher the price, the higher the demand. Hits happen, and it takes an appreciation for Brian Arthur's theory of increasing returns in order to understand *How Hits Happen.* (See Winslow Farrell, *How Hits*

Happen, HarperBusiness, New York, 1993, with a Foreword by Brian Arthur of the Santa Fe Institute.)

In all of these cases, on all of these different levels of emergent systems, we find different species of the genus *self-reflexivity.* These ideas aren't new. Think back to the symbolism of the *ourobouros,* the snake swallowing its own tail. Think of philosophers like Jean-Paul Sartre claiming that consciousness is not possible without self-consciousness. (Cf. Jean Paul Sartre, *Being and Nothingnes.*) Think of the weight borne in German idealism (and picked up by Sartre) by the distinction being *an sich* (in itself) and being *für sich* (for itself). It is this same burden of paradox that is picked up by Varela and Kauffman in their attempts to unlock the illogic of autonomous agents. *How can they get that way? How can they emerge?*

The burden of this book is to lift this very old but very new way of thinking by *its* own bootstraps by way of demonstrating its systematic and therefore self-supporting structure; to show how its several traits mutually entail one another, and in that mutual entailment, provide a web of support sufficiently strong to lift our Enlightenment-soaked intellects out of a paradigm of monological causal explanation and into a paradigm of cybernetic, systemic, evolutionary understanding.

To summarize the mutual entailments of the first four traits: Emergent systems don't have first instances because the whole must be present to lift itself, suddenly, by its own bootstraps before any part could claim the role of first. So it goes with words and languages, with thoughts and minds, with hits and markets.

Now, what can we say about the issues of prediction and reduction?

I.5 Fifth Trait: Emergent Systems are unpredictable from the features of their precursors.

I.6 Sixth Trait: Emergent Systems are irreducible to descriptions of their component parts.

This section, casual readers bewarmed, is fairly heavy on philosophy of science. But, hey, it doesn't hurt to learn something. The next section on Desire is juicier, but I encourage you to read this one despite the challenge.

The fifth trait and its symmetrical companion on irreducibility are the most familiar features of emergent systems. The problem has been that, up to now, defining emergent systems by these features *alone* has done little more than exile emergent systems from the realm of the discussible. If causal explanation under covering laws is the only way to give

a satisfactory account of some phenomenon, then the failure to conform to that model is a failure to provide understanding.

What, more precisely, is meant by a "covering law"? What is the relationship between causality, prediction and explanation? In order to explain complex phenomena like biological growth or human thought, it was thought necessary and sufficient to reduce those phenomena to their physical, constituent, simple parts and then plug state descriptions of those parts into equations representing well confirmed general laws whose generality "covers" all particular instances. One way to confirm general laws is to test their predictive power. Thus, if laws L1, L2. . .Lr enable one to predict events of type E from antecedent (initial) conditions C1, C2. . ., Ck, then those laws are confirmed, and event E can be scientifically explained. In a canonical statement by Carl Hempel and Paul Oppenheim:

> If *E* describes a particular event, then the antecedent circumstances described in the sentences C1, C2,. . ., Ck may be said jointly to "cause" that event, in the sense that there are certain empirical regularities, expressed by the laws L1, L2, . . ., Lr, which imply that whenever conditions of the kind indicated by C1, C2, . . ., Ck occur, an event of the kind in *E* will take place. (Carl G. Hempel and Paul Oppenheim, "The Logic of Explanation," *Philosophy of Science,* 15, 1948; reprinted in *Readings in the Philosophy of Science,* ed. Feigl and Brodbeck, New York, 1953, p. 32.)

This is how the first eye of monological science sees the interrelationships that bind together and mutually define the concepts of causality, law, prediction, and explanation. Just because we are now so familiar with this legacy of the eighteenth century Enlightenment, it can be difficult for some of us to imagine how (or when) anyone could have thought otherwise. *Of course* causes are "antecedent" to their effects. *Of course* explanation works by subsuming an event under a covering law, one of "certain empirical regularities." *Of course* prediction is possible "*whenever* conditions of the kind indicated by C1, C2, . . ., Ck occur."

To repeat: There is nothing wrong with this codification of monological science. It is not mistaken; merely incomplete. It is insufficient for an understanding of emergent systems and, if understanding limits itself to a Cyclopean vision that sees *all* phenomena through this single lens, then emergent systems will be either invisible, or their real nature will be distorted by attempts to explain them away by reducing them to descriptions of their simpler components or precursors.

The rhetoric of reductionism—the frequent use of the formula,

x is *nothing but* a, b, c—

makes some people uneasy. Cherished beliefs are cast in doubt by this formula. Think of some famous examples where reductionism amounts to debunking:

- Justice *is nothing but* the will of the stronger (Thrasymachus in the early books of Plato's *Republic*)

- Man *is nothing but* a descendent of the apes (Darwinism as seen through the eyes of the Christian Right)

- Truths are *nothing but* illusions whose origins we have forgotten (Nietzsche)

- Love *is nothing but* lust with an overlay of romance, or, as Paul Goodman once put it in conversation, "Love is a form of pathology from which we all emerge with luck."

- Mind *is nothing but* an epiphenomenon of the brain—the whistle on the train that contributes nothing to locomotion; the so-called *mental is nothing but* a read-out of physical interactions, as they say in the trade, its "neural correlates."

The rhetoric of reductionism should not be confused with the logic of reduction. The latter can be quite beautiful, as in Lavoisier's "reduction" of the apparently different phenomena of breathing, rusting, and combustion to the process of oxidation; or Boltzmann's reduction of the thermodynamics of gasses to statistical mechanics. Monological science has its moments of glory.

In sorting out the relationship between reductionism and emergent systems, it is important to be clear about this difference between the *logic* of reduction and the *rhetoric* of reductionism because we don't want to throw out the baby of the *logic* with the bath water of its historically associated *rhetoric*. Nor, conversely, do we want to follow the course of 20th century logical empiricism in over-extending the rhetoric of reductionism just because we're so impressed by the logic of reduction.

In what follows, it will be important to render unto monological science those realms of explanation it rightly owns, even as we defend monological science against attacks on its adequacy. Dennett is absolutely right to observe that, "Consciousness . . . leaves even the most sophisticated [monological] thinkers tongue-tied and confused." If all you have is a hammer, then everything looks like a nail. If all you have as a tool of explanation is reduction, then every explanation will have to be reductionistic. And you will be tongue-tied and confused when confronted with emergent systems that are irreducible to descriptions of the interactions of their component parts.

The strategy here is to untie the tongue by, at first, simply pointing to a number of instances of emergence that, by their range and variety, provide convincing existence

proofs for the possibility of the emergence of the unpredictable and irreducible. (See again the list of things that go *pop*.)

Classical rationalism postulated *the principle of sufficient reason.* Nothing happens without sufficient reason. Therefore, if something does happen—event E— then it should be possible for science to discover what that sufficient reason is. And once having discovered that cause, if it is indeed sufficient, then it would serve as sufficient grounds for *predicting* E. Whensoever and wheresoever those grounds (or initial conditions) occur, E *will* follow as night from day. To suggest that E might erupt in a way that is unpredictable, or to suggest that the occurrence of E is *not* reducible to its precursors or components, is to flout the principle of sufficient reason and open the door to magic and irrationality. Therefore—or so goes the argument against the possibility of emergent systems—emergence *can't happen.*

Only the predictable can happen—so goes the rationalist argument—because only the predictable can be explained by reduction. Reduction is the only valid form of explanation; so if it's not predictable, and not reducible, then it's not explicable and, according to the principle of sufficient reason, it couldn't have happened.

To those who think that emergence *doesn't happen,* that each particular instance of so-called emergence can be explained away by the closer application of monological science, it's worth replying: *Yes, a monological explanation may be possible . . . but it will amount to an epicycle on the old system. Just as there is not one planet but several calling out for a Copernican revolution on the old Ptolemaic system, so there are so many instances of emergent systems that it will prove more elegant to accept a science of emergent systems rather than to keep adding epicycles to the old monological model.*

Seventh trait: Desire

Desire happens. But what is it, and how does it help us to understand emergent systems?

As above, we begin with a fairly long list of plausible candidates, different instances of Desire on different levels of emergent complexity. Next, a quick literature review indicating the historical range of approaches to the phenomena of desire. Equipped with a vocabulary derived from the literature review, we return to the list of candidates to say a few more words about each one. The point is to show the similarities and differences among the various levels of desire, and the role desire plays in both breaking equilibrium at one level and then bonding the components of an emergent system at the next higher level of complexity.

This duplicity of desire—both bending equilibrium on one level, breaking it by transgressing its bounds, then seeking fusion on another level, the emergence of

something like love—this dialectical structure of desire makes desire a potential explainer of just about anything. While such explanatory power might at first seem appealing, a moment's thought should show that we run the risk of intellectual snake oil—Dr. Dizmo's elixir that will cure everything from cancer to hemorrhoids if you just buy and swallow enough of it. Rather than a simple vector in one direction in explanatory space, a dialectical understanding of desire will require at least two if not twenty vectors to describe its structure. So whatever you want to explain by finding a causal arrow to push in its direction, surely in the dialectical maelstrom of vectors that dance the dance called Desire you'll be able to find at least one to act as cause of just about any effect.

No wonder desire wears the rank of Principle. It truly does rival entropy in its reach and importance. Already in this penultimate section of Part One it should become clear that desire, the seventh trait of emergent systems, is as important to our understanding of the world around us as the second law of thermodynamics, the law of increasing disorder, or "entropy." But can we describe the role of desire with the same rigor that Claude Shannon gave to entropy in information theory?

Things fall apart; things come together. Is that *all* we're saying by talking about entropy and desire? If that were the whole story, then we would not have made any advance beyond the pre-Socratic philosopher, Empedocles, who maintained that "love and strife" were the two fundamental forces in the cosmos. This kind of "science" won't put a man on the moon or cure cancer.

The great achievement of complexity theory over recent decades has been to show how self-organizing systems can overcome entropy. But complexity theorists have not invoked anything like desire—for good reason given the preference for pushes rather than pulls, efficient causes rather than final causes, ever since Spinoza exiled teleological reasoning from Enlightenment rationality. The task of this section, then, is to assemble many of the candidates for explanation by desire—a list of the phenomena—together with a review of many of the theories developed to explain those phenomena, then to sort the salvageable and insightful from the wishful and magical.

Let's begin with another list, starting with the relatively simple and working through the list toward higher and more complex instances of desire.

(a) A uni-cellular organism moving upstream in a glucose gradient
(b) Evolution of reproduction from mitosis to sexual reproduction
(c) The role of desire in the development of consciousness
(d) The limbic system and the evolution of personal identity
(e) Consumer demand
(f) Values as decision criteria
(g) Social hopes: what do *we* want?
(h) Teleonomy: what does technology, or information, or the earth *want?*

(i) The sublimation of libido into love

As with the list of things that *pop*, this list represents a range of phenomena described in terms that are already theory-laden. What are some of the theories that have been developed to make sense out of the various phenomena of desire?

- As already mentioned, as early as the 5th century B.C. the pre-Socratic philosopher, Empedocles, identified "love and strife" as the most fundamental forces in the cosmos.

- Echoing Empedocles, Freud singled out Eros and Thanatos (sometimes called "the pleasure principle" and the "death instinct") as fundamental principles for explaining psychic life.

- In his *Symposium,* Plato put into the mouth of Diotima a "ladder of love" leading from base desires up to the love of wisdom —philosophy.

- Abraham Maslow's hierarchy of needs

- Ludwig Feuerbach is famous for saying, "You are what you eat," well before health food stores carved it on their walls.

- Hegel invoked desire as crucial to the development of consciousness; desire motivates the dialectic of mastery and slavery.

- Arthur Schopenhauer, in his *World as Will and Idea,* introduced volition as a complement to a purely cognitive interaction with the world.

- Nietzsche made much of what he called "the will to power."

- Heidegger introduced the concept of *care — Sorge —* as distinctive of human being-in-the-world (*Dasein).*

- Ernst Bloch wrote a three volume work, *Das Prinzip Hoffnung, (The Principle of Hope)* that rings the changes on many of the ways that hope and desire move the human spirit.

- Jacques Lacan, Julia Kristeva, and many other French intellectuals thematize desire as the red thread running through much of their writings.

- Judith Butler, in *Subjects of Desire,* and Martha Nussbaum, in *Upheavals of Thought,* offer a much needed feminine (if not feminist) perspective on the dialectics of desire from Hegel to French deconstruction.

Each of these lists could be much longer. For example, the first list could begin with magnetism and gravity. But as will be argued later in the chapter on *the phallusy of misplaced physics*, it may not be true that the simplest elements have the greatest range of

explanatory power, as if love, down deep, were magnetism in the same way that wood, down deep, is molecules arranged in fibers.

The strategy of assembling these lists has less to do with trying to find some single element or principle from which all others might be deduced, and more to do with cataloguing the range of phenomena and theories that lead us to believe that there's *something* worth puzzling over. In addition, the list of theorists should convince us that we're hardly the first to puzzle over these phenomena, and that many minds have found it useful to think about desire and its place in the world. And puzzled we should be. Again, we should take this puzzlement as evidence that we're looking in the right places, precisely where monological science comes up short so that a science of emergent systems has work to do.

As noted earlier, Abraham Maslow is famous for his hierarchy of needs, from survival needs, through security needs, to the need for esteem, and finally, the need for self-actualization through peak experiences. This section presents a hierarchy of desire that builds on and supplements Maslow's hierarchy of needs. Needs are universal. All god's chillun' need shoes, food, and water . . . All God's chillun' do not need anchovies, or SUVs, or socket wrenches. They may *want* anchovies. They may desire SUVs. And a particular mechanic may *need* a socket wrench at a particular moment. But the range of particular desires differs from the range of needs in not sharing the same degree of universality.

Desires tend to be highly differentiated, as in the desire for just one of 7 different varieties of mustard on the supermarket shelf. This particularity of desire has important consequences for economics and the kind of story we'll want to tell about the transition from an industrial economy, where increasing productivity is important for increasing the satisfaction of universal needs, to an information economy where information will assist in the identification and gratification of many different wants, whims and desires.

This much said by way of introduction, let's now work our way through the above list several items at a time. Expect a few digressions on method for, once again, we'll find that the attempt to explain some of these semi-mysterious phenomena will cause us to think again about causality and *what it is to offer a satisfactory explanation of anything at all.*

(a) A uni-cellular organism moving upstream in a glucose gradient

Let's begin to unpack the seventh trait and give it the rigor and specificity that Shannon gave to entropy by starting at the top of the first list with a uni-cellular organism

swimming upstream in a glucose gradient. Here we're already dealing with an emergent system—life. And already we're having to be very careful about our language. For is it really true that such an organism *wants* or *desires* more sugar? Wouldn't we be anthropomorphizing this uni-cellular organism if we said it *desired* anything? Doesn't the intentionality of desire—somebody desiring something—presuppose the intentionality of consciousness, that is, consciousness as always conscious *of something?* (The concept of the intentionality of consciousness as consciousness-of-something was first explored by Brentano and Husserl. The idea was later developed by Maurice Merleau-Ponty in his *Phenomenology of Perception.)* Sure, the organism behaves *as if* it wants more sugar. But maybe there's a perfectly mechanical explanation for its tendency to swim upstream in the gradient. There's no *subject* desiring an *object.* Just one object interacting with other objects according to the laws of physics and chemistry, or so goes the account of monological science.

Well, yes, we don't want to impute to this tiny organism a degree of consciousness that would render it more than an automaton. The very name, automaton, suggests patterns of behavior that are automatic rather than deliberative. But at the same time it's useful to see just how far down something *like* consciousness can be found. Already at this relatively low level of complexity, we see something like choice: Upstream or downstream? And we can make an argument from Darwin that those organisms that "choose", however automatically, to swim downstream are less likely to survive and reproduce than those that turn upstream.

(b) Evolution of reproduction from mitosis to sexual reproduction

Okay, they don't *choose* any more than a sunflower chooses to tip its petals toward the morning sun. (There's that tendency to anthropomorphize sneaking in again!) They just do it because the molecules of their membranes are set up in such a way that their cilia direct them sunward in the gradient. But they survive, and the ones that were set up to send them into shadows did not. Does that result justify us in saying that the ones that survived *wanted* to go sunward? Or that those paramecia *wanted* more glucose?

Consider a similar question at a higher order of complexity. When organisms engage in sexual reproduction rather than mitosis, does the little word 'sex' imply that we're witnessing the origins of romance?

(c) The role of desire in the development of consciousness

When we come to the level of consciousness, at however primitive a level, then desire appears to be crucially ingredient in making the source of its intentionality into

something we would call conscious. If you take the triad, C-desires-q, then the wanting of q is crucially ingredient in the definition of C as a conscious being. Simply *seeing* q or *sensing* q is not enough. Object of desire, q, can have an effect on C without C ever desiring it or being conscious of it. But once C *wants* q, then there is the emergence of a subjectivity that is also able to *see* q. (This point will be developed at greater length below in the section on "Hegel's depiction of desire.")

(d) The limbic system and the evolution of personal identity

In *A General Theory of Love,* three psychiatrists show how the definition of self-identity is largely based on the influence of the limbic system, the second layer of the tri-partite brain that sits above the reptilian brain stem, but nestles inside of the more advanced neo-cortex. The limbic system picks up pheromones. At the level of the limbic system, we sort out our likes from our dislikes in a way that has almost nothing to do with the complex judgments of the neo-cortex. You are drawn toward some people and repelled by others for "reasons" at the level of the limbic system. "Reasons" deserves quotes just because the limbic system is not rational. It is emotional, and important as such. Reptiles eat their young. They feel no limbic attachments to their young because they lack a limbic brain above their reptilian brain.

Antonio Damasio shows the importance of emotions in the development of consciousness. In *The Feeling of What Happens: Body and Emotion in the Making of Consciousness,* Damasio, a neuro-physiologist, shows how consciousness evolves in a series of self-reflexive stages in which feelings play a crucial role. Once consciousness comes on the scene, then, for us observers of the behavior of a conscious human being, it's easy enough to say that we're looking at the operation of desire. In the presence of a fully developed *anthropos* we're no longer so uneasy about indulging in *anthropomorphism*— treating that being as if it has the form—*morphe*—of *anthropos,* man.

But hold on a minute: If part of our effort in understanding emergent systems involves acquiring a better understanding of consciousness—that which distinguishes an automaton from a conscious human being—then it won't do to say desire is distinguished from animal need by the presence of consciousness, because we don't yet have an adequate understanding of what consciousness *is*. Indeed, by invoking desire as the seventh trait of emergent systems (like consciousness), we're hypothesizing that consciousness may emerge as a satellite of desire rather than the other way around. We cannot *assume* consciousness as a condition for anthropic desire if we are also assuming desire as a condition for consciousness. Or can we? Might not consciousness and desire mutually presuppose one another as parts of an emergent system in which neither is *first?*

Even in this first (phenomenological, inductive) pass on the syntax of desire, prior to exploring its semantics, its (theoretical, deductive) applications, we're bumping up against a Copernican revolution in the understanding of consciousness—the relocation of consciousness from the center of the universe of intelligence to its status as a satellite of desire. In the earlier "Ptolemaic-Cartesian" universe, we might imagine a representational consciousness that represents many possible objects of desire and then chooses one, thus generating desire from a representational classification system that divides the external world into two classes: objects of desire (e.g., food or sex) and everything else. Such a classification system should work to motivate a uni-cellular organism to swim upstream in a glucose gradient because that way it will get more food. Trouble is, a uni-cellular organism doesn't have the machinery of representation that would be necessary to make this a plausible story. That's why we're uneasy attributing a desire (that presupposes representational consciousness) to a uni-cellular organism.

But if we turn the tables, if we embrace a Copernican revolution that says that desire is a precondition for consciousness, then, when we come to explore the role of desire in the making of consciousness, we can't assume that we know what desire is. We may have to go back down the evolutionary tree to get a simpler, clearer picture of the nature of desire before we presume to know what desire *is* in the *making of human consciousness.* And on that descent back down the evolutionary tree, we're no longer anthropomorphising, because we've admitted that we don't yet know the form of anthropos. We're just phenomenologizing—just looking around to see what we can see, without yet being sure of just which categories we should use to classify what we see.

In a sense we're following a strategy very similar to the pioneers of the artificial life movement. They admit that they don't really know what life is in those organisms we find it easiest to call animate, so they set about to discover the rules that will animate inanimate matter. Here we admit that we don't really know what consciousness is, even in those organisms we find it easiest to call conscious, so we'll set about to discover the rules that will smarten up dumb matter. And when we do so, we'll find that desire—emotion, feeling, yen, urge, eros, hunger, will, care, love—all of the somewhat mysterious referents on the earlier list of phenomenologies of desire, are crucially ingredient in, and not just results of, the evolution of consciousness.

Just in case you think that the items on that list are not "somewhat mysterious," think about love. How many books have been written on the mysteries of love, starting with Ovid's *Metamorphoses* and Plato's *Symposium,* and running up through Freud, philosopher Robert Solomon, and most recently, *A General Theory of Love* (Lewis et al, Random House, New York, 2007). It's not simple. It's complex. It's emergent. There's no first instance. It's not reducible to sublimated libido. It's not predictable from animal hunger or sexual instinct. And just as we can learn more about life by exploring the ways

inanimate matter can be animated, so maybe we can learn more about love by exploring how a uni-cellular automaton gets food. We're *not* anthropomorphizing to explore desire at this level, because further up the tree, we're not going to presuppose consciousness as a precondition for desire. We're going to see desire as a precondition for consciousness.

(e) Consumer demand

Moving further down the list, on beyond those supposedly simple cases where we risk anthropomorphism in using the word, 'desire,' it will be less problematic to observe that desire is at the root of markets. It is at the heart of consumer demand. What do those consumers want? Whole armies of market researchers build careers around their acumen in discerning the desires of consumers.

The gratification of desire through the acquisition of possessions in the marketplace does much more than merely satisfy biological needs. Consumption (not just production) can serve as a medium for the creation and expression of character and individuality. Contrary to a tradition that privileges the activity of production and the value of productivity over the apparent passivity of mere consumption, the gratification of desire is a legitimate arena for human self-realization. As contemporary artist, Jenny Holzer, puts it in one of her carefully crafted aluminum plaques, "Finding extreme pleasure can make you a better person if you are careful about what thrills you."

The bad news implicit in the glorification of desire is the potential for desire run amok turning into greed. Where needs yield to finite and unambiguous satisfactions—a full stomach, a quenched thirst—desires, precisely to the extent that they become educated by advertisements and fed by creativity, can reel off into an infinite longing that can never be satisfied. (For a development of this thesis, see Colin Campbell, *The Romantic Ethic and the Spirit of Modern Consumerism,* Blackwell, Oxford, 1987.) Precisely to the extent that mind can always imagine some further variation on the object of desire, some further comparison with someone else's property, some further form of satisfaction . . . desire, strong desire, is ever susceptible to broaching over into the vice of greed.

Whatever role intense desire may have played in Adam Smith's day, its role today is heightened by the degree to which advanced economies are shifting from the satisfaction of universal needs to the gratification of particular wants and desires. This shift changes the balance between the need for heightened production to satisfy needs toward the need for heightened consumption to gratify desires.

Surprising as it may seem to macro-economists who see productivity as the primary driver of national economies, intense desire is playing an ever greater role in maintaining a healthy economy; its measurement might take the form of quantities of

"consumptivity." The marketplace is the best means of gratifying particular desires, however good central planning may be at satisfying universal needs. Note that this section is not about c*entral* planning or the trickling down of excess wealth produced by rampant greed. Instead it is about consumers satisfying highly differentiated desires, wants, and whims, whereas an industrial society tries to master mass production of necessities for a mass market. These two economies call for vastly different kinds of organization: the bureaucratic hierarchy for the industrial economy. CEOs have to give vision and direction to mid-managers as to how they should be directing the workers to do their jobs so that every consumer gets exactly the same thing when they open a box that says *Cheerios.* In an information economy, however, chances are that the retailer who gets face-to-face with the consumer, knows more about what consumers than high-priced market research consultants. So it's the "lowest" layer on the hierarchy that should be giving directions, not the higher level mid-managers, much less the high-salaried CEO.

(f) Values as decision criteria

One step above desire as a craving for some particular consumer good, it's worth considering *values* as sublimated desires. *What do you value?* This question has a different ring from, *What do you desire?* Values have a certain halo-effect. You might desire a chocolate éclair, but you would hardly say you *value* a chocolate éclair. No, you value things like love and leisure time and a sense of meaning in your work. Some of your values act like restraints on desires. You value good health, so you don't eat too many chocolate éclairs.

(g) Social hopes: what do *we* want?

At organizationally higher levels—whole organizations or communities, and not just the individuals that populate those communities—it's worth asking what a corporation wants. The words we use for this instance of desire are words like vision or strategy. *Where do we want to go? Who do we want to be?* Whole armies of consultants and strategic planning offices are engaged in answering these questions, just as armies of market researchers try to determine (and influence through advertising) what individual consumers want.

How do strategists figure out where a company wants to go? The old method, consistent with a "Ptolemaic-Cartesian" view of consciousness, consisted in predicting the future, and then calculating the best way to prevail in that future. Strategy followed a rationalistic, command and control paradigm. The model was *instructional*: Senior executives would figure out what to do and the best way to do it, and then instruct their

underlings to follow their directions. Trouble was, the future is not predicable. Time is real. Surprises occur. So the best laid plans of many a corporation, and the most careful and detailed instructions articulated by senior executives and passed down via middle managers, did not always lead to success. Many a vice-president of strategic planning and many a middle manager got fired.

By the turn of the millennium, an alternative approach to planning called scenario planning became the leading approach to strategic planning among Fortune 500 corporations. Scenario planning assumes that time is real, that the future is unpredictable, and that Murphy's Law will foil the most careful of instructions. Executives frame several possible scenarios of several possible futures and then train many employees to be sensitive to the signs that one or another scenario might be unfolding. This alternative approach is not so rationalistic, and not so instructional.

The business literature consistent with this approach talks about making decisions "closer to the coal face." Especially in information intensive business like software design or bio-technology, executives have to assume that their employees know more than they do about the work they do each day, so an instructional, down-from-the top model for directing a company toward what it wants is less viable than an up-from-the-bottom, entrepreneurial strategy. Let a thousand innovations bloom. Test them against different scenarios. Then let the strategy emerge from those options that prove to be most robust across the range of scenarios.

This kind of approach to the steering of a corporation begins to sound much more like the steering of a uni-cellular organism. We don't assume a representational intelligence that can observe and map its environment and then figure out the single best path to pursue. We assume a process of almost random variation and natural selection, a Darwinian process that achieves self-organization without a central organizer. In the corporate world, the variations won't be altogether random or blind. Lower level employees close to the coal face are not stupid. But we won't presume that they have a spectral overview of the entire corporation and its workings of the sort that we used to attribute to all-knowing CEOs like Harold Geneen or Jack Welsh.

The application of Darwinian principles to the anatomy of desire has its anticipations in the pursuit of artificial life by people like Rodney Brooks, who found that he could create smarter, quicker robots by giving up on the old down-from-the-top Artificial Intelligence paradigm.

> [The AI paradigm] stated that the robot first perceived its world and then began to think about it—it tried to build a little model of the world and then lay mental plans as to how it would achieve its goal in that world. Only then would the robot act, by translating its cognition into action. Brooks believed that there

> should only be two steps—perception and action. The robot should sense something and then act on it, without a cognitive bottleneck. . .
>
> Using this new idea, the robot would forgo the complicated planning, mapping, and cognition required by the AI paradigm. (Steven Levy, *Artificial Life: The Quest for a New Creation,* Pantheon Books, New York, 1992, pp. 277f.)

This begins to sound like business guru Tom Peters urging executives to "Ready, fire, aim." There's no time in real time for an elaborate process of prediction, mapping, modeling, aiming, and down-from-the-top instructional control. Instead, both corporations and much simpler organisms must try out a bunch of different behaviors and see what works. This evolutionary process of variation and selection is—to use a not very surprising word in this context—far more *adaptive* to an unpredictable environment, an environment where, as Stuart Kauffmann often says, we can't even predefine the parameters of the decision space.

Of course this kind of learning can be costly and inefficient. Think of the man whose wife asked, "Dear, can we afford another learning experience?" But everything we are learning from the rapidly advancing field of artificial life is telling us that evolutionary algorithms that spawn vast numbers of variations and then select the winners are, over the longer run, more resilient than specialized instruction sets composed by coders who think they know the answers in advance. In a world where the problems are never quite the same, it's much better to have a system that learns how to learn rather than a system that thinks it already knows the one best answer.

As we traverse our way up and down the hierarchy of desire, from uni-cellular organisms to vast corporations steering their ways into the future, we begin to see how the various instances of desire have less to do with a representational, cognitive, calculating intelligence than with Darwinian experimentation and selection of the most adaptive behaviors. Desire begins with an inchoate hunger that doesn't know precisely what it wants, then *desire learns by experimentation*. Like the man who didn't know what he thought until he heard what he said, desire doesn't always know what it wants until it feels most fully gratified.

Just as we rely on Edelman's "neural Darwinism" as a description of the way the brain creates a platform sufficient for the emergence of the experience of consciousness, so we find ourselves once again relieved of the charge of anthropomorphizing the biological precursors of intelligence. It's not that paramecia or corporations think the way humans do; it's that humans rely on Darwinian processes not unlike those that govern the success or failure of paramecia and corporations. Quite the opposite of presuming some cousin of the "ghost in the machine" residing in paramecia or corporations, we're using what we are learning about desire at the level of paramecia and corporations to help us deconstruct

the homunculus supposedly inside the representational/cognitivist model of the ghost in the machine.

Before moving up to the next rung on the hierarchy of desire—Teleonomy and story: What does technology, or information, or the earth *want?*—it may be worth pausing for a further reflection on method. While the overall strategy of this section (a) to (j) is to traverse a phenomenology of various grades of desire, from the archaic to the sublime, the reader may be reluctant to attribute *desire* to entities as diverse as paramecia and the earth. Rather than suspending judgment on whether the language of desire adds to or detracts from perspicuous representations of these phenomena, the reader may simply reject the phenomenological descriptions. Why? Because of an un-suspended conviction that the language of desire is appropriate only to those entities known as human beings. The last few pages may have gone part of the way toward deflecting the charge of anthropomorphism in attributing desire to entities other than *anthropoi.* But a further reflection on method may be needed to show just how radical is the reversal in priority of cognition and desire.

Put differently: Before the reader can accept a description of the lowest and highest rungs of this hierarchy as descriptions of *desire,* it may be necessary to deconstruct the cognitivist alternative. The new way of looking at things may not be available unless the old way is quite explicitly and self-consciously challenged. So that is what the next section attempts to do.

But further, the new way of looking at things, once available, may be rejected if it claims *too much.* Is it possible that we are expecting too much of desire? Is the lifting too heavy? If we paint everything with the brush of desire, do we lose all definition between what is and what is not a function of desire?

These are important questions. So before climbing the last two rungs on the ladder of desire, we'll take a detour to get around some obstacles that may be obstructing our view . . . and then state some caveats that will keep the new view of desire from claiming too much.

More on Method: Desire, Evolutionary Epistemology, and Explanation

According to a cognitivist epistemology, knowledge consists in the correspondence between subjective representations and objective things. Knowledge is defined as justified true belief. Time does not enter the equation. Desire is relevant only to the extent that it runs the danger of introducing subjective biases that may cloud our perception of objective truth.

American pragmatism—the philosophy developed by Peirce, James and Dewey and revived in the writings of Richard Rorty—puts time squarely into the equation that determines truth. To quote Rorty:

> As I see it, the link between Whitmanesque Americanism and pragmatist philosophy . . . is a willingness to refer all questions of ultimate justification to the future, to the substance of things hoped for. If there is anything distinctive about pragmatism it is that it substitutes the notion of a better human future for the notions of 'reality', 'reason' and 'nature'. One may say of pragmatism what Novalis said of Romanticism, that it is 'the apotheosis of the future'. (Richard Rorty, *Philosophy and Social Hope,* London, Penguin Books, 1999, p. 27.)

For pragmatists, *the future replaces objectivity as the horizon of validation.*

Truth is not about correspondence; truth is about survival. Justified true belief? "There is no justification, ever," writes W. W. Bartley III in the opening essay of an anthology entitled *Evolutionary Epistemology.* (W. W. Bartley III, *Evolutionary Epistemology,* La Salle, Illinois, Open Court, 1987, p. 24.) "The process that began with unjustified variations ends with unjustified survivors." The progress of knowledge is Darwinian in a sense that Karl Popper and Imre Lakatos capture with the language of their title: "Conjectures and refutations" (read "variation and selection"). Each new truth claim is a conjecture; if it is not refuted, it survives, but it is never, strictly speaking, *justified* in the sense of establishing a correspondence between subjective representation and objective reality. What is true? As parents say to impatient children, *"We'll see."* To repeat: *The future has replaced objectivity as the horizon of validation.*

As we peel back the onion of desire still further, we find that desire, in the form of shared hope, plays a role in influencing the outcome of inquiry. 'Influence' is the operative word. 'Determine' would be too strong. Wanting something doesn't necessarily make it so. Desire can lead to delusion. But the shared hopes of a community, as elicited and expressed through a process of planning and implementation of those plans, can *influence* the unfolding of history in such a way that hopes come true. Ernst Bloch surely shows the efficacy of hope in his three volumes on the principle of hope, *Das Prinzip Hoffnung.*

If what we want influences the future, and it is the future, not objectivity, that determines the truth of what we know today, then what we want will influence what we can know. The future will determine truth, so what we *will,* the products of our volition, will influence what becomes, in the future, true—what "comes true."

In this case *will* might become something close to what Schopenhauer meant by 'will' in his multi-volume work, *The World as Will and Idea.* Schopenhauer's sense of the importance of will is about more than stubborn volition, as in "She is a strong-willed young woman." Schopenhauer was getting at something else, something about the role of

desire in the constitution of the cosmos. But Schopenhauer's account does not hold up under analysis. His concept of will falls into the fallacy of Dr. Dizmo's elixir—explaining too little by explaining too much.

On the one hand it's worth giving credit where credit is due: To Schopenhauer's intuition that the world *as idea* alone is insufficient. Schopenhauer wanted to distinguish a purely spectatorial representation of the world from some sort of grasp of its dynamism, its drive, its creativity, its motive force. This he tried to do by distinguishing between world as will and world as idea. But when he came to describing what he meant by 'will,' he explained too much:

> [I]t is throughout not idea, but *toto genere* different from it; it is that of which all idea, all object, is the phenomenal appearance, the visibility, the objectification. It is the inmost nature, the kernel, of every particular thing, and also of the whole. It appears in every blind force of nature and also in the preconsidered action of man; and the great difference between these two is merely in the degree of the manifestation, not in the nature of what manifests itself. (Arthur Schopenhauer, *The World as Will and Idea,* Section 21; in *Schopenhauer Selections,* ed. DeWitt H. Parker, New York, Charles Scribner's Sons, 1956, p. 73.)

What is wrong with this kind of metaphysical talk? Simply that it is vacuous. Dr. Dizmo might say that the whole world is made of tomatoes. You object that the chair you are sitting on is not a tomato, and Dr. Dizmo replies that he's talking about very *special* tomatoes that can take on any shape whatever, and that you happen to be sitting on a chair-shaped tomato. This is precisely the tactic Schopenhauer employs in the section following the passage just quoted. He states, "It is, however, well to observe that here . . . the concept of the will receives a greater extension than it has hitherto had." I.e., these are very *special* tomatoes unlike any of the ordinary ones you have yet seen. "I therefore name the genus after its most important species, the direct knowledge of which lies nearer to us and guides us to the indirect knowledge of all other species." I.e., the generic Will he's talking about is a very *special will* of which the experience of human volition is only a particular species.

> But the word *will,* which, like a magic spell, discloses to us the inmost being of everything in nature, is by no means an unknown quantity, something arrived at only by inference, but is fully and immediately comprehended, and is so familiar to us that we know and understand what will is far better than anything else whatever. The concept of will has hitherto commonly been subordinated to that of force, but I reverse the matter entirely, and desire that every force in nature should be thought as will. It must not be supposed that this is mere verbal quibbling or of no consequence; rather, it is of the greatest significance and importance. (*Ibid.,* Section 22, pp. 74f.)

Well . . . maybe.

Nietzsche was much influenced by Schopenhauer. Though he studied Schopenhauer's writings carefully enough to find disagreements with them, he nonetheless paid them the respect of an admirer. One of the clearest marks of Schopenhauer's influence is surely on Nietzsche's idea of *the will to power.* But Nietzsche, too, succumbed to the logic of Dr. Dizmo by claiming too much for his "discovery":

> [T]his, my *Dionysian* world of the eternally self-creating, the eternally self-destroying, the mystery world of the twofold voluptuous delight, my "beyond good and evil," without goal, unless the joy of the circle is itself a goal; without will, unless a ring feels good will toward itself—do you want a *name* for this world? A *solution* for all its riddles? A *light* for you, too, you best-concealed, strongest, most intrepid, most midnightly men?—*This world is the will to power—and nothing besides!* And you yourselves are also this will to power—and nothing besides! (Friedrich Nietzsche, *The Will to Power,* ed. and trans. by Walter Kaufmann, New York, Vintage Books, 1967, p. 550, #1067. For a good discussion of Nietzsche's *will to power,* see Robert C. Solomon and Kathleen M. Higgins, *What Nietzsche Really Said,* New York, Schocken Books, 2000, pp. 215-222.)

And so concludes, with this exclamatory cadence, Nietzsche's posthumously published book, *The Will to Power,* which Heidegger claimed, on the first page of his four volumes of commentary, to be Nietzsche's "chief philosophical work." (Martin Heidegger, *Nietzsche,* trans. David Farrell Krell, San Francisco, Harper, 1984, Vol. I p. 3.) Yet when Heidegger is pressed to explain the meaning of Nietzsche's phrase, "the will to power," the lack of traction, the vacuousness, the ultimately tautological nature of the concept is manifest in Heidegger's formulation:

> [W]hat does Nietzsche himself understand by the phrase "will to power"? What does "will" mean? What does "will to power" mean? For Nietzsche these two questions are but one. For in his view will is nothing else than will to power, and power nothing else than the essence of will. Hence, will to power is will to will, which is to say, willing is self-willing. (*Ibid.,* p. 37.)

Okay . . . but in order to gain some traction while building on this lineage from Schopenhauer, through Nietzsche, we do better to turn to Heidegger's concept of care (*Sorge*). (Heidegger, *Being and Time,* trans. Macquarrie, New York, Harper & Row, 1962, esp. paragraphs 39 ff.) The German word, *Sorge,* has connotations that are stronger than the less emphatic, English word, 'care.' Part of what Heidegger is getting at might be better translated as 'giving a damn.' To be a human being, for Heidegger, is to *really* care about some things. Human subjectivity is not just a subject of cognition, not just a *Cogito* or a Kantian "transcendental unity of apperception." Human subjectivity is not just a spectral observation post on a passing scene of objects about which human

subjects are disinterested or indifferent. Nor is care to be confused with "will, wish, addiction, and urge. Care cannot be derived from these, since they themselves are founded upon it." (*Ibid.,* p. 227, H182.) In taking pains to distinguish a "pre-ontological" from an "ontological" understanding of care, Heidegger risks introducing us to a very *special* kind of care . . . but at least he brings us a little way down out of the clouds of vacuous metaphysics towards something recognizably closer to human experience and intelligiblity, a little closer to the practical rather than the purely theoretical.

Kantians talk about "the primacy of the practical" in a way that suggests Kant cared more about his second critique, which is his ethics, than his first, which is his epistemology. Kantians speak of the primacy of the practical as a way of saying that Kant's ethics takes precedence over his epistemology. The meaning of this phrase, 'the primacy of the practical,' can be clarified by its application to debates over the ethics of human cloning. Is it a question of finding a "right answer" through dispassionate, theoretical reason (the subject of Kant's first Critique)? Or is the question of cloning (and/or stem cell research) a question of *what we do or don't want?* The ethics of bio-technology boils down to the question: What sort of humans do we *want* to be? The ethics of bio-technology will not be settled by a theoretical inquiry asking: What *are* we such that bio-technological interventions are or are not ethical?

Just as history is written by the victors, so ethics will be written by the survivors. This proposition is not as cynical as it sounds. The point is not simply that might makes right. In the present context, the point is instead that a capacity for cooperative co-evolution makes for might, which only then makes right. Short term selfishness, or a survival of the fittest mentality that sees nature red in tooth and claw, will not prevail over the combined might of cooperative, emergent systems. Even an aggressive and well-armored snapping turtle—a veritable fighting machine—is no match for a more highly evolved pack of humans with sharp sticks. As opposed to a Spencerian understanding of evolution as a Hobbesian survival of the fittest, the survival value of cooperation is now a well-accepted principle, thanks to such works as Peter Corning's meticulously argued book, *The Synergism Hypothesis,* (Peter Corning, *The Synergism Hypothesis,* New York, McGraw-Hill, 1983.*)* Robert Axelrod's *The Evolution of Cooperation,* and Robert Wright's *Non-zero.* (Robert Axelrod, *The Evolution of Cooperation,* New York, Basic Books, 1984. See also Robert Wright, *Nonzero: The Logic of Human Destiny,* New York, Pantheon, 2000.*)*

Evolutionary epistemology assumes interested agents who care, not disinterested observers. In the context of an evolutionary epistemology, care, interest, desire, are not to be despised as engines of bias, distortion and falsehood. They are instead the agents of *what which will come true.* Nor need we restrict desire or interest to higher primates. As we travel up and down the hierarchy of desires, it makes sense to look for the origins of

desire (if not first instances) in simpler organisms at lower levels of emergent complexity. Hence Stuart Kauffman's interest in "autonomous agents."

> Consider a bacterium swimming upstream in a glucose gradient, its flagellar motor rotating. If we naively ask, "What is it doing?" we unhesitatingly answer something like, "It's going to get dinner." That is, without attributing consciousness or conscious purpose, we view the bacterium as acting on its own behalf in an environment. . .
>
> An autonomous agent is a physical system, such as a bacterium, that can act on its own behalf in an environment. (Stuart Kauffman, *Investigations, op. cit.,* pp. 7f.)

Now comes the surprising part, surprising both to Kauffman, by his own report, and surprising to the rest of us raised on centuries of Newtonian science. According to Kauffman, a proper understanding of autonomous agents calls for *a new way of doing science* because, "the concept of an agent is, inherently, a non-equilibrium concept." According to Kauffman, "Life is doing something far richer than we may have dreamed, literally something incalculable. What is the place of law if, as hinted above, the variables and configuration space cannot be prespecified for a biosphere, or perhaps a universe? Yet, I think, there are laws. And if these musings be true, we must rethink science itself."

This is the task we set ourselves in seeking laws or traits for emergent systems. Whe confronted with emergent phenomena, it's not that we need to surrender the search for understanding according to laws. We needn't yield to the impulse to say, as in the opening cartoon, "And here's where a miracle happens." But we need different laws, a set of traits for emergent systems, a set of traits or laws or lenses to supplement the single eye of monological—or monocular—science. So as Kauffman suggests, we may need to "rethink science itself." And this is no small order.

So much for this detour into method. The point was to acknowledge the danger of attributing desire to entities that cannot possibly *want* anything, at least not as long as those entities are seen through the lenses of monological science. We know this. But rather than shrink back from attributing desire to such entities, this reflection on method leads us to agree with Stuart Kauffman that we need to "rethink science itself." In doing so, it's worth worrying about Dr. Dizmo: Desire can lift a lot. But it cannot lift *everything.*

On this next rung of the ladder of desire, we'll see how it makes sense to talk about what the earth *wants* . . . but only if we have a fairly rigorous definition for *preference* on such a scale. Kauffman gives the idea clarity through his concept of "the adjacent possible." But Kauffman's account will be misconstrued unless set in the sort of context this detour

has provided, namely, a self-conscious meditation on the need to "rethink science itself." Yet again, no small order. This book is radical!

(h) Teleonomy and story: what does technology, or information, or the earth *want?*

Kauffman takes us right up to the brink of this new science and, quite correctly, points to the need for *story* as a means for understanding autonomous agents. "Our inability to prestate the configuration space of a biosphere foretells a deepening of science, a search for story and historical contingency, yet a place for natural laws." (*Ibid.,* p. 135.)

What is it about story that helps us comprehend historical contingency? Here we need to supplement Kauffman's account with an understanding of narrative that goes back to Hegel and one of his interpreters. Lacking lucid quotes from Hegel, who was a terrible writer, we can turn to Hayden White, literary critic and historiographer. Giving full credit to Hegel, whom he is interpreting, White writes: "The reality that lends itself to narrative representation is the conflict between desire and the law." (Hayden White, *The Content of the Form,* The Johns Hopkins University Press, Baltimore, 1987, p. 12.)

The law that Hegel and Hayden White are talking about is the law of the State, but if Kauffman is right about the need for story to comprehend what is otherwise incalculable under monological causality, then the principle should hold true for the conflict between desire and *natural law* as well. Synthesizing the insights of Kauffman and Hegel (as interpreted by White) we arrive at a systemic set of mutual implications among four concepts: (1) law, representing the realm of monological *necessity;* (2) desire, which wants what it wants in the face of what supposedly *must be;* (3) narrative, or story, which tells the tale of the conflict between desire and law; between freedom and necessity; and, finally, (4) *the emergence of subjectivity as constituted by desire's struggle with necessity, as told in the form of biography or history, and self-referentially re-told as the self-constituting auto-biography of an emergent subject.* (Cf. Kerby, Anthony Paul, 1991: *Narrative and the Self,* Bloomington, Indiana University Press.*)*

Because this struggle is not just the playing out of necessity under the covering law model—a description using equations stating generalities about what *must be*—but a struggle between necessity and *desire*, monological science must be supplemented by narrative or history. The only way to make sense of this struggle is to see it as the struggle of a subject who *cares about what happens* against an indifferent realm of law that doesn't give a damn. White continues his rendering of Hegel in a way that links back to both agency and law. "Where there is no rule of law, there can be neither a subject nor the kind of event that lends itself to narrative representation." Not only is monological science not simply wrong; it is a necessary condition for the emergence of subjectivity from the conflict between desire and that law.

An agent that acts on its own behalf is exercising desire in the face of necessity, whether it is going to get dinner or realizing the hopes of a community. By doing so, it constitutes itself *as* an agent. It is not the case that *first* there were agents and *then* they acted desirously. No, the emergence of subjectivity does not allow of such a first instance or ontological priority. Only in the struggle between desire and the law does there emerge—or co-evolve—a subject or "the kind of event that lends itself to narrative representation." The desiring subject and the object of desire are systematically linked by an ontological reciprocity that emerges all of a piece, all of a sudden, synchronically.

This all-of-a-piece-ness is best captured by a story that has a beginning, a middle, and *an end* that ties up all of the pieces in a nice narrative knot, a happy ending—or, alternatively, a tragedy—that looks equally like a telos or final cause. Teleology, however, is the bête-noir of monological science. Ever since Spinoza, science has had no place for arguments from design. We do not explain things by saying they were *meant* to be so, or that God *designed* them that way. Darwinian evolution has replaced the argument from design. You don't need a designer when you see how much can be explained by billions of years of variation and selection. But you do need some way of accounting for the degree of complexity and organization that has in fact evolved.

Kauffman casts doubts on the adequacy of our current understanding of evolution to account for the degree of complexity we see before us. According to Kauffman, who is well schooled in both philosophy and biology, it is not at all clear from classical evolutionary theory why complexity should have ever evolved, much less any particular complex organism like an eye.

> We do not understand evolution. We live it with moss, fruit, fin, and quill fellows. We see it since Darwin. We have insights of forms and their formation, won from efforts since Aristotle codified the embryological investigations that over twenty-five centuries ago began with the study of deformed fetuses in sacrificial animals.
>
> But we do not understand evolution. . .
>
> Darwin's theory of evolution is a theory of descent with modification. It does not yet explain the genesis of forms, but the trimmings of the forms, once they are generated. "Rather like achieving an apple tree by trimming off all the branches," said a late-nineteenth-century skeptic. (Kauffman, *Investigations,* pp. 16f.)

According to the usual understanding of Darwinian evolution, the genesis of forms we see before us—*morphogenesis,* to use Rene Thom's word—occurs as a result of random mutation and genetic variation. But Kauffman and his colleagues at the Santa Fe Institute have done the math. You can't get from there to here, from a primal soup of relatively

simple molecules to today's ecosystem of complex and diverse species, by assuming the physico-chemical equivalent of the million monkeys with a million typewriters who, given enough time, could supposedly bang out *Hamlet.*

The million monkeys with enough time hypothesis assumes an ergodic, which is to say *repeating,* universe, a billiard-ball universe of physical interactions which, sooner or later, exhaust all the possibilities and then trace the same path on the pool table. But our universe is radically non-ergodic, non-repeating. Time moves in only one direction, toward increasing complexity. And there is no way that in the history of our universe, all of the combinatorial possibilities of the simple elements could have been nearly enough exhausted to yield the complex combinations we now see before us.

The formulas of Newtonian physics are reversible; they work backwards as well as forwards. Real time, the time we live in every day, is irreversible. But why? It stands as something of an embarrassment to monological science that it does not have an answer to this simple sounding question. Many physicists and philosophers have spent a lot of time and chalk trying to figure out an explanation for what has come to be called "time's arrow." What is the explanation for time's arrow?

Some have fixed on the second law of thermodynamics as the source of time's arrow: entropy. But that explanation flies in the face of the manifest facts of increasing order and increasing complexity. If entropy told the whole story, then things would be coming apart, running down, headed toward universal heat death as surely as one's cup of coffee gets cooler rather than warmer while sitting on the desk. Something else is at work, some other force or field or phenomenon running contrary to entropy, and it's not that easy to discover what it is once you reject the usual suspects.

What are the usual suspects? First, the argument from God's design. This is not science, but suspect theology.

Second, Darwinian evolution is based on random variation and selection over billions of years. This answer is appealing for all the reasons that evolutionary theory is superior to religious creationism . . . but, for reasons given by Kauffman and others, it just won't work by itself. Granted, once you start with a few simple elements shortly after the Big Bang, there was a lot of time to try one combination and another and another to get the ball of increasing order rolling. But random variation and selection alone will not account for the *amount* of increasing order we see before us.

It's as if we set those million monkeys to typing, and after a billion years one got as far as: "To be, or not to be, that is the quesxoek . . ," and then they all had to start all over again. There has to be some ratchet effect such that order, once achieved, enjoys a degree of stability that can be built upon. Call it heritability.

Eschewing any argument from design, and granting the inadequacy of random variation and selection, the creators of complexity theory—Stuart Kauffman, John Holland, Brian Arthur, Terrence Deacon and others—have been hard at work trying to find the traits of self-organizing systems to explain how and why time has an arrow, and it doesn't just point toward entropic heat death.

Terrence Deacon and Ty Cashman have tackled the emergence of complexity, and in a way that might obviate the need for finding time's arrow. Their approach does not assume the predetermination of the future. ("Eliminativism, Complexity, and Emergence," The Routledge Companion to Religion and Science, 2010; see also 'Chapters 2 and 3 of Deacon's forthcoming book, *Falling Up,* "Living Complexity" and "Complexification.")

Kauffman "senses" a fourth law of thermodynamics. His explanation presupposes a concept he labels, "the adjacent possible" which amounts to a kind of global variation on the world as we know it. Before drawing implications from Kauffman's argument, it will be best to quote him at some length. He defines the adjacent possible as follows:

> The adjacent possible consists of all those molecular species that are not members of the actual, but are *one reaction step away from the actual.* That is, the adjacent possible comprises just those molecular species that are not present in the vicinity of the Earth out to twice the radius to the moon, but can be synthesized from the actual molecular species in a single reaction step from substrates in the actual to products in the adjacent possible. (*Ibid.,* p. 142.)

Kaufmann then observes:

> The biosphere has expanded, indeed, more or less persistently exploded, into the ever-expanding adjacent possible. . .
>
> It is more than slightly interesting that this fact is clearly true, that it is rarely remarked upon, and that we have no particular theory for this expansion. . .
>
> Now a second simple point. The molecular species of the actual exist. Those in the adjacent possible do not exist—at least within the volume of the universe we are talking about, which we can expand in a moment to be the actual molecular diversity of the entire universe, not just our tiny patch of it.
>
> The chemical potential of a single reaction with a single set of substrates and no products is perfectly definable. . .The substrates are present in the actual, and the products are not present in the actual, but only in the adjacent possible. It follows that every such reaction couple is displaced from its equilibrium in the direction of an excess of substrates compared to its products. This displacement constitutes a chemical potential driving the reaction toward equilibrium. The simple conclusion is that there is a real chemical potential from the actual to the

> adjacent possible. Other things being equal, the total system "wants" to flow into the adjacent possible. (*Ibid.,* p. 143.)

These last two sentences, especially that little word, "wants", call for comment from the perspective of desire, but not before following Kauffman's argument toward a fourth law of thermodynamics that seems to follow from these thoughts.

> It seems reasonable to think of the "workspace" of the biosphere, that is, what can happen next, as its actual plus its real adjacent possible. It seems likely, and I do conjecture, that the biosphere is expanding its workspace, on average, as fast as it can do so without destroying itself in the process. . .
>
> The greater the current diversity of matter, processes, and sources of energy, the more ways there are for these to couple to generate yet further novelty, further symmetry breakings. For this to be correct, time would have to have a directionality toward persistently broken symmetries. And an arrow of time would lie in this directionality . . .
>
> If one could ever show such a law, a law in which the diversity and complexity of the universe naturally increases in some optimal manner, that would be impressive. Some fourth law of thermodynamics? An arrow of time? In short, one intriguing hypothesis about the arrow of time is that the nonergodic universe as a whole constructs itself persistently into an expanding adjacent possible, persistently expanding its workspace. This is in sharp contrast to the familiar idea that the persistent increase in entropy of the second law of thermodynamics is the cause for the arrow of time. But the second law only makes sense for systems and time scales for which the ergodic hypothesis holds. (*Ibid.,* p. 151.)

Fourth law of thermodynamics? Or seventh trait of emergent systems? In either case it is a *natural* law, not magic. This "wanting," even in quotes, is the phenomenon we're after—desire at the level of the entire cosmos, possibly the referent of Nietzsche's ill-defined phrase, "the will to power." Possibly what Schopenhauer intuited but over-stated. The quotes around Kauffman's use of 'wants' are called for because we know perfectly well that the "total system" is not a conscious subject, not an *anthropos* like Fred or Alice. We need not anthropomorphize the cosmos any more than we need to anthropomorphize the corporation or the paramecium. Quite to the contrary, by looking at the full range of instances on different levels, up and down the hierarchy of desire, perhaps we can discern a *structure of desire* that does not presuppose a Cartesian subject or *Cogito,* but instead helps to *constitute* some sort of subject through a process of emergent co-evolution.

Kauffman posits "a real chemical potential from the actual to the adjacent possible." How is that potential produced? By an excess of substrates compared with their products. What actually *has happened* in the world as we know it is a small fraction of what might have

happened or might yet happen, as is clear from a glance at the adjacent possible. Kauffman conjectures, "that the biosphere is expanding its workspace, on average, as fast as it can do so without destroying itself in the process." Kauffman formalizes his concept of the adjacent possible in ways that lend precision to the concept of desire. Without reproducing his entire argument, it's safe to say that the concept of the adjacent possible is at the very least highly suggestive for lending rigor to the structure of desire without a *Cogito*.

Let's follow the suggestion. Let's be suggestible. Who among us has not felt desire for "an adjacent possible?" Call it the grass is greener syndrome. Recall the commandment against coveting thy neighbor's wife. Without assuming a Cartesian subject, we might describe this kind of desire as *the appeal of the adjacent possible* where the appeal has less to do with some image of the object of desire in the representational consciousness of the desirer, and more to do with the desirability of the *other.* What constitutes that desirability? In part, its very *otherness*, its *difference* from the actual, the simple fact that it is one step *away.* Here we are talking about desire in its capacity to break bonds, to transgress a given order (e.g. matrimony), and shatter the contentment and equilibrium of an existing system.

Clearly, we're taking liberties with Kauffman's language. He never meant to be describing the psychology of extra-marital affairs. We know this. But let us follow the phenomenology of desire where it wants to go, toward a description of this structure of "wanting." If it is not dependent on a representational image of what is wanted, but is instead a function of an entire regime or niche at one step removed from the actual—what Kauffman calls 'the adjacent possible'—then it appears that we are dealing with a kind of explanation or account that is quite different from monological causality under a covering law. We are dealing with a kind of explanation that Bateson describes as familiar to mathematicians, namely, that kind of explanation that doesn't explain *q* by invoking the covering law, *if p then q*, then supplying evidence for initial conditions, *p;* but proceeds instead by explaining *q* by talking about why everything *else* in the world *could not have happened.* In short, by citing all of the traits and conditions that would thwart the desire for anything *other* than *q.*

By asking us to think about the adjacent possible, however, Kauffman is suggesting the possibility of a relaxation of those laws. He is prepared to tell stories about the conflict between desire and necessity, stories in which desire might escape the necessities of the actual by transgressing those necessities to render the *adjacent* possible. This rendering of the adjacent possible is not the result of following instructions dictated by a representational image of some singular object of desire. The adjacent possible is not the kind of singular entity beloved of monological reasoning. It is an entire regime, a world, an *umwelt,* at the very least a niche. The explanation is relational and contextual rather

than monological. Granting the limits of what is possible in *this* world, the actual, what might be possible in *that* world of the adjacent possible?

What will that world afford? To use a phrase Kauffman likes, both in its literal applications in what he calls the "econosphere," and in its metaphorical applications in biology, it's worth asking: Are there different ways of "making a living" in the adjacent possible? What are the affordances of *that* world? What does that world *want?*

When Stewart Brand makes the gnomic pronouncement, "Information wants to be free," he is not assuming a Cartesian consciousness on behalf of the vast disaggregated sea of bits and bytes we call information. Instead he is saying that today's economy is the sort of context and information is the way it is, such that information is neither easy to measure, nor easy to price, nor easy to warehouse under lock and key. Information technology and the economy are so structured that the transfer of information tends to be rapid, unmeasurable and uncontrollable. The economy is such that it's hard to buy and sell information the way we buy and sell tangible commodities. This statement, "Information wants to be free," is as much a statement about the economic context as about information.

If it is the context and its affordances that account for the structure of desire, then this will help us to dismiss the traditional objection against teleological reasoning, namely that monological causality only works in one direction, from past to present, not from future to present. Causes precede effects. The trouble with a telos or goal as an explainer is that it follows rather than precedes its effect. If someone argues that effect, *q,* is as it is because of *p,* where *p* is a telos or goal that post-dates, *q,* as in the argument that May flowers are the cause of or reason for April showers, then we should reject that line of reasoning because it violates time's arrow. What is later cannot cause what is earlier. But if we are not explaining *q* by some single cause, *p,* whether earlier or later than *q,* but instead by the affordances of the entire context that is the adjacent possible to the regime in which we find *q,* then we are not guilty of the kind of teleological reasoning to which Spinoza and the rest of the rationalistic Enlightenment rightfully objected. We are offering *a different kind of explanation,* one that, as Kauffman rightly declares, will call upon our capacity for telling stories; one that, as Hegel and Hayden White add, will play upon the tension between desire and the law. And yet again, this is no small order, supplementing the way we usually explain *things*—causally—with a mode of explanation that gives *reasons,* not *causes.*

Desire, it turns out, *is* time's arrow. But not the desire of a Cartesian *cogito*. By placing the emphasis on the desirability of the adjacent possible, on the object of desire rather than on the intentionality of some subject of desire, this account avoids both the fallacy of teleology and the mystery of the Cartesian ghost in the machine. Instead this account

presumes an emergent system of mutually implicated parts of a whole: an initially inchoate desire—an ill defined will to power—which, with Darwinian profligacy and Dionysian delight, experiments with many objects, and learns through some version of pleasure and pain what is most gratifying. As this initially inchoate desire becomes more focused, as it gets educated by gratification, both the subject of desire and the object of desire gain, simultaneously or synchronically, greater definition. The subject is constituted through the successive stages of the gratification of desire (as suggested by Hegel, explicated by Kojève, and, under the influence of Hegel and Kojève, played out by French intellectuals like Lacan, Kristeva, Deleuze and Guattari. More on this below). The object of desire is likewise and reciprocally constituted as a goal, purpose, or telos. Before delving deeper into the constitution of the subject of desire, it's worth dwelling further on this ontological reciprocity, this co-evolution and systemic co-determination of the subject and object of desire—a desirer and a telos which, because it is co-constituted with desire, is not, as Thom worried, "gratuitous and otiose."

What's wrong with many teleological explanations? And how does the co-determination of subjectivity and teleonomy give us just enough purposiveness in the universe, no more, no less? The kind of teleology deemed otiose, from Spinoza to Rene Thom, is the kind of teleology that puts the cause *after* the effect, the May flowers that supposedly cause April showers, or, as Teilhard du Chardin supposed, an Omega point somewhere in the future that is supposedly sucking all of history towards itself like water down a drain. A predetermined telos denies the creativity of time. With the wrong kind of teleology, the end has already been written. The design has already been drawn. All that remains is to play out a plot whose every turn has been scripted beforehand in the mind of God or by omnipotent and omni-explanatory causality.

There's comfort in this kind of teleology. When bad things happen, believers can say that "everything happens for the best," because, eventually, God's design *will* be realized as surely as an acorn will become an oak. But such an essentialist worldview, in addition to denying the creativity of time, also undercuts our capacity to explain and thereby understand the world around us. Explanation by essence or by telos turns out to be no explanation at all. To explain an acorn's behavior by referring to its essential oakness is to explain nothing, but only to re-describe using what we already know: acorns become oaks. To invoke oak-essence as the explainer of the acorn's growth path is no more helpful than explaining how opium put a man to sleep by referring to its dormative powers, as Moliere's famous satire of Aristotelian science had it.

How does the opium work? *How* does what's inside an acorn work to direct the metabolism of earth, air, water and sunlight into an oak tree? In order to answer these questions, modern science turns away from final causes and pays attention to efficient causes, pushes rather than pulls. Neither Aristotle's essentialism nor the explanations of

the theologians help us answer these questions. Many teleological explanations may offer moral comfort, but they provide considerably less intellectual satisfaction than a careful description of the workings of neuro-transmitters or the role of DNA in directing the morphology of fibro-vascular growth. In order to gain such intellectual satisfaction, however, you need to traverse the long road of the patient scientist: observation, hypothesis, testing by controlled experiment, revision of hypothesis, further testing, further observation, all the while allowing the evolution of theory to inform a theory of evolution.

The long march of monological science marks a distinct advance over teleological explanation, at least as far as the development of testable hypotheses and the accretion of knowledge about how the world works. The price we've paid for this knowledge, however, is the loss of a sense of meaning or purpose. *We don't know where we're going anymore.* In the neo-Darwinian world of chance and necessity, we are devoid of purpose. And this is disconcerting. As Nietzsche observed, "Our will requires an aim; it would sooner have the void for its purpose than be void of purpose." (Nietzsche, *The Genealogy of Morals,* trans. Francis Golffing, Garden City, N.Y., Doubleday Anchor Books, p. 231; Third Essay, #1; *Werke in Drei Banden,* ed. Schlechta, Munich, Carl Hanser Verlag, Vol II, p. 839: "Dass aber uberhaupt das asketische Ideal dem Menschen so viel bedeutet hat, darin druckt sich die Grundtatsache des menschlichen Willens aus, sein *horror vacui: er braucht ein Ziel* — und eher will er noch *das Nichts* wollen als *nicht* wollen.") Ask a Zen Buddhist.

Darwin's critics, and Darwin himself, were right to worry about the loss of meaning that had been provided by a world created according to God's design. That Darwin's critics should have accused him of nihilism is entirely understandable. The play of chance and necessity can be interpreted as so much sound and fury signifying nothing. But it can also be interpreted in a way that threads a narrow path between nihilism on the one hand and teleological essentialism on the other. By articulating the co-evolution of subjectivity and teleonomy, of subject of desire and object of desire, of consciousness and purpose as the two sides of one emergent system, the present argument turns two mysteries into each other's solutions. This is radical!

Subjectivity and teleology, each mysterious when approached individually by monological science, stand together like two sides of an arch holding up the capstone of emergence. Take away the capstone, take away the science of emergent systems, and the arch collapses. By the same token, take away subjectivity, and teleology collapses; take away teleology, and subjectivity collapses. The emergence of subjectivity and teleology is all of a piece, all or none, no first part prior to the whole, not predictable from the parts by themselves, not reducible to the features of the parts by themselves, and inexplicable without desire.

What does desire contribute to this account? Why would subjectivity collapse if you took away teleonomy? These questions lead back to the earlier promisory note in the parenthetical reference to Hegel, Kojève, Lacan, Kristeva, Deleuze and Guatari. So let us retrace those steps and give credit where credit is due. The path is overgrown and obscure. The tradition of philosophy from Hegel to postmodern deconstruction is littered with texts as impenetrable as any ever written. But the basic insights are not all that difficult if you approach them in their proper historical sequence so you can see what's at stake and what's being said to whom and why.

Hegel's depiction of desire

The story begins long ago and far away with Descartes' doubt about the veracity of his knowledge. In the course of his *Meditations* it occurred to him to ask, "Could God be deceiving me?" Could the evidence of my senses be inadequate to the task of cobbling together enough knowledge to get through a life? After all, reasoned Descartes, I look at an oar in the water and it *looks* bent even though the eye of reason *knows* that it is straight. If my senses are so unreliable, perhaps I cannot trust them for anything. What can I trust? What remains indubitable? The only thing Descartes could find that would stand up to the challenge of radical doubt was the existence of the doubter. "I think, therefore I am," he declared. *Cogito, ergo sum.*

Descartes' doubt cast its shadow across the next two centuries of philosophy. Painting with a broad brush in order to cut to the chase, we can abbreviate the battle between the rationalists (Leibniz and Spinoza) and the empiricists (Locke and Hume) as a battle over how much weight to give to innate ideas prefigured in the mind (rationalism) or how much weight to give to the evidence of the senses (empiricism). The rationalists played up the role of innate ideas that give shape and order to the input of the senses. The empiricists thought that the mind began as a blank tablet *(tabula rasa)* and that there was nothing in the mind that did not first enter it through the senses.

Immanuel Kant cut the Gordion knot of the battle between the rationalists and the empiricists by answering Descartes' doubt as follows: True, I can have no knowledge of things in themselves (*noumena*). I can know only phenomena, that is, things *as they appear to me* through the lenses of my categories of understanding and forms of intuition. The rationalists were right about the role that categories (innate ideas) play in giving shape and form to our knowledge of phenomena. The empiricists were right about the need for external inputs through the senses. As Kant put it in one of his unusually lucid formulations, "Concepts without percepts are empty; percepts without concepts are blind." Both the rationalists and the empiricists had contributions to make to the story of how knowledge is possible. Both Hume and Leibniz were right to be skeptical about our

knowledge of things in themselves (noumena). But both were wrong to remain skeptical. Science *is* possible. We *can* overcome Descartes' doubt. But the science we end up with is a science of phenomena, of appearances, not of things-in-themselves.

The entire tradition from Descartes to Kant was a series of different attempts to overcome Descartes' doubt about the veracity of human knowledge. Philosophy became almost entirely preoccupied with epistemology, the theory of knowledge. Just a glance at the titles of many of the landmark books is enough to demonstrate the appropriateness of the label that Richard Rorty has given to this chapter in the history of philosophy, "the subjective turn:" Descartes' *Discourse on Method (1637)* and his *Meditations (1641);* Locke's *Essay Concerning Human Understanding (1690)*, Leibniz' *New Essays on the Understanding (1704);* Berkeley's *Treatise Concerning the Principles of Human Knowledge (1710),* Hume's *Inquiry Concerning Human Understanding (1758),* Kant's *Critique of Pure Reason (1781) —* these are attempts, one and all, to overcome the challenge of Descartes' doubt, his radical skepticism regarding the veracity of human knowledge.

The stage is now set for the entry of Hegel, who acknowledged that Kant provided his point of departure. Hegel's first great work, his *Phenomenology of Spirit (1807),* takes a fundamentally different turn from all of the works just cited as parts of "the subjective turn." He doesn't return to the pre-Cartesian ontology of the Greeks; nor does he make an objectivist turn toward a naïve empiricism; nor does he accept Kant's solution—that we know only appearances but not things in themselves. Instead he takes a *historical-phenomenological turn*. And what might that be?

Rather than trying to guarantee the veracity of ideas inside the mind by trying to justify their correspondence to their objects outside of the mind as the epistemological tradition had tried every which way to do, Hegel argued that we should take the whole history of such attempts as itself the datum. Rather than accepting Descartes' doubt, Hegel doubted the doubt. Rather than assuming that we are prisoners in a fortress of subjectivity, forced to test every message from the outside world for signs of its veracity—"Could the guards [the senses] be deceiving me? Could they be telling me the truth?"—Hegel jumps to a very different perspective. He says, in effect, "I will accept all reports at face value." This is what is meant by 'phenomenology'—accepting the phenomena as they appear rather than immediately interpreting them as messages from some other or "outside" world. This is the *phenomenological* part of "the historical-phenomenological turn." *I will take these reports and compare them, not with some risky and unverifiable inferences about an outside world, but instead I will compare them with one another. I will look at the succession of such ideas, and I will let them speak for themselves* about each other. *The record of their succession then becomes the record of the coming to be of consciousness.* (This is a summary of the argument contained in the Introduction to Hegel's

Phenomenology.) This is the *historical* part of the same mouthful, "the historical-phenomenological turn."

When you turn to the early pages of the *Phenomenology of Spirit*, to the section titled "Consciousness," and, more specifically, to the sub-section titled "Independence and Dependence of Self-Consciousness," what you find there is not one more chapter of the subjectivist turn. In place of the armchair of the perplexed and skeptical epistemologist trying to cogitate his way out of the prison of the Cogito, you find yourself wracked with *desire.* You find yourself locked in a "life and death struggle" over who shall be the "master" and who shall be the "slave" who must engage in labor that "shapes and fashions the thing" for the sake of the gratification of the master's desire. This is not one more armchair "inquiry into human understanding." This is a record of consciousness' tortuous climb up from the jungle of animal existence. There is more of Hobbes and Rousseau in Hegel than meets the epistemologist's eye, and hence a foretaste of Darwin. Hegel's dialectic traces the sequential record not of different species and their struggle for survival, but of different forms of consciousness (*Gestalten des Bewusstseins)* and *their* struggle for existence.

This dark passage, usually referred to as "the master-slave dialectic," was the subject of an influential set of lectures given by Alexandre Kojève. Between the two world wars. In Paris from 1933 to 1939, Kojève gave a series of lectures on Hegel that were attended by a remarkable group of students including Jean Paul Sartre, Maurice Merleau-Ponty, and Jacques Lacan. French theories of the self, from existentialism through post-structuralism, were much influenced by those lectures.

In his reading of Hegel, Kojève stressed the role of desire in the development of consciousness. He renders fully explicit a point that is only implicit in Hegel's text, namely, that the nature of the consciousness that is the source of desire will be determined by the nature of what is taken as the object of desire.

> The I of Desire is an emptiness that receives a real positive content only by negating action that satisfies Desire in destroying, transforming, and "assimilating" the desired non-I. And the positive content of the I, constituted by negation, is a function of the positive content of the negated non-I. If, then, the Desire is directed toward a "natural" non-I, the I, too, will be "natural." The I created by the active satisfaction of such a Desire will have the same nature as the things toward which that Desire is directed: it will be a "thingish" I, a merely living I, an animal I. (Alexandre Kojève, *Introduction to the Reading of Hegel,* trans. James Nichol, New York, Basic Books, 1969, p. 4.)

If, on the other hand, desire takes for its object another person, another self-consciousness, another desire, then that first desire will constitute an I that is not "thingish" but instead self-conscious. As Feuerbach later put it, to the eventual delight of

many a health food store owner, "You are what you eat." *Man ist was er isst.* This is the Hegelian basis for the adage of common sense that your character will be largely determined by the company you keep. Spend your time relating to things and your character will be "thingish." Spend your time in relation to other autonomous beings who are capable of both gratifying and resisting your wishes, and you will grow into a more highly developed, more fully self-conscious person. As tennis players say, if you want to improve your game, it's best to "play up."

The dynamic by which the self comes to be is *the play of recognition.* While this term (*Annerkennung)* has a very particular meaning in the context of Hegel's philosophy, it seems fair to say that what Hegel is describing is a familiar and pervasive phenomenon. In oriental cultures, people speak of *saving face.* In the argot of the 'hood, the brothers speak, and Aretha Franklin sings, about the importance of r-e-s-p-e-c-t. While there might be a temptation to write off respect as a sub-cultural peculiarity, the cross-cultural universality of the need for recognition should lead us to give it greater respect. Just because some Californian management consultants offer workshops on the importance of mutual "acknowledgment" in the workplace—always telling your work mates what a *great* job they are doing—we should not dismiss this stress on acknowledgement as so much pop psychology. Hegel and Kojève are making a claim about the *ontology* of the self, not its psychology. They are claiming that the self comes into existence as a result of the kind of recognition it receives from other selves. Not even the *Cogito* is indubitable. It is not *given.* Instead, consciousness that is self-conscious and not just a "thingish" consciousness comes to be as the result of the play of desire and the contest for recognition.

So now we must add a fifth component to the system of four mutually implied concepts: recognition. Recall from some pages ago, prior to the tour through the subjectivist turn to Hegel's deduction of desire:

Synthesizing the insights of Kauffman and Hegel (as interpreted by White) we arrive at a systemic set of mutual implications among four concepts: (1) law, representing the realm of *necessity;* (2) desire, which wants what it wants in the face of what supposedly *must be;* (3) narrative, or story, which tells the tale of the conflict between desire and law, freedom and necessity; and, finally, (4) the emergence of subjectivity as constituted by desire's struggle with necessity, as told in the form of biography or history, and self-referentially re-told as the self-constituting auto-biography of an emergent subject.

Now we add (5) the contest for recognition as a condition for the emergence of a non-thingish, truly human consciousness.

With Kojève's help, we have now retraced the steps starting long ago and far away, from Descartes' doubt through Hegel's dialectic of mastery and slavery to the very point that

sent us off on this long detour: the mutual interdependency between conscious subjectivity and an object of desire. We were looking for "a way that threads a narrow path between nihilism on the one hand and teleological essentialism on the other. By articulating the co-evolution of subjectivity and teleonomy, of subject of desire and object of desire, of consciousness and purpose as the two sides of one emergent system, the present argument turns two mysteries into each other's solutions."

Now that we have traced the path of this argument up through Kojève's lectures in the 1930s, what can we learn from what Kojève's listeners did with these insights? Did Sartre or Lacan thread that narrow path between nihilism on the one hand and teleological essentialism on the other?

Sadly, the answer is no. Sartre based his analysis of "Concrete relations with others" squarely on Hegel's master-slave dialectic, (Cf. Jean-Paul Sartre, *Being and Nothingness,* trans. Hazel Barnes, New York, Humanities Press, 1956, p. 370: "Up to this point our description would fall into line with Hegel's famous description of the Master and Slave relation.) but ended up concluding that the self can never confidently cross 'the reef of solipsism.'" Sartre turned Hegel's master-slave dialectic into a lengthy disquisition on the inescapability of sado-masochistic relationships. He defends the lengthiness of his treatment of sadism and masochism, "First because they are fundamental, and second because all of men's complex patterns of conduct toward one another are only enrichments of these two original attitudes (and of the third—hate . . .)" (*Ibid.,* p. 407; see also James Ogilvy, "Mastery and Sexuality," Hegel's Dialectic in Sartre and Post-Freudian Psychology," *Human Studies 3,*1980, pp. 201-219) What of love? "In fact, if the sadist is pleased upon obtaining a denial by means of torture, this is for a reason analogous to that which allows us to interpret the meaning of *Love.* We have seen in fact that Love does not demand the abolition of the Other's freedom but rather his enslavement as freedom; that is, freedom's self-enslavement." (*Ibid.,* p. 403.)

For Lacan the story is no better, just less accessible. Given the obscurity of Lacan's own language, allow a secondary commentator to tell the tale. After reviewing the case of Aimée, a young woman condemned for murder while suffering a paranoid delerium, Carolyn Dean writes:

> Lacan used Kojève's discourse about the constitution of human consciousness to transform what was *already* implicit in Aimée's crime into a theory of human development in which the struggle for freedom, for autonomy and subjecthood, is inseparable from an ever-unsatisfied desire for recognition. The crime propelled by the desire for the other's desire reveals that the struggle will never cease, because no "other" desire will ever compensate for the primordial lack at the heart of human consciousness. Aimée's crime, in other words, dramatizes the tragic human struggle to be free of others, whose recognition human beings always

> desire: the struggle to be free of the mirrors that are *at once* the source of human slavery and human pleasure. (Carolyn J. Dean, *The Self and Its Pleasures: Bataille, Lacan, and the history of the Decentered Subject,* Ithaca, Cornell University Press, 1992, p. 52.)

Those *"others"!* Can't live with 'em, can't live without 'em! The trail from Hegel's depiction of the role of desire in the constitution of self-consciousness leads through Kojève's lectures to a post-Freudian psychology that cannot escape a vicious circle of sadism and masochism. To make a very long story very short, the trail leads, to repeat for emphasis, to a "tragic human struggle to be free of others, whose recognition human beings always desire."

The human prospect was not much brighter for the master himself, Sigmund Freud. In the very concluding paragraphs of his very late work, *Civilization and its Discontents,* Freud averred, "that when one surveys the aims of cultural endeavor and the means it employs, one is bound to come to the conclusion that the whole effort is not worth the trouble." (Sigmund Freud, *Civilization and its Discontents,* trans. James Strachey, New York, W. W. Norton & Co., 1961, pp. 91f.) Why so glum? Several reasons, which are worth sorting out, for the fashionable malaise drifting westward from France and Germany may just be unnecessary, nothing more than a sad detour in a history that has more to offer.

One glaring reason for the sadness is the appalling bloodshed during the two world wars. How could the civilization that bred Kant and Hegel and Bach and Beethoven have brought upon itself the atrocities witnessed in the first half of the twentieth century? The Hegelian dream of absolute knowledge turned into a nightmare. Why? That question leads to the second reason for Freud's pessimism: An analysis of the fundamental forces of psychic construction that leaves no room for a happy ending to the human drama.

Freud's view of desire suffered from the phallusy of misplaced physics (see Part Four below). Freud regarded the play of erotic impulse, libido, as a zero-sum game. "Since a man does not have unlimited quantities of psychical energy at his disposal, he has to accomplish his tasks by making an expedient distribution of his libido. What he employs for cultural aims he to a great extent withdraws from women and sexual life." (*Ibid.,* pp. 50f.) There's no free lunch. You can't get more out of less. Freud thought of libido as strictly limited. His theory of the sublimation of libido into the works of culture and civilization rests on an economic or hydraulic metaphor that won't allow us to get more out of less (Holland's definition of emergence). Quite to the contrary, the sublimation or redirection of primitive libido into civilized culture is bound to suffer from leaky plumbing:

> If civilization requires such sacrifices, not only of sexuality but also of the aggressive tendencies in mankind, we can better understand why it should be so

> hard for man to feel happy in it. In actual fact, primitive man was better off in this respect, for he knew nothing of any restrictions on his instincts. As a set-off against this, his prospects of enjoying his happiness for any length of time were very slight. Civilized man has exchanged some part of his chances of happiness for a measure of security. (Ibid., p. 62.)

No wonder Freud was glum. With a zero-sum theory of sublimation, civilization reduces to a lot of sound and fury signifying if not nothing, then at most, less libidinal satisfaction than it started with in the caves.

Mix Hegel, Marx, Nietzsche and Freud and what do you get? Fascism and Communism—totalitarian ideologies that combine the totalizing impulse in Hegel, the class struggle from Marx, the will to power from Nietzsche, and a doomed battle between Eros and Thanatos in Freud. This is a devil's brew if ever there was one. No wonder a psychoanalyst as erudite in this tradition as Joel Whitebook would conclude his book, *Perversion and Utopia,* with a paragraph that begins, "A harmoniously integrated self and life history, as envisioned in the classical bourgeois ideal of *Erfahrung,* are undoubtedly impossible today." (Joel Whitebook, *Perversion and Utopia,* Cambridge, Mass., The MIT Press, 1995, p. 262.)

"Undoubtedly impossible?" Here, if anywhere, Descartes' doubt is worth resuscitating. The next and concluding section to Part One develops an alternative theory of sublimation, one that borrows from Hegelian sublation *(Aufhebung)* a movement that cancels, preserves, and transcends the forces it synthesizes. But, contrary to Hegel, dialectical synthesis does not require the kind of universal reach that leads to totalitarianism. Hegel was both a monotheist and an heir to monological science. If we sort out the monological grandiosity from the Hegelian tradition, and recast Freudian sublimation as non-zero-sum sublation, there's just a chance that we can find access to the romantic sublime. (Cf. Thomas Weiskel, *The Romantic Sublime,* Baltimore, Johns Hopkins University Press, 1976.)

(i) The sublimation of libido into love

The riddle of sublimation is equivalent to the riddle of emergence. In both cases, it is a question, as John Holland defines emergence, of "much coming from little." (John Holland, *Emergence,* Cambridge, Mass., Perseus Books, 1998, p. 1.) What's to puzzle over? Isn't it quite obvious that much can come from little? Just look around. Look at the diversity of species. Look at the richness of literature all built from just 26 letters.

We know that the complex can be created from the simple. The question is how *without* the help of a creator. Books are written by authors. How does natural order emerge from chaos? Evolution is the alternative to a creator god, but Stuart Kauffman says that we

don't really understand evolution, only the paring away due to natural selection, not the creation of new forms, not the expansion into the new workspace of the adjacent possible. To understand the evolution of genuinely new forms, we need to understand emergence. And that's what this book is radically all about.

Monological science will not allow an understanding of emergence —*in principle.* If to understand complex entity C is to reduce it to descriptions of its simple components, s,t,u . . . and z, then we're bound to see C as *nothing but* rearrangements of s,t,u . . . and z. Love is bound to be *nothing but* civilized window dressing on primitive libido. You can't get more out of less. Monological science is built upon the law of the constant conservation of mass and energy. You may get a great deal of energy out of very little mass in, say, a nuclear explosion. But even then, the constancy of mass and energy is preserved by that famous little equation, $e = mc^2$.

To the extent that monological science is built upon the symmetry of prediction and reduction, then it is in principle impossible for monological science to explain novelty, anything genuinely new under the sun. If explanation depends on reduction, then explaining the new amounts to reducing it to reconfigurations of the old. Prediction is possible because time doesn't really have any surprises, just a succession of state descriptions of the same old matter and motion in space and time.

There's no explaining love. Feyerabend was right. It's a gift. Nor, therefore, can you predict just who will fall in love with whom. This much is true to our experience. But does it follow that love is therefore impossible, that there is no such thing, that love is an illusion or, as Paul Goodman opined, "a pathological condition from which we all emerge with luck?"

To accept this cynical conclusion as it wafts westward from love's French deconstruction would be foolish. But to reject such cynicism without a pretty good argument is just callow. This section constructs a pretty good argument based on the hierarchy of desire we've traced thus far, and the possibility—*not* the predictable necessity—that it ascends upwards from hunger, animal attraction, and libido to something like love.

Why the hesitancy of the phrase, "something like love?" Because it is in the nature of the case that love, if it is to be an emergent phenomenon that is something new under the sun, should be *constructed* and not simply *given.* As La Rochfoucauld famously put it with a heavy dose of the subjunctive that French allows, "There are those who would not love as they do had they not read of love." Romantic love is a construct of culture, but not for that reason less real than, say, hate, or good manners for that matter. Love is something new under the sun. It did not exist in the Pleistocene era. There's reason to doubt whether what we now know as romantic love existed in Periclean Athens. When you compare the speeches in Plato's *Symposium* with what we have inherited from Shakespeare and the

French troubadours, the cultural history of love is unmistakable. Yes there is a constancy sufficient to make us suspect that Socrates, Shakespeare and Erich Segal (classicist and author of *Love Story)* are all talking about roughly *the same thing.* But, no, the use of similar words in the observation language should not lead us to believe that these words, either, are not "theory-laden" and context sensitive.

Love is a feature of what Karl Popper calls "World 3." Popper defines "World 1" as physical reality. "World 2" is the realm of subjective conscious experience. World 3 is "objective mind," which W. W. Bartley describes as, "the *logical contents* of books, libraries, computer memories, the logical structure of arguments, the objective problem situation at any time in a particular science." (W. W. Bartley, in *Evolutionary Epistemology, op. cit.,* p. 33. Cf. also pp. 87, 116, 149, 157, and 159-160 for comments on World 3 by Popper and others.)

The contents of World 3 are not natural; they are cultural; they are social constructs, but not for that reason any less real than, say, Beethoven's *Fifth Symphony.* Nor are the contents of World 3 any less *physical* for the fact that they are not natural. Beethoven's *Fifth* exists physically in the form of a written score, as performances played on instruments, as sound waves, and as recordings etched in vinyl and on CDs. Beethoven's *Fifth* is part of the inventory of the universe, along with the Golden Gate Bridge and the Sahara Desert. Beethoven's *Fifth* has an identity. It is recognizable by many in its several different performances and manifestations. You can see its score. You can hear it. You can play it on the piano. You can enjoy it. You can't smell it or taste it. You can, after a fashion, feel it. It is an *object* that can be sensed; it *objects* if you mistreat it, say, by playing it on an accordion. You can botch it. You can interact with it. It retains its own integrity, and by most accounts, it is quite beautiful. The universe would be less without it.

So likewise love is a construct, but not for that reason less real than hydrogen. It has come to be. It can pass away. Love has emerged and, like other emergent phenomena, what has come together can also come apart. Love can be strong; love can be fragile.

Love is an emergent level on the hierarchy of desire whose lower levels include the hunger of uni-cellular organism swimming upstream in glucose gradients, as well as the libido of Freud's uncivilized savages. Love is a sublimation, or dialectical sublation, of that lower level desire, but not for that reason reducible to *nothing but* redirected libido. As announced at the very opening of this section, "The riddle of sublimation is equivalent to the riddle of emergence." While Freud's theory of sublimation operated within the zero-sum logic of monological science, solving the riddle of emergence allows a non-reductionist reading of sublimation.

Paul Ricouer developed a non-reductionist reading of Freudian sublimation in his book, *Freud and Philosophy.* (Paul Ricoueur, *Freud and Philosophy: An Essay on Interprentation,* trans. Denis Savage, New Haven, Yale University Press, 1970*.)* Ricouer distinguishes between what he called "a hermeneutics of suspicion," and "a hermeneutics of belief." The former is reductionistic; the latter amounts to a more hopeful interpretation of the symbolic products of sublimation. The former will say that art (to put it very crudely) is *nothing but* a sublimation of the smearing of feces; the latter will say that art can create something new under the sun whose meaning and beauty cannot be reduced to rearrangements of the same old stuff. The symbols may be the same—a painting, a piece of music, a play—but the interpretations will be quite different depending on whether one approaches those symbols with a hermeneutics of suspicion or a hermeneutics of belief.

Ricouer's distinction can serve as a bridge toward yet higher levels on the hierarchy of desire, beyond romantic love to the shared hopes of a community for a better future. A better future is a beautiful future. The vision of a better future can motivate a community to action much in the way that an object of love motivates a lover. To the extent that the articulation of a normative scenario motivates a community toward the action required to realize that future, the articulation of a normative scenario acts like a telos. Scenario planning provides alternative interpretations of the present as the first chapter to several different futures, some good, some bad. A normative scenario of a better future therefore acts like a *telos* luring a community toward a better state of being. Teleological behavior is no mystery where the articulation of a normative scenario has efficacy in the present. Causality isn't running backwards from the future to the present. But it does make sense to speak of a "downward causality" where the emergence of symbolic consciousness exercises an influence on the manipulation of matter in the present—say, the building of a dam to create a water supply and electricity in the future.

Now it is worth recalling the point made earlier about evolutionary epistemology: that the future has replaced objectivity as the horizon of validation. As W. W. Bartley put it: "There is no justification, ever, . . The process that began with unjustified variations ends with unjustified survivors." The future will tell us which interpretation of the present was correct. *We'll see* whether the hermeneutics of suspicion or the hermeneutics of belief makes more sense.

Uncertainty regarding the interpretation of the present is surely well justified because the future is not predictable. Time is real. It is not just the moving image of eternity, or the unrolling of a predestined plan. But to the extent that communities can frame a more beautiful future in the form of a normative scenario, then their shared hopes can act like a desire that constitutes a telos. And this telos then exercises downward causality. It informs present action toward the end of that better future. But nothing is guaranteed.

Our blessedness consists precisely in the fact that a happy ending is not guaranteed, that no amount of monological science is sufficient to make beauty predictable. Just as it can all come together, so can it come apart. The center may not hold. Since the interpretation of current events is contingent on their outcome, what *will* "come true," and the outcome of current events is genuinely in doubt—uncertain, subject to the unpredictability of the future—we are therefore pressed nose to the glass against the *sublime*, which the romantic poets and philosophers found to consist not in some high plateau but in the contrast between the heights and the depths. That's why the romantic poets, particularly Wordsworth in the Simplon Pass passage of the *Prelude*, rely so heavily on the imagery of the Swiss Alps as a favorite medium in which to express their sense of the sublime. It's not just the heights that impressed them, but the cliffs, the steep descents from lofty peaks down into the abyss. High mountain waterfalls do it. You could watch them for hours. Big surf does it, with rhythm. The collapse of the World Trade towers did it as tragedy. Orgasm does it as ecstasy.

This nearness of the depths from the perspective of the heights, this sense of immanent peril that the romantics associated with the experience of the sublime, is also present in (a) the paradoxical combination of the strength of love and the fragility of love, and (b) the nearness of tragedy to comedy.

Without offering anything close to a deductive proof or argument, Plato hints at these associations toward the end of his dialogue on love, *The Symposium.* After several different definitions of love have been served up and discussed, Socrates tells a tale in which a wise woman named Diotima describes a ladder of love leading from the pleasures of the flesh, with one lover then with many, up through the pleasures of the mind, ending finally with the love of wisdom, philosophy. Then, at the very end of the evening . . .

> All of a sudden, just as Agathon was getting up to go and sit by Socrates, a whole crowd of revelers came to the door, and finding it open, as someone was just going out, they marched straight in and joined the party. No sooner had they sat down than the whole place was in an uproar; decency and order went by the board, and everybody had to drink the most enormous quantities of wine. By this time Eryximachus and Phaedrus and some of the others were beginning to leave, so Aristodemus told me, while he himself fell off to sleep.
>
> He slept on for some time, for this was in the winter and the nights were long, and when at last he woke it was near daybreak and the cocks were crowing. He noticed that all the others had either gone home or fallen asleep, except Agathon and Aristophanes and Socrates, who were still awake and drinking out of an enormous bowl which they kept passing round from left to right. Socrates was arguing with the others—not that Aristodemus could remember very much of what he said, for, besides having missed the beginning, he was still more than half

> asleep. But the gist of it was that Socrates was forcing them to admit that the same man might be capable of writing both comedy and tragedy—that the tragic poet might be a comedian as well. (Plato, *Syposium,* 223)

We face a fold, not unlike the "catastrophe" of the fight/flight reaction. The future is indeterminate. The significance of current events is a function of *what will come true*. Therefore, the significance of current events could be comic, it could be tragic. We don't know. We can't know. But we must act in any case. So best to plan our actions with both possibilities in mind. It's good to have some alternative scenarios at hand when trying to act with compassion and intelligence.

This endemic uncertainty that accompanies action in time reflects the structure of desire. Desire is initially inchoate. At first, it doesn't know what it wants. But in trying to find what it wants, desire drives toward the evolution of ever higher levels of emergence.

The structure of desire looks something like a cell (or cells) in a late stage of mitosis, the divided chromosomes in each half pulled almost together at their antipodes, but splaying apart as they approach the equator and soon-to-be waist of a cell undergoing mitotic division. This image provides a model for the linked evolution, the co-evolution of subjectivity and objectivity, driven, as Hegel said, by desire. The dark lines point toward a common origin in subjectivity, as well as to the unity of the object. But this image does not show a singular subject. There is no dot or nucleus to this process. The process *is* the nucleus. And the process is one of division, not unification. Mitosis is about one becoming two, not the other way around. The image is one of a unified *experience* dividing toward the directions of objectivity and subjectivity, without presupposing The Object or The Subject.

The "structure of desire" is actually a hierarchy of functions, wants fulfilled or denied. Inchoate desire that doesn't know what it wants, but acts nonetheless just because it *wants,* and then becomes *educated* by its experience of frustration and gratification. Gradually desire gets trained, in an altogether Darwinian fashion, to know which desires can survive, and which will perish in the jungle of human interaction. This *education of desire* is what a good upbringing provides by discouraging "thingish" gratification in favor of human love. A good upbringing, a good education, habituates the soul in the mutual recognition and respect of other self-consciousnesses.

In considering such an education, of course the idea of sublimation comes to mind, but not the zero-sum game of Freud's trade-off between civilization and instinctual gratification. There can be positive sum games in which we get more out of less, games in which desire figures centrally, and gratification is achievable. The education of desire, its sublimation into love, produces a kind of desireless desire: a desire for the pleasure of the other more than one's own pleasure. When two people *come together* with this kind

of mutual intentionality, when her pleasure is his greatest desire and his pleasure hers, then the results are quite different from what a selfish hedonism would produce. Instead a kind of Tantric bliss ensues, an extended play of eroticism without goal or end, a high plateau of ecstasy unparalleled in other human experience. . . with the possible exceptions of intense aesthetic gratification or spiritual rapture.

The more educated or sublimated the desire, the less episodic or immediate its gratification. At lower levels on the hierarchy of desire, satisfaction is simple: quenched thirst, sated hunger. At higher levels, the structure of desire is more mediated, e.g. through the sublimation of libido into love, or as the shared hopes of a community articulated in a normative scenario of a better future. No single episode will mark the realization of shared hopes. But a better future can emerge *if it all comes together*—health, wealth, good education, art, all of the components of a vibrant community. Likewise, no single event will establish love once and for all; but a life lived in love is not impossible, despite what Sartre and Lacan may have thought.

While the gratification of desire on the lower rungs of the ladder of love is fairly simple, on higher rungs it is, precisely, *complex.* We need complexity theory to appreciate those higher levels. We need a theory of self-organizing systems, because monological science cannot predict the unpredictable. Instead we need a theory of emergent systems that can accommodate the coming together of (1) a realm of necessity described by monological science; (2) desire that is at first inchoate, but learns what it wants through a Darwinian process of trial and gratification; (3) narrative, or story, which tells the tale of the conflict between desire and law; (4) the emergence of subjectivity as constituted by desire's struggle with necessity, as told in the form of biography or history, and self-referentially re-told as self-constituting autobiography or community history; (5) a struggle for recognition in which the subject gains the respect and love of another self-consciousness; and (6) a sense of the evanescence or fragility of the emergent system—that what has come together could just as well come apart, that comedy could turn into tragedy, and that nothing is guaranteed. Therein lies the sense of the sublime: in the nearness of the abyss to the experience of bliss.

We face folds within folds. Just as the sublime is a function of the nearness of comedy to tragedy, or bliss to the abyss, so comedy and tragedy, each in themselves, are structured around duplicity. Laughter is what the body does when its mind is caught in the cusp of a fold between two planes of a *double entendre.* If the mind were convinced of the definitiveness of either plane, if it were trapped in the single vision of literalism, then there would be no humor, no laughter. If the mind moved in a linear fashion from one interpretation as the first instance of meaning, and then on to the second instance of meaning as through the linear order of premises in a logical syllogism, again there would be no humor, no laughter. As they say, you can't explain a joke. If you have to explain it,

then they can't *get it.* If they do get it, then there is no first meaning followed by a second. There is instead an all-at-once, synchronic duplicity. This semiotic ambivalence shakes the body out of the smooth linearity of literal information processing into a shuddering oscillation between parallel planes of meaning. This semiotic oscillation of symbolic meanings manifests itself physically as laughter.

Likewise tragedy consists in more than mere pain. Whether as interpreted by Aristotle as revolving around a reversal in which the hero is brought down by a tragic flaw *(hamartia),* or as interpreted by Hegel as an irresolvable conflict between conflicting goods, in either case the structure of tragedy involves duplicity. We weep not simply from the experience of pain, but from a sense of the contrast between what is and what so nearly might have been.

In both cases, in laughter and in tears, there is downward causality: a direct and physical manifestation of a mental oscillation between alternative interpretations of the same old stuff. Laughter and tears are the incarnation of comic and tragic interpretations. Humor and despair both emerge from alternative interpretations of the literal. Humor practices a hermeneutics of belief; tragedy, a hermeneutics of suspicion.

With laughter, the body flaps as the mind flips between alternative interpretations of the same data. The body flapping is laughter. The mind flipping *is* the body in the very same constellation of neuropysiological states flipping between alternative semantic referents for one and the same constellation of signals. *Signals* are physical things: networks of synapses; landscapes of fitness thresholds; varying levels of saturation of different endorphins. But these physical things can exhibit semantic features. They can *refer.* They can and do manipulate information. They can store information. Laughter happens when the same constellation of neurophysiological states—one and the same body—simultaneously refer to two different semantic referents, two different interpretations of that same set of signals.

Folds within folds. One doesn't know whether to laugh or cry when contemplating the thousands of pages wasted on 'the problem of other minds.' Once defined as *thinker,* the Cartesian subject got written down into a bottomless pit of solipsism and skepticism about the Other. Cartesian doubt—the main product of his *Meditations* —worked like a corrosive on the subject's links with the objective world.

If, instead, you accomplish a Copernican revolution in the relationship between consciousness and desire, if you see consciousness as a satellite of desire rather than desire as a satellite of consciousness, suddenly many of those skeptical disconnects between subject and object that resulted from starting with the Cogito simply disappear. When you stop looking through Cartesian lenses you see a limbic self that is constituted by its relationships such that the question of access to others, much less questions about

the existence or knowability of *the other,* simply never come up. Are Castor and Pollux twins? Could either ever doubt it?

Only if you'd been utterly taken in by the rendering of subjectivity that ran from Rene Descartes to Immanuel Kant would you fall for a story about subjectivity whose plot centered on the solitude of the Cogito of Descartes' famous line, *Cogito ergo sum,* "I think, therefore I am." In Sartre's *Being and Nothingness (L'Etre et le Neant),* you see a subject who, relying on thinking alone, sinks into a skepticism that makes it impossible to be certain about knowledge regarding any *Other.* Relying only on the cogitation that qualifies it for indubitable existence, Sartre's For-itself *(Pour soi)* is irrevocably alone.

But see how this whole story turns on the lack of multimedia: The single medium of cogitation—thinking—unaided by such sensual and intuitive talents as smell, feeling, and perception of cues and signals given off by "the other" and received by "the self." Note that even here—and hence the quotes around 'other' and 'self'—we presuppose a monological ontology in which *being* belongs to the particular, to the individual, to the node at the end of a relationship, rather than to the web, to the field, to the network of relationships. What we see when we look through the lenses of a relational worldview is a self that is constituted by and does not precede those sensual and intuitive interactions like smell, feeling, and the perceptual cues and signals. To the extent that the self is constituted by its relationships with "the other," skepticism regarding the existence of the other is absurd.

*Video, ergo sum? "*I perceive, therefore I am?" No, the relational worldview sees things such that, *Video,* therefore the other *is*. Seeing *is* believing, at least insofar as assuring us that there *is* some other. If there weren't, there would be no relationship, and if no relationship, then no self either. If I am, the other is just as surely. So: *Video, ergo alter est.* I perceive, therefore the Other *is.* What constitutes me—my relationship with the other—constitutes the Other as well. My boundary has an inside and an outside. If it had no outside, there would be no inside. My boundary is my relationships. My relationships are my boundaries.

When you see through the lenses of the relational worldview, you don't know whether to laugh or cry when you look at the millions of miles of print wasted on "the problem of other minds."

Cogito ergo sum? No no, Rene. I desire, therefore I am. I love, therefore I am. *Amo ergo sum.*

Eighth Trait: Coming Apart

Do all emergent systems face a fold? It seems so. Whatever comes together can just as easily, just as suddenly, come apart. From *pop* . . . to plop!

As mysterious as *life* is to monological science, so also *death.* The very same molecules that were at one moment alive . . . suddenly ceased to be alive. The living body became a dead corpse.

As mysterious)as consciousness may be to monological science, so also the moment of sinking into sleep. The very same brain that was perfectly conscious a moment ago . . . slides into unconsciousness.

Jared Diamond's *Collapse (Viking,* 2005) is an extended meditation on examples of social and ecological collapse, from a doomed society of Vikings in the early history of Greenland to the mysterious disappearing civilization that once thrived on Easter Island.

To be an emergent system is to be the sort of thing that can suddenly come apart. Coming apart is the flip side of emergence. Part of what it is to really *live* is to be the sort of thing that dies, and knows that it will die. This is the important truth contained in Heidegger's dark teaching of Being-towards-death. Living on the existential edge lends life a certain *gravitas.*

Sure, it's possible to appreciate death's role as the ultimate "pruning algorithm." (Cf. Harold Morawetz, *The Emergence of Everything.*) Without death, further evolution is all but impossible. Without some pruning, we're condemned to the perpetuity of the same old, same old. Death prevents that. When you see death as the pruning algorithm for the further emergence of life, it all makes a certain kind of sense . . . But such thoughts about the benefits of death for the furtherance of the species do little to lighten the leadenness of death for the single, living, existing individual.

To people familiar with the momentum of large institutions, their sudden dissolution always comes as a great surprise. Think of the fall of the Berlin Wall, the sudden dis-union of the Soviet Union. Think of the end of Enron. This capacity for sudden death is as clear an indicator as there could be of the extent to which such institutions *are* emergent systems, not just substances subject to the laws and dynamics of monological science. They do not *defy* the laws of monological science. Insofar as they are physical, they are subject to the law of gravity. They cannot levitate. But insofar as they are emergent systems, to understand them is to understand the way they fit together, and their vulnerability to coming apart.

Where monological mass admits of continuous accretion—more and more of the same—an emergent system like morality will tip over into decadence if pushed to an extreme.

Virtue is not additive beyond a certain point. Just look at the tendency toward scandal exhibited in every religious tradition that tips towards asceticism.

Health is an emergent system. It has no single cause. It's all about balance and vitality. But our health *system* has gone monological, and for just that reason our health is coming apart. Sickness is the coming apart of health. Depression is the dissolution, the coming apart of happiness.

To acknowledge the perpetual threat of devolution—death, disintegration, decadence, depression, illness, collapse—is to distinguish oneself from those who find too easy comfort in the idea of evolution as progressive. Yes, there is a ratchet effect built into the confluence of contingency, variation and selection, and the laws of emergence. Stir these pieces together and you're bound to get more complex forms evolving even more complex forms. But the general direction toward greater complexity is a very jagged path, nothing as direct as the so-called "march of progress." And who's to know when you're approaching a peak and about to plunge into a valley? Even if there's a taller peak on the other side of that valley, you don't know while on the way down whether you'll ever get as high up as you've just been ever again. So much for evolution as progress. Maybe for the species, but your entire lifespan may be spent on the downside of some epochal peak of human progress. You could have the misfortune of living in a decadent era. The very idea of progress many seem woefully inappropriate to anyone caught in a temporary sink of history.

The fact that death—coming apart—is a necessary feature of emergent systems lends a certain existential exigency to life. "To be, or not to be, that is the question." This either/or, stressed by Shakespeare, Kierkegaard, and Heidegger, puts a perpetual fork in the path of so-called progress, a fold in the path of human history. However powerful the hope for higher planes of order and complexity, nothing is guaranteed, not life, not salvation, not redemption, not the Omega point. Nonetheless, the emergence of novelty is possible. It has happened before. It could happen again.

Part Two: Orders of Emergence

Having completed Part One's pass through the eight traits of emergent systems, now it is time to take a series of horizontal cuts across the rows in figure one, first life, then evolution, then language.

What is life?

First, life. What *is* life? The thought paths of monological science lead almost inevitably to the illusion of an *élan vital,* a spirit of life, a life force whose simple presence would cause life much as meteors cause the cavities called craters. But so-called *élan vital* does not exist. It is an illusion created by the will to see the physical world as the effect of some hidden cause that lies behind the actual world we see. (For an anatomy of this impulse to illusion, see F. Nietzsche, *Twilight of the Idols,* tr. Walter Kaufmann, in *The Portable Nietzsche,* Viking, New York, 1954, pp. 485f. On this passage in Nietzsche, see Heidegger, *Nietzsche,* Neske, Pfullingen, 1961, Vol. 1, pp. 235-242. See also Nietzsche, *Will to Power,* tr. Kaufmann, Vintage, New York, 1968, ## 507, 553-569, pp. 275f., 300-307.) This quest for a single presence to explain life is a theological way of looking at things, creating a creator who lives above and behind the created. *Elan vital*—the *spirit* of life—is a *spiritual* thing. Isn't it funny how the rails of monological rationality lead so inevitably toward a spiritual entity? Isn't it funny how so-called science should take us on a path toward monotheism?

Describing life in terms of the eight traits of emergent systems amounts to transferring the discourse about life from the realm of religion to the discipline of science. Extirpating magic is the means; explanation the goal. Showing how life exhibits the eight traits, and how applying the eight traits explains life, amounts to both an explanation of life *and* a ratification of the eight traits. This book is coherent.

First Trait: No first instance

There had to be an ecology before there was life. There had to come together a great many parts of a complex whole before there could be life. But if many of the parts have to be there in order that there *be* life, then none of the parts can be, strictly speaking, *first.*

Second Trait: Life pops

Have you ever attended a *birth*? Isn't it perfectly remarkable how life pops? Look at plants, at the way seeds pop and green shoots push above the soil. Look at how buds open. These are not simply instances *of;* these are paradigms, models, exemplars for what it means to *pop.*

Third Trait: The Whole influences the nature of the part.

Life is that sort of thing: It is not so much one particular part of a whole, but something about how the whole fits together, namely, in a way that is bounded, self-reproducing, and capable of acting in its own interests. Such 'wholes' influence the nature of their parts as 'organs' with some degree of functional specialization in the interest of the final goal of action on the whole's behalf.

Fourth Trait: Self-reflexivity

Living things act self-reflexively, autonomously, "acting on its own behalf," as Kaufmann would say; or *für sich,* as Hegel would say; or *pour soi,* as Sartre would say.

Fifth Trait: Life is not *reducible* to a re-description of any one of its parts . . .

because it is fundamentally about all of its parts. Life is less about the melody, more about the harmony.

Sixth Trait: Nor is life *predictable* from the perspective of any one of its parts.

Life does not exhibit the symmetry of prediction and reduction characteristic of monological science.

Seventh Trait: Desire drives life's metabolism.

Life cannot maintain itself without *desire* in at least some of its dozen forms—*need, want, appetite, hunger, craving, longing, lust, will, care, love, yearning, hope*. Here's where the biologist's concern comes in: that every form of life must *take in energy* and convert it to its ends. As entropy aims toward death, so desire drives toward life.

Eighth Trait: Coming apart.

Living things die. Death is a part of life. Whole societies collapse. See the record from Gibbon's *Decline and Fall of the Roman Empire* to Jared Diamond's *Collapse.* Bubbles burst. Markets tumble. Some jokes are dead on arrival.

Life conforms to the eight traits of emergent systems. It behaves like what we are coming to understand as typical behavior for emergent systems. On the one hand, careful observation of living things offers lots of hints toward the answer to the question, What is

life? On the other hand, some rudimentary hypotheses concerning an answer to the question of life may be enough to admit several kinds of systems into the category of the living: Not just the familiar plants and animals, but the earth (the Gaia hypothesis first put forth by Lovelock and Margulis), companies (the *living company* as described by Arie de Geus), cities (Cf. Stephen Johnson's *Emergence,* and Jane Jacobs, *The Life and Death of Cities.*), the internet . . .

Candidates abound, but we need some sort of rigorous definition that will keep us from Dr. Dizmo's striking but useless claim that *everything* is alive. If *everything* is alive, then 'life' has no definable, demarcatable meaning. If *everything* is alive, then we rob our language and our conceptual toolbox of a useful distinction, that between the living and the dead. If *everything* is alive, then we lose our ability to differentiate between a plant and a rock, or between Aunt Sally's condition as she speaks her last words and her condition *after* she has "passed away." *Animism,* also known as *panpsychism,* if taken as a religious or philosophical thesis claiming that *everything* (*pan*-) is alive, turns out to be selling more of those extraordinary tomatoes.

What we are looking for, then, is an understanding of life that allows us to see more than the traditional candidates as living, but less than everything. Likewise, when we come to consciousness, we will be looking for a definition that allows us to admit more candidates than just *homo sapiens*, but less than every species.

It is important to be clear about just how high and how wide we are opening and extending certain "umbrella" terms. In pursuing Part Two's strategy of tracing the paths of several emergent levels across Part One's eight traits of emergent systems, we will find that the umbrella terms, 'life' and 'consciousness', have legitimate extensions that cover more candidates than traditional usage might have allowed. But further, in tracking the applications of the eight traits across *several* levels of emergent systems, we are also extending the umbrella of *emergence* further than it has been extended before . . . but once again, not so far as to cover *everything.* Not *all* systems are emergent systems. Monological science is not wrong, just incomplete—finite or limited in *its* legitimate extension.

More on Method: The Need for Natural Philosophy

This sort of reflection on the proper use and extension of umbrella terms is what places this book closer to the tradition of philosophy than to the tradition of science. No laboratory experiments have been performed to gather new evidence for this argument. This discourse belongs in the very old tradition of *natural philosophy. Natural* because the subject is very often nature; but *philosophy* because the work performed here has less to do with hypothesis, experimental method, measurement and evidence, and more to do

with the careful ordering of concepts and their legitimate applications. As observed earlier, when it comes to explaining evolution, we end up learning as much about *explanation* as we learn about evolution. So likewise, when it comes to explaining life or consciousness as emergent systems, we'll end up learning as much about *emergence* as we learn about life or consciousness.

This discourse is philosophical because the questions we are asking and the ways we are answering them require that we go back to first principles and ask: What is it to *explain* anything at all?—This is the question that motivates the discipline known as the philosophy of science.

What is it to *know* anything at all? —This is the question of *epistemology.*

What is it to *be* anything at all?—This is the question of *ontology.*

Monological science presupposes one approach to each of these questions. The science of emergent systems cannot simply assume that same set of presuppositions. A science of emergent systems must be radical in its questioning of monological assumptions. If we look at nature only through the lenses of monological science, we will see singular entities *and only later their relationships.* We will tend to explain changes among those entities in terms of monolinear lines of causality, not in terms of structural shifts or evolutionary dynamics. We will try to know those things and their causes by way of disengaged, objective representation rather than interaction and caring involvement. These differences in *explanatory approach, epistemology* and *ontology* are deep. To be aware of them and self-conscious about them is to make the shift from monological natural science to a stereoscopic vision that also includes a natural philosophy of emergent systems.

These methodological reflections introducing Part Two prepare for the *kind of discourse* being undertaken here. It's not science if monological science sets the paradigm. Nor is it pure philosophy, if by that is meant a logical analysis of pure concepts in their syntactical relationships with one another. Instead it is a kind of natural philosophy that makes a move similar to that which Hegel attempted in his *Phenomenology,* namely, to step back from the question of how our ideas relate to things in themselves and to ask instead: How do the several forms of consciousness, the several sciences that have come on the scene, relate to one another?

We will be asking of several approaches to the question, *What is life?* not so much *whether they got it right,* that is, whether their accounts correctly represent the facts (as if we had some other, separate access to "the facts"). We will be asking, instead, How do these different approaches to the question, *What is life?* reflect monological assumptions about being, knowing and explaining? Or, alternatively, how do these different approaches to the question, *What is life?* reflect an alternative set of assumptions about being, knowing, and what constitutes a legitimate explanation? In this sense, we will be

conducting a *phenomenology of life* very much in the spirit of Hegel's *Phenomenology of Spirit.*

Unlike Descartes' doubt, whose radicality consists in its will to question everything until some bedrock indubitability has been reached, the radicality of this kind of phenomenological approach consists in its avowed circularity, a little like a dog chasing its own tail. We confess that we will never catch the indubitable, that the best we can do is compare one account against another, one philosophy against another, one interpretation, one hermeneutics, against another; one form of life against another, never claiming to have reached unmediated access to either pure subjectivity or pure objectivity.

Let it be said at the outset that there is reason for doubting the adequacy of our current worldview, dominated as it is by the assumptions of monological science. If the Aristotelian worldview erred in the direction of giving biological life too large a role in shaping ideas about being, knowing and explaining, then our current, monological science errs in the opposite direction of giving biological life too small a role in shaping our ontology, epistemology, and philosophy of science. Where the Aristotelian worldview took the acorn's growth into an oak as the model or paradigm for change, and therefore sought for teleological explanations in mechanics, where none are to be found, then our current worldview is conversely too quick to reduce organism back into mechanism. Our current worldview tries to understand living organisms in terms of the mechanics oof their matter. Putting that error right is no small thing, so that is why I keep describing this book as radical.

Just as you'll never put a man on the moon with an Aristotelian science that sees change in terms of the actualization of potentialities, so you'll never understand life with a science that sees nothing but dead matter in motion in space and time. We need a stereoscopic vision that sees *both* the mechanics of efficient causes and effects, *and* the dynamics of emergent systems. Only through a phenomenological comparison of both ways of knowing, together with a careful and disciplined account of their proper domains and applications, will we learn to see ourselves and the natural world around us in all their beauty and orderliness.

How do these philosophical, methodological reflections help us to get on with the question, What is life? Before looking at instances of the eight traits, it would be nice to know just what it is that is bearing those traits. What, precisely, are we talking about? Would that we had a clear and simple definition. But we don't, or so Lovelock discovered when he looked for one. And this is just the point: Neither the epistemology nor the ontology nor the philosophy of science assumed by monological science seem adequate to grasp this phenomenon called life. If you look for life through monological lenses, you will try to find something like *élan vital*—some singular cause behind the

effect called *life.* You will try to find some single part of the organism that is responsible for the life of the whole of the organism. (Spinoza, *The Ethics,* tr. R. H. M. Elwes, Dover, New York, 1955, p. 46; Part I, Definitions: III; Cf. also I.Prop. 6 corr.; and I. Prop. 8 n2.) Life is not the sort of thing that monological science is good at seeing, knowing, or explaining. If all you've got is a hammer, then everything looks like a nail. If all you've got is monological science, then everything will look dead. Life will elude you, as it has eluded most of post-Aristotelian science.

Life's elusiveness—from a monological perspective—may be taken by *us* (who are casting a phenomenological eye back at a comparison between monological science and a science of emergent systems) as a mark of its worthiness to be considered as a candidate for being a genuinely emergent system. If life were *not* so elusive, if monological science had an easy time seeing, knowing and explaining it, then we would not need a science of emergent systems. So rather than forcing life to conform to the assumptions of monological science, or pushing it out of mind altogether if it does not easily conform, we want to seize on evidence for life's elusiveness as support for our needing a science of emergent systems.

Such evidence is not hard to find. To the extent that life, by its very nature, namely, to grow, tends to reach beyond its previous boundaries, then life, by its very nature, is not easy to capture in the boundaries of a definition. As Margulis and Sagan say, "Defying definition—a word that means 'to fix or mark the limits of'—living cells move and expand incessantly." (Margulis and Sagan, *What is Life?,* p. 14.) (Schrödinger wrote of our "obvious inability" to define life. Quoted by Margulis and Sagan, p. 12.) In his authoritative article on the subject in the *Encyclopedia of Britannica*, Dorion Sagan's famous father, Carl, baldly states, "There is no generally accepted definition of life." Why, we may ask? Because we lack a science of emergent systems. Looking back, phenomenologically, from the perspective of a science of emergent systems, it *should* be the case that life looks elusive from a monological perspective.

Let us then take this elusiveness as evidence for the hypothesis that life is a good candidate for a science of emergent systems, and get on with looking for the eight traits of emergent systems in order to see whether they allow the phenomenon of life to come into clearer focus.

First trait: No first instance

Is it the case that we'll never understand life until we know how and where it began? Is life the kind of thing whose origins we must understand before we understand its nature or value? Before buying a painting, you want to know its *provenance:* Who painted it?

How was it produced? But we don't know this much about life. "The origin of the tiniest bacterial cell, the first autopoietic system, is obscure," write Margulis and Sagan. (Ibid., p. 61.) They point out that our understanding of life is like our understanding of fire 50 thousand years ago: "We can maintain and play with it, but we can't yet start it." (Ibid., p. 60.)

Okay, but the question here is not simply whether we know where life comes from, but whether we *need* to know where life comes from in order to understand it. Is the question of origin really all that important? Reading Robert Wright, you wouldn't think so: "How did life begin? Beats me." (Robert Wright, *Non-Zero,* p. 252.) Behind the humor, there may be wisdom in his off-hand dismissal of the question. *It's the wrong question to be asking.* Life isn't the sort of thing about which we *need* to know the first instance before we can know its nature.

In his own way, Aristotle was on to the irrelevance of the question about life's first instance. In addressing the question of just where the line might be drawn that divides the living from the dead, he wrote so many years ago, "Nature proceeds little by little [and it is] impossible to determine the exact line of demarcation." (Aristotle, *The History of Animals,* VIII, 1.) Nor, given the nature of life as being more about the harmony than the melody, can one possibly say with precision exactly *when* the harmony of life began. "The establishment of an autopeoietic system cannot be a gradual process. Either a system is an autopoietic system or it is not." (Francisco Varela, *Principles of Biological Autonomy,* p. 27) According to Varela, there cannot be intermediaries—first instances prior to second instances.

Rather than looking for life's *first instance* as the key to our understanding of what life *is,* we do better to approach it as an emergent system which, at its emergence, is all there all at once. Life is not an effect of some prior, monological cause, whether *élan vital* or a creator God. "Life on Earth is not a created hierarchy but an emergent holarchy arisen from the self-induced synergy of combination, interfacing, and recombination," (Margulis and Sagan, p. 18.). And again, "Life on Earth is a holarchy, a nested fractal network of interdependent beings." (Ibid. p. 71)

From their many and slightly different runs at defining life at the conclusions to each of their many chapters, you get the sense that Margulis and Sagan are in need of a philosophy to supplement their command of the science. They take a bold stab at linguistic analysis:

> Life is distinguished not by its chemical constituents, but by the behavior of its chemicals. The question, 'What is life?' is thus a linguistic trap. To answer according to the rules of grammar, we must imply a noun, a thing. But life on Earth is more like a verb. It repairs, maintains, recreates, and outdoes itself. (Ibid., p. 22.)

Here, from the perspective of natural philosophy, in the course of doing a phenomenology of life that looks at and compares different approaches to the question of life, *we,* who stand back and watch these scientists struggling to describe an emergent system with the language and assumptions of monological science, must sympathize. Not to sound condescending, for we don't deserve to, nonetheless we can appreciate the need for a science of emergent systems that would allow us to say: You don't *need* to find a first instance. Life isn't that sort of thing. It is an emergent system, not the effect of some prior monological cause.

Second Trait: Life pops

Look at the desert in spring. Look at buds bursting into bloom. Look at the miracle of birth, only hold the thought that just because it looks so miraculous to the eye of monological rationality, that does not mean that it must be forever shrouded in mystery. The science of emergent systems, equipped as it is with a list of traits exhibited by any and all emergent systems, can give operational specificity to the claim that life is an emergent system. A science of emergent systems leads us to expect life to pop. That's just the sort of thing that emergent systems do. If it doesn't pop—if it comes into being in a series of incremental steps like the orderly building of a house—first the foundation, then the frame, then the walls, then the roof, and then the doors and windows—then it isn't a living thing at all. It's an artifact, an artificial thing, a product of artifice, not life.

Which is not to say that there cannot be artificial life. But if the product of artifice is to qualify as being truly alive, then that artificial thing will have to be the sort of thing that is capable of popping, not the sort of thing that is built in gradual increments by some other artificer, a programmer with a separate blueprint drawn by a separate architect.

Third Trait: The whole, the living organism, influences the nature of the part, as part of a living organism.

The strongest argument for the holistic nature of life is the long series of arguments for the possibility of artificial life—that it is possible to create life from parts which, by themselves, would be regarded as dead. The artificial life movement has a long history. It was Thomas Hobbes who wrote, in the very first sentences of his *Leviathan:*

> Nature (the art whereby God hath made and governs the world) is by the art of man, as in many other things, so in this also imitated, that it can make an artificial animal. For seeing that life is but a motion of limbs, the beginning whereof is in some principal part within, why may we not say that all automata (engines that move themselves by springs and wheels as doth a watch) have an artificial life? (Thomas Hobbes, *Leviathan,* in *Hobbes Selections,* ed. F. J. E. Woodbridge, Charles Scribner's, New York, 1930, p. 136.)

Hobbes' basic idea holds true for the dozens of contemporary scientists in the dynamic field known as A-life—people like Rodney Brooks, Doyne Farmer, Danny Hillis, John Holland, Chris Langton, and Stephan Wolfram. (This is, of course, a very partial list. For a much fuller list, and a review of the field, see Stephen Levy, *Artificial Life: The Quest for a New Creation,* Pantheon, New York, 1992.) Mitchell Waldrop echoes Hobbes when he writes:

> . . . the connectionist idea shows how the capacity for learning and evolution can emerge even if the nodes, the individual agents, are brainless and dead . . . by putting the power in the connections and not the nodes, it points the way to a very precise theory of what Langton and the A-lifers mean when they say that the essence of life is in the organization and not in the molecules. (M. Mitchell Waldrop, *Complexity: The Emerging Science at the Edge of Order and Chaos,* Simon & Schuster, New York, 1992, p. 292.)

Fourth Trait: Life is a self-reflexive system.

Life is a holarchy, a whole of parts that are related to one another, not only in a linear sequence of cause-and-effect, but in a way that doubles back on itself self-reflexively. Life, in every one of its forms, achieves some degree of closure, whether it is the membrane of a unicellular organism, or the enclosure of a blastula, or the skin of a mammal, or the boundaries of an ecosystem.

Life is self-contained. Every living thing erects a boundary or membrane between self and not-self. For animals, this is the job of the immune system: determining the difference between self and other, inside and outside. For living companies it is the job of employment agreements, patents establishing ownership—what is ours as opposed to not-ours; non-disclosure clauses to keep what is ours *ours;* and *succession* plans that establish not only the question of who *we* are today, but how the company will retain its identity when we are gone and others take our places tomorrow.

Without getting into the technical details of just how autopoiesis works—whether in the form of Varela's formalism in his *Principles of Biological Autonomy,* or in the molecular biology of RNA and DNA—it's perfectly clear at a fairly gross level of analysis that an important feature of autopoeisis is the *auto-,* the *self*-creative, *self*-referential structure of autopoeisis. "Autopoiesis in the physical space is necessary and sufficient to characterize a system as a living system," (F. Varela, *Principles of Biological Autonomy,* Elsevier North Holland, Inc., New York, 1979, p. 41.) and further: "Self-reference is the hinge upon which levels of serial inclusiveness intercross." (*Ibid.,* p. 169.)

As von Neumann articulated the conditions for their possibility, it is not enough for self-reproducing automata simply to have the capacity to *produce* themselves; they must also

produce things that contain maps or descriptions of themselves so that they can *re-produce* themselves yet again. (John von Neumann, *Theory of Self-Reproducing Automata,* completed and edited by Arthur W. Burks, University of Illinois Press, Champaign-Urbana, 1966.) To fully appreciate the importance of closing this loop of self-referentiality, consider Chris Langton's intellectual and emotional epiphany upon coming to understand von Neumann's insight.

> All told, says Langton, it was about two months from the time he first read von Neumann until he finally got what he wanted. One night, he says, the pieces just finally came together. He sat staring at loops that extended their arms, curled those arms around to form new, identical loops, and went on to form still more loops ad infinitum. It looked like the growth of a coral reef. He had created the simplest self-reproducing cellular automaton ever discovered. "I had this incredible—volcano of emotion," he says. "This *is* possible. It *does* work. This *is* true. Evolution made sense now. This wasn't an external program that just manipulated a table. This had closure on itself, so that the organism *was* the program. It was complete. And now all these things that I'd been thinking of that might be the case if I could do this—well, they were all possible, too. It was like a landslide of possibilities. The dominoes fell, and just keep falling and falling and falling." (As quoted in Waldrop, *Complexity, op. cit.,* pp. 221f.)

And because life is just *one* of the levels of emergent systems, those dominoes, with their emotional concomitants, will keep falling and falling and falling.

Fifth Trait: Life is unpredictable from the features of its precursors

There is nothing in the molecular physics of hydrogen, oxygen, carbon or nitrogen that would allow you to predict life. Nor is there anything in the physical chemistry of more complex molecules, the polymers and amino acids that come closer to the complexity of DNA.

Sixth Trait: Life is irreducible to description in terms of the features of its components.

Some assemblages gain their properties from the properties of their components. For a movie studio to have a successful season, a significant number of its releases must be successful releases. If all of the novels in a library are written in French, then it's safe to say that it's a library of French literature. Inferences from the properties of the whole to the properties of the parts are not always justified, even where it's not an issue of emergence. Logicians have a name for "the fallacy of composition": Just because an engine is heavy, it does not follow that all of its parts are heavy. Without calling it an

emergent property, heaviness can be the result of a mere summation of the weights of very many un-heavy parts.

Granting such a range of examples of whole-part relationships, still there is something else being said when A-lifers make it a point of their creed that,

> Instead of being an effort to understand life by *analysis*—dissecting living communities into species, organisms, organs, tissues, cells, organelles, membranes, and finally molecules—artificial life is an effort to understand life by *synthesis:* putting simple pieces together to generate lifelike behavior in man-made systems. Its credo is that life is not a property of matter per se, but the organization of that matter. Its operating principle is that the traits of life must be traits of dynamical form, independent of the details of a particular carbon-based chemistry that happened to arise here on Earth four billion years ago. (Waldrop, *op. cit.,* p. 277.)

Because life is an emergent property of the whole organism, any attempt to reduce it to the presence or absence of some carboniferous component will be regarded by the A-lifers as a form of un-scientific prejudice in favor of carbon, something akin to saying that quarterbacks on football teams have to be white. We will meet such accusations of prejudice again when we come to the question of whether consciousness can or cannot be resident in silicon. Must consciousnesses reside only in meat?

Seventh Trait: All life exhibits desire.

Every living thing, from the uni-cellular protist swimming upstream in a glucose gradient to a living company, must take in energy in some form or another. And it must metabolize that energy in a way that nourishes and maintains its integrity. The wonder of it is the incredible range of sources of energy, and the variety of metabolisms and forms that then qualify as living.

Eighth Trait: Coming apart

If a living thing ceases to take in and metabolize energy, it dies, simple as that. But there are other causes of death as well: disease, gunshots, asphyxiation. That death marks the end of life—not its purpose, not its telos, but its cessation—says a lot about the sort of thing that life is: life is such that that it can cease to be, not so much by the addition or subtraction of some entity (though a bullet can be a proximate cause), but rather by a change in the way that all its parts are related to one another. All the parts that came together in a way that worked together . . . suddenly cease to work together.

Let's review this first pass across the columns of the chart. We've asked whether the eight traits of emergent systems shed light on the question of the nature and origin of life. Does it help us to understand life by seeing that it's not the kind of thing about which it helps to know its first instance? Yes. Does life pop? Yes. In the case of living things, does the organization of the whole influence the nature of the part? Yes. Is it a feature of living things that they achieve self-containment and integrity by virtue of some form of self-reflexivity? Yes. Is life unpredictable from the features of its precursor components? Yes. Is life reducible to the features of its components? No. Does life exhibit desire? Emphatically and universally, *yes.*. Do living things die? Is life subject to death? Often sadly, yes.

The eight traits all apply. They gain confirmation in this application to the question, What is life? By applying the eight traits and working out their application to life, we can conclude that we know what we mean when we say that life is an emergent phenomenon. We have given some operational specificity to this attribution of emergence. We are not just waving a wand of inarticulable miraculousness. This is no small thing.

II. Evolution of species

Before diving into the application of each of the eight traits to the question of evolution, it's worth observing that this entire section is, in effect, an interpretation of evolutionary theory that follows the red thread of emergence through a range of issues in evolutionary theory. The closing decades of the 20th century witnessed an explosion of interest in evolutionary theory—and an explosion of controversy. The controversies over evolution extend far beyond the crude conflict between creationists and evolutionists. Even among the sons and daughters of Darwin there rage intense sibling rivalries.

Is the course of evolution best explained by reference to "the selfish gene"—the genotypic paradigm favored by Dawkins and others—or by reference to natural selection operating on phenotypes—the phenotypic paradigm—or by adaptive dynamics operating on whole species and their environments—the paradigm of population genetics? Or consider another hotly fought battle: does evolution take place in small increments, more or less continuously, as Darwin argued, or does evolution unfold in fits and starts as described by Eldredge and Gould in their theory of *punctuated equilibrium?* Or, third, can evolution be said to have a direction toward higher levels of complexity? Or is evolution, as Stephen Jay Gould argues, an utterly random walk, a completely non-teleological play of chance and necessity such that the shape of the human soul is just the result of dumb luck signifying nothing?

These are big questions. A science of emergent systems, it turns out, can contribute to their answers, not by the introduction of new evidence, but by virtue of the application of the discipline of natural philosophy to the sorting and comparing of different truth claims. We can cut the Gordion knots of some of these long-fought disputes by showing how their adversaries are making different presuppositions about knowing, being and explaining. We can reframe these disputes by looking at them through not just one, but both lenses of a stereoscopic vision that combines monological science with a science of emergent systems. We can frame these disputes as antinomies which, like the antinomies resolved by Kant in his *Critique of Pure Reason,* can be resolved by a shift of perspective comparable to what Kant called his "Copernican revolution." (Immanuel Kant, *Critique of Pure Reason, tr.* Norman Kemp Smith, St. Martin's Press, New York, 1961, p. 22 (Bxvi-xvii) for his Copernican revolution; pp. 384ff for his resolution of the antinomies. For some sense of the momentousness of Kant's achievement in resolving those antinomies by reframing them through the lenses of his Copernican revolution, see the following commentaries: T.D. Weldon, *Kant's Critique of Pure Reason,* The Clarendon Press, Oxford, 1958, pp. 78f., 89-93, 202-4, 208-10; H. W. Cassirer, *Kant's First Critique,* Macmillan, New York, 1954, pp. 267f., 271ff., 302-308; Norman Kemp Smith,

A Commentary to Kant's 'Critique of Pure Reason', Humanities Press, New York, 1962, pp. 18f, 22-25, 478ff., 519f.; Nathan Rotenstreich, *Experience and its Systematization: Studies in Kant,* Martinus Nijhoff, The Hague, 1965, pp. 68ff., 122, 157f., 162-168; A. C. Ewing, *A Short Commentary on Kant's Critique of Pure Reason,* The University of Chicago Press, 1938, pp. 16, 208-27; Herman-J. de Vleeschauwer, *The Development of Kantian Thought, tr.* A. R. C. Duncan, Thomas Nelson & Sons Ltd., London, 1962, pp. 49-61.) Just as Kant was able to resolve some of the antinomies that seemed to arrest the advance of reason in his day, so a science of emergent systems can resolve the antinomies of evolutionary theory.

No first instance of a new species

From a monological perspective, the story of the evolution of a new species begins with a random mutation in a single genotype that grows into a phenotype that just happens to have a certain advantage over its peers when it comes to survival and reproduction. Imagine a random mutation that allowed an individual to sense impending earthquakes, asks Stuart Kauffman. Such an individual would be more likely to survive and reproduce than peers that would perish in earthquakes. That survivor's genes would enjoy a very slight advantage in the casino of subsequent survival such that, given enough time and enough earthquakes, that first mutant's progeny, thousands of generations hence, would eventually outnumber the progeny of their earthquake-insensitive peers. All because, once upon a time, there was a single point mutation that turned out to have an advantage in the earthquake-prone environment we homo sapiens inhabit.

Well, not *all* because. For, after all, those who lacked the earthquake sensing gene might have been clever enough to come up with seismographs. Architects might have improved the safety of buildings built near known faults. The gene for sensing earthquakes might have found itself linked, by virtue of its location on a given chromosome, to other genes that turned out to be evolutionarily disadvantageous for survival . . . But the monological perspective will be inclined to screen out all those ifs, ands and buts in order to focus on a single cause of survival: the original single point mutation. The monological perspective will try to find that first instance, then argue for its propagation through just a few progeny who begat more progeny who begat still more progeny down through the countless generations charted through the branching diagrams of genealogy and cladistics.

Now consider a very different type of explanation for variations of populations in successive generations, one that does not rely on the first instance of a single-point mutation, but is based instead on almost atmospheric shifts in the environment that are cited to explain the favoring of one tendency or another in a population that contains a

normalized distribution of a given property like size or color. An oft-sited example of this sort is the differential survivability of a species of moth in which, without any reliance on mutation, there is a distribution of pigmentation from black to white. (This oft-cited example has been subjected to criticism in recent years, but for the purposes of the present argument, the recent disputes over the precise mechanisms involved in this example don't cut against the force of the argument being made here.) When a great deal of coal was being burned in England, those moths that tended toward black were better camouflaged against their sooty surroundings and therefore enjoyed an advantage over their lesser pigmented peers. Later when coal burning was restricted, the population of white pigmented moths bounced back.

This second line of reasoning relies more on the changing environment's ability to favor one variant or another in a population that contains a normalized distribution of phenotypic variety without reliance on genotypic mutation. According to this second type of explanation, it makes sense to say that a sooty, coal-burning environment "wants" black moths, or that a coal-free energy policy "affords" white moths. In the context of this second type of explanation, it's not necessary to locate the first black moth or the first white moth in order to explain a given distribution of black or white moths. Finding the first black moth to explain genetic drift toward pigmentation is no more necessary than finding the first raindrop to explain the onset of a shower. Just as atmospheric conditions became such as to encourage the almost simultaneous condensation of many droplets, so a sooty, coal-burning London gave simultaneous protection to many pigmented moths.

Let us now park this argument, while granting that both types of explanation make a certain amount of sense. Single point mutations do happen, and some of them (though, to the extent that they are random, only very few) convey reproductive advantages. In cases where a new species is the result of such mutation, it makes sense to look for first instances, single point mutations that, had they not occurred, a given species would not exist. Equally, selective pressures on whole populations that are exerted by changes in the fitness landscape also happen, and in those cases, the search for first instances seems misplaced.

Apart from any reference to genetic mutation, the play of adaptation between changing populations and changing environments can, over time, produce new species—that is, populations that have accrued sufficient differences that they do not or cannot mate to produce common offspring. Rather than insist on a knock-down-drag-out argument to prove the superiority of the first or the second type of account of speciation, let us grant that both types of speciation occur. And then let us see what difference it would make to a broader understanding of evolutionary theory by seeing how both types of explanation play out on other planes of some familiar evolutionary controversies.

Second trait: Evolution pops

Back in Part One, while following the column of the third trait down through the several layers of emergence, we already came across Eldredge and Gould's theory of punctuated equilibrium. Recall how Galton's polyhedron doesn't roll smoothly across the surface of time, but tips in fits and starts from one face to another, from one ecosystem to another, from one equilibrium to another. This point alone is sufficient to ratify evolution's satisfaction of the second trait of emergent systems. Evolution pops in examples like the so-called "Cambrian explosion," when many different species emerged all at once in the blink of a paleontological eye.

Third Trait: The whole influences the nature of the part, or, the niche The fitness landscape, influences speciation as much as the gene.

This trait helps to make sense of the second type of explanation. The play of random sporting at the level of the genotype can produce a first instance that propagates successfully over many generations. Genetic drift as a function of phenotypic survival is as much about the nature of the niche as it is about the genotype. Without relying on a first instance of variation, a given environment—sooty London—will favor all black moths in general, not just a first black moth. The whole system—species plus environment—favors one or another of its parts.

Fourth Trait: All emergent systems, including species, are self-reflexive.

If a new species emerges through the course of evolution, it must, as von Neumann argued, contain a map of its own territory so that it can reproduce itself true to type. The fact that every single cell in the human body contains the same genetic blueprint of the whole organism stands as one of the most remarkable examples of self-referentiality on a long list that includes everything from feedback to auto-catalytic closure. No wonder Chris Langton had to work so hard to make his programs obey von Neumann's requirement for life. It is no small thing to make a thing that can make the very same thing over and over again. The thing you make will have to reflect back upon itself, picture itself, map itself with sufficient accuracy that it can produce another of itself. This is no small thing.

The process of evolution exhibits this fourth trait of emergent systems in the form of lineages of organisms that reproduce true to type, but with variations. If reproduction were *perfectly* true to type, if there were *no* variation, then there would be no evolution of species, just monotonous repetition. While perfect reproduction is a feature of computer

algorithms, such that variation is noise rather than signal, variation is a feature rather than a bug in evolution.

Fifth Trait: Emergent systems are unpredictable from their precursors.

New species are not predictable from their precursors where random mutation is involved. If evolution were entirely explicable by virtue of what the niche "wants," then evolution might be predictable. It would simply be a matter of doing the right kind of "market research" to discover what the environment will afford. With enough understanding of differential advantages, you could predict that black moths would prosper in a sooty environment, just as you can predict that wireless telephony will prevail over wireline in an environment like the 12,000 islands of Indonesia. To the extent that environmental affordances dictate the superiority or inferiority of one or another variation from the norm, evolution might well settle down to some calculable optimum.

But then along comes random mutation as the joker in this otherwise calculable deck. Chance has its hour in evolution. It is not the case that a calculation of environmental affordances will dictate some optimal solution to winning the game of survival in that environment, because mutation will spontaneously and unpredictably alter the list of players. Mutation will eventually disturb any temporary equilibrium, any happy eco-system where a group of species has co-evolved into a so-called "climax ecology" of mutually symbiotic relationships (a thory that is now discredited).

From the larger, longer perspective of evolutionary "progress" (quotations are required because the concept of directionality or progress in evolution is still in question), the joker of random mutation is a good thing, not a bad thing, for without it, the imperious demands of environmental determinism might become tyrannical. We need the upstart spring of random mutation to counter the success of success. Otherwise we might settle into some equilibrium or other, a million year Reich that would halt the proliferation of experiments, the play of freedom, and shut down the laboratory where *even better* solutions might be brewed.

From the larger, longer perspective of evolutionary "progress" (and we will eventually get to the question of whether this word is appropriate or not), the controversy over whether mutation is utterly random, or whether, as Rene Thom put it, "the fish already 'knew', before they became amphibious, that a life on land would be possible for them, and what new organs they would need," (Rene Thom, *Structural Stability and Morphogenesis, op. cit.,* p. 294.) takes on a new light. From the point of view of the individual fish, it might seem nicer to imagine that knowing and choosing ones future is better than being jostled into it by a series of random mutations and ruthless selections.

So it might seem. But from the larger, longer perspective of evolutionary "progress," it really doesn't make any difference whether a given fish knew what it was doing when it traded its gills for proto-lungs, or its fins for proto-feet. From the larger, longer perspective of evolutionary "progress," all that makes a difference is that we got enough variation so that some fish somewhere got a chance to try on proto-feet and/or proto-lungs. It's the *opportunity for variation* that is important, not any particular variety, which is why, many rungs higher on the evolutionary ladder, an Oliver Wendell Holmes will eventually say, "I don't like what you are saying, but I will defend to the death your right to say it." The protection of civil liberties and the freedom of speech is like the gift of random mutation: an insurance policy against locking into any one equilibrium with a totalitarian grasp on utopian optimality.

Sixth Trait: Emergent systems are irreducible to their components.

In the context of evolutionary theory, the issue of reductionism can cut a number of different ways: Some theorists want to reduce evolution to the play of chance and necessity; others want to reduce evolution to the script of a divine creator. Creationism might not sound like reductionism to believers, who see in the divine plan a "higher" rationale than the meaningless play of chance and necessity. But from the point of view of the syntax of explanation, creationism is every bit as reductionistic as a nihilistic spin on Darwin. In both cases, there is an equal if opposite confidence: evolution is (a) *nothing but* the play of meaningless chance and harsh necessity, or, equally reductionistic in syntax, evolution is (b) *nothing but* the sequential manifestation of God's eternal, read-only design.

The read on evolution taken here, through the stereoscopic lenses of both monological *and* emergent sciences, is that evolution cannot be reduced to either of these equally elegant but equally inadequate explanations. Evolution is an emergent explanandum that *cannot* be reduced to either of the usual explananda. Which is not to say that it cannot be explained. But the explanation will require *both* lenses: The monological lens that allows us to see the play of spontaneous (or, through the other lens, *meaningless)* single-point mutations; *and* the lens of a science of emergent systems that allows us to see the higher purpose served by seemingly meaningless bankruptcies, deaths and mutations, namely, increasing the likelihood of new experiments, new varieties, so that we can get improvements on the current models.

The question repeatedly and parenthetically begged by the account of evolution thus far is that of the criteria by which new varieties could be judged as "improvements" that are "better" than current models—in short, the question of "progress" in evolution. The massive extinction of the dinosaurs may have cleared the way for smaller, tastier

mammals like us, but it did not mark an "improvement" of the dinosaurs. We did not get "better" dinosaurs. And by some criteria—sheer size, speed, strength—many of the dinosaurs were superior to us. So by what standard, and by what justification of that standard, can evolution be said to have a direction, or make progress? Surely we are more intelligent than the dinosaurs. But who says that intelligence is a "higher" or "better" standard than sheer strength?

These are not simple questions. They drive us on toward a consideration of the seventh trait of emergent systems and the role of desire in the course of evolution. But before going there, let's add to our account of evolution. Let's look at and compare the ways different approaches to evolution play out the prejudices of monological rationality. Let's see how those pre-judgments obscure the emergence of complexity and "higher" order beneath reductionist accounts that see evolution as "nothing but" the play of chance and necessity.

The perceptive reader will have noticed, and wondered, about the repeated quotation remarks around "progress." Why the qualifying quotes? Isn't it Darwinian dogma that variation is completely random? Jacques Monod wrote a famous book called *Chance and Necessity.* Theorists like Dan Dennett (*Darwin's Dangerous Idea)* and Stephen Jay Gould (in many books and essays) insist that evolution is not directional. It doesn't know where it's going. It is non-teleological, and all attempts to claim that evolution makes *progress* are, according to them, just so much wishful thinking. We human beings might like to think that we are higher on some evolutionary ladder than the banana slug . . . but evolution did not bother to create *better* dinosaurs, and we, too, could be headed for extinction under the non-teleological hands of random chance and harsh necessity.

Another variant on what might be called the *brutal* reading of evolution is Richard Dawkins' famous phrase, in his book so named: *The Selfish Gene.* There's nothing moral or altruistic written into the mechanism of evolution. Living organisms (or phenotypes) are just the gene's (or genotype's) way of making sure that its code, its message, will survive and prosper. Those little genes get a ride on the bus—the phenotype—they help to build and steer. But the point of calling the gene "selfish" is to make clear that the genes will catch the next bus, and the next, and the next, and from the selfish gene's point of view, continuing the trip down through successive generations is far more important than the well-being of any particular bus that happens to carry it on its way.

Yet another aspect of this brutal read of evolution is captured in such phrases as "survival of the fittest," and "nature red in tooth and claw." Herbert Spencer is the name generally associated with a view of evolutionary theory as describing—and ultimately justifying—a dog eat dog world. To those bleeding hearts who want to advocate altruism in human affairs and kindness toward our animal bretheren, the brutal read of evolution to which Spencer, Dawkins, Dennett, Monod, and Gould all contribute (in importantly different

ways to be sure) says: Wake up and smell the blood. Evolution isn't set up to favor liberal ideology or Christian charity. The world is a sometimes cruel and generally heartless place headed nowhere in particular. Get used to it.

Before enlisting support from other theorists to counter the brutalists with a kinder, gentler theory of evolution, it's worth acknowledging the salutary influence of brutalism. Both creationism and anthropocentrism cry out for critique. God did not create man and all the rest of the species in six days. Nor is humanity the reason why life has been at work so hard evolving for 4.5 billion years. The brutalists stand in a noble tradition from Galileo and Copernicus through Marx, Darwin, Nietzsche and Freud. It's altogether worth deflating the pomposity of those who see man (in God's image by the way) as sitting smugly at the center of the universe; as better than the beasts; as specially and uniquely endowed with a mind that transcends and rules matter; as intrinsically moral, as essentially nice rather than naughty (that is to say, sexual). There's a prissy view of human nature, often but not only rooted in religion, that the great deflators have done well to criticize. The brutalists borrow strength from this noble tradition, and bid fair to add to it.

But have babies been flushed with the bathwater of creationism and pompous anthropocentrism? Very possibly. There may be a third way, a third read of evolution between the brutal and the pompous that a science of emergent systems can reveal. Fortunately we are not alone. We can find friends on this third path as diverse as Stuart Kauffman (author of *At Home in the Universe),* Gregory Bateson, Robert Wright (author of *Non-Zero),* Rene Thom (*Structural Stability and Morphogenesis),* Peter Corning (*The Synergism Hypothesis),* and a range of other writers who add stepping stones to this third path. They neither sacrifice the cold-eyed rigor of the brutalists, nor yield to the wishful thinking of the pompous and prissy.

Both brutalism and creationism are too simple. The creationists preach faith in a single creator God and a simplistic sense of divine purpose and destiny. "Don't worry if things look bad in the moment. It will all turn out for the best. So was it meant to be." This kind of teleological thinking justifies all meandering missteps by a blind faith in a divine telos that sweeps everything toward it like water toward a drain. Whether Augustine's City of God or Teilhard de Chardin's Omega Point, such teloi reduce complexity by sucking out all contingency.

For the brutalists, complexity disappears *down* instead of *up.* Life is a meaningless dance which, in the last analysis, signifies nothing. (Note the qualifier: "as read by the cynics." Marx, Nietzsche and Freud are richer, deeper, and subtler than the cynics make them out to be. For a demonstration of the range of readings possible for each of them, see Paul Ricouer's distinction between "the hermeneutics of suspicion" and "the hermeneutics of belief" in his *Freud and Philosophy: An Essay on Interpretation,* trans. Denis Savage,

New Haven, Yale University Press, 1970.) And if you read Darwin through those same reductionist lenses, evolution comes down to a play of blind chance and harsh necessity signifying nothing (Monod), or so many manipulations of selfish genes (Dawkins), or one damn thing after another with no sense of progress (Gould).

The brutalists ape the creationists, even as they simplify *down* rather than *up.* This was Hegel's insightful point in his passage on Enlightenment and Superstition in the *Phenomenology.* Between both of these paths that seek explanation by reductionistic simplification, we need to find a third way that preserves—and even explains—*complexity*.

According to Stuart Kauffman, nothing in the neo-Darwinian synthesis or in physics or chemistry explains complexity. (Stuart Kaufmann, *Investigations,* Oxford University Press, New York, 2000.) Certainly not entropy, which corrodes complex order into the less informed simplicity of heat death. So how does complexity happen? Not by design. We've rejected that path. Not by brutalist reduction. That way lies nothing but rearrangements of the same old simple stuff, whether the means of production in dialectical materialism, or will to power in an endless and non-progressive eternal recurrence, (See Friedrich Nietzsche, *Will to Power,* tr. Walter Kauffman, New York, Vintage Books, 1967.) or libido in zero-sum sublimations that lead civilization to discontent. (See Sigmund Freud, *Civilization and its Discontents,* tr. Strachey, New York, W. W. Norton & Co., 1961.)

Now that the contrast between the creationists and the brutalists has been drawn, it's easier to see what "progress" does and does not mean, and why those little quotation marks are so important. What *does* "progress" mean? That variation is for the most part random, but not utterly blind. Chance is real. Contra Einstein, God *does* play dice. But the dice are loaded. There *is* a small but, over many generations, very influential margin of directionality in evolution.

Before giving too much succor to the advocates of intelligent design, what does "progress" *not* mean? That directionality is equivalent to teleology. While there may be a tendency toward greater complexity and more order, this directionality is best seen as away from disorder and simplicity, but not toward any particular, predetermined telos.

Now that we know what "progress" does and does not mean, what arguments can be advanced to show that the claim for directionality toward greater complexity is true or, at the very least, not crazy. Consider four sources of support for the third path we're seeking.

1. We saw in Part One how French mathematical biologist Rene Thom challenged the idea that variation was utterly random at the level of single-point mutations of genes. Thom argues that, "Metazoa have, located in the

kinetic configuration of the metabolism of their gametocytes, a model of their actual conditions of existence."

Thom said that the fish contains within itself a kind of map, not only of its existing environment—water—but also, by virtue of the fact that the fish is part of a universe, an "ambient metabolism," containing more than water, the fish can model in its "research laboratory" another possible environment—land—and on the basis of that modeling capacity, can mutate in a way that is "not the more-or-less hazardous game of a mad molecular combination," but a more directed, more Lamarckian "aiming" toward that new global order.

2. Switching now from the level of mutation in the genotype, consider the role of "genetic drift" in the evolution of species. Here the source of variation may have nothing to do with (almost) random genotypic mutations, but with the distribution of certain traits among phenotypes. When coal burning was more prevalent in England than it is today, the black-winged members of the species were better camouflaged. Here the diachronic evolution of a species is influenced by synchronic relationships with its surroundings. The ratio of black moths to white moths "co-evolves" with the coal burning habits of humans. This kind of co-evolution produces ecologies in which the wealth of mutually beneficial symbiotic relationships is such that observers will be moved to believe that things were designed "for the best." But a closer attention to the role of genetic drift will show that neither a designer nor a telos need be invoked to explain this co-evolution of mutually functional relationships. An understanding of synergy is sufficient. (Cf. Peter Corning, *The Synergism Hypothesis,* New York, McGraw-Hill, 1983.)

3. In his book, *Non-Zero,* science writer Robert Wright makes a persuasive case for the directionality of evolution toward greater complexity and higher order. His palette extends beyond biological species to the successive stages of human history, from primitive hunter gatherers, through the invention of agriculture and from ancient city-states to the invention of empire. At each stage in the long march of human history, Wright notes a kind of ratchet effect by which zero-sum conflict of the kind the brutalists see is succeeded by a non-zero sum solution. Rather than zero-sum rearrangements of the same old stuff—whether wheat, wealth, power, or libido—Wright shows how the evidence of history shows an unsteady march toward more complex systems, from tribes to cities to states to empires to . . . who knows what? In his final chapters Wright allows himself to speculate on the possibility that we human beings, together with our invented technologies like the Internet, might be evolving toward a more peaceful, more self-conscious realization of a fuller human potential. (Robert

Wright, *Non-Zero: The Logic of Human Destiny,* Pantheon Books, New York, 2000, esp. 301ff.)

4. Stuart Kauffman extends his arguments about the evolution of complexity in the biosphere to what he calls the "econosphere" where he makes the incontrovertible observation that, "[A]mong us mere humans, the diversity of ways of making a living has increased dramatically over the past 3 million years, the past hundred thousand years, and even over the past thousand years. If you wanted a rabbit for dinner thirty thousand years ago, you bloody well went out and caught a rabbit. Now most of us can go buy a rabbit dinner. Something again has happened. At the level of species and ways of making a living in the 'econosphere,' the actual has expanded into a persistent adjacent possible." (Kauffman, *Investigations, op. cit.,* p. 143.) His point is simply that complexity begets greater complexity. The invention of the automobile begets gas stations, highways, suburbs, and shopping malls, none of which would have evolved had there been no cars.

Now consider some conclusions that can be drawn from these four sources of evidence for a third path that interprets "progress" as neither guaranteed (as by the creationists) nor the product of silly wishful thinking (as by the brutalists):

1. It *is* possible to get more out of less. Indeed, this notion of getting more out of less is the very definition of emergence according to Stuart Kauffman's colleague at the Santa Fe Institute, John Holland. (More precisely, Holland defines emergence as, "much coming from little." Cf. Holland, *Emergence,* Cambridge, Mass., Perseus Books, 1998, p. 1.) The concept of emergence is central to the thesis of this book. After years of neglect, since the early and somewhat wooly work of figures like C. Lloyd Morgan, Samuel Alexander and Bergson, emergence is now attracting considerable and much more rigorous attention. (For a popular treatment see Stephen Johnson, *Emergence: The Connected Lives of Ants, Brains, Cities, and Software,* Scribner, New York, 2001. For an extremely ambitious but less accessible treatment of the same topic, see Stephen Wolfram, *A New Kind of Science,* self-published, 2002, over 1,200 pages.) Don't be fooled by the brutalists and reductionists who take the law of constant conservation of mass and energy from the realm of physics and chemistry and mistakenly extend it into the realms of the biological and information sciences. Contingency happens. But emergence happens too.

2. There are no guarantees. Contrary to the creationists, happy endings are not foreordained. The best of intentions can yield unintended consequences. For any single actor, tribe, species or company, there is always the distinct possibility of tragedy, defeat, extinction or bankruptcy. Indeed if there are no failures, there are

no successes. The logic of Darwinian selection is based on profligate variation and ruthless selection.

3. Achieving higher order is often a matter of attending to synchronic rather than diachronic relationships. Wright's ratchet toward non-zero-sum solutions depends on co-evolutionary dynamics more than the diachronic evolution of any single tribe or species. If we're going to learn to get along *better,* then we have to learn to get along better *together.* As the philosopher Leibniz saw long ago, *real* possibility, not just *logical* possibility, is a function of *compossibility* with the rest of the real world.

4. Sometimes, not always, the more orderly system, or at least the conditions for its possibility, can be foreseen. Just as "the fish already 'knew', before they became amphibious, that a life on land would be possible for them, and what new organs they would need;" just as increasingly large associations of human beings required the yearning for a greater peace; so it is possible to transcend past and present to imagine better futures. (Cf. James Ogilvy, *Creating Better Futures: Scenario Planning as a Tool for a Better Tomorrow,* Oxford University Press, New York, 2002.)

This last conclusion carries the argument on to the seventh trait of emergent systems, the role of desire in determining what it would mean to call a future "better."

The Seventh Trait: Desire, and the reciprocity of purposes and desires in the course of evolution: the question of teleology and teleonomy

Evolution has evolved. There was a time when living things did not die, viz. fungi and bacteria. "Aging and death, in which living cells disintegrate with predictable timing, first evolved in sexual protoctists. 'Programmed' death as the final stop of a lifelong metabolism was absent at the origin of life—and for a very long time afterward. Unlike us, bacteria are immortal; they will live until external conditions prevent autopoeisis." (Margulis and Sagan, *What is Life, op. cit.,* p. 113.)

There was a time prior to the evolution of sex when cells reproduced by mitosis rather than meiosis. "Once upon a time, we think, eating and mating were the same. Terminal microbial indigestion may seem rather unromantic as the source of the human sex drive. But Cleveland's picture of hungry, serendipitously mating hypermastigotes presents a mix of comedy and terror appropriate for the origin of sex." (*Ibid.,* p. 114.)

How variation is achieved—whether by successive generations of clones that differ by virtue of adaptation to different environments, or by sexual reproduction that shuffles

genes with each generation—makes a big difference to the pace of experimentation. Pondering the evolution of evolution leads Margulis and Sagan to what is perhaps the most profound of their many definitions of seemingly indefinable life:

> Life is the representation, the "presencing" of past chemistries, a past environment of the early Earth that, because of life, remains on the modern Earth. It is the watery, membrane-bound encapsulation of spacetime. Death is part of life because even dying matter, once it reproduces, rescues complex chemical systems and budding dissipative structures from thermodynamic equilibrium. Life is a nexus of increasing sensitivity and complexity in a universe of parent matter that seems stupid and unfeeling in comparison. Life must maintain itself against the universal tendency of heat to dissipate with time. This thermodynamic view explains, in a way, the determination, the purposefulness of life—for billions of years it has been stuck in a pattern which, even if it wanted to, it can't get out of, of upping the stakes as it goes. For life itself is—are—these patterns of chemical conservation in a universe tending toward heat loss and disintegration. Preserving the past, making a difference between past and present, life binds time, expanding complexity and creating new problems for itself. (*Ibid.,* p. 67.)

Here, not unlike Kant in his Copernican revolution, by which he solved the antinomies of space and time, Margulis and Sagan, too, are so bold as to broach the metaphysics of space and time. Where Kant followed Copernicus, Margulis and Sagan follow Kant, who wrote:

> Hitherto it has been assumed that all our knowledge must conform to objects. But all attempts to extend our knowledge of objects by establishing something in regard to them *a priori,* by means of concepts, have, on this assumption, ended in failure. We must therefore make trial whether we may not have more success in the tasks of metaphysics, if we suppose that objects must conform to our knowledge. This would agree better with what is desired, namely, that it should be possible to have knowledge of objects *a priori,* determining something in regard to them prior to their being given. We should then be proceeding precisely on the lines of Copernicus' primary hypothesis. Failing of satisfactory progress in explaining the movements of the heavenly bodies on the supposition that they all revolved round the spectator, he tried whether he might not have better success if he made the spectator to revolve and the stars to remain at rest. A similar experiment can be tried in metaphysics, as regards the *intuition* of objects. If intuition must conform to the constitution of the objects, I do not see how we could know anything of the latter *a priori;* but if the object (as object of the senses) must conform to the constitution of our faculty of intuition, I have no difficulty in conceiving such a possibility. (Kant, *Critique of Pure Reason, op. cit.,* p. 22 (B xvi-xvii); Preface to the Second Edition.)

Just as Kant was able to cut through the antinomies of space and time—scholastic questions like whether "The world has a beginning [no first instance] in time, and is also limited as regards space," or whether "The world has no beginning and no limits in space

[but] is infinite as regards both time and space" (*Ibid.,* p. 396 (A 426; B 454).)—so Margulis and Sagan come close to resolving some of the controversies over the nature and origin of life by entertaining their own Kantian/Copernican hypothesis, namely, that life is not something fixed at the center of a spacetime that moves and revolves around it, but rather that life moves. *"Preserving the past, making a difference between past and present, life binds time."* (Margulis and Sagan, *op. cit.,* p. 67.) And in so doing, life is not just an anomaly that is struggling, ultimately in vain, against the second law of thermodynamics. Heat death is not life's destiny. Instead *Eros* is, as Freud maintained, an equal adversary to *Thanatos*, death. (Freud, the very last paragraph of *Civilization and its Discontents, op. cit.)* The principle of desire, the seventh trait of emergent systems, is equal to, and in reciprocity with, the second law of thermodynamics: entropy. Life is not some temporary anomaly inside of time. "Life binds time."

"Life binds time." What can this mean but that care, *Sorge,* binds past to future. Life wants *not* to die, even as entropy drives toward death. *Dasein* cares. To *be, to live,* is to resist death.

Life is hot. Death is cold.

Care creates the story of life. Without care, there is no story, just meaningless events that bear no relationship to one another. Where Kant comes up short is in seeing the subject as uncaring, as nothing more than a "transcendental unity of apperception," to quote the not so catchy language of his *Critique of Pure Reason.* Hegel put care back into consciousness. This allowed him to take a different cut on the antinomies.

Once the future replaces objectivity as the horizon of validation, then care and hope become constitutive of reality. Recall how Richard Rorty parsed the message of pragmatism as, "a willingness to refer all questions of ultimate justification to the future, to the substance of things hoped for. If there is anything distinctive about pragmatism it is that it substitutes the notion of a better human future for the notions of 'reality', 'reason' and 'nature'." To the extent that the future that finally comes to pass was actually hoped for and planned, that planning amounts to the kind of "downward causality" that Karl Popper invoked in his refutation of reductionism.

Our cares about the future inform our present, thus changing the future from what it might have been, and thereby changing the present in its potentiality. This caring is thus constitutive of the *meaning* of the present—what it *is* in Popper's World Three—the world of symbols and symbolic constructs like culture and language and song.

To the extent that our care is informed by a sense of beauty—our aesthetic sense of how things could and should fit together—this care becomes the vehicle of teleological directionality. In human time, time that is bound by life and care, things *are* likely to turn

out for the best precisely because that's how we *want* them to be and therefore work to make them to be.

There is an ontological reciprocity or autocatalytic closure between a locus of care on the one hand and a telos on the other. Consider a telos as an object of desire. Is it desired because it is a telos? Or is it a telos because it is desired? Neither by itself, but both in a dialectical structure of ontological reciprocity: you don't get one without the other, the telos or the desiring subject.

Heidegger was onto something when he declared that care, *Sorge,* was constitutive of *Dasein*—that "being-in-the-world" we know as human consciousness. Computers don't care. Much less do they desire. And this is why, for all their computational competence, they nevertheless lack consciousness. A computer can't care. Hence there is no telos for computation. Hence there is no 'I am' for a computer, no locus of desire, no desiring subject. It takes desire to give time directionality, and once time has a direction—from a suffering or acceptable present toward a better future—then conscious life comes on the scene as the bearer of that desire toward a better future, toward a telos.

Heidegger was also sensitive to the importance of the eighth trait of emergent systems as it relates to life, namely death. For Heidegger, being-towards-death is constitutive of the nature of *Dasein* (the kind of being that is *human* being).

Eighth Trait: Coming Apart — the extinction of species

Evolution cannot work without extinction. Some variants must be selected *out.* From the point of view of the individual fish, death is not a good thing. Death looks very much like a failure at living. When life fails, death results. But this view is short-sighted. Without the death of individuals and the extinction of species, we would be condemned to a stagnation of the gene-pool. The current distribution of genotypes and phenotypes would be, as it were, frozen in aspic. Whether optimal or not, we would be perpetual. Where's the opportunity for improvement if the current models gain eternal tenure?

Consider an economic analogy, and here we must anticipate the later treatment of wealth and markets as emergent systems, but it is not too soon to allow different levels of emergence to illuminate one another. The argument has been made, both with respect to centrally planned Communist economies, as well as to cronyism in Japan, that economies cannot prosper without those opportunities for *disinvestment* known as bankruptcies. Some companies, some economic activities, add value. Some do not. If a central government props up failing (often state owned) enterprises simply for the sake of providing jobs or helping friends, cronyism, then, over the longer run, those enterprises will sap finite resources from other activities that might have added more value. Clearly, from the point of view of a given enterprise, bankruptcy is a bad thing. But from the larger, longer perspective of a whole country over a period of decades, the insistence on

propping up unproductive enterprises is a worse thing, as the history of the Soviet Union and the long recession in Japan have demonstrated.

Extinction is the ultimate disinvestment. As such, like Schumpeter's "creative destruction," it clears the way for new, and possibly better, life. From the point of view of the individual person, death looks as meaningless as random mutation. But from the larger, longer perspective of evolution, bankruptcy, extinction, and random mutation all serve to enrich the range of experiments, and thereby increase the likelihood of improvements on current models. Thus it can be said, without falling into the creationist version of the argument from design or relapsing into an un-reconstructed Aristotelian teleology, that bankruptcy, extinction, and random mutation do in fact *serve a purpose*: that of increasing the likelihood of improvements on current models. There's no magic here, no future telos pulling the present like water down a drain; just the statistics of the opportunity space.

Evolution feeds on death and desire, and desire evolves. From its inchoate beginnings when desire doesn't quite know what it wants, through its sublimation and education, desire evolves toward love (See Part Four).

Language

Language is important. Language is something new under the sun since the Pleistocene Era. What *is* language? How did it emerge? How does it emerge each time a child learns to speak?

In this section it will be shown that language is not what we thought it was. It is not a set of verbal pictures that name things. It is not a code for linking words and things in one-to-one bonds. This picture of language, as something built up from the primitive and seemingly indisputable bonds between words and things, is a very plausible, common sense way of thinking about language. This seemingly straightforward way of thinking about language has been rendered unbelievably complex and arcane by the discipline known as generative linguistics. In this section it will be shown that not only is language *not* what common sense thought it was; it will also be shown that language is not what most linguists think it is. For it will be shown that the vast superstructure of generative linguistics is built on a picture of language that is fundamentally wrong. Language is neither what common sense thought it was *nor* what most linguists thought it is.

Bold statements are needed in order to frame this part of the discourse lest the reader mistake what is at stake here. Once again we are in danger of looking under the light at the corner when our quarry is in much darker places. Language is complex. It is a complex adaptive system. But it is not as *complicated* as most linguists make it seem. The linguists—mainly Chomsky and his many followers—*start* with *too simple* a picture of language, and then render it unbelievably complicated in order to account for all of the exceptions and irregularities that their too simple foundations cannot seem to account for. Generative linguistics is like a vast Rube Goldberg machine erected to spit out a few relatively simple sentences. It is a classic case of what Part Four will describe as the phallusy of misplaced physics. Generative linguistics is a misguided attempt to analyze an emergent system using the tools of monological science. As such, it has created a picture of language that is unbelievably complicated, so complicated that no child could ever master it, certainly not a child of only one or two years old. Yet children do.

How do they do it? How can little children master this complex emergent system called language? Because at a certain point *it all comes together*—the life, the love, the evolution of consciousness necessary for mere sounds to *mean* anything at all.

In this section we shall see how *symbolic reference* actually works, and how it could have evolved. It is not simple. It is complex. But it is not as complicated as the linguists make it appear. The linguists had to erect a tremendously complicated machine because they were not willing to allow themselves the feminine comforts of love and intentional consciousness. They insisted on building their language machine from the kind of simple, hard, cold elementary particles that physicists could respect.

The exciting thing about the emerging picture of language as a complex adaptive system is that, from this perspective, a lot of the old puzzles in linguistics simply disappear. They turn out to be problems we don't have to solve—like "the problem of other minds" that turn out to be not that *other*, but intimately related from the get go. Looking back from the perspective of language as an emergent system, we can *save the phenomena.* We can see how language *might have appeared* to be a very complicated machine built up out of very simple elements. We will be able to see how and why linguists should have been forced to invent all of the parts to the machine that they did: deep structure, generative grammar, complicated rules for manipulating phonemes. But from the perspective of a science of emergent systems, we will be able to see how all of this mechanical clap-trap is really unnecessary—how a small child that is loved by its mother can evolve a consciousness sufficient to allow its burblings and gurglings to one day, all of a sudden, *mean* something.

We are dealing here with two fundamentally different paradigms for what language is. The first, the old paradigm, was embraced and promoted by the early Wittgenstein. It is called the picture theory of language, and it suggests the possibility of an ideal

language—a perfect language that is so crystal clear and unambiguous in the way that it pictures facts that the truth of a proposition would be evident from the mere form of a proposition. Read Wittgenstein's early and very brilliant *Tractatus Logico Philosophicus* and you will see this fantasy of language worked out in breathtaking detail.

Then Wittgenstein grew up. He saw the light. He realized the error of his ways, his commission of the phallusy of misplaced physics. He turned. There is no more brilliant critique of the old paradigm of language than Wittgenstein's later work, his *Philosophical Investigations,* where he incisively criticizes the picture theory of language. He shows how very different *ordinary* language is from the *ideal* language of the *Tractatus.* He deconstructs the monological picture of unambiguous *meanings.* He shows how there are many different language games other than the kind of simple picturing that was supposed to serve as the foundational layer of the edifice constructed by the old paradigm.

Chomsky, too, has turned . . . but not as completely. In his later work he has abandoned a lot of the complicated clap-trap of generative linguistics. He has adopted what he calls "a minimalist theory." But he still insists that he can reduce linguistics to a natural science, where what he means by 'natural science' is a science devoid of the kind of intentionality characteristic of consciousness. His natural science is not a science of emergent systems. The incompleteness of Chomsky's turn has been shown in reviews of his work by philosopher John Searle, whose access to the new paradigm of language allows him to nail not only Chomsky's mistakes, but also the inadequacies of Steven Pinker's formidable arguments in *The Language Instinct.*

The point of invoking some of these names at the outset of this section—Chomsky, Pinker, Searle, Wittgenstein—is doubly twofold, two reasons for the sake of two different sets of readers. For the reader relatively untutored in the subtleties of linguistics, it's worth making two points: first, you may be so much the better off for a certain naivete; you may have saved yourself a lot of trouble at mastering a lot of clap-trap. But second, the approach to language being presented here—as an emergent system, as a complex adaptive system, as a system presupposing the coming together of life, love, and the evolution of consciousness—is not some wacky, unprecedented, totally new theory without scholarly support. There are allies already in the field: E.g., the late Wittgenstein, John Searle, Ilkka Tuomi, and Terrence Deacon, whose remarkable book, *The Symbolic Species,* will serve as a major source for the argument of this section.

For a second group of readers, for those already familiar with the literature of linguistics and the philosophy of mind, it's worth invoking some of these names, first, as a short-hand way of, as it were, choosing up sides. Let's be very clear at the outset what sort of game is being played here—new paradigm, not old; ordinary language as an emergent system presupposing consciousness, not ideal language as an elaborate code for picturing

things with words. Second, it's worth giving credit where credit is due. Footnotes are a form of gratitude.

Now that the playing field has been delimited, for both ordinary mortals and scholarly linguists alike, let's see how the eight traits of emergent systems show up in and shed light on the complex system known as language.

Trait One: No first instance of language; no first word

We are back to where we began, the first and most obvious instance of the first trait of emergent systems. Once again, the basic point is that the sound that was supposedly the first word could not have been a word if it weren't in the company of a whole lot of other words. Without the context of a language, a sound is just a sound. It doesn't *mean* anything. It may *sound like* a word to someone who has already mastered language—hence the conundrums of people trying to figure out whether parrots and/or dolphins and/or whales communicate with each other or with us in ways that count as language (on which more later). But a sound that *sounds like* a word cannot actually *be* a word unless it is in the context of a language that allows the sound to function not simply as an iconic sign but as a symbol.

What's the difference between an iconic sign and a symbol? This distinction was first made by the philosopher Charles Sanders Peirce. It is fairly elementary, but very important to our argument. Peirce used as an example of an icon the literal signs that appear in front of shops. For people who could not read, it wouldn't work to put the word 'cobbler' over the shoe-maker's shop. Better to put a picture, or model, of a big shoe. The big shoe was not really a shoe. It was way too big for any person to wear. But it worked as a *sign* for the place where shoes were made and sold and repaired. And it worked because it *looked like* the thing it represented.

An icon has a non-arbitrary relationship to the things it represents; there is a one-to-one mapping between the parts of the icon and the parts of the things it represents. The mapping is very simple. In the case of the big shoe over the cobbler's shop, the transform that operates between the icon and the things it represents is simply one of size. The big laces represent little laces; the big toe represents little toes; the big sole represents little soles. And the parts of the big shoe are geometrically related to one another in just the same way that the parts of actual shoes are related to one another. Iconic signs are similar to, or isomorphic with, the things they represent.

Not so symbols. Symbols refer to things in the world in ways that are a good bit more complex than simple similarity. They don't picture their referents with literal verisimilitude. They don't work by way of the one-to-one, monological relationships characteristic of *code.* Think of Morse Code. In a code, a given sign design *always* means just one thing. Dot-dash, always means A. Ordinary language doesn't work that way. Sometimes the sign-design 'bank' refers to the bank of a river, sometimes to a bank shot in billiards, sometimes to an institution for holding money. Just exactly what the sign-design 'bank' refers to in a given context is a function of the context, and of the intentions of the speaker who uses that sign-design in that context.

Trait Two: Language pops

You can't have half a language. You have it or you don't. You can or you can't. Sure, there are many degrees of eloquence, grade upon grade of fluency, as anyone who has thoroughly learned a second language knows. But the ability to speak, to use words *as* symbols – this one has . . . or doesn't.

It is hard to imagine the dawn of language, but it must have looked very like the steep S-curve that mapped the number of buttons picked up against the ratio of strings to buttons – that is, steep to the point of looking like a step-jump. One day there was no language, just a lot of sounds strewn about like so many unconnected buttons. Then the next day there was a shimmering of communication and mutual recognition like nothing ever experienced by hominids before. What a day that must have been! For it could not be the case that the hunting band sat around communicating with just one word, the first word. That makes no sense. As soon as there was language, there were quite a few words. Hard to tell how many. Hard to tell the minimum number of differences sufficient to make a language. But it has to have been more than five or ten if any of them was to function symbolically and not just performatively like the roar of a lion, or iconically like 'bow-wow' for the bark of a dog.

Danny Hillis, inventor of the parallel processing super-computer, speculates, in an essay on the concept of emergence and language, that language emerged from the singing of early hominids. He imagines a time when hominids used their evolving vocal chords to sing for the sheer delight of the sounds, somewhat like birds. Such an origin helps satisfy the requirement for sufficient complexity before a new order can emerge—enough buttons prior to a system for their interconnection.

Ontogeny recapitulates phylogeny. Babies indulge in sing-song sounds before they can speak. They play with their voices, their lips, their tongues, their larynxes. They learn to

make a lot of different sounds before any one sound can *mean* anything at all . . . and then, *pop,* they start talking.

They must be conscious to do so. Consciousness is a prerequisite of language and, interestingly enough from the point of view of the way emergent systems *come together,* language is a prerequisite for consciousness. Language and consciousness co-evolved as emergent systems during the pre-history of the human species, and they co-evolve again and again in particular cases every time a young child learns to speak.

Trait Three: The Whole influences the nature of the part

This is de Saussure's point about the importance of seeing language as a synchronic whole rather than as the product of some sort of diachronic progression from first instances, and it feeds back on the impossibility of a first word.

This is also Derrida's needlessly obscure, typically French point about the priority of writing to speech. Okay, an "arche-writing," a very special kind of writing that is "inscribed" in something other than bound and printed books. His real point is simply that enough of language must be always already there (*toujour deja la*) before any particular sign-design can be taken as a *word,* and thus *spoken.* Derrida is trying to wean us from the idea of a first spoken word. It couldn't have happened that way. You had to have a self-organized, complex array of differences before one difference could make a difference linguistically.

Trait Four: Language is self-reflexive

The science of emergent systems has gifts to bear to the discipline of linguistics. In the old paradigm of linguistics, language was not self-reflexive; it was denotational or referential. It did not refer to itself so much as it referred to things in the world. It pictured things—here (in language) the picture; there (in the world) the thing pictured. Here, the word; there, the denoted thing.

According to this fairly primitive but much elaborated picture of language, it made sense to imagine that the way we *learn* language is by ostension, by *pointing at* the things that words are supposed to *mean.* And surely ostension does play a role in the learning of language. People were not stupid to seize upon ostension as a means for learning language, and therefore as a crucial component of the meaningfulness of words.

What is then wrong with this picture of language and the role of ostension? How and why do we need to liberate ourselves from this picture? Here there is no better guide than the late Wittgenstein of the *Philosophical Investigations* where he asks us to engage in a number of different thought experiments intended to reveal the inadequacies of the old

paradigm of language. When I point to a chair and say ‘chair,’ how do you know that I am referring to the chair and not to my pointing finger? When I point to a red square and say ‘red’, how do you know I’m pointing to the color and not to the shape? When I point to you and then to me, and say first “you” and then “me,” how do you then learn that ‘you’ refers to me when you are using it, and to you when I use it? (In the language of the trade, words like ‘you’ and ‘me’, ‘this’ and ‘that’, ‘now’ and ‘then’ are called *indexicals;* their meaning in any particular case is a function of who is speaking, where and when, and they are the occasion for no end of riddles for a linguistics that doesn’t want to countenance the existence of conscious subjects as language users.)

Wittgenstein’s elusive and aphoristic thought experiments are designed to deconstruct the picture of language as a perfect code signifying or picturing a set of things or facts in the world—the very picture set up by his early *Tractatus.* As long as the reader remains trapped inside that earlier picture, Wittgenstein’s later thought experiments are like so many Zen koans: Riddles that are insoluble within the confines of that old picture; questions that are posed not so much to elicit answers, but to demonstrate the inadequacy of the context in which they are being posed.

Once one steps out of that old paradigm of language, however, once one has, to use one of Wittgenstein’s more graphic images, “allowed the fly to find the way out of the fly bottle,” then there is less senseless smashing of heads against glass walls. Once one understands language as an emergent system, as a complex structure of symbols and not as a complicated congeries of simple, monological codings, then the riddles of ostension simply evaporate. *Of course* you know I’m pointing to the chair because you already appreciate the role of pointing within the larger context of learning a language. *Of course* you know I mean the color when I point to the red square and say ‘red’ because you are a conscious being capable of comparing that gesture with many others, including my pointing to a green square and saying ‘green’, to a green circle and saying ‘green’, to a blue square and saying ‘blue’, to a blue circle and saying ‘blue’, to a green circle and saying ‘green’, and so on. You have learned which differences make a difference to the interpretation of different sounds and gestures.

This learning about how language works is, however, highly self-reflexive. You find yourself using words to talk about words. Certain words mean what they mean because of the way other words have been used in their company. This self-reflexivity of language is such that, with typical French over-statement and panache, Derrida is drawn to make the extreme statement, “There is nothing outside the text.” I.e. *no* referentiality, no pointing to things or facts that are independent of the text. This is, of course, solipsistic nonsense. But insofar as Derrida’s gnomic statement served to disabuse some of his countrymen of their Cartesian fixation on a representational paradigm of consciousness and language, it was, in its time and place, *useful* nonsense.

Looking back from the perspective of a science of emergent systems and a new paradigm of language, we can see how it made sense to deconstruct the picture theory of meaning and the representational theory of consciousness where language is conceived as the (silent or audible) soundtrack in a "Cartesian theater" of representational consciousness. The "Cartesian theater" is an invention, an image created to describe a certain kind of mistake in the philosophy of mind. Think about the difference between the image capture of a camera, and the image capture of a conscious human being. What is the difference between them? One way of thinking about the difference is to imagine that the conscious human being has something that the camera doesn't, namely a Cartesian Cogito that is *conscious of* the image in a way that the camera isn't. The camera *records* the image—either digitally, as a complex sequence of 1s and 0s, or as chemical changes in silver halide film—but it is not *conscious of* it. The camera does not know, for example, that the image of a tree on its film *represents* a real tree out there in the world, whereas the Cogito in the Cartesian theater *does* know as much. Or at least it *thinks* it knows as much. Sometimes, of course, it gets obsessed with Cartesian doubt and begins to wonder whether the images inside its mind, its Cartesian theater, *really do* represent things accurately, for it also knows that it has been fooled, e.g., by oars in water that *looked* bent even when the rational mind *knew* that they were straight.

What's wrong with this image of the Cartesian theater? For one thing it just relocates the problem it was meant to solve. The problem is the problem of conscious (rather than un-conscious) representation. How can the mind—or language, which is the main subject here—*represent* anything at all? The invention of the Cartesian theater tries to answer this question by differentiating the mind from the camera by positing a conscious observer inside the mind, the Cogito inside the Cartesian theater. The trouble with this invention is that it just relocates the problem to be solved by taking the relationship between represented and representer and introjecting it into the interior of the Cartesian theater. Once there, it's worth asking: How is the Cogito any different from a camera? How do we know that the Cogito accurately captures the image imported through the retina and optic nerve? What does it mean to say that the Cogito *knows* this image or is *conscious of* it in a way that a camera is not? Perhaps we should posit a COGITO inside the Cogito . . . but that way lies an infinite regress.

Let's instead dispense with the Cartesian theater altogether and solve the riddle of symbolic reference not by way of a representational theory of consciousness with language as its interior soundtrack, but rather by way of an account of *language in use among a community of social beings communicating with one another and with their day to day world.* In the terms coined by Charles Saunders Peirce, let's get beyond syntax *and* semantics to pragmatics. 'Syntax' refers to the rules by which signs and symbols interact with one another, their grammar. 'Semantics' refers to the way signs and symbols relate to things in the world outside language. 'Pragmatics' refers to the uses to which we put

language, the practical tasks to which we put our speech. Pragmatics include not only representation and description but exhortation, questioning, planning, scolding, inquiring, demanding—what Wittgenstein called the many different "language games."

So to say that language is self-reflexive is to acknowledge the tremendous complexity and inter-relatedness among the syntactical, semantic and pragmatic dimensions of language. It is not a complicated contraption built up out of simple, monological, ostensively learned representations. It is instead an emergent system that comes together only when there is a sufficiently rich set of interactions among sounds, signs, things, acts, and conscious intentions.

Trait Five: Language is not reducible

To what would you reduce it? Squawking for survival? So the brutalists might claim. But what a stretch to assert the survival value of poetry.

According to John Searle, even the Chomsky of minimalist theory is still intent on reducing linguistics to a natural science, a science of objective happenings devoid of subjective intentions.

Trait Six: Language is unpredictable

You never know what people will say! Or how they will say it.

Such a simple and stunningly obvious point . . . but it's worth recalling and pondering when thinking about the kind of emergent system that language is. While engaged in the practice of trying to understand what language is and how it works, it's important to be aware of what would *not* be a satisfactory outcome of the inquiry, namely, an account so rigorous that you could claim to be able to predict just what *will* be said or *can* be said.

One of the astonishing things about language that is all too easily forgotten is its infinite creativity and unpredictability. Unlike, say, the game of tic-tac-toe where the number of moves, and possible games, may be large but nonetheless finite, the permutations of letters and sounds and the uses to which they may be put is infinite. It will always be possible to say something that no one has ever said before. From just 26 letters, new jokes, new comedies, new tragedies, new poetry will be born.

Once you have learned a language, you don't just repeat the sentences you have learned by rote. You make up new sentences appropriate to new situations. Because the uses of language are not reducible to a natural science of stimulus-response reactions, you cannot

usually predict how someone will respond to a given conversational stimulant. If you could, conversation would be very boring (which it sometimes is if you are talking with someone who always responds the way they think they *ought to* respond).

Real dialogue is alive. It moves in unpredictable ways. Things get said that have never been said before. Real dialogue is not formulaic. You cannot *train* someone with enough formulas to make them predictably scintillating.

Trait Seven: Language and desire

As earlier, this seventh sub-section will be longer than preceding sub-sections, partly because the theme is unfamiliar, partly because it offers an opportunity to weave together a number of the points already made about the emergence of language, and partly because it calls for the presentation and elaboration of the fairly complex argument presented by Terrence Deacon in his important book, *The Symbolic Species.* Once again, it's worth giving credit to good work already done.

The Symbolic Species is a very big book that makes a novel and very complex argument. Several parts of that argument are worth traversing in considerable detail. But just to get the argument going by pre-figuring its ultimate destination, think of Deacon's book as making a very elaborate case for the proposition that *language evolved so that one hominid could say to another, "I love you," and mean it.* Nowhere in Deacon's book will you find his conclusion stated quite so concisely, but that is what he is ultimately saying. And that's why Deacon's argument belongs in this sub-section on language and desire. Deacon understands the inadequacy of a purely representational or computational account of language and consciousness. He sees what's wrong with old paradigm linguistics. He senses the need for a richer pragmatics that situates language in the context of an evolving human condition.

His book begins with a section on the differences between symbolic reference and simpler forms of iconic and indexical reference. He shows how language exhibits the first two traits of emergent systems. He begins by calling attention to the third trait, the role of the whole in influencing the nature of the parts. He calls attention to the fact that there are no simple languages, "with simple forms of nouns, verbs, and sentences." (Terrence Deacon, *The Symbolic Species: The Co-evolution of Language and the Brain,* W. W. Norton & Co., New York, 1997, p. 12.) Why is this so? Because the way symbolic reference works, as distinct from iconic or indexical reference, is irreducibly complex. Symbols are not linked to things in the world by one-to-one relationships of reference. Instead, symbolic reference presupposes a whole system of representation that is distinctly different from, though built upon, iconic and indexical reference. Symbolism

emerges from structured patterns of iconic and indexical reference without being reducible to those simpler forms of reference. Symbolic reference is no more the result of simply adding indexical and iconic references than the taste of salt is an additive result of the characteristics of sodium and chlorine. *Something else* happens in the neurophysiology of signal processing to produce symbolic reference, just as *something else* happens to produce the taste of salt. "When we interpret the meaning and reference of a word or sentence, we produce something more than what a parrot produces when it requests a cracker or what a dog produces when it interprets a command. This 'something more' is what constitutes our symbolic competence." (p. 68)

In a chapter appropriately titled, "Symbols Aren't Simple," Deacon builds his account using C. S. Peirce's taxonomy: "icons are mediated by a similarity between sign and object, indices are mediated by some physical or temporal connection between sign and object, and symbols are mediated by some formal or merely agreed-upon link irrespective of any physical characteristics of either sign or object." (70) Indexical relationships are constructed from multiple iconic relationships. Repeated instances of iconic reference induce patterns of association based on correlation or contiguity. Where iconic representation relies on the physical similarity of things and signs, indexical representation is based not on such direct similarity, but on repeated associations among iconic representations. The relationship between icons and indexes is in this way hierarchical. "Prior iconic relationships are necessary for indexical reference, but prior indexical relationships are not in the same way necessary for iconic reference." (78) The similarity between the big shoe over the cobbler's door does not presuppose frequent associations between ordinary shoes and the big shoe on the sign. The relationship between the big shoe on the sign and the ordinary shoes on a traveler's feet will be evident the first time a peasant comes to town from the country. When he walks to a second town and a third town and sees a big shoe over a door, however, his inference to a cobbler behind the door will be based on his indexical association with prior iconic relationships.

It's important to appreciate this relationship between icons and indices just so that we can see what symbolic reference is *not.* "The common sense idea is that a symbolic association is formed when we learn to pair a sound or typed string with something else in the world. But in the terms we have been developing, this is what we mean by an *indexical* association." (79) Symbols don't work this way. "Symbolic reference derives from *combinatorial* possibilities and impossibilities, and we therefore depend on combinations both to discover it (during learning) and to make use of it (during communication). Thus the imagined version of a nonhuman animal language that is made up of isolated words, but lacking regularities that govern possible combinations, is ultimately a contradiction in terms." (83)

Deacon takes over ten pages to articulate “the symbolic threshold” by describing a series of experiments on animals and then summarizing the results in a fairly complicated diagram showing both the hierarchical relationship between symbols and indices, and the *difference* between indexical reference and iconic reference. For our purposes, reproducing his diagram and quoting his conclusions will suffice:

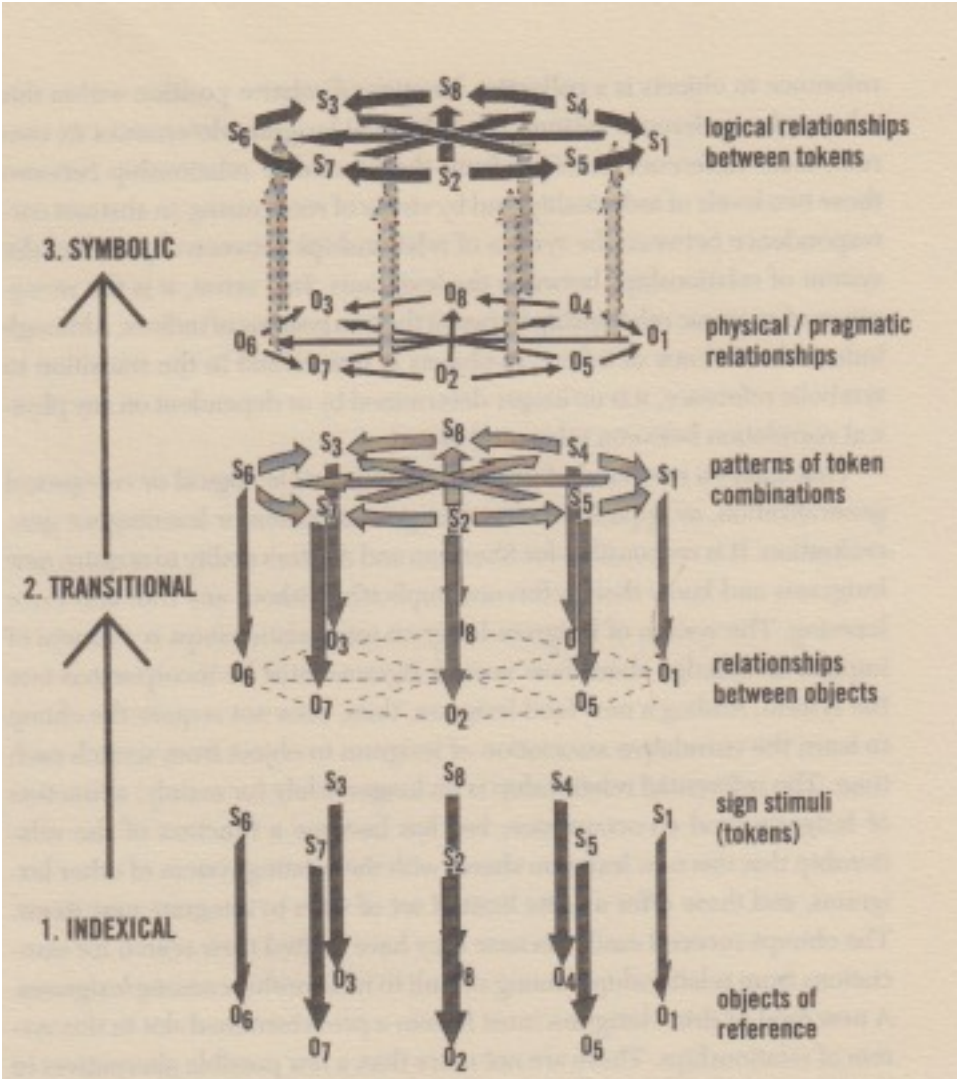


Figure 3.3 *A schematic depiction of the construction of symbolic referential relationships from indexical relationships. This figure builds on the logic depicted in Figure 3.2, but in this case the iconic relationships are only implied and the indexical relationships are condensed into single arrows. Three stages in the construction of symbolic relationships are shown from bottom to top. First, a collection of different indices are individually learned (varying strength indicated by darkness of arrows). Second, systematic relationships between index tokens (indexical stimuli) are recognized and learned as additional indices* (gray arrows linking indices). *Third, a shift* (reversal of indexical arrows) *in mnemonic strategy to rely on relationships between tokens* (darker arrows above) *to pick out objects indirectly via relationships between objects* (corresponding lower arrow system). *Individual indices can stand on their own in isolation, but symbols must be part of a closed group of transformations that links them in order to refer, otherwise they revert to indices.*

> What I am suggesting here is that the shift from associative predictions to symbolic predictions is initially a change in mnemonic strategy, a recoding. It is a way of offloading redundant details from working memory, by recognizing a higher-order regularity in the mess of associations, a trick that can accomplish the same task without having to hold all the details in mind. Unfortunately, nature seldom offers such nice neat logical systems that can help organize our associations. There are not many chances to use such strategies, so not much selection for this sort of process. We are forced to create artificial systems that have the appropriate properties. The crucial point is that when such a systematic set of tokens becomes available, it allows a shift in mnemonic strategy that results in a radical transformation in the mode of representation. What one knows in one way gets recoded in another way. It gets *re-represented.* We know the same associations, but we know them also in a different way. You might say we know them both from the bottom up, indexically, and from the top down, symbolically. (*Symbolic Species,* p. 89)

In order to grasp the significance of the symbolic threshold and the way it differs from indexical reference, consider an analogy. Imagine that you are a Martian conducting an analysis of all of the members of the U. S. Congress—all of the senators and all of the representatives. You have before you their names and their voting records recorded in the form of Yeses and Nos on a whole series of different bills that have been brought to a vote. You make an effort to get to know each individual congressman and congresswoman by memorizing his or her voting record. You get so good at this task that, when confronted with a given voting record, you can invariably bring up the right name. Martians, it turns out, have very good memories. By sheer brute force of memory (rather than iconic similarity) you can identify one string of Ys and Ns with the name Strom Thurmond, and another string of Ys and Ns with the name Edward Kennedy—and so on with every member of Congress, every time.

Now you begin to notice some correlations among voting records. You notice that the people who usually vote in favor of spending more money on education tend to vote against increased appropriations for the military. You needed to have the brute force (iconic) correlations between names and voting records before you could move from correlations of voting records to correlations of names. In order to "offload" some of the detail associated with the long lists of Ys and Ns, you decide to call the class of people voting for education and against military spending "liberals." You have now created a category based on something like an indexical association based on repeated contiguity. Certain folks "stick together" in the way they vote. Let's call them "liberals."

Now you are ready for a third level of insight. While no two members of Congress share *exactly* the same voting record, and the boundaries around the category of "liberals" is fuzzy, with some folks closer to the center of the clump and some folks often breaking ranks, you discover that you can save yourself a lot of brute force effort in memorizing names and voting patterns if you invent just two symbols: 'Democrat' and 'Republican'.

Now if you can link the name of each congressperson with a D or an R, you can predict with a fairly high degree of certainty whether they will vote liberal or conservative on any given issue. Granted, there are some conservative Democrats from the South, and some liberal Republicans in the coastal states . . . but mastery of these new symbols, Democrat and Republican, allows you to release a lot of your Martian gray matter to other mental tasks calling for your attention. As Deacon puts it, "You might say we know them both from the bottom up, indexically," from their liberal or conservative voting records, "and from the top down, symbolically," by virtue of their memberships in the Democratic or Republican parties.

What is "a Democrat"? When Bill Clinton tried to capture middle-of-the-road voters by co-opting some traditionally Republican issues by the strategy labeled "triangulation," — e.g., his advocacy of welfare reform that led to the resignation of some die-hard liberals—there were those on the left who said he had ceased to be a *real* Democrat. Yet the term 'Democrat' retained some meaning. The symbol of the donkey retained its indexical association with Democrats, while the symbol of the elephant remained stuck to the Republicans. And for many, but not all, voters in the 2000 elections, a mental labor-saving device when it came to casting votes consisted in throwing the master switch under the donkey or the elephant—unless a confusing ballot forced them back down to iconic details, mental doubt, and hanging chads.

This extended analogy is intended to make two points: first, the labor-saving virtue of symbolic reference, and second, the difference between top-down "symbols" like Democrat and Republican, and "indices" like the bottom-up voting records of proven liberals and conservatives. Just as a congressperson might have to prove himself or herself to be a liberal or a conservative by consistently voting yea or nay on issue after issue in a liberal or a conservative pattern, so indexical reference relies on the consistency of constant correlations. Nonetheless, if a congressperson bears the symbolic label, 'Democrat' or 'Republican', the Martian observer can bypass all the brute force of memorizing voting records in order to predict the next vote of any given senator or representative. And once symbolic reference has been achieved, it's unlikely that the Martian would slip back toward a brute force iconic association and indexical strategy of prediction.

Symbols provide a much more efficient strategy for knowing the world. As for so many voters, thank heavens for the Republican and Democratic parties! Having those symbols makes it so much easier to enter the voting booth. You don't have to involve yourself in the messy (iconic) details of issues; nor do you have to think hard about the ideological (indexical) abstractions of liberal or conservative policy. Just look for the (symbolic) donkey or elephant and throw the master switch to vote the straight ballot.

In order to drive home the strategic difference between indexical and symbolic reference, and the challenge of crossing the divide between the two types of reference, consider yet another analogy: calculating the volume of an irregular solid. We know how to calculate the volume of a cube: $V = S^3$ We know how to calculate the volume of a cylinder $V = \pi R^2 \times H$. Likewise we know how to calculate the volumes of many other regular solids like cones and spheres, etc. When confronted with an irregular solid, like the human body, we can approximate the shape and volume of that body by stacking enough regular solids together to approximate the precise shape of that body. But how much simpler—how much labor could be saved—by adopting a totally different strategy for determining volume! When Archimedes suddenly *got* the idea that the volume of water displaced by his body when stepping into a full tub of water was equivalent to the volume of his body, the revelation was so great that, as the apocryphal tale has it, he shouted *Eureka, ("I've got it!"),* and ran naked through the streets. *Such* is the power of the leap to symbolism. Just as it is *possible* to calculate the volume of a human body by the brute force summation of the volumes of hundreds of densely packed regular solids, so it is *possible* to go through life relying on the literalism of iconic and indexical reference. And just as it took the genius of Archimedes to make the leap to calculating volume by measuring the volume of displaced liquid, so it took a leap of genius to leave behind the strategy of indexical reference and adopt the wholly different mnemonic strategy of symbolic reference. But once the leap was made, how much easier it was to *talk.*

> The problem with symbol systems, then, is that there is both a lot of learning *and unlearning* that must take place before even a single symbolic relationship is available. Symbols cannot be acquired one at a time, the way other learned associations can, except *after* a reference symbol system is established. A logically complete system of relationships among the set of symbol tokens must be learned before the symbolic association between any one symbol token and an object can even be determined. The learning step occurs prior to recognizing the symbolic function, and this function only emerges from a system; it is not vested in any individual sign-object pairing. For this reason, it's hard to get started. To learn a first symbolic relationship requires holding a lot of associations in mind at once while at the same time mentally sampling the potential combinatorial patterns hidden in their higher-order relationships. (92f.)

In other words, it all has to come together. A symbol system must be *always already there—toujour deja la,* as Derrida likes to say—before a *single* symbol can be learned or used to communicate. Hence, no first instance (Trait 1). The whole influences the symbolic nature of any part, any single sign-design (Trait 3). The system is self-reflexive in the sense that symbolism depends on signs that refer to other signs (Trait 4). "[S]ymbols cannot be understood as an unstructured collection of tokens that map to a collection of referents because symbols don't just represent things in the world, they also represent each other. Because symbols do not directly refer to things in the world, but indirectly refer to them by virtue of referring to other symbols, they are implicitly

combinatorial entities whose referential powers are derived by virtue of occupying determinate positions in an organized system of other symbols." (p. 99)

Because the leap to symbolic reference is holistic, all of a piece, it *pops, viz.* the joy of parents when their toddler *gets it* (Trait 2). You can't predict what a child's first word will be, nor what else they will say when they learn to speak (Trait 5). Nor is symbolic reference reducible to summations of indexical reference any more than displacement is reducible to the summation of the volumes of regular solids (Trait 6). And what of the seventh trait of emergent systems, the principle of desire? What role does desire play in the evolution of language? We'll get to that by way of showing how the evolution of the brain kept pace with the evolution of symbolic reference, *not* vice versa.

Deacon's surprising and brilliant argument is that the human brain evolved to accommodate symbolic reference, *not* as has usually been assumed, that symbolic reference followed from the evolution of the brain. To describe Deacon's argument as a simple reversal of causal priority is too simple though. He does not replace one monological account with another. Instead he shows, as the sub-title of his book suggests, "The *co-evolution* of Language and the Brain." (emphasis added) In co-evolution, it all comes together. "Languages have adapted to human brains and human brains have adapted to languages, but the rate of language change is hundreds or thousands of times more rapid than biological change. . . The brain has co-evolved with respect to language, but languages have done most of the adapting." (122)

Part One of Deacon's book, "Language," concludes with an articulation of just what it is about language that sets the task for Part Two, "Brain." Once we appreciate the difference between indexical and symbolic reference, once we see the breadth of the gulf that must be crossed between summations and associations of iconic links on the one hand and the very different strategy of combinatorial systems among signs on the other, then we have a better idea about the kind of neuronal platform that will be required to achieve symbolic reference. Deacon concludes Part One with some surprising observations: "Immaturity of the brain is a learning handicap that greatly aids language acquisition."(141) It is as if our impatience or incapacity for summing the volumes of all those regular solids "greatly aids" our leap to noticing the displacement of water when we step in a hot tub.

How can a "learning handicap" aid language acquisition? Isn't the usual attitude toward language acquisition one of awe and wonder at the speed with which small children can acquire a working vocabulary? Deacon's understanding of symbolic reference runs contrary to the approach that has dominated most of 20th century linguistics, namely Chomsky's "Cartesian Linguistics" with its innate ideas about grammar, as well as Pinker's "Language Instinct" with its hard-wired rules for sentence formation. Because Deacon appreciates the complexity of the combinatorial logic that characterizes the

systematic structure of symbolism, he doesn't get caught in the usual game of trying to figure out how to combine icons and indexes into symbols. Rather than assuming that we must start with simple elements and then work our way inductively toward successively more complex combinations of those simple elements, Deacon sees that you can't get there—to symbolic reference—from here—iconic and indexical reference—by any direct or continuous accretion of similar instances, inductive leaps, or innate rules, any more than you can get to the *exact* volume of a gold crown (Archimedes' problem) by summing the simple solids that approximate its shape. You need a fundamentally different approach to reference, and for that, it's better that you handicap your old approach in order to drive you to a new one.

It's a little like the private in boot camp who learns you have to be smart to be lazy. Show that you are energetic and there will always be more work to do. Learn how to work the angles, though, and you can wangle the cushy jobs. (*Cf.* Joseph Heller's *Catch 22)*

Chomsky and Pinker are driven to posit innate ability and instinct because they conceive the job of symbolic reference to be *so* hard that, without presupposing such a solution to the problem before solving it, no solution could ever be found, certainly not by a two-year-old. Deacon's approach is radically different. Just as Archimedes' final solution to the volume problem is, once you know it, much *easier* than summing all those regular solids, so likewise, symbolic reference, once achieved, is much easier than accruing indexical references the way Comsky, et al., be damned. And the leap to symbolic reference, the *Eureka* of language, is actually rendered more likely to the extent that the human brain is biased against trying to get from indexical reference to symbolic reference by brute force.

> It is this insight—based on an appreciation of the logical gulf separating indexical from symbolic reference, and the labor-saving efficiency of symbolism once achieved—that sets the stage for studying the evolution of the human brain. So instead of looking for neural language modules, or postulating some global increase in general learning abilities, we need to begin to think of human brain evolution in terms of changes that could have produced certain biases in how we tend to learn. (*Symbolic Species,* p. 142)

What must a human brain be such that it can unlearn literalism in order to master symbolism? Sheer size is not the answer. Whales and elephants have brains that are larger than ours. Nor is the measure as simple as the ratio of brain weight to body weight. There is something about the *structure* of the human brain that differentiates it from other brains, and something about the evolution of that structure that is unique. As it turns out, the way Deacon describes the evolution of the human brain and its eventual structure has a great deal in common with the way Stephen Johnson describes the emergent order of ant colonies in his book, *Emergence.* In both cases, researchers have been challenged by similar mysteries. For the ant colony, the question takes the form, "How did they *know*

that the food was over *there?"* For the growth of the brain, which depends on neurons extending axons to "the right" locations so that the eventual architecture of connectivity generates the sort of structure adequate to symbolic reference, the question could similarly be posed: "How did so many axons *know* just how and where to grow?"

The old way to answer these questions, the monological way, would be to posit some sort of instructional *program,* whether instructions from on high, e.g., the more intelligent and authoritative "queen ant" of the colony, or instructions from, say, the DNA or various growth hormones in the embryonic brain. The new way, the way that works much better for evolving emergent systems, depends on a Darwinian logic of profligate variation and ruthless selection. Hundreds of ants wander *everywhere.* They don't *know* anything, much less where they are going. Some ants stumble across food. On their way back home they deposit pheromones. Other ants pick up the pheromone trail. What was at first a random walk, an undirected wandering in the wilderness, now takes on the structure of a highly directed march toward nutrients.

Now read Deacon on axon growth:

> Neurons overcome the problem of underdetermined target specificity by the same sort of logic that is used to match cell populations: selective elimination. They tend to overproduce [profligately] branches of their growing axons, and these sample a large number of potential targets during the early stages of development, though only a fraction of these connections are retained into adulthood. The remainder are eliminated in a competition between axons from different neurons over the same synaptic targets. This Darwinian-like process is responsible for much of the fine-tuning of neural connection patterns that accounts for the adaptive precision of brain functions. Like Darwinian evolution, the adaptive structure of neural circuitry emerges out of the selective promotion and elimination of specific variant patterns. By initially overproducing connections that have been spread to a wide variety of targets, and then selecting from among these on the basis of their different functional characteristics, highly predictable and functionally adaptive patterns of connectivity can be generated with minimal prespecification of the details. (p. 202)

"With minimal prespecification of the details." No marching orders. Very few instructions. Neither the ants nor the axons had to *know in advance* where they were going. "Such a strategy, while appearing somewhat wasteful of material, is highly efficient in its use of information. It circumvents the difficulties of planning ahead and allows development to proceed with a minimum of design or regulatory mechanisms." (195f.) Read that sentence over again and think about its applicability to the difference between centrally planned economies and free markets.

Deacon is describing the evolution of a brain that is adequate to handling the combinatorial logic of symbolic reference—a much more daunting task than the simple

coding of one-to-one iconic similarities or indexical references. Yet the effectiveness of symbolic reference, once achieved, is far more elegant and easy to work than the brute force accretion of iconic links or the chancy inductions from indexical associations. So the trick to be turned in the evolution of the human brain is to achieve a structure adequate to support the combinatorial logic of symbolic reference: Signs about signs, not just signs about things. How could this have come to pass?

Deacon develops an elaborate argument based on comparative studies of different animals, anthropological studies of early primates, and physiological studies of the larynx and the brain. This is not the place to summarize his entire argument. For the present purposes the important points are as follows: first, he finds it necessary to reverse the usual understandings of causes and consequences when it comes to thinking about the relationship between the brain and the evolution of language. “The remarkable expansion of the brain that took place in human evolution, and indirectly produced prefrontal expansion, was not the cause of symbolic language but a consequence of it.” (340)

How can this be? Because symbolic reference was so much more effective than indexical reference for solving the problems early hominids had to solve. What problems? Deacon makes a case, based on the helplessness of human infants, the need for women to stay with their offspring, the discovery of meat as a form of nutrition, the need for male hunting bands, and the need for stable social contracts between mothers and males who want to provide for their own genetic offspring. It is an elaborate argument that is to some extent inevitably conjectural . . . but it makes a very particular kind of sense, namely, a kind of sense that exhibits the traits of emergent systems.

Time and again Deacon argues for forms of “downward causality” based on the survival value of whole systems that determine the evolution of their parts: “the physical changes that make us human are the incarnations, so to speak, of the process of using words. . . Or, to put this miracle in simple terms, I suggest that an idea changed the brain.” (322)

Deacon is careful to distinguish his argument from Lysenko’s ideas about the inheritance of acquired characteristics. He relies instead upon the work of James Mark Baldwin. According to Baldwinian evolution, it is entirely possible for ideas to effect changes in the environment which in turn exert selective pressures on the survival of certain genes. “Baldwin suggested that learning and behavioral flexibility can play a role in amplifying and biasing natural selection because these abilities enable individuals to modify the context of natural selection that affects their future kin.” (322) Consequently, “More than any other group of species, hominids’ behavioral adaptations have determined the course of their physical evolution, rather than vice versa.” (345)

Both the first and third traits of emergent systems—the impossibility of first instances and the influence of the role of the part by the whole—are evident in Deacon’s argument

for the co-evolution of symbolic reference and the human brain: “to learn symbols it’s necessary to invest immense effort in learning associations that aren’t much use until the whole system of interdependent associations is sorted out. In other words for a long time in this symbol-learning process nothing useful can come of it. Only after a complete group (in the logical sense) of interdefined symbols is assembled can any one be used symbolically.” (378)

The achievement of symbolic reference is thus the result of the *coming together* of both the need for symbolic communication *and* a physiological platform capable of the necessary combinatorial tasks: “the evolution of language took place neither inside nor outside brains, but at the interface where cultural evolutionary processes affect biological evolutionary processes.”

Just as he made the case that children are better equipped to make the transition from indexical reference to symbolic reference by virtue of the immaturity of their brains, so he makes the case that the species was forced to make the transition to symbolic communication by virtue of the unique challenges posed by the helplessness of babies and the foraging habits of meat-eaters. Women needed food for their babies, and men needed to know that the babies they were feeding were *their own* babies.

“How does one tell who is obligated to whom?” asks Deacon.

> The first requirement, then, is that there must be a means for marking exclusive sexual relationships in a way that all members of the group recognize. Sexual access and a corresponding obligation to provide resources are not just habits of behavior; they cannot be more or less predictable patterns, or just predictions of probable future behaviors. Sexual access is a *prescription* for future behaviors. No index or memory of past behaviors can represent this. Nor can any index of present social status or reproductive state mark it. Even the refusal or avoidance of sexual activity only indicates a current state and is not necessarily predictive. Sexual or mating displays are incapable of referring to what might be, or should be. This information can only be given expression symbolically. The pair-bonding relationship in the human lineage is essentially a promise, or rather a set of promises that must be made public. These not only determine what behaviors are probable in the future, but more important, they implicitly determine which future behaviors are allowed and not allowed; that is, which are defined as cheating and may result in retaliation. (399)

Deacon thus arrives at an insight very similar to Nietzsche’s in the opening line to the Second Essay of his *Genealogy of Morals:* “To breed an animal with the right to make promises—is not this the paradoxical problem nature has set itself with regard to man?” (Friedrich Nietzsche, *The Genealogy of Morals,* trans. Francis Golffing, Doubleday, New York, 1956, p. 189.) Except that Deacon stresses not just the *right* to make promises, but the *capacity* to do so. And it is, truly, a paradoxical problem, for prior to the evolution of

symbolic reference, there could be no social contracts that would determine obligations between people.

> Without symbols that refer publicly and unambiguously to certain abstract social relationships and their future extension, including reciprocal obligations and prohibitions, hominids could not have taken advantage of the critical resources available to habitual hunters. The need to mark these reciprocally altruistic (and reciprocally selfish) relationships arose as an adaptation to the extreme evolutionary instability of the combination of group hunting/scavenging and male provisioning of mates and offspring. This was the question for which symbolization was the only viable answer. Symbolic culture was a response to a reproductive problem that only symbols could solve: the imperative of representing a social contract. (401)

In short, as was said at the outset of this section on the role of desire in the evolution of language, symbolism was necessary so that one hominid could say to another, "I love you," and *mean* it.

Let me pause a moments: "so that one hominid could say to another, "love you," *and mean it!"* Of course "I love you" might have meant quite a different thing from what it means today—'I love you,' today may mean something like, 'I adore you. I want to spend the rest of my life with. I think you're beautiful. Please accept this ring;" "I love you back then may have meant something more like, "I want your body, and if you give your body to any other man, I'll knock your block off!" You never know.

Like life itself, according to Margulis and Sagan, love binds time. Nor is love *just* desire. Love is loquacious. It speaks to make promises that bind the present to the past and future. Yesterday's pledge of loyalty secures tomorrow's meal, and for the male, a confidence in the continuity of his lineage.

Life, love, and language come together in the ritual of marriage and the rearing of a family. They are mutually implicated, mutually emergent parts of the human drama. The emergence of language and symbolic reference enables the emergence of love from animal desire. A stable, loving marriage enables and sustains human life in the face of an environment too hostile for suckling babies and nursing mothers left to their own devices. What now remains is to see how this coming together of life, evolution, and language suffice further for an account of the emergence of consciousness.

Trait 8: Death

Less important than the actual death of a language as the result of disuse or by the extinction of its users is the effective death of any and all languages by their descent into what Heidegger calls "idle chatter." When certain terms become meaningless as a result

of their misuse, or hackneyed usage—eg. “sustainable,” which is currently used not to refer only to a sustainable ecology but to sustainable profits”—then we suffer a mini-death—the uselessness of those terms to express anything meaningful.

Part Three: The Nature of Consciousness

Introduction

The time has come to scale the peak, the last great mystery, the reigning conundrum for philosophy, psychology and the cognitive sciences: the nature of consciousness. It's safe to say that all previous approaches have failed. If there's one thing that most current accounts agree on it is that no one has yet conquered the pinnacle of consciousness. It's therefore worth asking, like climbers in some Himalayan base camp: Might there be some new *approach* more promising than any yet attempted?

The approach taken here is different from others in three main respects: First, in order to circumvent the crevasse of ineffable interiority—the unreliability of introspective reports on purely private experience—this account will relocate the locus of questioning from the first person singular—the *I* of subjective experience—to the first person plural, *we.* We will ask: What is necessary in order for a group of people to be able to say, and mean, "We the people . . ."? What is necessary for a group to function as an autonomous, intentional, thoughtful agent?

Second, once we've made our way around that initial crevasse, then we will take a path to the summit that is not altogether unprecedented—Hume and Spinoza both started out in this direction—but never before has such a well-equipped hiking party mounted such an assault. The path in question follows a Copernican revolution that puts *desire* at the center of consciousness, and cognition as its satellite.

Third, what is the extra equipment this hiking party will carry? The reason we'll be able to succeed where Hume and Spinoza failed is that we'll be equipped with a *fluency* they lacked, a fluency gained from discourse about emergent phenomena like life, love and language.

Once life climbs the ladder of desire up to the level of love, and can articulate its desire in language, then consciousness emerges. This emergent property of the coming together of life, love and language is as it is largely as the result of the ways that life, love and language are as they are. We'll see how the nature of consciousness can be illuminated by the emergence of life, love and language . . . but first we must make our way around the crevasse of interiority, or what has recently been labeled "the hard problem."

The Elusiveness of Consciousness—The Hard Problem

The so-called "hard problem of consciousness" has been well articulated by David Chalmers. (David Chalmers, "Facing Up to the Problem of Consciousness," in *Explaining*

Consciousness—The 'Hard Problem,' ed. Johanthan Shear, MIT Press, Cambridge, 1998, pp. 9-30.) He argues that many attempts to explain consciousness simply miss the mark. They may explain one or another *function* associated with consciousness—the ability to discriminate environmental stimuli, the focus of attention, or the difference between wakefulness and sleep—but what they miss is the fairly obvious fact that each of these functions is somehow present to a *subject* of consciousness. "The really hard problem of consciousness is the problem of *experience.* When we think and perceive, there is a whir of information-processing, but there is also a subjective aspect. As Nagel has put it, there is *something it is like* to be a conscious organism. This subjective aspect is experience." (*Idem,* p. 10. Chalmers is referring to Thomas Nagel's paper, "What is it like to be a bat?," *Philosophical Review, **4,*** 1974, pp. 435-50. See also, T. Nagel, *The View From Nowhere,* Oxford University Press, New York, 1986, esp. Chapters 1 and 2.)

Chalmers quite rightly observes that you can marshal all sorts of scientific explanations to account for perception, wakefulness, agency, etc., and still lack an account of subjective experience. "This further question is the key question in the problem of consciousness. Why doesn't all this information-processing go on 'in the dark', free of any inner feel? Why is it that when electromagnetic waveforms impinge on a retina and are discriminated and categorized by a visual system, this discrimination and categorization is experienced as a sensation of vivid red?" (Chalmers, *op. cit.,* p. 13.)

While Chalmers is justly praised for having done a marvelous job of *formulating* the hard problem*,* neither he nor anyone else has satisfactorily *solved* it. Some even reject his formulation of the hard problem. Dan Dennett, for example, protests that, once you've given accounts of all of the functions of consciousness, you've explained everything there is to explain. There simply *isn't* anything else. "In the course of making an introspective catalogue of evidence, I wouldn't know what I was thinking about if I couldn't identify them for myself by these functional differentia. Subtract them away, and nothing is left beyond a weird conviction (in some people) that there is some ineffable residue of 'qualitative content' bereft of all powers to move us, delight us, annoy us, remind us of anything." (Dan Dennett, "Facing Backwards on the Problem of Consciousness," in the Shear anthology, *op. cit.,* p. 35.)

Dennett's approach hasn't convinced many, however. His book, *Consciousness Explained* (1991) has become known to many in search of a solution to the mystery as *Consciousness Explained Away.* John Searle says it best:

> The main point of Dennett's book is to deny the existence of inner mental states and offer an alternative account of consciousness, or rather what *he* calls 'consciousness.' The net effect is a performance of *Hamlet* without the Prince of Denmark." (John Searle, *The Mystery of Consciousness,* New York Review Books, New York, 1997, p. 100.) Further, "I regard Dennett's denial of the

> existence of consciousness not as a new discovery or even as a serious possibility but rather as a form of intellectual pathology." (*Idem,* p. 112.)

Dennett isn't crazy though. He simply wants to avoid the crevasse of ineffable interiority. Early in his career he came under the influence of the later Wittgenstein and Gilbert Ryle, who quite rightly warned us against making claims, based on introspective reports, about "the dogma of the Ghost in the Machine." (Gilbert Ryle, *The Concept of Mind,* Barnes & Noble, New York, 1949, pp. 15ff.) In his influential book, *The Concept of Mind* (1949), Ryle attacked "Descartes' Myth," according to which:

> Human bodies are in space and are subject to the mechanical laws which govern all other bodies in space. Bodily processes and states can be inspected by external observers. So a man's bodily life is as much a public affair as are the lives of animals and reptiles and even as the careers of trees, crystals and planets. But minds are not in space, nor are their operations subject to mechanical laws. The workings of one mind are not witnessable by other observers; its career is private. Only I can take direct cognizance of the states and processes of my own mind. (Ryle, *op. cit.,* p. 11.)

The dogma of Descartes' myth had a fatal flaw: It ran afoul of Wittgenstein's argument against the possibility of an altogether private language.

It's worth reviewing Wittgenstein's argument in order to see what it proves and what it doesn't. The problem is that Dennett and many others—the whole tradition known as psychological behaviorism—ended up throwing the baby out with the bathwater. They ended up throwing out consciousness along with private language. This was a mistake from which the philosophy of mind took about forty years to recover. From the late 1940s (around the time when Wittgenstein's later philosophy was exercising its influence on Ryle and others) until around 1992 (when John Searle published *The Rediscovery of the Mind)* both philosophy and psychology suffered a long drought during which inner experience was out of bounds. Talk of *mind* or *consciousness* or *private experience* was inherently suspect. The ghost in the machine had been exorcised. Subjectivity had been absorbed without remainder into objective, public, observable, behavioral science—to such an extent that, as one wag put it, when one behaviorist met another on the street she asked, "You're fine. How am I?"

Dennett is not alone, not crazy, not pathological in his denial of consciousness. He is part of a community, a decades-long tradition that convinced itself that most of our talk about subjective experience is as mistaken as talk about phlogiston. Under the influence of Ryle and the late Wittgenstein, philosophers like Wilfrid Sellars, Richard Rorty, Paul Armstrong and J. J. C. Smart convinced themselves and others of a view known as 'eliminative materialism,' according to which many of our first person singular reports of inner states would someday be eliminated and replaced by perfectly scientific descriptions of physical events in the brain, much as statements about humors have been replaced by statements about complexes among psychologists, or statements about

phlogiston have been replaced by statements about oxidation among chemists. (For an excellent account of various attempts to deny the existence of consciousness—and why they fail—see Owen Flanagan, *Consciousness Reconsidered,* MIT Press, Cambridge, Mass., 1992, especially Chapter Two, "Quining Consciousness.")

Suspicions about the acceptability of introspective reports have been circulating in the literature of psychology ever since William James's essay, "Does Consciousness Exist?." But Wittgenstein's argument against private language sealed the tomb of interiority in a way that took forty years to pry open again. So what was his argument? And what did it *not* prove?

As the preceding chapter on language pointed out, Wittgenstein was entirely right to criticize his own earlier picture theory of meaning. Despite the very real possibility of giving ostensive definitions of words once the complex structure of symbolic reference is already in place—e.g., "*This* [said while pointing] is a *disk-drive; that* is a *keyboard"*—the more fundamental job of establishing the nature symbolic reference relies, as Deacon argues, on a much more complex lattice-work of language and behavior: an un-learning of simple indexical reference, and a new learning of the way some symbols refer to other symbols that refer, in turn, to things in the world.

As Wittgenstein put it, "When one says 'He gave a name to his sensation' one forgets that a great deal of stage-setting in the language is presupposed if the mere act of naming is to make sense." (Wittgenstein, *Philosophical Investigations,* trans. G. E. M. Anscombe, Blackwell, Oxford, 1963, p. 92e, #257.) Wittgenstein's argument against the adequacy of ostensive definition and the picture theory of meaning applies to objective phenomena as well as to subjective phenomena. But it seemed to be particularly devastating in the case of introspective reports:

> 258. Let us imagine the following case. I want to keep a diary about the recurrence of a certain sensation. To this end I associate it with the sign 'S' and write this sign in a calendar for every day on which I have the sensation. —I will remark first of all that a definition of the sign cannot be formulated. —But still I can give myself a kind of ostensive definition. —How? Can I point to the sensation? Not in the ordinary sense. But I speak, or write the sign down, and at the same time I concentrate my attention on the sensation—and so, as it were, point to it inwardly. —But what is this ceremony for? For that is all it seems to be! A definition surely serves to establish the meaning of a sign. —Well, this is done precisely by the concentrating of my attention; for in this way I impress on myself the connection between the sign and the sensation. —But 'I impress it on myself' can only mean: this process brings it about that I remember the connection *right* in the future. But in the present case I have no criterion of correctness. One would like to say: whatever is going to seem right to me is right. And that only means that here we can't talk about 'right'. (*Idem,* #258.)

Wittgenstein's further images for the inadequacy of this ceremony are clever and persuasive. He likens someone trying to fix the right word of a private language to a particular inner sensation as being like someone trying to remember the time of departure of a train by recalling how a page of the time-table looked.

> "Isn't it the same here?" —No; for this process has got to produce a memory which is actually *correct.* If the mental image of the time-table could not itself be *tested* for correctness, how could it confirm the correctness of the first memory? (As if someone were to buy several copies of the morning paper to assure himself that what it said was true.) (*Idem,* #265.)

And further, "Imagine someone saying: 'But I know how tall I am!' and laying his hand on top of his head to prove it. (*Idem,* #279.)

Wittgenstein's very clever argument effectively destroyed the very notion of a private language—and so much the better. A language, in order to *be* a language, *must* be inter-subjective. It *must* be a means of communication *between* individuals. It *must* be shared. But the collateral damage of his devastating blast against private *language* turned out to include the very idea of the privacy of *interior experience*. For decades afterwards, the best and brightest of academic philosophers and psychologists were disinclined to attach any significance to reports from the interior. First person reports were off limits. For forty years, behaviorism ruled.

Philosophers returned to their senses in the 1990s, partly as a result of reactions to Dennett's influential but unacceptable explaining *away* of consciousness, partly as a result of John Searle's *Rediscovery of the Mind,* and partly as a result of assaults on Mt. Consciousness by computer scientists in the artificial intelligence community. By the mid-nineteen-nineties, *the race for consciousness was on!* But it has not been won.

Method Once Again

We must become self-conscious about method once again. This entry in the race for consciousness is less a deductive argument than an immersion in a set of examples drawn from a range of emergent systems. When it comes to consciousness, it is less a matter of *explaining* consciousness (as Dennett tried to do) and more a matter of *describing* consciousness in a language, fluency in which depends on practice in discourses about life, love, and language.

From the perspective taken here, it makes perfect sense that Dennett should have failed to *explain* consciousness in his book, *Consciousness Explained.* As Colin McGinn so rightly informs us, consciousness is not the sort of thing that *can be* explained by the ordinary

sort of science applicable to physical objects with volumes and locations and inertial masses that allow of measurement and mathematical calculation. Consciousness is not that sort of thing. You can't find it with a scalpel or weigh it on a scale.

McGinn does us a favor by clarifying just why consciousness is such a hard problem. His strategy starts off from a platform worth sharing: That of the intractability of emergent systems to monological science. McGinn takes off from the impossibility of *explaining* consciousness with ordinary science toward a skeptical conclusion known as "mysterian." He argues, in effect, *if ordinary science cannot explain consciousness, then consciousness cannot be explained*, period.

> My thesis is that consciousness depends upon an unknowable natural property of the brain . . . The materialists are right to think that it is *some* property of the brain that is responsible for consciousness, but they are wrong in the *kind* of brain property they select. The dualists are right to doubt that the brain as currently conceived can explain the mind, but they are wrong to infer that *no* brain property can do the job. Both views overestimate our knowledge of mind and brain, presupposing that our current conceptions are rich enough to capture the essence of the mind-brain link. I maintain that we need a qualitative leap in our understanding of mind and brain, but I also hold that this is not a leap our intellectual legs can take. (Colin McGinn, *The Mysterious Flame: Conscious Minds in a Material World,* Basic Books, New York, 1999, pp. 28f.)

So many "problems" associated with mind and consciousness! Isn't it time for some *solutions?* The tools are at hand. We now have at least a way to attack the hardest of the problems. What is it *like* to be conscious, ask a string of philosophers from Thomas Nagel to David Chalmers? Answer: *It is like life, evolution, and language, each in eight ways.* That makes no less than 24 parts of an answer; 36 when, later, we give an account of the emergence of love. If monological science is the *only* science at our disposal, then McGinn is surely right, that we can't explain consciousness. But if the monocular vision of monological science is supplemented by a science of emergent systems such that stereoscopic vision sees a new depth to experience, then, to switch anatomical metaphors mid-sentence, we might acquire the intellectual legs we need to take the leap of understanding that McGinn rightly sees as required.

The science of emergent systems is, to use the title of Stephen Wolfram's tome, *A New Kind of Science.* Wolfram, too, does us a favor by pointing out the limitations of the old kind of science, and Wolfram, too, leaves the same platform worth sharing in a direction not worth following. *His* new kind of science (and he never tires of telling us that it's *his),* mistakes the map for the territory. He over-estimates the explanatory power of models based on cellular automata. His basic insight, that complexity can and does derive from very simple beginnings, is a point worth appreciating. But Wolfram's attempts to apply that insight to different emergent systems often suffer from forced fits that work harder at fitting to the features of the model than to features of reality. The syntax is more impressive than the semantics.

Moving beyond McGinn's mysterianism, and respecting Wolfram's call for a new kind of science, this assault on Mt. Consciousness is not a deductive argument. Neither is it a progressive construction that builds on a sure foundation. Instead we seek a growing fluency in a language of emergent systems such that, when we come to consciousness, we are able to speak easily and fluently about its nuances. We are not "tongue-tied and flabbergasted" as Dennett puts it.

To what problems will this new fluency by applied? In addition to "the hard problem" of subjective experience, there are at least six other aspects of consciousness that call for attention:

i. The binding problem, namely, how do the several streams of sensory awareness—hearing, vision, touch, etc.—combine into a single field of sensory awareness;

ii. "The epi-phenomenalist suspicion," namely, the idea that consciousness may contribute no more to intelligent behavior than the train's whistle contributes to the speed of a locomotive;

iii. The mind-body problem: How does mind, which is non-physcal, exercise causal efficacy over physical behavior;

iv. The problem of other minds: How can one mind know another if both are as holed up in their own private interiority as some accounts of subjectivity would make it seem.

v. The question about *grades* of consciousness: How far down the phylogenetic tree can we justifiably attribute consciousness? Bonobos? Chimps? Dogs? Cats? Banana slugs? Protists?

And finally what most acknowledge as the *hardest* problem: What is it *like* to be conscious? We have no less than 24 parts of an answer to that question . In Parts One and Two we have constructed a kind of idea machine, represented in Figure One, The Table of Elements of Emergent Systems. We will now turn the crank on that idea machine and watch it issue forth answers to the question, *What is it like to be conscious?*

There are precedents for this approach. Terrence Deacon asks, in a manner sounding much like the students of Fritz Perls in the tradition of Gestalt Psychology, "What would it feel like to *be* evolution?" And he answers, revealingly, "It would feel very much like what it would feel like to be conscious." There's the same recursive, self-reflexive sampling of what has gone before in order to constitute what is now in one's field of experience; a similar structure of reproduction or re-presentation with variation and subsequent selection; a similar dynamic of anticipatory pre-presentation that has been described as "intentionality;" a similar demand for coherence in which the demand for

co-evolutionary adaptation in an ecology maps onto the demand for a neurological platform, a brain, adequate to the integration of different sensory streams solving the so-called binding problem of consciousness

Consciousness, as we shall see, also feels very much like what it is to be alive—like life. One of the likenesses that is most interesting is the likeness of consciousness to that aspect of life that Chris Langton got from von Neuman, namely, that it wasn't enough to build another of oneself. No, you have to go one better: You have to build another that can build *yet another.* In order to accomplish *that,* then life, whether artificial or real, needs something like DNA. It needs some sort of mechanism or program that can *re-present.* Representation is the key both to reproduction *and* to consciousness. Therein lies part of the paradigmatic likeness between consciousness and life.

You can also see how language plays on representation. The fact is obvious with every declarative sentence that purports to describe any state of affairs. Just *how* language does this, though, is far from obvious, in part because language is not obvious. Just as our impression of consciousness has been recently obscured by the computational metaphor, so has language. But when we're done sweeping away both of these mutual obscurations, we may be able to understand what Lacan meant when he said that the Unconscious is structured like a language. And *intentionality* will be a big part of the puzzle when we solve it.

All of these likenesses will be revealed by simply turning the crank on the idea machine built by weaving the warp of the eight traits across the woof of life, evolution, and language, then using that lattice to look very closely at consciousness. Insights derived from observing the traits of emergent systems in the evolution of life, evolution, and language will point the way.

The next three sections take three successive passes at applying the lessons of Parts One and Two to the question of consciousness. First, drawing on Part One, we'll offer a kind of executive summary from an altitude of 50,000 feet. We'll review some of the insights from the eight columns representing the eight traits of emergent systems. Next, following the high-level structure of Part Two, the horizontal slices across The Table of Elements of Emergent Systems, we'll ask what lessons there are to be learned from life, evolution, and language on a fairly general level—from 20,000 feet, about the nature of consciousness and money, love and artistic reativity. Finally, we'll drill down to a finer level of granularity by asking of some cells in the Table of Traits of Emergent Systems: What insights into the nature of consciousness can be derived from crossing each row and each column? How, for example, does the intersection of Trait 2: "Emergent systems pop," and the row of Evolution—punctuated equilibrium—shed light on the nature of consciousness? What is it *like* to be conscious? asks Thomas Nagel. We have three rows

times eight traits—24 answers to his question. We'll get to love later, and then we'll have four rows and 32 answers.

Lessons to be Learned from Life, Evolution, and Language

From the origins of life we learn the importance of self-containment—the cell wall, the membrane that marks the boundary between *inside* and *outside—that* is an important lesson for students of consciousness to learn: the boundary between me and not-me. Surely this has something to do with defining the boundaries of self and identity.

That which is self-contained becomes—or must be if it is to survive—self-interested. It learns to get what it wants, and it wants at a minimum food and energy. It must maintain itself against the ravages of entropy and heat loss. And if the germ-line is to continue, which a ratchet-effect based on preserving the winners and winnowing out the losers requires, then those self-interested cells are going to want to reproduce, at first mitotically by one becoming two identical copies of itself, and then as a way of fostering innovation and possible improvement, sexual reproduction mixes germ-lines so that two join to create one who is unique and importantly *unlike* both father and mother.

Evolution evolves, from mitosis to meiosis, and the drama of sexual reproduction evolves further, from the fertilization of fish eggs, through the elaborate plumage of birds and the mating rituals of higher mammals, to the love songs of the romantic poets and troubadours. From love there are lessons to learn about the importance of *desire,* and the different levels and grades of desire, all the way from the evident hunger of a protist swimming upstream in a glucose gradient, through the more developed appetites of progressively larger and more complicated animals, and finally to the refined tastes of highly educated human beings, and to the shared hopes of well constituted communities strategy. At every level on the hierarchy of desire you will find corresponding grades of consciousness.

Finally, language. What lessons does the emergence of language have for understanding the emergence of consciousness? Surely the answer has something to do with the importance of the whole for linfluencing the nature of the part—the need for a *system* always already *there* before any part of the system can function as it must function. Language provides an image of the kind of lattice-work, with feedback loops that bind the system back into a self-contained whole. Language does this, viz. de Saussure's stress on the *synchronic* rather than the diachronic dimension of etymology.

So a first pass through these several lessons for learning about the emergence of consciousness yields:

- Self-containment (Traits 3 and 4)
- Desire (Trait 7)
- Synchrony (Trait 2)

Each of these features of consciousness has a clear precedent in another emergent systems. Like life, consciousness must be self-contained (Trait 3). Subjectivity presupposes some sort of boundary between an inside and an outside, me and not-me. This self-containment requires some form of reflexivity (Trait 4), some self-reference, some of self-organizing system. It's loopy. There's a lot of cybernetics going on in consciousness. No wonder so many cognitive scientists got caught up in the metaphor of the mind as a computer, no matter how misleading that metaphor turned out to be. (Cf. Hubert Dreyfus, *What Computers Can't Do,* and John Searle, *The Rediscovery of Mind,* and *The Mystery of Consciousness.***)**

Consciousness represents itself to itself. It's not only a matter of perception, but of what Kant called *apperception.* Consciousness exhibits reflexivity in several ways on several different levels. On a fairly low level of body awareness, Antonio Damasio isolates several different feedback loops. Then on higher levels, Terrence Deacon diagrams feedback loops in several of his figures in *Symbolic Species* and in his papers on the three levels of emergence. Then there are feedback loops in the systems diagrams charting the agencies and companies and different actors and institutions that comprise the geo-political, economic, social landscape we occupy. Policies implementing adaptive strategies are as loopy as Damasio's and Deacon's maps of body awareness and the emergence of consciousness.

Further, consciousness is a satellite of desire (Trait 7). Computation alone doth not consciousness make. There must be some interestedness. Something has to *care.* Someone has to give a damn. And in that very caring, the someone-who-has-an-interest is born. So likewise with consciousness: It is precisely in the caring that the *subject* of all that computation is born, so that all of the *functions* of consciousness—perception, attention, intention—all have *a common purpose in the fulfillment of desire,* and the commonality, the commons created by that common purpose, *is* precisely the subject so necessary to the coming together—the community—of consciousness. Thus does the centrality of desire solve the so-called *binding problem.*

Still further on the lessons from Part Two, consciousness is surely *as complex as language.* Just as it takes many different sounds to make a language, so it takes many synapses, many connections, many signals to achieve consciousness. It takes many phonemes to form a language. And consciousness not only models itself on language in its degree of requisite complexity; further, consciousness as we know it in human beings *needs* language in order to represent itself to itself in narrative form. Consciousness

requires narrative in order to tell its own tale. There is an aspect of consciousness that is best captured in the question, “Who is the hero of this story? Who is the main protagonist? Whose point of view are we occupying? The satisfaction of *whose* desires count as a happy ending?”

Consciousness must be self-contained, it must be interested, and it must be complex. This much we know about consciousness from a first pass through what we learn from the emergence of life, evolution, and language. And this is only a *first* pass. Let us turn the crank on our idea machine yet further, through the eight traits of Part One as they apply to consciousness.

The Eight Traits as they Apply to Consciousness

No first instance

When you first wake up every morning, it is not the case that first one part of consciousness awakes, and then another, and then another. Just try to imagine the contrary, e.g., via the whimsical verse, “Toesies, toesies, please wake up. Kneesies, kneesies, please wake up. . .” and on to brain stem, limbic system, neo-cortex, and each of the functionalities of vision, hearing, taste, touch, memory and so on. No, it is not like that at all. When you wake up, when you become conscious each morning, all hands are on deck . . .

All at once! Consciousness pops!

No first instance, but all at once, like the onset of rain. It’s not about the first drop. With a slight shift in the atmosphere, it’s raining all over the place. First it wasn’t raining, then, POP! It’s raining.

This is how emergent systems begin. Synchrony is such that when it happens, it happens all over. That’s what synchrony *is*—happening all over, lots of instances, not just one first instance. As with the onset of rain, the important thing is not the first drop but the onset of conditions such that many drops will form almost simultaneously because the same dew point and humidity exist in a volume of space, not just at a single point. All over that volume, droplets of rain will condense more or less simultaneously.

Synchrony is easy to misconstrue as some kind of magical action at a distance—so-called *synchronicity*. Let it not be so. Let synchronicity not connote some kind of magical monological cause that defies the laws of physics, e.g., by acting at a distance even faster than the speed of light. No, let synchrony connote instead events that are meaningfully connected in such a way

that one has to wonder whether they are not the several effects of the same cause somewhere in their common past.

Against Jung's argument that seven sightings of something to do with a fish in a single day connote a proof of synchronicity: If you're looking around for proofs of synchronicity, you won't start counting until three or four show up, and once you start with something like 'fish' or 'square' . . . your chances of seeing at least two or three more in 24 hours are pretty good . . . whereas if you were to start with 'differential' or 'sprocket' or 'mauve,' then the chances of getting another six would truly be minimal. The mendacity of Jung's argument consists in its suggesting that wracking up three or four more instances of 'fish' or 'square' has the unlikelihood and therefore the high information value that a run of six instances of 'sprocket' or 'mauve' would have.

Consciousness pops! Singly in the form of waking up, and collectively in the form of social transfigurations of consciousness. The use of the term 'consciousness' to describe collective mentalities may be dismissed by some as merely metaphorical . . . but that would be a mistake. Here we want to invoke the move from first person singular, the consciousness of an *I,* to the first person plural, the consciousness of a *we.* There are borderline cases: Was "consciousness raising" in the women's movement about the *I* or the *we?* Conversations in women's groups led to transformations in individuals.

Less ambiguous is Hegel's use of "forms of consciousness," (*Gestalten des Buwusstseins).* Or Charles Reich's descriptions of Consciousness I, Consciousness II and Consciousness III in his now almost forgotten book, *The Greening of America.* That book was about the kind of abrupt transformations of consciousness that were readily accepted, perhaps *too* readily accepted, by the culture of the Sixties. Think of the kind of collective euphoria in 1967 known as "The Summer of Love" in the San Francisco Bay Area. Or Paris in 1968. Ask people who experienced these phenomena, and they will talk to you about a collective transformation of consciousness that *popped!* If they were not *too* affected by the experience, and still have the vocabulary, they will use words like *metanoia.*

It matters not for the present argument that the transformation did not stick, that after Altamount and the end of the Vietnam War there was a kind of collective return to somnolence and the putting of right feet in front of left feet as a generation temporarily doused in *love* turned its hands back to *work.* As we will explore at greater length when we get to Trait 8, consciousness, collective as well as individual, is the sort of thing that goes to sleep even as it is the sort of thing that wakes up. The fact is, *it happened,* and even at work there are still traces of that brief *pop,* as in the distinction between "Getting it," and those who just don't get it. (Cf. *Cluetrain Manifesto,* a book published in 2000 by Rick Levine, Christopher Locke, Doc Searls, and David Weinberger, Basic Books.)

Holism, Systematicity

What if the internet *woke up*? What would that look like? What would it be like to find oneself as part of a much larger intellect? What would it be like to discover that you are less than a tiny lobe on a brain much larger than any puny individual could imagine? What then? Might you discover that you are but a tiny part of a much larger whole? Would you collapse under the weight of security? Or would you relish the support of a system?

Consciousness is whole. It is all of a piece, whether in its collective first person plural, or in its individualized, personalized form. It is greater than the sum of its parts. It is holistic. Just as Terry Deacon launched his book, *The Symbolic Species* on the startling insight that there is no such thing as half a language, so it makes no sense to speak of half a consciousness.

Talk of bi-laterality—left brain and right brain—does not constitute a counter-example to this argument. Nor does the Freudian split between Consciousness and the Unconscious. Both right and left, and "up" and "down" (if you'll accept the imagery) are aspects of *one consciousness.*

Consciousness is greater than the sum of its parts, the sum of its several functions. Once the binding problem is solved, then the whole influences the behavior and functioning of the parts, which is a more precise way of saying that consciousness operates holistically.

Recursivity: Apperception

Kant had an answer to the binding problem. He called it the "transcendental unity of apperception" . . . and it's not any clearer in German than it is in English translation. If those words do not conjure a clear image in your mind's eye, you are healthier than you thought. If the phrase, 'transcendental unity of apperception' actually means something to you on first reading, then you may be very sick indeed.

With the definition of the binding problem—how to bind together tactile, olfactory, auditory and visual data streams into one experience—and the ability to measure at least some of those data-streams (though we are still not very good at measuring the olfactory . . .), it should now be possible to understand what Kant was trying to say with his infelititous express, 'transcendental unity of apperception."

Let's read the expression backwards:

Apperception: not just perception—not just looking at something—but *ap*-perception: looking at looking at something. What the camera doesn't do. The camera can't perceive, because the camera can't perceive itself perceive.

Unity: The issue is one of creating—or emerging—some unity out of what otherwise might seem a disaggregated multiplicity, a manifold, a mess. Finding, or establishing a one from the

many, or a one *in* the many. Law and custom and a common epistemology define the boundaries of a collective consciousness, and thus the unity of the many, e.g., the "mind" of Consciousness I, or Consciousness II, or Consciousness III, or the medieval mind, or the Islamic mind . . .

Transcendental: A technical term of the trade, esp. in German idealist philosophy, esp. Kant; Not equivalent to 'transcendent,' quite the opposite. Where 'transcendent' denotes the far side of experience, e.g. God or Heaven, 'transcendental' denotes the near side: that which is necessary in order that experience be possible. E.g., to jump from the first person singular to the first person plural, what is necessary such that "we the people" makes sense? The answer to which will define the basis upon which the legitimacy of the law is founded. It's the quest for the transcendental that drove Kant to chart the twelve Categories of Understanding and the two forms of intuition, space and time. It's a similar quest for the transcendental that drives philosophers of constitutional law and jurisprudence to explore the conditions for the possibility of legitimacy of authority in the State. (See in this regard Philip Bobbitt's remarkable 900 page tome, *The Shield of Achilles,* New York, Knopf, 2002, esp. pp. 584-664, where he reviews the arguments of different philosophies of law.)

Taking the terms one by one, "transcendental unity of apperception" actually means something when you think about it, and it's very pertinent to the binding problem . . . and to the problem about qualia. What is it *like* to be conscious? First of all, reflexive: I am me. There is an immediate awareness of my source of subjectivity that only philosophers as smart as Dan Dennett can figure out ways to deny.

Different philosophers have given different names to this subjective locus of reflection. Descartes called it the Cogito, Latin for *I think*, and the grammatical subject for Descartes' famous sentence, "Cogito ergo sum," "I think, therefore I am." Sartre talks about the *pour-soi*, the for-itself, as opposed to the *en-soi,* the in-itself. Hegel wrote about the *für sich* as opposed to the *an sich*. Aristotle projected human purposiveness further down into the animal and organic tree of life than modern science would like to take it. He saw final causes, *entelechia,* everywhere. Not only is an acorn seeking to become an oak—and surely there *is* a legitimate sense in which it makes sense to say that the acorn *wants* to be an oak—but for Aristotle an inorganic falling body is "seeking its proper place." This kind of explanation-by-destination prevailed in natural philosophy from Aristotle all the way up to the rationalist Enlightenment, until Newton, Bacon, Galileo and Spinoza expunged all final causes from our ontology. But now even naturalists are finding that function and functionality are not eliminable from either our epistemology or our ontology. Granted, there may be no transcendent Final Cause, no Design such that everything that *is* so was *meant to be so.* But instances of functionality *have emerged.* Evolution, not The Creator, has seen to that.

For all his rigorous naturalism, note the topological similarities in a number of the diagrams in Terry Deacon's *Symbolic Species*, from the hierarchy of semiotics from icons to indexes to symbols (p. 87), to the diagram describing Baldwinian evolution (p. 323), to the evolution of a

snowflake in his article on the three levels of emergence. A similar diagram could be constructed for multi-level selection in evolution . Also the self-regulation and evolution of the earth's eco-system—the Gaia hypothesis of Lovelock and Margulis. In all of these different realms it is necessary to appreciate a certain "topology of causality," to use Deacon's language—a topology that loops back and forth between 'outside' and 'inside' of emergent systems with Möbius-strip-like moves. (See Douglas Hofstedter, *Gödel, Escher, Bach, passim* for more on this topology.)

When it comes to consciousness, reflexivity is a feature as familiar as the phenomena of *self-consciousness* and *memory.* For all his rigorous, post-Enlightenment refusal to recognize the seemingly teleological *subjectivity* of consciousness, Dennett's editing metaphor for consciousness exhibits the very same loopiness with which Edelman ("re-entry"), Deacon ("recursive topology"), and Damasio ("self-reference") are all quite comfortable.

Unpredictability: Free Will

If you try to think through the computational metaphor for consciousness, you soon discover that, by nature, *it's not a reliable program*. If it were, there would be no such thing as *free will* (see the very final pages of Part Four, below).

On a global level, the succession of collective consciousnesses are not predictable from their precursors. This is the significance of Foucault's advance over Hegel. Where Hegel claimed that different forms of consciousness succeeded one another according to a dialectical logic that was, at least in retrospect, *necessary,* Foucault accepted the notion of systematically coherent forms of consciousness—he called them *epistemes*—but he argued that their succession was purely *contingent,* just one damn thing after another with no hint of necessity or progress. Marxist dialectical materialism got caught up in the illusion that the next phase of history was *inevitable,* as if iron laws of history were bound to bring on the revolution and, after the revolution, a classless society. Now we know better. The illusion of inevitability has been critiqued in word (Cf. Karl Popper, *The Poverty of Historicism,* Cf. Theador Adorno, *Negative Dialectic.)* as well as in deed, with the fall of communism. On a local level, precisely what distinguishes conscious behavior from *unconscious* behavior is its autonomy, its creativity, its novelty—*its unpredictability*. It is this very familiar aspect of consciousness that so-called "folk psychology" tries to capture with the phrase, 'free will.'

Free will is a vexed topic. Many in the cognitive sciences community would claim that free will is a bit of folk psychology better off forgotten. At a fairly crude level of analysis, free will is opposed to causal determinism. The argument against folk psychology has it that free will can be defended only at the cost of denying determinism; but to deny determinism is to abandon the rule of physical law . . . and that way lies magic and madness.

Part of the payoff that will come from an adequate account of consciousness is that it will save the phenomenon of free will by virtue of an account of time for which the future is far less determinate than the past. The future is not wholly indeterminate. It is not utterly unpredictable. And the past is not wholly determinate. Revisionism is not completely epistemological.

Perched in the path of progressive determination, the present dries the wet clay of potentiality. That clay is never as indeterminate as *prote hyle* (Aristotle's "prime matter" which is to wood as wood is to the form of a bed). Nor is it ever as dry and hard as the block universe of LaPlace's omniscient demon, who, if he reliable data on the position and momentum of every particle in the universe at a given, with Newton's Laws and a lot of high-speed computer time, calculate the position and velocity state descriptions for any entity out to any reach in the future. Both prediction and redemption are *partially* possible. But neither can ever be carried to completion. Such is the nature of time.

Consciousness cannot be predictable. If behavior is completely predictable, then it looks "robotic," too much like the behavior of an automaton. It fails the Turing test. Predictable behavior is bound to be very boring. Consciousness is not boring.

At the end of the day, the issue comes down to this: Can one make a difference? Is there any point to understanding and struggle? Is caring worth it? Or is consciousness just so much noise like the whistle of the locomotive that does nothing to add to its speed, direction, or power? (This is the so-called "epi-phenomenalist suspicion.")

If our consciousness is not steering but only along for the ride, if consciousness is just a read-out of electro-chemical signals driven from below, then it's hard to see how those physical and chemical reactions could be anything but noise on a semiotic level. But if sequences of thought follow a logico-linguistic sequence of ideas and propositions—a "train of thought" drawn by *reasons* toward a logical consequence, not a train of loco-motion driven by electro-chemical *causes*—and if those logical consequences lead to action which in turn leads to changes on the physical level, then consciousness cannot be said to be just along for the ride. Consciousness is driving. Free will is a reality.

A clear case of consciousness *mattering* is the role of fantasy in sexuality. It is a matter of the mind, a "mental thing," that invests a pair of shoes or a pair of panties or a leather vest with sexual meaning? And then the heart beats faster, the cock gets stiffer, the vagina gets wetter. Sex is the laboratory where the mind/body problem submits to solution. The power of sexual fantasies proves that ideas matter.

Irreducibility: Eliminating Eliminative Materialism

Nor is consciousness reducible to desire. That way lies pop-Freudo-Marxism: man as a bundle of erotic and economic drives. Desire is important. But consciousness is not *reducible* to desire.

Consciousness will not be reducible to neural-net descriptions, however closely correlated sequences of state descriptions may be. This suggests that Nick Humphrey is mistaken in *How to Solve the Mind-Body Problem* when he so blithely accepts the identity hypothesis.

To say that consciousness is irreducible is to evoke a vast literature in the cognitive sciences and the philosophy of mind, namely the whole debate over so-called eliminative materialism—the idea that someday, when we know just how the brain works, we will eliminate all talk relating to mind, all "mentalese," and we will refer only to 'c-firings' and e-waves, etc. As Jaegwon Kim puts it, reductionism must reduce the number of things there are in the universe. Once we see consciousness as reducible to a purely physical account describing various dispositions of matter in ways completely covered by physics, chemistry and biology, then our ontology is trimmed of all those things we used to call 'ideas' or 'qualia.' They disappear like ghosts, gods, Santa Claus and the tooth fairy. They simply don't exist, and neither does what we used to call consciousness.

To the extent that this argument is wielded against a view of consciousness already too steeped in monological thinking, it finds a large target: Consciousness as the ghost in the machine, consciousness as monological unity of apperception, consciousness as Cartesian theater, consciousness as Cogito. Monological thinking will lead you to posit an *élan conscience* just as it led Bergson and the vitalists to posit and *élan vital* to explain life.

But so long as you see consciousness as an emergent phenomenon, then its relation to the brain is such that none of the classic arguments of reductionistic materialism apply. Once we accept that consciousness emerges at the interface between brains and their environments—their physical and social surroundings—and once we start working with an epistemology and an ontology that are avowedly more than monological, then many of the conundrums worming about under the umbrella called 'the mind/body problem' simply disappear.

In Damasio's *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain,* a fine book giving us the contemporary science describing the "ladder of Conatus" directly comparable to our hierarchy of desire, Damasio builds on Maslow's "hierarchy of needs," and then credits Spinoza with the basic insight, namely*, property dualism*: regarding mind and body as two aspects of one and the same substance, ambivalently named *deus sive natura*—god or nature. Damasio articulates a contemporary property dualism. But he aspires to reductionism, as he hides it in one footnote: "I hope a reductionist research strategy eventually will allow us to explain how we get from the 'neural map' level to the 'mental' level." But then he hastens to add, "although the mental level will not 'reduce to' the neural-map level because it possesses emergent properties

created from the neural-map level." (p. 325, n. 21) So what is he saying? Is he, or is he not, a reductionist?

This sort of property dualism makes some sense. It is an expression of reality's own ambivalence: to appear once as object and then again subjectively. Spinoza had some very profound intuitions which he then articulated in an artificially rational and geometric way. His architechtonic of definitions, axioms, theorems, postulates, lemmas, and explanations is all a rhetorical ploy. The essence of Spinozism is almost a feeling: A certain conviction in the implacable rightness of things. Granted there are tragic side-effects of contingencies – shit happens. But Spinoza's *Ethics* exudes a positively Buddhist equanimity in the face of tragedy.

Desire

Buddhism damps down desire to make (too much) room for tranquility. But it's worth defending desire, even at risk of the tragedy of unrequited love, or love that is consummated and then lost. There's no possibility of tragedy in Zen Buddhism. Whatever is just *is*. If non-attachment—desirelessness—is the norm, then you cannot experience the loss that tragedy requires. Who cares? So Oedipus goes blind. Worse things have happened. Next?

If it is true that there's no place for tragedy in Buddhism, then is that a strike against Buddhism or against a sense of tragic possibility?

Coming Apart: Falling Asleep

Before looking at the application of Trait 8 to consciousness, recall the ways it shows up in other emergent systems:

- What is death on the level of love? Divorce. Cf. Updike's terrific line in the firstedition "The Music School" in *The New Yorker* and inexplicably dropped in its reprint, namely: "Divorce has replaced death as the redemptive horizon."
- And what is the application of "facing the fold" on the level of language? There are several possibilities: The devolution of signal into noise, for one. Another is "idle chatter"—*Gerrede* (as Maquarrie translates Heidegger), what "they say," received wisdom, the expected—bullshit.
- Re: governance: the coming apart of communism with Perestroika, Glasnost, and the fall of the Berlin Wall; the tumbling of communism throughout Eastern Europe, and the coming apart, the dis-union, of the Soviet Union.

- Re: wealth: the popping of a bubble, market crashes. The "hidden hand" *must* be invisible: if all could see it, you wouldn't have a market, but a rigged system redolent of Soviet communism. This necessary invisibility is what drives cyclicality (along with the fog induced by futurity, uncertainty, and interlocking lag times). Because prices are unpredictable, bubbles are inevitable. Sooner or later, people will become subject to "irrational exuberance." They will bid up the prices on tulips, or dot.com stocks, or cyber-currencies to levels eventually recognized as ridiculous . . . and then comes the crash.

Recalling all of these manifestations of *coming apart* in the contexts of other emergent systems suggests that, at the level of consciousness, *coming apart* is clearly falling asleep—ceasing to be conscious. Consciousness is the sort of thing that comes to be and passes away . . . just like *that!* That is a peculiar sort of thing, very unlike most physical things, which persist continuously through time and space. The nightstand does not cease to be when you leave the room.

Consciousness does not march in step to the "march of progress." Consciousness quite regularly goes to sleep. What might it be doing while it is sleeping? Or is that, too, the wrong question? When you go to sleep, does consciousness simply disappear, like your lap when you stand?

What is going on with *you* when your consciousness goes on strike and you sleep? We don't really know, is the short answer. There's a longer answer giving what we do know about the neurophysiology of sleep. But as many books and papers as we can cite, the big question remains unanswered, e.g., the main function or functions that sleep fulfills. What's it good for? And what might be lost over the long term by those who use drugs like Medafinil as a substitute for sleep?

In the vast literature on consciousness, you find far too little attention to the obvious fact that consciousness *goes to sleep***.** (Of course there are exceptions, e.g., Owen Flanagan's excellent book, *Dreaming Souls: Sleep, Dreams, and the Evolution of the Conscious Mind,* Oxford University Press, New York, 2000.) What is sleep? What happens to consciousness when we sleep? *Where does it go?* Why is this a stupid question? What does it nonetheless show about how we are inclined to misconstrue consciousness by such monological habits as those we follow in dealing with physical objects?

When a small child learns to imagine the existence of a toy when it is out of sight, that amounts to an intellectual achievement worthy of graduation to a new stage of cognitive development. This is an achievement that a camera never attains no matter how fine its lenses. But this intellectual achievement is misapplied if devoted to the question, "Where does consciousness go when you go to sleep?"

That consciousness goes to sleep says a great deal about the sort of thing consciousness is . . . and is not. It is *not* the sort of thing that persists like a rock or a night table or a toy out of sight.

And it certainly isn't something that miraculously comes to be and disappears like a rare ghost in the attic, or the ball that the baby cannot imagine to persist when out of sight. No, consciousness is not *that* sort of thing. But what sort of thing is it that comes to be and passes away the way consciousness does every time you to to sleep and wake again? This everyday happening, this glaringly obvious fact about consciousness, gains far too little attention in the literature.

The simple fact of sleep shows that it's not about the meat. The same meat, the same molecules, very close to the same mix of chemicals is present when you are awake as when you are asleep. Attention to the material cause—what it's made of, wood or stone or soft tissue; magnesium, silenium—these sorts of inventories aren't going to get us very far along the path toward understanding the difference between waking and sleeping. In both states, the same physical stuff is present. Consciousness is about the way those physical things interact and function, it is not about the presence or absence of some additional thing, some *élan conscience*.

Most things we interact with do not come to be and pass away the way consciousness does in waking and sleeping. There are exceptions: daytime, tornados, clouds, moods. But most things persist . . . until they disintegrate.

Consciousness—or what Heidegger called *Dasein*— comes into being only to the extent that the fragility of its being, the very real possibility of its *non*-being, is apparent to it in what Heidegger called being-towards-death.

In an earlier section on evolution we saw the evolutionary advantage that accrued from death and extinction: Without a capacity to get rid of the older models, a species could not introduce new and improved models. What lesson might there be for consciousness in this application of the eighth trait of emergent systems?

Implications of Life for Consciousness, trait by trait

The last section drew implications for consciousness from a consideration of each of the eight traits of emergent systems—the columns in Table One: The Eight Traits of Emergent Systems. The section before the last drew implications from the preceding considerations of life, evolution, and language—the rows in Table One. Now it's time for a closer look at several other cells in the table. We'll start with the first row, Life, and ask how the realization of each of the eight traits in their application to Life sheds light on the nature of consciousness. Then we'll turn to evolution, then language. Not all of the cells will yield insights of equal importance. But each cell has *something* to say. And together, the 24 cells in their totality, yield a fairly fine-grained answer to Thomas Nagel's famous question, "What is it *like* to be conscious?"

Trait One: Origins of Life/No first instance: Implications for Consciousness

If you go back and read pages on "no first instance" the lesson you will draw relates to a process of questioning, not to a specific answer. From Aristotle to Margulis and Sagan and Robert Wright, the question, "What is the first instance of life," was dismissed as *the wrong question.* The nature of life is not to be determined by knowledge of its origins; it's not that kind of thing.

So likewise with consciousness. As Wright wrote, "How did life begin? Beats me," (Robert Wright, *Non-Zero,* p. 252.) so we do well to brush aside with equal insouciance the quest for the *origins* of consciousness. It's the wrong quest. We won't find *the cause* of consciousness any more than we can find *the cause* of life. Nor will we be able to find a first instance, not simply because we lack a time machine, or an epistemological analog to the fossil record; rather, we won't find a first instance because, as an emergent system, consciousness comes on the scene all of a piece, as part of a system, always already in relation to other consciousnesses. Forget about the first consciousness. It's no more important to understanding the history of consciousness than your first memory is important to understanding your own biography, the record of your conscious life. Forget about the first conscious being. She's no more important to understanding the nature of consciousness than your first waking moment is important to understanding the meaning of your whole day.

Trait Two: How life pops: implications for consciousness

Look at the desert in spring. Look at buds bursting into bloom. . . .A science of emergent systems leads us to expect life to pop. That's just the sort of thing that emergent systems do. If it doesn't pop—if it comes into being in a series of incremental steps like the orderly building of a house: first the foundation, then the frame, then the walls, then the roof, and then the doors and windows—then it isn't a living thing at all. It's an artifact, an artificial thing, a product of artifice, not organic life.

Which is not to say that there cannot be artificial life. But if the product of artifice is to qualify as being truly alive, then that artificial thing will have to be the sort of thing that is capable of popping, not the sort of thing that is built in gradual increments by some other artificer, a programmer with a separate blueprint drawn by a separate architect. (I know I'm repeating myself. As my friend, Alfonso Moneouri keeps telling me, "You've got to bang the reader over the head with just how revolutionary your book is.")

Which raises the question of artificial intelligence, and why the quest for artificial intelligence, at least at this writing in 2024, has taught us so little about the nature of consciousness. The AI research program tries to develop computer programs that would mimic human intelligence. The Turing Test asked whether a computer behind a screen, fed inputs from conscious subjects, could convince us by its outputs that it was a person and not a computer. Big Blue may have beaten Boris Kasparov at chess, but no computer

has passed the Turing Test. As Terry Deacon maintains, the Turing test is just a gullibility test. The Turing test does not *test for consciousness*. Terry likes to call the AI movement a search for *simulated* intelligence, not the artificial creation of *human* intelligence.

Is it alive? Is it intelligent? Is it conscious? –These questions are different, but they share more than their grammatical structure. They all ask us to imagine a threshold, a line of demarcation, a set of criteria for a pass/fail judgment. Where the questioning over Trait 1 probed for a before which/after which line of demarcation, here there's a quest for an above which/below which demarcation. And once again, a certain insouciance is appropriate. Wrong question.

You won't learn more about consciousness by finding the line through which it pops and above which it lives, whether you look at the phylogenetic tree of species, or at the emergence of consciousness in an individual. It makes no sense to ask where in the fossil record consciousness pops. Nor does it make sense to ask where on the spine, or where in the brain-stem, or where in the amygdala consciousness pops. It's not that kind of thing.

Trait Three: Life's holism and its implications for consciousness

Mitchell Waldrop wrote:

> . . . the connectionist idea shows how the capacity for learning and evolution can emerge even if the nodes, the individual agents, are brainless and dead . . . by putting the power in the connections and not the nodes, it points the way to a very precise theory of what Langton and the A-lifers mean when they say that the essence of life is in the organization and not in the molecules. (M. Mitchell Waldrop, *Complexity: The Emerging Science at the Edge of Order and Chaos,* Simon & Schuster, New York, 1992, p. 292.)

The essence of life is in the organization and not in the molecules. It's about the harmony, not the individual notes of the melody.

So likewise consciousness: it's not, as the young Wittgenstein and the early Russell seemed to think it was, about the picturing of atomic facts. Nor is it, as the empiricists seemed to believe, about assemblages of individual ideas or percepts or sense impressions. Your camera doesn't become conscious once it has snapped its 10,000th picture. Your camera can't be conscious because it isn't self-conscious, which brings us to . . .

Trait Four: Life's reflexivity and its implications for consciousness

Life is loopy. Recall Langton's epiphany:

> One night, he says, the pieces just finally came together. He sat staring at loops that extended their arms, curled those arms around to form new, identical loops, and went on to form still more loops ad infinitum. . . He had created the simplest self-reproducing cellular automaton ever discovered. "I had this incredible—volcano of emotion," he says. "This *is* possible. It *does* work. This *is* true. Evolution made sense now. . . This had closure on itself, so that the organism *was* the program. It was complete. And now all these things that I'd been thinking of that might be the case if I could do this—well, they were all possible, too. It was like a landslide of possibilities. The dominoes fell, and just keep falling and falling and falling." (As quoted in Waldrop, *Complexity, op. cit.,* pp. 221f., and repeated a second time, knowingly. Bang on the head!).

Consciousness is the last domino. The varieties of recursive experience evident in life, evolution, and language lead inexorably, like a line of falling dominoes, toward consciousness. . . except that the mechanism of a "domino theory" belies the non-necessary, stochastic logic of emergence. *Almost* inevitable, but not quite. Dominoes are deterministic. Emergence is not. Emergence starts from the thermal hum of entropic energy, not a set of rigid mini-billiard balls. So the evolution of consciousness may be highly likely once you get atoms, molecules, crystals, complex polymers, and life. Once we appreciate the role of recursion in the auto-catalytic sets that constitute the autogen (cf. Deacon), and see how such proto-life leads toward further recursions through evolutionary lineages that lead toward life, then we have the basic moves for bringing purpose into existence.

One of the marks of consciousness is surely its capacity to cast purposes before itself. Consciousness is teleological in a way that a camera is not. *You* can aim a camera, but the camera cannot aim itself. Consciousness *can* aim itself. Consciousness is self-conscious. Like life, it is loopy.

Trait Five: Life's unpredictability

"There is nothing in the molecular physics of hydrogen, oxygen, carbon or nitrogen that would allow you to predict life. Nor is there anything in the physical chemistry of more complex molecules, the polymers and amino acids that come closer to the complexity of DNA." (From the earlier section on life, yet again, redundantly. The method of this book is emergent. These arguments have not been made before in the extensive literature on consciousness.)

Likewise there is nothing in the molecular biology of membranes, neurons or synapses that would allow you to predict consciousness. Nor is there enough in the anatomy of

more complex physical systems—the eye, the ear, the tongue—that would allow you to comprehend the complexity of perception.

On the one hand we are, as Stuart Kauffman tells us, "at home in the universe." (The title of his book.) This is the kind of universe where life is not just a meaningless accident, but downright probable. And yet it is not necessary. It might have been otherwise. It was not predictable, even if it was likely. If, as evolutionary biologists like Stephen J. Gould are fond of saying, we could "replay the tape," we would probably get something very close to what we now have. But given sensitivity to initial conditions, and the non-deterministic fact of a non-ergodic (non-repeating) universe . . . there's enough wiggle room in the system to warrant a certain amount of gratitude for what we got

Likewise with consciousness: One of its remarkable characteristics is that no two are exactly alike. Because the contents of consciousness accrue in unique complexes for which path-dependency and memory are important, no two minds are identical. Successive cycles of determination by something like sense data and something like intentionality build minds with different histories. What you saw and felt and experienced in the past will determine in part how you see and sense and experience in the future. If you "replay the tape" of any given consciousness, you'll get something very close to what you now have . . . but not identical. Conscious in unpredictable (Trait 5).

Trait Six: Life's irreducibility and its implications for consciousness

Just as A-Lifers scoff at meat-chauvinism and claim that a silicon platform can sustain life, so the proponents of artificial intelligence believe that consciousness—or at least intelligence—can be resident in silicon. But scratch an AI theorist and you will find a computationalist—someone who believes that consciousness *can* in fact be reduced to a series of programs that happen to be running on meat rather than on silicon. The computational metaphor for consciousness is part of the materialist reductionist claim that the mind is *nothing but* the brain. According to so-called "eliminative materialism," all of our talk in *mentalese*—all talk of intentionality, desire, or purposive functionality—will someday be replaced by talk about episodic events in the material brain--c-firings, etc.,--and that these in turn can be explained in terms of rules and algorithms that are the "software" running on the brain's wetware.

But just as the question, *Is it alive or dead?* turns out to be much more difficult to answer than, say, the question, *Is it solid or liquid?,* so the question, *Is it conscious or unconscious?* turns out to be similarly difficult. Indeed, where much of the debate about consciousness seems to presuppose that the what Howard Pattee cals "the epistemic cut" occurs where unconscious life gives rise to conscious life, while the difference between inorganic matter and organic life is just a matter of chemistry, the account that comes out

of Terry Deacon's reflections on emergence, the origins of life, and the semiotic theory of evolution suggest that the epistemic cut occurs at the break between the inorganic and organic life, and that successive grades of consciousness are almost bound to evolve once what Terry calls *teleodynamics* have emerged with the origins of life.

Trait Seven: Life, Desire, and Consciousness

From an earlier section: "Every living thing, from the uni-cellular protist swimming upstream in a glucose gradient to a living company, must take in energy in some form or another. And it must metabolize that energy in a way that nourishes and maintains its integrity. The wonder of it is the incredible range of sources of energy, and the variety of metabolisms and forms that then qualify as living." (Yet another redundancy. Bang!)

Likewise the wonder of consciousness is the incredible range of desirables given the gift of consciousness. The point to be appreciated about the emergence of life as a pre-figuration of the emergence of consciousness is *the polymorphousness of desire.* What feeds one thing may poison another. What appeals to Peter doesn't necessarily appeal to Paul.

How does the emergence of life shed light on the nature of consciousness when it comes to the role of desire? By disabusing us of the monological quest for some single set of necessary and sufficient conditions for the sustenance of consciousness. Like life, consciousness can feed on a spectacular range of sensations, perceptions, ideas, categories, stimuli, intuitions, concepts, instincts, drives, impulses, etc. in multiple combinations and forms.

Trait Eight: Life, Death, and Consciousness

Living things die. Whatsoever comes together can come apart, but consciousness "comes apart" in at least two significantly different ways: Death, and the simple cessation of consciousness called *sleep.* (See the above section on sleep.)

Implications of Evolution for Consciousness, trait by trait

The last section marched cell by cell, trait by trait, across the row in Table of Traits marked "Life." Now we want to march cell by cell, trait by trait, across the row marked "Evolution." Again, the point is to see whether there are any hints to be derived from the application of the language of emergence to evolution that can help us in answering

Thomas Nagel's famous question, "What is it *like* to be conscious?" We have 24 answers: it's like life in 8 ways, evolution in 8 ways, and language in 8 ways. Now it's evolution's turn to shed light on the nature of consciousness.

Evolution, No First Instance, and Consciousness

In reflecting on the way trait one, no first instance, applies to evolution, we contrasted single-point-source mutation to the kind of population genetics exemplified by a sooty London offering higher affordance to black moths than white moths who would be revealed to predatory birds by their contrast with a sooty background.

The question may sound bizarre when so baldly stated: How is consciousness more like moths in a coal-burning London than like single-point-source mutations? A bizarre sounding question perhaps, but balder is better when it comes to the lessons that evolution has for what it is like to be conscious. Recall the earlier quote from Deacon and the answers suggested for his question: "What would it feel like to *be* evolution?" And he answers, revealingly, "It would feel very much like what it would feel like to be conscious." There's the same recursive, self-reflexive sampling of what has gone before in order to constitute what is now in one's field of experience; a similar structure of reproduction or re-presentation with variation and subsequent selection; a similar dynamic of anticipatory pre-presentation that has been described as "intentionality;" a similar demand for coherence in which the demand for co-evolutionary adaptation in an ecology maps onto the demand for the integration of different sensory streams solving the so-called binding problem of consciousness.

See how attention to the entire fitness landscape (like London with more or less soot) does more to illuminate the emergence of the next thought than some notion of a single idea springing *de novo,* from nowhere, like a single-point mutation. First instances of so-called "original ideas" occur only very rarely. Edison's light bulb was the result of grueling experimentation in which metal after metal was tested until finally a tungsten filament proved to be the best. Einstein may have been the first individual to formulate the theory of relativity, but that theory did not just fall into his brain from nowhere. To focus on first instances of "new" ideas is to risk falling into the trap of framing "free will" as an unmoved mover, a causeless cause; or "genius" as a kind of inner homunculus that spews forth brilliant ideas from nowhere. Appreciating the logic of no first instances in evolution will save us from erroneous thinking about the origins of thoughts in consciousness.

Punctuated Equilibrium as a model for flashes of insight

Earlier you'll find:

> Recall how Galton's polyhedron doesn't roll smoothly across the surface of time, but tips in fits and starts from one face to another, from one ecosystem to another, from one equilibrium to another. This point alone is sufficient to ratify evolution's satisfaction of the second trait of emergent systems. Evolution pops in examples like the so-called "Cambrian explosion," when many different species emerged all at once in the blink of a paleontological eye. (Redundancy. *Bang!)*

And so it feels when you solve a puzzle, when it all comes together, when you "get" a joke, when you quite suddenly *see* what all the fuss was about. What is it *like* to be conscious? Well, when it comes to the experience of one state of consciousness giving way to another, it's rarely a matter of just one thing giving way to another like a field of red replacing a field of blue. Nicholas Humphrey's *Seeing Red* is a brilliant essay, beautifully written and deliciously erudite. But its prime example—a monochrome field of pure redness—is as distant from the felt realities of everyday consciousness as a cyclotron particle collider is distant from dropping a brick on your toe. We learn less about the nature of consciousness by constructing arcane experiments at a great distance from ordinary experience than we do by attending to the rich panoply of competing and conflicting signals and ideas and intentions and memories that mix and clamor for attention at every moment of waking life. When we truly honor this everyday, ordinary mess, rather than enter the laboratory where we screen out most of its complexity, then we gain an appreciation for the way successive states of consciousness feel more like synchronic co-evolution rather than diachronic survival of the singularly fittest, or effect of a *cause* rather than a *reason.*

So-called "*flashes* of insight" are like Cambrian explosions of synaptic connections. *It all comes together* in a new way. The diachronic march of logical sequences, right foot after left, left foot after right, may characterize some of our logico-analytic mental processes . . . but the effort it takes to teach logic to school children is a measure of just how rare such purely deductive diachronic sequences are in human consciousness. We seem to "intuit" more than we deduce—quotes around "intuit" because, lacking a science of emergence, we don't really know what we mean by "intuition." It's another one of those black boxes inserted into our language in a space where monological reasoning failed to find its footing. Why do we intuit more than we deduce? Because the contents of consciousness are so various, so complex, that the simplicity of deduction does violence to that complexity. (

Monological science, with its mono-linear sequences of 9-ball hitting 7-ball nudging the 2-ball into the side-pocket, has given us a paradigm for how the world works. It has proven a useful paradigm. We could build neither clocks nor rocket ships without it. But when it comes to comprehending the nature and contents of consciousness, it is a

dreadfully inadequate and misleading paradigm for modeling what consciousness is and the way it works. The tangled bank of the evolutionary fossil record provides a model that is much more adequate. Precisely in its richness, in its sedimentary layers, and in its co-evolutionary complexity, punctuated equilibrium provides a better model for the succession of states of consciousness than does the tick-tock sequence of a mechanical clock.

In condtions of pain, slavery, discomfort of any kind, consciousness must be so filled with pain, consciousness might be scoured so badly that it could wish to be extinguished. But for most of the time, most of the day, I find consciousness mostly enjoyable, e.g. at this moment I am not far from a nicely burning woodfire, burning away is a fireplace built from carefully chosen river rocks, on a high hearth that throws off pleasant heat.

Inside, Outside, and the Triggering of Consciousness

Recall the new evolutionary thinking known as evo-devo. The basic idea can be abbreviated as follows: rather than relying on a gene-centric account of embryological inheritance and development in which DNA supposedly holds a "blueprint" for the eventual phenotype, evo-devo focuses on the role of environmental "triggers" that are responsible for turning on or turning off certain genes over the course of both embryological development and species adaptation. Rather than taking the genotype as the main driver of phenotypes, evo-devo says that both the phenotype and its environment play a role in influencing whether and how genes "express" themselves.

During the last half of the twentieth century, the vibrant tradition of business strategy underwent a decades-long transition from what was known as "producer push" to "consumer pull." Rather than allowing new product developers to decide what customers should have, building it, and then "pushing the metal" through their sales forces, more and more manufacturers decided to "get close to the customer" and find out what they *wanted* rather than tell them what they "needed."

So likewise with respect to consciousness: Once upon a time there was a story about consciousness—sometimes known as subjective idealism—that reversed the adage, "seeing is believing," and proposed, in effect, "you have to believe it in order to see it." "Those who have eyes to see shall see . . ." etc. Innate ideas are supposedly necessary if knowledge is to be possible. Or, with Kant, the categories of understanding and the forms of intuition give shape to *a priori* knowledge.

Clearly there's some wisdom in this tradition. A truly "naïve" empiricism gives too little credit to the role of intentionality in shaping experience. But just as clearly, a deep commitment to subjective idealism leaves us trapped within an interior consciousness without any direct recourse to the world outside. Descartes' doubt leads to a hopeless

skepticism according to which we cannot know things in themselves, only their appearances. The veil of Maya falls. We find ourselves trapped in a hopeless solipsism wondering whether we have any real access to "other minds." The lure of this error is strong. Even the brilliant Wittgenstein got seduced by its siren song.

Following the tradition of evolutionary theory from the gene-centric approach of the neo-Darwinian synthesis to the more recent rise of evo-devo, an intellectual historian cannot help but notice parallels with the long and earlier history of epistemology (the theory of knowledge). Time and again—from Descartes' doubt to Spinoza's rationalism, From Kant's subjective idealism to Hegel's objective idealism, from Wittgenstein's skepticism to Searle's more sensible realism—philosophers keep dragging themselves out of the pit of intellectual interiority to re-engage with the world outside. Seductive as it may be to introverted intellectuals, the absurdity of solipsism cannot long endure.

The transition from the gene-centric neo-Darwinian synthesis toward the insights of evo-devo remind us that consciousness is not such an interior thing. It is more of a "membrane" relating inside to outside than a "nucleus" tucked deep inside. In the development of the person, consciousness is to be found as much in its "expressions" as in its "blueprint."

Von Neumann, Recursion and Consciousness

Recall, yet again, von Neumann's law: that it is not enough that a living thing be able to produce something like itself; it must be able to produce something that can also produce something like itself. It is not enough that **A** produce **A';** **A** must be able to produce and **A'** that can produce **A'',** and so on.

While von Neumann's law holds for evolution, earlier we observed:

> The process of evolution exhibits this fourth trait of emergent systems in the form of lineages of organisms that reproduce true to type, but with variations. If reproduction were *perfectly* true to type, if there were *no* variation, then there would be no evolution of species, just monotonous repetition. While perfect reproduction is a feature of computer algorithms, such that variation is noise rather than signal, variation is a feature rather than a bug in evolution. (Bang!!)

The significance of this point for consciousness is just this: The computational metaphor of strong AI is misleading at best, dead wrong at worst. Brains are not computers. Yes, they process information, but however true it may be that digital computation can translate analog computation to any degree of precision, it does not follow that the analog processes in the brain are equivalent to analog *computation.*

Part of what is so extraordinary about consciousness is precisely its imprecision: that it does *not* follow rules precisely. Yes, as a result of learning and memory, today's consciousness tends to replicate yesterday's in the broad outlines of a consistent level of intelligence, accrued memory, and an abiding personality much as species reproduce true to type. "Plant a carrot, get a carrot, not a brussel sprout," as that song from the *Fantastiks* has it. But no single sprout is *exactly* like any other. Likewise, part of what is so extraordinary about consciousness is precisely its fecundity. Like evolution, it evolves, it does not repeat in endless loops or closed circuits of recursion. Consciousness learns.

Yes, consciousness is recursive in the sense that consciousness requires self-consciousness. Perception is impossible without apperception. Consciousness is not a camera. But what we learn from the cell that crosses trait four, recursion, with evolution is that recursion need not be perfect. Indeed, it's better, it's more fecund, if it's not. A little noise in the system, a little variation, is a feature, not a bug.

Unpredictability

Just as evolution is unpredictable, so likewise consciousness—and so much the better. Just as evolution may be *directional* in the sense that complexity produces greater complexity as the biosphere searches adjacent possibilities for ever more ways to make a living, so consciousness seeks ever new ways to entertain itself.

Irreducibility

The earlier section on irreducibility in evolution from which this section borrows insights, is long. Its task was to find a "third way" between the reductionism *up* of creationism or intelligent design and the reduction *down* of the "brutalists." A similar task faces those of us trying to understand consciousness. For consciousness the two paths between which we want to find a third way are easier to describe than the length of the earlier section on the irreducibility of evoluion might lead us to believe: The low road leads to eliminative materialism; the high road to dualism.

On the low road, consciousness is nothing but a bunch of brain processes—c-firings and d-firings of various synapses—for which a complete description will eventually be given in physico-chemical terms. On the low road, the epi-phenomenalist suspicion proves true. Mentalese is as eliminable as talk of phlogiston.

On the high road consciousness is, per Descartes, a separate substance, unextended, ethereal. Those on the low road have an annoying habit of imagining that any and all of their critics *must*

be traveling this high road toward a disembodied soul. Those on the low road, from Dan Dennett and J. J. C. Smart to the Churchlands and Richard Rorty, cannot imagine that anyone who doesn't agree with them could find a path any different from the increasingly implausible high road. And that is precisely the challenge for a theory of consciousness that is at once firmly rooted in a naturalistic account, yet irreducibly emergent. We needn't flee to dualism in order to resist reductionism. We needn't invoke an ethereal soul in order to deny the epi-phenomenalist suspicion.

Evolutionary theory thus bears a double gift for illuminating the nature of consciousness: First, the *logic* of evolution—the fecundity of reproduction with variation and selection—gives us an existence proof for the possibility of a consciousness that also reproduces (in semantic representation) intentional contents of consciousness in a stream of never ending variation and selection. Second, the *fact* of evolution shows us how life has evolved from primal soup through viral and bacterial growths to multi-cellular organisms and on up the phylogenetic tree toward ever more conscious beings. Rather than simply assuming an ontological dualism separating mind from matter, unextended from extended substance, the *fact* of evolution shows us how consciousness could emerge from purely naturalistic origins, no dualism required.

Desire, Evolution, and Consciousness

The exposition of the "cell" in Table One from which this section would draw insight is also long: "The Seventh Trait: The reciprocity of purposes and desire in the course of evolution: teleology and the evolution of purpose." There was talk of the evolution of evolution from mitosis to meiosis; of "life binding time;" of a "Copernican revolution" according to which life was not so much an anomaly that seemed to violate the Second Law of Thermodynamics *within* time, but rather the very source of *care* that constitutes the *story* of life; and finally, a recognition of the role of care in casting forth purposes that give meaning to life.

Care was not present at the origin of the universe. Care has come to be. A large part of the burden—and the significance—of Deacon's work in demonstrating the possibility of a purely naturalistic emergence of desire lies in its rational reconstruction of teleology *without* a God-given or pre-scripted telos. Teleological striving has come to be in the universe. It *has emerged*. And now that it's here among us, it's worth observing its role in the constitution of consciousness.

Consciousness is not *only* calculation or computation, though it is at least that. It is not *just* the processing of information according to rules, though rule-following is surely part of it. Just as care binds time, so desire binds the several streams of information processing that are surely present in consciousness. For an autonomous agent to act on its own behalf, there must *be* something like its *behalf.* Consciousness is information

processing for a *purpose*—the furtherance of an agent, the gratification of its desires, both short-term and long-term, both episodic and enlightened, both individual and collective.

Coming Apart, Death, Extinction and Consciousness

Whatsoever can come together can also come apart. But as we saw earlier, death and extinction are not all bad. They serve a purpose: that of increasing the opportunity space for the improvement. Like Schumpeterian "creative destruction," death and extinction are "disinvestments" that make room for new and improved models of species and corporations.

Consciousness learns, but in order to learn it must sometimes forget or unlearn. Recall Deacon's articulation of the leap from iconic and indexical representation to symbolic representation: "The problem with symbol systems, then, is that there is both a lot of learning *and unlearning* that must take place before even a single symbolic relationship is available." (yet again, Deacon, *Symbolic Species,* p. 92f.)

One of the features of consciousness that calls out for clarification is its curious capacity for persistence, intermittence, change, and development. It comes to be and passes away both in sleep and in death *and* across the boundary separating conscious from unconscious mental processes. What a very *odd* sort of existence consciousness exhibits, so central to human life, and so obscure.

The theory of evolution can illuminate the nature of consciousness precisely to the extent that the theory of evolution can show how *coming apart* can be as *functional* as coming together; how death and extinction can serve evolution; and, therefore, how going to sleep and forgetting and ceasing to be conscious can serve the further life of consciousness.

Part Four: The Phallusy of Misplaced Physics

The title is a play on Whitehead's "Fallacy of Misplaced Concreteness." In his 1925 book, *Science and the Modern World,* Whitehead developed an idea later discussed using the term, *Reification.* The basic idea is that we lead ourselves astray when we treat abstract ideas as if they were concrete.

The crucial idea behind the fallacy of misplaced concreteness is the notion of *simple location.* The long tradition of *materialism* featured the idea that material entities occupy a single finite volume of space-time. A table is just *here.* Not so justice, friendship, or any number of other abstract concepts that we mistreat when we imagine that they are simply located.

Where the fallacy of misplaced concreteness is about entities considered under the supposedly universal reach of deterministic materialism, the phallusy of misplaced physics is about *the interactions among such entities*: deterministic causality is not the only way that progression unfolds.

In an excellent review of Kevin Mitchell's *Free Agents: How Evolution Gave us Free Will* (Princton University Press, 2023)*,* James Gleick, one of the pioneers in complexity theory, summarizes the issue thus:

> For physicists, the problem is that we are made of matter, like every particle and planet in the universe, and matter is governed by physical laws. According to the physicist and bestselling author Briane Greene, "We need to recognize that although the sensation of free will is real, the capacity to exert free will—the capacity for the human mind to transcend the laws that control physical progression—is not." We do not and cannot cause anything; we are caused. "Our choices are the result of our particles coursing one way or another through our brains," . . .
>
> Sam Harris, a neuroscientist and philosopher who wrote the popular book *Free Will* (2012), insisted not only that free will is an illusion but that the concept "cannot be made conceptually coherent." Consider it a challenge: "No one has ever described a way in which mental and physical processes could arise that would attest to the existence of such freedom." (*New York Review of Books*, 18, Jan 2024)

Greene is not alone. Indeed the majority of contemporary philosophers and physicists grappling with the issue of free will are convinced of its illusoriness. (Cf. Robert Sapolsky's books, *Determined: A Science of Life without Free Will, (*Bodley Head, 2023); *Behave: The Biology of Humans at our Best and Worst;* Daniel Wegner*, The*

Illusion of Conscious Will, together with dozens of other books and essays preaching the illusoriness of free will.)

Why *Phallusy?* Because there is an unmistakable gender bias in approaches to the issue. Einstein, the male genius inventor of relativity theory, "envisioned the universe as a four dimensional space-time continuum. 'Everything is determined,' he said, the beginning as well as the end, by forces over which we have no control. It is determined for the insect as well for the star. Human being, vegetables or cosmic dust, we all dance to an invisible tune, intoned in the distance by a mysterious player." (Gleick)

Likewise, Nobel Prize winning physicist, Steven Weinberg, (a man) understood the entire universe to be all push, push, push. No pull. No teleology. No purposes of the sort that characterize free will and give meaning to life.

Several women, on the other hand, have no problem avoiding the phallusy of misplaced physics, most prominently Evelyn Fox Keller, e.g in her marvelous book, *Making Sense of My Life in Science* (Modern Memoirs, 2023) where the longest chapter, "Women, Gender, and Science" runs 72 pages. Allow me to review her argument.

On the very first page (95) she asks, "Why were there so few women in science?" On p. 103 she launches her answer: "I found that I could no longer put to one side the fact that physics was such an overwhelmingly male-dominated discipline." And on the very next page, " . . . the most significant obstacle facing women scientists lay not in their own nature, but rather in popular mythology in the widespread belief that the nature of women was antithetical to the nature of science—that science itself was, somehow, inherently masculine." In other words, the problem is not some inadequacy in women but in the nature of science.

She grants the controversial claim, "we must accept the possibility of ineradicable cognitive differences between the sexes." (108) She distinguishes between sex—the biological differences between humans who have a uterus and those who don't—and gender, which is largely a matter of cultural mythologies reinforced generation after generation. And she grants that the differences between men and women derive from both, biological differences *and* cultural mythologies. " . . . it seemed perfectly plausible to think of sex differences as simultaneously real *and* as socially *constructed*." (11)

> We may not have made the objects that scientists study, but we are certainly responsible for the making of what counts as science. And if science has historically developed in the context of strong beliefs about natural (inborn) affinities between men and mind on the one hand and between women and nature on the other; between masculinity and objectivity, femininity and subjectivity; it

> is not unreasonable to expect the emergence of a division of emotional and intellectual labor between actual men and women that confirms such beliefs. (118)

One of Keller's books, her first major book, was an intellectual biography of Barbara Mclintock: *A feeling for the Organism : The Life and Work of Barbara McClintock* (W. H. Freeman, 1983).

> McLintock's story provided the perfect foil for my arguments about gender. She was a woman; she routinely emphasized the importance of qualities traditionally labeled as "feminine" (especially, feeling, identification, and intuition); and she produced an apparently different kind of science, forging a radically different picture of the role of genetics in development from that of her mainstream (primarily male) colleagues. (138)

As far as Keller's own contributions to science go, it's important to recall her early work on slime molds. Slime molds are peculiar. They exhibit two very different states. When food is plentiful, they are multiple—a swarm of separate cells. But when food is scarce, they coalesce into a single organism capable of searching for food. The question that stumped researchers for years was *how* they coalesce. For decades the assumption was that there must be, among the swarm of disconnected cells, a few that could function as *master cells*—cells that could exercise some sort of executive control that would organize the rest of the cells to generate a single, multi-cellular organism. The trouble is, they could not find such master cells. Instead, as Keller and her colleague, Lee Segal, finally proposed (under the influence of an earlier paper by Alan Turing) that, rather than being organized by executive master cells, the many separate cells *self organize.*

This was a radical idea: that rather than being subject to executive organization *from above,* the many separate cells of slime mold in its plural state could self organize *from below.* Nothing like this had ever been explicitly theorized. True, once the science of self-organization was more mature, retrospection revealed many examples of self organization, from Friedrich Engels' descriptions of the streets of Manchester, to Jane Jacobs' analysis of the formation of neighborhoods, to Marvin Minsky's studies of distributed networks in the human brain. But all that retrospective discovery of emergence came later when the science of self-organization was more mature. Keller and Segal's proposal of self-organization in slime molds was a *breakthrough*—a stunning recognition of organization *up from the bottom,* rather than organization by executive function *down from the top.* (The story of Keller and Segal's breakthrough is nicely told in the first chapter of Steven Johnson's book, *Emergence: The Connected Lives of Ants, Brains, Cities, and Software* (Scribner, 2001)).

As Keller puts it:

> Prevailing views of gene-centric (or DNA-centric) biological dynamics provided a good example of what I had come to call "master molecule theories." By this term, I sought to include kinds of theories that assumed the existence of a single, unitary, and primary cause. McClintock's conceptions of cellular organization seemed to provide an excellent alternative to such theories. (140)

Further:

> I remained committed to the examination of all that had been left out of both science and science studies by virtue of the historic definition of science as masculine, by the cleansing from our ideals of masculinity of all things feminine, and hence by the exclusion from science of everything associated with femininity (such as feeling, subjectivity, corporeality). (147f.)

Example: "Faced with the charge that 'women always get personal,' [Mary] Ellman counters, 'I'd say, men always get impersonal." (143)

So, as already claimed, the reason that there are so few women in science has less to do with some inadequacies in women's brains and more to do with the masculine character of science. Most women tend not to fall for a masculine science. Hence the gender bias of the phallusy of misplaced physics.

Terry Deacon's take on reality is more than compatible with Keller's. His substitution of *constraints* for *causes* represents a different, less deterministic rationale for the way reality unfolds over time. The world of causes is deterministic; Terry's world of constraints is stochastic. Definition of 'stochastic' in *Webster's New World Dictionary* (The World Publishing Company, 1966): "having a random probability distribution or pattern that may be analysed statistically but may not be predicted precisely."

The philosopher, Charles Saunders Peircen had a concept he called *Tychism.* The idea is that chance is real. The future is not determined. Indeterminacy is objective (Heisenberg), not just subjectively "uncertain". Late in his life, he wished that he'd called his great discovery, "The Indeterminacy Principle" rather than "The Uncertainty Principle." The latter sounds too subjective. As Perirce's concept of *Thychism* has it, randomness is real, not just a function of human ignorance. It's not just that I happen to be uncertain about the position or momentum of any given particle; the position and momentum of the particle are uncertain. You can know one of them at one time, but not the other, or *both.* That's the way they *are,* not just the way you see them.

# # #

Founding editor of *WIRED* magazine and author of several big and important books, Kevin Kelly writes:

> We tend to interpret the mysteries surrounding life in imagery suggested by the most complex system we are aware of at the time. Once nature was described as a body, then as a clock in the age of clocks, then a machine in the industrial age. Now, in the "digital age," we apply the computational metaphor. To explain how our minds work, or how evolution advances, we apply the pattern of a very large software program processing bits of information. (*What Technology Wants,* Penguin Group, 2010, pp. 63f.)

I would like to suggest that the leading metaphor for a deterministic world is the clock, that can give you tremendous accuracy, whereas the leading metaphor for Deacon's non-deterministic world of constraints should be a swamp. While not deterministic, all is not utterly random in a swamp. Plants recur true to their species type. Acorns will never produce dandelions; nor will the DNA of dandelions ever produce oak trees. But unlike the regular consistency of a Swiss watch, every dandelion is just a little bit different from every other dandelion. As someone put it, "There will never be a Newton for a blade of grass."

Now I would like to exhibit some examples of emergent systems that don't commit the phallusy of misplaced physics: money and love.

Money, markets, and the meaning of wealth

Like life, evolution, and consciousness, money is something so familiar that we sometimes forget just how peculiar it is. When the stock market suffers a sharp selloff and commentators say things like, "500 billion dollars of wealth evaporated last week!" it's not unreasonable to ask, "Where did the money go?" Since normal everyday exchanges in the marketplace exhibit something close to a constant conservation of cash—what one person pays, another receives, and vice versa—how can money simply disappear? Where does it go? And if money *can* simply disappear, can it be created *ex nihilo?*

The new economy of advanced informational capitalism is built on top of an old economy that includes raw materials, commodities and real estate. New wealth presupposes old just as surely as love presupposes sexual desire or consciousness presupposes perception. The new economy is no more free of physical reality than mind is free of matter. But an adequate understanding of the new economy—with its currency swaps, options, hedge funds and derivatives—can no more be reduced to purely physical

fundamentals (like gold ingots simply located) than consciousness can be reduced to neurophysiology. To reduce wealth to piles of gold ingots would be to commit the phallusy of misplaced physics.

The economy is an emergent system. It works when it all comes together—not just the fundamentals, but the whole complex adaptive system that includes everything from credit and interest rates to liquidity levels, rates of technological innovation, consumer confidence, trade agreements and well regulated markets. No single cause drives the economy. Monological attempts to reduce economic performance to dependence on any single variable will fail. Monetarists, gold bugs, supply siders—all fall prey to the phallusy of misplaced physics.

To gain a better sense of *money* as an emergent system, no better source is available than James Buchan's remarkable book, *Frozen Desire: the Meaning of Money.* Buchan's book on money will serve for this section much as Terry Deacon's did for Language: not simply as a source of good thinking, but as an illustration of the dangerously huge range of references necessary to capture the full richness and complexity of emergent systems.

A taste for emergent systems is the mark of a promiscuous mind. Drawn like a moth toward immolation in the flame of complexity, a mind that embarks on the task of understanding life, or language, or consciousness, is almost sure to wreck itself in confusion. And any book that claims to address all three is bound to burst its seams with references to a range of fields that no single reader can follow, much less any single writer could master. An incapacity to limit the discussion to any single discipline, an incontinence that cannot help but spew forth new ideas, these are the marks of minds that are drawn toward emergent systems.

It will be said of this book, as it has been said of books like James Buchan's *Frozen Desire,* that it wanders. But that is *precisely the point*: where emergent systems are concerned, any attempt to reduce their complexity to the comforting formulas of monological science is bound to betray the very richness and complexity of the subjects at hand.

Take money. Buchan's remarkable book devotes surprisingly few pages to the discipline of economics where you would think its discourse belongs. Instead, he gives us history, philosophy, autobiography, literary criticism—an impossible to classify mixture of genres. Where is a poor bookseller to locate it in her store?

In order to plumb the mysteries of money, Buchan takes us on a tour of the novel from Cervantes' *Don Quixote,* to Tom Wolfe's tour down Wall Street in *Bonfire of the Vanities.* Buchan's prodigal talents take us from Jane Austen to Henry James; from Dostoyevsky's Raskolnikov, wracked with guilt over slaying an old lady moneylender in

Crime and Punishment, to Sherman McCoy's incapacity to explain to his young son just what Daddy does. Oh, the elusiveness of money!

Sherman and his wife and son try to fathom the nature of money:

> 'Papa! what's money?'
>
> and then, unsatisfied with gold, silver and copper, guineas, shillings and half-pence, and his father not trusting himself to explain circulating medium, currency, depreciation, paper, bullion and rates of exchange, the boy asks again,
>
> 'I mean what's money after all?'
>
> It is a dangerous model, because the counterposition of money and childhood, artifice and innocence, evokes powerful emotions that may needs be diluted on the page. But there is no sentimentality in Wolfe's version, which unfolds at a beach club on Long Island.
>
> 'Daddy. . . . What are bonds? What is deal?'
>
> Now his mother began laughing. 'You've got to do better than that, Sherman!'
>
> 'Well, honey, bonds are – a bond is – well, let me see, what's the best way to explain it to you . . .'
>
> 'You build roads and hospitals, Daddy? That's what you do.'
>
> 'No, I don't actually build them, sweetheart. I handle the bonds, and the bonds are what makes it possible---'
>
> 'You *help* build them?'
>
> 'Well, in a way.'
>
> 'Which ones?'
>
> . . . 'Well, not any one specifically.'
>
> 'The road to Maine?'
>
> The Master of the Universe ties himself in knots. For even poor Sherman has grasped the imperative, which is categorical enough to be worth repeating once a century [and note the graceful reference by resonance to ponderous Kant's "categorical imperative"]: that even in great cities of finance, if you *cannot* explain your job to your child, you probably *shouldn't* be doing it. (Buchan, p. 169)

And so, from there, we're drawn off from the elusiveness of money into the complexities of morality.

But Buchan, to his everlasting credit, never shirks his task of appreciating the true complexity of his subject, so instead of trying to *nail down some precise definition of money,* he treats us to a feast of erudition and intellectual delight. He swoops through biblical scholarship on the famous thirty pieces of silver for which the body of Christ would be bartered. He devotes an admiring chapter to John Law, the Scot who murdered a man in his twenties, escaped from prison, and went on to create a financial empire in 18th Century France that rivaled Buchan's later portrait of Michael Milken's equally fragile edifice in 20th Century America. He paints a heart wrenching portrait of poor Karl and Jenny Marx, huddled in their two rooms in London while one of their daughters is dying of poverty. See Chapter Eight: "Death in Dean St." which, by the by, reveals Buchan's mastery of the monstrous and Talmudic literature known as Marxism. Is this, as claimed earlier, a "promiscuous mind" or what?

As if biblical studies, the history of the novel, and the sweep of Marxism were not enough, he devotes Chapter 7, "Coined Liberty: His and Hers," to a discourse on gender studies whose density and grace are exemplified in passages like:

> The reticence of romantic love is three-fold. Money conveys desire: it cares not at all where that desire finds satisfaction. To give money to a woman—and here I must speak as a man—is to deny her special quality, her *irreplaceability,* and reduce her unique amiability to a commodity. Money takes away her name, while transforming her lover into a nameless customer of a market of appetites. No person likes to be thought interchangeable in love. Finally, a gift of money allows the beloved to deploy wishes in every direction, whereas her lover wants to concentrate them on himself. The English satirists of female vanity of the early eighteenth century cannot conceal a certain trepidation at the new opportunities money offered women: Addison, in *The Spectator* of February 7, 1712, and only half in joke, warned husbands against pin-money—a wife's allowance of money—'as furnishing her with Arms against himself, and in a Manner becoming accessory to his own Dishonour.' Money therefore meets a triple resistance. (The same arguments apply to women's gifts to men and, *mutatis mutandis,* to presents of money to children at Christmas and on their birthdays.) (268 ff.)

In addition to erudition and intellectual delight, we get wisdom and plain good sense about everyday family life.

Buchan's profligate, promiscuous, incontinent mind spews forth at least a hundred tightly wound paragraphs like the one just quoted, bristling with cosmopolitanism contrasted with the vernacular, and so resonant you want to reread them aloud, like the following:

> The System, as Law called his invention, may be tackled at different levels of difficulty, like a video game. At its most elementary, it was a breathtaking device to convert the debts of France's bloody and capricious past into charges on a brilliant future, secured on the simplest fiscal regime ever devised and the potential of what is now the United States. At a somewhat more demanding level, it sought to create the most powerful nations on earth—Scotland, England, Savoy, France, Denmark, Russia, depending on the client—without strain and above all without violence. . . At a level more ambitious still, it sought, through the institution of a great commercial trust in which every citizen would risk his savings and enjoy his reward under a disciplined monarchy, to enfranchise society. And yet further, between the lines of Law's writings in English and French, for all their lucidity, one senses something else, wide and slow and undamnable as the Mississippi itself: a river not so much of money as of happiness, of which we saw but the springs in Saikaku and Defoe. The activities and artifacts of the Regency are drenched in happiness: Manon and Des Grieux are reunited on the Mississippi, Watteau's lovers are preserved for ever embarking for their inflationary Cythera, the Italian Comedy returns from its twenty-year exile, *gallant défunts* smile in secret recollection from Rosalba Carriera's pastels. The moment passes, the river flows on, past the graves of martyred Jesuits and *sauvages nobles* thigh-deep in the wild rice, the gaunt financial ruins of Turgot and Necker, and the corpses of kings and revolutionaries and prostitutes, till it issues at last at the famous city that Law helped found and that still carries his air of risk and magic and incorrigible delight, the city of New Orleans. (269f.)

Believe it or not, there are *many* paragraphs like this that sing and dance and close with cadences as ominous as those in Verdi's *Requiem.*

And at the close of Buchan's book there is a short chapter, a reprise, a kind of cadenza in which he summarizes, in relatively simple sentences, all that we have learned on his whirlwind tour of disciplines and genres so various.

> We have learned that money, far from existing for all eternity like Melchizedek in the Bible, has a history. For from obscure beginnings, money has spread out to colonize the world, both in its forms as coin or banknote or book entry and as a notion of happiness penetrating the minds of men and women. Money was probably not invented in a particular place and at a particular era, but came into manifold being, for manifold purposes. Money permitted human beings to expand not only their possessions but their wishes beyond limits held ultimate by predecessors. By now, money has become a system, which we understood, by way of the simile of the railway shunting yard, as gathering the wishes of the most estranged and scattered populations and dispatching them to unimagined destinations. We also saw that money became indifferent as to its physical form, and whether human beings revert to gold or cattle as money of account and payment, or pass into a realm of pure electricity, is a matter of indifference also to us. (Buchan,pp. 268ff.)

Now, what is so extraordinary about this summary paragraph is the way it bangs out five or six of the eight traits of emergent systems in a succession so emphatic that, on the one hand, it proves the point that money *is* indeed an emergent system, and on the other hand ratifies the value of the list of traits as an analytic tool. The relevance of the remaining traits follows just a few pages later.

First Trait: The impossibility of first instances

There could not have been a first word, for what made a barely articulate grunt into a word was the presence of other words constituting a rudimentary language. Likewise, there could not have been a *first coin.* There had to be *many coins* before any one of them could be counted *as* a coin.

Buchan captures this paradoxical lack of a first instance of money in his references to money's "obscure beginnings," and in his statement that, "Money was probably not invented in a particular place and at a particular era." Searching for the first invention of money is as foolish as searching for the first invention of language, or the first instance of life. As an emergent system, money simply isn't the sort of thing, like the automobile or the computer, about which it makes sense to seek the *first.*

Second Trait: Emergent systems pop.

Because they come together all of a piece, emergent systems come about all of a sudden. In retrospect, their emergence has the appearance of discontinuity, however gradual and continuous the processes leading up to emergence may have been. As Buchan describes it in the above paragraph, money "came into manifold being, for manifold purposes." One year it wasn't there, the next year it was—time and time again, in different cultures, different forms, different coinages, different geographies—money *popped.* It wasn't *invented.* It *emerged.*

Third Trait: Holism—The Whole influences the nature of the part.

What was purportedly the first word could not *be* a word if it lacked the context of a language. A language had to be already there for a sound to be a word. So likewise with money. There has to be a fair amount of it for there to be any of it. As a means of trade, it is intrincsically relational, and very much a matter of beliefs. As Buchan says in the paragraph above, "By now, money has become a *system*."

Fourth Trait: Emergent Systems are self-reflexive.

So likewise, yet again, with money. Now that we are off the fundamentalism of the gold standard, and have realized the relativity of the value of one currency against the value of other currencies in what we call "the float," the self-referential valuation of money by money is evident. I see this recognition in Buchan's "simile of the railway shunting

yard." All railway systems must form closed loops, else a train sent away could never get home again.

Fifth Trait: An emergent system is unpredictable from the features of its constituent parts.

Sixth Trait: Emergent systems are irreducible to configurations of their component parts.

Fact is, the trait of unpredictability is not touched on in the dense paragraph quoted above, perhaps because its relevance is perhaps the most obvious of all. If prices are predictable, then it's not a market you're looking at, but as Buchan elsewhere describes it, "a rigged system" redolent of Soviet communism.

And as for irreducibility, just as Professor of linguistics, Ferdinand de Saussure, introduced us to the notion of the "arbitrariness of the sign"—that you can't reduce the meaning of a word to anything about its physical shape or sound (with the rare exception of onamatopoietic words like the 'bow-wow' of a dog)—so Buchan writes in the paragraph above: "We also saw that money became indifferent to its physical form, and whether human beings revert to gold or cattle as money of account and payment, or pass into a realm of pure electricity, is a matter of indifference also to us."

Seventh Trait: The principle of Desire

Is the principle of desire relevant to money? Consider Buchan's title, *Frozen Desire: the Meaning of Money.* As reflected in the above pregnant paragraph: "Money permitted human beings to expand not only their possessions but their wishes beyond limits held ultimate by predecessors . . . gathering the wishes of the most estranged and scattered populations and dispatching them to unimagined destinations."

Eighth Trait: Coming Apart

Emergent systems can de-cohere. Love can end in divorce. Consciousness falls asleep. And wealth evaporates when a bull market goes bust. As Buchan puts it just two pages after the pregnant paragraph, "when states disintegrated their moneys vanished as completely as their laws."

QED: Money is an emergent system, and as Buchan's remarkable account shows so well, the eight traits of emergent systems fit money as a Saville Row suit fits the man.

But to return to the title of this chapter, "The Phallusy of Misplaced Physics," what we have seen is that wealth is not to be reduced to some physical, deterministic entity. Money and wealth emerge; they do not follow causal laws.

Likewise love:

Love

First Trait—Love comes in twos

At the level of the individual—or individuals we must say in the case of each of two individuals who must *be in love with one another* if love is to be requited and therefore truly love: Love comes in twos, not single units. Love is not like a shirt—something to be acquired and worn by one person alone. More like gloves, loves come in pairs. As such, it is often hard to say which came first—who fell in love with whom first. (Repetition *again. Bang!!!)*

Sometimes it is obvious, as when one but not the other is able to say that it was "love at first sight," an expression that is *so* familiar that we can site it as a case of an "exception that proves the rule." For most of the time love is something that emerges when two people achieve an intimacy not available to other human beings. Such an intimacy is by its very nature dual, very like a cell-almost-cells, well along in the process of mitosis when two nuclei have amost formed and the cell's outer wall has begun to pinch toward the profile of a peanut. Love is like that, and it's silly to say one end of a peanut is first, the other second. Love is dual.

Second Trait: Love pops

The phenomenology is obvious: The experience of "falling in love." Love hits you "like a ton of bricks." Sure, there are cases where friendship grows until it gradually turns into love . . . but when love finally emerges, it is likely to do so with a rush of feeling that marks a discontinuity from what has gone before.

Third Trait: Holism

You can't have the part without the whole. You can't have one person in love without the other. If you do, as all great stories of unrequited or lost love will tell you, that one person in love will suffer a broken heart. You can't sustain it alone. Love takes two. It is twoish—dual in essence, a whole of which each person's love is a part. Yes, there is particularity: two nuclei, not one. But the intimacy two lovers share is one intimacy, not two. And as such it is difficult if not impossible to say who entered that intimacy first. It was created by both lovers when they came together, when—suddenly or gradually—

they both found themselves face-to-face with one another without flinching. And next thing you know comes the heart-to-heart.

Love is self-reflexive.

Love talks about itself. It declares itself. It sings aloud. It writes poetry. Love is so articulate as to be positively effusive.

Love is self-referential. It feeds on itself, not like an ulcer but like a positive feedback loop that amplifies itself by its own achievement of overtones in resonance with one another.

Love is unpredictable.

You never know who is going to fall in love with whom. That is why arranged marriages are not a good idea. You cannot predict who will fall in love with whom on the basis of any component of personality or character because . . .

Love is not reducible to any constellation of such components of personality or character. Almost anyone can fall in love and, just as remarkably, anyone can be loved. And while they are not loved *for* any particular component of their personality or character, they are surely loved in and through each and every feature of their personalities. Love is unconditional. You don't love this and this about someone while hating that and that. You don't love someone *unless* . . . This is not to say that what has come together cannot come apart. All too easily it can. And all too often it does. Don't just look at the divorce rate. Look also at the untold number of dead marriages.

Love lives on desire, and all too often desire fades, or drifts off toward others. The cell wall ruptures. Intimacy dies that two may live as separately as two cells after the process of mitosis is complete.

Just as consciousness goes to sleep and then wakes again, so too does love wax and wane even when two people are *in love* over a long period of time. Part of the rhythm has to do with the ebb and flow of libido, a slippery concept if ever there was one, but not altogether without usefulness. Perhaps it's best to talk more simply about desire.

To show how love lives on desire, and how desire figures into the emergent system that is love, consider the hierarchy of desire from the simplest level of animal hunger up to the higher levels of care and love. Animal hunger is not love. Nor are the first stirrings of sexual desire in an adolescent, as strong as they may be. What differentiates love from sexual desire? And how can we answer this question in a way that does not differentiate the two so thoroughly that sexual desire gets expurgated from some too lofty, too romantic understanding of love? The title of this section, once again, is "Love feeds on

desire." The point is to show how the emergent system called love exhibits the trait of desire. And yet it is *something more* than desire.

Posing the question in this way is liable to elicit all of the old saws about love as "loftier" than desire; love as more about giving than taking; love as *agape* rather than *eros.* This edification of love over desire runs the danger of bowdlerizing love into something saintly and non-bodily, and very soon we see a love that flees from rather than lives on desire.

Philosopher Martha Nussbaum does not make this mistake. In her magnificent book, *Upheavals of Thought: The Intelligence of Emotions (*Cambridge University Press, 2001*),* she devotes hundreds of pages to the various approaches to the "ascents of love." In a series of chapters on Plato, Augustine, Spinoza, Dante, Emily Bronte, Gustav Mahler, Walt Whitman, Proust and Joyce, she shows how each adds something new to our understanding and therefore our experience of love. Earlier in the book she dismisses what philosopher Robert Solomon has already refuted as "the myth of the emotions," namely, that they are the very opposite of intelligence: brute forces that sweep over us and do battle with the better lights of reason and intelligence. Building on Solomon's earlier work on the emotions as judgments, she shows how emotions act as *evaluative judgments.*

In the case of love especially, there's a lot of detritus to be cleared away about love's blindness, its passionate upheaval of reason's cooler perspective. Before following Nussbaum up the ascents of love, it's worth pausing to note just how radical the revolution that is currently underway in the re-appropriation of love and the emotions as legitimate topics for academic discussion. Quite independent of its contents, the very fact that Martha Nussbaum, who has served as President of the American Philosophical Association, could publish a 700 page book on the subject marks a significant triumph for a movement that, as recently as 1990, was a largely neglected undercurrent in philosophy. In an essay first published in 1992, Solomon pointed out that, "Insofar as emotions provide us with a topic in philosophy, it has been understood that their place is tangential, their analysis a side issue." (Solomon, "Beyond Reason: The Importance of Emotion in Philosophy," in *Revisioning Philosophy,* ed. James Ogilvy, SUNY Press, Albany, p. 26.) And further, "The introduction of emotion in philosophy is not just the rediscovery of a neglected topic. It is also the recognition that something has gone dreadfully wrong with the narrow way that we define and do philosophy." (Solomon, p.47)

In addition to the pioneering work of Robert Solomon, and the triumphant work of Martha Nussbaum, see Keith Oatley, *Best Laid Schemes: The Psychology of Emotions,* Cambridge University Press, 1992, and Richard Wollheim, *On the Emotions,* Yale University Press, New Haven, 1999, together with many other recent books listed in Nussbaum's bibliography.

The revisioning of philosophy now afoot (see *The Revisioning of Philosophy,* ed. Ogilvy, State University of New York Press, 1992) amounts to a fundamental recasting of the relationships between reason and passion, mind and body, and the sense of what it is to be a good person. There is a tradition going all the way back to Plato that sees the emotions in general and love in particular as some kind of disease of reason. This tradition, which has surely dominated most of 20th Century Anglo-American philosophy, pits the emotions against reason as defender of man's higher aspirations. The practice of philosophy, like science, then becomes a dispassionate pursuit of pure rationality untainted by the distractions of mere feelings. The practice of philosophy in this tradition leads toward a philosophy of mind that is inevitably and badly distorted toward a pure cognitivism.

Only from the perspective of such a philosophy of mind, preoccupied as it is with visual perception, representation and belief, could the claims of "strong AI" be entertained as remotely plausible. Only from this perspective could one believe that computers—artificial intelligence—could pass the Turing test, that is, prove indistinguishable from human intelligence. The problem with the so-called "artificial intelligence" movement has less to do with what computers can or cannot do. The problem with "strong AI" is a fundamental misunderstanding about the nature of *human* intelligence, a misunderstanding that is only now being overcome by a new interest in emotions not as enemies of reason but as located at the very heart of a truly human rationality. While following what might appear to be "merely" literary lights up the ascents of love, it will be important to keep in mind that the emotions we are tracking are not, as most of 20th Century philosophy would have it, some alien and confusing adversaries to reason's higher calling.

As the Copernican revolution that will put desire at the heart of consciousness will show, love is at the very heart of reason. Where philosophers have been willing to quote Pascal in granting, if condescendingly, that "the heart has its reasons," we are only now coming to the more revolutionary claim that reason has a heart; that desire, and its ascent toward love, is constitutive of *Dasein,* a truly *human* rationality.

But what kind of love? And wither desire? With these preliminaries on the radicality of this enterprise in place, let's follow Nussbaum's course up the ascents of love. She begins with a comment on the elusiveness of love that, by now, should alert us to the fact that, once again, we are looking in the right place for an emergent system: "Precisely *because* love is more mysterious than the other passions, precisely because we cannot easily catalogue the reasons for our loves, we look to narratives for the understanding we lack, or at least for a confirmation of our sense that there is a great mystery here. (Martha Nussbaum, *Upheavals of Thought: The Intelligence of Emotions,* Cambridge University Press, 2001, p.68) She then notes that, "What we find emerging . . . in consequence of this perceived tension between love's energy for good and its subversive power, is a

recurrent attempt to reform or educate erotic love, so as to keep its creative force while purifying it of ambivalence and excess, and making it more friendly to general social aims."

The education of love is thus her recurrent theme as she follows the literary and philosophical tradition from the ancients to the moderns; and her recurrent question is whether the student, erotic love, does not become over-educated—so elevated beyond so-called "lower" desire that a higher love loses the juice that first motivated the quest. Among her clearest statements of the overall quest occurs at its very outset:

> I shall focus on three distinct types of ascent story that form their own continuous traditions within that larger tradition: an account of the ascent that focuses on contemplation of the good and beautiful; a Christian account of the ascent that investigates the role of humility, longing, and grace; and a Romantic account that rejects a static telos for ascent, holding that striving itself is love's transcendence.

The guides on the first of Nussbaum's three paths—contemplative creativity—are Plato, Spinoza and Proust. They teach liberation from "the bondage of the passions," as Spinoza would put it. "For a man at the mercy of his emotions is not his own master but is subject to fortune." (Spinoza, *Ethics,* Praface, IV; quoted by Nussbaum, p. 502.) Spinoza teaches that understanding brings freedom from this bondage. But the kind of understanding that is sought on the contemplative path from Plato through Spinoza to Proust is an almost "god-like omnipotence" that "closes off all possible sources of pain and uncontrol." (Nussbaum, 525f.)

On the second of the three paths, The Christian Ascent, Augustine and Dante acknowledge the sin of pride in this aspiration to omnipotence. Nussbaum shows how both "set themselves the task of rewriting and correcting the pagan ascent of love. For each there are deep psychological links between earthly and heavenly love; for each it is important to argue that the good Christian life is more volatile and erotic than the Platonic tradition has wished love to be." (592f.) But the Christian Ascent ascends too far. Heaven is not earth. The telos of heavenly love ultimately denies the sanctity of earthly particularity and sweeps all passion too swiftly upwards toward a redemption that expunges more than it redeems.

For an appreciation of particularity, Nussbaum turns to the third path, the Romantic Ascent as exemplified by Emily Bronte and Gustav Mahler.

> In Romantic conceptions of love's ascent, striving itself, and the peculiarly human movements of embodied erotic effort, become an ascent and an end in themselves, in no need of redemption by a static and extratemporal telos. . . Romantic love will claim that it uncovers deep sources of spiritual richness and personal authenticity without which any morality of human concern is dead. But the question must be what happens next—whether love can find a way back to

> compassion, or whether its absorption in the particular is so deep that it must simply depart from the world. (681)

In order to reclaim contact with the world and society, Nussbaum turns to Walt Whitman for a more inclusive love, a "democratic desire."

After traversing all three paths of ascent, Nussbaum concludes that all three paths share a common weakness: "All repudiate daily life. The very metaphor of ascent suggests to us that there is something low about where we usually live and are." (688f.) So to conclude her exegesis on love, Nussbaum turns to James Joyce for a respectful descent into the quotidian details of everyday life. In Joyce's *Ulysses,* which follows its protagonist, Bloom, through the course of just one single day, Nussbaum finally finds a transfiguration of everyday life.

Looking back on the three ascents, she observes, "Nobody has a menstrual period in Plato. Nobody excretes in Spinoza. Nobody masturbates in Proust (though in a certain sense also, nobody does anything else). Augustine and Dante record such moments, but leave them behind in Hell." (688f.) Nor do the romantics show the same respect for the everyday that Nussbaum finds in Joyce.

> The astonishing concreteness of Bloom's day, of the fragmentary and complexly interwoven texture of his musings, in which the past jostles against the present and actions against memories, compels assent. "Is that Boylan well off? He has money . . . He tore away half the prize story sharply and wiped himself with it. Then he girded up his trousers, braced and buttoned himself. He pulled back the jerky shaky door of the jakes and came forth from the gloom into the air . . . Quarter two. There again: the overtone following through the air, the third. Poor Dignam!" As Bloom says, "Life might be so." Grief for a friend follows a trip to the outhouse, and both, as present moments, are wound round by snaking strands of memory and fantasy and expectation, which crawl through the mind's day, leaving no moment single, no love exclusive, no logical deduction uncolored by wish and regret. In these ways the text says, here, here in this confusion is the really whole cosmos (or noncosmos), here and not in those ordered clarified probabilified well-plotted texts in which we are accustomed to look for our lives. Even the reader to whom a focus on consciousness is a familiar novelistic device—the reader, say, of Henry James, or of Proust—would still be arrested by the surprising multiplicity and daily disorderliness of consciousness in this world. (688)

The issue, finally, is the irreducibility of contingency. At essence, Joyce's *Ulysses* is a celebration of the inessential. In Bloom's love for Molly and Molly's love for Bloom "it all comes together," in an irreducible congeries of everyday detail.

Love emerges non-monologically. Joyce draws from the Christian, the contemplative and the romantic traditions, but none supervenes over the others. This single day that fills the

pages of *Ulysses* is not the first instance of love (Trait 1), nor the last. It is one more day in everyday life. Love pops with a single kiss on Mollie's ass (Trait 2). The whole influences the nature of each part as Joyce draws on the whole history of life and literature to tell his tale (Trait 3), and through Joyce's well wrought words, love speaks itself with loquacious, self-referential erudition (Trait 4). Given the richness of contingencies that fill the novel's pages, it is impossible to predict what will happen next (Trait 5). And the love that is represented is irreducible to any essence from any of the several traditions reviewed by Nussbaum or constantly drawn on by Joyce (Trait 6). Desire—yearning, craving, ravenous desire—remains un-sublimated, undignified by any monological essence that would steal its heat (Trait 7). And of course love can die. It lacks the permanence of a rock.(Trait 8) Love *emerges* in *Ulysses.*

Joyce's transgression of propriety and the legal battles over the novel's status as great literature or pornography are not to be dismissed as incidental to the novel's greatness.

> The novel's sexual explicitness and its insistent sexual focus can now be seen to have political significance. For, first of all, they are a linchpin of the project of restoring the reader to acceptance and love of the body, with all of its surprises, with precisely that disobedient ungovernable character that leads Augustine to find there our original equality in evil. Such love, the novel suggests, with Whitman, is necessary if we are to take the body's needs as seriously as a compassionate politics requires. And a focus on the body's universal needs is an essential step on the way to the repudiation of localism, therefore of ethnic hatred. Second, by showing Molly Bloom as the one character in the novel who never entertains thoughts of revenge, by showing how Bloom's own impulse to revenge is cut short by his arousal as he kisses Molly's bottom, the novel suggests, again with Whitman, that the root of hatred is not erotic need, as much of the ascent tradition repeatedly argues. It is, rather, the refusal to accept erotic neediness and unpredictability as a fact of human life. Saying yes to sexuality is saying yes to all in life that defies control—to passivity and surprise, to being one part of a very chancy world. (709)

In love we face the fold. Nothing is guaranteed, not fidelity, not eternity, not undying reciprocity. And for precisely that reason, vows of permanence are all the more honorable, and necessary. Love emerges, again and again, driven in part by libido's heat and not necessarily transcending desire's ungovernable passion.

Because we face the fold, and nothing is guaranteed, sentimental love songs do not contain the answers to life's big questions. It is not the case that "All you need is love." But without love, life is less than complete.

The education of desire from its inchoate beginnings, up through the several monological ascents, and down through the non-monological descent into contingency is well represented by Joyce's *Ulysses* and millions of real-life romances. Love must speak its

own name. And the sentence, “I love you,” is not to be reduced to a description of my physiology. To use J. L. Austin’s categories, “I love you” is not a declarative proposition; it is *performative.* You are not just describing your condition; you are not just announcing the firing of certain synapses; you are *doing something* when you say, “I love you.” Through the use of language, you are creating an intimacy that could not exist without language.

One of the things you are doing when you say “I love you” is *recognizing* the other. And you are asking for recognition from the other. When love speaks its name it is bringing into being the kind of mutual recognition that is, according to Hegel, the soup from which self-consciousness emerges. Recall the long and elaborate history recounted, the section titled *Hegel’s Depiction of Desire.* Recall that it concluded with a revision of Descartes’ famous dictum, *Cogito ergo sum,* to read instead, *Amo ergo sum.* I love therefore I am.

But what *is* love? After seeing how love exhibits the eight traits of emergent systems, we are in a better position to see how the emergence of love *comes together* with the emergence of self-consciousness. You cannot have one without the other. Inchoate desire, inarticulate desire, is not yet love. It doesn’t know what it wants. Only with the addition of mind, of consciousness, can desire know what it wants and therefore climb the hierarchy of desire to become love.

But only with the assistance of love can there *be* a climber—this is the additional part of the story that comes to light with an appreciation for emergent systems. This is the part of the story that was missed by Sartre and Lacan in their appropriations of Hegel’s dialectic in their own depictions of erotic relations with others. Because they assume that the climber is a naked Cogito, a pure cogitator devoid of a limbic system or senses of taste or touch or smell or erotic arousal, they have real trouble establishing a relationship with *the other.* The other is *very other* because they are not in a multimedia relationship with that other.

Oddly enough (because they should have known better) their analyses of relations with others suffers from the same sort of illicit assumption of pre-established personhood that they and their peers criticize in social contract theories of political philosophy. These heirs of Hegel and Marx are very quick to ridicule the idea that a bunch of Robinson Crusoes can come together to agree on a social contract. The French heirs of Hegel and Marx are quick to point out that mankind came on the scene as a herd and only later did possessive individuals emerge from the primordial social condition. But when it comes to analyzing sexual contracts (not social contracts) these heirs of Hegel and Marx somehow forget their criticisms of Locke and Rousseau. Descartes’ Cogito stands in for Robinson Crusoe as the signer of a sexual contract, a contract that veers toward the sadistic or masochistic precisely because its signatories are presumed to be geeks—computer programmers unconnected to their hearts or their bodies.

This section is titled, “Love feeds on desire.” Just to recall the larger context, we’re coming to the close of a section where we’re seeing how the seventh trait of emergent systems, The Principle of Desire, applies to the emergence of love. We’ve seen how love is twoish, the whole rendering neither lover first (Traits 1 and 2); we’ve recalled how love pops (Trait 3); we’ve noted love’s loquaciousness, it’s self-reflexive need to speak its own name (Trait 4); we’ve acknowledged that love is neither predictable (Trait 5) nor reducible to its components (Trait 6); granting that love cannot be *reduced* to nothing but circuitous and often pathological redirections of desire, still it is important to follow Martha Nussbaum through the “ascents of love” to her final insight that love cannot *transcend* desire. As this section is titled, “Love feeds on desire.”(Trait 7) And love can suddenly end. (Trait 8)

Now, how does this exposition of the eight traits in their application to love relate back to Hegel’s depiction of desire, which concluded with the line, *Amo ergo sum*?

Amo ergo sum. The ontological power of love is a condition for the possibility of solving (by dissolving) the ridiculous epistemological puzzle known as “the problem of other minds.” The tradition from Descartes to Sartre got trapped behind the “reef of solipsism” only because it conceived the subject as geek—pure *Cogito* without heart or body or limbic brain. If instead we see the subject as both lover and beloved, then we will not end up concluding with Sartre that “Love is an impossible project.” Impossible for a Cartesian geek, perhaps, but not for a consciousness that was cradled in its mother’s arms before it knew how to talk or think. Both onto-genetically and ontologically *we start with love* and the cogitation comes later. If you try to start with nothing but pure cogitation, no wonder you will have difficulty getting to love.

Granted, the love that a baby experiences for its mother is not yet articulate or self-reflexive. It does not yet speak its own name, it gurgles and burps. But it does smile. There is a connection between mother and baby. There is no reef of solipsism. The limbic connection between mother and child is an essential prerequisite to the progressive development of consciousness and self-consciousness. But it does not yet amount to the elaborate dialectic of mutual recognition and mutual respect that will characterize romance.

What is love? By now the question has taken on the eye-rolling richness that accrues to its equally brief cousin, *What is life?* Disabused of the monological penchant for locating a single crucial component—an *élan amour* in place of *élan vital*—we can now accept that, as an emergent phenomenon, love is first of all *complex.* It is not *just* desire; yet it cannot live without desire. The dialectics of mutual recognition tell us that the lover must know and respect the beloved. As opposed to inchoate desire, love must know what it wants; it must *know* the beloved. And simple common sense—not a sophisticated psychology—tells us that this is no small order. How easy it is to misconstrue the beloved

in a way that appears to serve one's own needs and desires better. How easy it is to slip into the kind of sado-masochistic dynamics that so easily follow and have in fact been derived from the dialectics of mutual recognition. But that slippage is no more inevitable than the plunge from the hermeneutics of belief to the hermeneutics of suspicion.

The possibility of the plunge is ever-present. We face the fold. And still the high plateau has been seen and occupied from time to time. Love has emerged. The romantic sublime *has been experienced*, and only the hermeneutics of suspicion will insist that it was a delusion.

Before leaving this section on love, we need to climb yet two further rungs on the hierarchy of desire: First, love of the earth and our natural environment; second, a Whitmanesque "democratic desire" that will foster a mutual recognition among cultures.

Love at large

This section on love began the statement: "No first instance applies in two senses, first on the level of the individual, and second, the level of the culture." Everything said about love so far refers to love among individuals. Now it is time to fetch a wider compass: Whitman's "democratic desire" at the cultural level.

With the possible exception of Rachel Carson's *Silent Spring (1963),* no other book has been more important to promoting the environmental movement than *Limits to Growth (1972)* by Donella Meadows, Dennis Meadows and Jorgen Randers. Twenty years later the same authors published a book called *Beyond the Limits: Confronting Global Collapse, Envisioning a Sustainable Future.* In this second book they track the course of their earlier projections, and lay out the trajectories of several different scenarios for our collective future. Their first book was hailed by many, but others condemned it as a doom and gloom forecast giving no credit to human creativity. "Don't worry. Be happy," said some industrialists. "We human beings can always count on technology and creativity to get us out of pretty much any jam."

In the years since *Limits to Growth,* human creativity and ingenuity have come into play. The alarm sounded by its authors *has* had an effect, much as they hoped it would. Theirs was intended to be a self-disconfirming prophecy. They *wanted* to be wrong. They *hoped* that their worst case scenarios would not come to pass. This logic of alarm comes through loud and clear in their second book, which holds out more clearly than the first the vision of a sustainable future. But just as clearly as ever, they show us the contours of a fold. They paint a picture of global environmental collapse if we exceed the carrying capacity of our natural environment.

This is not the place to review the science or track the trends in resource depletion and pollution that threaten our future. This *is* the place to quote at some length a remarkable passage that appears just pages from the conclusion to their later book:

One is not allowed in the modern culture to speak about love, except in the most romantic and trivial sense of the word. Anyone who calls upon the capacity of people to practice brotherly and sisterly love is more likely to be ridiculed than to be taken seriously. The deepest difference between optimists and pessimists is their position in the debate about whether human beings are able to operate collectively from a basis of love. In a society that systematically develops in people their individualism, their competitiveness, and their cynicism, the pessimists are in the vast majority.

> That pessimism is the single greatest problem in the current social system, we think, and the deepest cause of unsustainability. A culture that cannot believe in, discuss, and develop the best human qualities is one that suffers from a tragic distortion of information .
>
> The sustainability revolution will have to be, above all, a societal transformation that permits the best of human nature rather than the worst to be expressed and nurtured. (Donella Meadows, Dennis Meadows, Jorgen Randers, *Beyond the Limits,* Chelsea Green Publishing, White River Junction, Vermont, 1992, p. 233.)

This call for love in the context of a text on environmental sustainability is entirely consistent with an argument that has us climbing, historically and biologically, ever higher on a ladder of love, a hierarchy of desire from biota through libido to the love of conscious beings who respect and recognize one another. The structure of desire—at first omnivorous and indiscriminate, then educated to discrimination and the mutual recognition of another self-consciousness—can extend still further to family, community, country, or, as the Buddhists say, to all sentient beings.

This expansion of the radius of love is subject to limits, however. Love, too, has limits to growth. Just ask your spouse, or your neighbor, about your love for your neighbor's spouse. Because love creates intimacy—an *inside* distinct from its outside—there is an exclusivity to loving relationships that stands in contrast with attempts to expand the inclusivity of love. Yes, love thy neighbor; but, no, thou shalt not covet thy neighbor's wife. You dare not love *everybody* with the same love that you reserve for *somebody.*

What Nussbaum called "the particularity of love" cannot be ignored by well-meaning attempts to expand the radius of love. But who is the particular? Who is the lover? What is the radius of the "self"? The section on evolution concluded with reflections on a third reading of evolution, a narrow path between the brutalist story about "selfish genes" on the one hand, and the creationists' reliance on the divine self of a personal god who would design the cosmos to our liking. Between those two extremes, we traced a third

path: an interpretation of evolution that hinged on the role of desire in giving directionality to time, without relying on a predestined telos.

It takes desire to give time directionality, and once time has a direction—from a suffering or acceptable present toward a better future—then conscious life comes on the scene as the bearer of that desire toward a better future, toward a self-defined telos.

Evolution feeds on desire, and desire evolves. From its inchoate beginnings when desire doesn't quite know what it wants, through its sublimation and education, desire evolves toward love.

And so we embarked on this section's discourse on love. Now that we have followed the hierarchy of desire up Nussbaum's several "ascents of love," and seen the need for holding on to an abiding attachment to the everyday, the commonplace, where do we stand on the locus of care, the subject of *Sorge*?

Nussbaum's final descent into the text of *Ulysses* tells us that its eroticism has political significance: "Such love, the novel suggests, with Whitman, is necessary if we are to take the body's needs as seriously as a compassionate politics requires. And a focus on the body's universal needs is an essential step on the way to the repudiation of localism, therefore of ethnic hatred." How so? At first glance the argument would seem to be suggesting an expansion of the radius love beyond the local to encompass other ethnic groups. But the argument of previous sections—particularly the section on *Hegel's depiction of desire* and the section: *The principle of desire, and the reciprocity of purposes and desires in the course of evolution: the question of teleology and teleonomy*—suggest a different take on "the repudiation of localism, therefore of ethnic hatred." To the extent that the self is defined by the process of mutual recognition of and with an Other, then it is the lover, not just the beloved, that is non-local. *We* love, therefore we are.

On this higher rung of the ladder of love, ethnicities are preserved, along with all of the specifics of everyday life. The expansion of the radius of love is not a mystical quest for the white light that transcends all differentiation. But the subject of love, the lover, can be a limited collective. Not all, not one, but some—a family, a community in quest of a better future, even an entire ethnic group.

The dialectic of mutual recognition delivered to us by Hegel, extended to concrete relations with others but pessimistically misinterpreted by Sartre and Lacan, can be further extended to cultural relations. In the post-9/11 world of tensions between Christianity and Islam, it is essential that the "Clash of Civilizations" predicted by Samuel Huntington be averted by a further education of desire. We must care, not only for our natural environment, but also for our cultural environment: other civilizations. In place of dis-respect and humiliation, mutual respect and recognition can enhance a

cultural ecology that preserves differences and specificity. Perhaps love is too grand a word to use in this context. But perhaps not.

Trait 8: Divorce and Perversion

Love ebbs and flows. Like life, it lives and dies. Love evolves. Like life it cannot quite sit still. Like life, it defies demarcation, is forever over-reaching its boundaries and resisting simple definition. The spicy parts of adult romance are an indication of love's need for transgression. Even if you leave behind a Christian-Platonic revulsion for bodily desire as sinful and unclean and grant a healthy innocence to uncivilized desire, you may still want to allow for a certain quotient of naughtiness in adult love. Forget about original sin. We don't need it. But we may need a certain adventurousness in our love life, a capacity for pushing out the edges, for transgressing the norms presented in good Christian marriage manuals. The temptations of so-called perversions are not the residue of original sin but the necessary palette for erotic creativity. *What mutation is to evolution, perversion is to love.*

The language of love is rich. It includes hearts and flowers, poetry and love songs, symbols like wedding rings and tattoos, leather and lingerie. There's hardly an end to what can be eroticized given the will and the desire, so no wonder that the mind is an erogenous zone. Love is not just a social construction designed to channel the exigencies of physical desire. Love may feed on physical desire, but its metabolism is so elaborately symbolic in nature that the end product is ill-described as socially sanctioned urge. Love thinks and talks and sings. Love is loquacious, and thrives on language and symbolism.

And to return to the title of this chapter, "The Phallusy of Misplaced Physics," it should by now be obvious that love is not subject to laws of deterministic causality. It is far too complex for that. Love is emergent, and exhibits all eight traits of emergence.

Love and Consciousness

Consciousness, too, is twoish

Contrary to the tradition of possessive individualism that takes a solitary Robinson Crusoe as its model for the original human condition, or the isolated *Cogito* of Descartes, or the libertarian individual dating from John Locke, we now know better. Whether from empirical studies of

wolf children and abandoned orphans that receive too little human contact, or from theoretical accounts like Hegel's analysis of the struggle for recognition, we know that consciousness comes to be only where there is an interaction between one proto-consciousness and another. Perhaps that interaction takes the form of a life and death struggle issuing in a master/slave relationship; perhaps it is condemned to some version of sado-masochism or Lacanian despair; but just as likely if not more so, the struggle for recognition climbs the hierarchy of desire to the ascents of love and a greater, fuller consciousness emerges, whether individually or collectively.

Consciousness, we discover from reflections on love, is inevitably *mutual,* more of a membrane between selves than separated nuclei within each.

Falling in love and Waking to consciousness

Like falling in love, there is often a suddenness to the emergence of consciousness. Shazam! You're awake, all of you.

If love is the model, then so-called *enlightenment* may be the analogue for consciousness. The scales fall from your eyes. If not white light, then, according to many texts in the mystic tradition, some other form of *transformation* attends the kind of ascent characteristic of entry to a new level of consciousness. Perhaps the peace that passeth all understanding?

Whole and part

Obviously, and at risk of repetition, you can no more have half a consciousness than you can have half a language. But there's more to the trait of holism that is uniquely illuminated by the example of love. Just as being in love colors all of existence, so the nature of one's consciousness colors all of one's experience. It's not just a matter of the digital on/off of being awake or asleep, conscious or unconscious. There are *kinds* of consciousness, in addition to being in love, that color all aspects of consciousness, e.g. depression or elation; degrees of alertness that are pervasive; moods.

Love, Recursion, and Consciousness

Love talks about itself. It declares itself. It sings aloud. It writes poetry. Love is so articulate as to be positively effusive. Love is self-referential. It feeds on itself.

So likewise consciousness, even to the extent that Jean-Paul Sartre would write, with a typically French taste for the paradoxical, that, "There is no consciousness without self-consciousness."

In addition to the *general* applicability of Trait Four to consciousness, the gift we are given by focusing on the example of love is more specific: it is the combination of self-reference or recursion *and* the role of symbolic representation. Cats and dogs have a certain degree of consciousness. But lacking the gift of language, they can't enjoy (or suffer) the same degree of consciousness that humans have. Dogs mate with other dogs, and they genuinely seem to love their masters. Their affection is palpable and unmistakable. But they are incapable of romance. They don't write or read poetry. So likewise when it comes to consciousness. Highly sentient beings like dogs, chimps and Bonobos clearly have some degree of consciousness. They exhibit teleological behavior. They can plot and scheme after a fashion, even deceive. But for all their mastery of iconic and indexical expression, they lack genuinely symbolic communication and therefore neither love nor possess the same degree of consciousness that humans have.

The unpredictability of love, and consciousness

You never know who is going to fall in love with whom. That is why arranged marriages are not a good idea. You cannot predict who will fall in love with whom on the basis of any component of personality or character.

Once again, the general applicability of Trait Five to consciousness is fairly obvious: You can't predict what will occur to consciousness *next.* Like evolution, it is fecund and unpredictable. But what, more specifically, does the unpredictability of love have to tell us about the unpredictability of consciousness?

Earlier it was observed that not only does the heart have its reasons, but reason has a heart. Once we appreciate the degree to which consciousness is a satellite of desire rather than the other way around, then we stand forewarned that so-called *rationality* will not always follow the kind of strictly logico-analytic deductive sequences that rational choice theorists suggest. (See the work of Tversky and Kahneman as described in Michael Lewis, *The Undoing Project,* Norton, 2016.) The predictions of rational choice theorists, based as they may be on careful calculations of marginal utility, will be confounded by desires that desire what they desire despite what the models may predict.

One feels a certain sense of reluctance at stating the obvious in making these remarks . . . but this is one of those places of vindication for the stereoscopic vision that comes with supplementing monological science with a fluency in the traits of emergence. We all *knew* that rational choice theory was wacky, that people just don't behave only in quest of "marginal utility." But nothing in our theory enabled us to refute those smug theorists. Now, with a deeper

appreciation for the role of desire in the emergence of consciousness, we can see more clearly just how and why the calm calculations that robots and computers perform bear so little resemblance to the behavior of living breathing human beings.

The irreducibility of love, and consciousness

Try to complete the sentence, "I love her *because . . ."* and you will sound like an idiot. Likewise, any attempt to locate *the key to consciousness,* the single cause that serves as a necessary and sufficient condition . . . will end in foolishness. Consciousness, like love, is an emergent phenomenon that is dependent upon the coming together of *many* factors, not the mono-linear effect of one singular cause.

Desire, the Ascents of Love, and Levels of Consciousness

The long section following Martha Nussbaum's portrayal of three ascents of love was followed by a descent toward the everyday as represented in Joyce's *Ulysses.* The first ascent was the contemplation of the true and the beautiful in Plato, Spinoza and Proust. The second ascent was the Christian account of humility, longing and grace as represented in Augustine and Dante. Third, the Romantic Ascent was represented by Gustav Mahler and Emily Bronte.

To return to the world, Nussbaum takes a plunge into quotidian reality as represented in a single day of Bloom's richly textured and erotic life.

In seeking illumination about the nature of consciousness from the emergence of love, it would be too much to ask that we find parallels to every one of the steps of ascent and descent traversed by Nussbaum. But the fact of such a path, and the multiple steps of its trajectory, are significant in themselves, independent of the content of each of the steps. *There are different levels of consciousness, different developmental stages, both in the individual first person singular, and in the collective first person plural.*

Hegel was the master at articulating multiple *Gestalten des Buwusstseins* (forms of consciousness) for the first person plural—the collective consciousness that has come down to us in the term *Weltgeist,* or World-historical spirit. As opposed to all philosophers prior to Hegel, for whom the quest was to identify timeless truths and the eternal structure of consciousness, Hegel showed how the consciousness of the ancient Greeks was different from the consciousness of the early Christians, from which the consciousness of the rationalist Enlightenment differed yet again.

With respect to the first person singular, there is a long tradition dating from Plato's Divided Line to the works of developmental psychology articulating various stages of ascent for consciousness. From Piaget and Erik Erikson to Dan Levinson, Kohlberg,

various theorists and mystics have mapped many paths charting many levels up Mt. Consciousness. . . but let us beware of the power of this metaphor, "Mt. Consciousness." Mountains like K2 or Everest have a single, magestic, highest peak. All paths of ascent must necessarily converge on that one peak. So monological this metaphor of the single peak!

Consider another metaphor that might do more justice to both the prolific pluralism of evolution and the quotidian reality of Bloom's love for Mollie: substitute trees for mountains. Surely the upward striving of trees carries the image of verticality and ascent with which mountains seduced us. But with trees we find two advantages: first, there are so many different *kinds* of them. Second, we can use their differences to express a more multi-dimensional field than simply higher and lower. E.g., as much as one has to admire the iconoclastic courage of a Ken Wilber who has skirted the academy his whole life long while scrupulously cataloguing *all* of the developmental schemes, East *and* West, I come away saying that while his map for the ascent of consciousness is like a mighty redwood, mine is a mere maple, gnarly, and round at its crown rather than pointed heavenward. But observe the brilliant oranges and reds those maples turn in autumn, and taste the sweetness of their syrup! And see how the proliferation of branches in many directions mimics the proliferation of species in evolution: no *highest* species, as if evolution had a single telos, but still an obvious if multifarious directionality as thousands of twigs seek the sun like species investigating adjacent niches.

This whole business of different grades of consciousness, or levels on the ascent, is fraught with peril. There is the peril of heights and the peril of depths: levels of

enlightenment, and questions about how far down the phylogenetic tree it makes sense to attribute consciousness—to dogs and cats? To banana slugs? To protists?

Rather than get caught up in border skirmishes over conflicting taxonomies—What are the criteria for graduation from the third to the fourth level of Plato's Divided Line, or from the fifth to the sixth level in Wilber's color-coded hierarchy? If bacteria are *not* conscious but Bonobos *are,* then precisely where on the phylogenetic tree does consciousness emerge?—it seems more prudent to say something like the following: Yes, there are different levels of consciousness, different degrees. And sometimes the transition from one level to another occurs with the kind of dramatic discontinuity characteristic of emergent systems—a *metanoia,* a transformation, a *pop!* (Ttrait 2) Some such transformation are individualch—a characteristic of a whole system (Trait 3) suddenly coming together in a recursive closing of a circle (Trait 4). But, given the sheer complexity of the systems involved, and the gnarly variety of loops and relationships, it is unwise to give too much credence to pompous claims about precisely *where* certain boundary lines are *first* crossed (trait 1). Graduations and transitions are unpredictable from their precursors (trait 5), and the features of higher planes are irreducible to the properties of the component parts that preceded them (Trait 6). Yes, there is a yearning, an aspiration to ascent (trait 7). But if you put up your stores in any orthodoxy about the precise boundaries that define precise levels on the ascent of consciousness, you commit the sin of idolatry, the error of literalism, and will almost certainly suffer a dark night of the soul (trait 8). Let us only hope that the precise enumeration of 8, count them 8, traits of emergence is not equally guilty of the same error of literalism. Is it simple self-contradiction or sacred paradox to attempt precision in explaining the dangers of too much precision?

Divorce, Perversion, and Consciousness

What is it *like* to be conscious? Rather than focusing only on the very appealing if very difficult questions relating to the emergence and the ascents of consciousness, we should not neglect the lessons of Dante and Freud, or Joyce and Jung. There's much to be gained by following Dante into the dark wood, or Jung into the Orphic underworld. Sometimes you can learn a lot about how something fits together by taking it apart. It's called *analysis.*

The prevalence of erotic imagery in the practice of psychopathology bears witness to the centrality of desire at the heart of consciousness. The way love ebbs and flows and sometimes ends in divorce offers a possible model for the way consciousness comes and goes in wakefulness and sleep and sometimes disintegrates altogether in madness, sleep, or death. Further, the rich variety of species of madness and the colorful range of

perversions should give us some hint of the gnarly twists and turns that desire can take in seeking gratification.

Consciousness is not just about the processing of information. Hunger, desire, love and lust, yearning, aspiration and hope—the whole hierarchy of desire sits inside whatever hierarchy of levels of consciousness the cartographers of transcendence may wish to delineate. And as we lean our ladders ever higher, we do well to attend to the many chutes toward descent. They tell us as much about inner cauldrons of consciousness as do our aspirations toward enlightenment.

Emergence and Artistic Creativity

What does the concept of emergence have to contribute to our understanding of artistic creativity? Just this: By prying our comprehension of change loose from the notion of monolinear causality and wedding it instead to the image of many factors *coming together,* the concept of emergence renders otiose the idea of the singular, uncreated creator who, like the monotheistic god of Christianity, can act like Aristotle's unmoved mover.

Under the influence of—even intoxicated by—the brooding of the European Romantic tradition, we run the danger of being over-impressed by the myth of *genius.* Genius is to creativity as *élan vital* is to life: a mystifying black box that draws into itself, like a black hole, all of the many factors that must come together in acts of creativity, just as the singular but obscure cause called *élan vital* was supposed to account for life. Neither life nor creativity has a single cause. To seek that sort of singular cause for complex phenomena is to give in to what Heidegger called the onto-theology of the western

metaphysical tradition. It is to project upon every act of artistic creativity the *imago dei* of the Christian creator God, the *genius* who could create something out of nothing.

Emergence has been defined more modestly as the creation of more out of less, thereby seeming to defy the principle of the constant conservation of mass and energy. Of course there's the danger of the concept of emergence functioning as its own black box, for that is how it first came on the scene. In the hands of its earliest users—philosophers like Samuel Alexander and Henri Bergson, who foisted upon us the notion of *élan vital*—emergence stood in for those words in the famous cartoon I copied up top of physicists standing before a blackboard with the words, "And here a miracle occurs." In the hands of its first progenitors, well-forgotten theorists like C. Lloyd Morgan, the concept of emergence was invoked to "explain" phenomena for which ordinary science simply lacked the means. How does the taste of salt emerge from the combination of sodium and chlorine? Ordinary physics and chemistry won't tell us. Invoke emergence. How does the liquidity of water follow from the combination of hydrogen and oxygen? Ordinary physics and chemistry won't tell us. Invoke emergence. But such invocations provide no explanation at all. They just give a name to our ignorance.

But that was then, during the early years of the 20th century. This is now, after decades of scientific study in biology, linguistics, and complexity theory; after structuralism and its emphasis on complex structures of relationships; after the intoxication with singular causes and well into the hangover called deconstruction. Now, with this book, the black box has been pried open and we have peered within it to find all manner of traits and features of emergent systems. What was only implicit in the black box of emergence as first invoked, it has now been rendered explicit, and the results have implications for our understanding of creativity.

The Relevance of Emergence to Creativity

How do the eight traits shed light on artistic creativity? Let us count the ways:

1) **No first instance:** Acts of creativity do not spawn 2 from 1, 4 from 2, as on the genealogical trees of cladistics. Yes, there are lines of influence, but those lines are more likely twisted by what Harold Bloom has called "the anxiety of influence." Where the physicist stands on the shoulders of earlier physicists, the artist is as likely to stand in a relationship of profound ambivalence toward her elders, on the one hand inspired, but on the other in desperate need of something new and *different.* And in a particular act of creativity, say the creation of an abstract painting, there is no proper place to *begin*—as if the success of the act could be guaranteed by performing step one before step two, step two before step three . . .

foundation before first floor, first floor before second. Creativity doesn't work like that. Instead, one mucks about for a bit trying first this, then that, erasing, starting again, painting over, starting yet again . . . and then . . .

2) **Pop!** It all comes together! The pieces fall into place. Their interrelationships are not *caused* by any one *first* piece. Instead it is the interrelationships among all of the pieces that come into view all of a sudden with an *Ah-ha* that we sometimes attribute to "inspiration." But this latter concept, *inspiration,* may be as misleading as that of *genius*—one more attempt to find a single cause, a single source, a single origin for what is always and inevitably multiple. Likewise "the muse" may be called upon as a singular cause *outside* the creator just as simple and misleading as genius *inside.*

3) **Holism:** The success of an act of creation cannot be attributed to any one part, first or otherwise. Instead, the success of an act of artisistic creation—what, in the realm of the aesthetic, we often call *beauty*—is always a function of the way all of the pieces fit together. Try looking at a great work of art, then completing the sentence, "It is beautiful because _____," where you fill in the blank by naming some single part of the painting . . . and you will sound like an idiot. Definitions and attributions of beauty are famously difficult just because there is no one element that accounts for beauty. Instead it is always a matter of how the whole gives meaning and significance to each of the parts.

4) **Recursivity:** Think of the *leit-motif* that returns again and again. Think of the melodic fragment on which the jazz musician *riffs.* Think of the *ricorso,* the *cadenza,* the gestures of doubling back and repetition that save so many pieces of music from the linearity of a mere list. Think of the complex relationships between form and content in Plato's Dialogues where the dramatic action belies the logical conclusion, thus introducing an ironic depth that reflects in its form the very irony discussed as part of the dialogic content. Works of artistic creativity are characteristically loopy in this way, with their form somewhow reflecting their content, though not formulaicly so, for works of art are nothing if not . . .

5) **Unpredictable:** If you are painting by the numbers, if you are following a formula, then you are not creating. You are not *playing.* You are *working.* You can neither predict the next creation of an individual artist, nor the course of the next step in the history of art. The idea of *novelty* that is so central to what we think of as artistic creativity derives from this fifth trait of emergent systems. If you can predict it, then it's neither all that new nor is it creative.

6) **Irreducibility:** Likewise, and again symmetrically, that which is unpredictable cannot be reduced to the features of its antecedent components. The work of art is

not beautiful *because* of any one of its components. Nor can its beauty be analyzed into being *nothing but* the arrangement of its parts.

7) **Desire:** What is it that so delights us about the products of artistic creativity? Why do they give us such pleasure? Doesn't it have something to do with the way they satisfy some deeply felt need? The satisfaction of desire is evident at the level of the whole, but can be more precisely analyzed in some of the parts, e.g., the way the dissonance of the Neapolitan sixth toward the conclusion of the second movement of Bach's Italian Concerto just *has to* resolve into the cadence. There is a *nisus* toward closure in all works of art, even those whose service to novelty and unpredictability holds them open in ways that can't be closed in the old ways. Hence irony and the perpetual dissonances of post-modernism—a paradoxical closure upon openness. The work of creative art *wants* to be just the way it is, and thus elicits a feeling of inevitability and rightness even as it exposes us to something altogether new under the sun.

8) **Death:** Beauty is famously evanescent. It doesn't last. It can't be monumental, and therein lies its preciousness. Artistic creativity takes place in time, and time, *pace* Plato, is *not "*the moving image of eternity." Time is the medium in which things come to be and pass away, and artistic creativity has no cure for the ruthlessness of becoming save the balm of beauty in the moment.

So there we have at least a first cut at the ways the traits of emergent systems are instantiated in acts of creativity.

Specific Admonitions to the Aspiring Creator

What more can be said by way of shedding light on creativity from the reflected glare of the concept of emergence? Even at the risk of leading readers toward "painting by the numbers," can we conclude with some relatively simple, relatively straightforward dos and don'ts? Let's emphasize the don'ts, and leave the dos to the reader's creativity.

1) **No first instance:** Don't expect to find a sure foundation for creativity, some *first work* after which everything you do will be creative in ways that nothing previous ever was. Don't expect that, after years of practicing, one day you will sit down to *play* and then, henceforth, you will never have to practice again.

2) **Creativity pops!** Don't schedule it, but be ready for it when it happens. What is right about the dangerous reliance on the muse? Precisely the suddenness with which creativity sometimes happens when it does, and the sense that it is coming from outside oneself unscheduled, not from within by one's own intention. But don't try to "summon the muse" in some literal sense, as if all you need to do is

light just the right candles and play just the right music, then surely she will descend upon you. No, you must muck about, trying first this, then that, one draft and then another, one melody and then another, one dialogue among your characters and then another until, suddenly, *pop!* It all comes together and the novel starts writing itself.

3) **Holism:** And it must *all* come together. Don't pin your hopes for creativity on any *one* gesture, first, last, or otherwise. All the parts must make a whole that gives new meaning and significance to each of the parts.

4) **Recursivity:** Don't worry about repetition. Don't be over-impressed by Aristotle's dictum that every narrative has a beginning, a middle, and an end. Sure, every plot, every melody has its linear dimension. But creativity tends to be loopy, circling back with variations on a theme.

5) **Unpredictability:** Don't settle for the banal point that you can't schedule creativity. Indeed, you must schedule some time in the studio or some hours at the keyboard. Don't use the excuse of unpredictability to wait for the muse to bestow upon you the gift of creativity unscheduled. Instead, use what you know about the unpredictability of creativity to forgive yourself for those hours and days and months when, god knows, you worked at it but nothing happened. Use what we know about the unpredictability of emergent systems to defend yourself against all the manuals and the hucksters with their ten steps to creativity, guaranteed!

6) **Irreducibility:** In trying to learn something about creativity from the lives and works of the masters, don't try to reduce their gifts to some formula for success.

7) **Desire:** Don't confuse what you want with what the work wants. In all great works of art there is a palpable inevitability that transcends the wishes of the creator. Wherever the wishes of the creator are evident, there the promise of art surrenders to the pettiness of propaganda or the woodenness of didacticism. Great art achieves a life of its own, and its creator finds himself or herself swept up in a momentum that transcends his own desires. The work has its own needs, its own desires, and the act of creativity is a surrender to that transcendent desire.

8. Death: Don't try to create immortal art. Let your creations die. Once the first draft is complete, be prepared to burn it. After you've indulged your passion for alliteration or sexual imagery or esoteric references, go back over your text with a blue pencil that is ruthless. As F. Scott Fitzgerald once said about the act of editing and rewriting, you must be prepared, "to slay all your little darlings."

Creativity must know how to conclude. One of the hardest things an artist needs to judge is *when is this painting (or sculpture, etc.) finished.*

Artistic creation demands a coherence in which *it all comes together.* Emergent systems like life, love, language and consciousness exhibit the coming together of many elements in new wholes. The study of emergent systems can show us how many elements can coalesce in creations that embody something new under the sun. Those who aspire to creativity have something to learn from emergence.

In keeping with the title of this part, *The Phallusy of Misplaced Physics,* we can certainly conclude that creativity does not follow from deterministic causality. You cannot *cause* creativity. But you can create conditions under which it is likely to *emerge.*

Summary

This book began with a discussion of how much we *don't* know: about the origins of life, language, consciousness and love chief among them. It turns out that the key to knowing more is the concept of *emergence*.

Part One reviewed eight traits of emergent systems. These eight traits were discovered by my immersion in the literature about the origins of life and the origins of language. Most important to my learning were the books of Terrence Deacon: *The Symbolic Species* and *Incomplete Nature.* Could there have been seven traits, or nine? Possibly, but my reading of the literature on the origins of life and language revealed to me just eight.

I then used those eight traits to study the literature on evolution and on consciousness and also money, love (with Martha Nussbaum's help), and artistic creativity. I also studied the limited literature on emergence, a topic that has obsessed me ever since I read C. Lloyd Morgan's *Emergent Evolution* (1923) as an undergraduate in the early 1960s. More recently I've benefitted from Steven Johnson's excellent book, *Emergence,* and John Holland's book by the same name.

What I think I have discovered, as is suggested by the title of the last part, "The Phallusy of Misplaced Physics," is that despite the fact that everything in the universe, from neurons to planets, is made of matter, reductionistic materialism is not the last word on causality. What I call monological science is not *wrong,* just incomplete. Largely because of the evolution of information, some material things which obey the laws of physics are *also* capable of emergence.

This is a lucky thing because, as Kevin Mitchell's book, *Free Agents: How Evolution Gave Us Free Will* (Princeton University Press, 2023) shows us, without the emergence of information, the concept of free will is incoherent. It seems to defy the laws of physics, as a significant literature argues. But even those who argue for the finality of materialistic reductionism are pretty much flummoxed by our altogether obvious command of free will. In an anthology on the topic (*Free Will, ed.* Gary Watson, Oxford, 2003), a very bright philosopher, Thomas Nagel, stumbles around for 27 pages before concluding with the sentence, "As I have said, it seems to me that nothing approaching the truth has been said on this subject." (p. 256)

Fortunately, the emergence of information, which did not exist 5 billion years ago prior to the emergence of life, the concept of emergence also allows us to make sense of the origins of life, language, consciousness, love, money, and artistic creativity—all topics about which reductionistic materialism has little to offer, as is evident in the literature on free will.

I'm sure that there are other topics that can be enlightened by the concept of emergence. Cf. the sub-title of Steven Johnson's book: *The Connected Lives of Ants, Brains, Cities, and Software*. I leave it to you, the reader, to find still other topics that can be illuminated by the concept of emergence.

Acknowledgments

I've been working on this book, off and on, for over twenty years. In that long time, I'm sure there were influencers that I'll neglect here. But I cannot neglect my good friend, Terry Deacon, with whom I've been in a discussion group for most of that time. Other crucial members of that discussion group include Ty Cashman, Jeremy Sherman, Julie Hui, and James Haag, plus others who joined our discussions intermittently. I want to thank David Judson, who elicited from me many essays for the website on strategic forecasting called *Stratfor.* Nor can I neglect Alfonso Montouri, who elicited from me several essays for the jounal he edits, *World Futures: The Journal of New Paradigm Research.* Nor, most emphatically, can I neglect my other best friend, Peter Schwartz, who lifted me over the walls of academe and taught me everything I know about scenario planning. I should also add Steven Johnson, not only for his book, *Emergence,* but also for attending a conference that I hosted back in 2003. Also in attendance at that conference were Stuart Kauffman, Owen Flanagan, and Michael Murphy. Michael has been a close friend and stimulating influence for over 40 years. I can't thank him enough for his friendship and for hosting many stimulating conferences at Esalen Institute. His wife, Dulce, has also been a close friend and gentle influence, together with her lifelong friend, the late Mac Mqouwn. Credit goes to my editor, David Dennen, who changed the nature of this book by informing me, after reading an earlier draft, that it was not so much a book about consciousness as about emergence, as applied to several topics including consciousness. Though I don't know him personally, I want to thank Mitchell Waldrop for his book, *Complexity,* that was an early stimulus for writing down my long simmering thoughts about emergence. Finally, I want to thank my wife, Tricia, and my sons, David and Jonathan, for all of those hours when I was researching and writing this book, rather than attending to the family.