The INTERACTIVITY Reader

I: Interactivity as a paradigm for understanding how the world works (or, is the universe really indeterminate?):

II. Activity as a compositional value for artists; complex systems and automata:

What are generative systems?
Metacreation by Michael Whitelaw
Cellular Automata and Art by Brian P. Hoke

III. Interactivity as a compositional value for artists:

Why is interface design important in interactive art systems?
An Abstract Model of Instrument Control by N.B.Aldrich
The History of the Interface in Interactive Art by Soke Dinkla
The Interface by Lev Manovich

Why would artists create an interactive work?
The History of Electronic Music as a Reflection of Structural Paradigms by Joel Chadabe
Interaction by Caitlin Rubin
The Myth of Interactivity by Lev Manovich
Healing Interactions and Interactive Digital Art by Barbara Buckner

How has our historical perception of the physical universe changed?

OUR PICTURE OF THE UNIVERSE
Stephen W. Hawking

A well-known scientist (some say it was Bertrand Russell) once gave a public lecture on astronomy. He described how the earth orbits around the sun and how the sun, in turn, orbits around the center of a vast collection of stars called our galaxy. At the end of the lecture, a little old lady at the back of the room got up and said: "What you have told us is rubbish. The world is really a flat plate supported on the back of a giant tortoise." The scientist gave a superior smile before replying, "What is the tortoise standing on." "You're very clever, young man, very clever," said the old lady. "But it's turtles all the way down!"

Most people would find the picture of our universe as an infinite tower of tortoises rather ridiculous, but why do we think we know better? What do we know about the universe, and how do we know it? Where did the universe come from, and where is it going? Did the universe have a beginning, and if so, what happened before then? What is the nature of time? Will it ever come to an end? Can we go back in time? Recent breakthroughs in physics, made possible in part by fantastic new technologies, suggest answers to some of these longstanding questions. Someday these answers may seem as obvious to us as the earth orbiting the sun – or perhaps as ridiculous as a tower of tortoises. Only time (whatever that may be) will tell.

As long ago as 340 BC the Greek philosopher Aristotle, in his book On the Heavens, was able to put forward two good arguments for believing that the earth was a round sphere rather than a Hat plate. First, he realized that eclipses of the moon were caused by the earth coming between the sun and the moon. The earth's shadow on the moon was always round, which would be true only if the earth was spherical. If the earth had been a flat disk, the shadow would have been elongated and elliptical, unless the eclipse always occurred at a time when the sun was directly under the center of the disk. Second, the Greeks knew from their travels that the North Star appeared lower in the sky when viewed in the south than it did in more northerly regions. (Since the North Star lies over the North Pole, it appears to be directly above an observer at the North Pole, but to someone looking from the equator, it appears to lie just at the horizon. From the difference in the apparent position of the North Star in Egypt and Greece, Aristotle even quoted an estimate that the distance around the earth was 400,000 stadia. It is not known exactly what length a stadium was, but it may have been about 200 yards, which would make Aristotle's estimate about twice the currently accepted figure. The Greeks even had a third argument that the earth must be round, for why else does one first see the sails of a ship coming over the horizon, and only later see the hull?

Aristotle thought the earth was stationary and that the sun, the moon, the planets, and the stars moved in circular orbits about the earth. He believed this because he felt, for mystical reasons, that the earth was the center of the universe, and that circular motion was the most perfect. This idea was elaborated by Ptolemy in the second century AD into a complete cosmological model. The earth stood at the center, surrounded by eight spheres that carried the moon, the sun, the stars, and the five planets known at the time, Mercury, Venus, Mars, Jupiter, and Saturn. The planets themselves moved on smaller circles attached to their respective spheres in order to account for their rather complicated observed paths in the sky. The outermost sphere carried the so-called fixed stars, which always stay in the same positions relative to each other but which rotate together across the sky.

What lay beyond the last sphere was never made very clear, but it certainly was not part of mankind's observable universe.

Ptolemy's model provided a reasonably accurate system for predicting the positions of heavenly bodies in the sky. But in order to predict these positions correctly, Ptolemy had to make an assumption that the moon followed a path that sometimes brought it twice as close to the earth as at other times. And that meant that the moon ought sometimes to appear twice as big as at other times! Ptolemy recognized this flaw, but nevertheless his model was generally, although not universally, accepted. It was adopted by the Christian church as the picture of the universe that was in accordance with Scripture, for it had the great advantage that it left lots of room outside the sphere of fixed stars for heaven and hell.

A simpler model, however, was proposed in 1514 by a Polish priest, Nicholas Copernicus. (At first, perhaps for fear of being branded a heretic by his church, Copernicus circulated his model anonymously.) His idea was that the sun was stationary at the center and that the earth and the planets moved in circular orbits around the sun. Nearly a century passed before this idea was taken seriously. Then two astronomers – the German, Johannes Kepler, and the Italian, Galileo Galilei – started publicly to support the Copernican theory, despite the fact that the orbits it predicted did not quite match the ones observed. The death blow to the Aristotelian/Ptolemaic theory came in 1609. In that year, Galileo started observing the night sky with a telescope, which had just been invented. When he looked at the planet Jupiter, Galileo found that it was accompanied by several small satellites or moons that orbited around it. This implied that everything did not have to orbit directly around the earth, as Aristotle and Ptolemy had thought. (It was, of course, still possible to believe that the earth was stationary at the center of the universe and that the moons of Jupiter moved on extremely complicated paths around the earth, giving the appearance that they orbited Jupiter. However, Copernicus's theory was much simpler.) At the same time, Johannes Kepler had modified Copernicus's theory, suggesting that the planets moved not in circles but in ellipses (an ellipse is an elongated circle). The predictions now finally matched the observations.

As far as Kepler was concerned, elliptical orbits were merely an ad hoc hypothesis, and a rather repugnant one at that, because ellipses were clearly less perfect than circles. Having discovered almost by accident that elliptical orbits fit the observations well, he could not reconcile them with his idea that the planets were made to orbit the sun by magnetic forces. An explanation was provided only much later, in 1687, when Sir Isaac Newton published his Philosophiae Naturalis Principia Mathematica, probably the most important single work ever published in the physical sciences. In it Newton not only put forward a theory of how bodies move in space and time, but he also developed the complicated mathematics needed to analyze those motions. In addition, Newton postulated a law of universal gravitation according to which each body in the universe was attracted toward every other body by a force that was stronger the more massive the bodies and the closer they were to each other. It was this same force that caused objects to fall to the ground. (The story that Newton was inspired by an apple hitting his head is almost certainly apocryphal. All Newton himself ever said was that the idea of
gravity came to him as he sat "in a contemplative mood" and "was occasioned by the fall of an apple.") Newton went on to show that, according to his law, gravity causes the moon to move in an elliptical orbit around the earth and causes the earth and the planets to follow elliptical paths around

the sun.
The Copernican model got rid of Ptolemy's celestial spheres, and with them, the idea that the universe had a natural boundary. Since "fixed stars" did not appear to change their positions apart from a rotation across the sky caused by the earth spinning on its axis, it became natural to suppose that the fixed stars were objects like our sun but very much farther away.

Newton realized that, according to his theory of gravity, the stars should attract each other, so it seemed they could not remain essentially motionless. Would they not all fall together at some point? In a letter in 1691 to Richard Bentley, another leading thinker of his day, Newton argued that this would indeed happen if there were only a finite number of stars distributed over a finite region of space. But he reasoned that if, on the other hand, there were an infinite number of stars, distributed more or less uniformly over infinite space, this would not happen, because there would not be any central point for them to fall to.

This argument is an instance of the pitfalls that you can encounter in talking about infinity. In an infinite universe, every point can be regarded as the center, because every point has an infinite number of stars on each side of it. The correct approach, it was realized only much later, is to consider the finite situation, in which the stars all fall in on each other, and then to ask how things change if one adds more stars roughly uniformly distributed outside this region. According to Newton's law, the extra stars would make no difference at all to the original ones on average, so the stars would fall in just as fast. We can add as many stars as we like, but they will still always collapse in on themselves. We now know it is impossible to have an infinite static model of the universe in which gravity is always attractive.

It is an interesting reflection on the general climate of thought before the twentieth century that no one had suggested that the universe was expanding or contracting. It was generally accepted that either the universe had existed forever in an unchanging state, or that it had been created at a finite time in the past more or less as we observe it today. In part this may have been due to people's tendency to believe in eternal truths, as well as the comfort they found in the thought that even though they may grow old and die, the universe is eternal and unchanging.

Even those who realized that Newton's theory of gravity showed that the universe could not be static did not think to suggest that it might be expanding. Instead, they attempted to modify the theory by making the gravitational force repulsive at very large distances. This did not significantly affect their predictions of the motions of the planets, but it allowed an infinite distribution of stars to remain in equilibrium – with the attractive forces between nearby stars balanced by the repulsive forces from those that were farther away. However, we now believe such an equilibrium would be unstable: if the stars in some region got only slightly nearer each other, the attractive forces between them would become stronger and dominate over the repulsive forces so that the stars would continue to fall toward each other. On the other hand, if the stars got a bit farther away from each other, the repulsive forces would dominate and drive them farther apart.

Another objection to an infinite static universe is normally ascribed to the German philosopher Heinrich Olbers, who wrote about this theory in 1823. In fact, various contemporaries of Newton had raised the problem, and the Olbers article was not even the first to contain plausible arguments against it. It was, however, the first to be widely noted. The difficulty is that in an infinite static universe nearly every line of sight would end on the surface of a star. Thus one would expect that the whole sky would be as bright as the sun, even at night. Olbers' counter-argument was that the light from distant stars would be dimmed by absorption by intervening matter. However, if that happened the intervening matter would eventually heat up until it glowed as brightly as the stars. The only way of avoiding the conclusion that the whole of the night sky should be as bright as the surface of the sun would be to assume that the stars had not been shining forever but had turned on at some finite time in the past. In that case the absorbing matter might not have heated up yet or the light from distant stars might not yet have reached us. And that brings us to the question of what could have caused the stars to have turned on in the first place.

The beginning of the universe had, of course, been discussed long before this. According to a number of early cosmologies and the Jewish/Christian/Muslim tradition, the universe started at a finite, and not very distant, time in the past. One argument for such a beginning was the feeling that it was necessary to have "First Cause" to explain the existence of the universe. (Within the universe, you always explained one event as being caused by some earlier event, but the existence of the universe itself could be explained in this way only if it had some beginning.) Another argument was put forward by St. Augustine in his book The City of God. He pointed out that civilization is progressing and we remember who performed this deed or developed that technique. Thus man, and so also perhaps the universe, could not have been around all that long. St. Augustine accepted a date of about 5000 BC for the Creation of the universe according to the book of Genesis. (It is interesting that this is not so far from the end of the last Ice Age, about 10,000 BC, which is when archaeologists tell us that civilization really began.)

Aristotle, and most of the other Greek philosophers, on the other hand, did not like the idea of a creation because it smacked too much of divine intervention. They believed, therefore, that the human race and the world around it had existed, and would exist, forever. The ancients had already considered the argument about progress described above, and answered it by saying that there had been periodic floods or other disasters that repeatedly set the human race right back to the beginning of civilization.

The questions of whether the universe had a beginning in time and whether it is limited in space were later extensively examined by the philosopher Immanuel Kant in his monumental (and very obscure) work Critique of Pure Reason, published in 1781. He called these questions antinomies (that is, contradictions) of pure reason because he felt that there were equally compelling arguments for believing the thesis, that the universe had a beginning, and the antithesis, that it had existed forever. His argument for the thesis was that if the universe did not have a beginning, there would be an infinite period of time before any event, which he considered absurd. The argument for the antithesis was that if the universe had a beginning, there would be an infinite period of time before it, so why should the universe begin at any one particular time? In fact, his cases for both the thesis and the antithesis are really the same argument. They are both based on his unspoken assumption that time continues back forever, whether or not the universe had existed forever. As we shall see, the concept of time has no meaning before the beginning of the universe. This was first pointed out by St. Augustine. When asked: "What did God do before he created the universe?" Augustine didn't reply: "He was preparing Hell for people who asked such questions." Instead, he said that time was a property of the universe that God created, and that time did not exist before the beginning of the universe.

When most people believed in an essentially static and unchanging universe, the question of whether or not it had a beginning was really one of metaphysics or theology. One could account for what was observed equally well on the theory that the universe had existed forever or on the theory that it was set in motion at some finite time in such a manner as to look as though it had existed forever. But in 1929, Edwin Hubble made the landmark observation that wherever you look, distant galaxies are moving rapidly away from us. In other words, the universe is expanding. This means that at earlier times objects would have been closer together. In fact, it seemed that there was a time, about ten or twenty thousand million years ago, when they were all at exactly the same place and when, therefore, the density of the universe was infinite. This discovery finally brought the question of the beginning of the universe into the realm of science.

Hubble's observations suggested that there was a time, called the big bang, when the universe was infinitesimally small and infinitely dense. Under such conditions all the laws of science, and therefore all ability to predict the future, would break down. If there were events earlier than this time, then they could not affect what happens at the present time. Their existence can be ignored because it would have no observational consequences. One may say that time had a beginning at the big bang, in the sense that earlier times simply would not be defined. It should be emphasized that this beginning in time is very different from those that had been considered previously. In an unchanging universe a beginning in time is something that has to be imposed by some being outside the universe; there is no physical necessity for a beginning. One can imagine that God created the universe at literally any time in the past. On the other hand, if the universe is expanding, there may be physical reasons why there had to be a beginning. One could still imagine that God created the universe at the instant of the big bang, or even afterwards in just such a way as to make it look as though there had been a big bang, but it would be meaningless to suppose that it was created before the big bang. An expanding universe does not preclude a creator, but it does place limits on when he might have carried out his job!

In order to talk about the nature of the universe and to discuss questions such as whether it has a beginning or an end, you have to be clear about what a scientific theory is. I shall take the simpleminded view that a theory is just a model of the universe, or a restricted part of it, and a set of rules that relate quantities in the model to observations that we make. It exists only in our minds and does not have any other reality (whatever that might mean). A theory is a good theory if it satisfies two requirements. It must accurately describe a large class of observations on the basis of a model that contains only a few arbitrary elements, and it must make definite predictions about the results of future observations. For example, Aristotle believed Empedocles's theory that everything was made out of four elements, earth, air, fire, and water. This was simple enough, but did not make any definite predictions. On the other hand, Newton's theory of gravity was based on an even simpler model, in which bodies attracted each other with a force that was proportional to a quantity called their mass and inversely proportional to the square of the distance between them. Yet it predicts the motions of the sun, the moon, and the planets to a high degree of accuracy.

Any physical theory is always provisional, in the sense that it is only a hypothesis: you can never prove it. No matter how many times the results of experiments agree with some theory, you can never be sure that the next time the result will not contradict the theory. On the other hand, you can disprove a theory by finding even a single observation that disagrees with the predictions of the theory. As philosopher of science Karl Popper has emphasized, a good theory is characterized by the fact that it makes a number of predictions that could in principle be disproved or falsified by observation. Each time new experiments are observed to agree with the predictions the theory survives, and our confidence in it is increased; but if ever a new observation is found to disagree, we have to abandon or modify the theory. At least that is what is supposed to happen, but you can always question the competence of the person who carried out the observation.

In practice, what often happens is that a new theory is devised that is really an extension of the previous theory. For example, very accurate observations of the planet Mercury revealed a small difference between its motion and the predictions of Newton's theory of gravity. Einstein's general theory of relativity predicted a slightly different motion from Newton's theory. The fact that Einstein's predictions matched what was seen, while Newton's did not, was one of the crucial confirmations of the new theory. However, we still use Newton's theory for all practical purposes because the difference between its predictions and those of general relativity is very small in the situations that we normally deal with. (Newton's theory also has the great advantage that it is much simpler to work with than Einstein's!)

The eventual goal of science is to provide a single theory that describes the whole universe. However, the approach most scientists actually follow is to separate the problem into two parts. First, there are the laws that tell us how the universe changes with time. (If we know what the universe is like at any one time, these physical laws tell us how it will look at any later time.) Second, there is the question of the initial state of the universe. Some people feel that science should be concerned with only the first part; they regard the question of the initial situation as a matter for metaphysics or religion. They would say that God, being omnipotent, could have started the universe off any way he wanted. That may be so, but in that case he also could have made it develop in a completely arbitrary way. Yet it appears that he chose to make it evolve in a very regular way according to certain laws. It therefore seems equally reasonable to suppose that there are also laws governing the initial state.

It turns out to be very difficult to devise a theory to describe the universe all in one go. Instead, we break the problem up into bits and invent a number of partial theories. Each of these partial theories describes and predicts a certain limited class of observations, neglecting the effects of other quantities, or representing them by simple sets of numbers. It may be that this approach is completely wrong. If everything in the universe depends on everything else in a fundamental way, it might be impossible to get close to a full solution by investigating parts of the problem in isolation. Nevertheless, it is certainly the way that we have made progress in the past. The classic example again is the Newtonian theory of gravity, which tells us that the gravitational force between two bodies depends only on one number associated with each body, its mass, but is otherwise independent of what the bodies are made of. Thus one does not need to have a theory of the structure and constitution of the sun and the planets in order to calculate their orbits.

Today scientists describe the universe in terms of two basic partial theories – the general theory of relativity and quantum mechanics. They are the great intellectual achievements of the first half of this century. The general theory of relativity describes the force of gravity and the large-scale structure of the universe, that is, the structure on scales from only a few miles to as large as a million million million million (1 with twenty-four zeros after it) miles, the size of the observable universe. Quantum mechanics, on the other hand, deals with phenomena on extremely small scales, such as a millionth of a millionth of an inch. Unfortunately, however, these two theories are known to be inconsistent with each other – they cannot both be correct. One of the major endeavors in physics today, and the major theme of this book, is the search for a new theory that will incorporate them both – a quantum theory of gravity. We do not yet have such a theory, and we may still be a long way from having one, but we do already know many of the properties that it must have. And we shall see, in later chapters, that we already know a fair amount about the predications a quantum theory of gravity must make.

Now, if you believe that the universe is not arbitrary, but is governed by definite laws, you ultimately have to combine the partial theories into a complete unified theory that will describe everything in the universe. But there is a fundamental paradox in the search for such a complete unified theory. The ideas about scientific theories outlined above assume we are rational beings who are free to observe the universe as we want and to draw logical deductions from what we see. In such a scheme it is reasonable to suppose that we might progress ever closer toward the laws that govern our universe. Yet if there really is a complete unified theory, it would also presumably determine our actions. And so the theory itself would determine the outcome of our search for it! And why should it determine that we come to the right conclusions from the evidence? Might it not equally well determine that we draw the wrong conclusion.? Or no conclusion at all?

The only answer that I can give to this problem is based on Darwin's principle of natural selection. The idea is that in any population of self-reproducing organisms, there will be variations in the genetic material and upbringing that different individuals have. These differences will mean that some individuals are better able than others to draw the right conclusions about the world around them and to act accordingly. These individuals will be more likely to survive and reproduce and so their pattern of behavior and thought will come to dominate. It has certainly been true in the past that what we call intelligence and scientific discovery have conveyed a survival advantage. It is not so clear that this is still the case: our scientific discoveries may well destroy us all, and even if they don't, a complete unified theory may not make much difference to our chances of survival. However, provided the universe has evolved in a regular way, we might expect that the reasoning abilities that natural selection has given us would be valid also in our search for a complete unified theory, and so would not lead us to the wrong conclusions.

Because the partial theories that we already have are sufficient to make accurate predictions in all but the most extreme situations, the search for the ultimate theory of the universe seems difficult to justify on practical grounds. (It is worth noting, though, that similar arguments could have been used against both relativity and quantum mechanics, and these theories have given us both nuclear energy and the microelectronics revolution!) The discovery of a complete unified theory, therefore, may not aid the survival of our species. It may not even affect our lifestyle. But ever since the dawn of civilization, people have not been content to see events as unconnected and inexplicable. They have craved an understanding of the underlying order in the world. Today we still yearn to know why we are here and where we came from. Humanity's deepest desire for knowledge is justification enough for our continuing quest. And our goal is nothing less than a complete description of the universe we live in.

CHAPTER

1

The Particle Zoo

The fourth translation of Wu Li is "I Clutch My Ideas." This is appropriate to a book on physics since the history of science in general often has been the story of scientists vigorously fighting an onslaught of new ideas. This is because it is difficult to relinquish the sense of security that conies from a long and rewarding acquaintance with a particular world view.

The value of a physical theory depends upon its usefulness. In this sense the history of physical theories might be said to resemble the history of individual personality traits. Most of us respond to our environment with a collection of automatic responses that once brought desirable results, usually in childhood. Unfortunately, if the environment that produced these responses changes (we grow up) and the responses themselves do not adapt, they become counterproductive. Showing anger, becoming depressed, flattering, crying, and bullying behavior are response patterns appropriate to times often long past. These patterns change only when we are forced to realize that they are no longer productive. Even then change is often painful and slow. The same is true of scientific theories.

Not one person, except Copernicus, wanted to accept the Copernican idea that the earth revolves around the sun. Goethe wrote about the Copernican revolution:

> Perhaps a greater demand has never been laid upon mankind; for by this admission[that the earth is not the center of the universe], how much else did not collapse in dust and smoke: a second paradise, a world of innocence, poetry, and piety, the witness of the senses, the convic-

> tions of a poetic and religious faith; no wonder that men had no stomach for all this, that they ranged themselves in every way against such a doctrine . . .[1]

Not one physicist, not even Planck himself, wanted to accept the implications of Planck's discovery, for to do so threatened a scientific structure (Newtonian physics) over three hundred years old. Heisenberg wrote about the quantum revolution:

> . . . when new groups of phenomena compel changes in the pattern of thought . . . even the most eminent of physicists find immense difficulties. For the demand for change in the thought pattern may engender the feeling that the ground is to be pulled from under one's feet. . . . I believe that the difficulties at this point can hardly be overestimated. Once one has experienced the desperation with which clever and conciliatory men of science react to the demand for a change in the thought pattern, one can only be amazed that such revolutions in science have actually been possible at all.[2]

Scientific revolutions are forced upon us by the discovery of phenomena that are not comprehensible in terms of the old theories. Old theories die hard. Much more is at stake than the theories themselves. To give up our privileged position at the center of the universe, as Copernicus asked, was an enormous psychological task. To accept that nature is fundamentally irrational (governed by chance), which is the essential statement of quantum mechanics, is a powerful blow to the intellect. Nonetheless, as new theories demonstrate superior utility, their adversaries, however reluctantly, have little choice but to accept them. In so doing, they also must grant a measure of recognition to the world views that accompany them.

Today, particle accelerators, bubble chambers and computer printouts are giving birth to another world view. This world view is as different from the world view at the beginning of this century as the Copernican world view was from its predecessors. It calls upon us to relinquish many of our closely clutched ideas.

In this world view there is no substance.

The most common question that we can ask about an object is, "What is it made of?" That question, however, "What is it made of?", is based upon an artifical mental structure that is much like a hall of mirrors. If we stand directly between two mirrors and look into one, we see our reflection, and, just behind ourselves, we see a crowd of "us"s, each looking at the back of the head in front of it, stretching backward as far as we can see. These reflections, all of them, are illusions. The only real thing in the whole setting is *us (we)*.

This situation is very similar to what happens whenever we ask of something, "What is it made of?" The answer to such a question is always another something to which we can apply the same question.

Suppose, for example, that we ask of an ordinary toothpick, "What is it made of?" The answer, of course, is "wood." However, the question itself has taken us into a hall of mirrors because now we can ask about the wood, "What is it made of?" Closer examination reveals that wood is made of fibers, uut what the fibers are made of is another question, and so on.

Like a pair of parallel mirrors, reflecting reflections, gives the illusion of an unending progression to nowhere, the idea that a thing can be different from what it is made of creates an infinite progression of answers, leaving us forever frustrated in an unending search. No matter what something—anything—is "made of, we have created an illusion which forces us to ask, "Yes, but what is *that* made of?"

Physicists are people who have pursued tenaciously this endless series of questions. What they have found is startling. Wood fibers, to continue the example, are actually patterns of cells. Cells, under magnification, are revealed to be patterns of molecules. Molecules, under higher magnification, are discovered to be patterns of atoms, and, lastly, atoms have turned out to be patterns of subatomic particles. In other words, "matter" is actually a series of *patterns out of focus*. The search for the ultimate stuff of the universe ends with the discovery that there *isn't any*.

If there is any ultimate stuff of the universe, it is pure energy, but subatomic particles are not "made of energy, they *are* energy. This is what Einstein theorized in 1905. Subatomic interactions, therefore, are interactions of energy with energy. At the subatomic level there is no longer a clear distinction between what is and what happens, between the actor and the action. At the subatomic level the dancer and the dance are one.

According to particle physics, the world is fundamentally dancing energy; energy that is everywhere and incessantly assuming first this form and then that What we have been calling matter (particles) constantK is being created annihi lated and created again This happens as particles interact and it also happens literally out of nowhere

Where there was nothing there suddenly is something and then the something is gone again often changing into something else before vanishing In particle physics there is no distinction between empty as in empty space and not emptv or between something and not-something The world of particle physics is a world of sparkling energy forever danc mg with itself in the form of its particles as they twinkle in and out of existence collide transmute and disappear again

The world view of particle physics is a picture of *chaos beneath order* At the fundamental level is a confusion of continual creation annihilation and transformation Above this confusion limiting the forms that it can take are a set of conservation laws (page 156) They do not specify what must happen as ordinary laws of physics do rather they specify what can*not* happen They are permissive laws At the subatomic level absolutely everything that is not forbidden by the conservation laws actually happens (Quantum theory describes the probabilities of the possibilities permitted by the conservation laws)

As Jack Sarfatti wrote

> Particles no longer move stiffly and formally if not majestically in predetermined paths Rather it is Marx Brothers hyperkmetic pandemonium Charlie Chaplin slapstick Helter Skelter now you see it now vou don t In fact it is not even clear what it is that has a path It s psychedelic confusion—until one sees the subtle order [3]

The old world view was a picture of order beneath chaos It assumed that beneath the prolific confusion of detail that constitutes our dailv experience he systematic and rational laws which relate them one and all This was Newton s great in sight The same laws which govern falling apples govern the motion of planets There is still of course much truth in this but the world view of particle physics is essentially the oppo site

The world view of particle physics is that of a world without ' stuff, where what is = what happens and where an unending tumultuous dance of creation annihilation and transformation

runs unabated within a framework of conservation laws and probability

High energy particle physics is the study of subatomic particles It usually is shortened to particle physics Quantum theory and relativity are the theoretical tools of particle phys ics The hardware of particle physics is housed in unimaginably expensive facilities which couple particle accelerators and computers

The original purpose of particle physics was to discover the ultimate building blocks of the universe This was to be ac comphshed bv breaking matter into smaller and smaller pieces, eventually arriving at the smallest pieces possible The experimental results of particle physics however have not been so simple Today most particle physicists are engaged in making sense out of their copious findings *

In principle particle physics hardly could bt simpler Physicists send subatomic particles smashing into each other as hard as they can They use one particle to shatter another particle so that they can see what the remains are made of The particle that does the smashing is called the projectile and the particle that gets smashed is called the target The most advanced (and expensive) particle accelerators send both the projectile and the target particles flying toward a common collision point

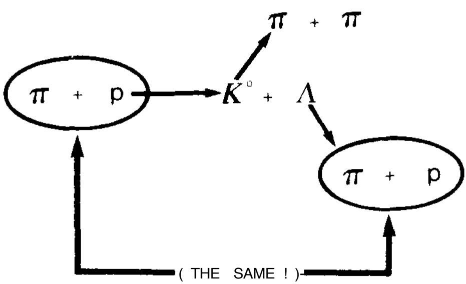
The collision point usually is located inside a device called a bubble chamber As charged particles move through a bubble chamber they leave trails similar to the vapor trails that jet liners leave in the atmosphere. The bubble chamber is located inside a magnetic field This causes particles with a positive charge to curve in one direction and particles with a negative charge to curve in the opposite direction The mass of the particle can be determined by the tightness of the curve that the particle makes (lighter particles curve more than heavier particles with the same velocity and charge) A computer-triggered camera makes a photograph every time a particle enters the bubble chamber

---

* The present state of high encrgv theory is similar to Ptolemaic astronomy before its collapse under the pressure of the new Ooernican world view The discovery of new particles and new quantum nu ibers e g charm (to be discussed later) is analogous to the addition of epicycles piled on an already unwicldv theoretical structure

This elaborate arrangement is necessary because most particles live much less than a millionth of a second and are too small to be observed directly,t In general, everything a particle physicist knows about subatomic particles, he deduces from his theories and from photographs of the tracks that particles leave in a bubble chamber, tt

Bubble chamber photographs, thousands and thousands of them, show clearly the frustrating situation which early particle physicists encountered in their search for "elementary" particles. When the projectile strikes the target, both particles are destroyed at the point of impact. In their place, however, are created *new* particles, all of which are as "elementary" as the original particles and often as massive as the original particles!



The schematic diagram above shows a typical particle interaction. A particle called a negative pi meson (IT") collides with a proton (p). Both the pie meson and the proton are destroyed and in their place are created two new particles, a neutral K meson (K°) and a lambda particle (A). Both of these particles decay spontaneously (no collision necessary) into two additional particles, leaving four new particles. Of these four particles, two of them are the same particles that we started with! It is as though, wrote Finkelstein, we fling two clocks together, they shatter, and out of them come flying not gears and springs but more clocks, some of them as large as the originals.

How can this happen? The answer is partly given by

---

t The dark-adapted eye can detect single photons. All of the other subatomic particles must be detected indirectly.

tt In addition to bubble-chamber physics there is emulsion {photographic plate) physics, counter physics, etc. However, the bubble chamber is probably the most commonly used detection device in particle physics.

Einstein's special theory of relativity. *The new particles are created from the kinetic energy (energy of motion) of the projectile particle* in addition to the mass of the projectile particle and the mass of the target particle. The faster the projectile particle is traveling, the more kinetic energy is available to create new particles at the point of impact. For this reason, governments have spent more and more money to construct larger and larger particle accelerators which can push projectile particles to higher and higher velocities. If both the projectile particle and the target particle are accelerated to the point of impact, so much the more kinetic energy is available to create new particles to study.

Every subatomic interaction consists of the annihilation of the original particles and the creation of new subatomic particles. The subatomic world is a continual dance of creation and annihilation, of mass changing to energy and energy changing to mass.* Transient forms sparkle in and out of existence creating a never-ending, forever-newly-created reality.

Mystics from both the East and the West who claim to have beheld "the face of God" speak in terms so similar to these that any psychologist who professes an interest in altered states of awareness scarcely can ignore this obvious bridge between the disciplines of physics and psychology.

The first question of particle physics is, "What collides?"

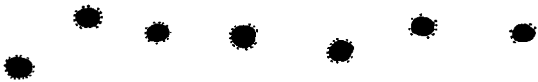
According to quantum mechanics, a subatomic particle is not a particle like a particle of dust. Rather, subatomic particles are "tendencies to exist' (page 32) and "correlations between macroscopic observables" (page 70). They have no objective existence. That means that we cannot assume, if we are to use quantum theory, that particles have an existence apart from their interactions with a measuring device (page 95). As Heisenberg wrote:

> In the light of the quantum theory . . . elementary particles are no longer real in the same sense as objects of daily life, trees or stones . . .[4]

---

* The mass/energy dualism of our ordinary conceptualizations does not exist in the formalism of relativity or quantum theory. According to Einstein's $E = me^2$, mass does not change into energy or vice versa: Energy is mass. Wherever energy, E, is present, mass, m, is present and the amount of mass, m, is given by $E = me^2$. The total amount of energy. E, is conserved, and hence the total amount of mass, m, also is conserved. This mass, m, is defined by the fact that it is a source of the gravitational field.

When an electron, for example, passes through a photographic plate it leaves a visible "track" behind it This "track," under close examination, is actually a series of dots Each dot is a grain of silver formed by the electron's interaction with atoms in the photographic plate When we look at the track under a microscope, it looks something like this



Ordinarily we would assume that one and the same electron, like a little baseball, went streaking through the photographic plate and left this trail of silver grains behind it This is a mistake Quantum mechanics tells us the same thing that Tantrie Buddhists have been saying for a millenium *The connection between the dots (the moving object") is a product of our minds* and it is not really there In rigorous quantum mechanical terms, the moving object—the particle with an independent existence-—is an unprovable assumption

According to our customarv wav of reasoning," wrote David Bohm, a professor of phvsics at Birkbeck College, University of London

> we could suppose that the track of grains of silver indicates that a real electron moves continuously through space in a path somewhere near these grains, and by interaction caused the formation of the grams But according to the usual interpretation of the quantum theory, it would be incorrect to suppose that this really happened All that we can say is that certain grains appeared, but we must not try to imagine that these grains were produced by a real object moving through space in the way in which we usuallv think of objects moving through space For although this idea of a continuously moving obiect is good enough for an approximate theory, we would discover that it would break down in a very exact theory [5]

The natural assumption that objects, like particles, are real things that run their course in space and time according to causal laws regardless of whether we are around to observe them or not is repudiated bv quantum mechanics This is especially significant because quantum mechanics is *the theory* of physics It has explained successfully everything from subatomic particles to stellar phenomena There never has been a more successful theory It has no competition

Therefore, when we look at the tracks in a bubble chamber, we are left with the question, 'What made them? The best answer that physicists have so far is that "particles are actually interactions between fields A field, like a wave, is spread out over a much larger area than a particle (a particle is restricted to one point) A field, moreover, completelv fills a given space, like the gravitational field of the earth fills all of the space immediately around it

When two fields interact with each other they interact neither gradually nor at all their areas of contact Rather, when two fields interact, thev do it *instantaneously* and at one single point in space (' instantaneouslv and locally ') These instantaneous and local interactions make what we call particles In fact, according to this theory, these instantaneous and local interactions *are* "particles " The continual creation and annihilation of particles at the subatomic level is the result of the continual interaction of different fields

This theory is called quantum field theory Some major cornerstones of the theory were laid in 1928 by the English physicist, Paul Dirac Quantum field theory has been highly successful in predicting new types of particles and in explaining existing particles in terms of field interactions According to this theory, a separate field is associated with each tvpe of particle Since only three types of particles were known in 1928, only three different fields were required to explain them The problem today, however, is that there are over one hundred known particles, which, according to quantum field theory, require over one hundred different fields This abundance of theoretical fields is somewhat awkward, not to mention embarrassing, to physicists whose goal is to simplify nature Therefore, most physicists have given up the idea of a separate field existing for each tvpe of particle

Nevertheless, quantum field theory is still an important theory not only because it works, but also because it was the first theory to merge quantum mechanics and relativity, albeit in a limited way All physical theories, including quantum theory, must satisfy the requirement of relativity theory that the laws of physics be independent of the state of motion of the observer Attempts to integrate the theory of relativity with quantum theory, however, have been generally unsuccessful Nonetheless, both relativity and quantum theory are required, and routinely used, in the understanding of particle physics Their forced relationship is best described as strained but necessary In this regard, one of the most successful integrations of the two is quantum field theory, although it covers only a relatively small range of phenomena *

Quantum field theorv is an *ad hoc* theory That means that, like Bohr's famous specific-orbits-only model of the atom, quantum field theory is a practical but conceptually inconsistent scheme Some parts of it don t fit together mathematically It is a working model designed around the available data to give physicists a place to stand in the exploration of subatomic phenomena The reason that it has been around so long is that it works so well (Some physicists think that it may work *too* well Thev fear that the pragmatic success of quantum field theory impedes the development of a consistent theory)

Even with these well-known shortcomings, the fact is that quantum field theorv is a successful phvsical theory, and it is premised on the assumption that *physical reality is* ***essentıally nonsubstantıal*** According to quantum field theory, fields alone are real *They* are the substance of the universe and not "matter Matter (particles) is simply the momentary manifestations of interacting fields which, intangible and insubstantial as they are, are the onlv real things in the universe Their interactions seem particle-like because fields interact very abruptly and in very minute regions of space

Quantum field theory ' is, of course, an outrageous contradiction in terms A quantum is an indivisible whole It is a small piece of something, while a field is a whole area of something A quantum field is the juxtaposition of two irreconcilable concepts In other words, it is a paradox It defies our categorical imperative that something be either *this* or *that,* but not both

The major contribution of quantum mechanics to western thought, and there arc many, mav be its impact on the artificial categories by which we structure our perceptions, since ossified structures of perception are the prisons in which we unknowingly become prisoners Quantum theory boldly states that something can be this *and* that (a wave *and* a particle) t

---

* S Matrix theoiv mtiges quantum theoiv and lelativity but it provides limited information on the details of subatomic phenomena and it currentK is restricted to hadron interactions (S Vlatnx theory is discussed in the next chapter)

t The language of quantum theory is precise but tiicky Quantum theory does not state that something—like light for example[1]—can be wave like and particle

It makes no sense to ask which of these is really the true description Both of them are required for a complete understanding

In 1922, Werner Heisenberg, as a student, asked his professor and friend-to-be, Niels Bohr, "If the inner structure of the atom is as closed to descriptive accounts as you say, if we really lack a language for dealing with it, how can we ever hope to understand atoms'*'

Bohr hesitated for a moment and then said, I think we may yet be able to do so But in the process we may have to learn what the word understanding really means '6

In human terms it means that the same person can be good *and* evil, bold *and* timid a lion *and* a lamb

All of the above notwithstanding, particle physicists of necessity analyze subatomic particles as if they *were* like little baseballs that fly through space and collide with each other When a particle physicist studies a track on a bubble-chamber photograph of a particle interaction, he assumes that it *was* made by a little moving object and that the other tracks on the photograph likewise were made by small moving objects In fact, particle interactions are analyzed in much the same terms that can be applied to the collision of billiard balls Some particles collide (and are annihilated in the process) and other newly created particles come flying out of the collision area In short, particle interactions are analyzed essentially in terms of masses velocities, and momenta These are the concepts of Newtonian physics and they also apply to automobiles and streetcars

Phvsicists do this because they have to use these concepts if they are to communicate at all What is available to them is usually a black photograph with white lines on it They know that, (1) according to quantum theory, subatomic particles have no independent existence of their own, (2) subatomic particles have wave-like characteristics as well as particle-like characteristics, and (3) subatomic particles actually may be manifestations of interacting fields Nonetheless, these white

---

like *at the same tune* According to Bohr s comptementai ity (page 93) light reveals either a particle like aspect 01 a wave like aspect depending upon the context i e the experiment It is not possible to observe both the wave like aspect and the particle like aspect in the same situation However *both of these mutually exclusive (complementary) aspects are needed to understand hsht* In this sense light is both particle like and wave like lines (more patterns) lend themselves to analysis in classical terms, and so that is how particle physicists analyze them

This dilemma, the dilemma of having to talk in classical terms about phenomena which cannot be described in classi cal concepts is the basic paradox of quantum mechanics It pervades every part of it It is like trying to explain an LSD experience We try to use familiar concepts as points of departure, but beyond that, the familiar concepts do not fit the phenomena The alternative is to say nothing at all

'Physicists who deal with the quantum theory, wrote Heisenberg,

> are also compelled to use a language taken from ordinary life We act as if there really were such a thing as an electric current [or a particle] because, if we forbade all physicists to speak of electric current [or particles] they could no longer express their thoughts [7]

Therefore, physicists talk about subatomic particles as if they were real little objects that leave tracks in bubble chambers and have an independent (objective) existence This convention has been extremely productive Over the last forty years almost one hundred particles have been discovered They constitute what Kenneth Ford calls the particle zoo *

[redacted]

4

WHILE QUANTUM PHYSICS THUS DEFINES with great accuracy the mathematical relationships governing the basic units of radiation and matter, it seems to obscure our picture of the true nature of both. Most modern physicists, however, consider it rather naïve to speculate about the true nature of anything. They are "positivists"—or "logical empiricists"—who contend that a scientist can do no more than report his observations. And so if he performs two experiments with different instruments and one seems to reveal that light is made up of particles and the other that light is made up of waves, he must accept both results, regarding them not as contradictory but as complementary. By itself neither concept suffices to explain light, but together they do. Both are necessary to describe reality and it is meaningless to ask which is really true. For in the abstract lexicon of quantum physics there is no such word as "really."

It is futile, moreover, to hope that the invention of more delicate tools may enable man to penetrate much farther into the microcosm. There is an indeterminacy about all the events of the atomic universe which refinements of measurement and observation can never dispel. The element of caprice in atomic behavior cannot be blamed on man's coarse-grained implements. It stems from the very nature of things, as shown by Heisenberg in 1927 in a famous statement of physical law known as the "Principle of Uncertainty." To illustrate his thesis Heisenberg pictured an imaginary experiment in which a physicist attempts to observe the position and velocity* of a moving electron by using an immensely powerful supermicroscope. Now, as has already been suggested, an individual electron appears to have no definite position or velocity. A physicist can define electron behavior accurately enough so long as he is dealing with great numbers of them. But when he tries to locate a particular electron in space the best he can say is that a certain point in the complex superimposed wave motions of the electron group represents the *probable* position of the electron in question. The individual electron is a blur—as indeterminate as the wind or a sound wave in the night—and the fewer the electrons with which the physicist deals, the more indeterminate his findings. To prove that this indeterminacy is a symptom not of man's immature science but of an ultimate barrier of nature, Heisenberg presupposed that the imaginary microscope used by his imaginary physicist is optically capable of magnifying by a hundred billion diameters—i.e., enough to bring an object the size of an electron within the range of human visibility. But now a further difficulty is encountered. For inasmuch as an electron

* In physics the term "velocity" connotes direction as well as speed.

is smaller than a light wave, the physicist can "illuminate" his subject only by using radiation of shorter wave length. Even X-rays are useless. The electron can be rendered visible only by the high-frequency gamma rays of radium. But the photoelectric effect, it will be recalled, showed that photons of ordinary light exert a violent force on electrons; and X-rays knock them about even more roughly. Hence the impact of a still more potent gamma ray would prove disastrous.

The Principle of Uncertainty asserts therefore that it is impossible with any of the principles now known to science to determine the position and the velocity of an electron at the same time—to state confidently that an electron is "right here at this spot" and is moving at "such and such a speed." For by the very act of observing its position, its velocity is changed; and, conversely, the more accurately its velocity is determined, the more indefinite its position becomes. And when the physicist computes the mathematical margin of uncertainty in his measurements of an electron's position and velocity he finds it is always a function of that mysterious quantity—Planck's Constant, *h*.

* * *

Quantum physics thus appears to shake two pillars of the old science, causality and determinism. For by dealing in terms of statistics and probabilities it abandons all idea that nature exhibits an inexorable sequence of cause and effect between individual happenings. And by its admission of margins of uncertainty it yields up the ancient hope that science, given the present state and velocity of every material body in the universe, can forecast the history of the universe for all time. One by-product of this surrender is a new argument for the existence of free will. For if physical events are indeterminate and the future is unpredictable, then perhaps the unknown quantity called "mind" may yet guide man's destiny among the infinite uncertainties of a capricious universe. Another conclusion of greater scientific importance is that in the evolution of quantum physics the barrier between man, peering dimly through the clouded windows of his senses, and whatever objective reality may exist has been rendered almost impassable. For whenever he attempts to penetrate and spy on the "real" objective world, he changes and distorts its workings by the very process of his observation. And when he tries to divorce this "real" world from his sense perceptions he is left with nothing but a mathematical scheme. He is indeed somewhat in the position of a blind man trying to discern the shape and texture of a snowflake. As soon as it touches his fingers or his tongue it dissolves. A wave electron, a photon, a wave of probability, cannot be visualized; they are simply symbols useful in expressing the mathematical relationships of the microcosm.

To the question, why does modern physics employ such esoteric methods of description, the physicist answers: because the equations of quantum physics define more accurately than any mechanical model the fundamental phenomena beyond the range of vision. In short, *they work,* as the calculations which hatched the atomic bomb spectacularly proved. The aim of the practical physicist, therefore, is to enunciate the laws of nature in ever more precise mathematical terms. Where the nineteenth century physicist envisaged electricity as a fluid and, with this metaphor in mind, evolved the laws that generated our present electrical age, the twentieth century physicist tries to avoid metaphors. He knows that electricity is not a physical fluid, and he knows that such pictorial concepts as "waves" and "particles," while serving as guideposts to new discovery, must not be accepted as accurate representations of reality. In the abstract language of mathematics he can describe how things behave though he does not know—or need to know—what they are.

Yet there are present-day physicists to whom the void between science and reality presents a challenge. Einstein more than once expressed the hope that the statistical method of quantum physics would prove a temporary expedient. "I cannot believe," he wrote, "that God plays dice with the world." He repudiated the positivist doctrine that science can only report and correlate the results of observation. He believed in a universe of order and harmony. And he believed that questing man may yet attain a knowledge of physical reality. To this end he looked not within the atom, but outward to the stars, and beyond them to the vast drowned depths of empty space and time.

Is the biological landscape an interactive design?

EARTH AS A SPACE SHIP

by Kenneth E. Boulding

May 10, 1965
Washington State University
Committee on Space Sciences

In the imagination of those who are sensitive to the realities of our era, the earth has become a space ship, and this, perhaps, is the most important single fact of our day. For millennia, the earth in men's minds was flat and illimitable. Today, as a result of exploration, speed, and the explosion of scientific knowledge, earth has become a tiny sphere, closed, limited, crowded, and hurtling through space to unknown destinations. This change in man's image of his home affects his behavior in many ways, and is likely to affect it much more in the future.

It is not only that man's image of the earth has changed; the reality of the world social system has changed. As long as man was small in numbers and limited in technology, he could realistically regard the earth as an infinite reservoir, an infinite source of inputs and an infinite cesspool for outputs. Today we can no longer make this assumption. Earth has become a space ship, not only in our imagination but also in the hard realities of the social, biological, and physical system in which man is enmeshed. In what we might call the "old days," when man was small in numbers and earth was large, he could pollute it with impunity, though even then he frequently destroyed his immediate environment and had to move on to a new spot, which he then proceeded to destroy. Now man can no longer do this; he must live in the whole system, in which he must recycle his wastes and really face up to the problem of the increase in material entropy which his activities create. In a space ship there are no sewers.

Let me suggest, then, some of the consequences of earth becoming a space ship. In the first place, it is absolutely necessary for man now to develop a technology that is different from the one on which he now bases his high-level societies. High-level societies are now based on the consumption of fossil fuels and ores, none of which, at present rates of consumption, are likely to last more than a few hundred years. A stable, circular-flow high-level technology is conceivable in which we devote inputs of energy to the concentration of materials into useful form, sufficient to compensate for the diffusion of materials which takes place in their use. At the moment we take fuels and burn them, we take concentrated deposits of iron ore for instance, and phosphates, and we spread these throughout the world in dumps, and we flush them out to the oceans in sewers. The stable high-level technology will have to rely on the oceans and the atmosphere as a basic resource from which materials may be concentrated in sufficient quantity to overcome their diffusion through consumption. Even this, of course, will require constant inputs of energy. There is no way for the closed system to prevent the increase of entropy. Earth, fortunately, has a constant input of energy from the sun, and by the time that goes, man will probably have abandoned earth; and we have also the possibility of almost unlimited energy inputs from nuclear fusion, if we can find means of harnessing it usefully.

Man is finally going to have to face the fact that he is a biological system living in an ecological system, and that his survival power is going to depend on his developing symbiotic relationships of a closed-cycle character with all the other elements and populations of the world of ecological systems. What this means, in effect, is that all the other forms of life will have to be domesticated, even if on wildlife preserves.

The consequences of earth becoming a space ship for the social system are profound and little understood. It is clear that much human behavior and many human institutions in the past, which were appropriate to all infinite earth, are entirely inappropriate to a small closed space ship. We cannot have cowboys and Indians, for instance, in a space ship, or even a cowboy ethic. We cannot afford unrestrained conflict, and we almost certainly cannot afford national sovereignty in an unrestricted sense. On the other hand, we must beware of pushing the analogy too far. In a small ship, there would almost have to be a dictatorial political system with a captain, and a planned economy. A voyaging space ship, like a battleship, almost has to be a centrally planned economy. A large space ship with three billion passengers, however, or perhaps ten billion, may have a very different social structure. Large social organizations are very different from small. It may be able to have much more individual freedom, a price system and a market economy of a limited and controlled kind, and even democratic political institutions. There must be, however, cybernetic or homeostatic mechanisms for preventing the overall variables of the social system from going beyond a certain range. There must, for instance, be machinery for controlling the total numbers of the population; there must be machinery for controlling conflict processes and for preventing perverse social dynamic processes of escalation and inflation. One of the major problems of social science is how to devise institutions which will combine this overall homeostatic control with individual freedom and mobility. I believe this problem to be not insoluble, though not yet solved.

Once we begin to look at earth as a space ship, the appalling extent of our ignorance about it is almost frightening. This is true of the level of every science. We know practically nothing, for instance, about the long-run dynamics even of the physical system of the earth. We do not understand, for instance, the machinery of ice ages, the real nature of geological stability or disturbance, the incidence of volcanism and earthquakes, and we understand fantastically little about that enormously complex heat engine known as the atmosphere. We do not even know whether the activities of man are going to make the earth warm up or cool off. At the level of the biological sciences, our ignorance is even greater. Ecology as a science has hardly moved beyond the level of bird-watching. It has yet to become quantified, and it has yet to find an adequate theory. Even to an economist, its existing theoretical structures seem fantastically naive, and when it comes to understanding the world social system or the sociosphere, we are not only ignorant but proud of our ignorance. There is no systematic method of data collection and processing, and the theory of social dynamics is still in its first infancy.

The moral of all this is that man must be made to realize that all his major problems are still unsolved, and that a very large and massive intellectual effort is still necessary to solve them. In the meantime we are wasting our intellectual resources on insoluble problems like unilateral national defense and on low-priority achievements like putting a man on the moon. This is no way to run a space ship.

Kenneth E. Boulding Papers, Archives (Box # 38), University of Colorado at Boulder Libraries.

4
Collective Wisdom, Slime-Mold-Style

4.1 Slime Time

It is the spring of 1973, and the weather has been unseasonably wet. As you gaze out the window into your yard, your eye is caught by a proliferation of deep yellow blob-like masses. What could they be? Puzzled, you return to work but are unable to settle down. A while later you return to the window. The yellow jelliform masses are still in evidence, but you would swear they have moved. You are right. The newcomers are slowly but surely creeping around your yard, climbing up the nearby telephone pole—moving in on you. In a panic, you phone the police to report a likely sighting of alien life forms in the USA. In fact, what you (and many others) saw was a fully terrestrial being, but one whose life cycle is alien indeed: *Fuligo septica*, a type of acellular slime mold. [1]

Slime molds come in many varieties[2] and sizes., but all belong to the class of Mycetozoa. The name is revealing, combining 'mycet' (fungus) and 'zoa' (animal). They like moist surroundings and are often found on rotting logs, tree stumps, or piles of decaying plant matter. They are widely distributed geographically, and do not seem bound to specific climates. As one handbook puts it, "many species are apt to pop up most anywhere, unexpectedly" (Farr 1981, p. 9).

Of special interest is the life cycle of the "cellular" slime mold. Take, for instance, the species *Dictyostelium discoideum*,[3] first discovered in 1935 in North Carolina. The life cycle of *D. discoideum* begins with a so-called vegetative phase, in which the slime-mold cells exist individually, like amoeba (they are called *myxamoebae*). While local food sources last (the myxamoebae feed on bacteria) the cells grow and divide. But when

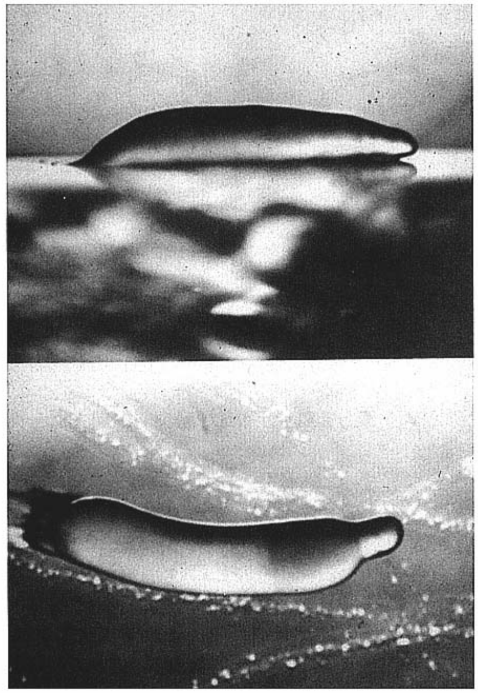


Figure 4.1 Migrating shrugs (pseudoplasmodia) of acellular slime mold. Source: Morrissey 1982. Used by permission of Academic Press.

food sources run out, a truly strange thing happens. The cells begin to cluster together to form a tissue-like mass called a *pseudoplasmodium*. The pseudoplasmodium, amazingly, is a mobile collective creature—a kind of miniature slug (figure 4.1)—that can crawl along the ground. [4] It is attracted to light, and it follows temperature and humidity gradients. These cues help it to move toward a more nourishing location. Once such a spot is found, the pseudoplasmodium changes form again, this time differentiating into a stalk and a fruiting body—a spore mass comprising about two-thirds of the cell count. When the spores are propagated, the cycle begins anew with a fresh population of myxamoebae.

How do the individual slime-mold cells (the myxamoebae) know to cluster? One solution—the biological analogue of a central planner (see chapter 3)—would be for evolution to have elected "leader cells." Such cells would be specially adapted so as to "call" the other cells, probably by chemical means, when food ran low. And they would somehow orchestrate the construction of the pseudoplasmodium. It seems, however, that nature has chosen a more democratic solution. In fact, slime-mold cells look to behave rather like the ants described in section 2.3. When food runs low, each cell releases a chemical (cyclic AMP) which attracts other cells. As cells begin to cluster, the concentrations of cyclic AMP increases, thus attracting yet more cells. A process of positive feedback thus leads to the aggregation of cells that constitutes a pseudoplasmodium. The process is, as Mitchel Resnick (1994, p. 51) notes, a nice example of what has become known as *self-organization*. A self-organizing system is one in which some kind of higher-level pattern emerges from the interactions of multiple simple components without the benefit of a leader, controller, or orchestrator.

The themes of self-organization and emergence are not, I shall suggest, restricted to primitive collectives such as the slime mold. Collectives of human agents, too, exhibit forms of emergent adaptive behavior. The biological brain, which parasitizes the external world (see chapter 3) so as to augment its problem-solving capacities, does not draw the line at inorganic extensions. Instead, the collective properties of groups of individual agents determine crucial aspects of our adaptive success.

4.2 Two Forms of Emergence

There are at least two ways in which new phenomena can emerge (without leaders or central controllers) from collective activity. The first, which I will call *direct emergence*, relies largely on the properties of (and relations between) the individual elements, with environmental conditions playing only a background role. Direct emergence can involve multiple homogeneous elements (as when temperature and pressure emerge from the interactions between the molecules of a gas), or it can involve heterogeneous ones (as when water emerges from the interactions between hydrogen and oxygen molecules). The second form of emergence, which I will call *indirect emergence*, relies on the interactions of individual elements but requires that these interactions be mediated by active and often quite

complex environmental structures. The difference thus concerns the extent to which we may understand the emergence of a target phenomenon by focusing largely on the properties of the individual elements (direct emergence), versus the extent to which explaining the phenomenon requires attending to quite specific environmental details. The distinction is far from absolute, since all phenomena rely to some extent on background environmental conditions. (It can be made a little more precise by casting it in terms of the explanatory roles of different kinds of "collective variables"—see chapter 6). But we can get a working sense of the intuitive difference by looking at some simple cases.

A classic example of direct emergence is the all-too-familiar phenomenon of the traffic jam. A traffic jam can occur even when no unusual external event (such as a collision or a broken set of traffic lights) is to blame. For example, simple simulations recounted by Mitchel Resnick [5] show that bunching will occur if each car obeys just two intuitive rules: "If you see another car close ahead, slow down; if not, speed up (unless you are already moving at the speed limit)" (Resnick 1994, pp. 69, 73). Why, given just these two rules and no external obstacles, doesn't the traffic simply accelerate to the speed limit and stay there? The answer lies in the initial placements. At the start of the simulation, the cars were spaced randomly on the road. Thus, sometimes one car would start close to another. It would soon need to slow down, which would cause the car behind it to slow, and so on. The upshot was a mixture of stretches of fast-moving traffic and slow-moving jams. Every now and then a car would leave the jam, thus freeing space for the one behind it, and accelerate away. But as fast as the jam "unraveled" in one direction, it grew in the other direction as new cars reached the backmarkers and were forced to slow. Although each car was moving forward, the traffic jam itself, considered as a kind of higher-order entity, was moving backward! The higher-order structure (which Resnick calls the *collective structure*) was thus displaying behavior fundamentally different from the behavior of its components. Indeed, the individual components kept changing (as old cars left and new ones joined), but the integrity of the higher-order collective was preserved. (In a similar fashion, a human body does not comprise the same mass of matter over time—cells die and are replaced by new ones built out of energy from food. We, too, are higher-order collectives whose constituting matter is in constant flux.) Traffic jams count as cases of direct emergence because the necessary environmental backdrop (varying distances between cars) is quite minimal—random spacing is surely the default condition and requires no special environmental manipulations. The case of indirect emergence, as we shall now see, is intuitively quite different.

Consider the following scenario: You have to remember to buy a case of beer for a party. To jog your memory, you place an empty beer can on your front doormat. When next you leave the house, you trip over the can and recall your mission. You have thus used what is by now a familiar trick (recall chapter 3)—exploiting some aspect of the real world as a partial substitute for on-board memory. In effect, you have used an alteration to your environment to communicate something to yourself. This trick of using the environment to prompt actions and to communicate signals figures in many cases of what I am calling indirect emergence.

Take the nest-building behavior of some termites. A termite's building behavior involves modifying its local environment in response to the triggers provided by previous alterations to the environment—alterations made by other termites or by the same termite at an earlier time. Nest building is thus under the control of what are known as *stigmergic algorithms*. [6]

A simple example of stigmergy is the construction of arches (a basic feature of termite nests) from mudballs. Here is how it works[7]: All the termites make mud balls, which at first they deposit at random. But each ball carries a chemical trace added by the termite. Termites prefer to drop their mudballs where the chemical trace is strongest. It thus becomes likely that new mudballs will be deposited on top of old ones, which then generate an even stronger attractive force. (Yes, it's the familiar story!) Columns thus form. When two columns are fairly proximal, the drift of chemical attractants from the neighboring column influences the dropping behavior by inclining the insects to preferentially add to the side of each column that faces the other. This process continues until the tops of the columns incline together and an arch is formed. A host of other stigmergic affects eventually yield a complex structure of cells, chambers, and tunnels. At no point in this extended process is a plan of the nest represented or followed. No termite acts as a construction leader. No termite "knows" anything beyond how to respond when confronted with a specific patterning of its local environment. The termites do not talk to one another in any way, except through the environmental products of their own

activity. Such environment-based coordination requires no linguistic encoding or decoding and places no load on memory, and the "signals" persist even if the originating individual goes away to do something else (Beckers et al. 1994, p. 188).

To sum up: We learn important lessons from even these simple cases of emergent collective phenomena. Such phenomena can come about in either direct or highly environmentally mediated ways. They can support complex adaptive behaviors without the need for leaders, blueprints, or central planners. And they can display characteristic features quite different in kind from those of the individuals whose activity they reflect. In the next section, we see these morals in a more familiar, human guise.

4.3 Sea and Anchor Detail

In the most successful and sustained investigation of the cognitive properties of human groups to date, Edwin Hutchins—anthropologist, cognitive scientist, and open-ocean racing sailor and navigator—has described and analyzed the role of external structures and social interactions in ship navigation. Here is his description of how some of the necessary tasks are performed and coordinated (Hutchins 1995, p. 199; my note):

> In fact, it is possible for the [navigation] team to organize its behavior in an appropriate sequence without there being a global script or plan anywhere in the system. [8] Each crew member only needs to know what to do when certain conditions are produced in the environment. An examination of the duties of members of the navigation team shows that many of the specified duties are given in the form "Do X when Y." Here are some examples from the procedures:

A. Take soundings and send then to the bridge on request.

B. Record the time and sounding every time a sounding is sent to the bridge.

C. Take and report bearings to the objects ordered by the recorder and when ordered by the recorder.

Each member of the navigation team, it seems, need follow only a kind of stigmergic[9] procedure, waiting for a local environmental alteration (such as the placing of a specific chart on a desk, the arrival of a verbal request, or the sounding of a bell) to call forth a specific behavior. That behavior, in turn, affects the local environment of certain other crew members and calls forth further bursts of activity, and so on until the job is done.

Of course, these are *human* agents, who will form ideas and mental models of the overall process. And this general tendency, Hutchins observes, makes for a more robust and flexible system, since the individuals can monitor one another's performance (e.g., by asking for a bearing that has not been supplied on time) and, if need be (say, if someone falls ill), try to take over aspects of other jobs. Nonetheless, no crew member will have internalized all the relevant knowledge and skills.

Moreover, a large amount of work is once again done by external structures: nautical slide rules, alidades, bearing record logs, hoeys, charts, fathometers, and so on. [10] Such devices change the nature of certain computational problems so as to make them more tractable to perceptual, pattern-completing brains. The nautical slide rule, Hutchins's favorite example, turns complex mathematical operations into scale-alignment operations in physical space.[11]

Finally, and again echoing themes from chapter 3, the navigational work space itself is structured so as to reduce the complexity of problem solving. For example, the charts that will be used when entering a particular harbor are preassembled on a chart table and are laid one on top of the other in the order of their future use (the first-needed on top).

All these factors, Hutchins argues, unite to enable the overall system of artifacts, agents, natural world, and spatial organization to solve the problem of navigation. The overall (ship-level) behavior is not controlled by a detailed plan in the head of the captain. The captain may set the goals, but the sequence of information gatherings and information transformations which implement the goals need not be explicitly represented anywhere. Instead, the computational power and expertise is spread across a heterogeneous assembly of brains, bodies, artifacts, and other external structures. Thus do pattern-completing brains navigate the unfriendly and mathematically demanding seas.

4.4 The Roots of Harmony

But how does such delicate harmonization of brains, bodies, and world come about? In the cases of what I have called direct emergence the problem is less acute, for here the collective properties are determined directly by the mass action of some uniform individual propensity. Thus, if

nature were (heaven forbid) to evolve cars and roads, then (given random initial distribution and the two rules rehearsed in section 4.2) traffic jams would immediately result.

Indirect emergence presents a superficially greater puzzle. In these cases, the target property (e.g., a termite nest or successful navigation of a ship) emerges out of multiple and often varied interactions between individuals and a complexly structured environment. The individuals are apparently built or designed so that the coupled dynamics of the agents and these complex environments yield adaptive success. No single individual, in such cases, needs to know an overall plan or blueprint. Yet the total system is, in a sense, well designed. It constitutes a robust and computationally economical method of achieving the target behavior. How does such design come about?

For the nervous systems of the individual termites, an important part of the answer [12] is clearly "through evolution." Hutchins suggests that a kind of quasi-evolutionary process may be at work in a navigation team too. The key feature is simply that small changes occur without prior design activity, and these changes tend to be preserved according to the degree to which they enhance biological success. Evolutionary change thus involves the gradual accretion of small "opportunistic" changes: changes which themselves alter the "fitness landscape" for subsequent changes both within the species and in other species inhabiting the same ecosystem.

Now, still following Hutchins, consider the case in which some established cognitive collective (such as a navigation team) faces a new and unexpected challenge. Suppose that this challenge calls for a fast response, so there is no time for the group to meet and reflect on how best to cope.[13] How, under such conditions, is the group to discover a new social division of labor that responds to the environmental demand? What actually happens, Hutchins shows, is that each member of the group tries to fulfill the basic functions necessary to keep the ship from going aground, but in so doing each member constrains and influences the activity of the others in what amounts to a collective, parallel search for a new yet computationally efficient division of labor. For example, one crew member realizes that a crucial addition must be performed but does not have enough time. That crew member therefore tells a nearby person to add the numbers. This in turn has effects further down the line. The solution to the problem of averting disaster emerges as a kind of equilibrium point in an iterated series of such local negotiations concerning task distribution—an equilibrium point that is determined equally by the skills of the individuals and the timing and sequence of incoming data. No crew member reflects on any overall plan for redistributing the tasks. Instead, they all do what each does best, negotiating whatever local help and procedural changes they need. In such cases there is a fast, parallel search for a coherent collective response, but the search does not involve any explicit and localized representation of the space of possible global solutions. In this sense, as Hutchins notes, the new solution is found by a process more akin to evolutionary adaptation than to global rationalistic design.

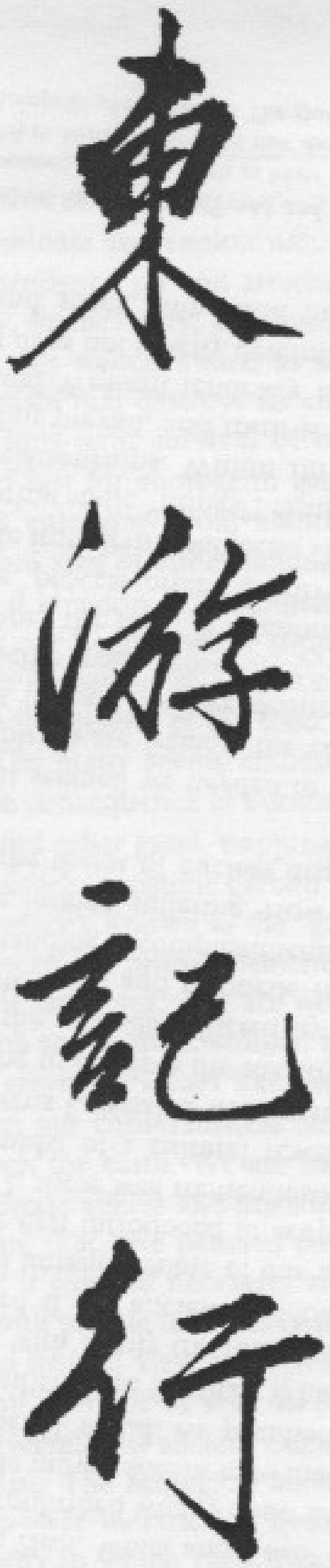
Here is a somewhat simpler version of the same idea [14] : Imagine that your task is to decide on an optimum placement of footpaths to connect a complex of already-constructed buildings (say, on a new university campus). The usual strategy is global rationalistic design, in which an individual or a small group considers the uses of the various buildings, the numbers of pedestrians, etc. and seeks some optimal pattern of linkages reflecting the patterns of likely use. An alternative solution, however, is to open the campus for business without any paths, and with grass covering all the spaces between buildings. Over a period of months, tracks will begin to emerge. These will reflect both the real needs of the users and the tendency of individuals to follow emerging trails. At the end of some period of time the most prominent trails can be paved, and the problem will have been solved without anyone's needing to consider the global problem of optimal path layout or needing to know or represent the uses of all the various buildings. The solution will have been found by means of an interacting series of small individual calculations, such as "I need to get from here to the refectory—how shall I do it?" and "I need to get to the physics lab as fast as possible—how shall I do it?" The overall effect of these multiple local decisions is to solve the global problem in a way that looks more like a kind of evolution than like classical, centralized design.

The need to account for the origins of collective success does not, it seems, force us back to the image of a central planning agency that knows the shape of the overall problem space. Instead, we may sometimes structure our own problem-solving environment as a kind of by-product of our basic problem-solving activity. On our hypothetical campus, the early

walkers structure the environment as a by-product of their own actions, but subsequent walkers will then encounter a structured environment that may help them, in turn, to solve the very same problems. [15]

4.5 Modeling the Opportunistic Mind

These first few chapters have, I hope, conveyed a growing sense of the opportunistic character of much of biological cognition. For example: faced with the heavy time constraints on real-world action, and armed only with a somewhat restrictive, pattern-completing style of on-board computation, the biological brain takes all the help it can get. This help includes the use of external physical structures (both natural and artifactual), the use of language and cultural institutions (see also chapters 9 and 10 below), and the extensive use of other agents. To recognize the opportunistic and spatiotemporally extended nature of real problem solving is, however, to court a potential methodological nightmare. How are we to study and understand such complex and often non-intuitively constructed extended systems?



Our Journey East:
Exploring Foreign Territory

CHAPTER ONE

Our Journey East: Exploring Foreign Territory

Lying motionless, gazing at a chart on the wall showing streams of force connecting the little toe with the corner of the eye in a web of continuous loops, I feel my breath soften and my vision sharpen.

Delicate pins protrude from my elbow, ankle, and knee. My arms and legs are flooded by tiny rivulets of current. It's not like a hypodermic needle that injects a foreign substance—what I'm feeling is simply more of myself. It's strange, in the sense of odd and unfamiliar, but as a sensation not unpleasant. In fact, the edges of my mouth are cradling a silly smile.

The skin is stretched less tightly over my bony frame as my pores relax. I sense movement as I lie quiet, aware of impulse within my mind at rest. Thoughts tumble into consciousness, roll over, and shuffle off.

With my eyes closed, and perception directed inward, simultaneous layers of activity play like instruments in concert with each other. I am the composer and the composed, the musician and the listener, the instrument and its player.

SUBTLE YET PALPABLE, MY INITIAL ENCOUNTER WITH ACUPUNCTURE LEFT me tantalized by mystery and promise. Mystery, in that tiny needles could extend my field of awareness and completely alter the state of my being. Promise, in that by burrowing into the conceptual soil of this system, I could deepen my self-understanding.

As the daughter of a surgeon and the granddaughter of two surgeons, my early life was steeped in the cauldron of medicine brewed over several generations. Enthusiasm for healing was contagious, and I became infected. As a child I was impressed by my father's devotion and satisfaction. He rushed to the hospital day or night to operate on a man lacerated in a motorcycle accident or a child threatened by a ruptured appendix. Lives would have been lost without his heroic intervention.

The role of doctor and the appeal of medicine came naturally—but why Chinese medicine? The ideology of Chinese medicine immediately captivated me by its stark contrast to the perspective of Western medical science. I had never been comfortable thinking of myself in my father's language of electrolytes and blood-gas ratios, a collection of quantities and statistics. The Chinese medical vocabulary contained metaphors from nature like *Wood, Fire, Earth, Metal,* and *Water, Heat, Wind,* and *Cold.** This cosmological description of human process confirmed what I knew intuitively to be so—that what moves the world outside moves within me—that subject and object are two aspects of one phenomenal world. As peculiarly outside my cultural context as it was, Chinese medicine felt familiar. What enticed me even more than my sense of continuity with family tradition was the affinity I felt with its concepts, and I wondered if the ancient wisdom embedded within its construction of reality could untangle some of our modern predicaments.

When Efrem and I were first introduced to acupuncture at a seminar at Esalen in the spring of 1972, there was tremendous upheaval in the world. The Chinese were in the midst of a cultural revolution, and so were we. During the sixties the concerns I wrestled with were more social than medical. Many of us were seeking to antidote the toxicity of racism in the American social body and heal the wounds inflicted by a decade of violence in Vietnam. I struggled to understand and reconcile how Western civilization, having achieved some outstanding accomplishments, could so often contribute to rather than alleviate human suffering. How could it perpetuate vast environmental insult and the threat of nuclear disaster and yet be building a better future?

To remake the world it seemed we needed to rethink it. After all, solutions depend on how problems are framed, the context within which they exist. At issue for me was in part how we defined reality—and the reality assumed by Chinese medicine made sense.

Chinese medicine echoes the logic of quantum physics, which suggests that we exist in a relative, process-oriented universe in which there is no "objective" world separable from living subjects. The essential questions cannot be resolved by measuring static "things"; rather, answers become stories about interactions and relationships. Within this paradigm contradictions are not only sanctioned but prevail, and truth is purely contextual. In contrast with our conventional Western tendency to draw sharp lines of distinction, Chinese thought does not strictly determine the boundaries between rest and motion, time and space, mind and matter, sickness and health.

*Specially defined Chinese medical terms are italicized and capitalized throughout the text to differentiate them from their common meaning, such as heart as an anatomically located organ in the chest versus *Heart* as an *Organ Network,* or wind that dances through the trees versus Wind that represents an adverse pathological phenomenon. The glossary will assist in clarification.

Chinese medicine transcends the illusion of separation by inhabiting the reality of a unified field, an interwoven pattern of inseparable links in a circular chain.

The Ancient Chinese Worldview

The ancient Chinese perceived human beings as a microcosm of the universe that surrounded them, suffused with the same primeval forces that motivated the macrocosm. They imagined themselves as part of one unbroken wholeness, called Tao, a singular relational continuum within and without. This thinking predates the dissection of mind from body and man from nature that Western culture performed in the seventeenth century.

From when I first encountered it, this fusing of the mind with the body and people with their world had a magic, seductive lure. I was eager to imagine myself as integrated and interrelated, an active participant in a system in the continuous process of transforming itself. I learned that Chinese philosophers and physicians have studied nature over thousands of years, divining how to interact with it to cultivate and guide *Qi* (pronounced "chee"), life's animating force and substance. The Chinese concept of *Qi* symbolizes life in all its forms: thoughts and emotions, tissue and blood, inner life and outer expression. I wanted to reap and share the benefits of their knowledge.

I was startled to find that my father considered Chinese medicine quackery. Even though I could appreciate that surgery and antibiotics were often necessary, my father could not entertain the value of acupuncture and herbs. Though I could appreciate the worth of a microscope to identify tissue alterations and bacteria, my father could not fathom the notion of *Qi*. It probably should not have surprised me, since Chinese medicine does not rest solely upon the tangible, measurable phenomena upon which his medicine relies. That I was drawn to a system outside his Western conceptual framework unsettled and offended him. Yet I recognized the Chinese view as inclusive, not exclusive. Within that view the subtle, ephemeral, and invisible were as significant as what could be seen, touched, and counted, and caring for the human spirit was no less essential or real than caring for the human structure, nor separate from it.

The Modern Predicament

Western philosophy was more akin to Eastern until the Renaissance of the seventeenth century, when our civilization revolutionized its thinking. It was then that the scientific philosophy of the 1600s began to consider people as independent of the living systems that surround them and assumed that we could dominate and exploit nature without being affected by it. We escaped from dependency on and attachment to the natural world, pursuing invulnerability, invincibility, and immortality. Four hundred years later many of us regret this stance, aware, as anthropologist Gregory Bateson puts it, that "an organism that destroys its environment destroys itself."

The long-term survival of our species has been placed at risk by an unbridled lust for short-term gains coupled with a false postulation of civilization's autonomy from nature. Human beings have befouled their own nest, a sure sign of either madness or disease when observed in the rest of the animal kingdom. Contamination of rivers and seas puts fish in peril and threatens our potable water supply, the ozone layer is nibbled by fluorocarbons, rain forests perish so that cattle can graze, soil depletion and pesticides diminish the quality of our food, urban horizons shimmer with heavy ocher air . . . The litany seems endless. The disruption of our global respiration seems the consequence of burning fossil fuels (which fill the air with carbon dioxide and other gases, trapping heat) and of denuding forests (whose trees release oxygen, consume carbon dioxide, and attract water vapor to cool the earth's surface). Known as the "greenhouse effect," this upset in the relationship between the primal elements of heat and cold, fire and water, produces a fever of land, sea, and air with disastrous repercussions—upon weather conditions and marine, animal, and human life as we know it.

The gravity of issues like this has motivated burgeoning numbers of us to re-form our consciousness about who we are and how to live with each other upon the earth. We are being obliged to see how that which appears to be separate affects and fits within a pattern, and we are seeing, as Bateson points out, that "the patterns connect." Reluctantly we are entertaining the idea that it may be necessary to transcend thinking of ourselves solely as entities with private interests (nation-states, corporate bodies, individual persons) and instead view ourselves and our world as one organic system.

Today many of us seek to reclaim the sense of connectedness that existed universally in ancient culture, when human fate was wholly entwined with nature. The ecological assumption within these cosmologies held that all things were inextricably bound together. The world was seen as a symbiotic entity in which each living system interacted with and mutually supported every other. The growing popularity of Chinese medicine, which embodies these values, is itself part of a pattern. Chinese medicine instructs us to perceive the way the world functions and re-create harmony within the context of the whole.

TAOIST PHILOSOPHY IN ACTION

Protecting human life by preserving the conditions within which it thrives is the purpose of Chinese medicine. Each of us is pictured as an ecosystem as well as living within one. The balance of forces within us (*Yin-Yang, Heat-Cold, Blood-Qi*) determines our internal climate, our health or disease.

Chinese medicine embraces the logic that the best remedy for calamity is to avert it—the best cure for sickness is prevention. The *Nei Jing*, a medical classic written in the second century B.C., states:

> Maintaining order rather than correcting disorder is the ultimate principle of wisdom. To cure disease after it has appeared is like digging a well when one already feels thirsty, or forging weapons after the war has already begun.

The true physician teaches the Tao—how to live. Traditional Chinese doctors are trained to cultivate wellness as well as to correct ill health. Planning ahead, Chinese medicine knows that storms interrupt clear weather, that illness stalks and gains a foothold when we are vulnerable. Its strategy is to enable us to withstand the storm without becoming disabled by it and to accumulate resources in times of good weather, peace, and plenty.

The technology of Chinese medicine—simple, inexpensive, and highly portable—was what first ignited Efrem's imagination. A traditional doctor in China need carry only a few needles and gather local herbs from the countryside to minister to his patients. Because this medicine was so accessible, after the revolution in 1949 many thousands of "barefoot doctors" were trained to serve the unmet needs of the Chinese people for medical care. This equipped ordinary people with the tools to gain control of their own lives.

In contrast with this model, American medical doctors are assigned the active (powerful) roles while patients are resigned to passive (powerless) ones. Efrem felt a kinship with the Taoist physician who performed skillfully to mitigate distress but also acted as teacher, sharing knowledge and power. Doctor and patient then engaged in the mutual endeavor of grasping the problem and accomplishing the healing.

Sages describe a state without suffering and a way to get there. That way is not to search for a single remedy, a panacea, a "magic bullet," but to engage in the ongoing process of learning to become more animated, more connected: more charged with life. Both Efrem and I were struck by the notion that Chinese medicine was about just that—striving toward greater integration through the cultivation of *Qi*. The excitement was and remains that Chinese medicine offers an approach that includes, yet moves beyond, issues of physical health.

By the summer of 1972, at a time when few Americans had heard of acupuncture, and at the risk of being considered cultural renegades, we struggled to cross over ethnic, social, and intellectual barriers, riveted by the idea that with Chinese medicine we could both understand the world and change it. Just as my younger sister was finishing medical school, we were exploring foreign territory, making our journey east.

THE NATURE OF REALITY SHIFTS

After the acupuncture seminar at Esalen, the intrigue escalated. Efrem and I read what we could, but it wasn't enough. Like hounds on the scent of a fox, we were pulled irresistibly by the magnet of Chinese philosophy and driven to discover whether these ancient healing techniques really worked. At that time no schools offered study programs in the United States, and China had not yet been opened to American students. So we traveled to England in the fall of 1972 to enroll in the College of Traditional Chinese Acupuncture.

We were part of the first class of Americans to study at this college. Half of our group were Western medical doctors. We all went through major readjustments during our course of study as our presumptions about the nature of reality, ourselves, health, and healing were pulled apart, stretched, and rewoven into new cloth.

I began to experience myself as a little tao, a galaxy of internal planets spinning through time and space. These planets became named as five transformative phases: *Wood*, *Fire*, *Earth*, *Metal*, and *Water*. Each one functioned as an energetic field linked with the visceral organs of *Liver*, *Heart*, *Spleen*, *Lung*, and *Kidney*, whose lines of force coursed through channels in my body like a lattice of waterways crosses the earth. Physiological functions, personality characteristics, even single impulses could be classified within this new language.

We came to recognize that each class member could be characterized by the attributes of one *Organ* and *Phase*. For example, an enthusiastic cardiologist became easily excited, had a nervous laugh, and liked to give running commentary about his experience as he lived it. We identified him as expressing the influence of *Fire*. A methodical and serious graduate student who dressed fastidiously and set high standards for herself and others reflected the influence of *Metal*. A dynamic, impatient psychologist who took the initiative to set up the school's first training clinic exhibited typical *Wood* characteristics. A socially reticent, intense nonconformist who plunged into the learning, absorbing all he could in his fertile mind, manifested a strong *Water* aspect. And an internist able to heal with his compassionate touch and smile, who assumed care of us when the stress of our studies took its toll, had the [illegible]

Just as I could see myself as a little tao, our class too became a projection of a single body. We each performed our own function, a distinct role appropriate to our individual nature. Our physical shape, the quality of our movement and manner, our symptoms, the tastes and flavors we preferred, the odor of our flesh, the color of our tongue, the sound of our voice, the feeling of our pulse, revealed ways our persona organized and expressed itself within the context of these new categories.

THE REALITY OF MEDICINE SHIFTS

We experienced a radical shift of focus from what we'd known. We learned to feel the pulse in six positions on the radial artery of both wrists, each position corresponding to a channel associated with an *Organ Network*. This was particularly dramatic for the doctors. Since medical school they had accepted as an article of faith and an assumption of science the existence of only one pulse, which reflected the function of the heart alone. The notion that the shape, force, and rhythm of the blood pulsing through the radial artery was not merely determined by the activity of the heart, but was the outcome of the interaction of all forces coalescing in the human body at any given time, diverged sharply from their medical education.

Alongside our training, we were experiencing acupuncture for our own maladies. In the midst of a particularly intense week of uninterrupted study, Efrem felt flushed, agitated, and had the unpleasant sensation of his pulse pounding throughout his body. His blood pressure was abnormally high, 140/90. Our teacher inserted one needle along the outer edge of his hand and another along the front border of his neck muscle. The change was striking—within minutes his blood pressure normalized, the pounding ceased, and he felt an expansive calm. What had been a pervasive mood of anxiety and melancholy became one of exhilaration and repose. He experienced acupuncture as simultaneously alleviating the distress in his body and restoring his good humor. This imprinted upon him that within this medical system, mental discomforts need not be ignored nor invalidated as being "only in your mind."

CHANGING DEFINITIONS

Although I had always considered myself to be in perfect health, my self-image was chipped away by the chisel of this new perspective. As a child, whenever I had a sniffle my father gave me a powerful pill to squelch it, and as a result, I never got sick. Later, even though my stamina was capricious to worry about high blood pressure. In my family, as long as you weren't sick, you were healthy. But bothersome aspects of my physical and emotional life that I'd previously dismissed as idiosyncratic began to appear as lifelong signs of a pattern that betrayed less than optimal function. It dawned on me that although I was not sick, I could feel better.

For years my hands and feet were ice blocks under the covers at night and my lower back ached periodically. Cuts became easily infected, and frequent canker sores burned in my mouth, aggravated by the sweets that I craved. Sometimes a lingering hunger persisted after meals, and in the evenings when I was tired, my hearing diminished to the hum of a dim tuner. Restful sleep eluded me if I remained awake past eleven P.M., and rising the next morning could seem an Olympian feat. Yet more limiting than any of these minor body afflictions was a nagging sense of fatigue. There was constantly more on my agenda than I had energy to accomplish. In the early afternoon I could be easily overcome by a muddled feeling that left me sluggish, unable to focus, and unproductive for several hours of prime time during the work day. I felt hopeful and optimistic but lacked the steady strength to fulfill my ambitions.

I had never assembled these vague complaints. My father would have dismissed them since they didn't make particular sense within his framework, and besides, there was nothing in his repertoire to "fix" any of them. From his point of view I was normal. But within the context of Chinese medicine, each sign and symptom assumed new meaning.

Although Western lab tests would have indicated normal kidney function, the coldness, tiredness, and hearing loss represented a weakness of *Water*, the *Kidney Network*. The sores in my mouth and surge of energy late at night were characteristic of an imbalance of *Wood*, the *Liver Network*. Sleeping problems were associated with a deficiency of *Fire*, the *Heart Network*. Overall, my character was most dominated by *Earth*, the *Spleen Network*. A sunny disposition, an eagerness to help other people, and constant reflection upon how to generally improve the human condition are traits of *Earth*. A fixation with details, a sometimes obsessive need to get things done, and that sudden soggy, muddled feeling are associated with disharmonies of *Earth*.

As I received acupuncture treatment, and later herbal medicine, my resources increased and steadied; in addition, I could relieve my monthly irritability, water retention, and insomnia. As my complaints receded, I experienced myself as a more effective captain of my vessel, guiding fate by altering the givens, assuming control over tendencies and reactions that I had presumed came with my territory.

All of us established relationships between physical and emotional patterns, connecting what had otherwise seemed unrelated. In Western medicine, cold extremities, back pain, hearing loss, and fatigue are not associated with each other or with the kidney. Troubled sleep, excitability, and laughter are not associated with the heart. Menstrual problems are unrelated to the liver. But in Chinese medicine, a weak *Kidney* may mean not only having a backache, poor hearing, and fatigue, but also being fearful, stingy, withdrawn, and apathetic. The *Organ Networks* of Chinese medicine were not equivalent to the anatomical organs of Western medicine. *Organ Networks* were defined functionally, not structurally. As such, they included emotional and mental function as well as physiological performance, and even their physiology was defined differently.

Furthermore, within this logic, problems are not isolated from the context in which they occur, so a single symptom is understood in relation to the entire body environment. The strategy is to adjust the field, reorganizing the process out of which symptoms grow, coercing them into retreat by eliminating the milieu in which they thrive.

Stimulated and unsettled, many of us underwent personal as well as professional transformation. The sun still rose each morning and traded places with the moon at night, but our perception of ourselves and the world around us altered.

After reading these chapters, you will also learn how to interpret yourself within the context of these concepts so that you can say, "I am an *Earth* type who has conflict with *Wood* and weakness in *Water*. My primary therapeutic focus will be in these areas." That is a purpose of this book.

THE CROSS-FERTILIZATION OF MEDICINE: EAST GOES WEST AND WEST GOES EAST

From the late Middle Ages on, traditional Chinese medicine trickled into foreign lands. A comprehensive Chinese herbal text written in 1587 found its way into the hands of German physicians around 1605. In 1810 a French physician named Louis Berlioz reported successfully treating neurogenic diseases with acupuncture. In 1929, after serving as French consul in China, Dr. Soulie de Mourant translated a Chinese text, furthering the practice of acupuncture in France.

During the mid-1800s Chinese immigrants building railroads and mining gold in the United States brought their medicine with them. For the most part it remained within the Chinese community. However, a few exceptions occurred in the Northwest, where local settlers benefited from the care of Chinese doctors. Historical records from the towns of John Day, Oregon, and Boise, Idaho, indicate that Chinese doctors successfully treated respiratory, digestive, and reproductive infections, as well as arthritis and symptoms of cardiovascular disease. Pioneer women who were unable to conceive were helped to bear healthy children.

Meanwhile, European and American missionaries had been exporting Western medicine to China. The Chinese were very impressed with surgical technique, vaccinations, and the means of hygiene and sanitation that helped reduce the spread of infectious diseases, wound sepsis, and the mortality of women following childbirth. Western industrial and medical technology came to be associated with progress. By the time penicillin was developed in the 1940s, traditional medical institutions in China had been significantly replaced by the more highly prized Western medical schools.

In 1949, the year of the Chinese revolution, there were 10,000 Western-trained doctors for 400 million people. At this time the 550,000 traditional doctors recovered their stature within the medical establishment because of the dire need for more doctors and a fervent desire to reassert faith in China's own cultural legacy. Since 1949 traditional medicine has been revived and incorporated as a mandatory part of the training in many modern medical schools. Researchers there now investigate the biochemistry of herbal constituents and the electrophysiology of acupuncture points to trace their impact on the healing mechanisms of the body.

Still, in many life-threatening situations, only Western medical and surgical techniques are appropriate. Functional disorders constitute a middle ground where either system may be elected. In chronic degenerative illnesses for which Western medicine has little to offer, traditional medicine is a welcome option. Some Chinese hospitals accommodate alternate floors for Western and traditional care, with doctors selecting what best suits their patients' needs. In some circumstances both systems work in concert. For instance, Western surgery, radiation, or chemotherapy is used to control cancer, while traditional Chinese medicine is used to strengthen the body and enhance immunity, enabling the person to better tolerate treatment.

During the 1950s and 1960s when traditional medicine was being reintegrated in its homeland, travel, commerce, and exchange between the United States and China was highly restricted. In 1972, the year when then-President Nixon unstrapped the political restraints upon our relationship with the People's Republic, acupuncture needles passed from East to West. At this time our class of Americans at the College of Acupuncture was preparing to use these needles on American bodies.

While in China, *New York Times* journalist James Reston had an emergency appendectomy, and his postoperative abdominal pain was relieved by an acupuncture needle in his leg. It appeared particularly spectacular because it was so little understood. Realizing that Qi traverses the body through channels, it becomes less enigmatic that toothaches can be relieved by the stimulation of a point along a pathway that connects the mouth to the hand

or foot. Learning that white fungus reduces blood cholesterol levels, it no longer seems a miracle that people who live in regions in China where it is part of the diet have lower incidences of heart attacks. It is common knowledge among Chinese that pseudoginseng powder stops hemorrhage, chrysanthemum flower tea will relieve a bilious headache, and black mushrooms increase immunity.

Since the dawn of this century the West has been exporting medical technology and pharmaceuticals to the East. In the last quarter of this century the medicinal treasures of the Orient have been imported to the West. In 1976 Efrem and I were among the first two hundred and ten practitioners to receive an acupuncture license from the California State Board of Medical Quality Assurance. Nationally there are now between six and ten thousand practitioners of Chinese medicine, and two dozen other states have instituted licensure programs. Two decades ago there were no colleges of Chinese medicine in this country; now there are more than twenty. What began as a handful of books on the subject has grown to a library of more than four hundred titles, many of which are translations of Chinese texts.

Miracles East and West

The advantages of integrating Western and Chinese medicine were dramatized for Efrem and me following the birth of our son in 1975. Born with two gaping holes in a swollen heart, Bear's life would have been eclipsed without surgical intervention. After diagnosing his problem by injecting a dye into his heart that projected an image on a video screen, our cardiologist determined it was possible for the holes to be closed with Dacron patches. If he had been born twenty years earlier when open-heart surgery for infants had not yet been developed, his heart could not have been repaired. We were grateful that miracles existed in the West.

Following surgery, because of his fragile constitution, he was subject to countless bouts of bronchitis and upper-respiratory infection. In San Francisco we studied herbs with local Chinese doctors. When Bear was able to swallow herbal pills, they became incorporated into his routine along with breakfast and lunch. We used certain formulas for overcoming acute illness and others as a tonic for strengthening him, improving his resistance. Eventually he rebuffed the onslaught of winter colds even when his schoolmates succumbed. Now as a teenager Bear radiates the health we had hoped for from his birth. We were grateful for miracles of the East as well.

In 1980, after eight years of practicing acupuncture, Efrem went to China to further study the use of herbs and experience Chinese medicine in its native habitat. A Western-style dispensary stood alongside the herbal pharmacy at the Kunming Railway Workers Hospital. A special kitchen brewed individual herb teas, delivering fresh medicine to both inpatients and outpatients.

Efrem's teacher, Dr. Zheng Wen Tao, was first trained as a Western physician. Later he was asked to return to school to study traditional medicine. He combined classical knowledge with the findings of modern research, producing herbal formulas particularly effective for treating cardiovascular and immune disorders. Dr. Zheng Wen Tao was particularly renowned for his treatment of intractable diseases. Efrem witnessed the gradual remission of scleroderma in a young woman after one year of herbal treatment. Scleroderma, an autoimmune disease, is marked by the progressive hardening of connective tissue, considered both fatal and untreatable by Western medicine. This young woman had been told that she had one year to live. Instead, during that year she reversed the course of her illness and regained health. Zheng Wen Tao is the epitome of the modern Chinese medical hero, able to unite modern science with classical tradition. When Efrem returned home, we expanded our use of herbal medicine in the treatment of our acupuncture patients, setting up an herbal pharmacy in our acupuncture clinic.

A Foreign Culture Planted on Home Soil

Some people tend to dismiss Chinese medicine for not conforming to the modern medical model, while others tend to romanticize and mystify it. Some discount it because it is not modern, others revere it because it is old. The former group considers it eccentric quackery, and the latter considers it a panacea. It is clear to us that it is a system unto itself, altogether distinct from a trendy health fad or other therapies considered to be holistic or alternative. It has endured for centuries not only because its techniques produce tangible results, but because it embodies a coherent philosophy that integrates many aspects of human life.

Western medicine has had a monopoly not only on our medicine cabinets, but also on our minds. Now other possibilities exist. More and more people are receiving acupuncture, so far several million Americans, and Chinese herbs are finding the way onto their shelves. Within the next decade medicinal foods will be simmering in stainless-steel pots, making the exotic aromas and flavors of China as familiar to us as chicken soup. Americans, like the Chinese, are quite pragmatic—Chinese medicine meets their need, so people try it; it helps, so they come back for more.

By familiarizing yourself with its concepts, you will be able to judge if

and when it is a fitting option for you. Spotting yourself within its categories will help you better understand your behavior and feelings, the ways you express your health, and the ways you fall ill. By the end, we hope you will be able to interpret yourself amply enough to acquire the rudimentary know-how to initiate personal change.

How is interactivity structured in a networked system?

The End of Books

By ROBERT COOVER

In the real world nowadays, that is to say, in the world of video transmissions, cellular phones, fax machines, computer networks, and in particular out in the humming digitalized precincts of avant-garde computer hackers, cyberpunks and hyperspace freaks, you will often hear it said that the print medium is a doomed and outdated technology, a mere curiosity of bygone days destined soon to be consigned forever to those dusty unattended museums we now call libraries. Indeed, the very proliferation of books and other print-based media, so prevalent in this forest-harvesting, paper-wasting age, is held to be a sign of its feverish moribundity, the last futile gasp of a once vital form before it finally passes away forever, dead as God.

Which would mean of course that the novel, too, as we know it, has come to its end. Not that those announcing its demise are grieving. For all its passing charm, the traditional novel, which took center stage at the same time that industrial mercantile democracies arose -- and which Hegel called "the epic of the middle-class world" -- is perceived by its would-be executioners as the virulent carrier of the patriarchal, colonial, canonical, proprietary, hierarchical and authoritarian values of a past that is no longer with us.

Much of the novel's alleged power is embedded in the line, that compulsory author-directed movement from the beginning of a sentence to its period, from the top of the page to the bottom, from the first page to the last. Of course, through print's long history, there have been countless strategies to counter the line's power, from marginalia and footnotes to the creative innovations of novelists like Laurence Sterne, James Joyce, Raymond Queneau, Julio Cortazar, Italo Calvino and Milorad Pavic, not to exclude the form's father, Cervantes himself. But true freedom from the tyranny of the line is perceived as only really possible now at last with the advent of hypertext, written and read on the computer, where the line in fact does not exist unless one invents and implants it in the text.

"Hypertext" is not a system but a generic term, coined a quarter of a century ago by a computer populist named Ted Nelson to describe the writing done in the nonlinear or nonsequential space made possible by the computer. Moreover, unlike print text, hypertext provides multiple paths between text segments, now often called "lexias" in a borrowing from the pre-hypertextual but prescient Roland Barthes. With its webs of linked lexias, its networks of alternate routes (as opposed to print's fixed unidirectional page-turning) hypertext presents a radically divergent technology, interactive and polyvocal, favoring a plurality of discourses over definitive utterance and freeing the reader from domination by the author. Hypertext reader and writer are said to become co-learners or co-writers, as it were, fellow-travelers in the mapping and remapping of textual (and visual, kinetic and aural) components, not all of which are provided by what used to be called the author.

Though used at first primarily as a radically new teaching arena, by the mid-1980's hyperspace was drawing fiction writers into its intricate and infinitely expandable, infinitely alluring webs, its green-limned gardens of multiple forking paths, to allude to another author popular with hypertext buffs, Jorge Luis Borges.

Several systems support the configuring of this space for fiction writing. Some use simple randomized linking like the shuffling of cards, others (such as Guide and HyperCard) offer a kind of do-it-yourself basic tool set, and still others (more elaborate systems like Storyspace, which is currently the software of choice among fiction writers in this country, and Intermedia, developed at Brown University) provide a complete package of sophisticated structuring and navigational devices.

Although hypertext's champions often assail the arrogance of the novel, their own claims are hardly modest. You will often hear them proclaim, quite seriously, that there have been three great events in the history of literacy: the invention of writing, the invention of movable type and the invention of hypertext. As hyperspace-walker George P. Landow puts it in his recent book surveying the field, "Hypertext": "Electronic text processing marks the next major shift in information technology after the development of the printed book. It promises (or threatens) to produce effects on our culture, particularly on our literature, education, criticism and scholarship, just as radical as those produced by Gutenberg's movable type."

Noting that the "movement from the tactile to the digital is the primary fact about the contemporary world," Mr. Landow observes that, whereas most writings of print-bound critics working in an exhausted technology are "models of scholarly solemnity, records of disillusionment and brave sacrifice of humanistic positions," writers in and on hypertext "are downright celebratory. . . . Most poststructuralists write from within the twilight of a wished-for coming day; most writers of hypertext write of many of the same things from within the dawn."

Dawn it is, to be sure. The granddaddy of full-length hypertext fictions is Michael Joyce's landmark "Afternoon," first released on floppy disk in 1987 and moved into a new Storyspace "reader," partly developed by Mr. Joyce himself, in 1990.

Mr. Joyce, who is also the author of a printed novel, "The War Outside Ireland: A History of the Doyles in North America With an Account of their Migrations," wrote in the on-line journal Postmodern Culture that hyperfiction "is the first instance of the true electronic text, what we will come to conceive as the natural form of multimodal, multisensual writing," but it is still so radically new it is hard to be certain just what it is. No fixed center, for starters -- and no edges either, no ends or boundaries. The traditional narrative time line vanishes into a geographical landscape or exitless maze, with beginnings, middles and ends being no longer part of the immediate display. Instead: branching options, menus, link markers and mapped networks. There are no hierarchies in these topless (and bottomless) networks, as paragraphs, chapters and other conventional text divisions are replaced by evenly empowered and equally ephemeral window-sized blocks of text and graphics -- soon to be supplemented with sound, animation and film.

As Carolyn Guyer and Martha Petry put it in the opening "directions" to their hypertext fiction "Izme Pass," which was published (if "published" is the word) on a disk included in the spring 1991 issue of the magazine Writing on the Edge:

"This is a new kind of fiction, and a new kind of reading. The form of the text is rhythmic, looping on itself in patterns and layers that gradually accrete meaning, just as the passage of time and events does in one's lifetime. Trying the textlinks embedded within the work will bring the narrative together in new configurations, fluid constellations formed by the path of your interest. The difference between reading hyperfiction and reading traditional printed fiction may be the difference between sailing the islands and standing on the dock watching the sea. One is not necessarily better than the other."

I must confess at this point that I am not myself an expert navigator of hyperspace, nor am I -- as I am entering my seventh decade and thus rather committed, for better or for worse, to the obsolescent print technology -- likely to engage in any major hypertext fictions of my own. But, interested as ever in the subversion of the traditional bourgeois novel and in fictions that challenge linearity, I felt that something was happening out (or in) there and that I ought to know what it was: if I were not going to sail the Guyer-Petry islands, I had at least better run to the shore with my field glasses. And what better way to learn than to teach a course in the subject?

Thus began the Brown University Hypertext Fiction Workshop, two spring semesters (and already as many software generations) old, a course devoted as much to the changing of reading habits as to the creation of new narratives.

Writing students are notoriously conservative creatures. They write stubbornly and hopefully within the tradition of what they have read. Getting them to try out alternative or innovative forms is harder than talking them into chastity as a life style. But confronted with hyperspace, they have no choice: all the comforting structures have been erased. It's improvise or go home. Some frantically rebuild those old structures, some just get lost and drift out of sight, most leap in fearlessly without even asking how deep it is (infinitely deep) and admit, even as they paddle for dear life, that this new arena is indeed an exciting, provocative if frequently frustrating medium for the creation of new narratives, a potentially revolutionary space, capable, exactly as advertised, of transforming the very art of fiction, even if it now remains somewhat at the fringe, remote still, in these very early days, from the mainstream.

With hypertext we focus, both as writers and as readers, on structure as much as on prose, for we are made aware suddenly of the shapes of narratives that are often hidden in print stories. The most radical new element that comes to the fore in hypertext is the system of multidirectional and often labyrinthine linkages we are invited or obliged to create. Indeed the creative imagination often becomes more preoccupied with linkage, routing and mapping than with statement or style, or with what we would call character or plot (two traditional narrative elements that are decidedly in jeopardy). We are always astonished to discover how much of the reading and writing experience occurs in the interstices and trajectories between text fragments. That is to say, the text fragments are like stepping stones, there for our safety, but the real current of the narratives runs between them.

"The great thing," as one young writer, Alvin Lu, put it in an on-line class essay, is "the degree to which narrative is completely destructed into its constituent bits. Bits of information convey knowledge, but the juxtaposition of bits creates narrative. The emphasis of a hypertext (narrative) should be the degree to which the reader is given power, not to read, but to organize the texts made available to her. Anyone can read, but not everyone has sophisticated methods of organization made available to them."

The fictions developed in the workshop, all of which are "still in progress," have ranged from geographically anchored narratives similar to "Our Town" and choose-your-own-adventure stories to parodies of the classics, nested narratives, spatial poems, interactive comedy, metamorphic dreams, irresolvable murder mysteries, moving comic books and Chinese sex manuals.

IN hypertext, multivocalism is popular, graphic elements, both drawn and scanned, have been incorporated into the narratives, imaginative font changes have been employed to identify various voices or plot elements, and there has also been a very effective use of formal documents not typically used in fictions -- statistical charts, song lyrics, newspaper articles, film scripts, doodles and photographs, baseball cards and box scores, dictionary entries, rock music album covers, astrological forecasts, board games and medical and police reports.

At our weekly workshops, selected writers display, on an overhead projector, their developing narrative structures, then face the usual critique of their writing, design, development of character, emotional impact, attention to detail and so on, as appropriate. But they also engage in continuous on-line dialogue with one another, exchanging criticism, enthusiasm, doubts, speculations, theorizing, wisecracks. So much fun is all of this, so compelling this "downright celebratory" experience, as Mr. Landow would have it, that the creative output, so far anyway, has been much greater than that of ordinary undergraduate writing workshops, and certainly of as high a quality.

In addition to the individual fictions, which are more or less protected from tampering in the old proprietary way, we in the workshop have also played freely and often quite anarchically in a group fiction space called "Hotel." Here, writers are free to check in, to open up new rooms, new corridors, new intrigues, to unlink texts or create new links, to intrude upon or subvert the texts of others, to alter plot trajectories, manipulate time and space, to engage in dialogue through invented characters, then kill off one another's characters or even to sabotage the hotel's plumbing. Thus one day we might find a man and woman encountering each other in the hotel bar, working up some kind of sexual liaison, only to return a few days later and discover that one or both had sex changes. During one of my hypertext workshops, a certain reading tension was caused when we found that there was more than one bartender in our hotel: was this the same bar or not? One of the students -- Alvin Lu again -- responded by linking all the bartenders to Room 666, which he called the "Production Center," where some imprisoned alien monster was giving birth to full-grown bartenders on demand.

This space of essentially anonymous text fragments remains on line and each new set of workshop students is invited to check in there and continue the story of the Hypertext Hotel. I would like to see it stay open for a century or two.

However, as all of us have discovered, even though the basic technology of hypertext may be with us for centuries to come, perhaps even as long as the technology of the book, its hardware and software seem to be fragile and short-lived; whole new generations of equipment and programs arrive before we can finish reading the instructions of the old. Even as I write, Brown University's highly sophisticated Intermedia system, on which we have been writing our hypertext fictions, is being phased out because it is too expensive to maintain and incompatible with Apple's new operating-system software, System 7.0. A good portion of our last semester was spent transporting our documents from Intermedia to Storyspace (which Brown is now adopting) and adjusting to the new environment.

This problem of operating-system standards is being urgently addressed and debated now by hypertext writers; if interaction is to be a hallmark of the new technology, all its players must have a common and consistent language and all must be equally empowered in its use. There are other problems too. Navigational procedures: how do you move around in infinity without getting lost? The structuring of the space can be so compelling and confusing as to utterly absorb and neutralize the narrator and to exhaust the reader. And there is the related problem of filtering. With an unstable text that can be intruded upon by other author-readers, how do you, caught in the maze, avoid the trivial? How do you duck the garbage? Venerable novelistic values like unity, integrity, coherence, vision, voice seem to be in danger. Eloquence is being redefined. "Text" has lost its canonical certainty. How does one judge, analyze, write about a work that never reads the same way twice?

And what of narrative flow? There is still movement, but in hyperspace's dimensionless infinity, it is more like endless expansion ; it runs the risk of being so distended and slackly driven as to lose its centripetal force, to give way to a kind of static low-charged lyricism -- that dreamy gravityless lost-in-space feeling of the early sci-fi films. How does one resolve the conflict between the reader's desire for coherence and closure and the text's desire for continuance, its fear of death? Indeed, what is closure in such an environment? If everything is middle, how do you know when you are done, either as reader or writer? If the author is free to take a story anywhere at any time and in as many directions as she or he wishes, does that not become the obligation to do so?

No doubt, this will be a major theme for narrative artists of the future, even those locked into the old print technologies. And that's nothing new. The problem of closure was a major theme -- was it not? -- of the "Epic of Gilgamesh" as it was chopped out in clay at the dawn of literacy, and of the Homeric rhapsodies as they were committed to papyrus by technologically innovative Greek literati some 26 centuries ago. There is continuity, after all, across the ages riven by shifting technologies.

Much of this I might have guessed -- and in fact did guess -- before entering hyperspace, before I ever picked up a mouse, and my thoughts have been tempered only slightly by on-line experience. What I had not clearly foreseen, however, was that this is a technology that both absorbs and totally displaces. Print documents may be read in hyperspace, but hypertext does not translate into print. It is not like film, which is really just the dead end of linear narrative, just as 12-tone music is the dead end of music by the stave.

Hypertext is truly a new and unique environment. Artists who work there must be read there. And they will probably be judged there as well: criticism, like fiction, is moving off the page and on line, and it is itself susceptible to continuous changes of mind and text. Fluidity, contingency, indeterminacy, plurality, discontinuity are the hypertext buzzwords of the day, and they seem to be fast becoming principles, in the same way that relativity not so long ago displaced the falling apple.

Osama bin Laden and The Advent of Netwar

John Arquilla and David Ronfeldt are directors of the "Networks and Netwars" project sponsored by the office of the US assistant secretary of defense for command, control, communications and intelligence. this article is an excerpt of their introduction to that forthcoming report.

The "age of networks," now dawning with such promise, has just yielded an astounding "attack on America," heralding the onset of an archetypal netwar of the darkest kind. Transnational terrorists organized in widely dispersed, networked nodes have shown how it is possible to swarm together swiftly, on cue, then pulse to the attack simultaneously. They relied on the Internet, communicating via encrypted messages-sometimes even embedding them in photographic and other images on the world-wide Web. But what really distinguishes them-particularly Osama bin Laden's al-Qaeda ("the Base")-is the highly networked organizational form that they have built, based on their unusual social, religious, and kinship ties. US Secretary of State Colin Powell has put it aptly: To win against terror, this network must be "ripped apart."

The league of hierarchical nation-states forming to fight this terrorism will have to build its own set of nimble networks. In the military realm, this means relying more on networks of agile special forces (e.g., the US's Delta Force; Britain's Special Air Service; France's Commando Hubert; and Germany's Grenzschutzgruppe-Neun) than on the missiles, tanks, bombers and aircraft carriers that, until now, have been the sine qua non of national power. Just as the terrorists' power derives more from their organizational form than from technology, so too must the military power to defeat them become more reliant upon organization and doctrine than upon advanced technical systems.

The intelligence world faces an equally urgent need for institutional redesign-away from notions of "central" intelligence, toward the construction of transnational intelligence networks able to share what they have on a real-time basis. Swift movement of important information has played a major role in the success of networked businesses over the past decade. Now it is time for networking to redefine the approach to intelligence-the quality and timeliness of which will determine whether bin Laden's or any other terror network can indeed be "ripped apart."

Improved international networking among military and intelligence organizations can help win this war against terror. But this will not suffice in the long run. A balanced strategy for countering networked terror should also involve a much improved capacity to work with networks of civil-society NGOs around the world, many of which are engaged in social netwars to advance human rights, pressure authoritarian regimes, and foster ethical norms of behavior. Nurturing this emergent global civil society offers the best chance to create an "integral security system" that could free all of us, ultimately, from terror. For in a truly networked world, joined together by common values rather than just common "wires," there will simply be little space left for such a scourge.

Above all, US strategy should avoid getting mired in a "clash of civilizations." The war against terror is not a war of Western values against Islam. Rather, it is a "time war," in this case between an emerging global civilization of the 21st century and a xenophobic religious fanaticism of the 14th century (or earlier). Osama bin Laden and his cohorts are tribal, medieval, absolutist and messianic. The more clearly terrorists are revealed as such, the sooner they will be rejected by the vast majority *of the Muslim world for which they purport to be fighting.*

Los Angeles - The information revolution is altering the nature of conflict across the spectrum. We call attention to two developments in particular. First, this revolution is favoring and strengthening network forms of organization, often giving them an advantage over hierarchical forms. The rise of networks means that power is migrating to nonstate actors, because they are able to organize into sprawling multiorganizational networks (especially "all-channel" networks, in which every node is connected to every other node) more readily than can traditional, hierarchical, state actors. This means that conflicts may increasingly be waged by "networks," perhaps more than by "hierarchies." It also means that whoever masters the network form stands to gain the advantage.

Second, as the information revolution deepens, the conduct and outcome of conflicts increasingly depend on information and communications. More than ever before, conflicts revolve around "knowledge" and the use of "soft power." Adversaries are learning to emphasize "information operations" and "perception management"-that is, media-oriented measures that aim to attract or disorient rather than coerce, and that affect how secure a society, a military or other actor feels about its knowledge of itself and of its adversaries. Psychological disruption may become as important a goal as physical destruction.

These propositions cut across the entire conflict spectrum. Major transformations are thus coming in the nature of adversaries, in the type of threats they may pose, and in how conflicts can be waged. Information-age threats are likely to be more diffuse, dispersed, multidimensional, nonlinear and ambiguous than industrial-age threats. Metaphorically, then, future conflicts may resemble the Oriental game of Go more than the Western game of chess. The conflict spectrum will be remolded from end to end by these dynamics.

An illustrative case of netwar was the effort by Serbia's reformist Radio b-92, along with a supportive network of United States and European government agencies and NGOs, to broadcast its reportage back into Serbia over the Internet, after b-92's transmitters were shut down by the Milosevic regime in 1998 and again in 1999. For a seminal case of a worldwide netwar, one need look no further than the International Campaign to Ban Landmines. This unusually successful movement consists of a loosely internetted array of NGOs and governments, which rely heavily on the Internet for communications. Through the personage of one of its many leaders, Jody Williams, this netwar won a well-deserved Nobel peace prize.

DEFINING NETWAR | To be precise, the term netwar refers to an emerging mode of conflict (and crime) at societal levels, short of traditional military warfare, in which the protagonists use network forms of organization and related doctrines, strategies and technologies attuned to the information age. These protagonists are likely to consist of dispersed organizations, small groups and individuals who communicate, coordinate and conduct their campaigns in an internetted manner, often without a precise central command. Thus, netwar differs from modes of conflict and crime in which the protagonists prefer to develop formal, stand-alone, hierarchical organizations, doctrines and strategies as in past efforts, for example, to build centralized movements along Leninist lines. Thus, for example, netwar is about the Zapatistas more than the Fidelistas, Hamas more than the Palestine

Liberation Organization (PLO), the American Christian Patriot movement more than the Ku Klux Klan, and the Asian Triads more than the Cosa Nostra.

The term netwar is meant to call attention to the prospect that network-based conflict and crime will become major phenomena in the decades ahead. Various actors across the spectrum of conflict and crime are already evolving in this direction. This includes familiar adversaries who are modifying their structures and strategies to take advantage of networked designs-e.g., transnational terrorist groups, black-market proliferators of weapons of mass destruction (WMD), drug and other crime syndicates, fundamentalist and ethnonationalist movements, intellectual-property pirates, and immigration and refugee smugglers. Some urban gangs, back-country militias and militant single-issue groups in the US have also been developing netwar-like attributes. The netwar spectrum also includes a new generation of revolutionaries, radicals and activists who are beginning to create information-age ideologies, in which identities and loyalties may shift from the nation state to the transnational level of "global civil society." New kinds of actors, such as anarchistic and nihilistic leagues of computer-hacking "cyboteurs," may also engage in netwar.

Many-if not most-netwar actors will be nonstate, even stateless. Some may be agents of a state, but others may try to turn states into their agents. Also, a netwar actor may be both subnational and transnational in scope. Odd hybrids and symbioses are likely.

Furthermore, some bad actors (terrorist and criminal groups) may threaten US and other nations' interests, but other actors (NGO activists in Burma or Mexico) may not-indeed, some actors who at times turn to netwar strategies and tactics, such as the New York-based Committee to Protect Journalists (CPJ), may have salutary liberalizing effects. Some actors may aim at destruction, but more may aim mainly at disruption and disorientation. Again, many variations are possible.

The full spectrum of netwar proponents may thus seem broad and odd at first glance. But there is an underlying pattern that cuts across all variations: the use of network forms of organization, doctrine, strategy and technology attuned to the information age.

MORE ABOUT ORGANIZATIONAL DESIGN | In an archetypal netwar, the protagonists are likely to amount to a set of diverse, dispersed "nodes" who share a set of ideas and interests and who are arrayed to act in a fully internetted "all-channel" manner.

Networks come in basically three types or topologies:

--The chain or line network, as in a smuggling chain where people, goods or information move along a line of separated contacts, and where end-to-end communication must travel through the intermediate nodes.

-- The hub, star or wheel network, as in a franchise or a cartel where a set of actors is tied to a central (but not hierarchical) node or actor and must go through that node to communicate and coordinate with each other.

-- The all-channel or full-matrix network, as in a collaborative network of militant peace groups where everybody is connected to everybody else.

Each node may be an individual, a group, an organization, part of a group or organization, or even a state. The nodes may be large or small, tightly or loosely coupled, and inclusive or exclusive in membership. They may be segmentary or specialized-that is, they may look alike and engage in similar activities, or they may undertake a division of labor based on specialization. The boundaries of the network, or of any node included in it, may be well-defined, or blurred and porous in relation to the outside environment. Many variations are possible.

Each type may be suited to different conditions and purposes, and all three may be found among netwar-related adversaries-e.g. the chain in smuggling operations; the hub at the core of terrorist and criminal syndicates; and the all-channel type among militant groups that are highly internetted and decentralized. There may also be hybrids of the three types, with different tasks being organized around different types of networks. For example, a netwar actor may have an all-channel council or directorate at its core but use hubs and chains for tactical operations. There may also be hybrids of network and hierarchical forms of organization.

For example, traditional hierarchies may exist inside particular nodes in a network. Some actors may have a hierarchical organization overall but use network designs for tactical operations; other actors may have an all-channel network design overall but use hierarchical teams for tactical operations. Again, many configurations are possible, and it may be difficult for an analyst to discern exactly what type characterizes a particular network.

Of the three network types, the all-channel has been the most difficult to organize and sustain, partly because it may require dense communications. But it is the type that gives the network form its new, high potential for collaborative undertakings and that is gaining new strength from the information revolution. Pictorially, an all-channel netwar actor resembles a geodesic "Bucky ball" (named for Buckminster Fuller); it does not look like a pyramid. The organizational design is flat. Ideally, there is no single, central leadership, command or headquarters-no precise heart or head that can be targeted. The network as a whole (but not necessarily each node) has little to no hierarchy; there may be multiple leaders. Decision making and operations are decentralized, allowing for local initiative and autonomy. Thus the design may sometimes appear acephalous (headless) and at other times polycephalous (Hydra-headed).

The capacity of this design for effective performance over time may depend on the existence of shared principles, interests and goals-perhaps an overarching doctrine or ideology-which spans all nodes and to which the members subscribe in a deep way. Such a set of principles, shaped through mutual consultation and consensus-building, can enable members to be "all of one mind" even though they are dispersed and devoted to different tasks. It can provide a central ideational and operational coherence that allows for tactical decentralization. It can set boundaries and provide guidelines for decisions and actions so that the members do not have to resort to a hierarchy because "they know what they have to do."

The network design may depend on having an infrastructure for the dense communication of functional information. This does not mean that all nodes must be in constant communication; that may not make sense for a secretive, conspiratorial actor. But when communication is needed, the network's members must be able to disseminate information promptly and as broadly as desired within the network and to outside audiences.

CAVEATS ABOUT THE ROLE OF TECHNOLOGY | Netwar is a result of the rise of network forms of organization, which in turn is partly a result of the computerized information revolution. To realize its potential, a fully interconnected network requires a capacity for constant, dense information and communications flows, more so than do other forms of organization (e.g., hierarchies). This capacity is afforded by the latest information and communication technologies-cellular telephones, fax machines, electronic mail (e-mail), Web sites and computer conferencing. Such technologies are highly advantageous for netwar actors whose constituents are geographically dispersed.

But two caveats are in order. First, the new technologies, however enabling for organizational networking, are not absolutely necessary for a netwar actor. Older technologies, like human couriers, and mixes of old and new systems may do the job in some situations. The late Somali warlord, Mohamed Farah Aidid, for example, proved very adept at eluding those seeking to capture him while at the same time retaining full command and control over his forces by means of runners and drum codes. Similarly, the first Chechen War (1994-1996), which the Islamic insurgents won, made wide use of runners and old communications technologies like ham radios for battle management and other command and control functions. So, netwar may be waged in high-, low-, or no-tech fashion.

Second, netwar is not simply a function of "the Net"; it does not take place only in "cyberspace" or the "infosphere." Some battles may occur there, but a war's overall conduct and outcome will normally depend mostly on what happens in the "real world"-it will continue to be, even in information-age conflicts, generally more important than what happens in cyberspace or the info-sphere.

Netwar is not solely about Internet war (just as cyberwar is not just about "strategic information warfare"). Americans have a tendency to view modern conflict as being more about technology than organization and doctrine. In our view, this is a misleading tendency. For example, social netwar is more about a doctrinal leader like Subcomandante Marcos than about a lone, wild computer hacker like Kevin Mitnick.

SWARMING | This distinctive, often ad-hoc design has unusual strengths, for both offense and defense. On the offense, networks tend to be adaptable, flexible and versatile vis-à-vis opportunities and challenges. This may be particularly the case where a set of actors can engage in swarming. Little analytic attention has been given to swarming, which is quite different from traditional mass- and maneuver-oriented approaches to conflict. Yet swarming may become the key mode of conflict in the information age, and the cutting edge for this possibility is found among netwar protagonists.

Swarming is a seemingly amorphous, but deliberately structured, coordinated, strategic way to strike from all directions at a particular point or points, by means of a sustainable pulsing of force and/or fire, close-in as well as from stand-off positions. This notion of "force and/or fire" may be literal in the case of military or police operations, but metaphorical in the case of NGO activists, who may, for example, be blocking city intersections or emitting volleys of e-mails and faxes. Swarming will work best-perhaps it will only work-if it is designed mainly around the deployment of myriad, small, dispersed, networked maneuver units. Swarming occurs when the dispersed units of a network of small (and perhaps some large) forces converge on a target from multiple directions. The overall aim is sustainable pulsing-swarm networks must be able to coalesce rapidly and stealthily on a target, then sever and redisperse, immediately ready to recombine for a new pulse. The capacity for a "stealthy approach" suggests that, in netwar, attacks are more likely to occur in "swarms" than in more traditional "waves." The Chechen resistance to the Russian army and the Direct Action Network's operations in the anti-World Trade Organization "Battle of Seattle" both provide excellent examples of swarming behavior.

Swarming may be most effective, and difficult to defend against, where a set of netwar actors do not "mass" their forces, but rather engage in dispersion and "packetization" (for want of a better term). This means, for example, that drug smugglers can break large loads into many small packets for simultaneous surreptitious transport across a border, or that NGO activists, as in the case of the Zapatista movement, have enough diversity in their ranks to respond to any discrete issue that arises-human rights, democracy, the environment, rural development, whatever.

In terms of their defensive potential, networks tend to be redundant and diverse, making them robust and resilient in the face of attack. When they have a capacity for interoperability and shun centralized command and control, network designs can be difficult to crack and defeat as a whole. In particular, they may defy counter leadership targeting-a favored strategy in the drug war as well as in overall efforts to tamp organized crime in the United States. Thus, whoever wants to attack a network is limited-generally, only portions of a network can be found and confronted. Moreover, the deniability built into a network affords the possibility that it may simply absorb a number of attacks on distributed nodes, leading an attacker to believe the network has been harmed and rendered inoperable when, in fact, it remains viable and is seeking new opportunities for tactical surprise.

The difficulty of dealing with netwar actors deepens when the lines between offense and defense are blurred or blended. When blurring is the case, it may be difficult to distinguish between attacking and defending actions, particularly where an actor goes on the offense in the name of self-defense. For example, the Zapatista struggle in Mexico demonstrates anew the blurring of offense and defense. The blending of offense and defense will often mix the strategic and tactical levels of operations. For example, guerrillas on the defensive strategically may go on the offense tactically, as in the war of the mujahideen in Afghanistan during the 1980s, and in both recent Chechen wars with the Russians.

OPERATING IN THE SEAMS | The blurring of offense and defense reflects another feature of netwar (albeit one that is exhibited in many other policy and issue areas): It tends to defy and cut across standard boundaries, jurisdictions and distinctions between state and society, public and private, war and peace, war and crime, civilian and military, police and military, and legal and illegal. This makes it difficult if not impossible for a government to assign responsibility to any single agency-e.g. military, police or intelligence-to be in charge of responding.

Thus, the spread of netwar adds to the challenges facing the nation state in the information age. Its sovereignty and authority are usually exercised through bureaucracies in which issues and problems can be sliced up and specific offices can be charged with taking care of specific problems. In netwar, things are rarely so clear. A protagonist is likely to operate in the cracks and gray areas of a society, striking where lines of authority crisscross and the operational paradigms of politicians, officials, soldiers, police officers and related actors get fuzzy and clash. Moreover, where transnational participation is strong, a netwar's protagonists may expose a local government to challenges to its sovereignty and legitimacy by arousing foreign governments and business corporations to put pressure on the local government to alter its domestic policies and practices.

NETWORKS VERSUS HIERARCHIES: CHALLENGES FOR COUNTERNETWAR | Hierarchies have a difficult time fighting networks. There are examples of this across the conflict spectrum.

Some of the best are found in the failings of many governments to defeat transnational criminal cartels engaged in drug smuggling, as in Colombia. The persistence of religious revivalist movements, as in Algeria, in the face of unremitting state opposition, shows both the defensive and offensive robustness of the network form.

The Zapatista movement in Mexico, with its legions of supporters and sympathizers among local and transnational NGOs, shows that social netwar can put a democratizing autocracy on the defensive and pressure it to continue adopting reforms.

It takes networks to fight networks. Governments that want to defend against netwar may have to adopt organizational designs and strategies like those of their adversaries. This does not mean mirroring the adversary, but rather learning to draw on the same design principles that he has already learned about the rise of network forms in the information age. These principles depend to some extent on technological innovation, but mainly on a willingness to innovate organizationally and doctrinally, perhaps especially by building new mechanisms for interagency and multijurisdictional cooperation.

Whoever masters the network form first and best will gain major advantages. In these early decades of the information age, adversaries who are advanced at networking (be they criminals, terrorists or peaceful social activists, including ones acting in concert with states) are enjoying an increase in their power relative to state agencies. While networking once allowed them simply to keep from being suppressed, it now allows them to compete on more nearly equal terms with states and other hierarchically oriented actors. The histories of Hamas and of the Cali cartel illustrate this; so do the Zapatista movement in Mexico and the International Campaign to Ban Landmines.

Counternetwar may thus require very effective interagency approaches, which by their nature involve networked structures.

It is not necessary, desirable or even possible to replace all hierarchies with networks in governments. Rather, the challenge will be to blend these two forms skillfully, while retaining enough core authority to encourage and enforce adherence to networked processes. By creating effective hybrids, governments may become better prepared to confront the new threats and challenges emerging in the information age, whether generated by ethnonationalists, terrorists, militias, criminals or other actors.

Recent netwar conflicts feature an uneven split between those about globalist issues-aimed at fostering the rise of a rights- and ethics-based civil society-and the more frequent, somewhat darker "autonomist" variety of netwar, featuring nonstate actors trying to get out from under state controls. Most of the limited successes that have been achieved thus far are globalist in orientation, while most of the substantial successes (save for the Battle of Seattle and Serbia) have been autonomist. It will be interesting, as the instances of netwar increase over time, to see whether this pattern holds. The outcomes of the globalist cases suggest the prevalence of negotiated solutions, while the autonomist conflicts may, in general, have a much more inherently desperate character that drives them to greater violence and less willingness to reach accommodation. All this we will watch in the years to come. For now, these early cases have helped us to develop this taxonomy of netwar, further refining the concept.

Will netwar continue to empower nonstate actors, perhaps reducing the relative power advantage enjoyed by nation states?

Civil society networks have already made much use of social netwar as a tool for advancing a globalist, ethics-based agenda focused on broadening and deepening human rights regimes-often in the context of an ongoing effort to foster movement from authoritarian rule to democracy (e.g., Burma). But there is another side of nonstate-actor-oriented netwar, characterized not by globalist impulses, but rather by the desire to avoid state control of a network's criminal, terrorist or ethnic-separatist agenda (e.g., Hamas and Chechens). While the globalist netwars seem devoted to nonviolent tools of struggle, the autonomists may employ both means of engagement-often with a greater emphasis on violence.

VARIETIES OF NETWAR-DUAL PHENOMENA | Netwar can be waged by "good" as well as "bad" actors and through peaceful as well as violent measures. From its beginnings, netwar has appealed to a broad cross-section of nonstate actors who are striving to confront or cope with their state authorities.

Ethnonationalists, criminals and terrorists-all have found new power in networking. But so too have emerging global civil society actors who have emphasized nonviolent efforts to win the "battle of the story"-a more purely informational dimension of netwar-rather than the violent swarming characteristic of its darker side.

The duality of netwar in the real world-dark-side criminals and terrorists on the one hand, but enlightening civil society forces on the other, is mirrored in the virtual world of cyberspace, which is increasingly utilized for crime and terror, along with social activism.

At present, social activism is far more robust and established in the cyber realm than is crime or terror. Will this continue to be the case? We think so. Activists will become more adept at integrating the mobilizing force of the Internet with the power and appeal of messages aimed at spreading and protecting human rights. Even so, criminal and terrorist organizations will learn how to manipulate the infosphere with increasing skill.

Thus, netwar has two faces, like the Roman god Janus. Janus was the god of doors and gates, and thus of departures and returns, and new beginnings and initiatives. This, in a sense, meant he was the god of communications, too. His double face, one old and looking back, the other younger and peering forward, conveyed that he was an inherently dual god. At the beginning of creation, he partook in the separation of order from chaos. In Roman times, he was identified with the distinction between war and peace, for the gate to his temple at the Forum was kept ceremoniously closed in times of peace and open in times of war-which meant the gates were rarely closed. At the start of the 21st century, the world is again at a new beginning. It is uncertain whether it will be an era of peace or conflict; but how matters turn out will depend to some degree on which face of netwar predominates.

2 The Extended Mind

Andy Clark and David J. Chalmers[1]

Introduction

Where does the mind stop and the rest of the world begin? The question invites two standard replies. Some accept the demarcations of skin and skull, and say that what is outside the body is outside the mind. Others are impressed by arguments suggesting that the meaning of our words "just ain't in the head," and hold that this externalism about meaning carries over into an externalism about mind. We propose to pursue a third position. We advocate a very different sort of externalism: an *active externalism*, based on the active role of the environment in driving cognitive processes.

1 Extended Cognition

Consider three cases of human problem-solving:

(1) A person sits in front of a computer screen which displays images of various two-dimensional geometric shapes and is asked to answer questions concerning the potential fit of such shapes into depicted "sockets." To assess fit, the person must mentally rotate the shapes to align them with the sockets.
(2) A person sits in front of a similar computer screen, but this time can choose either to physically rotate the image on the screen, by pressing a rotate button, or to mentally rotate the image as before. We can also suppose, not unrealistically, that some speed advantage accrues to the physical rotation operation.
(3) Sometime in the cyberpunk future, a person sits in front of a similar computer screen. This agent, however, has the benefit of a neural implant which can perform the rotation operation as fast as the computer in the previous example. The agent must still choose which internal resource to

use (the implant or the good old-fashioned mental rotation), as each resource makes different demands on attention and other concurrent brain activity.

How much *cognition* is present in these cases? We suggest that all three cases are similar. Case (3) with the neural implant seems clearly to be on a par with case (1). And case (2) with the rotation button displays the same sort of computational structure as case (3), although it is distributed across agent and computer instead of internalized within the agent. If the rotation in case (3) is cognitive, by what right do we count case (2) as fundamentally different? We cannot simply point to the skin/skull boundary as justification, since the legitimacy of that boundary is precisely what is at issue. But nothing else seems different.

The kind of case just described is by no means as exotic as it may at first appear. It is not just the presence of advanced external computing resources which raises the issue, but rather the general tendency of human reasoners to lean heavily on environmental supports. Thus consider the use of pen and paper to perform long multiplication (McClelland, Rumelhart, and Hinton 1986; Clark 1989), the use of physical rearrangements of letter tiles to prompt word recall in Scrabble (Kirsh 1995), the use of instruments such as the nautical slide rule (Hutchins 1995), and the general paraphernalia of language, books, diagrams, and culture. In all these cases the individual brain performs some operations, while others are delegated to manipulations of external media. Had our brains been different, this distribution of tasks would doubtless have varied.

In fact, even the mental rotation cases described in scenarios (1) and (2) are real. The cases reflect options available to players of the computer game Tetris. In Tetris, falling geometric shapes must be rapidly directed into an appropriate slot in an emerging structure. A rotation button can be used. David Kirsh and Paul Maglio (1994) calculate that the physical rotation of a shape through 90 degrees takes about 100 milliseconds, plus about 200 milliseconds to select the button. To achieve the same result by mental rotation takes about 1,000 milliseconds. Kirsh and Maglio go on to present compelling evidence that physical rotation is used not just to position a shape ready to fit a slot, but often to help *determine* whether the shape and the slot are compatible. The latter use constitutes a case of what Kirsh and Maglio call an “epistemic action.” *Epistemic* actions alter the world so as to aid and augment cognitive processes such as recognition and search. Merely *pragmatic* actions, by contrast, alter the world because some physical change is desirable for its own sake (e.g., putting cement into a hole in a dam).

Epistemic action, we suggest, demands spread of *epistemic credit*. If, as we confront some task, a part of the world functions as a process which, *were it done in the head*, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world *is* (so we claim) part of the cognitive process. Cognitive processes ain't (all) in the head!

2 Active Externalism

In these cases, the human organism is linked with an external entity in a two-way interaction, creating a *coupled system* that can be seen as a cognitive system in its own right. All the components in the system play an active causal role, and they jointly govern behavior in the same sort of way that cognition usually does. If we remove the external component the system's behavioral competence will drop, just as it would if we removed part of its brain. Our thesis is that this sort of coupled process counts equally well as a cognitive process, whether or not it is wholly in the head.

This externalism differs greatly from standard variety advocated by Putnam (1975) and Burge (1979). When I believe that water is wet and my twin believes that twin water is wet, the external features responsible for the difference in our beliefs are distal and historical, at the other end of a lengthy causal chain. Features of the *present* are not relevant: if I happen to be surrounded by XYZ right now (maybe I have teleported to Twin Earth), my beliefs still concern standard water, because of my history. In these cases, the relevant external features are *passive*. Because of their distal nature, they play no role in driving the cognitive process in the here-and-now. This is reflected by the fact that the actions performed by me and my twin are physically indistinguishable, despite our external differences.

In the cases we describe, by contrast, the relevant external features are *active*, playing a crucial role in the here-and-now. Because they are coupled with the human organism, they have a direct impact on the organism and on its behavior. In these cases, the relevant parts of the world are *in the loop*, not dangling at the other end of a long causal chain. Concentrating on this sort of coupling leads us to an *active externalism*, as opposed to the passive externalism of Putnam and Burge.

Many have complained that even if Putnam and Burge are right about the externality of content, it is not clear that these external aspects play a causal or explanatory role in the generation of action. In counterfactual cases where internal structure is held constant but these external features are changed, behavior looks just the same; so internal structure seems to be doing the crucial work. We will not adjudicate that issue here, but we note that active externalism is not threatened by any such problem. The external features in a coupled system play an ineliminable role—if we retain internal structure but change the external features, behavior may change completely. The external features here are just as causally relevant as typical internal features of the brain.[2]

By embracing an active externalism, we allow a more natural explanation of all sorts of actions. One can explain my choice of words in Scrabble, for example, as the outcome of an extended cognitive process involving the rearrangement of tiles on my tray. Of course, one could always try to explain my action in terms of internal processes and a long series of "inputs" and "actions," but this explanation would be needlessly complex. If an isomorphic process were going on in the head, we would feel no urge to characterize it in this cumbersome way.[3] In a very real sense, the rearrangement of tiles on the tray is not part of action; it is part of *thought*.

The view we advocate here is reflected by a growing body of research in cognitive science. In areas as diverse as the theory of situated cognition (Suchman 1987), studies of real-world robotics (Beer 1989), dynamical approaches to child development (Thelen and Smith 1994), and research on the cognitive properties of collectives of agents (Hutchins 1995), cognition is often taken to be continuous with processes in the environment.[4] Thus, in seeing cognition as extended one is not merely making a terminological decision; it makes a significant difference to the methodology of scientific investigation. In effect, explanatory methods that might once have been thought appropriate only for the analysis of "inner" processes are now being adapted for the study of the outer, and there is promise that our understanding of cognition will become richer for it.

Some find this sort of externalism unpalatable. One reason may be that many identify the cognitive with the conscious, and it seems far from plausible that consciousness extends outside the head in these cases. But not every cognitive process, at least on standard usage, is a conscious process. It is widely accepted that all sorts of processes beyond the borders of consciousness play a crucial role in cognitive processing: in the retrieval of memories, linguistic processes, and skill acquisition, for example. So the mere fact that external processes are external where consciousness is internal is no reason to deny that those processes are cognitive.

More interestingly, one might argue that what keeps real cognition processes in the head is the requirement that cognitive processes be *portable*. Here, we are moved by a vision of what might be called the Naked Mind: a package of resources and operations we can always bring to bear on a cognitive task, regardless of the local environment. On this view, the trouble with coupled systems is that they are too easily *decoupled*. The true cognitive processes are those that lie at the constant core of the system; anything else is an add-on extra.

There is something to this objection. The brain (or brain and body) comprises a package of basic, portable, cognitive resources that is of interest in its own right. These resources may incorporate bodily actions into cognitive processes, as when we use our fingers as working memory in a tricky calculation, but they will not encompass the more contingent aspects of our external environment, such as a pocket calculator. Still, mere contingency of coupling does not rule out cognitive status. In the distant future we may be able to plug various modules into our brain to help us out: a module for extra short-term memory when we need it, for example. When a module is plugged in, the processes involving it are just as cognitive as if they had been there all along.[5]

Even if one were to make the portability criterion pivotal, active externalism would not be undermined. Counting on our fingers has already been let in the door, for example, and it is easy to push things further. Think of the old image of the engineer with a slide rule hanging from his belt wherever he goes. What if people always carried a pocket calculator, or had them implanted? The real moral of the portability intuition is that for coupled systems to be relevant to the core of cognition, *reliable* coupling is required. It happens that most reliable coupling takes place within the brain, but there can easily be reliable coupling with the environment as well. If the resources of my calculator or my Filofax are always there when I need them, then they are coupled with me as reliably as we need. In effect, they are part of the basic package of cognitive resources that I bring to bear on the everyday world. These systems cannot be impugned simply on the basis of the danger of discrete damage, loss, or malfunction, or because of any occasional decoupling: the biological brain is in similar danger, and occasionally loses capacities temporarily in episodes of sleep, intoxication, and emotion. If the relevant capacities are generally there when they are required, this is coupling enough.

Moreover, it may be that the biological brain has in fact evolved and matured in ways which factor in the reliable presence of a manipulable external environment. It certainly seems that evolution has favored onboard capacities which are especially geared to parasitizing the local environment so as to reduce memory load, and even to transform the nature of the computational problems themselves. Our visual systems have evolved to rely on their environment in various ways: they exploit contingent facts about the structure of natural scenes (e.g., Ullman and Richards

1984), for example, and they take advantage of the computational shortcuts afforded by bodily motion and locomotion (e.g., Blake and Yuille 1992). Perhaps there are other cases where evolution has found it advantageous to exploit the possibility of the environment being in the cognitive loop. If so, then external coupling is part of the truly basic package of cognitive resources that we bring to bear on the world.

Language may be an example. Language appears to be a central means by which cognitive processes are extended into the world. Think of a group of people brainstorming around a table, or a philosopher who thinks best by writing, developing her ideas as she goes. It may be that language evolved, in part, to enable such extensions of our cognitive resources within actively coupled systems.

Within the lifetime of an organism, too, individual learning may have molded the brain in ways that rely on cognitive extensions that surrounded us as we learned. Language is again a central example here, as are the various physical and computational artifacts that are routinely used as cognitive extensions by children in schools and by trainees in numerous professions. In such cases the brain develops in a way that complements the external structures, and learns to play its role within a unified, densely coupled system. Once we recognize the crucial role of the environment in constraining the evolution and development of cognition, we see that extended cognition is a core cognitive process, not an add-on extra.

An analogy may be helpful. The extraordinary efficiency of the fish as a swimming device is partly due, it now seems, to an evolved capacity to couple its swimming behaviors to the pools of external kinetic energy found as swirls, eddies, and vortices in its watery environment (see Triantafyllou and Triantafyllou 1995). These vortices include both naturally occurring ones (e.g., where water hits a rock) and self-induced ones (created by well-timed tail flaps). The fish swims by building these externally occurring processes into the very heart of its locomotion routines. The fish and surrounding vortices together constitute a unified and remarkably efficient swimming machine.

Now consider a reliable feature of the human environment, such as the sea of words. This linguistic surround envelops us from birth. Under such conditions, the plastic human brain will surely come to treat such structures as a reliable resource to be factored into the shaping of on-board cognitive routines. Where the fish flaps its tail to set up the eddies and vortices it subsequently exploits, we intervene in multiple linguistic media, creating local structures and disturbances whose reliable presence drives our ongoing internal processes. Words and external symbols are thus paramount among the cognitive vortices which help constitute human thought.

3 From Cognition to Mind

So far we have spoken largely about "cognitive processing," and argued for its extension into the environment. Some might think that the conclusion has been bought too cheaply. Perhaps some *processing* takes place in the environment, but what of *mind*? Everything we have said so far is compatible with the view that truly mental states—experiences, beliefs, desires, emotions, and so on—are all determined by states of the brain. Perhaps what is truly mental is internal, after all?

We propose to take things a step further. While some mental states, such as experiences, may be determined internally, there are other cases in which external factors make a significant contribution. In particular, we will argue that *beliefs* can be constituted partly by features of the environment, when those features play the right sort of role in driving cognitive processes. If so, the mind extends into the world.

First, consider a normal case of belief embedded in memory. Inga hears from a friend that there is an exhibition at the Museum of Modern Art, and decides to go see it. She thinks for a moment and recalls that the museum is on 53rd Street, so she walks to 53rd Street and goes into the museum. It seems clear that Inga believes that the museum is on 53rd Street, and that she believed this even before she consulted her memory. It was not previously an *occurrent* belief, but then neither are most of our beliefs. The belief was sitting somewhere in memory, waiting to be accessed.

Now consider Otto. Otto suffers from Alzheimer's disease, and like many Alzheimer's patients, he relies on information in the environment to help structure his life. Otto carries a notebook around with him everywhere he goes. When he learns new information, he writes it down. When he needs some old information, he looks it up. For Otto, his notebook plays the role usually played by a biological memory. Today, Otto hears about the exhibition at the Museum of Modern Art, and decides to go see it. He consults the notebook, which says that the museum is on 53rd Street, so he walks to 53rd Street and goes into the museum.

Clearly, Otto walked to 53rd Street because he wanted to go to the museum and he believed the museum was on 53rd Street. And just as Inga had her belief even before she consulted her memory, it seems reasonable to say that Otto believed the museum was on 53rd Street even before consulting his notebook. For in relevant respects the cases are entirely analogous:

the notebook plays for Otto the same role that memory plays for Inga. The information in the notebook functions just like the information constituting an ordinary non-occurrent belief; it just happens that this information lies beyond the skin.

The alternative is to say that Otto has no belief about the matter until he consults his notebook; at best, he believes that the museum is located at the address in the notebook. But if we follow Otto around for a while, we will see how unnatural this way of speaking is. Otto is constantly using his notebook as a matter of course. It is central to his actions in all sorts of contexts, in the way that an ordinary memory is central in an ordinary life. The same information might come up again and again, perhaps being slightly modified on occasion, before retreating into the recesses of his artificial memory. To say that the beliefs disappear when the notebook is filed away seems to miss the big picture in just the same way as saying that Inga's beliefs disappear as soon as she is no longer conscious of them. In both cases the information is reliably there when needed, available to consciousness and available to guide action, in just the way that we expect a belief to be.

Certainly, insofar as beliefs and desires are characterized by their explanatory roles, Otto's and Inga's cases seem to be on a par: the essential causal dynamics of the two cases mirror each other precisely. We are happy to explain Inga's action in terms of her occurrent desire to go to the museum and her standing belief that the museum is on 53rd street, and we should be happy to explain Otto's action in the same way. The alternative is to explain Otto's action in terms of his occurrent desire to go to the museum, his standing belief that the Museum is on the location written in the notebook, and the accessible fact that the notebook says the Museum is on 53rd Street; but this complicates the explanation unnecessarily. If we must resort to explaining Otto's action this way, then we must also do so for the countless other actions in which his notebook is involved; in each of the explanations, there will be an extra term involving the notebook. We submit that to explain things this way is to take *one step too many*. It is pointlessly complex, in the same way that it would be pointlessly complex to explain Inga's actions in terms of beliefs about her memory. The notebook is a constant for Otto, in the same way that memory is a constant for Inga; to point to it in every belief/desire explanation would be redundant. In an explanation, simplicity is power.

If this is right, we can even construct the case of Twin Otto, who is just like Otto except that a while ago he mistakenly wrote in his notebook that the Museum of Modern Art was on 51st Street. Today, Twin Otto is a physical duplicate of Otto from the skin in, but his notebook differs. Consequently, Twin Otto is best characterized as believing that the museum is on 51st Street, where Otto believes it is on 53rd. In these cases, a belief is simply not in the head.

This mirrors the conclusion of Putnam and Burge, but again there are important differences. In the Putnam/Burge cases, the external features constituting differences in belief are distal and historical, so that twins in these cases produce physically indistinguishable behavior. In the cases we are describing, the relevant external features play an active role in the here-and-now, and have a direct impact on behavior. Where Otto walks to 53rd Street, Twin Otto walks to 51st. There is no question of explanatory irrelevance for this sort of external belief content; it is introduced precisely because of the central explanatory role that it plays. Like the Putnam/Burge cases, these cases involve differences in reference and truth conditions, but they also involve differences in the dynamics of *cognition.*[6]

The moral is that when it comes to belief, there is nothing sacred about skull and skin. What makes some information count as a belief is the role it plays, and there is no reason why the relevant role can be played only from inside the body.

Some will resist this conclusion. An opponent might put her foot down and insist that as she uses the term “belief,” or perhaps even according to standard usage, Otto simply does not qualify as believing that the museum is on 53rd Street. We do not intend to debate what is standard usage; our broader point is that the notion of belief *ought* to be used so that Otto qualifies as having the belief in question. In all *important* respects, Otto’s case is similar to a standard case of (non-occurrent) belief. The differences between Otto’s case and Inga’s are striking, but they are superficial. By using the “belief” notion in a wider way, it picks out something more akin to a natural kind. The notion becomes deeper and more unified, and is more useful in explanation.

To provide substantial resistance, an opponent has to show that Otto’s and Inga’s cases differ in some important and relevant respect. But in what deep respect are the cases different? To make the case *solely* on the grounds that information is in the head in one case but not in the other would be to beg the question. If this difference is relevant to a difference in belief, it is surely not *primitively* relevant. To justify the different treatment, we must find some more basic underlying difference between the two.

It might be suggested that the cases are relevantly different in that Inga has more *reliable* access to the information. After all, someone might take away Otto’s notebook at any time, but Inga’s memory is safer. It is not implausible that constancy is relevant: indeed, the fact that Otto always uses his notebook played some role in our justifying its cognitive status. If Otto were consulting a guidebook as a one-off, we would be much less likely to ascribe him a standing belief. But in the original case, Otto's access to the notebook is very reliable—not perfectly reliable, to be sure, but then neither is Inga's access to her memory. A surgeon might tamper with her brain, or more mundanely, she might have too much to drink. The mere possibility of such tampering is not enough to deny her the belief.

One might worry that Otto's access to his notebook *in fact* comes and goes. He showers without the notebook, for example, and he cannot read it when it is dark. Surely his belief cannot come and go so easily? We could get around this problem by redescribing the situation, but in any case an occasional temporary disconnection does not threaten our claim. After all, when Inga is asleep, or when she is intoxicated, we do not say that her belief disappears. What really counts is that the information is easily available when the subject needs it, and this constraint is satisfied equally in the two cases. If Otto's notebook were often unavailable to him at times when the information in it would be useful, there might be a problem, as the information would not be able to play the action-guiding role that is central to belief; but if it is easily available in most relevant situations, the belief is not endangered.

Perhaps a difference is that Inga has *better* access to the information than Otto does? Inga's "central" processes and her memory probably have a relatively high-bandwidth link between them, compared to the low-grade connection between Otto and his notebook. But this alone does not make a difference between believing and not believing. Consider Inga's museum-going friend Lucy, whose biological memory has only a low-grade link to her central systems, due to nonstandard biology or past misadventures. Processing in Lucy's case might be less efficient, but as long as the relevant information is accessible, Lucy clearly believes that the museum is on 53rd Street. If the connection was too indirect—if Lucy had to struggle hard to retrieve the information with mixed results, or a psychotherapist's aid were needed—we might become more reluctant to ascribe the belief, but such cases are well beyond Otto's situation, in which the information is easily accessible.

Another suggestion could be that Otto has access to the relevant information only by *perception*, whereas Inga has more direct access—by introspection, perhaps. In some ways, however, to put things this way is to beg the question. After all, we are in effect advocating a point of view on which Otto's internal processes and his notebook constitute a single cognitive system. From the standpoint of this system, the flow of information between notebook and brain is not perceptual at all; it does not involve the impact of something outside the system. It is more akin to information flow within the brain. The only deep way in which the access is perceptual is that in Otto's case, there is a distinctly perceptual phenomenology associated with the retrieval of the information, whereas in Inga's case there is not. But why should the nature of an associated phenomenology make a difference to the status of a belief? Inga's memory may have some associated phenomenology, but it is still a belief. The phenomenology is not visual, to be sure. But for visual phenomenology consider the Terminator, from the Arnold Schwarzenegger movie of the same name. When he recalls some information from memory, it is "displayed" before him in his visual field (presumably he is conscious of it, as there are frequent shots depicting his point of view). The fact that standing memories are recalled in this unusual way surely makes little difference to their status as standing beliefs.

These various small differences between Otto's and Inga's cases are all *shallow* differences. To focus on them would be to miss the way in which for Otto, notebook entries play just the sort of role that beliefs play in guiding most people's lives.

Perhaps the intuition that Otto's is not a true belief comes from a residual feeling that the only true beliefs are occurrent beliefs. If we take this feeling seriously, Inga's belief will be ruled out too, as will many beliefs that we attribute in everyday life. This would be an extreme view, but it may be the most consistent way to deny Otto's belief. Upon even a slightly less extreme view—the view that a belief must be *available* for consciousness, for example—Otto's notebook entry seems to qualify just as well as Inga's memory. Once dispositional beliefs are let in the door, it is difficult to resist the conclusion that Otto's notebook has all the relevant dispositions.

4 Beyond the Outer Limits

If the thesis is accepted, how far should we go? All sorts of puzzle cases spring to mind. What of the amnesic villagers in *100 Years of Solitude*, who forget the names for everything and so hang labels everywhere? Does the information in my Filofax count as part of my memory? If Otto's notebook has been tampered with, does he believe the newly installed information? Do I believe the contents of the page in front of me before I read it? Is my cognitive state somehow spread across the Internet?

We do not think that there are categorical answers to all of these questions, and we will not give them. But to help understand what is involved in ascriptions of extended belief, we can at least examine the features of our central case that make the notion so clearly applicable there. First, the notebook is a constant in Otto's life—in cases where the information in the notebook would be relevant, he will rarely take action without consulting it. Second, the information in the notebook is directly available without difficulty. Third, upon retrieving information from the notebook he automatically endorses it. Fourth, the information in the notebook has been consciously endorsed at some point in the past, and indeed is there as a consequence of this endorsement.[7] The status of the fourth feature as a criterion for belief is arguable (perhaps one can acquire beliefs through subliminal perception, or through memory tampering?), but the first three features certainly play a crucial role.

Insofar as increasingly exotic puzzle cases lack these features, the applicability of the notion of "belief" gradually falls off. If I rarely take relevant action without consulting my Filofax, for example, its status within my cognitive system will resemble that of the notebook in Otto's. But if I often act without consultation—for example, if I sometimes answer relevant questions with "I don't know"—then information in it counts less clearly as part of my belief system. The Internet is likely to fail on multiple counts, unless I am unusually computer-reliant, facile with the technology, and trusting, but information in certain files on my computer may qualify. In intermediate cases, the question of whether a belief is present may be indeterminate, or the answer may depend on the varying standards that are at play in various contexts in which the question might be asked. But any indeterminacy here does not mean that in the central cases, the answer is not clear.

What about socially extended cognition? Could my mental states be partly constituted by the states of other thinkers? We see no reason why not, in principle. In an unusually interdependent couple, it is entirely possible that one partner's beliefs will play the same sort of role for the other as the notebook plays for Otto.[8] What is central is a high degree of trust, reliance, and accessibility. In other social relationships these criteria may not be so clearly fulfilled, but they might nevertheless be fulfilled in specific domains. For example, the waiter at my favorite restaurant might act as a repository of my beliefs about my favorite meals (this might even be construed as a case of extended desire). In other cases, one's beliefs might be embodied in one's secretary, one's accountant, or one's collaborator.[9]

In each of these cases, the major burden of the coupling between agents is carried by language. Without language, we might be much more akin to discrete Cartesian “inner” minds, in which high-level cognition relies largely on internal resources. But the advent of language has allowed us to spread this burden into the world. Language, thus construed, is not a mirror of our inner states but a complement to them. It serves as a tool whose role is to extend cognition in ways that on-board devices cannot. Indeed, it may be that the intellectual explosion in recent evolutionary time is due as much to this linguistically enabled extension of cognition as to any independent development in our inner cognitive resources.

What, finally, of the self? Does the extended mind imply an extended self? It seems so. Most of us already accept that the self outstrips the boundaries of consciousness; my dispositional beliefs, for example, constitute in some deep sense part of who I am. If so, then these boundaries may also fall beyond the skin. The information in Otto’s notebook, for example, is a central part of his identity as a cognitive agent. What this comes to is that Otto *himself* is best regarded as an extended system, a coupling of biological organism and external resources. To consistently resist this conclusion, we would have to shrink the self into a mere bundle of occurrent states, severely threatening its deep psychological continuity. Far better to take the broader view, and see agents themselves as spread into the world.

As with any reconception of ourselves, this view will have significant consequences. There are obvious consequences for philosophical views of the mind and for the methodology of research in cognitive science, but there will also be effects in the moral and social domains. It may be, for example, that in some cases interfering with someone’s environment will have the same moral significance as interfering with their person. And if the view is taken seriously, certain forms of social activity might be reconceived as less akin to communication and action, and as more akin to thought. In any case, once the hegemony of skin and skull is usurped, we may be able to see ourselves more truly as creatures of the world.

Notes

This essay was originally published in *Analysis* 58 (1998): 10–23. Reprinted in P. Grim (ed.), *The Philosopher’s Annual*, vol. 21 (1998); reprinted in D. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings* (Oxford University Press, 2002).

1. The authors are listed in order of degree of belief in the central thesis.

2. Much of the appeal of externalism in the philosophy of mind may stem from the intuitive appeal of active externalism. Externalists often make analogies involving external features in coupled systems, and appeal to the arbitrariness of boundaries between brain and environment. But these intuitions sit uneasily with the letter of standard externalism. In most of the Putnam/Burge cases, the immediate environment is irrelevant; only the historical environment counts. Debate has focused on the question of whether mind must be in the head, but a more relevant question in assessing these examples might be: is mind in the present?

3. Herbert Simon (1981) once suggested that we view internal memory as, in effect, an external resource upon which "real" inner processes operate. "Search in memory," he comments, "is not very different from search of the external environment." Simon's view at least has the virtue of treating internal and external processing with the parity they deserve, but we suspect that on his view the mind will shrink too small for most people's tastes.

4. Philosophical views of a similar spirit can be found in Haugeland 1995, McClamrock 1995, Varela, Thompson, and Rosch 1991, and Wilson 1994.

5. Or consider the following passage from a fairly recent science fiction novel (McHugh 1992, p. 213): "I am taken to the system's department where I am attuned to the system. All I do is jack in and then a technician instructs the system to attune and it does. I jack out and query the time. 10:52. The information pops up. Always before I could only access information when I was jacked in, it gave me a sense that I knew what I thought and what the system told me, but now, how do I know what is system and what is Zhang?"

6. In the terminology of Chalmers's "The Components of Content" (2002): the twins in the Putnam/Burge cases differ only in their *relational* content, but Otto and his twin can be seen to differ in their *notional* content, which is the sort of content that governs cognition. Notional content is generally internal to a cognitive system, but in this case the cognitive system is itself effectively extended to include the notebook.

7. The constancy and past-endorsement criteria may suggest that history is partly constitutive of belief. One might react to this by removing any historical component (giving a purely dispositional reading of the constancy criterion and eliminating the past-endorsement criterion, for example), or one might allow such a component as long as the main burden is carried by features of the present.

8. Might this sort of reasoning also allow something like Burge's extended "arthritis" beliefs? After all, I might always defer to my doctor in taking relevant actions concerning my disease. Perhaps so, but there are some clear differences. For example, any extended beliefs would be grounded in an existing active relationship with the doctor, rather than in a historical relationship to a language community. And on the current analysis, my deference to the doctor would tend to yield something

like a true belief that I have some other disease in my thigh, rather than the false belief that I have arthritis there. On the other hand, if I used medical experts solely as terminological consultants, the results of Burge's analysis might be mirrored.

9. From the *New York Times*, March 30, 1995, p. B7, in an article on former UCLA basketball coach John Wooden: "Wooden and his wife attended 36 straight Final Fours, and she invariably served as his memory bank. Nell Wooden rarely forgot a name—her husband rarely remembered one—and in the standing-room-only Final Four lobbies, she would recognize people for him."

References

Beer, R. (1989). *Intelligence as Adaptive Behavior.* New York: Academic Press.

Blake, A., and Yuille, A. (eds.) (1992). *Active Vision*. Cambridge, MA: MIT Press.

Burge, T. (1979). Individualism and the mental. *Midwest Studies in Philosophy*, *4*, 73–122.

Chalmers, D. J. (2002). The components of content. In David J. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings*. Oxford: Oxford University Press.

Clark, A. (1989). *Microcognition*. Cambridge, MA: MIT Press.

Haugeland, J. (1995). Mind embodied and embedded. In Y. Houng and J. Ho (eds.), *Mind and Cognition*. Taipei: Academia Sinica.

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.

Kirsh, D. (1995). The intelligent use of space. *Artificial Intelligence*, *73*, 31–68.

Kirsh, D., and Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*, 513–549.

McClamrock, R. (1995). *Existential Cognition*. Chicago: University of Chicago Press.

McClelland, J. L., Rumelhart, D. E., and Hinton, G. E. (1986). The appeal of parallel distributed processing. In J. L. McClelland, D. E. Rumelhart, and PDP Research Group, *Parallel Distributed Processing* (vol. 2). Cambridge, MA: MIT Press.

McHugh, M. (1992). *China Mountain Zhang*. New York: Tom Doherty Associates.

Putnam, H. (1975). The meaning of "meaning." In K. Gunderson (ed.), *Language, Mind, and Knowledge*. Minneapolis: University of Minnesota Press.

Simon, H. (1981). *The Sciences of the Artificial*. Cambridge, MA: MIT Press.

Suchman, L. (1987). *Plans and Situated Actions*. Cambridge: Cambridge University Press.

Thelen, E. and Smith, L. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press.

Triantafyllou, M., and Triantafyllou, G. (1995). An efficient swimming machine. *Scientific American*, *272*(3), 64–70.

Ullman, S., and Richards, W. (1984). *Image Understanding*. Norwood, NJ: Ablex.

Varela, F., Thompson, E., and Rosch, E. (1991). *The Embodied Mind*. Cambridge, MA: MIT Press.

Wilson, R. (1994). Wide computationalism. *Mind*, *103*, 351–372.

Do contemporary media structures promote cultural interactivity?

From
Understanding Media: The Extensions of Man
by Marshall McLuhan

CHAPTER 1

The Medium is the Message

MARSHALL McCLUHAN

In a culture like ours, long accustomed to splitting and dividing all things as a means of control, it is sometimes a bit of a shock to be reminded that, in operational and practical fact, the medium is the message. This is merely to say that the personal and social consequences of any medium—that is, of any extension of ourselves—result from the new scale that is introduced into our affairs by each extension of ourselves, or by any new technology. Thus, with automation, for example, the new patterns of human association tend to eliminate jobs it is true. That is the negative result. Positively, automation creates roles for people, which is to say depth of involvement in their work and human association that our preceding mechanical technology had destroyed. Many people would be disposed to say that it was not the machine, but what one did with the machine, that was its meaning or message. In terms of the ways in which the machine altered our relations to one another and to ourselves, it mattered not in the least whether it turned out cornflakes or Cadillacs. The restructuring of human work and association was shaped by the technique of fragmentation that is the essence of machine technology. The essence of automation technology is the opposite. It is integral and decentralist in depth, just as the machine was fragmentary, centralist, and superficial in its patterning of human relationships.

The instance of the electric light may prove illuminating in this connection. The electric light is pure information. It is a medium without a message, as it were, unless it is used to spell out some verbal ad or name. This fact, characteristic of all media, means that the "content" of any medium is always another medium. The content of writing is speech, just as the written word is the content of print, and print is the content of the telegraph. If it is asked, "What is the content of speech?," it is necessary to say, "It is an actual process of thought, which is in itself nonverbal." An abstract painting represents direct manifestation of creative thought processes as they might appear in computer designs. What we are considering here, however, are the psychic and social consequences of the designs or patterns as they amplify or accelerate existing processes. For the "message" of any medium or technology is the change of scale or pace or pattern that it introduces into human affairs. The railway did not introduce movement or transportation or wheel or road into human society, but it accelerated and enlarged the scale of previous human functions, creating totally new kinds of cities and new kinds of work and leisure. This happened whether the railway functioned in a tropical or a northern environment, and is quite independent of the freight or content of the railway medium. The airplane, on the other hand, by accelerating the rate of transportation, tends to dissolve the railway form of city, politics, and association, quite independently of what the airplane is used for.

Let us return to the electric light. Whether the light is being used for brain surgery or night baseball is a matter of indifference. It could be argued that these activities are in some way the "content" of the electric light, since they could not exist without the electric light. This fact merely underlines the point that "the medium is the message" because it is the medium that shapes and controls the scale and form of human association and action. The content or uses of such media are as diverse as they are ineffectual in shaping the form of human association. Indeed, it is only too typical that the "content" of any medium blinds us to the character of the medium. It is only today that industries have become aware of the various kinds of business in which they are engaged. When IBM discovered that it was not in the business of making office equipment or business machines, but that it was in the business of processing information, then it began to navigate with clear vision. The General Electric Company makes a considerable portion of its profits from electric light bulbs and lighting systems. It has not yet discovered that, quite as much as A.T.&T., it is in the business of moving information.

The electric light escapes attention as a communication medium just because it has no "content." And this makes it an invaluable instance of how people fail to study media at all.

For it is not till the electric light is used to spell out some brand name that it is noticed as a medium. Then it is not the light but the "content" (or what is really another medium) that is noticed. The message of the electric light is like the message of electric power in industry, totally radical, pervasive, and decentralized. For electric light and power are separate from their uses, yet they eliminate time and space factors in human association exactly as do radio, telegraph, telephone, and TV, creating involvement in depth.

A fairly complete handbook for studying the extensions of man could be made up from selections from Shakespeare. Some might quibble about whether or not he was referring to TV in these familiar lines from *Romeo and Juliet:*

> But soft! what light through yonder window breaks?
> It speaks, and yet says nothing.

In Othello, which, as much as *King Lear*, is concerned with the torment of people transformed by illusions, there are these lines that bespeak Shakespeare's intuition of the transforming powers of new media:

> Is there not charms
> By which the property of youth and maidhood
> May be abus'd? Have you not read Roderigo,
> Of some such thing?

In Shakespeare's *Troilus and Cressida*, which is almost completely devoted to both a psychic and social study of communication, Shakespeare states his awareness that true social and political navigation depend upon anticipating the consequences of innovation:

The providence that's in a watchful state
Knows almost every grain of Plutus' gold,
Finds bottom in the uncomprehensive deeps,
Keeps place with thought, and almost like the gods
Does thoughts unveil in their dumb cradles.

The increasing awareness of the action of media, quite independently of their "content" or programming, was indicated in the annoyed and anonymous stanza:

In modern thought, (if not in fact)
Nothing is that doesn't act,
So that is reckoned wisdom which
Describes the scratch but not the itch.

The same kind of total, configurational awareness that reveals why the medium is socially the message has occurred in the most recent and radical medical theories. In his *Stress of Life*, Hans Selye tells of the dismay of a research co1league on hearing of Selye's theory:

> When he saw me thus launched on yet another enraptured description of what I had observed in animals treated with this or that impure, toxic material, he looked at me with desperately sad eyes and said in obvious despair: "But Selye try to realize what you are doing before it is too late! You have now decided to spend your entire life studying the pharmacology of dirt!"

(Hans Selye, T*he Stress of Life*)

As Selye deals with the total environmental situation in his "stress" theory of disease, so the latest approach to media study considers not only the "content" but the medium and the cultural matrix within which the particular medium operates. The older unawareness of the psychic and social effects of media can be illustrated from almost any of the conventional pronouncements.

In accepting an honorary degree from the University of Notre Dame a few years ago, General David Sarnoff made this statement: "We are too prone to make technological instruments the scapegoats for the sins of those who wield them. The products of modern science are not in themselves good or bad; it is the way they are used that determines their value." That is the voice of the current somnambulism. Suppose we were to say, "Apple pie is in itself neither good nor bad; it is the way it is used that determines its value." Or, "The smallpox virus is in itself neither good nor bad; it is the way it is used that determines its value." Again, "Firearms are in themselves neither good nor bad; it is the way they are used that determines their value." That is, if the slugs reach the right people firearms are good. If the TV tube fires the right ammunition at the right people it is good. I am not being perverse. There is simply nothing in the Sarnoff statement that will bear scrutiny, for it ignores the nature of the medium, of any and all media, in the true Narcissus style of one hypnotized by the amputation and extension of his own being in a new technical form. General Sarnoff went on to explain his attitude to the technology of print, saying that it was true that print caused much trash to circulate, but it had also disseminated the Bible and the thoughts of seers and philosophers. It has never occurred to General Sarnoff that any technology could do anything but add itself on to what we already are.

Such economists as Robert Theobald, W. W. Rostow, and John Kenneth Galbraith have been explaining for years how it is that "classical economics" cannot explain change or growth. And the paradox of mechanization is that although it is itself the cause of maximal growth and change, the principle of mechanization excludes the very possibility of growth or the understanding of change. For mechanization is achieved by fragmentation of any process and by putting the fragmented parts in a series. Yet, as David Hume showed in the eighteenth century, there is no principle of causality in a mere sequence. That one thing follows another accounts for nothing. Nothing follows from following, except change. So the greatest of all reversals occurred with electricity, that ended sequence by making things instant. With instant speed the causes of things began to emerge to awareness again, as they had not done with things in sequence and in concatenation accordingly. Instead of asking which came first, the chicken or the egg, it suddenly seemed that a chicken was an egg's idea for getting more eggs.

Just before an airplane breaks the sound barrier, sound waves become visible on the wings of the plane. The sudden visibility of sound just as sound ends is an apt instance of that great pattern of being that reveals new and opposite forms just as the earlier forms reach their peak performance. Mechanization was never so vividly fragmented or sequential as in the birth of the movies, the moment that translated us beyond mechanism into the world of growth and organic interrelation. The movie, by sheer speeding up the mechanical, carried us from the world of sequence and connections into the world of creative configuration and structure. The message of the movie medium is that of transition from lineal connections to configurations. It is the transition that produced the now quite correct observation: "If it works, it's obsolete." When electric speed further takes over from mechanical movie sequences, then the lines of force in structures and in media become loud and clear. We return to the inclusive form of the icon.

To a highly literate and mechanized culture the movie appeared as a world of triumphant illusions and dreams that money could buy. It was at this moment of the movie that cubism occurred and it has been described by E. H. Gombrich (*Art and Illusion*) as "the most radical attempt to stamp out ambiguity and to enforce one reading of the picture—that of a man-made construction, a colored canvas." For cubism substitutes all facets of an object simultaneously for the "point of view" or facet of perspective illusion. Instead of the specialized illusion of the third dimension on canvas, cubism sets up an interplay of planes and contradiction or dramatic conflict of patterns, lights, textures that "drives home the message" by involvement. This is held by many to be an exercise in painting, not in illusion.

In other words, cubism, by giving the inside and outside, the top, bottom, back, and front and the rest, in two dimensions, drops the illusion of perspective in favor of instant sensory awareness of the whole. Cubism, by seizing on instant total awareness, suddenly announced that the *medium is the message*. Is it not evident that the moment that sequence yields to the simultaneous, one is in the world of the structure and of configuration? Is that not what has happened in physics as in painting, poetry, and in communication? Specialized segments of attention have shifted to total field, and we can now say, "The medium is the message" quite naturally. Before the electric speed and total field, it was not obvious that the medium is the message. The message, it seemed, was the "content," as people used to ask what a painting was about. Yet they never thought to ask what a melody was about, nor what a house or a dress was *about*. In such matters, people retained some sense of the whole pattern, of form and function as a unity. But in the electric age this integral idea of structure and configuration has become so prevalent that educational theory has taken up the matter. Instead of working with specialized "problems" in arithmetic, the structural approach now follows the lines of force in the field of number and has small children meditating about number theory and "sets."

Cardinal Newman said of Napoleon, "He understood the grammar of gunpowder." Napoleon had paid some attention to other media as well, especially the semaphore telegraph that gave him a great advantage over his enemies. He is on record for saying that "Three hostile newspapers are more to be feared than a thousand bayonets."

Alexis de Tocqueville was the first to master the grammar of print and typography. He was thus able to read off the message of coming change in France and America as if he were reading aloud from a text that had been handed to him. In fact, the nineteenth century in France and in America was just such an open book to de Tocqueville because he had learned the grammar of print. So he, also, knew when that grammar did not apply. He was asked why he did not write a book on England, since he knew and admired England. He replied:

> One would have to have an unusual degree of philosophical folly to believe oneself able to judge England in six months. A year always seemed to me too short a time in which to appreciate the United States properly, and it is much easier to acquire clear and precise notions about the American Union than about Great Britain. In America all laws derive in a sense from the same line of thought. The whole of society, so to speak, is founded upon a single fact; everything springs from a simple principle. One could compare America to a forest pierced by a multitude of straight roads all converging on the same point. One has only to find the center and

everything is revealed at a glance. But in England the paths run criss-cross, and it is only by travelling down each one of them that one can build up a picture of the whole.

De Tocqueville in earlier work on the French Revolution, had explained how it was the printed word that, achieving cultural saturation in the eighteenth century, had homogenized the French nation. Frenchmen were the same kind of people from north to south. The typographic principles of uniformity, continuity, and lineality had overlaid the complexities of ancient feudal and oral society. The Revolution was carried out by the new literati and lawyers.

In England, however, such was the power of the ancient oral traditions of common law, backed by the medieval institution of Parliament, that no uniformity or continuity of the new visual print culture could take complete hold. The result was that the most important event in English history has never taken place; namely, the English Revolution on the lines of the French Revolution. The American Revolution had no medieval legal institutions to discard or to root out, apart from monarchy. And many have held that the American Presidency has become very much more personal and monarchical than any European monarch ever could be.

De Tocqueville's contrast between England and America is clearly based on the fact of typography and of print culture creating uniformity and continuity. England, he says, has rejected this principle and clung to the dynamic or oral common-law tradition. Hence the discontinuity and unpredictable quality of English culture. The grammar of print cannot help to construe the message of oral and nonwritten culture and institutions. The English aristocracy was properly classified as barbarian by Matthew Arnold because its power and status had nothing to do with literacy or with the cultural forms of typography. Said the Duke of Gloucester to Edward Gibbon upon the publication of his *Decline and Fall:* "Another damned fat book, eh, Mr. Gibbon? Scribble, scribble, scribble, eh, Mr. Gibbon?" De Tocqueville was a highly literate aristocrat who was quite able to be detached from the values and assumptions of typography. That is why he alone understood the grammar of typography. And it is only on those terms, standing aside from any structure or medium, that its principles and lines of force can be discerned. For any medium has the power of imposing its own assumption on the unwary. Prediction and control consist in avoiding this subliminal state of Narcissus trance. But the greatest aid to this end is simply in knowing that the spell can occur immediately upon contact, as in the first bars of a melody.

A Passage to India by E. M. Forster is a dramatic study of the inability of oral and intuitive oriental culture to meet with the rational, visual European patterns of experience. "Rational," of course, has for the West long meant "uniform and continuous and sequential." In other words, we have confused reason with literacy, and rationalism with a single technology. Thus in the electric age man seems to the conventional West to become irrational. In Forster's novel the moment of truth and dislocation from the typographic trance of the West comes in the Marabar Caves. Adela Quested's reasoning powers cannot cope with the total inclusive field of resonance that is India. After the Caves: "Life went on as usual, but had no consequences, that is to say, sounds did not echo nor thought develop. Everything seemed cut off at its root and therefore infected with illusion."

A Passage to India (the phrase is from Whitman, who saw America headed Eastward) is a parable of Western man in the electric age, and is only incidentally related to Europe or the Orient. The ultimate conflict between sight and sound, between written and oral kinds of perception and organization of existence is upon us. Since understanding stops action, as Nietzsche observed, we can moderate the fierceness of this conflict by understanding the media that extend us and raise these wars within and without us.

Detribalization by literacy and its traumatic effects on tribal man is the theme of a book by the psychiatrist J. C. Carothers, *The African Mind in Health and Disease* (World Health Organization, Geneva, 1953). Much of his material appeared in an article in *Psychiatry* magazine, November, 1959: "The Culture, Psychiatry, and the Written Word." Again, it is electric speed that has revealed the lines of force operating from Western technology in the remotest areas of bush, savannah, and desert. One example is the Bedouin with his battery radio on board the camel. Submerging natives with floods of concepts for which nothing has prepared them is the normal action of all of our technology. But with electric media Western man himself experiences exactly the same inundation as the remote native. We are no more prepared to encounter radio and TV in our literate milieu than the native of Ghana is able to cope with the literacy that takes him out of his collective tribal world and beaches him in individual isolation. We are as numb in our new electric world as the native involved in our literate and mechanical culture.

Electric speed mingles the cultures of prehistory with the dregs of industrial marketeers, the nonliterate with semiliterate and the postliterate. Mental breakdown of varying degrees is the very common result of uprooting and inundation with new information and endless new patterns of information. Wyndham Lewis made this a theme of his group of novels called *The Human Age*. The first of these, *The Childermass*, is concerned precisely with accelerated media change as a kind of massacre of the innocents. In our own world as we become more aware of the effects of technology on psychic formation and manifestation, we are losing all confidence in our right to assign guilt. Ancient prehistoric societies regard violent crime as pathetic. The killer is regarded as we do a cancer victim. "How terrible it must be to feel like that," they say. J. M. Synge took up this idea very effectively in his *Playboy of the Western World.*

If the criminal appears as a nonconformist who is unable to meet the demand of technology that we behave in uniform and continuous patterns, literate man is quite inclined to see others who cannot conform as somewhat pathetic. Especially the child, the cripple, the woman, and the colored person appear in a world of visual and typographic technology as victims of injustice. On the other hand, in a culture that assigns roles instead of jobs to people—the dwarf, the skew, the child create their own spaces. They are not expected to fit into some uniform and repeatable niche that is not their size anyway. Consider the phrase "It's a man's world." As a quantitative observation endlessly repeated from within a homogenized culture, this phrase refers to the men in such a culture who have to be homogenized Dagwoods in order to belong at all. It is in our I.Q. testing that we have produced the greatest flood of misbegotten standards. Unaware of our typographic cultural bias, our testers assume that uniform and continuous habits are a sign of intelligence, thus eliminating the ear man and the tactile man.

C. P. Snow, reviewing a book of A. L. Rowse (*The New York Times Book Review*, December 24, 1961) on *Appeasement* and the road to Munich, describes the top level of British brains and experience in the 1930s. "Their I.Q.'s were much higher than usual among political bosses. Why were they such a disaster?" The view of Rowse, Snow approves: "They would not listen to warnings because they did not wish to hear." Being anti-Red made it impossible for them to read the message of Hitler. But their failure was as nothing compared to our present one. The American stake in literacy as a technology or uniformity applied to every level of education, government, industry, and social life is totally threatened by the electric technology. The threat of Stalin or Hitler was external. The electric technology is within the gates, and we are numb, deaf, blind, and mute about its encounter with the Gutenberg technology, on and through which the American way of life was formed. It is, however, no time to suggest strategies when the threat has not even been acknowledged to exist. I am in the position of Louis Pasteur telling doctors that their greatest enemy was quite invisible, and quite unrecognized by them. Our conventional response to all media, namely that it is how they are used that counts, is the numb stance of the technological idiot. For the "content" of a medium is like the juicy piece of meat carried by the burglar to distract the watchdog of the mind. The effect of the medium is made strong and intense just because it is given another medium as "content." The content of a movie is a novel or a play or an opera. The effect of the movie form is not related to its program content. The "content" of writing or print is speech, but the reader is almost entirely unaware either of print or of speech.

Arnold Toynbee is innocent of any understanding of media as they have shaped history' but he is full of examples that the student of media can use. At one moment he can seriously suggest that adult education, such as the Workers Educational Association in Britain, is a useful counterforce to the popular press. Toynbee considers that although all of the oriental societies have in our time accepted the industrial technology and its political consequences: "On the cultural plane, however, there is no uniform corresponding tendency." (Somervell, I. 267) This is like the voice of the literate man, floundering in a milieu of ads, who boasts, "Personally, I pay no attention to ads." The spiritual and cultural reservations that the oriental peoples may have toward our technology will avail them not at all. The effects of technology do not occur at the level of opinions or concepts, but alter sense ratios or patterns of perception steadily and without any resistance. The serious artist is the only person able to encounter technology with impunity, just because he is an expert aware of the changes in sense perception.

The operation of the money medium in seventeenth century Japan had effects not unlike the operation of typography in the West. The penetration of the money economy, wrote G. B. Sansom (in *Japan*, Cresset Press, London, 1931) "caused a slow but irresistible revolution, culminating in the breakdown of feudal government and the resumption of intercourse with foreign countries after more than two hundred years of seclusion." Money has reorganized the sense life of peoples just because it is an *extension* of our sense lives. This change does not depend upon approval or disapproval of those living in the society.

Arnold Toynbee made one approach to the transforming power of media in his concept of "etherialization," which he holds to be the principle of progressive simplification and efficiency in any organization or technology. Typically, he is ignoring the *effect* of the challenge of these forms upon the response of our senses. He imagines that it is the response of our opinions that is relevant to the effect of media and technology in society, a "point of view" that is plainly the result of the typographic spell. For the man in a literate and homogenized society ceases to be sensitive to the diverse and discontinuous life of forms. He acquires the illusion of the third dimension and the "private point of view" as part of his Narcissus fixation, and is quite shut off from Blake's awareness or that of the Psalmist, that we become what we behold.

Today when we want to get our bearings in our own culture, and have need to stand aside from the bias and pressure exerted by any technical form of human expression, we have only to visit a society where that particular form has not been felt, or a historical period in which it was unknown. Professor Wilbur Schramm made such a tactical move in studying *Television in the Lives of Our Children*. He found areas where TV had not penetrated at all and ran some tests. Since he had made no study of the peculiar nature of the TV image, his tests were of "content" preferences, viewing time, and vocabulary counts. In a word, his approach to the problem was a literary one, albeit unconsciously so. Consequently, he had nothing to report. Had his methods been employed in 1500 A.D. to discover the effects of the printed book in the lives of children or adults, he could have found out nothing of the changes in human and social psychology resulting from typography. Print created individualism and nationalism in the sixteenth century. Program and "content" analysis offer no clues to the magic of these media or to their subliminal charge.

Leonard Doob, in his report *Communication in Africa*, tells of one African who took great pains to listen each evening to the BBC news, even though he could understand nothing of it. Just to be in the presence of those sounds at 7 P.M. each day was important for him. His attitude to speech was like ours to melody—the resonant intonation was meaning enough. In the seventeenth century our ancestors still shared this native's attitude to the forms of media, as is plain in the following sentiment of the Frenchman Bernard Lam expressed in *The Art of Speaking* (London, 1696):

> ‘Tis an effect of the Wisdom of God, who created Man to be happy, that whatever is useful to his conversation (way of life) is agreeable to him . . . because all victual that conduces to nourishment is relishable, whereas other things that cannot be assimulated and be turned into our substance are insipid. A Discourse cannot be pleasant to the Hearer that is not easie to the Speaker; nor can it be easily pronounced unless it be heard with delight.

Here is an equilibrium theory of human diet and expression such as even now we are only striving to work out again for media after centuries of fragmentation and specialism.

Pope Pius XII was deeply concerned that there be serious study of the media today. On February 17, 1950, he said:

> It is not an exaggeration to say that the future of modern society and the stability of its inner life depend in large part on the maintenance of an equilibrium between the strength of the techniques of communication and the capacity of the individual’s own reaction.

Failure in this respect has for centuries been typical and total for mankind. Subliminal and docile acceptance of media impact has made them prisons without walls for their human users. As A. J. Liebling remarked in his book *The Press*, a man is not free if he cannot see where he is going, even if he has a gun to help him get there. For each of the media is also a powerful weapon with which to clobber other media and other groups. The result is that the present age has been one of multiple civil wars that are not limited to the world of art and entertainment. In *War and Human Progress*, Professor J. U. Nef declared: “The total wars of our time have been the result of a series of intellectual mistakes . . .”

If the formative power in the media are the media themselves, that raises a host of large matters that can only be mentioned here, although they deserve volumes. Namely’ that technological media are staples or natural resources, exactly as are coal and cotton and oil. Anybody will concede that society whose economy is dependent upon one or two major staples like cotton, or grain, or lumber, or fish, or cattle is going to have some obvious social patterns of organization as a result. Stress on a few major staples creates extreme instability in the economy but great endurance in the population. The pathos and humor of the American South are embedded in such an economy of limited staples. For a society configured by reliance on a few commodities accepts them as a social bond quite as much as the metropolis does the press. Cotton and oil, like radio and TV, become “fixed charges” on the entire psychic life of the community. And this pervasive fact creates the unique cultural flavor of any society. It pays through the nose and all its other senses for each staple that shapes its life.

That our human senses, of which all media are extensions are also fixed charges on our personal energies, and that they also configure the awareness and experience of each one of us may be perceived in another connection mentioned by the psychologist C. G. Jung:

> Every Roman was surrounded by slaves. The slave and his psychology flooded ancient Italy, and every Roman became inwardly, and of course unwittingly, a slave. Because living constantly in the atmosphere of slaves, he became infected through the unconscious with their psychology. No one can shield himself from such an influence (*Contributions to Analytical Psychology*, London, 1928).

CHAPTER 7

Challenge and Collapse

The Nemesis of Creativity

It was Bertrand Russell who declared that the great discovery of the twentieth century was the technique of the suspended judgment. A. N. Whitehead, on the other hand, explained how the great discovery of the nineteenth century was the discovery of the technique of discovery. Namely, the technique of starting with the thing to be discovered and working back, step by step, as on an assembly line, to the point at which it is necessary to start in order to reach the desired object. In the arts this meant starting with the *effect* and then inventing a poem, painting, or building that would have just that effect and no other.

But the "technique of the suspended judgment" goes further. It anticipates the effect of, say, an unhappy childhood on an adult, and offsets the effect before it happens. In psychiatry it is the technique of total permissiveness extended as an anesthetic for the mind, while various adhesions and moral effects of false judgments are systematically eliminated.

This is a very different thing from the numbing or narcotic effect of new technology that lulls attention while the new form slams the gates of judgment and perception. For massive social surgery is needed to insert new technology into the group mind, and this is achieved by the built-in numbing apparatus discussed earlier. Now the "technique of the suspended judgment" presents the possibility of rejecting the narcotic and of postponing indefinitely the operation of inserting the new technology in the social psyche. A new stasis is in prospect.

Werner Heisenberg, in *The Physicist's Conception of Nature*, is an example of the new quantum physicist whose over-all awareness of forms suggests to him that we would do well to stand aside from most of them. He points out that technical change alters not only habits of life, but patterns of thought and valuation, citing with approval the outlook of the Chinese sage:

> As Tzu-Gung was traveling through the regions north of the river Han, he saw an old man working in his vegetable garden. He had dug an irrigation ditch. The man would descend into a well, fetch up a vessel of water in his arms and pour it out into the ditch. While his efforts were tremendous the results appeared to be very meager.
>
> Tzu-Gung said, "There is a way whereby you can irrigate a hundred ditches in one day, and whereby you can do much with little effort. Would you not like to hear of it?"
>
> Then the gardener stood up, looked at him and said, "And what would that be?"
>
> Tzu-Gung replied, "You take a wooden lever, weighted at the back and light in front. In this way you can bring up water so quickly that it just gushes out. This is called a draw-well."
>
> Then anger rose up in the old man's face, and he said "I have heard my teacher say that whoever uses machines does all his work like a machine. He who does his work like a machine grows a heart like a machine, and he who carries the heart of a machine in his breast loses his simplicity. He who has lost his simplicity becomes unsure in the strivings of his soul. Uncertainty in the strivings of the soul is something which does not agree with honest sense. It is not that I do not know of such things; I am ashamed to use them."

Perhaps the most interesting point about this anecdote is that it appeals to a modern physicist. It would not have appealed to Newton or to Adam Smith, for they were great experts and advocates of the fragmentary and the specialist approaches. It is by means quite in accord with the outlook of the Chinese sage that Hans Selye works at his "stress" idea of illness. In the 1 920s he had been baffled at why physicians always seemed to concentrate on the recognition of individual diseases and specific remedies for such isolated causes, while never paying any attention to the "syndrome of just being sick." Those who are concerned with the program "content" of media and not with the medium proper, appear to be in the position of physicians who ignore the "syndrome of just being sick." Hans Selye, in tackling a total, inclusive approach to the field of sickness, began what Adolphe Jonas has continued in *Irritation and Counter-irritation*; namely, a quest for the response to injury as such, or to novel impact of any kind. Today we have anesthetics that enable us to perform the most frightful physical operations on one another.

The new media and technologies by which we amplify and extend ourselves constitute huge collective surgery carried out on the social body with complete disregard for antiseptics. If the operations are needed, the inevitability of infecting the whole system during the operation has to be considered. For in operating on society with a new technology, it is not the incised area that is most affected. The area of impact and incision is numb. It is the entire system that is changed. The effect of radio is visual, the effect of the photo is auditory. Each new impact shifts the ratios among all the senses. What we seek today is either a means of controlling these shifts in the sense-ratios of the psychic and social outlook, or a means of avoiding them altogether. To have a disease without its symptoms is to be immune. No society has ever known enough about its actions to have developed immunity to its new extensions or technologies. Today we have begun to sense that art may be able to provide such immunity.

In the history of human culture there is no example of a conscious adjustment of the various factors of personal and social life to new extensions except in the puny and peripheral efforts of artists. The artist picks up the message of cultural and technological challenge decades before its transforming impact occurs. He, then, builds models or Noah's arks for facing the change that is at hand. "The war of 1870 need never have been fought had people read my *Sentimental Education,*" said Gustave Haubert.

It is this aspect of *new* art that Kenneth Galbraith recommends to the careful study of businessmen who want to stay in business. For in the electric age there is no longer any sense in talking about the artist's being ahead of his time. Our technology is, also, ahead of its time, if we reckon by the ability to recognize it for what it is. To prevent undue wreckage in society, the artist tends now to move from the ivory tower to the control tower of society. Just as higher education is no longer a frill or luxury but a stark need of production and operational design in the electric age, so the artist is indispensable in the shaping and analysis and understanding of the life of forms, and structures created by electric technology.

The percussed victims of the new technology have invariably muttered clichés about the impracticality of artists and their fanciful preferences. But in the past century it has come to be generally acknowledged that, in the words of Wyndham Lewis, "The artist is always engaged in writing a detailed history of the future because he is the only person aware of the nature of the present." Knowledge of this simple fact is now needed for human survival. The ability of the artist to sidestep the bully blow of new technology of any age, and to parry such violence with full awareness, is age-old. Equally age-old is the inability of the percussed victims, who cannot sidestep the new violence, to recognize their need of the artist. To reward and to make celebrities of artists can, also, be a way of ignoring their prophetic work, and preventing its timely use for survival. The artist is the man in any field, scientific or humanistic, who grasps the implications of his actions and of new knowledge in his own time. He is the man of integral awareness.

The artist can correct the sense ratios before the blow of new technology has numbed conscious procedures. He can correct them before numbness and subliminal groping and reaction begin. If this is true, how is it possible to present the matter to those who are in a position to do something about it? If there were even a remote likelihood of this analysis being true, it would warrant a global armistice and period of stocktaking. If it is true that the artist possesses the means of anticipating and avoiding the consequences of technological trauma, then what are we to think of the world and bureaucracy of "art appreciation"? Would it not seem suddenly to be a conspiracy to make the artist a frill, a fribble, or a Milltown? If men were able to be convinced that art is precise advance knowledge of how to cope with the psychic and social consequences of the next technology, would they all become artists? Or would they begin a careful translation of new art forms into social navigation charts? I am curious to know what would happen if art were suddenly seen for what it is, namely, exact information of how to rearrange one's psyche in order to anticipate the next blow from our own extended faculties. Would we, then, cease to look at works of art as an explorer might regard the gold and gems used as the ornaments of simple nonliterates?

At any rate, in experimental art, men are given the exact specifications of coming violence to their own psyches from their own counter-irritants or technology. For those parts of ourselves that we thrust out in the form of new invention are attempts to counter or neutralize collective pressures and irritations. But the counter-irritant usually proves a greater plague than the initial irritant, like a drug habit. And it is here that the artist can show us how to "ride with the punch," instead of "taking it on the chin." It can only be repeated that human history is a record of "taking it on the chin."

Emile Durkheim long ago expressed the idea that the specialized task always escaped the action of the social conscience. In this regard, it would appear that the artist is the social conscience and is treated accordingly! "We have no art," say the Balinese; "we do everything as well as possible."

The modern metropolis is now sprawling helplessly after the impact of the motorcar. As a response to the challenge of railway speeds the suburb and the garden city arrived too late, or just in time to become a motorcar disaster. For an arrangement of functions adjusted to one set of intensities becomes unbearable at another intensity. And a technological extension of our bodies designed to alleviate physical stress can bring on psychic stress that may be much worse. Western specialist technology transferred to the Arab world in late Roman times released a furious discharge of tribal energy.

The somewhat devious means of diagnosis that have to be used to pin down the actual form and impact of a new medium are not unlike those indicated in detective fiction by Peter Cheyney. In *You Can't Keep the Change* (Collins, London, 1956) he wrote:

A case to Callaghan was merely a collection of people, some of whom,—all of whom—were giving incorrect information, or telling lies, because circumstances either forced them or led them into the process.

But the fact that they *had* to tell lies; *had* to give false impressions, necessitated a reorientation of their own viewpoints and their own lives. Sooner or later they became exhausted or careless. Then, and not until then, was an investigator able to put his finger on the one fact that would lead lead him to a possible logical solution.

It is interesting to note that success in keeping up a respectable front of the customary kind can only be done by a frantic scramble back of the façade. After the crime, after the blow has fallen, the facade of custom can only be held up by swift rearrangement of the props. So it is in our social lives when a new technology strikes, or in our private life when some intense and, therefore, indigestible experience occurs, and the censor acts at once to numb us from the blow and to ready the faculties to assimilate the intruder. Peter Cheyney's observations of a mode of detective fiction is another instance of a popular form of entertainment functioning as mimic model of the real thing.

Perhaps the most obvious "closure" or psychic consequence of any new technology is just the demand for it. Nobody wants a motorcar till there are motorcars, and nobody is interested in TV until there are TV programs. This power of technology to create its own world of demand is not independent of technology being first an extension of our own bodies and senses. When we are deprived of our sense of sight, the other senses take up the role of sight in some degree. But the need to use the senses that are available is as insistent as breathing—a fact that makes sense of the urge to keep radio and TV going more or less continuously. The urge to continuous use is quite independent of the "content" of public programs or of the private sense life, being testimony to the fact that technology is part of our bodies. Electric technology is directly related to our central nervous systems, so it is ridiculous to talk of "what the public wants" played over its own nerves. This question would be like asking people what sort of sights and sounds they would prefer around them in an urban metropolis! Once we have surrendered our senses and nervous systems to the private manipulation of those who would try to benefit from taking a lease on our eyes and ears and nerves, we don't really have any rights left. Leasing our eyes and ears and nerves to commercial interests is like handing over the common speech to a private corporation, or like giving the earth's atmosphere to a company as a monopoly. Something like this has already happened with outer space, for the same reasons that we have leased our central nervous systems to various corporations. As long as we adopt the Narcissus attitude of regarding the extensions of our own bodies as really *out there* and really independent of us, we will meet all technological challenges with the same sort of banana-skin pirouette and collapse.

Archimedes once said, "Give me a place to stand and I will move the world." Today he would have pointed to our electric media and said, "I will stand on your eyes, your ears, your nerves, and your brain, and the world will move in any tempo or pattern I choose." We have leased these "places to stand" to private corporations.

Arnold Toynbee has devoted much of his *A Study of History* to analyzing the kinds of challenge faced by a variety of cultures during many centuries. Highly relevant to Western man is Toynbee's explanation of how the lame and the crippled respond to their handicaps in a society of active warriors. They become specialists like Vulcan, the smith and armorer. And how do whole communities act when conquered and enslaved? The same strategy serves them as it does the lame individual in a society of warriors. They specialize and become indispensable to their masters. It is probably the long human history of enslavement, and the collapse into specialism as a counterirritant, that have put the stigma of servitude and pusillanimity on the figure of the specialist, even in modern times. The capitulation of Western man to his technology, with its crescendo of specilized demands, has always appeared to many observers of our world as a kind of enslavement. But the resulting fragmentation has been voluntary and enthusiastic, unlike the conscious strategy of specialism on the part of the captives of military conquest.

It is plain that fragmentation or specialism as a technique of achieving security under tyranny and oppression of any kind has an attendant danger. Perfect adaptation to any environment is achieved by a total channeling of energies and vital force that amounts to a kind of static terminus for a creature. Even slight changes in the environment of the very well adjusted find them without any resource to meet new challenge. Such is the plight of the representatives of "conventional wisdom" in any society. Their entire stake of security and status is in a single form of acquired knowledge, so that innovation is for them not novelty but annihilation.

A related form of challenge that has always faced cultures is the simple fact of a frontier or a wall, on the other side of which exists another kind of society. Mere existence side by side of any two forms of organization generates a great deal of tension. Such, indeed, has been the principle of symbolist artistic structures in the past century. Toynbee observes that the challenge of a civilization set side by side with a tribal society has over and over demonstrated that the simpler society finds its integral economy and institutions "disintegrated by a rain of psychic energy generated by the civilization" of the more complex culture. When two societies exist side by side, the psychic challenge of the more complex one acts as an explosive release of energy in the simpler one. For prolific evidence of this kind of problem it is not necessary to look beyond the life of the teenager lived daily in the midst of a complex urban center. As the barbarian was driven to furious restlessness by the civilized contact, collapsing into mass migration, so the teenager, compelled to share the life of a city that cannot accept him as an adult, collapses into "rebellion without a cause." Earlier the adolescent had been provided with a rain check. He was prepared to wait it out. But since TV, the drive to participation has ended adolescence, and every American home has its Berlin wall.

Toynbee is very generous in providing examples of widely varied challenge and collapse, and is especially apt in pointing to the frequent and futile resort to futurism and archaism as strategies of encountering radical change. But to point back to the day of the horse or to look forward to the coming of antigravitational vehicles is not an adequate response to the challenge of the motorcar. Yet these two uniform ways of backward and forward looking are habitual ways of avoiding the discontinuities of present experience with their demand for sensitive inspection and appraisal. Only the dedicated artist seems to have the power for encountering the present actuality.

Toynbee urges again and again the cultural strategy of the imitation of the example of great men. This, of course, is to locate cultural safety in the power of the *will*, rather than in the power of adequate *perception* of situations. Anybody could quip that this is the British trust in character as opposed to intellect. In view of the endless power of men to hypnotize themselves into unawareness in the presence of challenge, it may be argued that will-power is as useful as intelligence for survival. Today we need also the will to be exceedingly informed and aware.

Arnold Toynbee gives an example of Renaissance technology being effectively encountered and creatively controlled when he shows how the revival of the decentralized medieval parliament saved English society from the monopoly of centralism that seized the continent. Lewis Mumford in *The City in History* tells the strange tale of how the New England town was able to carry out the pattern of the medieval ideal city because it was able to dispense with walls and to mix town and country. When the technology of a time is powerfully thrusting in one direction, wisdom may well call for a countervailing thrust. The implosion of electric energy in our century cannot be met by explosion or expansion, but it can be met by decentralism and the flexibility of multiple small centers. For example, the rush of students into our universities is not explosion but implosion. And the needful strategy to encounter this force is not to enlarge the university, but to create numerous groups of autonomous colleges in place of our centralized university plant that grew up on the lines of European government and nineteenth-century industry.

In the same way the excessive tactile effects of the TV image cannot be met by mere program changes. Imaginative strategy based on adequate diagnosis would prescribe a corresponding depth or structural approach to the existing literary and visual world. If we persist in a conventional approach to these developments our traditional culture will be swept aside as scholasticism was in the sixteenth century. Had the Schoolmen with their complex oral culture understood the Gutenberg technology, they could have created a new synthesis of written and oral education, instead of bowing out of the picture and allowing the merely visual page to take over the educational enterprise. The oral Schoolmen did not meet the new visual challenge of print, and the resulting expansion or explosion of Gutenberg technology was in many respects an impoverishment of the culture, as historians like Mumford are now beginning to explain. Arnold Toynbee, in *A Study of History*, in considering "the nature of growths of civilizations," not only abandons the concept of enlargement as a criterion of real growth of society, but states: "More often geographical expansion is a concomitant of real decline and coincides with a 'time of troubles' or a universal state—both of them stages of decline and disintegration."

Toynbee expounds the principle that times of trouble or rapid change produce militarism, and it is militarism that produces empire and expansion. The old Greek myth which taught that the alphabet produced militarism ("King Cadmus sowed the dragon's teeth, and they sprang up armed men") really goes much deeper than Toynbee's story. In fact, "militarism" is just vague description, not analysis of causality at all. Militarism is a kind of visual organization of social energies that is both specialist and explosive, so that it is merely repetitive to say, as Toynbee does, that it both creates large empires and causes social breakdown. But militarism is a form of industrialism or the concentration of large amounts of homogenized energies into a few kinds of production. The Roman soldier was a man with a spade. He was an expert workman and builder who processed and packaged the resources of many societies and sent them home. Before machinery, the only massive work forces available for processing material were soldiers or slaves. As the Greek myth of Cadmus points out, the phonetic alphabet was the greatest processer of men for homogenized military life that was known to antiquity. The age of Greek society that Herodotus acknowledges to have been "overwhelmed by more troubles than in the twenty preceding generations" was the time that to our literary retrospect appears as one of the greatest of human centuries. It was Macaulay who remarked that it was not pleasant to live in times about which it was exciting to read. The succeeding age of Alexander saw Hellenism expand into Asia and prepare the course of the later Roman expansion. These, however were the very centuries in which Greek civilization obviously fell apart.

Toynbee points to the strange falsification of history by archeology, insofar as the survival of many material objects of the past does not indicate the quality of ordinary life and experience at any particular time. Continuous technical improvement in the means of warfare occurs over the entire period of Hellenic and Roman decline. Toynbee checks out his hypothesis by testing it with the developments in Greek agriculture. When the enterprise of Solon weaned the Greeks from mixed farming to a program of specialized products for export, there were happy consequences and a glorious manifestation of energy in Greek life. When the next phase of the same specialist stress involved much reliance on slave labor there was spectacular increase of production. But the armies of technologically specialized slaves working the land blighted the social existence of the independent yeomen and small farmers, and led to the strange world of the Roman towns and cities crowded with rootless parasites.

To a much greater degree than Roman slavery, the specialism of mechanized industry and market organization has faced Western man with the challenge of manufacture by mono-fracture, or the tackling of all things and operations one-bit-at-a-time. This is the challenge that has permeated all aspects of our lives and enabled us to expand so triumphantly in all directions and in all spheres.

13

THE WORK OF ART IN THE ELECTRONIC AGE

Interview with *La Sept*

We are in the age of electronic communications; do you think that this communication has fundamentally changed the way in which we see and understand the world or is it merely an accelerated form of technical reproduction?

I think we have a kind of transformation, yes, perhaps not a revolution in the subversive sense of the term, but a transformation of the relations of exchange. It is often said that we are within communication and we are perhaps no longer exactly within exchange. At this time, what has changed is that the means of communication, the medium, is becoming a determinant element in exchange and quite often dominates its function, even technologically – as McLuhan has it – dominates the content, the message, the subject communicated, the very substance of communication. So there is an inversion of terms of some sort, where what was a means of communication gains a kind of finality, and possibly also a counter-finality, and then the strategies which revolve around the medium, the communication media, become more essential than the strategies which concern the contents.

The reality of images counts for more . . .

Yes, it seems that from this point a sort of proliferation occurs, a saturation of the field of the medium. That is, as long as one has things to say and there is something to be exchanged between speakers, there is meaning, so it gets said and the process is limited. The medium gets the upper hand, basically, when images – or discourses – begin to proliferate, apart even from the meaning which they bear. There is a sort of – I don't want to say cancer; that would be borrowing a bit too much from modern pathology – but something similar. There is a kind of autonomy now of the systems of images and systems of discourse which has got the upper hand over meaning.

So then, where is reality still situated?

That's the entire problem. Is there still reality? I would rather say that we are in hyper-reality. Effectively, everything can be an object of communication.

Communication is completely generalized; it is no longer only discourse, but everything, which is an object of communication: architecture, art, have as their first end, it seems, to communicate, before even that of bearing a meaning. So where is the reality of what must be said, of what one wants to say? One almost has the impression that communication is a collective enterprise: the medium must function. There is a kind of obligation, a collective categorical imperative to make the media function at all costs. It's of little consequence whether the contents are completely real or unreal, or hyper-real; the important thing is that the medium continues to roll. So communication is drawn into this cycle of panic. It seems to become immediately an unlimited, proliferating system. There is a kind of imperialism of communication.

When you say hyper-reality, you might specify, in relation to reality, the reality of my daily life . . .

Yes, we can say for example that for events, politics, history, from the moment where they only exist as broadcasts by the media and proliferate, nearly globally, their own reality disappears. In the extreme case the event could just as well not have taken place – there are examples. It has taken place on the level of the screen. They are screen events and no longer authentic events – I am not sure if one can speak of authenticity, but nevertheless an ordinary reality which has a historical actuality that disappears behind the mediating hyper-reality of things. And here then, one can question indefinitely the degree, the rate of reality which continues to be shown. It's something else which is taking place: circuits are functioning. They can nourish themselves with anything, they can devour anything and, as Benjamin said of the work of art, you can never really go back to the source, you can never interrogate an event, a character, a discourse about its degree of original reality. That's what I call hyper-reality. Fundamentally, it's a domain where you can no longer interrogate the reality or unreality, the truth or falsity of something. We walk around in a sphere, a megasphere where things no longer have a reality principle. Rather a communication principle, a mediatizing principle.

Therefore we are all becoming images?

Yes, in one way or another . . . not only are there screens and terminals in technical terms, but we ourselves, the listeners, the TV spectators, become the terminals of all this communications network. We ourselves are screens. Lastly, the interlocutors are no longer exactly human beings. That sounds pejorative, but it's like that. The play has settled to one from screen to screen. It is almost dialogues between terminals or between different media.In a way it is the medium conversing with itself, this intense circulation, this type of auto-referentiality of media which includes us in its network. But it's somewhat of an integrated man–machine circuit. And at the present the difference between man and machine is very difficult to determine.

And that goes for artists, works of art?

Absolutely. I believe a work of art no longer has any privilege as a singular object of breaking through this type of circuit, of interrupting the circuit in some sense – because that is what it would amount to, a singular, unique appearance of an object unlike any other. However, for the communication system and that of the media there is no privileged object. All are substitutable and the work of art has finished up in there, it has been plugged into the same circuit. That does not mean that it may not have an authentic origin, but this will become less and less retrievable through the consumption of media, perhaps on screens but equally the large exhibitions where thousands attend and the crowd itself constitutes a kind of medium; the perception which one has of greatness is even constituted today by a crowd of spectators who are no longer amateurs, lovers of culture, in the traditional sense of the term. We are obliged to take this massive consumption into account. It is not merely the irruption of television or reproducability, reprography, or things of that ilk. Rather, the irruption of the mass even into aesthetic consumption changes, in my opinion, the status of the art work. Can we still talk of a 'work'? It's becoming something else. It is not exactly a commodity but it passes into the condition of a sign which must be able to circulate like any other. Therefore its own time and place, its uniqueness, is effectively removed.

Thus change of medium – the work becomes image – change of consumption . . .

Yes, the change would perhaps be that while the work of art creates its own space, it invents itself, it takes its inspiration from itself, it has a quite unique reality, it becomes itself an object on the screen, it is transmitted by the screen. The change is perhaps this. One passes from what was a stage, the stage of the work, the stage of the theatre, the stage of representation, even the stage of politics – one can generalize this phenomenon – to the screen. But the screen has quite another dimension: it is superficial, it only communicates images, not a particular time and place. In the end it makes everything circulate in one space, without depth, where all the objects must be able to follow one after the other without slowing down or stopping the circuit. But the work of art is made for stopping, in the end it is made to interrupt something, to arrest the gaze, to arrest contemplation. If there is an aesthetic pleasure or aesthetic relationship there it must arise from this species of sublime moment which is a moment of immobility, of contemplation, which corresponds also to the moment of creation in the artist. But then for the media none of this works. It is not that which counts. You have to move on in any case. It is a bit like what was being said about the Mona Lisa in Japan, with five seconds to look at it.

But if the work of art has fallen, has been banalized like the other products of communication, how is it that we consume so many of them? The state, individuals, businesses ...

You have answered yourself. Because it's been banalized, so many are consumed, I would say. Because it is reproduced in an indefinite number of examples, it becomes a bit like holograms in the end; it has also cloned itself, in a way. And that is really the principle of consumption; it is as if everything becomes sign and image – in fact, it is more and more necessary for this to happen. There is no limit to consumption. There is a kind of limit to creation, but not to consumption. You enter into an indefinite series, the seriality of things, and then it is imperative for it not to stop. It is rather like the principle of television: the screen must always be filled, the void is not permitted, and as we are somewhat screens ourselves now, transformed into reflecting screens, there must always be images there, so they are filled in by something or other. And then, perhaps, there is the fact that banalization quite simply involves commercialization, entrance into the market, and that strategies of competition, of course, also enter into it, but there is more to it than simple market strategies, a profusion of images is needed. The screen, the image, the sign, the modern message, the media require profusion, proliferation; there must be plenty for everyone. This is not the case for the work of art, of course, which creates a personal, privileged relationship ....

There you seem to be returning to a kind of nostalgia for the aura ...

Not exactly. What I have said sounds like that and it is difficult to say it in another way because the discourse brings in an element of nostalgia. But you could conceive of ... well, many things; of course something has been lost, but Benjamin said this well. In my opinion, Baudelaire before him had a more modern vision of things – the idea of the absolute commodity. If we enter into the era of the commodity, or that of the media, one could say this of us. For him, the modern artist should not try to revalorize, resacralize traditional art or aesthetics, but go further into the commodity. You had to go some way towards the absolute object, the absolute commodity, and he had a lovely formula: the modern artist owes it to himself to give the commodity a heroic status while the bourgeoisie only gives it, in advertising and all that, sentimental status. So when I say that, yes, of course nostalgia is there, the discourse is necessarily ambivalent. There is a background of nostalgia, and it is not possible to say it in another way since we are in discourse and thus still in something not yet entirely mediated, and then I think that it is possible nevertheless to have a vision there, not a more cynical vision, no there is perhaps a new aesthetic as a result of mediation, of disappearance, really. To go further in this direction and to really play the commodity, but at the power of two or the power of ten. To play the media, but in a sense with an almost ironic strategy.

It makes me think of Andy Warhol ...

Yes, Andy Warhol, for instance. It's not a solution but it is the other road and the spirit, that's not nostalgic, but rather an ultra-radical practice of mediation. It has its grandiose, but also its sublime, which is nevertheless of the cool sort; it's a cool strategy, while one may suppose that the strategy of the traditional art work was more warm, of another order, really.

And so, in your opinion, what remains of the order of traditional aesthetic or nostalgic creation? There cannot be an avant-garde – more disenchantment is still needed.

There are perhaps two paths. There is perhaps a path which continues the history of art, if I can call it that, where artists like Bacon and all those continue to work as creative individuals. We can't abandon this path, it still exists, but in my view it is true to say that it's no longer exactly contemporary. It is anachronistic in some ways, but wonderful all the same. In my opinion our modern – or postmodern, I don't know – condition is really that of mediation and it is there that strategies are worked out or, indeed, another destiny, where one can talk of the disappearance of art. Art is perhaps in the process of playing out its own disappearance – in my opinion this is what it has been doing for the last century. The problem posed today is perhaps that we have reached the end of this process and that we are entering a period where art no longer does anything else than stimulate its own disappearance because it has already disappeared in reality, because the media have already carried it off, because the system has. So in my opinion it's difficult to say the game is up. Not exactly, but it is also difficult to think that there still might be a real history for art.

Therefore, instead of resistance do we accept the deed of mediation, banalization and go as far as possible?

Yes, that would seem a less banal strategy, let's say, than that of nostalgia.

So what about you in all this, your own identity?

Well, yes, obviously to say that is still to speak as the subject of a discourse. But I don't think that need be a contradiction, but it is a paradox, surely. Because if the media have done away with it and if the work of art itself is really threatened, so is critical discourse, analytical discourse, just as much. That is to say that my position, to the extent that I make an analysis like this, is also paradoxical, because it no longer normally has any reference itself. It is there on the screen, that's obvious: I am on the screen, you are too, so we are already in another stratosphere and we are really talking, perhaps, there of something which has been lost. But I don't see the possibility of passing to the other side. You have to accept the deed, in my opinion, and then there are ambivalent strategies. There are double strategies.

You mean more experts than mediators?

Yes, but there is perhaps a way of playing with the media, to accept the deed of this system, of this entire integrated circuit and set it in play and perhaps not disrupt it, yet make it reversible – it's possible – and perhaps still to obtain some critical results, or sublime ones. But you have to take into account that these are fragile strategies.

Difficult therefore to .... How are you ... are you an image?

We are in a false situation now, it's true; you have to work within the paradox, the paradox of communication, which is in effect that everywhere there is communication and no one any longer has anything to say to anyone, or almost, whilst the paradox of language or of the work of art would be that there is something to say but the medium is no longer there. The message is there but the medium no longer responds. Or else in the opposite sense the medium is there, but there is no longer any message. Finally, we will now be in this paradoxical situation perpetually.

Do you think that one can still think?

Because I think it, it must be true. But nothing will be able to prove it any more than one can elaborate a truth outside a system like this. One will have to work with the hyper-reality of this system and enter the sphere of the floating signifier, of floating meaning or non-meaning, with risky strategies, etc. This is what I believe must be done. One must abandon the objective radical position of the subject and of the message.

But economic power relations are determining, all the same?

I myself am in no position to speak about it. I don't think so but it's a bit stupid to say that. As for myself, they are not determining. They are themselves mediated, they themselves pass into another sphere. I no longer believe that there are objective power relations nor objective strategies of this kind. In a way, they themselves have lost their own reality principle. Before, we could have made an infrastructure of it, a determining causality and everything, but today, no, they are entering too into a hyper-real zone where it is all connected with the media, all connected with advertising, etc., and even power relationships, even objective strategies are obliged to do their own advertising, to make their own image before being decrypted or decoded, before taking action. So they no longer have the material force they might have had before, if they ever had any. But perhaps sides have already been taken. Myself, I privilege in analysis things other than the economic, but perhaps that is debatable.

The Filter Bubble

by Eli Pariser

INTRODUCTION (Abridged)

A squirrel dying in front of your house may be more relevant to your interests right now than people dying in Africa.

Mark Zuckerberg, Facebook founder

We shape our tools, and thereafter our tools shape us.

Marshall McLuhan, media theorist

FEW PEOPLE NOTICED the post that appeared on Google's corporate blog on December 4, 2009. It didn't beg for attention, no sweeping pronouncements, no Silicon Valley hype, just a few paragraphs of text sandwiched between a weekly roundup of top search terms and an update about Google's finance software.

Not everyone missed it. Search engine blogger Danny Sullivan pores over the items on Google's blog looking for clues about where the monolith is headed next, and to him, the post was a big deal. In fact, he wrote later that day, it was the biggest change that has ever happened in search engines. For Danny, the headline said it all: “Personalized search for everyone.”

Starting that morning, Google would use fifty-seven “signals”, everything from where you were logging in from to what browser you were using to what you had searched for before to make guesses about who you were and what kinds of sites you'd like. Even if you were logged out, it would customize its results, showing you the pages it predicted you were most likely to click on.

Most of us assume that when we google a term, we all see the same results, the ones that the company's famous Page Rank algorithm suggests are the most authoritative based on other pages' links. But since December 2009, this is no longer true. Now you get the result that Google's algorithm suggests is best for you in particular, and someone else may see something entirely different. In other words, there is no standard Google anymore.

It's not hard to see this difference in action. In the spring of 2010, while the remains of the Deepwater Horizon oil rig were spewing crude oil into the Gulf of Mexico, I asked two friends to search for the term “BP.” They're pretty similar, educated white left-leaning women who live in the Northeast. But the results they saw were quite different. One of my friends saw investment information about BP. The other saw news. For one, the first page of results contained links about the oil spill; for the other, there was nothing about it except for a promotional ad from BP.

Even the number of results returned by Google differed, about 180 million results for one friend and 139 million for the other. If the results were that different for these two progressive East Coast women, imagine how different they would be for my friends and, say, an elderly Republican in Texas (or, for that matter, a businessman in Japan).

With Google personalized for everyone, the query “stem cells” might produce diametrically opposed results for scientists who support stem cell research and activists who oppose it. “Proof of climate change” might turn up different results for an environmental activist and an oil company executive. In polls, a huge majority of us assume search engines are unbiased. But that may be just because they're increasingly biased to share our own views. More and more, your computer monitor is a kind of one-way mirror, reflecting your own interests while algorithmic observers watch what you click.

Google's announcement marked the turning point of an important but nearly invisible revolution in how we consume information. You could say that on December 4, 2009, the era of personalization began.

WITH LITTLE NOTICE or fanfare, the digital world is fundamentally changing. What was once an anonymous medium where anyone could be anyone, where, in the words of the famous New Yorker cartoon, nobody knows you're a dog, is now a tool for soliciting and analyzing our personal data. According to one Wall Street Journal study, the top fifty Internet sites, from CNN to Yahoo to MSN, install an average of 64 data-laden cookies and personal tracking beacons each. Search for a word like “depression” on Dictionary.com , and the site installs up to 223 tracking cookies and beacons on your computer so that other Web sites can target you with antidepressants. Share an article about cooking on ABC News, and you may be chased around the Web by ads for Teflon-coated pots. Open, even for an instant, a page listing signs that your spouse may be cheating and prepare to be haunted with DNA paternity-test ads. The new Internet doesn't just know you're a dog; it knows your breed and wants to sell you a bowl of premium kibble.

The race to know as much as possible about you has become the central battle of the era for Internet giants like Google, Facebook, Apple, and Microsoft. As Chris Palmer of the Electronic Frontier Foundation explained to me, “You're getting a free service, and the cost is information about you. And Google and Facebook translate that pretty directly into money.” While Gmail and Facebook may be helpful, free tools, they are also extremely effective and voracious extraction engines into which we pour the most intimate details of our lives. Your smooth new iPhone knows exactly where you go, whom you call, what you read; with its built-in microphone, gyroscope, and GPS, it can tell whether you're walking or in a car or at a party.

While Google has (so far) promised to keep your personal data to itself, other popular Web sites and apps, from the airfare site Kayak.com to the sharing widget AddThis, make no such guarantees. Behind the pages you visit, a massive new market for information about what you do online is growing, driven by low-profile but highly profitable personal data companies like BlueKai and Acxiom. Acxiom alone has accumulated an average of 1,500 pieces of data on each person on its database, which includes 96 percent of Americans, along with data about everything from their credit scores to whether they've bought medication for incontinence. And using lightning-fast protocols, any Web site, not just the Googles and Facebooks of the world, can now participate in the fun. In the view of the “behavior market” vendors, every “click signal” you create is a commodity, and every move of your mouse can be auctioned off within microseconds to the highest commercial bidder.

As a business strategy, the Internet giants' formula is simple: The more personally relevant their information offerings are, the more ads they can sell, and the more likely you are to buy the products they're offering. And the formula works. Amazon sells billions of dollars in merchandise by predicting what each customer is interested in and putting it in the front of the virtual store. Up to 60 percent of Netflix's rentals come from the personalized guesses it can make about each customer's movie preferences, and at this point, Netflix can predict how much you'll like a given movie within about half a star. Personalization is a core strategy for the top five sites on the Internet, Yahoo, Google, Facebook, YouTube, and Microsoft Live, as well as countless others.

In the next three to five years, Facebook COO Sheryl Sandberg told one group, the idea of a Web site that isn't customized to a particular user will seem quaint. Yahoo Vice President Tapan Bhat agrees: “The future of the web is about personalization ... now the web is about 'me.' It's about weaving the web together in a way that is smart and personalized for the user.” Google CEO Eric Schmidt enthuses that the “product I've always wanted to build” is Google code that will “guess what I'm trying to type.” Google Instant, which guesses what you're searching for as you type and was rolled out in the fall of 2010, is just the start. Schmidt believes that what customers want is for Google to “tell them what they should be doing next.”

It would be one thing if all this customization was just about targeted advertising. But personalization isn't just shaping what we buy. For a quickly rising percentage of us, personalized news feeds like Facebook are becoming a primary news source, 36 percent of Americans under thirty get their news through social networking sites. And Facebook's popularity is skyrocketing worldwide, with nearly a million more people joining each day. As founder Mark Zuckerberg likes to brag, Facebook may be the biggest source of news in the world (at least for some definitions of “news”).

And personalization is shaping how information flows far beyond Facebook, as Web sites from Yahoo News to the New York Times-funded startup News.me cater their headlines to our particular interests and desires. It's influencing what videos we watch on YouTube and a dozen smaller competitors, and what blog posts we see. It's affecting whose e-mails we get, which potential mates we run into on OkCupid, and which restaurants are recommended to us on Yelp, which means that personalization could easily have a hand not only in who goes on a date with whom but in where they go and what they talk about. The algorithms that orchestrate our ads are starting to orchestrate our lives.

The basic code at the heart of the new Internet is pretty simple. The new generation of Internet filters looks at the things you seem to like, the actual things you've done, or the things people like you like, and tries to extrapolate. They are prediction engines, constantly creating and refining a theory of who you are and what you'll do and want next. Together, these engines create a unique universe of information for each of us, what I've come to call a filter bubble, which fundamentally alters the way we encounter ideas and information.

Of course, to some extent we've always consumed media that appealed to our interests and avocations and ignored much of the rest. But the filter bubble introduces three dynamics we've never dealt with before.

First, you're alone in it. A cable channel that caters to a narrow interest (say, golf) has other viewers with whom you share a frame of reference. But you're the only person in your bubble. In an age when shared information is the bedrock of shared experience, the filter bubble is a centrifugal force, pulling us apart.

Second, the filter bubble is invisible. Most viewers of conservative or liberal news sources know that they're going to a station curated to serve a particular political viewpoint. But Google's agenda is opaque. Google doesn't tell you who it thinks you are or why it's showing you the results you're seeing. You don't know if its assumptions about you are right or wrong, and you might not even know it's making assumptions about you in the first place.

My friend who got more investment-oriented information about BP still has no idea why that was the case, she's not a stockbroker. Because you haven't chosen the criteria by which sites filter information in and out, it's easy to imagine that the information that comes through a filter bubble is unbiased, objective, true. But it's not. In fact, from within the bubble, it's nearly impossible to see how biased it is.

Finally, you don't choose to enter the bubble. When you turn on Fox News or read The Nation, you're making a decision about what kind of filter to use to make sense of the world. It's an active process, and like putting on a pair of tinted glasses, you can guess how the editors' leaning shapes your perception. You don't make the same kind of choice with personalized filters. They come to you, and because they drive up profits for the Web sites that use them, they'll become harder and harder to avoid.

THE STRUCTURE OF our media affects the character of our society. The printed word is conducive to democratic argument in a way that laboriously copied scrolls aren't. Television had a profound effect on political life in the twentieth century, from the Kennedy assassination to 9/11, and it's probably not a coincidence that a nation whose denizens spend thirty-six hours a week watching TV has less time for civic life.

The era of personalization is here, and it's upending many of our predictions about what the Internet would do. The creators of the Internet envisioned something bigger and more important than a global system for sharing pictures of pets. The manifesto that helped launch the Electronic Frontier Foundation in the early nineties championed a “civilization of Mind in cyberspace”, a kind of worldwide metabrain. But personalized filters sever the synapses in that brain. Without knowing it, we may be giving ourselves a kind of global lobotomy instead.

From megacities to nanotech, we're creating a global society whose complexity has passed the limits of individual comprehension. The problems we'll face in the next twenty years, energy shortages, terrorism, climate change, and disease, are enormous in scope. They're problems that we can only solve together.

Early Internet enthusiasts like Web creator Tim Berners-Lee hoped it would be a new platform for tackling those problems. I believe it still can be, but first we need to pull back the curtain, to understand the forces that are taking the Internet in its current, personalized direction. We need to lay bare the bugs in the code, and the coders, that brought personalization to us.

If “code is law,” as Larry Lessig famously declared, it's important to understand what the new lawmakers are trying to do. We need to understand what the programmers at Google and Facebook believe in. We need to understand the economic and social forces that are driving personalization, some of which are inevitable and some of which are not. And we need to understand what all this means for our politics, our culture, and our future.

Without sitting down next to a friend, it's hard to tell how the version of Google or Yahoo News that you're seeing differs from anyone else's. But because the filter bubble distorts our perception of what's important, true, and real, it's critically important to render it visible.

Why would artists create work algorithmically?

In the logician's voice:

an algorithm is

a finite procedure,

written in a fixed symbolic vocabulary,

governed by precise instructions,

moving in discrete steps, 1, 2, 3,...,

whose execution requires no insight, cleverness,
intuition, intelligence or perspicuity,

and that sooner or later comes to an end.

David Berlinski

the serial attitude

mel bochner

> *What order-type is universally present wherever there is any order in the world? The answer is, serial order. Any row, array, rank, order of precedence, numerical or quantitative set of values, any straight line, any geometrical figure employing straight lines, and yes, all space and all time.*
>
> - Joshua Royce, Principles of Logic

Serial order is a method, not a style. The results of this method are surprising and diverse. Edward Muybridge's photographs, Thomas Eakins' perspective studies, Jasper Johns' numerals, Alfred Jensen's polyptychs, Larry Poons' circles, dots and ellipsoids, Donald Judd's painted wall pieces, Sol LeWitt's orthogonal multi-part floor structures all are works employing serial logics. This is not a stylistic phenomenon. Variousness of the above kind is sufficient grounds for suggesting that rather than a style we are dealing with an attitude. The serial attitude is a concern with how order of a specific type is manifest.

Many artists work "in series." That is, they make different versions of a basic theme; Morandi's bottles or de Kooning's women, for example. This falls outside the area of concern here. Three basic operating assumptions separate serially ordered works from multiple variants:

1 — The derivation of the terms or interior divisions of the work is by means of a numerical or otherwise systematically predetermined process (permutation, progression, rotation, reversal).

2 — The order takes precedence over the execution.

3 — The completed work is fundamentally parsimonious and systematically self-exhausting.

Serial ideas have occurred in numerous places and in various forms. Muybridge's photographs are an instance of the serialization of time through the systematic subtraction of duration from event. Muybridge simultaneously photographed the same activity from 180°, 90°, and 45° and printed the three sets of photographs parallel horizontally. By setting up alternative reading logics within a visually discontinuous sequence he completely fragmented perception into what Stockhausen called, in another context, a "directionless time-field."

Robert Rauschenberg's *Seven White Panels* and Ellsworth Kelly's orthogonal eight-foot-square *Sixty-Four* are anomalous works of the early 1950s. Both paintings fall within a generalized concept of arrays, which is serial, although their concerns were primarily modular. Modular works are based on the repetition of a standard unit. The unit, which may be anything (Andre's bricks, Morris's truncated volumes, Warhol's soup cans) does not alter its basic form, although it may appear to vary by the way in which units are adjoined. While the addition of identical units may modify simple gestalt viewing, this is a relatively uncomplex order form. Modularity has a history in the "cultural methods of forming" and architectural practice. Frank Stella has often worked within a modular set, although in his concentric square paintings he appears to have serialized color arrangement with the addition of random blank spaces. Some of the early black paintings, like *Die Fahne Hoch*, employed rotational procedures in the organization of quadrants.

Logics which precede the work may be absurdly simple and available. In Jasper Johns' number and alphabet paintings the prime set is either the letters A-Z or the numbers 0–9. Johns chose to utilize convention. The convention happened to be serial. Without deviating from the accustomed order of precedence he painted all the numbers or letters, in turn, beginning again at the end of each sequence until all the available spaces on the canvas were filled. The procedure was self-exhausting and solipsistic. Other works of Johns are noteworthy in this context, especially his *Three Flags*, which is based on size diminution and, of course, the map paintings. His drawings in which all the integers 0–9 are superimposed are examples of a straightforward use of simultaneity.

An earlier example of simultaneity appears in Marcel Duchamp's *Nude Descending a Staircase*. Using the technique of superimposition and transparency he divided the assigned canvas into a succession of time intervals. Due to the slight variation in density it is impossible to visualize specific changes as such. Alternations are leveled to a single information which subverts experiential time. Duchamp has said the idea was suggested to him by the experiments of Dr. Etienne Jules Marey (1830–1904). Marey, a French physiologist, began with ideas derived from the work of Muybridge, but made a number of significant conceptual and mechanical changes. He invented an ingenious optical device based on principles of revolution similar to Gatling's machine gun. This device enabled him to photograph multiple points of view on one plate. In 1890 he invented his "chronophotograph," which was capable of recording, in succession, 120 separate photos per second. He attempted to visualize the passage of time by placing a clock within camera range, obtaining by this method a remarkable "dissociation of time and image."

Types of order are forms of thoughts. They can be studied apart from whatever physical form they may assume. Before observing some further usages of seriality in the visual arts, it will be helpful to survey several other areas where parallel ideas and approaches also exist. In doing this I wish to imply neither metaphor nor analogy.

> *My desire was for a* conscious control *over the new means and forms that arise in every artist's mind.*
>
> - Arnold Schoenberg

Music has been consistently engaged with serial ideas. Although the term "serial music" is relatively contemporary, it could be easily applied to Bach or even Beethoven. In a serial or Dodecaphonic (twelve tone) composition, the order of the notes throughout the piece is a consequence of an initially chosen and ordered set (the semitonal scale arranged in a definite linear order). Note distribution is then arrived at by permuting this prime set. Any series of notes (or numbers) can be subjected to permutation as follows: 2 numbers have only 2 permutations (1, 2; 2, 1); 3 numbers have 6 (1, 2, 3; 1, 3, 2; 2, 1, 3; 2, 3, 1; 3, 1, 2; 3, 2, 1); 4 numbers have 24; . . . 12 numbers have 479,001,600.

Other similarly produced numerical sequences and a group of pre-established procedures give the exact place in time for each sound, the coincidence of sounds, their duration, timbre and pitch.

The American serial composer Milton Babbit's *Three Compositions for Piano* can be used as a simplified example of this method (see George Perle's Serial Composition and Atonality for a more detailed analysis). The prime set is represented by these integers: P = 5, 1, 2, 4. By subtracting each number in turn from a constant of such value that the resulting series introduces no numbers not already given, an inversion results (in this case the constant is 6): I = 1, 5, 4, 2. A rotational procedure applied to P and I yields the third and fourth set forms: Rp = 2, 4, 5, 1; Ri = 4, 2, 1, 5.

> *Mathematics – or more correctly arithmetic – is used as a compositional device, resulting in the most literal sort of "programme music," but one whose course is determined by a numerical rather than a narrative or descriptive "programme."*
>
> - Milton Babbit

The composer is freed from individual note-to-note decisions which are self-generating within the system he devises. The music thus attains a high degree of conceptual coherence, even if it sometimes sounds "aimless and fragmentary."

> *The adaptation if the serial concept of composition by incorporating the more general notion of permutation into structural organization – a permutation the limits of which are rigorously defined inb terms if the restrictions place on its self-determination constitutes a logical and fully justified development, since both morphology and rhetoric are governed by one and the same principle.*
>
> -Pierre Boulez
>
> *The form itself is of very limited importance, it becomes the grammar of the total work.*
>
> - Sol Lewitt
>
> *Language can be approached in either of two ways, as a set of culturally transmitted behavior patterns shared by a group or as a system conforming to the rules which constitute its grammar.*
>
> - Joseph Greenberg, *Essays in Linguistics*

In linguistic analysis, language is often considered as a system of elements without assigned meanings ("uninterpreted systems"). Such systems are completely permutational, having grammatical but not semantic rules. Since there can be no system without rules of arrangement, this amounts to the handling of language as a set of probabilities. Many interesting observations have been made about uninterpreted systems which are directly applicable to the investigation of any array of elements obeying fixed rules of combination. Studies of isomorphic (correspondence) relationships are especially interesting.

Practically all systems can be rendered isomorphic with a system containing only one serial relation. For instance, elements can be reordered into a single line, i. e., single serial relation by arranging them according to their coordinates. In the following two-dimensional array, the coordinates of C are (1, 3), of T (3, 2):

R	P	D
L	B	T
C	U	O

Isomorphs could be written as: R, L, C, P, B, U, D, T, O or R, P, D, L, B, T, C, U, O.
An example of this in language is the ordering in time of speech to correspond to the ordering of direction in writing. All the forms of cryptography from crossword puzzles to highly sophisticated codes depend on systematic relationships of this kind.

> *The limits of my language are the limits of my world.*
>
> - Ludwig Wittgenstein

(. . .)

> *The structure of an artificial optic array may, but need not, specify a source. A wholly invented structure need not specify anything. This would be a case of structure as such. It contains information but not information* about*, and it affords perception but not perception of.*
>
> - James J. Gibson, *The Senses Considered as Perceptual Systems*

Perspective, almost universally dismissed as a concern in recent art, is a fascinating example of the application of prefabricated systems. In the work of artists like Ucello, Durer, Piero, Saendredam, Eakins (especially their drawings), it can be seen to exist entirely as methodology. It demonstrates not how things appear but rather the workings of its own strict postulates. As it is, these postulates are serial.

Perspective has had an oddly circular history. Girard Desargues (1593–1662) based his non-Euclidean geometry on an intuition derived directly from perspective. Instead of beginning with the unverifiable Euclidean axiom that parallel lines never meet, he accepted instead the visual evidence that they do meet at the point where they intersect on the horizon line (the “vanishing point” or “infinity” of perspective). Out of his investigations of “visual” (as opposed to “tactile”) geometry came the field of projective geometry. Projective geometry investigates such problems as the means of projecting figures from the surface of three-dimensional objects to two-dimensional planes. It has led to the solution of some of the problems in mapmaking. Maps are highly abstract systems, but since distortion of some sort must occur in the transformation from three to two dimensions, maps are never completely accurate. To compensate for distortion, various systems have been devised. On a topographical map, for example, the lines indicating levels (contour lines) run through points which represent physical points on the surface mapped so that an isomorphic relation can be established. Parallels of latitude, isobars, isothermal lines and other grid coordinate denotations, all serialized, are further cases of the application of external structure systems to order the unordered.

Another serial aspect of mapmaking is a hypothesis in topology about color. It states that with only four colors all the countries on any map can be differentiated without any color having to appear adjacent to itself. (One wonders what the results might look like if all the paintings in the history of art were repainted to conform to the conditions of this hypothesis.) (. . .)

This text was published in Artforum, 6:4 (December 1967), pp. 28–33.

"Music as a Gradual Process," by Steve Reich

I do not mean the process of composition, but rather pieces of music that are, literally, processes.

The distinctive thing about musical processes is that they determine all the note-to-note (sound-to-sound) details and the over all form simultaneously. (Think of a round or infinite canon.)

I am interested in perceptible processes. I want to be able to hear the process happening throughout the sounding music.

To facilitate closely detailed listening a musical process should happen extremely gradually.

Performing and listening to a gradual musical process resembles:

- pulling back a swing, releasing it, and observing it gradually come to rest;
- turning over an hour glass and watching the sand slowly run through the bottom;
- placing your feet in the sand by the ocean's edge and watching, feeling, and listening to the waves gradually bury them.

Though I may have the pleasure of discovering musical processes and composing the musical material to run through them, once the process is set up and loaded it runs by itself.

Material may suggest what sort of process it should be run through (content suggests form), and processes may suggest what sort of material should be run through them (form suggests content). If the shoe fits, wear it.

As to whether a musical process is realized through live human performance or through some electro-mechanical means is not finally the main issue. One of the most beautiful concerts I ever heard consisted of four composers playing their tapes in a dark hall. (A tape is interesting when it's an interesting tape.)

It is quite natural to think about musical processes if one is frequently working with electro-mechanical sound equipment. All music turns out to be ethnic music.

Musical processes can give one a direct contact with the impersonal and also a kind of complete control, and one doesn't always think of the impersonal and complete control as going together. By "a kind" of complete control I mean that by running this material through the process I completely control all that results, but also that I accept all that results without changes.

John Cage has used processes and has certainly accepted their results, but the processes he used were compositional ones that could not be heard when the piece was performed. The process of using the *I Ching* or imperfections in a sheet of paper to determine musical parameters can't be heard when listening to music composed that way. The compositional processes and the sounding music have no audible connection. Similarly in serial music, the series itself is seldom audible. (This is a basic difference between serial (basically European) music and serial (basically American) art, where the perceived series is usually the focal point of the work.)

What I'm interested in is a compositional process and a sounding music that are one and the same thing.

James Tenney said in conversation, "then the composer isn't privy to anything". I don't know any secrets of structure that you can't hear. We all listen to the process together since it's quite audible, and one of the reasons it's quite audible is, because it's happening extremely gradually.

The use of hidden structural devices in music never appealed to me. Even when all the cards are on the table and everyone hears what is gradually happening in a musical process, there are still enough mysteries to satisfy all. These mysteries are the impersonal, unattended, psycho-acoustic by-products of the intended process. These might include sub-melodies heard within repeated melodic patterns, stereophonic effects due to listener location, slight irregularities in performance, harmonics, difference tones, etc.

Listening to an extremely gradual musical process opens my ears to *it*, but *it* always extends farther than I can hear, and that makes it interesting to listen to the musical process again. That area of every gradual (completely controlled) musical process, where one hears the details of the sound moving out away from intentions, occurring for their own acoustic reasons, is *it*.

I begin to perceive these minute details when I can sustain close attention and a gradual process invites my sustained attention. By "gradual" I mean extremely gradual; a process happening so slowly and gradually that listening to it resembles watching a minute hand on a watch--you can perceive it moving after you stay with it a little while.

Several currently popular modal musics like Indian classical and drug oriented rock and roll may make us aware of minute sound details because in being modal (constant key center, hypnotically droning and repetitious) they naturally focus on these details rather than on key modulation, counterpoint and other peculiarly Western devices. Nevertheless, these modal musics remain more or less strict frameworks for improvisation. They are not processes.

The distinctive thing about musical processes is that they determine all the note-to-note details and the over all form simultaneously. One can't improvise in a musical process--the concepts are mutually exclusive.

While performing and listening to gradual musical processes one can participate in a particular liberating and impersonal kind of ritual. Focusing in on the musical process makes possible that shift of attention away from *he* and *she* and *you* and *me* outwards towards *it.*

The Cut-Up Method of Brion Gysin

William S. Burroughs

At a surrealist rally in the 1920s Tristan Tzara the man from nowhere proposed to create a poem on the spot by pulling words out of a hat. A riot ensued wrecked the theater. André Breton expelled Tristan Tzara from the movement and grounded the cut-ups on the Freudian couch.

In the summer of 1959 Brion Gysin painter and writer cut newspaper articles into sections and rearranged the sections at random. Minutes to Go resulted from this initial cut-up experiment. Minutes to Go contains unedited unchanged cut ups emerging as quite coherent and meaningful prose. The cut-up method brings to writers the collage, which has been used by painters for fifty years. And used by the moving and still camera. In fact all street shots from movie or still cameras are by the unpredictable factors of passers by and juxtaposition cut-ups. And photographers will tell you that often their best shots are accidents . . . writers will tell you the same. The best writing seems to be done almost by accident but writers until the cut-up method was made explicit— all writing is in fact cut ups. I will return to this point—had no way to produce the accident of spontaneity. You can not will spontaneity. But you can introduce the unpredictable spontaneous factor with a pair of scissors.

The method is simple. Here is one way to do it. Take a page. Like this page. Now cut down the middle and cross the middle. You have four sections: 1 2 3 4 . . . one two three four. Now rearrange the sections placing section four with section one and section two with section three. And you have a new page. Sometimes it says much the same thing. Sometimes something quite different—cutting up political speeches is an interesting exercise—in any case you will find that it says something and something quite definite. Take any poet or writer you fancy. Here, say, or poems you have read over many times. The words have lost meaning and life through years of repetition. Now take the poem and type out selected passages. Fill a page with excerpts. Now cut the page. You have a new poem. As many poems as you like. As many Shakespeare Rimbaud poems as you like. Tristan Tzara said: "Poetry is for everyone." And André Breton called him a cop and expelled him from the movement. Say it again: "Poetry is for everyone." Poetry is a place and it is free to all cut up Rimbaud and you are in Rimbaude is a Rimbaud poem cut up.

Visit of memories. Only your dance and your voice house. On the suburban air improbable desertions ... all harmonic pine for strife.

The great skies are open. Candor of vapor and tent spitting blood laugh and drunken penance.

Promenade of wine perfume opens slow bottle.

The great skies are open. Supreme bugle burning flesh children to mist.

Cut-ups are for everyone. Anybody can make cut ups. It is experimental in the sense of being something to do. Right here write now. Not something to talk and argue about. Greek philosophers assumed logically that an object twice as heavy as another object would fall twice as fast. It did not occur to them to push the two objects off the table and see how they fall. Cut the words and see how they fall.

Shakespeare Rimbaud live in their words. Cut the word lines and you will hear their voices. Cut-ups often come through as code messages with special meaning for the cutter. Table tapping? Perhaps. Certainly an improvement on the usual deplorable performance of contacted poets through a medium. Rimbaud announces himself, to be followed by some excruciatingly bad poetry. Cutting Rimbaud and you are assured of good poetry at least if not personal appearance.

All writing is in fact cut-ups. A collage of words read heard overhead. What else? Use of scissors renders the process explicit and subject to extension and variation. Clear classical prose can be composed entirely of rearranged cut-ups. Cutting and rearranging a page of written words introduces a new dimension into writing enabling the writer to turn images in cinematic variation. Images shift sense under the scissors smell images to sound sight to sound sound to kinesthetic. This is where Rimbaud was going with his color of vowels. And his "systematic derangement of the senses." The place of mescaline hallucination: seeing colors tasting sounds smelling forms.

The cut-ups can be applied to other fields than writing. Dr Neumann in his Theory of Games and Economic Behavior introduces the cut-up method of random action into game and military strategy: assume that the worst has happened and act accordingly. If your strategy is at some point determined . . . by random factor your opponent will gain no advantage from knowing your strategy since he can not predict the move. The cut-up method could be used to advantage in processing scientific data. How many discoveries have been made by accident? We can not produce accidents to order. The cut-ups could add new dimension to films. Cut gambling scene in with a thousand gambling scenes all times and places. Cut back. Cut streets of the world. Cut and rearrange the word and image in films. There is no reason to accept a second-rate product when you can have the best. And the best is there for all. "Poetry is for everyone" . . .

Now here are the preceding two paragraphs cut into four sections and rearranged:

ALL WRITING IS IN FACT CUT-UPS OF GAMES AND ECONOMIC BEHAVIOR OVERHEARD? WHAT ELSE? ASSUME THAT THE WORST HAS HAPPENED EXPLICIT AND SUBJECT TO STRATEGY IS AT SOME POINT CLASSICAL PROSE. CUTTING AND REARRANGING FACTOR YOUR OPPONENT WILL GAIN INTRODUCES A NEW DIMENSION YOUR STRATEGY. HOW MANY DISCOVERIES SOUND TO KINESTHETIC? WE CAN NOW PRODUCE ACCIDENT TO HIS COLOR OF VOWELS. AND NEW DIMENSION TO FILMS CUT THE SENSES. THE PLACE OF SAND. GAMBLING SCENES ALL TIMES COLORS TASTING SOUNDS SMELL STREETS OF THE WORLD. WHEN YOU CAN HAVE THE BEST ALL: "POETRY IS FOR EVERYONE" DR NEUMANN IN A COLLAGE OF WORDS READ HEARD INTRODUCED THE CUT-UP SCISSORS RENDERS THE PROCESS GAME AND MILITARY STRATEGY, VARIATION CLEAR AND ACT ACCORDINGLY. IF YOU POSED ENTIRELY OF REARRANGED CUT DETERMINED BY RANDOM A PAGE OF WRITTEN WORDS NO ADVANTAGE FROM KNOWING INTO WRITER PREDICT THE MOVE. THE CUT VARIATION IMAGES SHIFT SENSE ADVANTAGE IN PROCESSING TO SOUND SIGHT TO SOUND. HAVE BEEN MADE BY ACCIDENT IS WHERE RIMBAUD WAS GOING WITH ORDER THE CUT-UPS COULD "SYSTEMATIC DERANGEMENT" OF THE GAMBLING SCENE IN WITH A TEA HALLUCINATION: SEEING AND PLACES. CUT BACK. CUT FORMS. REARRANGE THE WORD AND IMAGE TO OTHER FIELDS THAN WRITING.

What is cybernetics?

7
Emergence of Control

In ancient Greece the first artificial self

The invention of autonomous control, like most inventions, has roots in ancient China. There, on a dusty windswept plain, a small wooden statue of a man in robes teeters upon a short pole. The pole is carried between a pair of turning wagon wheels, pulled by two red horses outfitted in bronze finery.

The statue man, carved in the flowing dresses of 9th-century China, points with outstretched hand towards a distant place. By the magic of noisy gears connecting the two wooden wheels, as the cart races along the steppes, the wooden man perched on the stick invariably, steadily, without fail, points south. When the cart turns left or right, the geared wheels calculate the change and swing the wooden man's (or is it a god's?) arm a corresponding amount in the opposite direction, negating the cart's shift and keeping the guide forever pointing to the south. With an infallible will, and on his own accord, the wooden figure automatically seeks south. The south-pointing chariot precedes a lordly procession, preventing the party from losing its way in the desolate countryside of old China.

How busy was the ingenious medieval mind of China! Peasant folk in the backwaters of southwestern China, wishing to temper the amount of wine downed in the course of a fireside toast, came upon a small device which, by its own accord, would control the rowdy spirits of the wine. Chou Ch'u-Fei, a traveler among the Ch'i Tung natives then, reported that drinking bouts in this kingdom had been perfected by means of a two-foot-long bamboo straw which automatically regulated wine consumption, giving large-throated and small-mouthed drinkers equal advantage. A "small fish made of silver" floated inside the straw. The downward weight of the internal metal float restricted the flow of warm plum wine if the drinker sucked too feebly (perhaps through intoxication), thereby calling an end for his evening of merriment. If he inhaled too boisterously, he also got nothing, as the same float became wedged upwards by force of the suction. Only a temperate, steady draw was profitable.

Upon inspection, neither the south-pointing carriage nor the wine straw are truly automatic in a modern (self-steering) sense. Both devices merely tell their human masters, in the most subtle and unconscious way, of the adjustment needed to keep the action constant, and leave the human to make the change in direction of travel or power of lung. In the lingo of modern thinking, the human is part of the loop. To be truly automatic, the south-pointing statue would have to turn the cart itself, to make it a south-*heading* carriage. Or a carrot would have to be dangled from the point of his finger so that the horses (now in the loop) followed it. Likewise the drinking straw would have to regulate its volume no matter how hard one sucked. Although not automatic, the south-pointing cart is based on the differential gear, a thousand-year-old predecessor to the automobile transmission, and an early prototype of modern self-pointing guns on an armored tank which aid the drivers inside where a magnetic compass is useless. Thus, these clever devices are curious stillbirths in our genealogy of automation. The very first truly automatic devices had actually been built long before, a millennia earler.

Ktesibios was a barber who lived in Alexandria in the first half of the third century B.C. He was obsessed with mechanical devices, for which he had a natural genius. He eventually became a proper mechanician—a builder of artifactual creations—under King Ptolemy II. He is credited with having invented the pump, the water organ, several kinds of catapults, and a legendary water clock. At the time, Ktesibios's fame as an inventor rivaled that of the legendary engineer Archimedes. Today, Ktesibios is credited with inventing the first honest-to-goodness automatic device.

Ktesibios's clock kept extraordinarily good time (for then) by self-regulating its water supply. The weakness of most water clocks until that moment was that as the reservoir of water propelling the drive mechanism emptied, the speed of emptying would gradually decrease (because a shallow level of water provides less pressure than a high level), slowing down the clock's movements. Ktesibios got around this perennial problem by inventing a regulating valve *(regula)* comprised of a float in the shape of a cone which fit its nose into a mating inverted funnel. Within the *regula,* water flowed from the funnel stem, over the cone, and into the bowl the cone swam in. The cone would then float up into the concave funnel and constrict the water passage, thus throttling its flow. As the water diminished, the float would sink, opening the passage again and allowing more water in. The *regula* would immediately seek a compromise position where it would let "just enough" water for a constant flow through the metering valve vessel.

Ktesibios's *regula* was the first nonliving object to self-regulate, self- govern, and self-control. Thus, it became the first *self* to be born outside of biology. It was a true *auto* thing—directed from within. We now consider it to be the primordial automatic device because it held the first breath of lifelikeness in a machine.

It truly was a *self* because of what it displaced. A constant autoregulated flow of water translated into a constant autoregulated clock and relieved a king of the need for servants to tend the water clock's water vessels. In this way, "auto-self" shouldered out the human self. From the very first instance, automation replaced human work.

Ktesibios's invention is first cousin to that all-American 20th-century fixture, the flush toilet. Readers will recognize the Ktesibios floating valve as the predecessor to the floating ball in the upper chamber of the porcelain throne. After a flush, the floating ball sinks with the declining water level, pulling open the water valve with its metal arm. The incoming water fills the vessel again, raising the ball triumphantly so that its arm closes the flow of water at the precise level of "full." In a medieval sense, the toilet yearns to keep itself full by means of this automatic plumbing. Thus, in the bowels of the flush toilet we see the archetype for all autonomous mechanical creatures.

About a century later, Heron, working in the same city of Alexandria, came up with a variety of different automatic float mechanisms, which look to the modern eye like a series of wildly convoluted toilet mechanisms. In actuality, these were elaborate party wine dispensers, such as the "Inexhaustible Goblet" which refilled itself to a constant level from a pipe fitted into its bottom. Heron wrote a huge encyclopedia (the *Pneumatica)* crammed with his incredible (even by today's standards) inventions. The book was widely translated and copied in the ancient world and was influential beyond measure. In fact, for 2,000 years (that is, until the age of machines in the 18th century), no feedback systems were invented that Heron had not already fathered.

The one exception was dreamed up in the 17th century by a Dutch alchemist, lens grinder, pyromaniac, and hobby submariner by the name of Cornelis Drebbel. (Drebbel made more than one successful submarine dive around 1600!) While tinkering in his search for gold, Drebbel invented the thermostat, the other universal example of a feedback system. As an alchemist, Drebbel suspected that the transmutation of lead into gold in a laboratory was inhibited by great temperature fluctuations of the heat sources cooking the elements. In the 1620s he jerry-rigged a minifurnace which could bake the initial alchemic mixture over moderate heat for a very long time, much as might happen to gold-bearing rock bordering the depths of Hades. On one side of his ministove, Drebbel attached a glass tube the size of a pen filled with alcohol. The liquid would expand when heated, pushing mercury in a connecting second tube, which in turn would push a rod that would close an air draft on the stove. The hotter the furnace, the futher the draft would close, decreasing the fire. The cooling tube retracted the rod, thus opening the draft and increasing the fire. An ordinary suburban tract home thermostat is conceptually identical—both seek a constant temperature. Unfortunately, Drebbel's automatic stove didn't make gold, nor did Drebbel ever publish its design, so his automatic invention perished without influence, and its design had to be rediscovered a hundred years later by a French gentleman farmer, who built one to incubate his chicken eggs.

James Watt, who is credited with inventing the steam engine, did not. Working steam engines had been on the job for decades before Watt ever saw one. As a young engineer, Watt was once asked to repair a small-scale model of an early working, though inefficient, Newcomen steam engine. Frustrated by its awkwardness, Watt set out to improve it. At about the time of the American Revolution, he added two things to the existing engines; one of them evolutionary, the other revolutionary. His key evolutionary innovation was separating the heating chamber from the cooling chamber; this made his engine extremely powerful. So powerful that he needed to add a speed regulator to moderate this newly unleashed machine power. As usual Watt turned to what already existed. Thomas Mead, a mechanic and miller, had invented a clumsy centrifugal regulator for a windmill that would lower the millstone onto the grain only when stone's speed was sufficient. It regulated the output but not the power of a millstone.

Watt contrived a radical improvement. He borrowed Mead's regulator from the mill and revisioned it into a pure control circuit. By means of his new regulator the steam machine gripped the throat of its own power. His completely modern *regula* automatically stabilized his now ferocious motor at a constant speed of the operator's choice. By adjusting the governor, Watt could vary the steam engine to run at any rate. This was revolutionary.

Like Heron's float and Drebbel's thermostat, Watt's centrifugal governor is transparent in its feedback. Two leaden balls, each at the end of a stiff pendulum, swing from a pole. As the pole rotates the balls spin out levitating higher the faster the system spins. Linkages scissored from the twirling pendulums slide up a sleeve on the pole, levering a valve which controls the speed of rotation by adjusting the steam. The higher the balls spin, the more the linkages close the valve, reducing the speed, until an equilibrium point of constant rpms (and height of spinning balls) is reached. The control is thus as dependable as physics.

Rotation is an alien power in nature. But among machines, it is blood. The only known bearing in biology is at the joint of a sperm's spinning hair propeller. Outside of this micromotor, the axle and wheel are unknown to those with genes. To the ungened machine, whirling wheels and spinning shafts are reasons to live. Watt gave machines the secret to controlling their own revolutions, which was his revolution. His innovation spread widely and quickly. The mills of the industrial age were fueled by steam, and the engines earnestly regulated themselves with the universal badge of self-control: Watt's flyball governor. Self-powered steam begat machine mills which begat new kinds of engines which begat new machine tools. In all of them, self-regulators dwelt, fueling the principle of snowballing advantages. For every one person visibly working in a factory, thousands of governors and self-regulators toiled invisibly. Today, hundreds of thousands of regulators, unseen, may work in a modern plant at once. A single human may be their coworker.

Watt took the volcanic fury of expanding steam and tamed it with information. His flyball governor is undiluted informational control, one of the first non-biological circuits. The difference between a car and an exploding can of gasoline is that the car's information—its design—tames the brute energy of the gas. The same amount of energy and matter are brought together in a car burning in a riot and one speeding laps in the Indy 500. In the latter case, a critical amount of information rules over the system, civilizing the dragon of fire. The full heat of fire is housetrained by small amounts of self-perception. Furious energy is educated, brought in from the wilds to work in the yard, in the basement, in the kitchen, and eventually in living rooms.

The steam engine is an unthinkable contraption without the domesticating loop of the revolving governor. It would explode in the face of its inventors without that tiny heart of a self. The immense surrogate slave power released by the steam engine ushered in the Industrial Revolution. But a second, more important revolution piggybacked on it unnoticed. There could not have been an industrial revolution without a parallel (though hidden) information revolution at the same time, launched by the rapid spread of the automatic feedback system. If a fire-eating machine, such as Watt's engine, lacked self-control, it would have taken every working hand the machine displaced to babysit its energy. So information, and not coal itself, turned the power of machines useful and therefore desirable.

The industrial revolution, then, was not a preliminary primitive stage required for the hatching of the more sophisticated information revolution. Rather, automatic horsepower was, itself, the first phase of the knowledge revolution. Gritty steam engines, not teeny chips, hauled the world into the information age.

Maturing of mechanical selfhood

Heron's regulator, Drebbel's thermostat, and Watt's governor bestowed on their vessels a wisp of self-control, sensory awareness, and the awakening of anticipation. The governing system sensed its own attributes, noted if it had changed in a certain respect since it last looked, and if it had, it adjusted itself to conform to a goal. In the specific case of a thermostat, the tube of alcohol detected the system's temperature, and then took action or not to tweak the fire in order to align itself with the fixed goal of a certain temperature. It had, in a philosophical sense, a purpose.

Although it may strike us as obvious now, it took a long while for the world's best inventors to transpose even the simplest automatic circuit such as a feedback loop into the realm of electronics. The reason for the long delay was that from the moment of its discovery electricity was seen primarily as power and not as communication. The dawning distinction of the two-faced nature of the spark was acknowledged among leading German electrical engineers of the last century as the split between the techniques of strong current and the techniques of weak current. The amount of energy needed to send a signal is so astoundingly small that electricity had to be reimagined as something altogether different from power. In the camp of the wild-eyed German signalists, electricity was a sibling to the speaking mouth and the writing hand. The inventors (we would call them hackers now) of weak current technology brought forth perhaps the least precedented invention of all time—the telegraph. With this device human communication rode on invisible particles of lightning. Our entire society was reimagined because of this wondrous miracle's descendants.

Telegraphers had the weak model of electricity firmly in mind, yet despite their clever innovations, it wasn't until August 1929, that telephone engineer H. S. Black, working at Bell Laboratories, tamed an electrical feedback loop. Black was hunting for a way to make durable amplifier relays for long-distance phone lines. Early amplifiers were made of crude materials that tended to disintegrate over use, causing the amp to "run away." Not only would an aging relay amplify the phone signal, it would mistakenly compound any tiny deviation from the range it expected until the mushrooming error filled and killed the system. What was needed was Heron's *regula,* a counter signal to rein in the chief signal, to dampen the effect of the perpetual recycling. Black came up with a *negative* feedback loop, which was designated negative in contrast to the snowballing positive loop of the amplifier. Conceptually, the electrical negative feedback loop is a toilet flusher or thermostat. This braking circuit keeps the amplifier honed in on a steady amplification in the same way a thermostat hones in on a steady temperature. But instead of metallic levers, a weak train of electrons talks to itself. Thus, in the byways of the telephone switching network, the first *electrical self* was born.

From World War I and after, the catapults that launched missiles had become so complicated, and their moving targets so sophisticated, that calculating ballistic trajectories taxed human talent. Between battles, human calculators, called computers, computed the settings for firing large guns under various wind, weather and altitude conditions. The results were sometimes printed in pocket-size tables for the gunmen on the front line, or if there was enough time and the missile-gun was common, the tables were mechanically encoded into an apparatus on the gun, known as the automaton. In the U.S., the firing calculations were compiled in a laboratory set up at the Navy's Aberdeen Proving Ground in Maryland, where rooms full of human computers (almost exclusively women) employed hand-cranked adding machines to figure the tables.

By World War II, the German airplanes which the big guns boomed at were flying as fast as the missiles themselves. Speedier on-the spot calculations were needed, ideally ones that could be triggered from measurements of planes in flight made by the newly invented radar scanner. Besides, Navy gunmen had a weighty problem: how to move and aim these monsters with the accuracy the new tables gave them. The solution was as close at hand as the stern of the ship: a large ship controlled its rudder by a special type of automatic feedback loop known as a servomechanism.

Servomechanisms were independently and simultaneously invented a continent apart by an American and a Frenchman around 1860. It was the Frenchman, engineer Leon Farcot, who tagged the device with a name that stuck: *moteur asservi,* or servo-motor. As boats had increased in size and speed over time, human power at the tiller was no longer sufficient to move the rudder against the force of water surging beneath. Marine technicians came up with various oil-hydraulic systems that amplified the power of the tiller so that gently swinging the miniature tiller at the captain's helm would move the mighty rudder, kind of. A repeated swing of the minitiller would translate into different amounts of steerage of the rudder depending on the speed of the boat, waterline, and other similar factors. Farcot invented a linkage system that connected the position of the heavy rudder underwater back to the position of the easy-to-swing tiller—the automatic feedback loop! The tiller then indicated the actual location of the rudder, and by means of the loop, moving the indicator moved the reality. In the jingo of current computerese, What you see is what you get!

The heavy gun barrels of World War II were animated the same way. A hydraulic hose of compressed oil connected a small pivoting lever (the tiller) to the pistons steering the barrel. As the shipmate's hand moved the lever to the desired location, that tiny turn compressed a small piston which would open a valve releasing pressurized oil, which would nudge a large piston moving the heavy gun barrel. But as the barrel swung it would push a small piston that, in return, moved the hand lever. As he tried to turn the tiller, the sailor would feel a mild resistance, a force created by the feedback from the rudder he wanted to move.

Bill Powers was a teenage Electronic Technician's Mate who worked with the Navy's automated guns, and who later pursued control systems as explanation for living things. He describes the false impression one gets by reading about servomechanism loops:

> The sheer mechanics of speaking or writing stretches out the action so it seems that there is a sequence of well-separated events, one following the other. If you were trying to describe how a gun-pointing servomechanism works, you might start out by saying, "Suppose I push down on the gun-barrel to create a position error. The error will cause the servo motors to exert a force against the push, the force getting larger as the push gets larger." That seems clear enough, but it is a lie. If you really did this demonstration, you would say "Suppose I push down on the gun-barrel to create an error...wait a minute. It's stuck."
>
> No, it isn't stuck. It's simply a good control system. As you begin to push down, the little deviation in sensed position of the gun-barrel causes the motor to twist the barrel up against your push. The amount of deviation needed to make the counteractive force equal to the push is so small that you can neither see nor feel it. As a result, the gun-barrel feels as rigid as if it were cast in concrete. It creates the appearance of one of those old-fashioned machines that is immovable simply because it weighs 200 tons, but if someone turned off the power the gun-barrel would fall immediately to the deck.

Servomechanisms have such an uncanny ability to aid steering that they are still used (in updated technology) to pilot boats, to control the flaps in airplanes, and to wiggle the fingers in remotely operated arms handling toxic and nuclear waste.

More than the purely mechanical self-hood of the other regulators like Heron's valve, Watt's governor, and Drebbel's thermostat, the servomechanism of Farcot suggested the possibility of a man-machine symbiosis—a joining of two worlds. The pilot merges into the servomechanism. He gets power, it gets existence. Together they steer. These two aspects of the servomechanisms—steering and symbiosis—inspired one of the more colorful figures of modern science to recognize the pattern that connected these control loops.

The toilet: archetype of tautology

OF ALL THE MATHEMATICIANS assigned during World War I to the human calculating lab in charge of churning out more accurate firing tables at the Aberdeen Proving Grounds, few were as overqualified as Private Norbert Wiener, a former math prodigy whose genius had an unorthodox pedigree.

The ancients recognized genius as something given rather than created. But America at the turn of the century was a place where the wisdom of the past was often successfully challenged. Norbert's father, Leo Wiener, had come to America to launch a vegetarian commune. Instead, he was distracted with other untraditional challenges, such as bettering the gods. In 1895, as a Harvard professor of Slavic languages, Leo Wiener decided that his firstborn son was going to be a genius. A genius deliberately made, not born.

Norbert Wiener was thus born into high expectations. By the age of three he was reading. At 18 he earned his Ph.D. from Harvard. By 19 he was studying metamathematics with Bertrand Russell. Come 30 he was a professor of mathematics at MIT and a thoroughly odd goose. Short, stout, splay-footed, sporting a goatee and a cigar, Wiener waddled around like a smart duck. He had a legendary ability to learn while slumbering. Numerous eyewitnesses tell of Wiener sleeping during a meeting, suddenly awakening at the mention of his name, and then commenting on the conversation that passed while he dozed, usually adding some penetrating insight that dumbfounded everyone else.

In 1948 he published a book for nonspecialists on the feasibility and philosophy of machines that learn. The book was initially published by a French publisher (for roundabout reasons) and went through four printings in the United States in its first six months, selling 21,000 copies in the first decade of its influence—a best seller then. It rivaled the success of the Kinsey Report on sexual behavior, issued the same year. As a *Business Week* reporter observed in 1949, "In one respect Wiener's book resembles the Kinsey Report: the public response to it is as significant as the content of the book itself."

Wiener's startling ideas sailed into the public mind, even though few could comprehend his book, by means of the wonderfully colorful name he coined for both his perspective and the book: *Cybernetics.* As has been noted by many writers, cybernetics derives from the Greek for "steersman"—a pilot that steers a ship. Wiener, who worked with servomechanisms during World War II, was struck by their uncanny ability to aid steering of all types. What is usually not mentioned is that cybernetics was also used in ancient Greece to denote a governor of a country. Plato attributes Socrates as saying, "Cybernetics saves the souls, bodies, and material possessions from the gravest dangers," a statement that encompasses both shades of the word. Government (and that meant self-government to these Greeks) brought order by fending off chaos. Also, one had to actively steer to avoid sinking the ship. The Latin corruption of *kubernetes* is the derivation of *governor,* which Watt picked up for his cybernetic flyball.

The managerial nature of the word has further antecedent to French speakers. Unbeknownst to Wiener, he was not the first modern scientist to reactivate this word. Around 1830 the French physicist Ampere (whence we get the electrical term amperes, and its shorthand "amp") followed the traditional manner of French grand scientists and devised an elaborate classification system of human knowledge. Ampere designated one branch the realm of "Noological Sciences," with the subrealm of Politics. Within political science, immediately following the sub-subcategory of Diplomacy, Ampere listed the science of Cybernetics, that is, the science of governance.

Wiener had in mind a more explicit definition, which he stated boldly in the full title of his book, *Cybernetics: or control and communication in the animal and the machine.* As Wiener's sketchy ideas were embodied by later computers and fleshed out by other theorists, cybernetics gradually acquired more of the flavor of Ampere's governance, but without the politics.

The result of Wiener's book was that the notion of feedback penetrated almost every aspect of technical culture. Though the central concept was both old and commonplace in specialized circumstances, Wiener gave the idea legs by generalizing the effect into a universal principle: lifelike self-control was a simple engineering job. When the notion of feedback control was packaged with the flexibility of electronic circuits, they married into a tool anyone could use. Within a year or two of *Cybernetics'*s publication, electronic control circuits revolutionized industry.

The avalanche effects of employing automatic control in the production of goods were not all obvious. Down on the factory floor, automatic control had the expected virtue of moderating high-powered energy sources as mentioned earlier. There was also an overall speeding up of things because of the continuous nature of automatic control. But those were relatively minor compared to a completely unexpected miracle of self-control circuits: their ability to extract precision from grossness.

As an illustration of how the elemental loop generates precision of out imprecise parts, I follow the example suggested by the French writer Pierre de Latil in his 1956 book *Thinking by Machine.* Generations of technicians working in the steel industry pre-1948 had tried unsuccessfully to produce a roll of sheet metal in a uniform thickness. They discovered about a half-dozen factors that affected the thickness of the steel grinding out the rolling-mill—such as speed of the rollers, temperature of the steel, and traction on the sheet—and spent years strenuously perfecting the regulation of each of them, and more years attempting their synchronization. To no avail. The control of one factor would unintentionally disrupt the other factors. Slowing the speed would raise the temperature; lowering the temperature would raise the traction; increasing traction lowers the speed, and so on. Everything was influencing everything else. The control was wrapped up in some interdependent web. When the steel rolled out too thick or too thin, chasing down the culprit out of six interrelated suspects was inevitably a washout. There things stalled until Wiener's brilliant generalization published in *Cybernetics.* Engineers around the world immediately grasped the crucial idea and installed electronic feedback devices in their mills within the following year or two.

In implementation, a feeler gauge measures the thickness of the just-made sheet metal (the output) and sends this signal back to a servo-motor controlling the single variable of traction, the variable to affect the steel last, just before the rollers. By this meager, solo loop, the whole caboodle is regulated. Since all the factors are interrelated, if you can keep just one of them directly linked to the finished thickness, then *you can indirectly control them all.* Whether the deviation tendency comes from uneven raw metal, worn rollers, or mistakenly high temperatures doesn't matter much. What matters is that the automatic loop regulates that last variable to compensate for the other variables. If there is enough leeway (and there was) to vary the traction to make up for an overly thick source metal, or insufficiently tempered stock, or rollers contaminated with slag, then out would come consistently even sheets. Even though each factor is upsetting the others, the contiguous and near instantaneous nature of the loop steers the unfathomable network of relationships between them toward the steady goal of a steady thickness.

The cybernetic principle the engineers discovered is a general one: if all the variables are tightly coupled, and if you can truly manipulate one of them in all its freedoms, then you can indirectly control all of them. This principle plays on the holistic nature of systems. As Latil writes, "The regulator is unconcerned with causes; it will detect the deviation and correct it. The error may even arise from a factor whose influence has never been properly determined hitherto, or even from a factor whose very existence is unsuspected." How the system finds agreement at any one moment is beyond human knowing, and more importantly, not worth knowing.

The irony of this breakthrough, Latil claims, is that technologically this feedback loop was quite simple and "it could have been introduced some fifteen or twenty years earlier, if the problem had been approached with a more open mind..." Greater is the irony that twenty years earlier the open mind for this view was well established in economic circles. Frederick Hayek and the influential Austrian school of economics had dissected the attempts to trace out the routes of feedback in complex networks and called the effort futile. Their argument became known as the "calculation argument." In a command economy, such as the then embryonic top-down economy installed by Lenin in Russia, resources were allotted by calculation, tradeoffs, and controlled lines of communication. Calculating, even less controlling, the multiple feedback factors among distributed nodes in an economy was as unsuccessful as the engineer's failure in chasing down the fleeing interlinked factors in a steel mill. In a vacillating economy it is impossible to calculate resource allotment. Instead, Hayek and other Austrian economists of the 1920s argued that a single variable—the price—is used to regulate all the other variables of resource allotment. That way, one doesn't care how many bars of soap are needed per person, or whether trees should be cut for houses or for books. These calculations are done in parallel, on the fly, from the bottom up, out of human control, by the interconnected network itself. Spontaneous order.

The consequence of this automatic control (or human uncontrol) is that the engineers could relax their ceaseless straining for perfectly uniform raw materials, perfectly regulated processes. Now they could begin with imperfect materials, imprecise processes. Let the self-correcting nature of automation strain to find the optima which let only the premium through. Or, starting with the same quality of materials, the feedback loop could be set for a much higher quality setting, delivering increased precision for the next in line. The identical idea could be exported upstream to the suppliers of raw materials, who could likewise employ the automatic loop to extract higher quality products. Cascading further out in both directions in the manufacturing stream, the automatic self became an overnight quality machine, ever refining the precision humans can routinely squeeze from matter.

Radical transformations to the means of production had been introduced by Eli Whitney's interchangeable parts and Ford's idea of an assembly line. But these improvements demanded massive retooling and capital expenditures, and were not universally applicable. The homely auto-circuit, on the other hand—a suspiciously cheap accessory—could be implanted into almost any machine that already had a job. An ugly duckling, like a printing press, was transformed into a well-behaved goose laying golden eggs.

But not every automatic circuit yields the ironclad instantaneity that Bill Power's gun barrel enjoyed. Every unit added onto a string of connected loops increases the likelihood that the message traveling around the greater loop will arrive back at its origin to find that everything has substantially changed during its journey. In particularly vast networks in fast moving environments, the split second it takes to traverse the circuit is greater than the time it takes for the situation to change. In reaction, the last node tends to compensate by ordering a large correction. But this also is delayed by the long journey across many nodes, so that it arrives missing its moving mark, birthing yet another gratuitous correction. The same effect causes student drivers to zigzag down the road, as each late large correction of the steering wheel overreacts to the last late overcorrection. Until the student driver learns to tighten the feedback loop to smaller, quicker corrections, he cannot help but swerve down the highway hunting (in vain) for the center. This then is the bane of the simple auto-circuit. It is liable to "flutter" or "chatter," that is, to nervously oscillate from one overreaction to another, hunting for its rest. There are a thousand tricks to defeat this tendency of overcompensation, one trick each for the thousand advance circuits that have been invented. For the last 40 years, engineers with degrees in control theory have written shelffuls of treatises communicating their latest solution to the latest problem of oscillating feedback. Fortunately, feedback loops can be combined into useful configurations.

Let's take our toilet, that prototypical cybernetic example. We install a knob which allows us to adjust the water level of the tank. The self-regulating mechanism inside would then seek whatever level we set. Turn it down and it satisfies itself with a low level; turn it up and it hones in on a high level of water. (Modern toilets do have such a knob.) Now let's go further and add a self-regulating loop to turn the knob, so that we can let go of that, too. This second loop's job is to seek the goal for the first loop. Let's say the second mechanism senses the water pressure in the feed pipe and then moves the knob so that it assigns a high level to the toilet when there is high water pressure and a lower level when the pressure is low.

The second circuit is controlling the range of the first circuit which is controlling the water. In an abstract sense the second loop brings forth a second order of control—the control of control—or a metacontrol. Our newfangled second-order toilet now behaves "purposefully." It adapts to a *shifting* goal. Even though the second circuit setting the goal for the first is likewise mechanical, the fact that the whole is choosing its own goal gives the metacircuit a mildly biological flavor.

As simple as a feedback loop is, it can be stitched together in endless combinations and forever stacked up until it forms a tower of the most unimaginable complexity and intricacy of subgoals. These towers of loops never cease to amuse us because inevitably the messages circulating along them cross their own paths. *A* triggers *B,* and *B* triggers *C,* and *C* triggers *A.* In outright paradox, *A* is both cause and effect. Cybernetician Heinz von Foerster called this elusive cycle "circular causality." Warren McCulloch, an early artificial intelligence guru called it "intransitive preference," meaning that the rank of preferences would cross itself in the same self-referential way the children's game of Paper-Scissors-Stone endlessly intersects itself: Paper covers stone; stone breaks scissors; scissors cuts paper; and round again. Hackers know it as a recursive circuit. Whatever the riddle is called, it flies in the face of 3,000 years of logical philosophy. It undermines classical everything. If something can be both its own cause and effect, then rationality is up for grabs.

Self-causing agencies

THE COMPOUNDED LOGIC OF STACKED loops which doubles back on itself is the source of the strange counterintuitive behaviors of complex circuits. Made with care, circuits perform dependably and reasonably, and then suddenly, by their own drumbeat, they veer off without notice. Electrical engineers get paid well to outfox the lateral causality inherent in all circuits. But pumped up to the density required for a robot, circuit strangeness becomes indelible. Reduced back to its simplest—a feedback cycle—circular causality is a fertile paradox.

Where does self come from? The perplexing answer suggested by cybernetics is: it emerges from itself. It cannot appear any other way. Brian Goodwin, an evolutionary biologist, told reporter Roger Lewin, "The organism is the cause and effect of itself, its own intrinsic order and organization. Natural selection isn't the *cause* of organisms. Genes don't *cause* organisms. There are no *causes* of organisms. Organisms are self-causing agencies." Self, therefore, is an auto-conspired form. It emerges to transcend itself, just as a long snake swallowing its own tail becomes Uroborus, the mythical loop.

The Uroborus, according to C. G. Jung, is one of those resonant projections of the human soul that cluster around timeless forms. The ring of snake consuming its own tail first appeared as art adorning Egyptian statuary. Jung developed the idea that the nearly chaotic variety of dream images visited on humans tend to gravitate around certain stable nodes which form key and universal images, much as interlinked complex systems tend settle down upon "attractors," to use modern terminology. A constellation of these attracting, strange nodes form the visual vocabulary of art, literature, and some types of therapy. One of the most enduring attractors, and an early pattern to be named, was the

Thing Eating Its Own Tail, often graphically simplified to a snakelike dragon swallowing its own tail in a perfect circle.

The loop of Uroborus is so obviously an emblem for feedback that I have trouble ascertaining who first used it in a cybernetic context. In the true manner of archetypes it was probably realized as a feedback symbol independently more than once. I wouldn't doubt that the faint image of snake eating its tail spontaneously hatches whenever, and wherever, the GOTO START loop dawns on a programmer.

Snake is linear, but when it feeds back into itself it becomes the archetype of non-linear being. In the classical Jungian framework, the tail-biting Uroborus is the symbolic depiction of the self. The completeness of the circle is the self-containment of self, a containment that is at the same time made of one thing and made of competing parts. The flush toilet then, as the plainest manifestation of a feedback loop, is a mythical beast—the beast of self.

The Jungians say that the self is taken to be "the original psychic state prior to the birth of ego consciousness," that is, "the original mandala-state of totality out of which the individual ego is born." To say that a furnace with a thermostat has a self is not to say it has an ego. The self is a mere ground state, an auto-conspired form, out of which the more complicated ego can later distinguish itself, should its complexity allow that.

Every self is a tautology: self-evident, self-referential, self-centered, and self-created. Gregory Bateson said a vivisystem was "a slowly self-healing tautology." He meant that if disturbed or disrupted, a self will "tend to settle toward tautology"—it will gravitate to its elemental self-referential state, its "necessary paradox."

Every self is an argument trying to prove its identity. The self of a thermostat system has endless internal bickering about whether to turn the furnace up or down. Heron's valve system argues continuously around the sole, solitary action it can take: should it move the float or not?

A system is anything that talks to itself. All living systems and organisms ultimately reduce to a bunch of regulators—chemical pathways and neuron circuits—having conversations as dumb as "I want, I want, I want; no, you can't, you can't, you can't."

The sowing of selves into our built world has provided a home for control mechanisms to trickle, pool, spill, and gush. The advent of automatic control has come in three stages and has spawned three nearly metaphysical changes in human culture. Each regime of control is boosted by deepening loops of feedback and information flow.

The control of energy launched by the steam engine was the first stage. Once energy was controlled it became "free." No matter how much more energy we might release, it won't fundamentally change our lives. The amount of calories (energy) require to accomplish something continues to dwindle so that our biggest technological gains no longer hinge on further mastery of powerful energy sources.

Instead, our gains now derive from amplifying the accurate control of materials—the second regime of control. Informing matter by investing it with high degrees of feedback mechanisms, as is done with computer chips, empowers the matter so that increasingly smaller amounts do the same work of larger uninformed amounts. With the advent of motors the size of dust motes (successfully prototyped in 1991), it seems as if you can have anything you want made in any size you want. Cameras the size of molecules? Sure, why not? Crystals the size of buildings? As you wish. Material is under the thumb of information, in the same handy way that energy now is—just spin a dial. "The central event of the twentieth century is the overthrow of matter," says technology analyst George Gilder. This is the stage in the history of control in which we now dwell. Essentially, matter—in whatever shape we want—is no longer a barrier. Matter is almost "free."

George Gilder tests the newest communication technology.

The third regime of the control revolution, seeded two centuries ago by the application of information to coal steam, is the control of information itself. The miles of circuits and information looping from place to place that administers the control of energy and matter has incidentally flooded our environment with messages, bits, and bytes. This unmanaged data tide is at toxic levels. We generate more information than we can control. The promise of more information has come true. But more information is like the raw explosion of steam—utterly useless unless harnessed by a self. To paraphrase Gilder's aphorism: "The central event of the twenty-first century will be the overthrow of information."

Genetic engineering (information which controls DNA information) and tools for electronic libraries (information which manages book information) foreshadow the subjugation of information. The impact of information domestication will be felt initially in industry and business, just as energy and material control did, and then later seep to the realm of individual.

The control of energy conquered the forces of nature (and made us fat); the control of matter brought material wealth within easy reach (and made us greedy). What mixed cornucopia will the blossoming of full information control bring about? Confusion, brilliance, impatience?

Without selves, very little happens. Motors, by the millions, bestowed with selves, now run factories. Silicon chips, by the billions, bestowed with selves, will redesign themselves smaller and faster and rule the motors. And soon, the fibrous networks, by the zillions, bestowed with selves, will rethink the chips and rule all that we let them. If we had tried to exploit the treasures of energy, material, and information by holding all the control, it would have been a loss.

As fast as our lives allow us, we are equipping our constructed world to bootstrap itself into self-governance, self-reproduction, self-consciousness, and irrevocable selfhood. The story of automation is the story of a *one-way* shift from human control to automatic control. The gift is an irreversible transfer from ourselves to the second selves.

The second selves are out of our control. This is the key reason, I believe, why the brightest minds of the Renaissance never invented another self-regulator beyond the obvious ones known to ancient Heron. The great Leonardo da Vinci built control machines, not out-of-control machines. German historian of technology Otto Mayr claims that great engineers in the Enlightenment could have built regulated steam power of some sort with the technology available to them at the time. But they didn't because they didn't have the ability to let go of their creation.

The ancient Chinese on the other hand, although they never got beyond the south-pointing cart, had the right no-mind about control. Listen to these most modern words from the hand of the mystical pundit Lao Tzu, writing in the *Tao Teh King* 2,600 years ago:

> Intelligent control appears as uncontrol or freedom.
> And for that reason it is genuinely intelligent control.
> Unintelligent control appears as external domination.
> And for that reason it is really unintelligent control.
> Intelligent control exerts influence without appearing to do so.
> Unintelligent control tries to influence by making a show of force.

Lao Tzu's wisdom could be a motto for a gung-ho 21st-century Silicon Valley startup. In an age of smartness and superintelligence, the most intelligent control methods will appear as uncontrol methods. Investing machines with the ability to adapt on their own, to evolve in their own direction, and grow without human oversight is the next great advance in technology. Giving machines freedom is the only way we can have intelligent control.

What little time left in this century is rehearsal time for the chief psychological chore of the 21st century: letting go, with dignity.

II

With these preliminary remarks, let me turn to the real theme of this little book.

There are at least three points in cybernetics which appear to me to be relevant to religious issues. One of these concerns machines which learn; one concerns machines which reproduce themselves; and one, the coordination of machine and man. I may say that such machines are known to exist. A program has been written by Dr. A. L. Samuel of the International Business Machines Corporation which allows a computer to play a game of checkers, and this computer learns, or at least appears to learn, to improve its game by its own experience.* There are certain statements

* Samuel, A. L., "Some Studies in Machine Learning, Using the Game of Checkers," *IBM Journal of Research and Development*, Vol. 3, 210–229 (July, 1959).

here which need confirmation, or at least clarification; and I shall devote one section of this book to this clarification.

Learning is a property that we often attribute exclusively to self-conscious systems, and almost always to living systems. It is a phenomenon that occurs in its most characteristic form in Man, and constitutes one of those attributes of Man which is most easily put in conjunction with those aspects of Man which are easily associated with his religious life. Indeed, it is hard to see how any non-learning being can be concerned with religion.

There is, however, another aspect of life which is naturally associated with religion. God is supposed to have made man in His own image, and the propagation of the race may also be interpreted as a function in which one living being makes another in its own image. In our desire to glorify God with respect to man and Man with respect to matter, it is thus natural to assume that machines cannot make other machines in their own image; that this is something associated with a sharp dichotomy of systems into living and non-living; and that it is moreover associated with the other dichotomy between creator and creature.

Is this, however, so? I shall devote a section of this book to certain considerations which, in my opinion, show that machines are very well able to make other machines in their own image. The subject upon which I am entering here is at once very technical and very precise. It should not be taken too seriously as an actual model of the process of biological generation, and even less as a complete model of divine creation; but neither is it negligible as to the light it throws upon both concepts.

These two parts of this book of lectures may be regarded as complementary the one to the other. The learning of the individual is a process that occurs in the life of the individual, in *ontogeny.* Biological reproduction is a phenomenon that occurs in the life of the race, in *phylogeny,* but the race learns even as the individual does. Darwinian natural selection is a kind of racial learning, which operates within the conditions imposed by the reproduction of the individual.

The third group of topics of this book is also related to problems of learning. It is concerned with the relations of the machine to the living being, and with systems involving elements of both kinds. As such, it involves considerations of a normative and, more specifically, of an ethical nature. It concerns some of the most important moral traps into which the present generation of human beings is likely to fall. It is also closely connected with a great body of human tradition and human legend, concerning magic and the like.

To begin with learning machines: an organized system may be said to be one which transforms a certain incoming message into an outgoing message, according to some principle of transformation. If this principle of transformation is subject to a certain criterion of merit of performance, and if the method of transformation is adjusted so as to tend to improve the performance of the system according to this criterion, the system is said to *learn*. A very simple type of system with an easily interpreted criterion of performance is a game, to be played according to fixed rules, where the criterion of performance is the successful winning of the game according to these rules.

Among such games are games with a perfect theory, which are not interesting. Nim, as defined by Bouton, and ticktacktoe are examples of such games. In these games, we not only can theoretically find a best policy for the playing of the game, but this policy is known in all its details. The player of such a game (either the first or the second) can always win, or at any rate draw, by following the policy indicated. In theory, any game can be brought to such a state—this is the idea of the late John von Neumann—but once a game has been brought to this state, it loses all interest, and need no longer be considered even as an amusement.

An omniscient being such as God would find chess and checkers (or draughts in England, or *dames* on the continent) to be examples of such von Neumann games, but as yet their complete theory has not been humanly worked out, and they still represent genuine contests of insight and intelligence. However, they are not played according to the manner suggested in the von Neumann theory. That is, we do not play them by making the best possible move, on the assumption that an opponent will make the best possible move, on the assumption that we shall make the best possible move, and so on, until one of the players wins or the game repeats itself. Indeed, to be able to play a game in the von Neumann manner is tantamount to possessing a complete theory of the game and to having reduced the game to a triviality.

The subject of learning, and in particular of machines that learn to play games, may seem somewhat remote from religion. Nevertheless, there is a religious problem to which those notions are relevant. This is the problem of the game between the Creator and a creature. This is the theme of the Book of Job, and of *Paradise Lost* as well.

In both these religious works the Devil is conceived as playing a game with God, for the soul of Job, or the souls of mankind in general. Now, according to orthodox Jewish and Christian views, the Devil is one of God's creatures. Any other supposition would lead to a moral dualism, savoring of Zoroastrianism and of that bastard offshoot of Zoroastrianism and Christianity which we call Manicheanism.

But if the Devil is one of God's creatures, the game that furnishes the content of the Book of Job and of *Paradise Lost* is a game between God and one of his creatures. Such a game seems at first sight a pitifully unequal contest. To play a game with an omnipotent, omniscient God is the act of a fool; and, as we are told, the Devil is a master of subtlety. Any uprising of the rebel angels is foredoomed to failure. It is not worth the Manfred-like rebellion of Satan to prove this point. Or else that omnipotence which needs to establish itself by celestial bombardments of thunderbolts is no omnipotence at all but merely a very great strength, and the Battle of the Angels might have ended with Satan on the celestial throne, and God cast down into eternal damnation.

Thus, if we do not lose ourselves in the dogmas of omnipotence and omniscience, the conflict between God and the Devil is a real conflict, and God is something less than absolutely omnipotent. He is actually engaged in a conflict with his creature, in which he may very well lose the game. And yet his creature is made by him according to his own free will, and would seem to derive all its possibility of action from God himself. Can God play a significant game with his own creature? Can *any* creator, even a limited one, play a significant game with his own creature?

In constructing machines with which he plays games, the inventor has arrogated to himself the function of a limited creator, whatever the nature of the game-playing device that he has constructed. This is in particular true in the case of game-playing machines that learn by experience. As I have already mentioned, such machines exist. How do these machines function? What degree of success have they had?

Instead of functioning after the pattern of the von Neumann game theory, they act in a manner much more closely analogous to the proceeding of the ordinary human game player. At each stage, they are subject to constraints that restrict the choice of the next move to one which is legal according to the rules of the game. One of these moves must be selected according to some normative criterion of good play.

Here, the experience of the human player of the game furnishes a number of clues to be used in picking out this criterion. In checkers or chess it is generally disadvantageous to lose pieces and generally advantageous to take an opponent's piece. The player who retains his mobility and right of choice, as well as the player who secures the command of a large number of squares, is usually better off than his opponent who has been careless in these respects.

These criteria of good play hold throughout the game, but there are other criteria that belong to a particular stage of the game. At the end of the game, when the pieces are sparse on the board, it becomes more difficult to close with the opponent for the kill. At the beginning of the game —and this is a far more important factor in chess than in checkers—the pieces are arranged in a way that tends to make them immobile and impotent, and a development is needed that will get them out of one another's way, both for offensive and defensive purposes. Furthermore, with the great variety of pieces in chess as compared with the poverty of checkers in this regard, there are in chess a large number of special criteria of good play, the importance of which has been proved by centuries of experience.

These considerations may be combined (either additively or in some more complicated way) to give a figure of merit for the next move to be played by the machine. This may be done in a somewhat arbitrary manner. Then the machine compares the figures of merit of the moves legally possible and chooses that move with the largest figure of merit. This gives one way of automatizing the next move.

This automatization of the next move is not necessarily, or even usually, an optimum choice, but it is a choice, and the machine can go on playing. To judge the merit of this way of mechanizing a game, one should divest oneself of all the images of mechanization belonging to the technical devices used, or the physical image of humanity as displayed by the ordinary game player. Luckily this is easy, for it is what we always do in correspondence chess.

In correspondence chess, one player sends the moves by mail to the other, so that the only connection between the two players is a written document. Even in this sort of chess, a skilled player soon develops an image of the personality of his opponent—of his *chess* personality, that is. He will learn if his opponent is hasty or careful; if he is easily tricked or subtle; if he learns the tricks of the other player, or can be fooled again and again by the same elementary strategy. All this is done, I repeat, with no further communication than the playing of the game itself.

From this point of view, the player, be he a man or a machine, who plays by a simple table of merit, chosen once for all and unalterable, will give the impression of a rigid chess personality. Once you have found out his weak point, you have found it out for all time. If a strategem has worked once against him, it will always work. A very small number of plays are enough to establish his technique.

So much for the mechanized player who does not learn. However, there is nothing to prevent a mechanized player from playing in a more intelligent way. For this he must keep a record of past games and past plays. Then, at the end of each game or each sequence of games of a determined sort, his mechanism is put to a totally different sort of use.

In building up the figure of merit, certain constants are introduced which might have been chosen differently. The relative importance of the command constant, the mobility constant, and the number-of-pieces constant might have been 10:3:2, instead of 9:4:4. The new use of the regulating machine is to examine games already played and, in view of the outcome of these, to give a figure of merit; not to the plays already made, but to the weighting chosen for the evaluation of these plays.

In this way, the figure of merit is continually being re-evaluated, in such a manner as to give a higher figure of merit for configurations occurring chiefly in winning games, and a lower figure of merit for situations occurring chiefly in losing games. The play will continue with this new figure of merit, which may be established in many ways differing in detail. The result will be that the game-playing machine will continually transform itself into a different machine, in accordance with the history of the actual play. In this, the experience and success, both of the machine and of its human opponent, will play a role.

In playing against such a machine, which absorbs part of its playing personality from its opponent, this playing personality will not be absolutely rigid. The opponent may find that strategems which have worked in the past, will fail to work in the future. The machine may develop an uncanny canniness.

It may be said that all this unexpected intelligence of the machine has been built into it by its designer and programmer. This is true in one sense, but it need not be true that all of the new habits of the machine have been explicitly foreseen by him. If this were the case, he should have no difficulty in defeating his own creation. This is not in accordance with the actual history of Samuel's machine.

As a matter of fact, for a considerable period Samuel's machine was able to defeat him rather consistently, after a day or so of working in. It must be said that Samuel, by his own statement, was no checker expert to begin with, and that with a little further instruction and practice he was able to win over his own creation. It will not do, however, to belittle the fact that there was a period when the machine was rather a consistent victor. It did win, and it did learn to win; and the method of its learning was no different in principle from that of the human being who learns to play checkers.

It is true that the choice of policies open to the checker-playing machine is almost certainly narrower than that open to the human checker player; but it is also true that the choice of policies effectively open to the human checker player is not unlimited. He may be restrained from a wider choice only by the bounds of his intelligence and imagination, but those are very real bounds indeed and not of a sort essentially different from the bounds of the machine.

Thus the checker-playing machine already plays a reasonably good game, which with a little further study of the end game and a little more skill in applying the *coup de grâce* may begin to approach master level. If it were not for the fact that the interest in checker championship has already been greatly diminished by the cut-and-dried nature of normal human play, the checker-playing machine might already be said to have destroyed the interest in checkers as a game. It is not surprising that people are already beginning to ask, will chess go the same way? And, when is this castastrophe to be expected?

Chess-playing machines, or machines to play at least an appreciable part of a chess game, are already in existence, but they are comparatively poor things. They do not, at their best, go beyond the level of a competent game between amateurs with no pretense to chess mastership, and they very seldom reach that level. This is largely due to the far greater complexity of chess than of checkers, both as to pieces and moves, and as to the greater discrimination between the policies suitable for the different stages of the game. The relatively small number of considerations necessary for mechanizing a checker game and the low degree of discrimination needed between its different stages are totally inadequate for chess.

Nevertheless, I find it to be the general opinion of those of my friends who are reasonably proficient chess players that the days of chess as an interesting human occupation are numbered. They expect that within from ten to twenty-five years, chess machines will have reached the master class, and then, if the efficient but somewhat machinelike methods of the Russian school have allowed chess to survive so long, it will cease to interest human players.

Be this as it may, there will be many other games that will continue to offer a challenge to the games engineer. Among these is Go, that Far Eastern game in which there are seven or more different levels of recognized mastery. Moreover, war and business are conflicts resembling games, and as such, they may be so formalized as to constitute games with definite rules. Indeed, I have no reason to suppose that such formalized versions of them are not already being established as models to determine the policies for pressing the Great Push Button and burning the earth clean for a new and less humanly undependable order of things.

In general, a game-playing machine may be used to secure the automatic performance of any function *if the performance of this function is subject to a clear-cut, objective criterion of merit.* In checkers and chess, this merit consists of the winning of the game according to the accepted rules of permissible play. These rules, which are totally different from the accepted maxims of *good* play, are simple and inexorable. Not even an intelligent child can be in doubt concerning them for longer than it takes to read them while facing a board. There may be great doubt as to how to win a game, but no doubt whatever as to whether it has been won or lost.

The chief criterion as to whether a line of human effort can be embodied in a game is whether there is some objectively recognizable criterion of the merit of the performance of this effort. Otherwise the game assumes the formlessness of the croquet game in *Alice in Wonderland,* where the balls were hedgehogs and kept unrolling themselves, the mallets were flamingoes, the arches cardboard soldiers who kept marching about the field, and the umpire the Queen of Hearts, who kept changing the rules and sending the players to the Headsman to be beheaded. Under these circumstances, to win has no meaning, and a successful policy cannot be learned, because there is no criterion of success.

However, given an objective criterion of success, the learning game may certainly be played, and is much closer to the way in which we learn to play games than the image of a game given in the von Neumann theory. Unquestionably the technique of the learning game is certain to be employed in many fields of human effort which have not yet been subjected to it. Nevertheless, as we shall see later, the determination of a sharp test for good performance raises many problems concerning learning games.

an interview with hans haacke

jeanne siegel

(. . .)

Jeanne Siegel: You have been called a naturalist because of your extensive interest in physical elements as well as grass, birds, ants, and animals.

Hans Haacke: I don't consider myself a naturalist, nor for that matter a conceptualist or a kineticist, an earth artist, elementalist, minimalist, a marriage broker for art and technology, or the proud carrier of any other button that has been offered over the years. I closed my little statement of 1965 with "articulate something Natural." That has an intended double meaning. It refers to "nature," and it means something self-understood, ordinary, uncontrived, normal, something of an everyday quality. When people see the wind stuff or the things I have done with animals, they call me a "naturalist." Then they get confused or feel cheated when they discover, for example, my interest in using a computer to conduct a demographic survey. This is inconsistent only for those with a naive understanding of nature—nature being the blue sky, the Rockies, Smokey the Bear. The difference between "nature" and "technology" is only that the latter is man-made. The functioning of either one can be described by the same conceptual models, and they both obviously follow the same rules of operation. It also seems that the way social organizations behave is not much different. The world does not break up into neat university departments. It is one supersystem with a myriad of subsystems, each one more or less affected by all the others.

If you take a grand view, you can divide the world into three or four categories—the physical, biological, the social and behavioral—each of them having interrelations with the others at one point or another. There is no hierarchy. All of them are important for the upkeep of the total system. It could be that there are times when one of these categories interests you more than another. So, for example, I now spend more thought on things in the social field, but simultaneously I am preparing a large water-cycle for the Guggenheim show that uses the peculiarities of the building.

J.S.: When did you first become aware of systems theory?

H.H.: Sometime in '65 or '66 I was introduced to the concept of systems. I heard about systems analysis, and the related fields of operational research, cybernetics, etc. The concepts used in these fields seemed to apply to what I had been doing and there was a useful terminology that seemed to describe it much more succinctly than the terminology that I and other people had been using until then, so I adopted it. But using a new terminology doesn't mean that the work described has changed. A new term is nothing holy, so it can't serve as a union label. On the other hand, a clear terminology can help to stimulate thinking.

J.S.: Jack Burnham has had a lot to say about systems and sculpture, yours in particular. When did you first meet him?

H.H.: I met Jack in 1962 when we were both isolated from people interested in what we were doing. Since then we have been in contact and have had a very fruitful exchange of ideas. It was Jack who introduced me to systems analysis.

J.S.: What is your definition of a system that is also a work of art?

H.H.: A system is most generally defined as a grouping of elements subject to a common plan and purpose. These elements or components interact so as to arrive at a joint goal. To separate the elements would be to destroy the system. The term was originally used in the natural sciences for understanding the behavior of physically interdependent processes. It explained phenomena of directional change, recycling, and equilibrium. I believe the term system should be reserved for sculptures in which a transfer of energy, material, or information occurs, and which do not depend on perceptual interpretation. I use the word "systems" exclusively for things that are not systems in terms of perception, but are physical, biological, or social entities which, I believe, are more real than perceptual titillation. (. . .)

A very important difference between the work of Minimal sculptors and my work is that they were interested in inertness, whereas I was concerned with change. From the beginning the concept of change has been the ideological basis of my work. All the way down there's absolutely nothing static—nothing that does not change, or instigate real change. Most Minimal work disregards change. Things claim to be inert, static, immovably beyond time. But the status quo is an illusion, a dangerous illusion politically. (. . .)

J.S.: Is there any difference in communication between social systems and physical or biological ones?

H.H.: For physical or biological processes to take their course, there is no need for the presence of a viewer—unless, as with some participatory works, his physical energy is required (he then becomes an indispensable part of the system's physical environment). However, there is no need for *anybody* to get mentally involved. These systems function on their own, since their operation does not take place in the viewer's mind (naturally this does not prevent a mental or emotional response).

The rigging of a social situation, however, usually follows a different pattern. There the process takes place exclusively in the minds of people. Without participants there is no social set. Take the "MOMA Poll" in last year's "Information" show: the work was based on a particular political situation circumscribed by the Indochina War, Nixon's and Rockefeller's involvement in it. MOMA's close ties to both, my own little quarrels with the museum as part of the Art Workers Coalition's activities, and then all the minds of the people who had a stake in this game—the Vietcong as much as the Scarsdale lady on her culture tour to the city. The result of the poll—approximately 2 to 1 against Rockefeller/Nixon and the war—is only the top of the iceberg. The figures are not quite reliable because MOMA, as usual, did not follow instructions, and the polls have to be taken with a grain of salt.

Emily Genauer gave us a little glimpse of the large base of the work in her review of the show. She wrote: "One may wonder at the humor (propriety, obviously, is too archaic a concept even to consider) of such poll-taking in a museum founded by the governor's mother, headed now by his brother, and served by himself and other members of his family in important financial and administrative capacities since its founding 40 years ago." With this little paragraph she provided some of the background for the work that was not intelligible for the politically less-informed visitors of the museum. She also articulated feelings that are shared by the top people at numerous museums. It goes like this: We are the guardians of culture. We honor artists by inviting them to show in *our* museum, we want them to behave like guests; proper, polite, and grateful. After all, we have put up the dough for this place.

The energy of information interests me a lot. Information presented at the right time and in the right place can be potentially very powerful. It can affect the general social fabric.

Such things go beyond established high culture as it has been perpetrated by a taste-directed art industry. Of course I don't believe that artists really wield any significant power. At best, one can focus attention. But every little bit helps. In concert with other people's activities outside the art scene, maybe the social climate of society can be changed. Anyway, when you work with the "real stuff" you have to think about potential consequences. A lot of things would never enter the decision-making process if one worked with symbolic representations that have to be weighed carefully. If you work with real-time systems, well, you probably go beyond Duchamp's position. Real-time systems are double agents. They might run under the heading "art," but this culturization does not prevent them from operating as normal. The MOMA Poll had even more energy in the museum than it would have had in the street—real sociopolitical energy, not awe-inspiring symbolism.

J.S.: Can you describe a social work that is not political?

H.H.: Probably all things dealing with social situations are to a greater or lesser degree political. Take *The Gallery-Goer's Residence Profile*. I asked the people that came to my exhibition to mark with a blue pin on large maps where they were living. After the show I traveled to all those spots on the Manhattan map that were marked by a blue pin and took a photograph of the building or approximately that location. I came up with about 730 photographs for Manhattan (naturally not every visitor participated in the game). The photographs were enlarged to 5″ by 7″. They will be displayed on the wall of the Guggenheim according to a geographical score. All those spots that were east of Fifth Avenue go upward on the wall from a horizontal center line, those west go downward. The respective distance from Fifth Avenue determines the sequence of pictures East and West. The Fifth Avenue spine takes up approximately 36 yards of wall space. Sometimes the photographs reach up to the ceiling, on other occasions (e.g., there is only one on the west side and none on the east side) it becomes a very jagged distribution. The "composition" is a composition determined by the information provided by the gallery-goers. No visual considerations play a role.

All this sounds very innocent and apolitical. The information I collected, however, is sociologically quite revealing. The public of commercial art galleries, and probably that of museums, lives in easily identifiable and restricted areas. The main concentrations are on the upper West Side (Central Park and adjoining blocks, and West End Avenue with adjoining blocks), the Upper East Side, somewhat heavier in the Madison–Park Avenue areas, then below 23rd Street on the East and West sides with clusters on the Lower East Side and the loft district. The photographs give an idea of the economic and social fabric of the immediate neighborhood of the gallery-goers. Naturally the Lower East Side pins were not put there by Puerto

Ricans. Puerto Ricans and blacks (Harlem is practically not represented) do not take part in an art scene that is obviously dominated by the middle- and upper-income strata of society or their drop-out children. I leave it up to you as far as how you evaluate this situation. You continue the work by drawing your own conclusions from the information presented.

This interview, conducted in early 1971, first appeared in *Arts Magazine,* 45:7 (May 1971), pp. 18–21.

What are Generative Systems?

An image is projected on a screen: it is abstract and intricately, impossibly detailed; luminous and smooth, its filigreed structures recede toward computer-graphic horizons. On another screen, a moving image is projected; a dense grove of synthetic foliage conceals a group of what appear to be insects. The viewer's image, mirrored within this artificial environment, causes the insects to recoil and retreat. On yet another screen, an animation shows an intricate, geometric flower unfurling, extending snaky tendrils into a digital void; it spins and writhes, filling the frame. In a gallery space a group of intricate white sculptures stand on a table, forms made up of masses of tiny cubes, three-dimensional pixels. On a nearby computer monitor, similar forms appear in an ever-changing series. An artists' statement describes how these forms arise as a cubic volume

1 INTRODUCTION

differentiates itself, splitting like a living cell but at ever finer scales. Elsewhere a bicycle-wheeled robot rolls around a room in a nervous interactive dance with the people gathered there: it advances and retreats, spindly body rocking back and forth. Another room is filled with loud, skeletal machines that shriek and flail, seemingly attacking each other, menacing passersby with blinding lights and horrendous noises. In yet another, quieter space, three arms made from grape vines and wire twist and pivot from the ceiling, "singing" to each other in telephone touch-tones.

These are strange objects, embodying a series of contradictions and ambiguities. They are technological objects, this much is clear; made from the glowing points of light on computer screens, or from metal, motors, and electronics. Yet unlike the technological objects we routinely encounter, they are unpredictable and apparently autonomous; something in their movement, their reactions, their structure, reminds us — is clearly intended to remind us — of living things. They are art objects; we find them in galleries, at symposia and conferences, among other art objects made of computer code and electronics. Yet some of them are hardly objects at all; they refuse to sit still and be observed, but hide from us, play with us, or invite us into their own virtual worlds. Clearly, these things are made — they are the works of artists or others working artfully — but the signs of the will of a creator are sometimes less palpable in these objects than the manifestation of a "will" of their own. And while these works can be found in galleries and festivals, under the banner of art, they might also appear with their creators at conferences of another stripe, alongside elaborate computer simulations of cellular biology and crawling, multilegged robots, the technological objects of the science of artificial life.

Artificial life, or a-life, is a young, interdisciplinary scientific field concerned with the creation and study of artificial systems that mimic or manifest the properties of living systems. It is a strange object in itself; its Promethean project to create new forms of life arouses scepticism, fascination, and alarm in equal measures. Having turned (in part) away from the task of analyzing nature and toward its synthesis, a-life seems unlike a science in the conventional sense. However, the objects just described are, if anything, stranger still. While they apply the techniques and ideas of artificial life in a variety of ways, they present themselves as art objects rather than as scientific artifacts. They are manifestations of a kind of transdisciplinary dissemination of artificial life, the results of its recent propagation through cyberculture and popular science writing. They arise where artificial life meets contemporary practice in the new media arts.

This meeting point provides the location for this project, and these strange objects, their makers, and the thought and writing around them are the objects of its attention. This complex is interesting and significant for a number of reasons. Most immediately, the artwork itself is striking. It evolves, responds, mutates, and forms complex, supple systems and cryptic alien artifacts. It offers engaging experiences, interactions with complex looking-glass worlds and embodied agencies, encounters with weird aesthetic objects. Beyond the immediate experience, these works tend to become more difficult, though no less interesting. The elegance of the engineering — the beguiling way these systems operate — counterpoints a sense of suspicion at the lines of metaphor and association they draw. How is this lifelike exactly? At times, this practice entails an expansive, godlike creative sweep, bringing whole worlds into being, populating them with virtual creatures. Ingenious, certainly, but is this also an extreme form of artistic hubris? At the same time, conventions of creative agency are stretched to breaking point: much of the work is made in such a way that it makes itself — it is somehow autonomous. Is this an abdication of creative will or its ultimate fulfillment?

More generally, this work is important in that the mapping around which it pivots, between living things and technological systems, is provocative, problematic, and highly current. Western culture is in the midst of an explosive development in the technologies of life and the living: the modeling, simulation, decomposition, engineering, and manipulation of biological life. Contemporary culture is slowly coming to grips with radical changes in its notions of life as medical and biological technologies reveal living matter as increasingly plastic and susceptible to engineering. Stem cells — unspecialized protocells, a kind of basic living material — are isolated and cultured. Reproduction, conventionally a unique and definitive capacity of the living organism, is ever more readily engineered and decomposed. Large mammals are cloned, and some strive to clone human beings, raising ethical rather than technological questions. Proprietary life forms, genetically manipulated species, are grown as food crops. Distinctions between natural and artificial, born and made, become unsupportable. Technoscience seems to have an ever-increasing command of living matter, and in an era of global capital, life is reshaped according to logics that are principally commercial.

However, at the same time, life itself continues to escape and evade technological control; it remains active, retains an agency. In fact, technological interventions seem to create unforeseen opportunities for living things. During the late 1990s, outbreaks of bovine Creutzfeldt-Jakob disease ("mad cow disease") demonstrated how even the most primitive organism can work through, and ultimately against, a technological network. The gruesome efficiency of industrial agriculture forms a feedback cycle; the protein-folding prion, barely even alive, crosses species barriers and infects the human consumers at the top of the food chain; entire industry sectors close down. Elsewhere, genetically modified crops escape their fenced-off testing grounds to compete and possibly interbreed with the surrounding biota. Antibiotic-resistant bacteria thrive in hospital wards.

The tangled counterpoints of biology and technology are ubiquitous. Computer viruses proliferate, with an impact as real as the biological variety — or in a technocentric business culture, more so. The industrial networks that host the technologization of life are simultaneously engaged in the decimation of biodiversity and the destruction of habitat. In April 2002 a Japanese institute launched the Earth Simulator, currently the most powerful supercomputer in existence; it runs massively detailed climatic and seismic simulations. If successful, it will improve weather prediction and warn of impending earthquakes; in the words of one of the scientists involved, the project aims to "keep a good relationship between nature and mankind, a symbiotic relationship. . . ."[1] A large-scale digital simulation seems to act here as a benign intermediary; yet at the same time the accelerated currents of digital media tend to pull us away from the difficult, polluted, outside space of the physical world and toward the clean, controllable "inside" of mediated experience and synthetic immersion. Life and technology, biology and information, hang in a tense articulation.

Against this background, a cultural practice that is engaged with both technological culture and biological science is in a particularly interesting position. New media art, the primary context for this practice, is already deeply enmeshed in a wider technoculture; its standard practice is to take up the products of the technology industries — focused recently on personal and networked computing — and apply them to its own diverse ends, in a cultural domain. It draws on these technical resources but also characteristically reflects on and critiques them. New media art self-consciously reworks technology into culture, and rereads technology as culture. What's more, it does so in a concrete, applied way; it manipulates the technology itself, with a nonindustrial latitude that admits misapplication and adaptation, rewiring and hacking, pseudofunctionality and accident. New media art also fractures that technocultural material into millions of heterogeneous interests and agendas, specific investigations, aesthetics, approaches, and projects.

When this practice begins adopting and adapting the technoscience of artificial life, it comes to grips with a troublesome constellation. A-life crystallizes the conjunction of biological life and technology into a handful of bold claims and images. The computer in this context seems to contain not only organisms but whole living systems in detailed articulation. Evolution, an idea that has become the most powerful organizing narrative of contemporary culture, appears to unfold on a screen. A-life proposes not a slavish imitation of this or that living thing but, at its strongest, an abstract distillation of aliveness, life itself, reembodied in voltage and silicon.

In appropriating (and altering) artificial life, the artists considered in this book are engaged in a crucial task: that of working through the implications of its concepts and techniques, testing their potential, deforming and transforming them. These operations are only partly technical; they are primarily and most importantly cultural. New media art provides a venue for the transformation and translation of the technical and conceptual artifacts of artificial life into cultural objects — conglomerates of rhetoric, metaphor, and aesthetics. Such translations are important in general because of the terms they articulate; at a time of rapid and dramatic technological change, the process of assimilating, debating, contesting, and reflecting on that change within cultural domains is crucial. The interface of artificial life and cultural practice is particularly significant for all these reasons; it opens a space for creative experimentation and debate around the increasing technologization of living matter as well as broader issues of life and autonomy, agency and evolution, genetics, code and matter. This work explores a practice in which we are all increasingly required to participate: the art of technologized life.

Artificial Life

Artificial life is a field of scientific research devoted to the simulation and synthesis of living things. It was founded in 1987 with a workshop at the Los Alamos National Laboratory, New Mexico. In subsequent years, interest in the field has grown: the artificial life workshops have become an ongoing international series, and the field has spawned dozens of other conferences; 1993 saw the publication of a journal dedicated to its work. The handful of scientists involved in the initial workshop has grown into a small international community.

Of course, efforts toward the simulation or synthesis of life are far older than this field. What distinguishes a-life from earlier work, and what unifies it currently, is a specific approach to this task. A-life begins with a notion of life that is wholly materialistic, involving no soul, vital force, or essence. In the words of the convenor of the first artificial life workshop, Christopher Langton, "Living organisms are nothing more than complex biochemical machines."[2] Langton contends that rather than being any special substance or force, life is "a property of the organization of matter."[3] Further, this organization is not simply a complex structure but a dynamic structure, a system active in time: for a-life, life is most importantly manifest in behavior. If, then, the "universal features" of life are in its abstract dynamic processes rather than inherent to a biological medium, we can consider the creation of such structures in another, artificial medium. Artificial life sets about creating such dynamic structures, almost always involving the most flexible, dynamic, and tightly controllable artificial medium at its disposal, computation.

It is this sense of living things as complex dynamic systems that informs the methodologies of artificial life. A-life's focus on the synthesis of such systems leads it to adopt the "bottom-up" approach that is one of the field's tenets. Influenced by theories of complex systems, a-life regards the complex dynamics of living things across all scales as phenomena that arise from the interaction of multitudes of smaller elements. Langton asserts that "natural life emerges out of the organized interactions of a great number of nonliving molecules, with no global controller responsible for the behavior of every part." Similarly,

> Artificial life starts at the bottom, viewing an organism as a large population of simple machines, and works upwards synthetically from there — constructing large aggregates of simple, rule-governed objects which interact with one another nonlinearly in the support of life-like, global dynamics.[4]

The process, known as emergence, by which these simple components interact to produce complex, lifelike results is another central concept in artificial life. Just as artificial life proposes that the complex behaviors of a living thing emerge from its nonliving parts, it seeks to recreate this process in artificial systems, so that an ensemble of simple computational parts interacts to spontaneously produce lifelike dynamic structures.

A useful way to briefly provide a sense of a-life's approach and its particular innovations is to examine the way it distinguishes itself from artificial intelligence (AI). It does so frequently, and tends to present itself as succeeding in its aims where AI has failed. Langton explains that in focusing on intelligence — the underlying mechanisms of which were (and are) poorly understood — AI was left without a model to follow and resorted to "serial computer programming," a methodology that "bore no demonstrable relationship to the method by which intelligence is generated in natural systems." Conventional AI strove without much success to make computer programs that could think; its approach was centralized, or "top-down," and focused on cognition. A-life, in contrast, deals with behavior that emerges from the bottom up. Langton describes a-life as remaining "true to natural life," following the "key insight" that "nature is fundamentally parallel" — that is, natural systems tend to be complex aggregates of parts, each of which has its own "behavioral repertoire"; behavior arises out of the parallel operation of these parts.

A-life has developed and adopted a repertoire of formal structures and techniques that apply this philosophy. While this repertoire is not fixed or static, there are a handful of key techniques, which bear introduction here.

Genetic algorithms, a central technique, roughly simulate biological genetics in digital computation. A genetic algorithm involves a "genotype," which is a string of code specifying a "phenotype." The phenotype can be any digital artifact: an artificial organism, a three-dimensional form, or a piece of software. By simulating the genetic variations caused by sexual reproduction and mutation, a genetic algorithm alters the genotype and the phenotype; and since this process is computational rather than biological, breeding is rapid and prolific. Wide ranges of possible phenotypes can be generated, which are often automatically evaluated for their "fitness," based on some formally specified criteria. In functional applications, an accelerated process of artificial evolution is applied to find a solution to a complex problem by searching within a large range of possible outcomes.

Agent-based systems often also apply artificial genetics. These systems model individuals interacting in an artificial world; their behaviors may be as basic as breeding and eating or as sophisticated as "communicating" or cooperating. Population dynamics may emerge, such as fluctuating predator/prey balances; with artificial genetics, agents' attributes may evolve, so that phenomena such as speciation, interbreeding, symbiosis, and coevolution become possible. Some simple agent-based systems involve no genetics yet exhibit a-life's characteristic bottom-up dynamics: in *flocks* agents follow simple rules for moving through space; each individual seeks to maintain a certain distance from the others while moving forward.[5] The result is the spontaneous formation of a flock of agents, with a supple coherence that resembles that of real-life flocks or shoals.

This architecture of decentralized control has also been applied in robotics: in *bottom-up robotics* multiple sensorimotor processes operate in parallel, in the absence of a controlling "brain", or an internal representation of the sensed world. As the work of Rodney Brooks and the MIT Robot Lab has shown, this architecture can generate simple, computationally efficient robots with surprisingly robust behaviors.[6]

Finally, *cellular automata* manifest this local-global transition in a purely formal domain. In these systems, an array of logical units or cells is computed with a set of simple rules for how each cell's future state is affected by the current states of its neighbors. In the best-known cellular automaton, the Game of Life, a two-dimensional array of one-bit, on/off cells and a handful of simple transition rules give rise to what seems to be a clockwork nanobestiary: cell formations blossom and disintegrate; oscillating, mobile formations crawl across the array.[7] This striking emergence of complexity from simplicity, and lifelike dynamics from formal rules, is frequently invoked in arguments for the merits of the a-life approach.

A-Life Art

Shortly after artificial life's self-declared inception in 1987, artists began to apply its techniques. Its earliest adopters were artists with interdisciplinary interests, followers of the biological and computational sciences who had the technical and conceptual means to begin experimenting with artificial life. William Latham and Karl Sims were prominent among these; their work was shown in major cultural institutions and on the new media festival circuit in the early 1990s. It demonstrated the viability and some of the potential of the conjunction of a-life and art making, and sparked the interest of other artists working in digital media. Since then increasing numbers of artists have taken up a-life concepts and techniques. While in terms of contemporary culture, or even contemporary art practice, a-life art remains a "fringe" activity, it has come to be recognized as an active area within new media practice. Publications such as *Leonardo* and festivals including Ars Electronica have devoted space to a-life art; in 1999 an annual competition for a-life art was inaugurated with *Life 2.0*.[8]

In a process mirroring the expansion and diversification of artificial life science, a-life art has come to encompass work in a wide range of forms, reflecting diverse intentions and perspectives. The early works in the field focused on a single key process — artificial evolution — and its application in generating aesthetic objects. In the following decade, artists began to draw on other elements and forms: ecosystem simulations, cellular automata, and behavioral robotics. These techniques are applied across the gamut of "new media" forms: digital image, animation, interactive installation and CD-ROM, on- and off-line virtual environments, and static, robotic and biological-robotic sculpture. Less obvious, though perhaps more important, is a corresponding diversity of conceptual approaches. Some artists endorse and play out a-life's aims for the synthesis of living systems; they reflect some of the progressive, futurist tendencies of a-life and the cultural discourses it has inspired. Others approach a-life critically, questioning the assumptions that underpin its techniques as they turn those techniques to creative ends. Still others draw on the technical resources of a-life only to alter them, reconfigure and reengineer them to serve particular aesthetic and conceptual concerns.

Contemporary new media artists use a-life in a variety of contexts, to a variety of ends: some works pursue an absolute, self-sufficient autonomy; others use an appearance of autonomy to provoke empathy or raise questions about human agency. Many of the artists using a-life strive for a supple, engaging form of interactivity and a work that draws the audience into an active relationship; others present aesthetic artifacts that arise through their own intense engagement with a-life processes. Some of the works considered in this book set about creating whole artificial worlds, and others seek out a complex, dynamic relationship with the physical "outside" world. In many works, the familiar appearances and behaviors of nature are imported and reproduced; the natural world is redrawn within computational space. In others, the process of rendering biology as computation comes under question; in still others, the familiar image of nature gives way to something else: a raw, blank sense of potential, of the unknown and of what could be.

Defining or delimiting a-life art is problematic, of course, though it is necessary for a review such as this. We can situate the field within a wider area of art practice that engages and applies technoscience in its form, content, and technique; Stephen Wilson has named this area "information arts" and made a comprehensive survey of it.[9] There is a range of work here that approaches a-life art in various ways, with related concerns, techniques, and approaches. Artists are working with biotechnology, another science of technologized life, though one that operates in the "wetware" of living tissue. Many other artists draw on artificial intelligence, which while it shares some of a-life's interests in autonomous agency, emphasizes mind and thought over behavior and life. A host of other tech-art forms touch on artificial life: work with robotics, avatars, or artificial agencies; generative processes or simulated worlds; and work addressing that central articulation of the natural and the technological. This work is very often concerned with a notion of artificial life in the broadest sense. Some of it may even resemble or seem to manifest artificial life forms. However, in this book, in the interest of clarity and focus, a-life art is defined quite strictly as work that specifically and deliberately takes up the techniques and processes of a-life science.

Precursors to A-Life Art

This definition gives the survey a very specific historical compass and a year zero, 1987, linked to the self-declared inception of artificial life. Yet in terms of developing an understanding of the field, it is essential to take a wider and longer view. As Lev Manovich has shown, the new media are not entirely new but have been anticipated and prefigured by old media practices and forms.[10] Such media prehistories enrich our knowledge of contemporary forms and guard against the technofuturistic rush that often characterizes new media culture. Similarly, in the interest of grounding a consideration of contemporary practice, it is useful to consider some precursors to a-life art practice. Yet how can we trace a-life art prior to a-life itself? Simply by suspending that definition momentarily and considering parallels in practice and theory. In contemporary work, artists apply and manipulate a-life's formal techniques for modeling (and perhaps instantiating) living systems. Broadly, art here seeks to mimic and apply the dynamic formal structures of life; and this book looks back to work that predates both a-life and a-life art.

The formal analogy that likens a work of art to a living organism is ancient, traceable to Plato and Aristotle, who use the body as a model of organization and coherence in discussions of rhetoric and drama.[11] That analogy reappears in the work of the German romantic poet and scholar Johann Wolfgang von Goethe around the turn of the nineteenth century. In fact Goethe's philosophy of nature parallels that of a-life in emphasizing an appreciation of the dynamic living whole over the constituent parts while also proposing a common underlying formal structure. This is exemplified in Goethe's notion of the *Urpflanze,* or "ur-plant," the archetype or template that underpins all real plant forms. Rather than a fixed template or Platonic ideal, the *Urpflanze* was, as one contemporary commentator remarks, "a vision of a dynamic pattern."[12] Goethe regarded the study of nature — based on an "intuitive awareness" of the organic whole — as a communion with the divine and the ultimate goal of art: "The highest demand made on an artist is this: that he be true to nature, that he study her, imitate her, and produce something that resembles her phenomena."[13] This resemblance is more than an image, however; it is procedural. In art "we can in the end rival nature only when we have learned, at least in part, her method of procedure in the creation of her works."

These ideas are echoed by Goethe's contemporaries and in a lineage of major nineteenth-century figures. Around 1800, August Wilhelm von Schlegel writes that art "must form living works, which are first set in motion, not by an outside mechanism, like a pendulum, but by an indwelling power."[14] Later, Samuel Taylor Coleridge developed this notion in his critical writing on Shakespeare; for Coleridge, the true work of art is organic, and organic form, unlike the arbitrary imposition of "mechanical" form, "is innate; it shapes as it develops itself from within."[15] This is the idealist and Romantic core of a line of organicist thinking that manifests itself across the arts, in literary criticism, architecture, and musical analysis. In the visual arts these ideas are taken up in the European and Russian avant-gardes during the early decades of the twentieth century.

It is clear that a-life art is engaged, in a very general way, with the underlying forms of living things; however, it is also engaged in the translation of those dynamic forms into technological media, into structures of code and engineering, into explicit and formal rules and processes. The clearest predecessors for a-life art practice, then, are those that combine these organic ideals with a tendency towards rigor and systematisation, where creative organisms arise not through the transfer of an ineffable vital essence but from the interactions of formal elements in a medium deliberately abstracted from nature.

The work of Paul Klee provides a rich example of exactly this combination of organic idealism with formalist thought. Klee's work expresses a Goethean sense of nature but manifests it in refined, considered abstraction. Once again, Klee begins with the study of nature and an understanding that moves from surface to dynamic formal structure. The artist's intuition "can transform outward impression into functional penetration. . . . Anatomy becomes physiology."[16] Here, too, this intimate, structural understanding enables the artist to "form free abstract structures which surpass schematic intention and achieve a new naturalness, the naturalness of the work." Examples of this process can be found in Klee's notebooks. In one 1923 lecture, Klee makes a graphic analysis of plant forms that abstracts them into general principles and "forces"; the leaf stem is an energetic vector that exhausts itself as it branches, and at its endpoint is contour, the outline of the leaf.[17] Beginning in the observation of nature, this becomes a lesson in the rules of an abstract, graphic cosmos, and in the relation of line to contour and plane. Klee set a creative exercise demonstrating the final, synthetic, or creative stage of his methodology: it was entitled "Imaginary leaves on the basis of the foregoing insight into basic rules."[18] The notebooks show Klee's own example: an artificial leaf made up of stem vectors and outline contours. While here inner dynamics crystallize into form, Klee cautions that "form is the end, death. Form-giving is movement, action. Form-giving is life."[19] The organic artwork must ultimately be alive: "Our work is given form in order that it may function, in order that it may be a functioning organism."[20]

Some of Klee's contemporaries in the Russian avant-garde pursued a similar vision of the organic artwork, though with more emphasis on the role of technology. In particular, Kasimir Malevich, founder of Suprematism, produced an expansive utopian discourse of the artwork as an autonomous organic machine. While Suprematism is widely known for its pursuit of abstract purity, emblematized by Malevich's black square, Malevich, like Klee, writes of abstract form as an approach to nature's underlying dynamics and forces. In "On New Systems in Art" (1919), the artist, observing a natural landscape, "stands and exults in the flow of forces and their harmony."[21] When these dynamics are transferred into the artwork, we find not a copy or a tracing, but "pure" or "absolute" creation, and "a work of pure, living art."[22] In "Infinity . . ." (1919), Malevich writes that the "highest and purest artistic, creative structure . . . does not possess a single form of the existent. It consists of elements of nature and forms an island, appearing anew."[23] While their exact constitution remains vague, it is clear that these autonomous islands are at once organic and technological: Malevich imagines the Suprematist machine as a spacecraft, propelled "not by means of engines, . . . but through the smooth harnessing of form to natural processes, through some magnetic interrelations within a single form."[24] These forms are so refined, so perfect, that they cleave away from the mundane

Earth and become new, autonomous, artificial worlds: "All technical organisms are nothing but small satellites, a whole living world ready to fly away into space and take up a particular position. Indeed, every such satellite is in fact equipped with reason and prepared to live out its own personal life."

If the details of Malevich's vision are indistinct, it must be partly because of limitations in its raw material; it was based primarily in the mechanical paradigm that defined the technology of his time. With the rise of electronics some fifty years later came a form of technology that miniaturized and internalized the dynamics of the machine. It was this technological shift that made it possible for the Soviet Union, in 1957, to fulfill one element of Malevich's vision, sending a tiny ball of electronic circuitry into orbit around the Earth. Meanwhile, during the preceding decade, a new scientific field had been emerging in the United States, through the Macy conferences on "Circular Causal and Feedback Mechanisms in Biological and Social Systems." This was cybernetics, named by Norbert Wiener in 1948.[25] A predecessor of contemporary complex systems science, and thus artificial life, cybernetics set out to address problems across living and nonliving systems by considering both in terms of abstract causal dynamics, inputs and outputs, and feedback loops. Moreover, like a-life, cybernetics was taken up in cultural as well as scientific practices: during the 1950s, artists began to encounter and apply cybernetics. Throughout the 1960s, as interest in electronic and kinetic art forms grew, it was taken up more widely and also began to appear in critical and theoretical art discourse. This period throws up some striking precursors for contemporary a-life art.

Among the early adopters of cybernetic techniques was Hungarian-born artist Nicholas Schöffer, who gained wide attention during the 1950s and 1960s with his kinetic and cybernetic sculptures. His 1956 *CYSP I* was an articulated tower that responded to sound and colored light by moving itself and its rotating metal vanes. Schöffer describes this work as "the first sculpture to have a human-like self-determined behavior."[26] The critic Jack Burnham writes that in *CYSP I* "ambiguous stimuli . . . produce the unpredictability of an organism."[27] For Schöffer, cybernetic techniques serve an aim of "nonredundancy," enabling art to keep pace with the perpetual novelties of the mass, electronic media. Moreover, Schöffer asserts, this is metacreation: "We are no longer creating a work, we are creating creation. . . . We are able to bring forth . . . results . . . which go beyond the intentions of their originators, and this in infinite number."

These ideas are echoed by James Seawright, a prominent American cyborg sculptor. He says of his works *Watcher* (1965–1966), *Searcher* (1966), and *Scanner* (1966), "My aim is not to 'program' them but to produce a kind of patterned personality. Just as a person you know very well can surprise you, so can these machines. That's the crux of what I want to happen."[28] All Seawright's works were cybernetic systems responding to environmental inputs; some, such as *Searcher* and *Scanner,* use feedback to dynamically modify their own programs. When grouped together, the works communicate among themselves: "The pieces interact and provide a continually varying pattern of independent and collective activity."

Artists in this cybernetic era also experimented with composite systems linking biological life with electronics in various ways. Anticipating Christa Sommerer and Laurent Mignonneau's *Interactive Plant Growing* (see chapter 3), Thomas Shannon and John Lifton experimented in the mid-1960s with living plants acting as electric pickups for robotic and sonic systems. A rare example of a warm-blooded composite is Nicholas Negroponte's *Seek* (1970), in which a robot arm transports and stacks two-inch cubes that form the "built environment" for a group of gerbils: the arm attempts to adaptively alter the structure to satisfy the desires of its rodent population.

With related work by Edward Ihnatowicz, Tsai Wen-Ying, and cybernetician Gordon Pask, and the animist kinetics of Robert Breer and Jean Tinguely, this period produced a strain of cyborg art that was very much concerned with the shared circuits within and between the living and the technological. A line of cyborg art theory also emerged during the late 1960s, and here again some striking premonitions of a-life art can be found. Writers including Jonathan Benthall and Gene Youngblood drew on cybernetics and cybernetic art, Benthall in his 1972 survey *Science and Technology in Art Today,* and Youngblood in *Expanded Cinema* (1970).[29] The most substantial contributor, however, was the American critic and theorist Jack Burnham. Burnham's *Beyond Modern Sculpture* (1968) builds cybernetic art into an expansive theory that centers on art's drive to imitate and ultimately reproduce life.[30]

Burnham begins at a point of artistic crisis: sculpture after World War II was apparently obliterating itself, abandoning traditional sculptural concerns for a dematerialized dynamism. This is a transition from object to system, Burnham argues, evident in forms such as kinetics, light art, cybernetic art, and environment art (13). With the rise of industrial capitalism, and the progress of science and technology, the modern environment is a sophisticated, interlocking artificial system, and this is reflected in art practice. This artificial system is, moreover, evolving; Burnham broadly invokes negentropy, or self-organization: it is "a common effect linking social, technical and biological evolution"; "each . . . moves towards a higher life form" (14). Art is inescapably involved: "sculpture . . . in a technological society must be regarded as a tiny microcosm of the entire . . . evolution." So, ultimately, if both art and technology are negentropic, then their common destiny is the creation of life. Burnham projects its arrival into the near future (our present): "The logical outcome of technology's influence on art before the end of the century should be a series of art forms that manifest true intelligence . . . with a capacity for reciprocal relationships with human beings" (15). This drive is at the core of *Beyond Modern Sculpture:* it organizes Burnham's historical account of modern sculpture, and the cybernetic art of the 1950s and 1960s is held out as its most complete realization. Twin art-historical threads of organicism and vitalism — for Burnham, the quest to convey life's metaphysical essence — converge: "[V]italism is a transitional step in this process from inanimate object to system" (76), and "the meaning of organicism . . . has already begun to converge toward a single end result — the understanding of living matter through its creation" (51).

What we find in cybernetic art, in Burnham, and in Klee and Malevich, suggests that a-life art is only the most recent addition to a modern creative tradition that seeks to imitate not only the appearance of nature but its functional structures, and that applies (or imagines) technological means to do so. More striking, though, is the Modernist-organicist drive that runs through this history, where artificial life is the very destiny of art making. This rearranges the terms of the present investigation: instead of art following technoscience and importing its techniques, a-life itself is an artistic project, even *the* artistic project. Can we understand contemporary a-life art as a continuation of this drive? Does it finally fulfill Burnham's vision of a living, cyborg art form? As it happens, Burnham later renounced his predictions in the wake of his experience as curator of the ambitious but troubled 1970 tech-art exhibition *Software.* Writing in 1974, he denounces the "archetypal desires" of science to create artificial intelligence as "Faustian myths of the highest order."[31] Later, he writes of *Beyond Modern Sculpture* that it "erred gravely . . . in its tendency to anthropomorphize the goals of technology."[32] He dismisses the cybernetic art of the 1950s and 1960s as "little more than a trivial fiasco," and the results of AI research (circa 1980) as "pale imitations." In this dramatic about-face is another possible reading of a-life art: that it, too, is replaying Faustian myths as well as myths of technological, evolutionary, and artistic progression, and that it, too, will come to be seen as a "trivial fiasco."

However this practice may be judged and rejudged in the future, it is highly significant in the present, for all the reasons outlined. Moreover, in terms of the field's development, the present moment is one where a critical examination of a-life art practice has become both possible and worthwhile. Artists have been using artificial life for around a decade, a short span in art history though a longer one in the accelerated time scale of new media practice. For most of that time, new media art has adopted a-life techniques experimentally, in scattered, initial encounters. However, activity in the field has increased in recent years, and at the same time experimentation has given way (in part) to more self-conscious, strategic engagements with artificial life. The field has developed to a point where a wide analytical account can be valuable. While a number of writers have made isolated forays, a-life art as a whole has received limited critical attention. So, my aim here is to simply provide that critical account of the field — the work itself and the conceptual and discursive structures that surround it. This book makes no claim to be an exhaustive catalog of a-life art, but it does aim to represent the range of practice in the field. Further, while the field will continue to change, and the work presented here will inevitably date, the intention is to address the broader themes and drives that it manifests and so to give an account that will remain useful even as its details age.

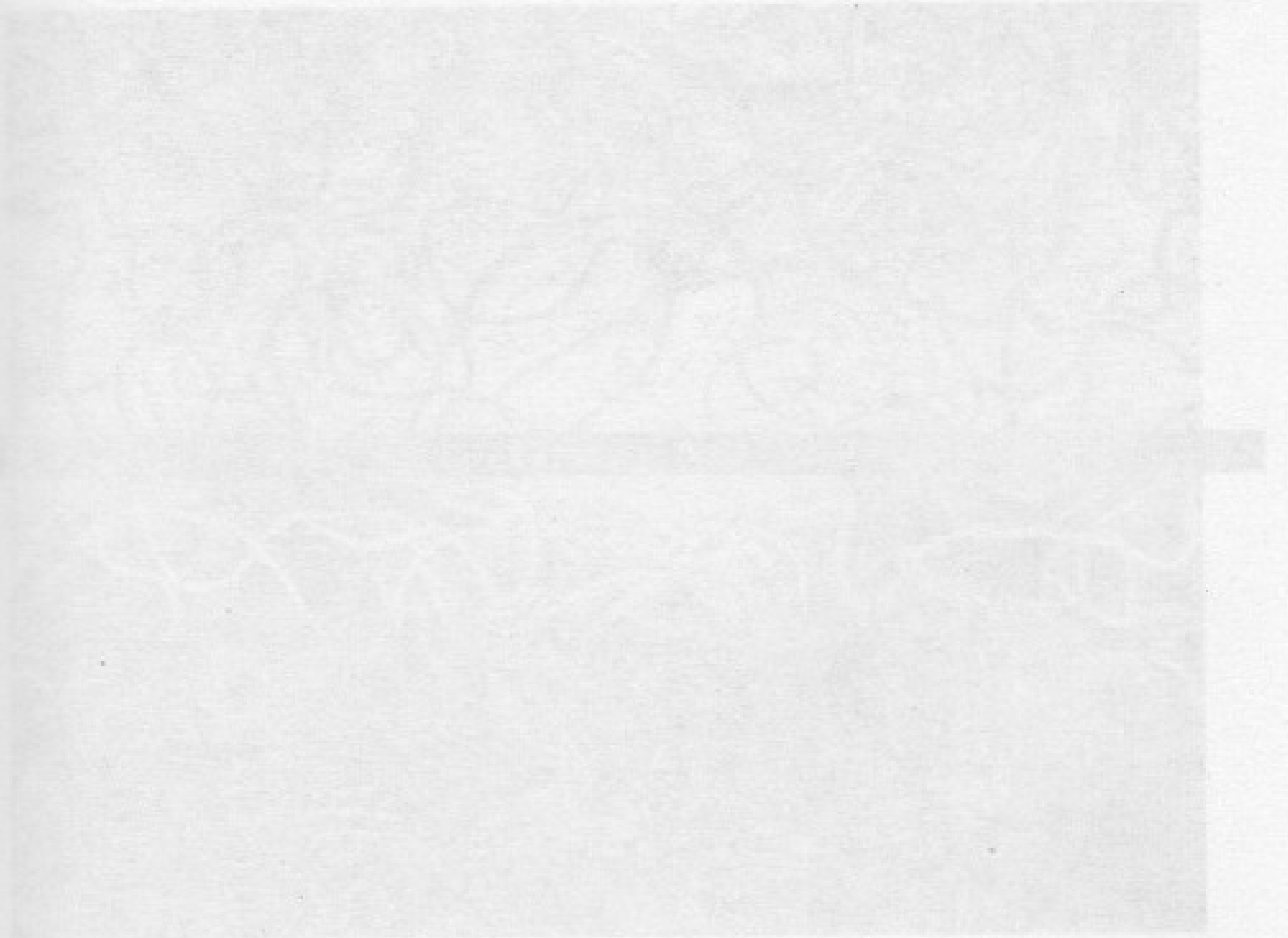
Chapters 2–5 deal with the primary material, the work itself. Here, a-life art practice is presented through a simple typology based on four of its prominent techniques and tendencies. The first of these, Breeders, focuses on processes of artificial evolution — a group that includes the earliest works of a-life art. The second, Cybernatures, expands the scope of the simulation: many of these works are interactive computational systems that mimic the tangled interrelations of organic life; all address the tension between organic life, or "nature," and its technological double. Chapter 4, Hardware, considers work that centers on a physical manifestation; as well as interactive robotic systems, this category includes biorobotic composites that involve a coupling between biological life forms and electromechanical systems. In chapter 5, Abstract Machines, the "life" in a-life recedes momentarily, in works in which the analogy implicit in these techniques is less important than their formal, generative properties.

Throughout these chapters the exposition of the work feeds directly into critical response and analysis. What is the work attempting? What does it achieve? What does it evoke or invoke? What does it exploit, critique, endorse, celebrate, or mourn? What are its implications? What does it suggest? This analysis also begins to abstract from individual instances, revealing commonalities and questions that bear on the field as a whole.

Chapter 6 pulls back to consider theoretical contexts for a-life art practice, which is not, of course, the only manifestation of artificial life in cultural thought. A-life has drawn the attention of some in fields such as cultural studies and anthropology, and their work makes some important contributions to an understanding of a-life art practice. Closer to that practice, there is a small cluster of writing addressing a-life art directly: artists and critics set out a range of aspirations, explanations, manifestos, proposals, and critiques. Reading these closely and analytically gives a sense of the variety of ways in which a-life art is being defined, justified, contextualized, and interrogated, and of the range of conceptual projects it contains.

Finally, Chapter 7 focuses on an elusive concept, emergence, which is at the core of both a-life science and a-life art practice. Emergence is the process by which complex systems seem to acquire new properties from one level of scale to another; centrally, how the complex interactions of inert matter at the microlevel give rise to life at the macrolevel. Emergence is central to a-life science's interests and its claims to be lifelike; a-life art, too, it will be argued, aspires to a state of emergence and to the surprise, the excess, the "something more" which that entails. Chapter 7 sets out to explore the concept of emergence in some detail, investigating its provenance and history, the claims for its manifestations in a-life and a-life art, and the forces and structures that act to limit and condition its operation.

Emergence is such a beguiling idea that it might not be too pat to apply it reflexively here. In fact, that might be essential if we are to take a-life's connectionist underpinnings seriously and regard culture itself as a system characterized by an interwoven and processual causality, by complex dynamics that are continuous with those not only of living systems but of their material substrate. While a text such as this is frozen, a static block, its aspirations must extend outward into those ongoing cultural dynamics, especially in a case where the subject matter is itself in flux. That is, this book is not intended to "cover" a-life art, to summarize the field once and for all. On the contrary, it is a starting point, an element to be taken up in wider systems, as the field's complex future unfolds.



CELLULAR AUTOMATA AND ART

By Brian P. Hoke

I. Introduction: A Brief History

...O fill me
With strength against those who would freeze my
humanity, would dragoon me into a lethal automaton,
would make me a cog in a machine...

-- Louis MacNeice, 1944

In the early 1950's, the well-known mathematician John Von Neumann was trying to develop what he termed a self-replicating automaton; that is, a machine whose computer brain was capable of devising instructions to effect the construction of itself. Von Neumann never actually intended for the machine to be built. Rather, he was interested in arriving at rules by which a computer could be programmed such that it could fashion an exact replica of itself. He envisioned initially a robot wandering around a warehouse littered with spare parts, identifying the relevant pieces, and fashioning an exact replica of itself. [1]

As the legend goes, Von Neumann fooled around with various approaches for a while but was not satisfied with the results. The physical movement of pieces by the robot did not seem to fit the mathematical theory he desired, nor did the complexity of various attempts at solution suit his liking. Von Neumann sought a solution that was simple, elegant, and general. [2]

Stanislaw Ulam, a fellow mathematician, suggested to Von Neumann a different approach. Consider a rectangular array of cells, much like a chessboard, in which each cell can exist in one of a finite number of states: 0, 1, 2, ... Time would progress discretely (i.e. in jumps, rather than continuously. Each increment of time would be a chance for cells to change their state. The rule governing the change of state for each cell would depend only upon the states of the cell's immediate neighbors and possibly upon the state of the cell itself. The rule would be the same for each cell and all cells would change (or not change) according to the rule with each time step. All cells would initially be in the 0, or quiescent, state; to start the automaton, place some cells in nonzero states and start the clock. Watch the cells and see how they changed according to their local - but common - rule. Ulam's grid was an example of a cellular automaton.

Von Neumann quickly saw that this system could fulfill his purpose in solving the self-replication problem. The grid of cells would be a computer; what is a computer, after all, if not a series of circuits coupled together such that the passage of current through them results in meaningful output? It is common knowledge that the fundamental language of a computer is just 1's and 0's. Interpreting these 1's and 0's allows people to interact with the computer. The cellular automaton computer would be programmed by placing some cells in nonzero states. The combination of initial condition - the states of the cells at time zero - and the rule for how each cell updates would be the means by which the computer operated. The resulting states of the cells at a later time would be interpreted as output.

This output could be used to instruct the movements of a mechanical arm. The mechanical arm could change a nearby grid of cells, all initially in the zero state and adjacent to the cells constituting the computer, so that this new grid would have the same initial conditions as the original grid. The computer would thus have replicated itself, and Von Neumann would have solved his problem. Von Neumann's solution required hundreds of thousands of cells, each of which could exist in one of 29 different states; **[3]** nevertheless, the fundamental design of the system was inherently simple and solved the replication problem elegantly. His result was published posthumously in 1966 as the *Theory of Self-Reproducing Automata.*

Cellular automata (CA) manifest one of the most intriguing ideas in mathematics: from simple rules and algorithms, complex patterns and behavior can result. Underlying this is the notion of scale. The rule for state-change of cells in Von Neumann's computer is local; each cell 'sees' only its immediate neighbors. Yet the combination of the right initial conditions and the right local rule produces a global pattern which, when interpreted correctly, can instruct the arm to construct a replica of the computer. Cellular automata, of which Von Neumann's self-replicating automaton is just one example, also transmit information in an interesting manner. There is no moving piece that carries data from one portion of the automaton to another. Cells convey information by referencing their neighbors; without movement, data is transmitted across the automaton.

This concept of transmitting information via only localized interaction and the remarkable complexity arising from simple, local rules are what I find most significant about cellular automata. This paper attempts to explain these concepts and to display some of the interesting results possible from cellular automata systems.

II. Sequences: Recursive and Explicit Definitions

For the rest of it, the last and greatest art is to limit and isolate oneself.

- Johann Wolfgang von Goethe, 1825

The simplest mathematical example of the information-transmission property of cellular automata is a recursively defined sequence. In general, sequences are written as strings of numbers (termed elements) with commas between each number. The sequences discussed here will be infinite in length and will be generated by some specific rule; that is, there exists an algorithm for generating each number in the sequence. The rule might be "double the last number and add three" to get the next number. A different sequence-generating algorithm might be "write down all the numbers that are one greater than the powers of two."

If a sequence is defined **recursively** then each successive element is defined as a function of an element (or several elements) preceding it. This manner of defining sequences is akin to setting up a row of dominoes; each falling domino knocks over the domino behind it. As with dominoes, one must have a place to start - there must be a domino to start the chain. With sequences this amounts to naming the first or first several element(s). This number is called (unsurprisingly) the initial or first element.

A typical recursive rule for a sequence might look like:

$S_n = 2S_{n-1}$

$S_1 = 1$. [4]

The rule is interpreted to mean: "Start with the number 1 as the first element. The next element is generated by multiplying the current element by 2." Thus, starting with 1 as the first element, one would produce the second element by multiplying 1 by 2 to get 2. The third element would be found by taking the second element (2) and multiplying it by 2 to get 4. The first few elements of the sequence defined by the above rule look as follows:

S: 1, 2, 4, 8, 16, 32, ...

Note the similarity between generating a sequence in this manner and the dominoes discussed above: the value of each successive element is found by looking only at the value of the element preceding it; no other elements are considered (obviously, though, each element depends indirectly upon all of its predecessors). The rule here is local.

An alternative means of defining a sequence is an **explicit** rule. In this rule, each element is defined in terms of its place in line, usually denoted n. The rule allows one to construct a sequence by specifying the value of the nth element as a function of n. An example might be:

$S_n = (2)^{n-1}$.

This rule can be interpreted to mean: "To generate the nth term, raise 2 to the power n-

1." To find the first term, one would substitute 1 for n, raise 2 to the power 0, and get 1 as a result. To find the third term, substitute 3 for n, raise 2 to the power 2, and arrive at 4. The beginning of the sequence defined by this rule would look as follows:

S: 1, 2, 4, 8, 16, 32, ...

Note that the two sequences are the same. This exhibits an important point: starting with very different conceptual ideas of how to generate a sequence, one can produce exactly the same sequence. The recursive formula specified a starting number and then gave a rule for how to arrive at the next element, given the value of the first element. The explicit formula gave a means by which any element could be found, without looking at the value of any other elements around it. The two formulas represent very different ways of thinking about producing or transmitting information, the former in a local manner, the latter in a global manner.

The key point of the above discussion is that complex behavior (i.e. a sequence of numbers related to each other through some complicated pattern) can be generated via a simple rule that involves only localized pieces of the system. None of the numbers in a recursively-defined sequence 'know' what the overall rule is for the sequence. Each number knows only its relationship to one other number - the elements are in a sense isolated from all elements other than their immediate predecessor. Yet the pattern is coherent from an overall viewpoint. At the same time, the pattern could have been generated by an explicit rule, in which the guiding principle, in effect, resides in the overall system - not in each element therein. This distinction between local and global viewpoints is central to the thesis of this paper, and will be discussed further below.

III. One-Dimensional Cellular Automata

[Hercule Poirot] tapped his forehead. 'These little grey cells. It is "up to them".'

- *Dame Agatha Christie, 1920*

The simplest type of cellular automata (CA) is the one-dimensional, two-state CA. It consists of two parts: a group of cells, all in either the 0-state or 1-state, and a rule specifying how each cell's state changes with time. The group of cells can be thought of a long row of boxes, each containing the number 0 or 1 (hence the name two-state.) The rule for how a cell changes (or does not change) is based entirely upon the state of the immediate neighbors of that cell.

Whereas each sequence element was a single number, for 1-D CA each element is a row of cells. One-dimensional CA elements are typically displayed as horizontal rows, with the first element at the top of the page or computer screen and subsequent elements stacked below it. The stacking allows the viewer to see the evolution of elements over time, since lower elements are later elements - elements resulting from applying the update rule to the cells of the row above. Cells are often displayed on a computer screen as pixels, with the 0 state shown as white and the 1 state shown as black.

Mathematically, one can describe this rule in terms of the sum of a cell's state and the states of its left and right neighbors. [6] Given the sum for any given cell (which means adding the value of the cell and its two neighbors), one can write a rule describing the state of the cell after the next time step. For the CA pictured in Figure 1, the rule would be described in terms of sums as:

if the sum is... 0 1 2 3

...then the state of the cell becomes 0 1 1 0

The above rule, along with the state of the cells in the first row, is all one would need to know to generate an exact copy of this picture.

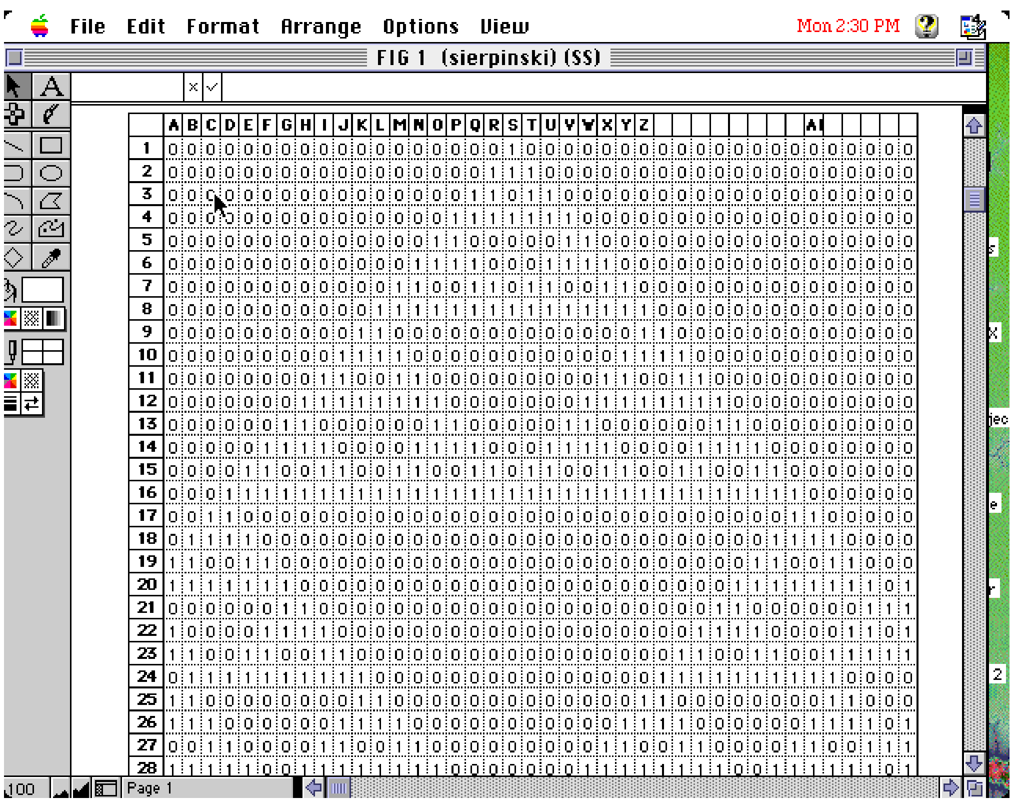


Figure 1

Interestingly, Figure 1 is an example of a geometric shape known as a Sierpinski triangle. A Sierpinski triangle is a triangle with the middle triangle, formed by connecting the midpoints of the three sides removed. Repeat this process (i.e. removing the middle triangle) on the three triangles remaining. Continue this indefinitely on the smaller and smaller triangles remaining after each removal. The result is, after infinite iterations, a figure in which any black point is a branching fork. Figure 1, while limited in resolution because of the large size of each cell, depicts this shape with 1's as black and 0's as white.

The top picture in Figure 2 is an example of a Sierpinski triangle created in the same manner, but using software **[8]** for which each cell is a pixel, rather than a box containing 0 or 1. The picture is much clearer since the size of each cell (a pixel in Figure 2, rather than a box in Figure 1) is smaller, hence the resolution is greater. As in Figure 1, the first row in the top picture in Figure 2 was 'seeded' with one black pixel. Subsequent rows were generated with the same rule used in Figure 1.

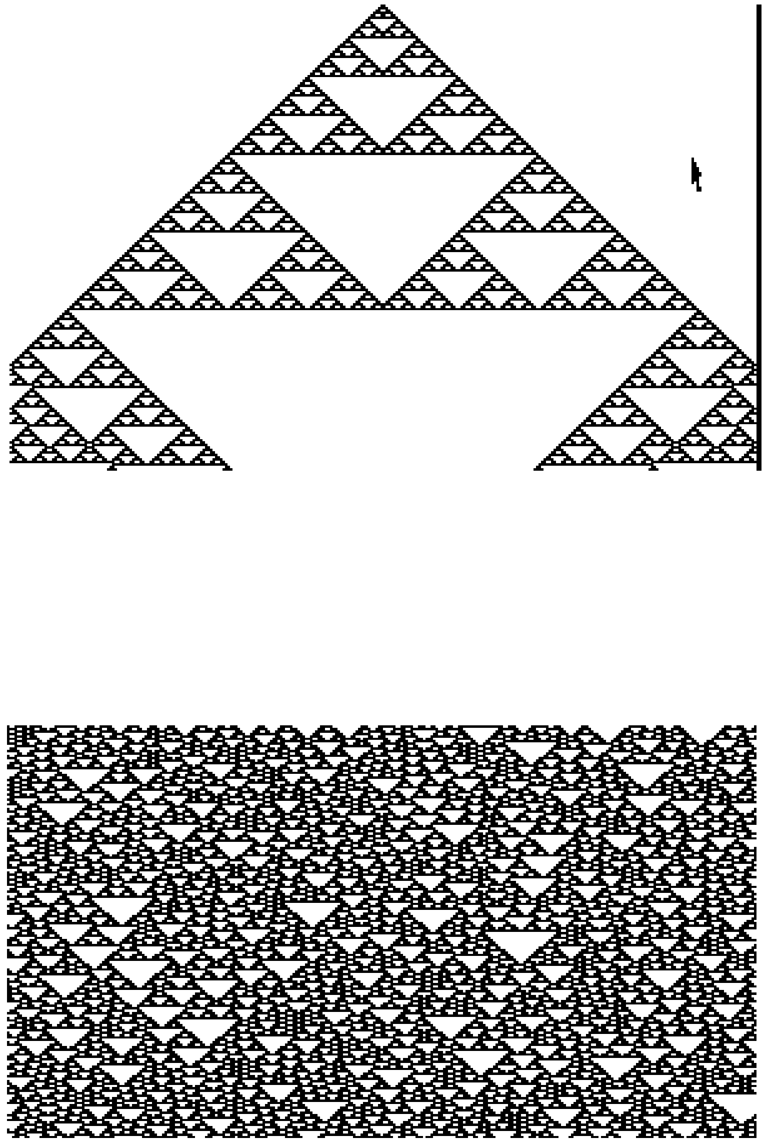


Figure 2: Two examples from Mac Cellular Automata software of 1-D, 2-state CA. Both are generated with Sierpinski rule. However, top starts with initial condition of one 'seeded' row (one black pixel in row; all other pixels white); bottom is top after many time-steps have passed.

The bottom picture in Figure 2 was constructed using the same rule and initial conditions as those of the top, but the process was allowed to run for a greater length of time. After colliding into each other, the Sierpinski triangles eventually form the netted, less ordered pattern shown. The bottom picture is an image from a later time in the life of the CA shown in the top picture; the result of applying the rule repeatedly - of numerous iterations - is the netting pattern you see. This pattern results because of the way the cells on the borders of the CA are defined, which is explained as follows.

Since each successive state of a cell is defined for this example in terms of its left and right neighbors, special rules must apply at the left and right borders of each row. This implies a choice between two alternatives. First, one could modify the rule so that the cells on the edge refer only to their one neighbor's state in deciding on how/if to change states; cells on the left border (row A in Figure 1) would change states by a rule depending only upon their right neighbors (row B), while the rule for these cells would not refer to any other cell. This presents a problem: the rows can no longer be thought as infinite and cells no longer all have the same rule. Edge effects due to these midified rules may interfere with the continuity of the pattern one would see in the absence of borders.

To solve these problems, there exists a second alternative. The rule for edge cells can refer to the opposite border cells - this is usually termed 'wraparound'. In Figure 1, cells in column A would see cells in row AN (the rightmost column) as their left neighbors and vice versa. Since the finite nature of any machine or method of displaying cellular automata precludes the possibility of portraying an infinite-length row, wraparound is in some ways a good substitute, since all cells have a right and left neighbor. As shown in the bottom picture of Figure 2, however, the left and right borders of the image tend to collide after many iterations and the figure will lose the true pattern it would have had if a computer were truly able to handle an infinitely-wide CA region. For many CA rules, though, wraparound does not interfere with the pattern formed by the cells.

An interesting question: how do the cells, each with its individual rules and each blind to the states of cells outside of its neighborhood, 'know' enough to cut out the middle one-fourth of every triangle, know enough to generate a Sierpinski triangle? First we will consider the pattern along the vertical axis of symmetry, running through the top vertex of the triangle. Note that there are horizontal rows of 1's running the width of the Sierpinski triangle at some rows. (This stops being true because of the wraparound effect mentioned above--the triangles wrap around, collide with each other, and break the pattern.) Immediately below these rows of 1's there are rows of 0's: since the sum for cells (except those near the edge of the triangle) in the long row of 1's is 3, the rule states that cells below should be in the 0 state.

Since these rows of 1's and 0's form part of the pattern we define as a Sierpinski triangle, the question of 'How do the cells know to form this pattern?' becomes 'When (i.e. at which rows) do these long rows occur?' Consider a row of all 1's running the width of the triangle and occurring at row n. This row will have 2n-1 ones, and the row below it (row n+1) will contain 2(n+1)-5 zeroes (since long rows of 0's contain 4 fewer 0's than the width of the row), or 2n-3 zeroes. Any long row of 0's bounded by 1's at right and left 'contracts' in width downward at a rate of 2 zeroes per row, since the two cells immediately below the rightmost and leftmost zeroes must become a 1 in the next generation. **[10]** A row of w zeroes will thus take (w+1)/2 rows to dissipate, or contract completely.[11]

Combining the formulae in the above paragraph, one gets the following result: after row n of all 1's, a long row of zeroes will take ((2n-3)+1)/2 rows to dissipate. Simplifying gives n-1 rows for the 0's to dissipate. The triangle formed from 0's just below the long row of 1's at row n will be gone by row n + (n-1). The next row of all 1's will thus be at n+n, or 2n. This result shows that long rows of 1's form at rows 1, 2, 4, 8, 16, ... Looked at another way, this means that triangles of smaller and smaller height, each one-half the height of its predecessor, will be removed (i.e. constructed out of 0's) as one nears the top vertex. It is through this process that the cells know to remove the middle one-fourth of any triangle - since removing the middle one-fourth involves joining the midpoints of the triangle.

A similar process occurs off the vertical axis through the top vertex, with smaller triangles of 0's forming and dissipating as above. In a 'real' Sierpinski triangle, the sequence ..., 8, 4, 2, 1 would be repeated into the fractions between 1 and 0: ..., 8, 4, 2, 1, 1/2, 1/4, 1/8, ... Since the cells forming the Sierpinski triangle in Figure 1 have a finite height and width, though, the process stops at 1. Furthermore, the wraparound effect ruins the pattern after row 19, since the left edge of the triangle wraps around, collides with the right side, and destroys the pattern. One does not see, for example, all 1's in row 32, as would be the case in a CA of infinite width.

Drawing a Sierpinski triangle using CA has important resonances in information theory. Consider the difference between describing the CA rule used to generate the triangle in Figure 1 or 2 and describing the picture itself. The former would be much more economical than specifying (as a computer must) the state of every pixel in a graphics image. When I made an informal test of the memory savings on my computer, I found that the text file containing the instruction required 26K, while the graphics file containing the Sierpinski triangle picture took up 39K worth of memory, a savings of 13K. The graphics image is much larger, yet the instructions in the text comprise all that one need know to exactly reproduce the image.

IV. Two-Dimensional Cellular Automata.

The chessboard is the world; the pieces are the phenomena of the universe; the rules of the games are what we call the laws of Nature.

- T. H. Huxley, 1870

Where the elements of sequences were numbers and those of one-dimensional CA were rows of cells, the elements of 2-D CA are planes of cells. Each cell can, as in the 1-D case, exist in a finite number of states. The rules for how a cell changes in 2-D CA typically reference the north, east, south, and west neighbors of the cell, reference the eight cells in a box around the cell, or reference nearby cells in some other similar pattern.

Since the time evolution of 2-D CA would have to be displayed as stacks of planes (each plane of which is an element of the automata), this type of cellular automata is typically shown in an animated fashion; each successive element occupies the screen as it is generated, giving the impression of animation as cells change color. A grid of cells represents the initial states of cells (usually by assigning each state a color) and subsequent elements are written over the original, showing how the cells change with each time step and giving the impression of movement. Figures 3 and 4 shows two snapshots **[12]** of a 2-D CA. [13]

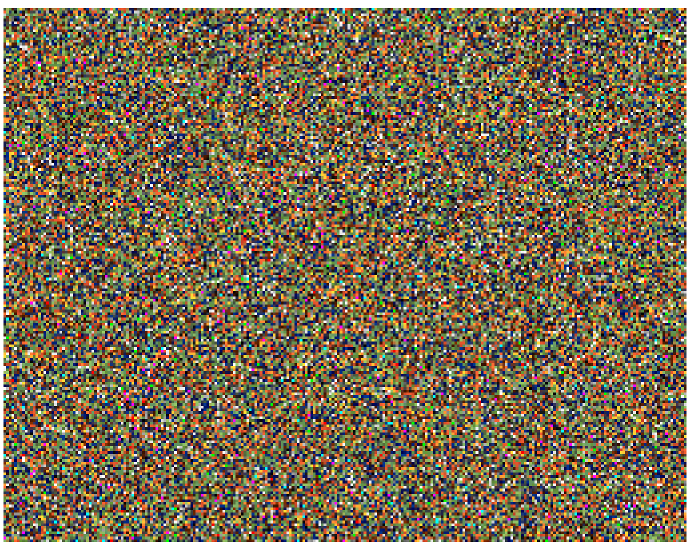


Figure 3

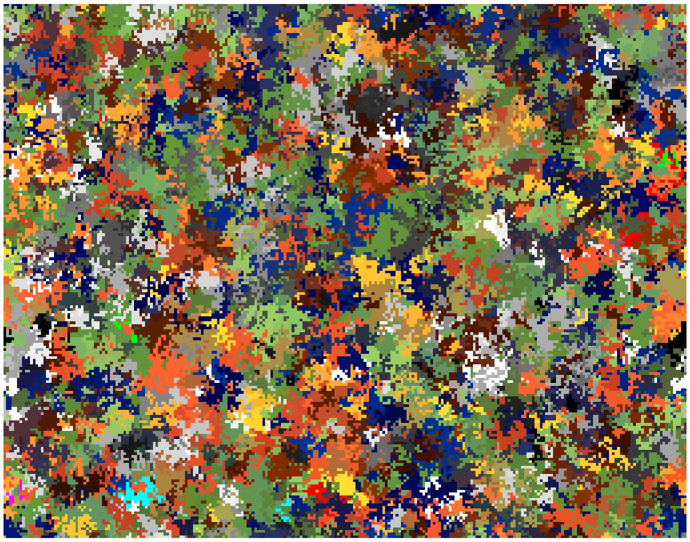


Figure 4

Figures 3 and 4 show examples of a 2-D CA rule named 'stepping stone.' Figure 3, showing the initial condition at time=0 (i.e. the first element), contains 256 colors distributed randomly across the rectangle. Figure 4 is the result of applying the rule repeatedly. The rule for each cell is as follows:

Choose a number between 0 and 1; this will be the update probability for all cells. For each cell in the array, generate a random number between 0 and 1 at every time step. If the random number generated for the given cell is higher than the update probability, the color of the cell changes to that of one of its neighbors selected uniformly at random. (Neighbor is defined as the four orthogonally adjacent cells: north, east, south, west.) [14]

Informally, this rule dictates that cells randomly 'eat' one of their neighbors. The image in Figure 4 has less than 256 colors, since some have been eliminated through iteration. As with the 1-D examples above, wraparound is again used so that the top border cells see bottom border cells as neighbors, left border cells see right border cells as neighbors, etc.

The result of this rule is the formation of planes of single color of larger and larger size. The interior of any region containing pixels of uniform color is stable under this rule, since cells that eat their neighbors won't change color. Random fluctuations, however, will enable some colors to win out over other colors. Colors will compete in the random, initial state for regions and then, when regions have been formed (as in Figure 3), they will compete at the edges of their regions. Interestingly, one color will always win in the end, taking over the entire rectangle. [15] Since the rule involves random numbers, two different colors may emerge as the eventual winner in 2 trials, even starting with exactly the same initial condition.

One can think of the stepping stone rule as a model for competition between selectively neutral genetic types. In fact, the stepping stone model for population ecology goes back to the population geneticist Sewall Wright who used it for exactly this purpose in the 1940's. **[16]** A similar use of cellular automata is that of John Conway, who developed the 'Game of Life'. Conway used 2-D cellular automata to model microorganism life. His 2-D, 2-state CA ran with rules such as "if a cell has 0 or 1 neighbors in the eight cells bordering it, then it dies of loneliness" and "a cell surrounded by 4 or more neighbors dies of overcrowding." [17] Guessing which initial conditions lead to stable, periodic, or vanishing future behavior in the Game of Life is a fascinating exercise. [18]

We now have three examples of cellular automata, with sequences included as the first. There is a very beautiful symmetry inherent in the transition from sequences to one-dimensional automata to two-dimensional automata. Sequences, which one could make a strong case for labeling 'zero-dimensional cellular automata', have single numbers, which can be thought of as points, as their constituent elements, and propagate to form strings of numbers - lines. In the typical representation of 1-D CA, each element is a horizontal line, and each successive iteration is another line: lines join to form planes. Similarly, each element in the two-dimensional automata is a plane; showing the time evolution of 2-D CA could involve stacked planes - a solid.

The various-dimension automata involve transition to increasingly higher Euclidian dimensions, from points to lines to planes to solids. Obviously, the process could be extrapolated infinitely, with the next case being solid (3-D) automata; each element would be a solid, comprising a set of cubical cells, and the changes over time would be presented as a hypersolid. This raises conceptual problems much greater than those of the examples above. (I, quite frankly, am up past my bedtime on this.) Luckily, there exists such a wealth of elegant and complex mathematics and behavior in the lower-dimensional cases that, restricting the scope of this paper to these cases, leaves a large body of material from which to draw.

V. Aspects of Cellular Automata in Art

My vegetable love should grow
Vaster than empires, and more slow....
But at my back I always hear
Time's wingèd chariot hurrying near.

- Andrew Marvell, 1681

At one point in his book, *Painting Techniques of the Impressionists*, Bernard Dunston relates Camille Pissarro's advice to a young painter; Dunston calls this "[p]robably the purest exposition of what Impressionism is all about." Since the themes addressed in this paper above appear so prominently in this passage, it is worth quoting at length.

> *Do not define too closely the outlines; it is the brushstroke of the right value and color which should produce the drawing... Paint the essential character of things; try to convey it by any means whatever, without bothering about technique. When painting, make a choice of subject, see what is lying at right and left, then work on everything simultaneously... [P]lace tones everywhere, with brushstrokes of the right color and value, while noticing what is alongside... One must always have only one master - nature; she is the one always to be consulted.* [19]

Pissarro maintains that it is the brushstroke - the detail - out of which the painting grows. The artist must work on all parts of the canvas at once, but concentrate on making the individual details true to the scene before him or her. The result will 'convey the essential character of things'. This idea is strikingly similar to the concepts underlying cellular automata: attention to individual details (the 'rules' established in the scene itself by nature) will bring about global results. Claude Monet expressed this idea even more explicitly: "... Merely think, here is a little square of blue, here is an oblong of pink, here is a streak of yellow, and paint it, just as it looks to you, the exact color and shape, until it gives your own naive impression of the scene." [20]

It is noteworthy that some of Pissarro's work is done in the Pointillist style. Dunston writes of a Pissarro painting from his Pointillist period that "broad color shapes are made up of separate, repetitious strokes, setting up a rhythm or a sense of contained movement throughout the painting." [21] Pointillist works such as those by Pissarro or Georges Seurat seem to express a cellular automata style of thinking about art, breaking a subject down into constituent blobs of color to render a scene or portrait. While Seurat and other Pointillists were intent upon creating subtle color gradations and shimmering color effect in their works, I cannot help but think that painted from a sensibility related to cellular automata.

Figures 5 and 6 show two pictures of the actress Sharon Stone. The first is a scanned image, the second the result of applying repeatedly the stepping stone rule to the scanned image. (David Griffeath, who wrote the software I used here, included Sharon Stone's picture as a play on the words 'stepping stone.') As in Figures 3 and 4, iterating the stepping stone rule has the effect of fostering competition between colors, with larger and larger planes of color formed. Note the similarity between the latter image and an impressionist painting. (Despite the similarity, I suspect that Renoir would have been against at the colors in the left side of the face. [22]) In Figure 6, I especially like the way in which the light coming from the right side of the picture is expressed through the planes of high light in the hair and background. Drawing too much of a conclusion from these two images is, I believe, stretching an analogy too far, but the reader will no doubt recognize some similarities between art of the late nineteenth century and the image presented.

Figure 5

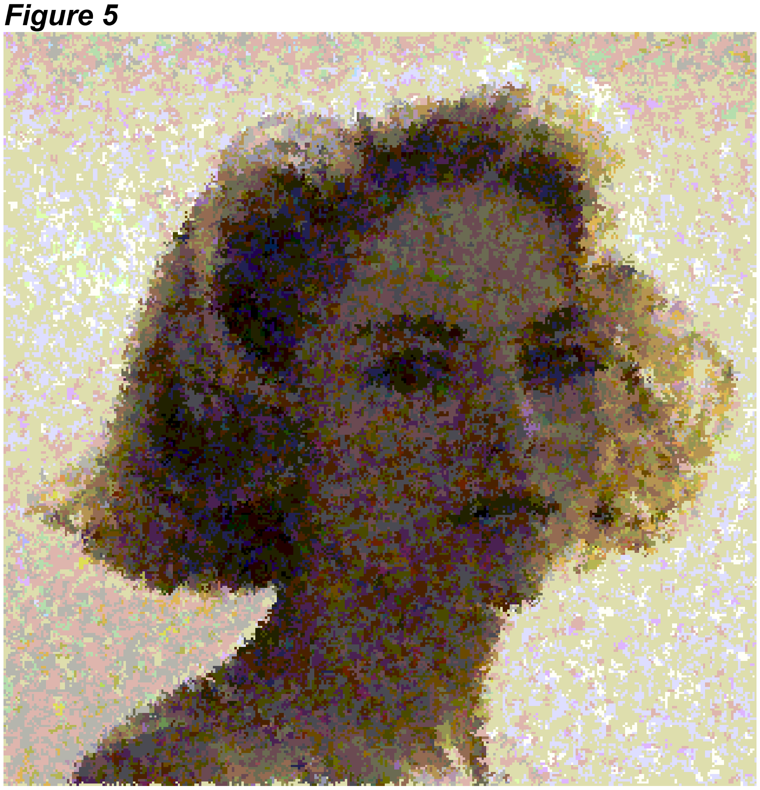


Figure 6

The desire of Friedensreich Hundertwasser, the Austrian-born painter and graphic artist, to find a happy medium between rigidity and freedom in art also reflects similarities with concepts inherent in cellular automata. Dismayed by, as Pierre Restany writes in his book *Hundertwasser*, both "the rational conventionalism of the apparent geometric rigor" and "the uncontrolled license of Tachist automatism," Hunderwasser sought a middle ground. **[23]** For him, both extremes were artistic dead ends. As such, Hundertwasser founded the school of Transautomatism (with himself as lone member).

Transautomatism stressed the "evolutionary slowness of vegetal-vegetative order." **[24]** Hundertwasser endeavored to think like a plant in producing his work; that is, to follow the biological rules (as he saw them) inherent in plant growth and life. [25] "I paint flat horizontally, without an easel," wrote Hundertwasser, "this is a vegetal, earthbound discipline. My colored lines are like the sap rings on trees, like sediments of nature, like organic growth." [26] Restany describes Transautomatism as "controlled automatism compatible with biological...determinism." [27] Spirals play an important role in Hundertwasser's art; he saw them as "organic, biological, and vegetative" [28] and the antithesis of the straight line, the ultimate symbol of all he hated about geometric rigor in art and architecture. Interestingly, spirals are often found in 2-D cellular automata; it is a stable pattern under many 2-D rules.

VI. Creating Cellular Automata Art

Rules and models destroy genius and art.

- William Hazlitt, 1839

As part of this project, I wanted to create a work developed through a CA outlook, using some of the techniques and media from earlier class assignments. My goal was to establish the rules and initial conditions, bury myself in the interrelationships between 'cells' and in the details of the work, and emerge at the completion to discover how the piece as a whole turned out. There would be no thought of any overall considerations during the creation - I would not let my desire for balance or coherence override the local rules. It would be as if I set up an experiment in a lab, left for the night, and returned the next day to examine the results.

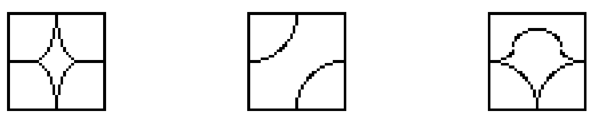
I chose printing as the medium, since the blocks I would use fit nicely into the CA notion of cells. The inspiration for the work was a mental picture of the Big Bang, the creation of the universe; postulating a limited number of particles and a rule for how those particles would propagate, I would start with a small square of four blocks in the middle of the paper and see how the surrounding blocks formed. This would be a 1-D

CA with the initial row bent around into a square. Successive rows would be rings or 'concentric squares' around the first. I would represent time by changing the value of one color. The darkest value of blue in the center would denote the earliest particles, while increasingly lighter values would express the passage of time, later and later.

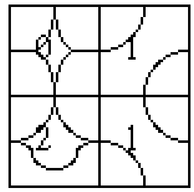
I chose three different square block designs, one for each of the three states in this 1-D CA. Each design contained arcs of some sort and touched the middle of three or four sides of its block, thus ensuring that lines would meet up between most adjacent blocks regardless of the arrangement of the blocks on the paper; thus large patterns would form from the small designs. I then constructed several different rules and proceeded, on a computer, to test the results of each one.

The three blocks shown below correspond to states 0, 1, and 2.

0 1 2

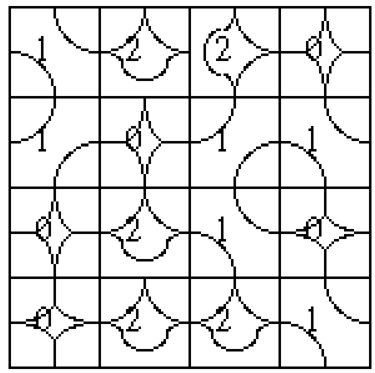


at time=0



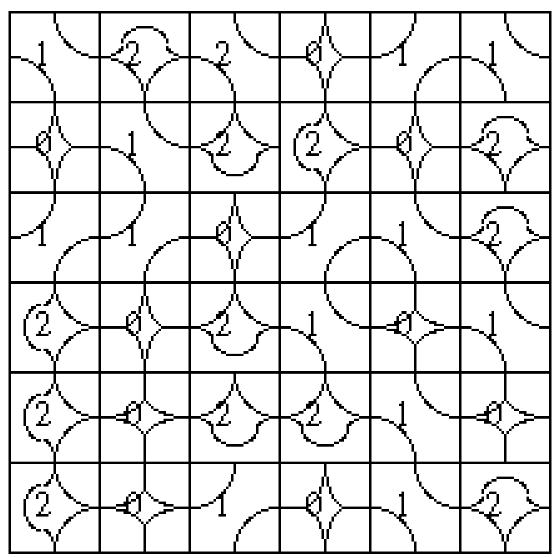
at time=1

rule has been applied 12 times to find states of cells in second 'ring'



at time=2

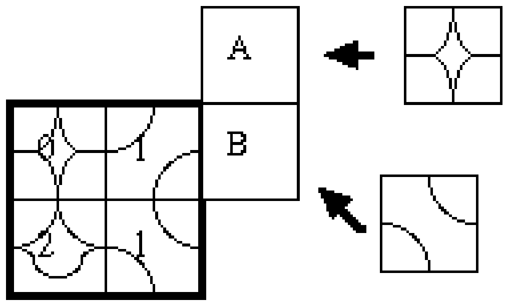
rule has been applied 20 times to find states of cells in third 'ring'



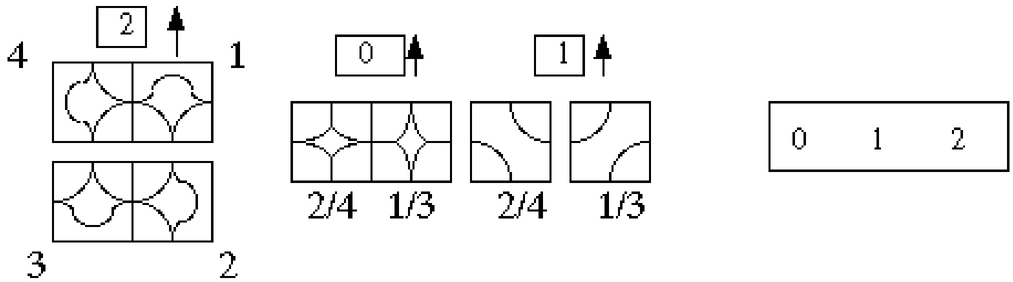
I started arbitrarily with four blocks formed in a square in the middle of the page; this would be the big bang itself. (Figure 7 shows the development of one sketch.) The rule for generating new blocks was as follows: add the states of a block on an edge and its interior neighbor; the block outside it will possess state according to the following rule:

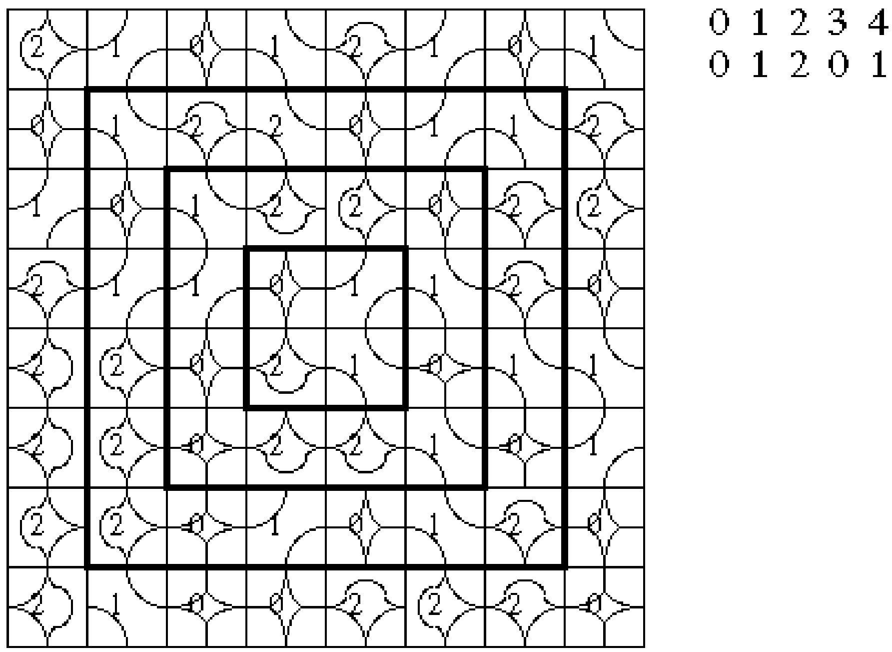
if the sum is...	**0**	**1**	**2**	**3**	**4**
...then the state of the block becomes	**0**	**1**	**2**	**0**	**1**

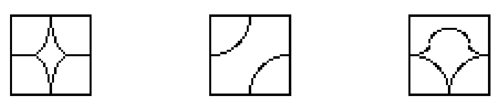
The drawing below illustrates the use of this rule. In addition to using the rule to determine the state of each cell, I randomly rotated each block before placing it on the paper. I assigned each block an arbitrary 'up' direction and rotated it by 0, 90, 180, or 270 degrees according to a list of random numbers.



For example, to find the state of cell A, add 1+2 to get 3, and look up 3 on the table above to get 0. Randomly pick a rotation for the cell and the result is as shown in the picture above. Cell B's state is determined by the 0- and 1-state cells to B's left; add to get 1 and the table gives 1 as cell B's state. Cell B's orientation (how many times it is rotated by a quarter circle) is again determined randomly. Each ring, starting with the one surrounding the initial four blocks, was formed in this manner. Figure 8 shows the sketch I used to construct the actual print.







After drawing the sketches on the computer and selecting one I liked, I cut the three blocks and printed according to the sketch. One thing I especially liked about the print was the way large patterns and structures formed from the white lines of the block designs, almost as if particles were combining to form different types of matter. I was also pleased with how the designs on the block seemed to float in front of the blocks themselves, so that the viewer could disassociate the overall pattern of thin white lines from the 'digitalness' of the blocks. Seemingly random, there was a determinism to the way in which the patterns were formed.

VII. Conclusion

Everything's got a moral, if only you can find it.

- Lewis Carroll, 1865

CA systems can generate output of amazing complexity from the repeated iteration of simple rules. The Sierpinski triangle alone would, for me, be proof of this. 2-D CA exhibit this property even more strikingly. **[29]** The variety of applications for which cellular automata are used is great, from neural networks to modeling life to counting coal dust particles. [30] And the manner in which information is propagated through space without physical movement is intellectually (at least for me) intriguing.

The aspect of CA which I find most interesting though, and an idea I had not considered before writing this paper, is the manner in which deterministic and stochastic processes are, in a sense, reconciled. Consider the stepping stone rule. The rule for each cell is well-defined, regions of color of greater and greater area will always form, and one color will always win out in the end. However, the random choice of which color a cell takes on means that the end result - which color will win out - may be different for every trial. Even in the automata with rules that are completely deterministic, different initial conditions and the length of time for which one iterated the rule of the automaton allow for great variety of output; witness the two images (Figure 3) from the Sierpinski rule.

One could perhaps even extrapolate that here is a case where mathematics suggests a reconciliation between moral absolutism and moral relativism: strict, independent rules governing phenomena, yet allowing leeway for considerable diversity. Deriving this from the mathematics discussed above may be a reach. Friendensreich Hundertwasser, however, certainly saw this in art. Walking the middle ground between rigidity and lack of any constraint did not have implications for art alone, but for how we lived out lives:

> *So I venture to say that the line described be my feet as I go walking to the museum is more important that the lines one finds hanging on the walls inside. And, I get enormous pleasure in seeing that the line is never straight, never confused, but has its reasons for being the way it is in every smallest part. Beware of the straight line and the drunken line, but especially of the straight one! The straight line leads to the loss of humanity.* **[31]**

NOTES

[1] Perry, p. 182
[2] Preston and Duff, p. 3
[3] Perry, p. 182.
[4] n is a variable representing an ordinal number, thus starting at 1. The 'nth term' denotes the term occupying the nth spot in the sequence. The n=2 term would be the second term, for example.
[5] Insert the disk "B Hoke Final Project" and double click on the "Begin Slideshow" icon to display Figure 1. Figure 1 is an image captured from an entire screen--you'll see mac desktop items (arrow cursor, icons).
[6] A good treatment of writing rules in terms of sums can be found in Perry, pp. 181-182.
[7] Click once on the play button (the right-pointing arrow) to view Figure 2.
[8] "Macintosh Cellular Automata".
[9] Click on the back button (left-facing arrow) to display Figure 1.
[10] If this is confusing, look at Figure 1 and note that triangles formed of 1's get bigger as one moves down the rows, while triangles composed of 0's get smaller. 0-triangles 'lose' two 0's after each row.
[11] Strictly speaking, a row of zeroes of width w will take x rows to dissipate, where x is the least integer greater than or equal to w/2. Since the rows in discussion here all contain odd numbers of 0's and 1's, I simplified the formula to avoid confusion.
[12] Remember that each picture is one moment in time for the CA--thus the pictures are snapshots of the state of the CA at a given time.
[13] These pictures can be generated using sstone.xpt in "WinCA".
[14] "WinCA" sstone.xpt notes
[15] "WinCA" sstone.xpt notes
[16] www.math.wisc.edu/~griffeat/logos.html. The palette I used for Figures 3 and 4 contained mostly biodiverse shades; Figure 4 looks to me like competition in a deciduous forest in autumn.
[17] www.research.digital.com/nsl/projects/life/rules.html
[18] The World Wide Web offers many interactive pages by which the user can try out the Game of Life. One I found particularly good was at www.research.com/nsl/projects/life/cgi-bin/life
[19] pp. 12-13.
[20] Dunston, p. 46.
[21] p. 75.
[22] The poor color is partly due to the fact that I had to translate the images several times between two different platforms, four different applications, and three different file types. The original was a bit better.
[23] Restany, p. 14.
[24] Restany, p. 13.

[25] Dorothy Wallace, lecture.
[26] Restany, p. 17.
[27] p. 13-14.
[28] Restany, p. 25.
[29] The reader wishing to browse some of the more interesting 2-D CA should delve into David Griffeath's WWW page, found at www.math.wisc.edu/~griffeat.
[30] Preston and Duff, pp. 3-4.
[31] Restany, p. 14.

WORKS CITED

Dunston, Bernard. *Painting Methods of the Impressionists.* New York: Watson-Guptill Publications, 1983.

Griffeath, David. Website. http://www.math.wisc.edu/~griffeat .

Perry, Kenneth E. "Abstract Mathematical Art." *Byte* December 1986: 181-190.

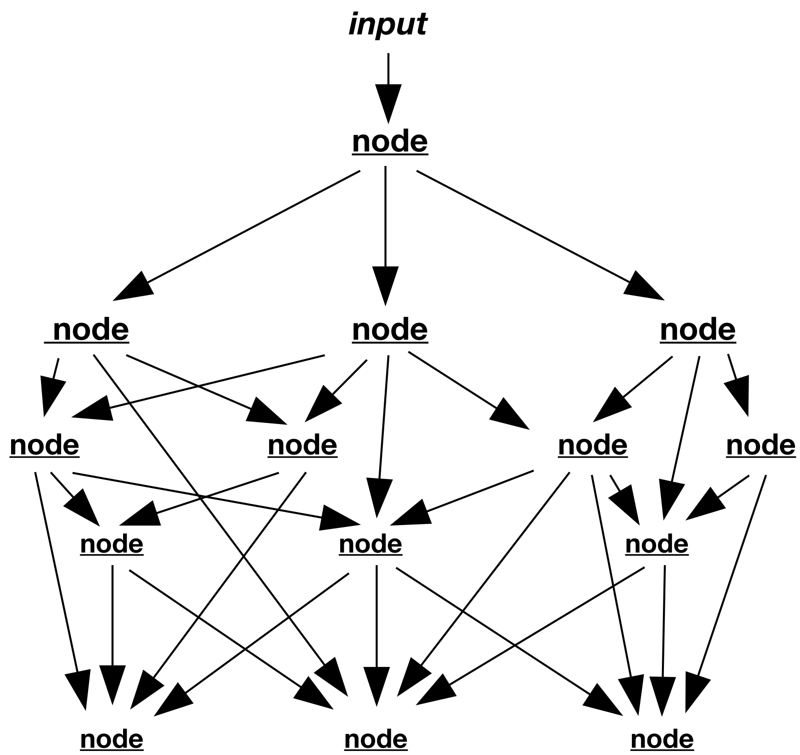
Preston, Kendall, Jr. and Michael J. B. Duff. *Modern Cellular Automata: Theory and Applications* . New York: Plenum Press, 1984.

Restany, Pierre. *Hundertwasser* . New York: Ballantine, 1975.

Wallace, Dorothy, lecture. MALS 300: Pattern. August 15, 1996.

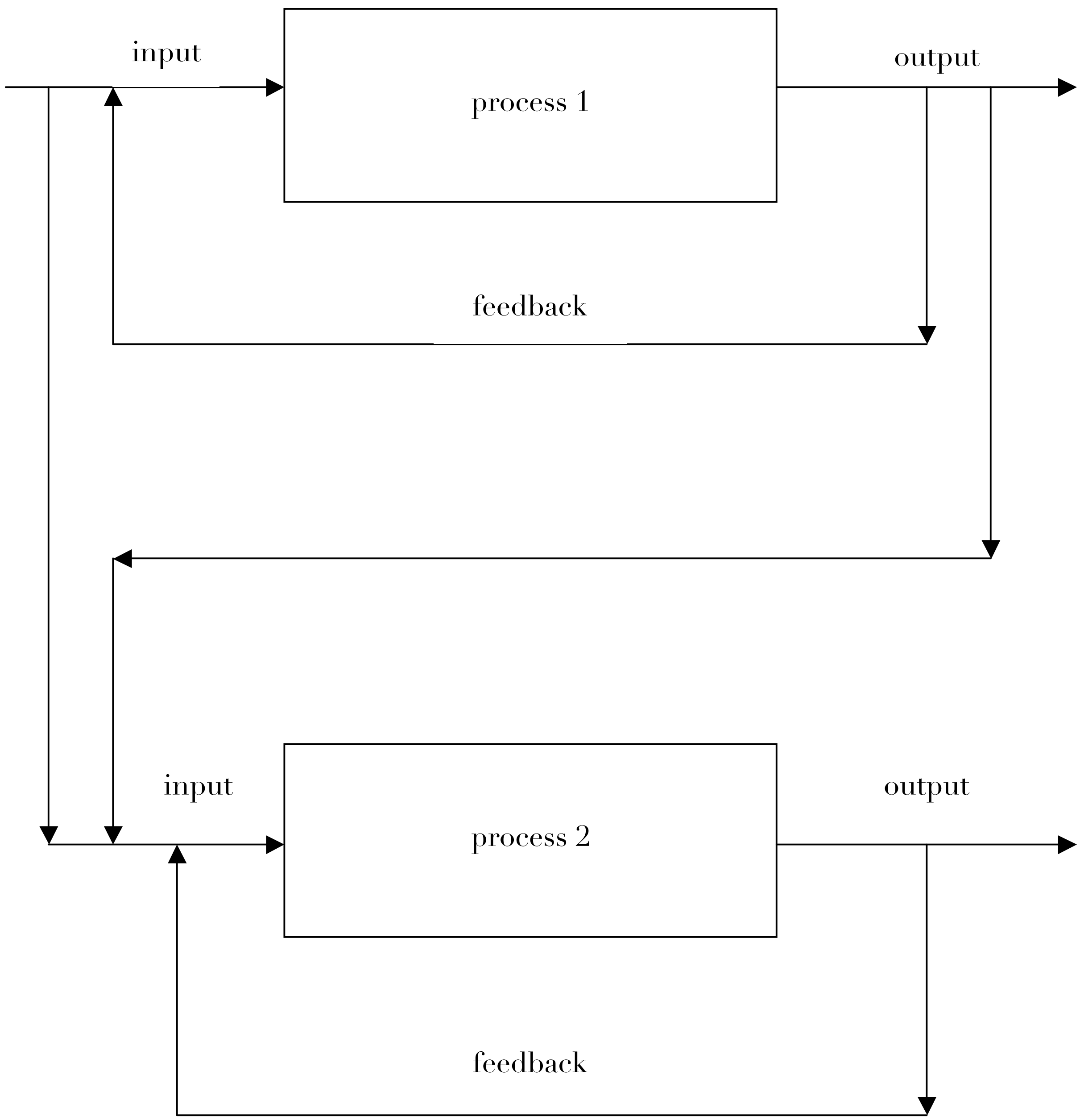
Website. http://www.research.digital.com/nsl/projects/life .

Why is interface design important in interactive art systems?

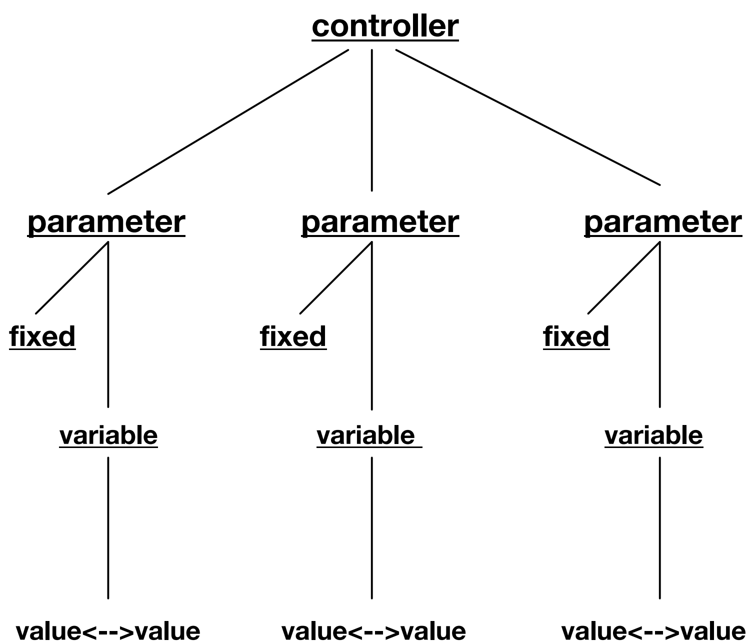


input
node
node
node
node
node
node
node
node
node
node
node
node
node
node

AN ABSTRACT MODEL OF SYSTEM DESIGN



AN ABSTRACT MODEL OF INSTRUMENT CONTROL STRUCTURE



N.B.Aldrich 2005

The History of the Interface in Interactive Art

Söke Dinkla, 1994

At the moment the catch word "interactivity" is common talk. Most often it is mentioned in connection with a revolution in television. Techno-prophets anticipate more than 200 TV channels for the near future in each home. Thus, viewers will not only be able to choose from an almost unlimited offer, they will also be able to determine the course and outcome of individual programs [1]. Proponents of these new opportunities are already praising interactivity as a means to change the passive reception of the viewer into an active one [2]. Thus, it seems as if Bertolt Brecht's Radio Theory [3], which he developed in the late twenties, is now to become reality. Brecht envisioned the transformation of broadcasting from a distribution machine into a communication device that offers listeners the opportunity to help create its content. And actually this development has been actively pursued for years by groups such as the Ponton Media Art Lab, by people like Myron Krueger, and by the communication structure of the internet.

This slightly anarchistic approach was notably absent from this year's Siggraph computer trade show in Orlando, Florida. The trade show showed that besides interactive TV games, the US-American entertainment industry is concentrating on the employment of interactive technologies in the scope of big theme parks. While the well-known Virtuality games by W-Industries individualize the player, the theme parks stress cooperation and team spirit. The company Evans & Sutherland, for example, presented the game *Virtual Adventures* in which six players search together for the eggs of the Loch Ness monster. The game appeals to typically athletic characteristics such as ambition and team spirit. It offers alternative experiences of pleasure and frustration that are the classic features of a game.

Computer games like this have a more than 20-year-old history of technological development that, remarkably, took place at the same time in military research and in art. In these years Interactive Art supplied many alternatives to the above-mentioned Loch Ness game and is essentially characterized by the attempt to "humanize" the interface between system and player. In addition, the use of technologies that address the viewer directly and involve her or him in a dialogue constitutes a decisive change in the tradition of the image. Therefore, my main attention focuses here on the reciprocal dialogue between user and system and on the design of the interface. In the following text I will distinguish six important implications of interactivity.

0. Historical Background and some Conflicts of Interactive Art

Its background in art consists of participational art forms from the late sixties like, for example, Happenings and reactive kinetic environments. Theoretical works like

Umberto Ecos *Opera Aperta* (1962) [4] contributed to our interpretation of the part played by the spectator. In German aesthetics this view was further developed, particularly by Wolfgang Kemp in the middle of the 1980s. His book *Der Betrachter ist im Bild (The Viewer is Inside the Picture)* [5], in which he describes the method of receptional aesthetics, seems to anticipate the perception principle we are experiencing today in virtual reality. But this line of tradition is not unbroken, as will be shown later on in this paper.

In a way, Interactive Art builds on the traditions of participational art forms by allowing the viewer to intervene in the action. However, in most works, unlike in Happenings, this interaction is not meant as an attack against the established art audience. Instead, it meets the needs of a media educated public. The implications of Interactive Art, though, go even further: this art also reflects the role played by computer technology [6]. This may seem complicated, because Interactivity uses the same technology it comments upon, meaning, there is a certain lack of distance. The situation of Interactive Art is therefore comparable with Video Art, which had to gain certain independence from the language of television. Both art forms demonstrate that today the role of the artist is changing significantly. Instead of being a commentator standing outside society, the artist now decides to take part in the socio-technological change and judge from within.

1. Power and Play

With the American Myron Krueger, the development of computer-controlled Interactive Art began. As early as 1969 he conceived of spaces in which the actions of visitors set off effects. In co-operation with Dan Sandin, Jerry Erdman and Richard Veneszky, he created *Glowflow* in 1969. *Glowflow* is a space with pressure sensitive sensors on its floor, loudspeakers in the four corners of the room, and tubes with colored suspensions on the walls. The visitor who steps on one of the sensors sets off either sound or light effects. In the scope of the Art & Technology movement in the late sixties, artists like Robert Rauschenberg and James Seawright created similar 'responsive environments'. But at that time no one in the 'art world' thought of creating a more complex computer-controlled dialogue and focusing the interaction itself.

In the computer sciences the situation was different. Almost simultaneously with *Glowflow*, Ivan Sutherland at the University of Utah developed the precursor of today's head-mounted-display (HMD). This display was worn like a pair of glasses and contained two small monitors, each of which showed was stereoscopic. Sensors register head movements and transmit the information to a computer that then calculates the perspective, giving the viewer the impression of moving within the image.

Thus, at the end of the sixties, two trends emerged independently of each other which have significantly influenced current Interactive Art and computer technology in general:

1. the development of 'responsive environments' in the scope of the US-American Art & Technology movement and

2. the development of the head-mounted-display

Krueger's work cannot be assigned to either of these trends. He did not participate in the projects of the Art & Technology movement, nor did he regard the head-mounted-display as a suitable interface. He thus used a different variant, which was also developed at the end of the sixties but in the scope of Video Art: it was the closed-circuit installation in which visitors are confronted by their own camera image. Krueger now combined this principle with computer technology.

In *Videoplace*, a work Krueger has been constantly developing since 1974, the visitors find themselves faced with their own projected video image that can be changed by the computer program. In *Videoplace* there are a number of different interactions in which Krueger subverts the rules of narcissistic self-reflection and self-control of the traditional video closed-circuit and lets the user play with constantly changing versions of themselves. In the most famous interaction, called *Critter*, a green figure appears on the screen and tries to make contact with the visitor. It steers towards an exposed part of the visitor's body and lands there. Then *Critter* begins to climb up the arm, shoulder and neck until it reaches the highest point of the head. Once there, it performs a joyful dance. Since *Critter* is programmed to reach the highest point of the visitor's outline, the aim of the player is to outwit *Critter*. That is, to subvert the program and develop their own rules. Thus, the interactions of *Videoplace* are not only a joyful game but are also concerned with the probing of power distribution between user and system.

Krueger's attitude towards the interface shows that he is opposed to the isolation of the user caused by the head-mounted-display. Instead, he creates an open space where it is the interaction and not the instrument that causes the proximity to the system. This has important consequences for the understanding of the interface. The technical interface - in this case the video camera - is, in a way, invisible and loses significance. It is substituted for by the application itself.

In Europe the approach to Interactive Art and the use of the interface was quite different. The situation at the beginning of the 1980s could be characterized by the catch phrase "Participation versus Interaction".

2. Participation versus Interaction

In Amsterdam in 1983 the Australian Jeffrey Shaw produced his first interactive installation. He transferred his participational concept of art, which he developed during the 1960s, to computer installations. In his first interactive installation, *Points of View*, Shaw takes up the joystick, an interface that is still customary for video games. Sitting on a chair, the spectator can move a projected video image of a stage with Egyptian Hieroglyphs. With a second joystick she or he can steer sound traces. In *Points of View*, the spectator turns into the director who individually selects the picture and sound material. The intention of *Points of View* as described by Shaw is: "the particular audio visual journey made by a spectator, who operates the joystick, which constitutes a 'performance' of this work. For the other spectators that performance becomes 'theater'. " [7]

Although in *Points of View* Shaw dispenses with the physical performance of the spectator, he still keeps his familiar terminology. The term "movement" does not any longer signify the movement of the performer in space, like in the former Happenings, but the movement of the image caused by the joystick. The projected scene can be changed in its perspective with very little physical expenditure. Thus, the computer-controlled system inverts the reception situation of the earlier Happenings. Formerly, the spectator had to change her or his position to perceive differently; now she or he induces the computer image to change its perspectives. Thus, the movement of the spectator is substituted for the movement of the image.

By means of the development in Shaw's oeuvre, the above-mentioned break in the tradition from participational art forms to Interactive Art becomes clear. New points of view are not formed by physical experience but with the help of new interactive media strategies. As presumed at the beginning, artists like Shaw address in their Interactive Art a media-educated audience, but nevertheless formulate an opposite position to the passive reception of technically produced moving images. At the same time Shaw is also criticizing certain potentials of interactive technology itself. He decides against the video camera as the interface with the system, perhaps because he considers it too invisible. Instead, he uses a bicycle in his most famous work, *The Legible City*, begun in 1988. With familiar pedaling and steering movements, the cyclist can move through a projected city of letters. The choice of this specific interface, on the one hand, aims at providing the visitors with familiar patterns of behavior. On the other hand, the bicycle as interface constitutes a refusal to do without physical activity altogether.

3. Proximity and Manipulation

At the same time as *Points of View* - in 1983 - the Canadian David Rokeby began to develop his interactive sound installation *Very Nervous System*, which originally he exhibited under various titles and changing technical equipment [8]. After Rokeby had experimented for a short time with light sensors as interface, and with analog electronics, he decided - without knowing the earlier works of Krueger - to use the video camera as interface. Rokeby's *Very Nervous System* has a much more suggestive effect than the works of Krueger and Shaw because he works with non-visual system effects as well as with an invisible interface. If one reacts intuitively to the sound, a closed circuit is created, in which music and movement are slowly becoming unified.

There is, however, a basic restriction: Krueger's *Videoplace* requires a contrasted background to distinguish the persons in space; Rokeby, on the other hand, is working only with a strong spotlight to achieve the same effect. Therefore, the causal relations between an actual movement and the sound are ambiguous. Although Rokeby employs the same interface as Krueger, their positions differ from each other. Krueger's dissatisfaction with the 'responsive environment' *Glowflow* was caused primarily by the fact that the visitors interpreted chance events as the response to their actions. While Krueger attempts a precise attribution of cause and effect to reveal the reactions of the system, Rokeby is playing with the irritation of the visitor. He hugely reduces the distance between visitor and system.

This is shown by his mot recent installation, titled *Silicon remembers Carbon*, from 1993. In this installation the visitor is even allowed to enter the image that is projected on the floor and change it with her or his movements. Infrared sensors and cameras are used as interface With this concept of reducing the distance Rokeby attempts a tightrope walk: on the one hand the visitor assumes that she or he can control the image or the sound, on the other hand the visitor is manipulated by these effects. This suggestive power of interactive correlation is only disturbed by the fact that Rokeby, as well as Shaw and Krueger, creates environments that allow the presence of more than one visitor.

The works of Shaw, Rokeby and Krueger are conceived as environments. This is not the case with contemporary works created in the United States. Most of them are conceived as installations, that is, the surrounding space is involved less strongly and the user often has direct access to the input instruments. The most common input instruments are the touch screen and the mouse. As the works of Krueger, Shaw and Rokeby have shown, the description of the interface is not restricted to its technology. The same holds true for videodisc installations.

4. Strategies of Seduction

Around the same time as similar works by the group associated with Glorianna Davenport at the Media Lab at MIT [9], Lynn Hershman from San Francisco developed her first interactive installation, *Lorna*, finished in 1984. *Lorna*, and Hershman's second installation, *Deep Contact* (1990), both work with verbal requests like "Help Lorna Leave Her Home!". The picture sequences and texts depict women in the world of media as passive objects of male desire. In *Deep Contact*, changes in a projected video image are triggered by touching a screen. Touching the body parts of the character Marion on a touch-screen sets off different strands of narration and, according to Hershman, 'entangles the viewers in meeting their own voyeurism' [10].

Her most recent work, *A Room of One's Own* (1992), also attributes this part to the spectator: the visitor looks through a little periscope into a small bedroom with a back wall where sequences of images are projected. The interaction in Hershman's work is being sexualized by the tactility of the touch screen (in *Deep Contact*) and by the intimacy of the observed situation (in *A Room of One's Own*). At the same time a fatal situation ensues. As soon as the spectator acts, he or she is caught in the role of voyeur. Hershman does not use interactivity to free the user from passivity, but to expose him or her as a voyeur. Put differently, the desires of the audience become the cause for the repressive depiction of women in media. Not even interactive technology can change that fact.

5. Nonlinear Narration

The New Yorker Grahame Weinbren produced his first interactive installation, *The Erlking*, in 1986. In this installation the interaction is mainly initiated and born by mysterious, almost static images. Weinbren - in co-operation with Roberta Friedman - works with distinctly cinematographic sequences.

The first picture shows the soprano Elisabeth Arnold singing Schubert's song *Der Erlkönig*. This picture functions as leitmotif and guideline assistance the user can return to again and again. The other pictures are partly based Goethe's ballad in which an old man narrates the saga of *The Erlking*. From the original sequence, the structure of the narration branches out. It goes not only into detail but also into additional aspects which are only loosely associated with the main plot or the backup picture. In addition to the storyline, Weinbren uses Freud's 1918 case study "From the History of an Infantile Neurosis" to try out a nonlinear access to the sequence of images.

Narration and song in *The Erlking* are being quoted as historical examples of oral tradition and are confronted with the nonlinear interactive form of narration. As a result, the interactive system takes over the role of pictorial memory. The user occupies the role of the director and cutter respectively, similar to Shaw's *Points of View*. Weinbren hopes that interactive technologies are a more appropriate means to tell these old stories [11]. This hope is problematic, though, since with originally linear storylines the fragmentation of content doesn't necessarily lead to a better understanding. Only if the stories were very well known today - meaning if they have a kind of social significance - the interactive access could possibly add new points of view.

6. Remembering, Forgetting, Reconstructing - The 'Surrogate Travel'

The New Yorker Ken Feingold is the first who uses a touchscreen as interface without integrating a second monitor. In his first interactive installation, *The Surprising Spiral* from 1991, the surface susceptible to touch is set in the cover of a book. Fingerprints and two open hands inside the book indicate that this object may be touched. Thus, the book functions as an interface to the pictorial action of *The Surprising Spiral*. A second contact point, depicting a mouth with a light source, makes sound manipulations possible.

On the videodisc of *The Surprising Spiral*, pictures and sound are stored that Feingold recorded in India, Japan, Argentina, Thailand, Scotland and the United States over a period of 12 years. The documentary pictures are contrasted with fast-moving Japanese TV advertisments and colored computer animations. Feingold makes a collage out of disparate film material from different contexts, such as ethnographical, cultural, historical, religious, aesthetic and medial contexts. Thus he combines nearly all the approaches that until now made the reconstruction of historical facts possible. It becomes clear that, despite the partly documentary film material and the mostly photorealistic video pictures, the aspect of documentary truth in the *The Surprising Spiral* is of no importance [12].

Because of the missing mise-en-scene, it is only the interaction, or to be more precise, the filling up of empty positions, which creates a new context for the user. Thus, her or his part in the reconstruction of reality seems to be autonomous to a large extent. Although Feingold - with the book as interface - is quoting from reading culture, his position differs fundamentally from Shaw's, who, in *Legible City*, tries to mediate between reading culture and interactive perception. In Feingold's *Surprising Spiral*, the book is a relict of times past - aurally charged, but nevertheless hollow and robbed of its original function. What today is preserved or forgotten as history does not follow the laws of written culture anymore, but instead is determined by the technological memory media. The reconstruction of the stored material is determined by the perception strategies of these new media.

The Surprising Spiral does not allow the purposeful approach of certain places, which is still possible in Shaw's *Legible City*. Its place is taken by the non-directional, intuitive exploration of images and texts. This gliding through the picture sequences is similar to the images of Feingold's travel impressions - short moments which are unrepeatable, which are always remembered, or reconstructed differently or sometimes even forgotten.

On the basis of this sketch showing the beginnings of Interactive Art, one cans see critical concepts about the role of interactivity in society. By discussing the works of Myron Krueger, Jeffrey Shaw, David Rokeby, Lynn Hershman, Grahame Weinbren and Ken Feingold, I have distinguished six important implications of interactivity:

1 Power and Play
2 Participation versus Interaction
3 Proximity and Manipulation
4 Strategies of Seduction
5 Nonlinear Narration and
6 Remembering, Forgetting, and Reconstructing

7. The Second Generation

In the past years, a second generation of interactive artists has emerged. As with every second generation, things are both easier and more difficult for them. On the one hand, the artists are able to build on what has already been achieved and on the other hand, they have to fulfill expectations of new developments. This young generation shows a clear geographical separation regarding the technologies used. While North American artists like Bill Seaman and Luc Courchesne are working with interactive installations and are using a touchscreen as interface, in Europe, and especially in Germany, the environment is asserting itself. The group Supreme Particles from Germany, for example, is working with the video camera as interface, like Krueger and Rokeby did. In *Architexture*, the recorded image of the visitor is reproduced as a metallic-organic color pattern on a moving projection screen. The computer-generated graphical alienation of the image is so pronounced that recognition is difficult. The fascination of a game is created by the inner life of the image that pulsates between its own morphology and the representation of the visitor.

The sea animals in *A-Volve* by Christa Sommerer and Laurent Mignonneau have an autonomous existence, too. In *A-Volve*, the visitors create little sea creatures with which they can then interact in a large water basin. The individual virtual creatures react very differently to the hand movements of the visitors. Some can be attracted, others try to flee. As their behavior is very difficult to predict, they create free play for the visitors who start to ascribe individual characteristics to the various animals. The interface Sommerer and Mignonneau worked with has completely lost its technoid character. This idea was previously employed by the artists in their 1992 work, *Interactive Plant Growing*. Here, the reaching for real plants causes the growth of computer-generated plants on a projection screen. Sommerer and Mignonneau trace the consequences of the increasing control computer technology has over our environment. To them, the so-called artificial and natural worlds do not oppose each other, but are closely interconnected. In dealing with both, a sensibility is required that has to be partly re-learned, partly found anew.

Agnes Hegedfs' work *Handsight* requires a similar sensibility. The externalized eye - as the interface with the system - gives the viewer access to a virtual world which, in the end, is to be explored by using the sense of touch. In Joachim Sauter's and Dirk Lysebrink's *Zerseher*, too, the eye acquires tactile qualities. Through eye movements recorded by an eye tracker, a monitor image be destroyed and newly generated.

These few examples show that already the concepts for designing the interface, and with it the design of the interaction, are getting more and more subtle and diverse. The feedback loop, most conspicuous in David Rokeby's *Very Nervous System*, is getting closer in the works of the young generation. The group Otherspace – that, by the way, like the Supreme Particles and Sommerer & Mignonneau, worked at the Institute for New Media in Frankfurt - uses brainwaves to set little beetle-like beings into motion. Only when the test person manages to relax do the solar-powered beetles start to move. Their movement in turn soothes the visitor so much that the result is a very intimate relationship. The debate on Artificial Life - or AI - that took place at last year's Ars Electronica seems to have created a sort of "Frankfurt School" that is decisively influencing the development of Interactive Art. In questioning the crucial differences between the first and the second generation of interactive artists helps clarify the following various aspects:

1 Through institutions such as the Institute for New Media in Frankfurt, the Media Art Academy in Cologne and the Karlsruhe Centre of Art and Media Technologies, SGI workstations are available to young artists, especially in Germany. This is one reason why the second generation favors interactive environments (and invisible or 'natural' interfaces) over installation work.
2 While in the work of the first generation a story or metaphors often influenced the content of the work, the content of the newer works is the interaction itself, which works without any form of traditional narration. Because of this new meaning placed on the interaction, the design of the interface becomes increasingly important.
3 At the same time, the antagonism between computer system and human being is overcome. It is not so much the antagonism but the forms of future co-existence that are being reflected. That is, the affirmation of interactive technology prevails over a critical distance, but this does not result in an unreflected use of technology. Here, the first generation does not differ greatly from the second.

All in all, the multi-layered, encoded levels of meaning in early interactive works, which disclose their actual content only after a sort of decoding, contributed to a certain acceptance of Interactive Art in the 'art world'. However, this strategy had its price: the narrational contents often do not come from contemporary social contexts, but from the safe context of history. In this way some artists of the first generation addressed the 'reading-habits' of the art critic's establishment. They negated the achievements of the avant-garde, which clearly saw that art only has a chance when talking to the masses and not only to a small bourgeois elite.

This trend is starting to change with the new generation. If they continue in this direction, Interactive Art will fulfill its promise of being the beginning of a new dialogue between the two ideologically separated sections of art and technology.

References

[1] Uwe Jean Heuser: Der Computer Sbernimmt (The Computer Takes Over), in:Die Zeit, 29. 10. 1993, pp. 41 -42

[2] See for example Lynn Hershman: Art-ificial Sub-versions, Inter-action, and the New Reality, in: Camerawork. A Journal of Photographic Arts, Vol. 20, Nr. 1, 1993, pp. 20-25, p. 22

[3] Bertolt Brecht, Radiotheorie (Radio Theory), in: Gesammelte Schriften, Vol. 18, Frankfurt/M. 1967, pp. 119-134

[4] Umberto Eco, Das offene Kunstwerk (The Open Work, Opera apena), Frankfurt/M. [2]. 21973

[5] Wolfgang Kemp, Der Betrachter ist im Bild. Kunstwissenschaft und Rezeptionsästhetik (The Viewer is Inside the Picture. Sciences of Art and Receptional Aesthetics), Köln 1985

[6] See Erkki Huhtamo: Commentaries on Metacommentaries on Interactivity, in: CAD Forum, 4th International Conference on Development and Use of ComputerSystems, MediaScape, Zagreb 1993, pp 229-236

[7] Jeffrey Shaw, information on Points of View 1, 2, 3, dated 1983, withletter to the author, 24. 6. 1991

[8] For a detailed discussion of the development of Rokeby's oeuvre and itsdating see Süke Dinkla: Interaktive computergestützte Installationen. Eineexemplarische Analyse (Interactive Computer-Controlled Installations. A Studyof Some Examples), unpublished MA thesis, University of Hamburg, 1992, pp. 73-78. Here you will find a more detailed discussion of Shaw's Legible City and Krueger's Videoplace, too. A kind of summary of the MA thesis is published under the title 'Interactive Computer-Supported Installations. Examples of a New Art Form', in: CAD Forum, 5th International Conference on Development and Use of Computer Systems, MediaScape, Zagreb 1994, pp. 29-36 (originally published in: Künstlerischer Austausch. Artistic Exchange. Conference Proceedings of the XXVIII. International Congress for Art History, Berlin 1992, pp. 283-294, ill. , german)

[9] The former Film/Video Group (now the Interactive Cinema Group) produced the Videodisc Elastic Movies Disc with pieces by Bill Seaman, Luc Courchesne, Russell Sasnett and Rosalyn Gerstein in 1984. They worked at that time withBenjamin Bergery and Glorianna Davenport in the work-shop in Elastic Movie Time. This information is based on interviews with Bill Seaman in Karlsruhe(8. /9. 2. 1994) and with Glorianna Davenport (11. 7. 1994) in Cambridge and on theviewing of the Elastic Movies Disc at the Media Lab at MIT.

[10] See Lynn Hershman: (note 2), pp. 23, 24 and Lynn Hershman: Some Thoughts on Deep Contact, unpublished statement, 1991

[11] Interview with Grahame Weinbren, 18. 7. 1994, New York City

[12] Timothy Druckrey made a similar observation concerning Feingold's work Childhood/Hot and Cold Wars/The Appearance of Nature in his article 'Revisioning Technology', in: Iterations. The New Image, ed. by Timothy Druckrey, International Center of Photography New York City, Cambridge/London 1993, pp. 17-38, p. 35

II. The Interface

In 1984 the director of Blade Runner Ridley Scott was hired to create a commercial which introduced Apple Computer's new Macintosh. In retrospect, this event is full of historical significance. Released within two years of each other, Blade Runner (1982) and Macintosh computer (1984) defined the two aesthetics which, twenty years, still rule contemporary culture. One was a futuristic dystopia which combined futurism and decay, computer technology and fetishism, retro-styling and urbanism, Los Angeles and Tokyo. Since Blade Runner release, its techno-noir was replayed in countless films, computer games, novels and other cultural objects. And while a number of strong aesthetic systems have been articulated in the following decades, both by individual artists (Mathew Barney, Mariko Mori) and by commercial culture at large (the 1980s "post-modern" pastiche, the 1990s techno-minimalism), none of them was able to challenge the hold of Blade Runner on our vision of the future.

In contrast to the dark, decayed, "post-modern" vision of Blade Runner, Graphical User Interface (GUI), popularized by Macintosh, remained true to the modernist values of clarity and functionality. The user's screen was ruled by strait lines and rectangular windows which contained smaller rectangles of individual files arranged in a grid. The computer communicated with the user via rectangular boxes containing clean black type rendered again white background. Subsequent versions of GUI added colors and made possible for users to customize the appearance of many interface elements, thus somewhat deluding the sterility and boldness of the original monochrome 1984 version. Yet its original aesthetic survived in the displays of hand-held communicators such as Palm Pilot, cellular telephones, car navigation systems and other consumer electronic products which use small LCD displays comparable in quality to 1984 Macintosh screen.

Like Blade Runner, Macintosh's GUI articulated a vision of the future, although a very different one. In this vision, the lines between human and is technological creations (computers, androids) are clearly drawn and decay is not tolerated. In computer, once a file is created, it never disappears except when explicitly deleted by the user. And even then deleted items can be usually recovered. Thus if in "meatspace" we have to work to remember, in cyberspace we have to work to forget. (Of course while they run, OS and applications constantly create, write to and erase various temporary files, as well as swap data between RAM and virtual memory files on a hard drive, but most of this activity remains invisible to the user.)

Also like Blade Runner, GUI vision also came to influence many other areas of culture. This influence ranges from purely graphical (for instance, use of GUI elements by print and TV designers) to more conceptual. In the 1990s, as the Internet progressively grew in popularity, the role of a digital computer shifted from being a particular technology (a calculator, a symbol processor, an image manipulator, etc.) to being a filter to all culture, a form through which all kinds of cultural and artistic production is being mediated. As a window of a Web browser comes to replace cinema and television screen, a wall in art gallery, a library and a book, all at once, the new situation manifest itself: all culture, past and present, is being filtered through a computer, with its particular human-computer interface.[57]

In semiotic terms, the computer interface acts as a code which carries cultural messages in a variety of media. When you use the Internet, everything you access — texts, music, video, navigable spaces — passes through the interface of the browser and then, in its turn, the interface of the OS. In cultural communication, a code is rarely simply a neutral transport mechanism; usually it affects the messages transmitted with its help. For instance, it may make some messages easy to conceive and render others unthinkable. A code may also provide its own model of the world, its own logical system, or ideology; subsequent cultural messages or whole languages created using this code will be limited by this model, system or ideology. Most modern cultural theories rely on these notions which I will refer to together as "non-transparency of the code" idea. For instance, according to Whorf-Sapir hypothesis which enjoyed popularity in the middle of the twentieth century, human thinking is determined by the code of natural language; the speakers of different natural languages perceive and think about world differently.[58] Whorf-Sapir hypothesis is an extreme expression of "non-transparency of the code" idea; usually it is formulated in a less extreme form. But then we think about the case of human-computer interface, applying a "strong" version of this idea makes sense. The interface shapes how the computer user conceives the computer itself. It also determines how users think of any media object accessed via a computer. Stripping different media of their original distinctions, the interface imposes its own logic on them. Finally, by organizing computer data in particular ways, the interface provides distinct models of the world. For instance, a hierarchical file system assumes that the world can be organized in a logical multi-level hierarchy. In contrast, a hypertext model of the World Wide Web models the world as a non-hierarchical system ruled by metonymy. In short, far from being a transparent window into the data inside a computer, the interface bring with it strong messages of its own.

As an example of how the interface imposes its own logic on media, consider "cut and paste" operation, standard in all software running under modern GUI. This operation renders insignificant the traditional distinction between spatial and temporal media, since the user can cut and paste parts of images, regions of space and parts of a temporal composition in exactly the same way. It is also "blind" to traditional distinctions in scale: the user can cut and paste a single pixel, an image, a whole digital movie in the same way. And last, this operation also renders insignificant traditional distinctions between media: "cut and paste" can be applied to texts, still and moving images, sounds and 3D objects in the same way.

The interface comes to play a crucial role in information society yet in a another way. In this society, not only work and leisure activities increasingly involve computer use, but they also converge around the same interfaces. Both "work" applications (word processors, spreadsheet programs, database programs) and "leisure" applications (computer games, informational DVD) use the same tools and metaphors of GUI. The best example of this convergence is a Web browser employed both in the office and at home, both for work and for play. In this respect information society is quite different from industrial society, with its clear separation between the field of work and the field of leisure. In the nineteenth century Karl Marx imagined that a future communist state would overcome this work-leisure divide as well as the highly specialized and piece-meal character of modern work itself. Marx's ideal citizen would be cutting wood in the morning, gardening in the afternoon and composing music in the evening. Now a subject of information society is engaged in even more activities during a typical day: inputting and analyzing data, running simulations, searching the Internet, playing computer games, watching streaming video, listening to music online, trading stocks, and so on. Yet in performing all these different activities the user in essence is always using the same few tools and commands: a computer screen and a mouse; a Web browser; a search engine; cut, paste, copy, delete and find commands. (In the introduction to "Forms" chapter I will discuss how the two key new forms of new media — database and navigable space — can be also understood in relation to work--leisure opposition.)

If human-computer interface become a key semiotic code of the information society as well as its meta-tool, how does this affect the functioning of cultural objects in general and art objects in particular? As I already noted ("Principles of New Media," 4.2), in computer culture it becomes common to construct the number of different interfaces to the same "content." For instance, the same data can be represented as a 2D graph or as an interactive navigable space. Or, a Web site may guide the user to different versions of the site depending on the bandwidth of her Internet connection. (I will elaborate on this in "Database" section where a new media object will be defined as one or more interfaces to a multimedia database.) Given these examples, we may be tempted to think of a new media artwork as also having two separate levels: content and interface. Thus the old dichotomies content — form and content — medium can be re-written as content — interface. But postulating such an opposition assumes that artwork's content is independent of its medium (in an art historical sense) or its code (in a semiotic sense). Situated in some idealized medium-free realm, content is assumed to exist before its material expression. These assumptions are correct in the case of visualization of quantified data; they also apply to classical art with its well-defined iconographic motives and representational conventions.

But just as modern thinkers, from Whorf to Derrida, insisted on "non-transparency of a code" idea, modern artists assumed that content and form can't be separated. In fact, from the 1910s "abstraction" to the 1960s "process," artists keep inventing concepts and procedures to assure that they can't paint some pre-existent content.

This leaves us with an interesting paradox. Many new media artworks have what can be called "an informational dimension," the condition which they share with all new media objects. Their experience includes retrieving, looking at and thinking about quantified data. Therefore when we refer to such artworks we are justified in separating the levels of content and interface. At the same time, new media artworks have more traditional "experiential" or aesthetic dimensions, which justifies their status as art rather than as information design. These dimensions include a particular configuration of space, time, and surface articulated in the work; a particular sequence of user's activities over time to interact with the work; a particular formal, material and phenomenological user experience. And it is the work's interface that creates its unique materiality and the unique user experience. To change the interface even slightly is to dramatically change the work. From this perspective, to think of an interface as a separate level, as something that can be arbitrary varied is to eliminate the status of a new media artwork as art.

There is another way to think about the difference between new media design and new media art in relation to the content — interface dichotomy. In contrast to design, in art the connection between content and form (or, in the case of new media, content and interface) is motivated. That is, the choice of a particular interface is motivated by work's content to such degree that it can no longer be thought of as a separate level. Content and interface merge into one entity, and no longer can be taken apart.

Finally, the idea of content pre-existing the interface is challenged in yet another way by new media artworks which dynamically generate their data in real time. While in a menu-based interactive multimedia application or a static Web site all data already exists before the user accesses it, in dynamic new media artworks the data is created on the fly, or, to use the new media lingo, at run time. This can be accomplished in a variety of ways: procedural computer graphics, formal language systems, Artificial Intelligence (AI) and Artificial Life (AL) programming. All these methods share the same principle: a programmer setups some initial conditions, rules or procedures which control the computer program generating the data. For the purposes of the present discussion, the most interesting of these approaches are AL and the evolution paradigm. In AL approach, the interaction between a number of simple objects at run time leads to the emergence of complex global behaviors. These behaviors can only be obtained in the course of running the computer program; they can't be predicted beforehand. The evolution paradigm applies the metaphor of the evolution theory to the generation of images, shapes, animations and other media data. The initial data supplied by the programmer acts as a genotype which is expanded into a full phenotype by a computer. In either case, the content of an artwork is the result of a collaboration between the artist/programmer and the computer program, or, if the work is interactive, between the artist, the computer program and the user. New media artists who most systematically explored AL approach is the team of Christa Sommerer and Laurent Mignonneau. In their installation "Life Spacies" virtual organisms appear and evolve in response to the position, movement and interactions of the visitors. Artist/programmer Karl Sims made the key contribution to applying the evolution paradigm to media generation. In his installation "Galapagos" the computer programs generates twelfth different virtual organisms at every iteration; the visitors select an organism which will continue to leave, copulate, mutate and reproduce.[59] The commercial products which use AL and evolution approaches are computer games such as Creatures series (Mindscape Entertainment) and "virtual pet" toys such as Tamagochi.

In organizing this book I wanted to highlight the importance of the interface category by placing its discussion right in the beginning. The two sections of this chapter present the examples of different issues raised this category -- but they in no way exhaust it. In "The Language of Cultural Interface" I introduce the term "cultural interfaces" to describe interfaces used by stand-alone hypermedia (CD-ROM and DVD titles), Web sites, computer games and other cultural objects distributed via a computer. I think we need such a term because as the role of a computer is shifting from being a tool to a universal media machine, we are increasingly "interfacing" to predominantly cultural data: texts, photographs, films, music, multimedia documents, virtual environments. Therefore, human-computer interface is being supplemented by human-computer-culture interface, which I abbreviate as "cultural interface." The section then discusses the how the three cultural forms -- cinema, the printed word, and a general-purpose human-computer interface — contributed to shaping the appearance and functionality of cultural interfaces during the 1990s.

The second section "The Screen and the User" discusses the key element of the modern interface — the computer screen. As in the first section, I am interested in analyzing continuities between a computer interface and older cultural forms, languages and conventions. The section positions the computer screen within a longer historical tradition and it traces different stages in the development of this tradition: the static illusionistic image of Renaissance painting; the moving image of film screen, the real-time image of radar and television; and real-time interactive image of a computer screen.

The Language of Cultural Interfaces

Cultural Interfaces

The term human-computer interface (HCI) describes the ways in which the user interacts with a computer. HCI includes physical input and output devices such a monitor, a keyboard, and a mouse. It also consists of metaphors used to conceptualize the organization of computer data. For instance, the Macintosh interface introduced by Apple in 1984 uses the metaphor of files and folders arranged on a desktop. Finally, HCI also includes ways of manipulating this data, i.e. a grammar of meaningful actions which the user can perform on it. The example of actions provided by modern HCI are copy, rename and delete file; list the contents of a directory; start and stop a computer program; set computer's date and time.

The term HCI was coined when computer was mostly used as a tool for work. However, during the 1990s, the identity of computer has changed. In the beginning of the decade, a computer was still largely thought of as a simulation of a typewriter, a paintbrush or a drafting ruler -- in other words, as a tool used to produce cultural content which, once created, will be stored and distributed in its appropriate media: printed page, film, photographic print, electronic recording. By the end of the decade, as Internet use became commonplace, the computer's public image was no longer that of tool but also that a universal media machine, used not only to author, but also to store, distribute and access all media.

As distribution of all forms of culture becomes computer-based, we are increasingly "interfacing" to predominantly cultural data: texts, photographs, films, music, virtual environments. In short, we are no longer interfacing to a computer but to culture encoded in digital form. I will use the term "cultural interfaces" to describe human-computer-culture interface: the ways in which computers present and allows us to interact with cultural data. Cultural interfaces include the interfaces used by the designers of Web sites, CD-ROM and DVD titles, multimedia encyclopedias, online museums and magazines, computer games and other new media cultural objects.

If you need to remind yourself what a typical cultural interface looked in the second part of the 1990s, say 1997, go back in time and click to a random Web page. You are likely to see something which graphically resembles a magazine layout from the same decade. The page is dominated by text: headlines, hyperlinks, blocks of copy. Within this text are few media elements: graphics, photographs, perhaps a QuickTime movie and a VRML scene. The page also includes radio buttons and a pull-down menu which allows you to choose an item from the list. Finally there is a search engine: type a word or a phrase, hit the search button and the computer will scan through a file or a database trying to match your entry.

For another example of a prototypical cultural interface of the 1990s, you may load (assuming it would still run on your computer) the most well-known CD-ROM of the 1990s — Myst (Broderbund, 1993). Its opening clearly recalls a movie: credits slowly scroll across the screen, accompanied by a movie-like soundtrack to set the mood. Next, the computer screen shows a book open in the middle, waiting for your mouse click. Next, an element of a familiar Macintosh interface makes an appearance, reminding you that along with being a new movie/book hybrid, Myst is also a computer application: you can adjust sound volume and graphics quality by selecting from a usual Macintosh-style menu in the upper top part of the screen. Finally, you are taken inside the game, where the interplay between the printed word and cinema continue. A virtual camera frames images of an island which dissolve between each other. At the same time, you keep encountering books and letters, which take over the screen, providing with you with clues on how to progress in the game.

Given that computer media is simply a set of characters and numbers stored in a computer, there are numerous ways in which it could be presented to a user. Yet, as it always happens with cultural languages, only a few of these possibilities actually appear viable in a given historical moment. Just as early fifteenth century Italian painters could only conceive of painting in a very particular way — quite different from, say, sixteenth century Dutch painters — today's digital designers and artists use a small set of action grammars and metaphors out of a much larger set of all possibilities.

Why do cultural interfaces — Web pages, CD-ROM titles, computer games — look the way they do? Why do designers organize computer data in certain ways and not in others? Why do they employ some interface metaphors and not others?

My theory is that the language of cultural interfaces is largely made up from the elements of other, already familiar cultural forms. In the following I will explore the contributions of three such forms to this language during its first decades -- the 1990s. The three forms which I will focus make their appearance in the opening sequence of the already discussed prototypical new media object of the 1990s — Myst. Its opening activates them before our eyes, one by one. The first form is cinema. The second form is the printed word. The third form is a general-purpose human-computer interface (HCI).

As it should become clear from the following, I use words "cinema" and "printed word" as shortcuts. They stand not for particular objects, such as a film or a novel, but rather for larger cultural traditions (we can also use such words as cultural forms, mechanisms, languages or media). "Cinema" thus includes mobile camera, representation of space, editing techniques, narrative conventions, activity of a spectator -- in short, different elements of cinematic perception, language and reception. Their presence is not limited to the twentieth-century institution of fiction films, they can be already found in panoramas, magic lantern slides, theater and other nineteenth-century cultural forms; similarly, since the middle of the twentieth century, they are present not only in films but also in television and video programs. In the case of the "printed word" I am also referring to a set of conventions which have developed over many centuries (some even before the invention of print) and which today are shared by numerous forms of printed matter, from magazines to instruction manuals: a rectangular page containing one or more columns of text; illustrations or other graphics framed by the text; pages which follow each sequentially; a table of contents and index.

Modern human-computer interface has a much shorter history than the printed word or cinema -- but it is still a history. Its principles such as direct manipulation of objects on the screen, overlapping windows, iconic representation, and dynamic menus were gradually developed over a few decades, from the early 1950s to the early 1980s, when they finally appeared in commercial systems such as Xerox Star (1981), the Apple Lisa (1982), and most importantly the Apple Macintosh (1984).[60] Since than, they have become an accepted convention for operating a computer, and a cultural language in their own right.

Cinema, the printed word and human-computer interface: each of these traditions has developed its own unique ways of how information is organized, how it is presented to the user, how space and time are correlated with each other, how human experience is being structured in the process of accessing information. Pages of text and a table of contents; 3D spaces framed by a rectangular frame which can be navigated using a mobile point of view; hierarchical menus, variables, parameters, copy/paste and search/replace operations -- these and other elements of these three traditions are shaping cultural interfaces today. Cinema, the printed word and HCI: they are the three main reservoirs of metaphors and strategies for organizing information which feed cultural interfaces.

Bringing cinema, the printed word and HCI interface together and treating them as occupying the same conceptual plane has an additional advantage -- a theoretical bonus. It is only natural to think of them as belonging to two different kind of cultural species, so to speak. If HCI is a general purpose tool which can be used to manipulate any kind of data, both the printed word and cinema are less general. They offer ways to organize particular types of data: text in the case of print, audio-visual narrative taking place in a 3D space in the case of cinema. HCI is a system of controls to operate a machine; the printed word and cinema are cultural traditions, distinct ways to record human memory and human experience, mechanisms for cultural and social exchange of information. Bringing HCI, the printed word and cinema together allows us to see that the three have more in common than we may anticipate at first. On the one hand, being a part of our culture now for half a century, HCI already represents a powerful cultural tradition, a cultural language offering its own ways to represent human memory and human experience. This language speaks in the form of discrete objects organized in hierarchies (hierarchical file system), or as catalogs (databases), or as objects linked together through hyperlinks (hypermedia). On the other hand, we begin to see that the printed word and cinema also can be thought of as interfaces, even though historically they have been tied to particular kinds of data. Each has its own grammar of actions, each comes with its own metaphors, each offers a particular physical interface. A book or a magazine is a solid object consisting from separate pages; the actions include going from page to page linearly, marking individual pages and using table of contexts. In the case of cinema, its physical interface is a particular architectural arrangement of a movie theater; its metaphor is a window opening up into a virtual 3D space.

Today, as media is being "liberated" from its traditional physical storage media — paper, film, stone, glass, magnetic tape — the elements of printed word interface and cinema interface, which previously were hardwired to the content, become "liberated" as well. A digital designer can freely mix pages and virtual cameras, table of contents and screens, bookmarks and points of view. No longer embedded within particular texts and films, these organizational strategies are now free floating in our culture, available for use in new contexts. In this respect, printed word and cinema have indeed became interfaces -- rich sets of metaphors, ways of navigating through content, ways of accessing and storing data. For a computer user, both conceptually and psychologically, their elements exist on the same plane as radio buttons, pull-down menus, command line calls and other elements of standard human-computer interface.

Let us now discuss some of the elements of these three cultural traditions -- cinema, the printed word and HCI -- to see how they have shaped the language of cultural interfaces.

Printed Word

In the 1980's, as PCs and word processing software became commonplace, text became the first cultural media to be subjected to digitization in a massive way. But already in the 1960's, two and a half decades before the concept of digital media was born, researchers were thinking about having the sum total of human written production -- books, encyclopedias, technical articles, works of fiction and so on -- available online (Ted Nelson's Xanadu project[61]).

Text is unique among other media types. It plays a privileged role in computer culture. On the one hand, it is one media type among others. But, on the other hand, it is a meta-language of computer media, a code in which all other media are represented: coordinates of 3D objects, pixel values of digital images, the formatting of a page in HTML. It is also the primary means of communication between a computer and a user: one types single line commands or runs computer programs written in a subset of English; the other responds by displaying error codes or text messages.[62]

If a computer uses text as its meta-language, cultural interfaces in their turn inherit the principles of text organization developed by human civilization throughout its existence. One of these is a page: a rectangular surface containing a limited amount of information, designed to be accessed in some order, and having a particular relationship to other pages. In its modern form, the page is born in the first centuries of the Christian era when the clay tablets and papyrus rolls are replaced by a codex — the collection of written pages stitched together on one side.

Cultural interfaces rely on our familiarity with the "page interface" while also trying to stretch its definition to include new concepts made possible by a computer. In 1984, Apple introduced a graphical user interface which presented information in overlapping windows stacked behind one another — essentially, a set of book pages. The user was given the ability to go back and forth between these pages, as well as to scroll through individual pages. In this way, a traditional page was redefined as a virtual page, a surface which can be much larger than the limited surface of a computer screen. In 1987, Apple shipped popular Hypercard program which extended the page concept in new ways. Now the users were able to include multimedia elements within the pages, as well as to establish links between pages regardless of their ordering. A few years later, designers of HTML stretched the concept of a page even more by enabling the creation of distributed documents, where different parts of a document are located on different computers connected through the network. With this development, a long process of gradual "virtualization" of the page reached a new stage. Messages written on clay tablets, which were almost indestructible, were replaced by ink on paper. Ink, in its turn, was replaced by bits of computer memory, making characters on an electronic screen. Now, with HTML, which allows parts of a single page to be located on different computers, the page became even more fluid and unstable.

The conceptual development of the page in computer media can also be read in a different way — not as a further development of a codex form, but as a return to earlier forms such as the papyrus roll of ancient Egypt, Greece and Rome. Scrolling through the contents of a computer window or a World Wide Web page has more in common with unrolling than turning the pages of a modern book. In the case of the Web of the 1990s, the similarity with a roll is even stronger because the information is not available all at once, but arrives sequentially, top to bottom, as though the roll is being unrolled.

A good example of how cultural interfaces stretch the definition of a page while mixing together its different historical forms is the Web page created in 1997 by the British design collective antirom for HotWired RGB Gallery.[63] The designers have created a large surface containing rectangular blocks of texts in different font sizes, arranged without any apparent order. The user is invited to skip from one block to another moving in any direction. Here, the different directions of reading used in different cultures are combined together in a single page.

By the mid 1990's, Web pages included a variety of media types — but they were still essentially traditional pages. Different media elements — graphics, photographs, digital video, sound and 3D worlds — were embedded within rectangular surfaces containing text. To that extent a typical Web age was conceptually similar to a newspaper page which is also dominated by text, with photographs, drawings, tables and graphs embedded in between, along with links to other pages of the newspaper. VRML evangelists wanted to overturn this hierarchy by imaging the future in which the World Wide Web is rendered as a giant 3D space, with all the other media types, including text, existing within it.[64] Given that the history of a page stretches for thousands of years, I think it is unlikely that it would disappear so quickly.

As Web page became a new cultural convention of its own, its dominance was challenged by two Web browsers created by artists — Web Stalker (1997) by I/O/D collective[65] and Netomat (1999) by Maciej Wisniewski.[66] Web Stalker emphasizes the hypertextual nature of the Web. Instead of rendering standard Web pages, it renders the networks of hyperlinks these pages embody. When a user enters a URL for a particular page, Web Stalker displays all pages linked to this page as a line graph. Netomat similarly refuses the page convention of the Web. The user enters a word or a phrase which are passed to search engines. Netomat then extracts page titles, images, audio or any other media type, as specified by the user, from the found pages and floats them across the computer screen. As can be seen, both browsers refuse the page metaphor, instead substituting their own metaphors: a graph showing the structure of links in the case of Web Stalker, a flow of media elements in the case of Netomat.

While the 1990's Web browsers and other commercial cultural interfaces have retained the modern page format, they also have come to rely on a new way of organizing and accessing texts which has little precedent within book tradition — hyperlinking. We may be tempted to trace hyperlinking to earlier forms and practices of non-sequential text organization, such as the Torah's interpretations and footnotes, but it is actually fundamentally different from them. Both the Torah's interpretations and footnotes imply a master-slave relationship between one text and another. But in the case of hyperlinking as implemented by HTML and earlier by Hypercard, no such relationship of hierarchy is assumed. The two sources connected through a hyperlink have an equal weight; neither one dominates the other .Thus the acceptance of hyperlinking in the 1980's can be correlated with contemporary culture’s suspicion of all hierarchies, and preference for the aesthetics of collage where radically different sources are brought together within the singular cultural object ("post-modernism").

Traditionally, texts encoded human knowledge and memory, instructed, inspired, convinced and seduced their readers to adopt new ideas, new ways of interpreting the world, new ideologies. In short, the printed word was linked to the art of rhetoric. While it is probably possible to invent a new rhetoric of hypermedia, which will use hyperlinking not to distract the reader from the argument (as it is often the case today), but instead to further convince her of argument's validity, the sheer existence and popularity of hyperlinking exemplifies the continuing decline of the field of rhetoric in the modern era. Ancient and Medieval scholars have classified hundreds of different rhetorical figures. In the middle of the twentieth century linguist Roman Jakobson, under the influence of computer's binary logic, information theory and cybernetics to which he was exposed at MIT where he was teaching, radically reduced rhetoric to just two figures: metaphor and metonymy.[67] Finally, in the 1990's, the World Wide Web hyperlinking has privileged the single figure of metonymy at the expense of all others.[68] The hypertext of the World Wide Web leads the reader from one text to another, ad infinitum. Contrary to the popular image, in which computer media collapses all human culture into a single giant library (which implies the existence of some ordering system), or a single giant book (which implies a narrative progression), it maybe more accurate to think of the new media culture as an infinite flat surface where individual texts are placed in no particular order, like the Web page designed by <u>antirom</u> for HotWired. Expanding this comparison further, we can note that Random Access Memory, the concept behind the group's name, also implies the lack of hierarchy: any RAM location can be accessed as quickly as any other. In contrast to the older storage media of book, film, and magnetic tape, where data is organized sequentially and linearly, thus suggesting the presence of a narrative or a rhetorical trajectory, RAM "flattens" the data. Rather than seducing the user through the careful arrangement of arguments and examples, points and counterpoints, changing rhythms of presentation (i.e., the rate of data streaming, to use contemporary language), simulated false paths and dramatically presented conceptual breakthroughs, cultural interfaces, like RAM itself, bombards the users with all the data at once.[69]

In the 1980's many critics have described one of key's effects of "post-modernism" as that of spatialization: privileging space over time, flattening historical time, refusing grand narratives. Computer media, which has evolved during the same decade, accomplished this spatialization quite literally. It replaced sequential storage with random-access storage; hierarchical organization of information with a flattened hypertext; psychological movement of narrative in novel and cinema with physical movement through space, as witnessed by endless computer animated fly-throughs or computer games such as <u>Myst</u>, <u>Doom</u> and countless others (see “Navigable Space.”) In short, time becomes a flat image or a landscape, something to look at or navigate through. If there is a new rhetoric or aesthetic which is possible here, it may have less to do with the ordering of time by a writer or an orator, and more with spatial wandering. The hypertext reader is like Robinson Crusoe, walking through the sand and water, picking up a navigation journal, a rotten fruit, an instrument whose purpose he does not know; leaving imprints in the sand, which, like computer hyperlinks, follow from one found object to another.

Cinema

Printed word tradition which has initially dominated the language of cultural interfaces, is becoming less important, while the part played by cinematic elements is getting progressively stronger. This is consistent with a general trend in modern society towards presenting more and more information in the form of time-based audio-visual moving image sequences, rather than as text. As new generations of both computer users and computer designers are growing up in a media-rich environment dominated by television rather than by printed texts, it is not surprising that they favor cinematic language over the language of print.

A hundred years after cinema's birth, cinematic ways of seeing the world, of structuring time, of narrating a story, of linking one experience to the next, are being extended to become the basic ways in which computer users access and interact with all cultural data. In this way, the computer fulfills the promise of cinema as a visual Esperanto which pre-occupied many film artists and critics in the 1920s, from Griffith to Vertov. Indeed, millions of computer users communicate with each other through the same computer interface. And, in contrast to cinema where most of its "users" were able to "understand" cinematic language but not "speak" it (i.e., make films), all computer users can "speak" the language of the interface. They are active users of the interface, employing it to perform many tasks: send email, organize their files, run various applications, and so on.

The original Esperanto never became truly popular. But cultural interfaces are widely used and are easily learned. We have an unprecedented situation in the history of cultural languages: something which is designed by a rather small group of people is immediately adopted by millions of computer users. How is it possible that people around the world adopt today something which a 20-something programmer in Northern California has hacked together just the night before? Shall we conclude that we are somehow biologically "wired" to the interface language, the way we are "wired," according to the original hypothesis of Noam Chomsky, to different natural languages?

The answer is of course no. Users are able to "acquire" new cultural languages, be it cinema a hundred years ago, or cultural interfaces today, because these languages are based on previous and already familiar cultural forms. In the case of cinema, it was theater, magic lantern shows and other nineteenth century forms of public entertainment. Cultural interfaces in their turn draw on older cultural forms such as the printed word and cinema. I have already discussed some ways in which the printed word tradition structures interface language; now it is cinema's turn.

I will begin with probably the most important case of cinema's influence on cultural interfaces — the mobile camera. Originally developed as part of 3D computer graphics technology for such applications as computer-aided design, flight simulators and computer movie making, during the 1980's and 1990's the camera model became as much of an interface convention as scrollable windows or cut and paste operations. It became an accepted way for interacting with any data which is represented in three dimensions — which, in a computer culture, means literally anything and everything: the results of a physical simulation, an architectural site, design of a new molecule, statistical data, the structure of a computer network and so on. As computer culture is gradually spatializing all representations and experiences, they become subjected to the camera's particular grammar of data access. Zoom, tilt, pan and track: we now use these operations to interact with data spaces, models, objects and bodies.

Abstracted from its historical temporary "imprisonment" within the physical body of a movie camera directed at physical reality, a virtualized camera also becomes an interface to all types of media and information beside 3D space. As an example, consider GUI of the leading computer animation software — PowerAnimator from Alias/Wavefront.[70] In this interface, each window, regardless of whether it displays a 3D model, a graph or even plain text, contains Dolly, Track and Zoom buttons. It is particularly important that the user is expected to dolly and pan over text as if it was a 3D scene. In this interface, cinematic vision triumphed over the print tradition, with the camera subsuming the page. The Guttenberg galaxy turned out to be just a subset of the Lumières' universe.

Another feature of cinematic perception which persists in cultural interfaces is a rectangular framing of represented reality.[71] Cinema itself inherited this framing from Western painting. Since the Renaissance, the frame acted as a window onto a larger space which was assumed to extend beyond the frame. This space was cut by the frame's rectangle into two parts: "onscreen space," the part which is inside the frame, and the part which is outside. In the famous formulation of Leon-Battista Alberti, the frame acted as a window onto the world. Or, in a more recent formulation of French film theorist Jacques Aumont and his co-authors, "The onscreen space is habitually perceived as included within a more vast scenographic space. Even though the onscreen space is the only visible part, this larger scenographic part is nonetheless considered to exist around it."[72]

Just as a rectangular frame of painting and photography presents a part of a larger space outside it, a window in HCI presents a partial view of a larger document. But if in painting (and later in photography), the framing chosen by an artist was final, computer interface benefits from a new invention introduced by cinema: the mobility of the frame. As a kino-eye moves around the space revealing its different regions, so can a computer user scroll through a window's contents.

It is not surprising to see that screen-based interactive 3D environments, such as VRML words, also use cinema's rectangular framing since they rely on other elements of cinematic vision, specifically a mobile virtual camera. It may be more surprising to realize that Virtual Reality (VR) interface, often promoted as the most "natural" interface of all, utilizes the same framing.[73] As in cinema, the world presented to a VR user is cut by a rectangular frame. As in cinema, this frame presents a partial view of a larger space.[74] As in cinema, the virtual camera moves around to reveal different parts of this space.

Of course, the camera is now controlled by the user and in fact is identified with his/her own sight. Yet, it is crucial that in VR one is seeing the virtual world through a rectangular frame, and that this frame always presents only a part of a larger whole. This frame creates a distinct subjective experience which is much more close to cinematic perception than to unmediated sight.

Interactive virtual worlds, whether accessed through a screen-based or a VR interface, are often discussed as the logical successor to cinema, as potentially the key cultural form of the twenty-first century, just as cinema was the key cultural form of the twentieth century. These discussions usually focus on the issues of interaction and narrative. So, the typical scenario for twenty-first century cinema involves a user represented as an avatar existing literally "inside" the narrative space, rendered with photorealistic 3D computer graphics, interacting with virtual characters and perhaps other users, and affecting the course of narrative events.

It is an open question whether this and similar scenarios commonly invoked in new media discussions of the 1990's, indeed represent an extension of cinema or if they rather should be thought of as a continuation of some theatrical traditions, such as improvisational or avant-garde theater. But what undoubtedly can be observed in the 1990's is how virtual technology's dependence on cinema's mode of seeing and language is becoming progressively stronger. This coincides with the move from proprietary and expensive VR systems to more widely available and standardized technologies, such as VRML (Virtual Reality Modeling Language). (The following examples refer to a particular VRML browser — WebSpace Navigator 1.1 from SGI.[75] Other VRML browsers have similar features.)

The creator of a VRML world can define a number of viewpoints which are loaded with the world.[76] These viewpoints automatically appear in a special menu in a VRML browser which allows the user to step through them, one by one. Just as in cinema, ontology is coupled with epistemology: the world is designed to be viewed from particular points of view. The designer of a virtual world is thus a cinematographer as well as an architect. The user can wander around the world or she can save time by assuming the familiar position of a cinema viewer for whom the cinematographer has already chosen the best viewpoints.

Equally interesting is another option which controls how a VRML browser moves from one viewpoint to the next. By default, the virtual camera smoothly travels through space from the current viewpoint to the next as though on a dolly, its movement automatically calculated by the software. Selecting the "jump cuts" option makes it cut from one view to the next. Both modes are obviously derived from cinema. Both are more efficient than trying to explore the world on its own.

With a VRML interface, nature is firmly subsumed under culture. The eye is subordinated to the kino-eye. The body is subordinated to a virtual body of a virtual camera. While the user can investigate the world on her own, freely selecting trajectories and viewpoints, the interface privileges cinematic perception — cuts, pre-computed dolly-like smooth motions of a virtual camera, and pre-selected viewpoints.

The area of computer culture where cinematic interface is being transformed into a cultural interface most aggressively is computer games. By the 1990's, game designers have moved from two to three dimensions and have begun to incorporate cinematic language in a increasingly systematic fashion. Games started featuring lavish opening cinematic sequences (called in the game business "cinematics") to set the mood, establish the setting and introduce the narrative. Frequently, the whole game would be structured as an oscillation between interactive fragments requiring user's input and non-interactive cinematic sequences, i.e. "cinematics." As the decade progressed, game designers were creating increasingly complex — and increasingly cinematic — interactive virtual worlds. Regardless of a game's genre — action/adventure, fighting, flight simulator, first-person action, racing or simulation — they came to rely on cinematography techniques borrowed from traditional cinema, including the expressive use of camera angles and depth of field, and dramatic lighting of 3D computer generated sets to create mood and atmosphere. In the beginning of the decade, many games such as <u>The 7th Guest</u> (Trilobyte, 1993) or <u>Voyeur</u> (1994) or used digital video of actors superimposed over 2D or 3D backgrounds, but by its end they switched to fully synthetic characters rendered in real time.[77] This switch allowed game designers to go beyond branching-type structure of earlier games based on digital video were all the possible scenes had to be taped beforehand. In contrast, 3D characters animated in real time move arbitrary around the space, and the space itself can change during the game. (For instance, when a player returns to the already visited area, she will find any objects she left there earlier.) This switch also made virtual words more cinematic, as the characters could be better visually integrated with their environments.[78]

A particularly important example of how computer games use — and extend — cinematic language, is their implementation of a dynamic point of view. In driving and flying simulators and in combat games, such as Tekken 2 (Namco, 1994 -), after a certain event takes place (car crashes, a fighter being knocked down), it is automatically replayed from a different point of view. Other games such as the Doom series (Id Software, 1993 -) and Dungeon Keeper (Bullfrog Productions, 1997) allow the user to switch between the point of view of the hero and a top down "bird's eye" view. The designers of online virtual worlds such as Active Worlds provide their users with similar capabilities. Finally, Nintendo went even further by dedicating four buttons on their N64 joypad to controlling the view of the action. While playing Nintendo games such as Super Mario 64 (Nintendo, 1996) the user can continuously adjust the position of the camera. Some Sony Playstation games such as Tomb Rider (Eidos, 1996) also use the buttons on the Playstation joypad for changing point of view. Some games such as Myth: The Fallen Lords (Bungie, 1997) go further, using an AI engine (computer code which controls the simulated “life” in the game, such as human characters the player encounters) to automatically control their camera.

The incorporation of virtual camera controls into the very hardware of a game consoles is truly a historical event. Directing the virtual camera becomes as important as controlling the hero's actions. This is admitted by the game industry itself. For instance, a package for Dungeon Keeper lists four key features of the game, out of which the first two concern control over the camera: "switch your perspective," "rotate your view," "take on your friend," "unveil hidden levels." In games such as this one, cinematic perception functions as the subject in its own right.[79] Here, the computer games are returning to "The New Vision" movement of the 1920s (Moholy-Nagy, Rodchenko, Vertov and others), which foregrounded new mobility of a photo and film camera, and made unconventional points of view the key part of their poetics.

The fact that computer games and virtual worlds continue to encode, step by step, the grammar of a kino-eye in software and in hardware is not an accident. This encoding is consistent with the overall trajectory driving the computerization of culture since the 1940's, that being the automation of all cultural operations. This automation gradually moves from basic to more complex operations: from image processing and spell checking to software-generated characters, 3D worlds, and Web Sites. The side effect of this automation is that once particular cultural codes are implemented in low-level software and hardware, they are no longer seen as choices but as unquestionable defaults. To take the automation of imaging as an example, in the early 1960's the newly emerging field of computer graphics incorporated a linear one-point perspective in 3D software, and later directly in hardware.[80] As a result, linear perspective became the default mode of vision in computer culture, be it computer animation, computer games, visualization or VRML worlds. Now we are witnessing the next stage of this process: the translation of cinematic grammar of points of view into software and hardware. As Hollywood cinematography is translated into algorithms and computer chips, its convention becomes the default method of interacting with any data subjected to spatialization, with a narrative, and with other human beings. (At SIGGRAPH '97 in Los Angeles, one of the presenters called for the incorporation of Hollywood-style editing in multi-user virtual worlds software. In such implementation, user interaction with other avatar(s) will be automatically rendered using classical Hollywood conventions for filming dialog.[81]) To use the terms from the 1996 paper authored by Microsoft researchers and entitled "The Virtual Cinematographer: A Paradigm for Automatic Real-Time Camera Control and Directing," the goal of research is to encode "cinematographic expertise," translating "heuristics of filmmaking" into computer software and hardware.[82] Element by element, cinema is being poured into a computer: first one-point linear perspective; next the mobile camera and a rectangular window; next cinematography and editing conventions, and, of course, digital personas also based on acting conventions borrowed from cinema, to be followed by make-up, set design, and the narrative structures themselves. From one cultural language among others, cinema is becoming the cultural interface, a toolbox for all cultural communication, overtaking the printed word.

Cinema, the major cultural form of the twentieth century, has found a new life as the toolbox of a computer user. Cinematic means of perception, of connecting space and time, of representing human memory, thinking, and emotions become a way of work and a way of life for millions in the computer age. Cinema's aesthetic strategies have become basic organizational principles of computer software. The window in a fictional world of a cinematic narrative has become a window in a datascape. In short, what was cinema has become human-computer interface.

I will conclude this section by discussing a few artistic projects which, in different ways, offer alternatives to this trajectory. To summarize it once again, the trajectory involves gradual translation of elements and techniques of cinematic perception and language into a de-contextualized set of tools to be used as an interface to any data. In the process of this translation, cinematic perception is divorced from its original material embodiment (camera, film stock), as well as from the historical contexts of its formation. If in cinema the camera functioned as a material object, co-existing, spatially and temporally, with the world it was showing us, it has now become a set of abstract operations. The art projects described below refuse this separation of cinematic vision from the material world. They reunite perception and material reality by making the camera and what it records a part of a virtual world's ontology. They also refuse the universalization of cinematic vision by computer culture, which (just as post-modern visual culture in general) treats cinema as a toolbox, a set of "filters" which can be used to process any input. In contrast, each of these projects employs a unique cinematic strategy which has a specific relation to the particular virtual world it reveals to the user.

In The Invisible Shape of Things Past Joachim Sauter and Dirk Lüsenbrink of the Berlin-based Art+Com collective created a truly innovative cultural interface for accessing historical data about Berlin's history.[83] The interface de-virtualizes cinema, so to speak, by placing the records of cinematic vision back into their historical and material context. As the user navigates through a 3D model of Berlin, he or she comes across elongated shapes lying on city streets. These shapes, which the authors call "filmobjects", correspond to documentary footage recorded at the corresponding points in the city. To create each shape the original footage is digitized and the frames are stacked one after another in depth, with the original camera parameters determining the exact shape. The user can view the footage by clicking on the first frame. As the frames are displayed one after another, the shape is getting correspondingly thinner.

In following with the already noted general trend of computer culture towards spatialization of every cultural experience, this cultural interface spatializes time, representing it as a shape in a 3D space. This shape can be thought of as a book, with individual frames stacked one after another as book pages. The trajectory through time and space taken by a camera becomes a book to be read, page by page. The records of camera's vision become material objects, sharing the space with the material reality which gave rise to this vision. Cinema is solidified. This project, than, can be also understood as a virtual monument to cinema. The (virtual) shapes situated around the (virtual) city, remind us about the era when cinema was the defining form of cultural expression — as opposed to a toolbox for data retrieval and use, as it is becoming today in a computer.

Hungarian-born artist Tamás Waliczky openly refuses the default mode of vision imposed by computer software, that of the one-point linear perspective. Each of his computer animated films The Garden (1992), The Forest (1993) and The Way (1994) utilizes a particular perspectival system: a water-drop perspective in The Garden, a cylindrical perspective in The Forest and a reverse perspective in The Way. Working with computer programmers, the artist created custom-made 3D software to implement these perspectival systems. Each of the systems has an inherent relationship to the subject of a film in which it is used. In The Garden, its subject is the perspective of a small child, for whom the world does not yet have an objective existence. In The Forest, the mental trauma of emigration is transformed into the endless roaming of a camera through the forest which is actually just a set of transparent cylinders. Finally, in The Way, the selfsufficiency and isolation of a Western subject are conveyed by the use of a reverse perspective.

In Waliczky's films the camera and the world are made into a single whole, whereas in The Invisible Shape of Things Past the records of the camera are placed back into the world. Rather than simply subjecting his virtual worlds to different types of perspectival projection, Waliczky modified the spatial structure of the worlds themselves. In The Garden, a child playing in a garden becomes the center of the world; as he moves around, the actual geometry of all the objects around him is transformed, with objects getting bigger as he gets close to him. To create The Forest, a number of cylinders were placed inside each other, each cylinder mapped with a picture of a tree, repeated a number of times. In the film, we see a camera moving through this endless static forest in a complex spatial trajectory — but this is an illusion. In reality, the camera does move, but the architecture of the world is constantly changing as well, because each cylinder is rotating at its own speed. As a result, the world and its perception are fused together.

HCI: Representation versus Control

The development of human-computer interface, until recently, had little to do with distribution of cultural objects. Following some of the main applications from the 1940's until the early 1980's, when the current generation of GUI was developed and reached the mass market together with the rise of a PC (personal computer), we can list the most significant: real-time control of weapons and weapon systems; scientific simulation; computer-aided design; finally, office work with a secretary as a prototypical computer user, filing documents in a folder, emptying a trash can, creating and editing documents ("word processing"). Today, as the computer is starting to host very different applications for access and manipulation of cultural data and cultural experiences, their interfaces still rely on old metaphors and action grammars. Thus, cultural interfaces predictably use elements of a general-purpose HCI such as scrollable windows containing text and other data types, hierarchical menus, dialogue boxes, and command-line input. For instance, a typical "art collection" CD-ROM may try to recreate "the museum experience" by presenting a navigable 3D rendering of a museum space, while still resorting to hierarchical menus to allow the user to switch between different museum collections. Even in the case of The Invisible Shape of Things Past which uses a unique interface solution of "filmobjects" which is not directly traceable to either old cultural forms or general-purpose HCI, the designers are still relying on HCI convention in one case — the use of a pull-down menu to switch between different maps of Berlin.

In their important study of new media Remediation, Jay David Bolter and Richard Grusin define medium as "that which remediates."[84] In contrast to a modernist view aims to define the essential properties of every medium, Bolter and Grusin propose that all media work by "remediating," i.e. translating, refashioning, and reforming other media, both on the levels of content and form. If we are to think of human-computer interface as another media, its history and present development definitely fits this thesis. The history of human-computer interface is that of borrowing and reformulating, or, to use new media lingo, reformatting other media, both past and present: the printed page, film, television. But along with borrowing conventions of most other media and eclectically combining them together, HCI designers also heavily borrowed "conventions" of human-made physical environment, beginning with Macintosh use of desktop metaphor. And, more than an media before it, HCI is like a chameleon which keeps changing its appearance, responding to how computers are used in any given period. For instance, if in the 1970s the designers at Xerox Park modeled the first GUI on the office desk, because they imagined that the computer were designing will be used in the office, in the 1990s the primary use of computers as media access machine led to the borrowing of interfaces of already familiar media devices, such as VCR or audio CD player controls.

In general, cultural interfaces of the 1990's try to walk an uneasy path between the richness of control provided in general-purpose HCI and an "immersive" experience of traditional cultural objects such as books and movies. Modern general-purpose HCI, be it MAC OS, Windows or UNIX, allow their users to perform complex and detailed actions on computer data: get information about an object, copy it, move it to another location, change the way data is displayed, etc. In contrast, a conventional book or a film positions the user inside the imaginary universe whose structure is fixed by the author. Cultural interfaces attempt to mediate between these two fundamentally different and ultimately non-compatible approaches.

As an example, consider how cultural interfaces conceptualize the computer screen. If a general-purpose HCI clearly identifies to the user that certain objects can be acted on while others cannot (icons representing files but not the desktop itself), cultural interfaces typically hide the hyperlinks within a continuous representational field. (This technique was already so widely accepted by the 1990's that the designers of HTML offered it early on to the users by implementing the "imagemap" feature). The field can be a two-dimensional collage of different images, a mixture of representational elements and abstract textures, or a single image of a space such as a city street or a landscape. By trial and error, clicking all over the field, the user discovers that some parts of this field are hyperlinks. This concept of a screen combines two distinct pictorial conventions: the older Western tradition of pictorial illusionism in which a screen functions as a window into a virtual space, something for the viewer to look into but not to act upon; and the more recent convention of graphical human-computer interfaces which, by dividing the computer screen into a set of controls with clearly delineated functions, essentially treats it as a virtual instrument panel. As a result, the computer screen becomes a battlefield for a number of incompatible definitions: depth and surface, opaqueness and transparency, image as an illusionary space and image as an instrument for action.

The computer screen also functions both as a window into an illusionary space and as a flat surface carrying text labels and graphical icons. We can relate this to a similar understanding of a pictorial surface in the Dutch art of the seventeenth century, as analyzed by art historian Svetlana Alpers in her classical The Art of Describing. Alpers discusses how a Dutch painting of this period functioned as a combined map / picture, combining different kids of information and knowledge of the world.[85]

Here is another example of how cultural interfaces try to find a middle ground between the conventions of general-purpose HCI and the conventions of traditional cultural forms. Again we encounter tension and struggle — in this case, between standardization and originality. One of the main principles of modern HCI is consistency principle. It dictates that menus, icons, dialogue boxes and other interface elements should be the same in different applications. The user knows that every application will contain a "file" menu, or that if she encounters an icon which looks like a magnifying glass it can be used to zoom on documents. In contrast, modern culture (including its "post-modern" stage) stresses originality: every cultural object is supposed to be different from the rest, and if it is quoting other objects, these quotes have to be defined as such. Cultural interfaces try to accommodate both the demand for consistency and the demand for originality. Most of them contain the same set of interface elements with standard semantics, such as "home," "forward" and "backward" icons. But because every Web site and CD-ROM is striving to have its own distinct design, these elements are always designed differently from one product to the next. For instance, many games such as War Craft II (Blizzard Entertainment, 1996) and Dungeon Keeper give their icons a "historical" look consistent with the mood of an imaginary universe portrayed in the game.

The language of cultural interfaces is a hybrid. It is a strange, often awkward mix between the conventions of traditional cultural forms and the conventions of HCI — between an immersive environment and a set of controls; between standardization and originality. Cultural interfaces try to balance the concept of a surface in painting, photography, cinema, and the printed page as something to be looked at, glanced at, read, but always from some distance, without interfering with it, with the concept of the surface in a computer interface as a virtual control panel, similar to the control panel on a car, plane or any other complex machine.[86] Finally, on yet another level, the traditions of the printed word and of cinema also compete between themselves. One pulls the computer screen towards being dense and flat information surface, while another wants it to become a window into a virtual space.

To see that this hybrid language of the cultural interfaces of the 1990s represents only one historical possibility, consider a very different scenario. Potentially, cultural interfaces could completely rely on already existing metaphors and action grammars of a standard HCI, or, at least, rely on them much more than they actually do. They don't have to "dress up" HCI with custom icons and buttons, or hide links within images, or organize the information as a series of pages or a 3D environment. For instance, texts can be presented simply as files inside a directory, rather than as a set of pages connected by custom-designed icons. This strategy of using standard HCI to present cultural objects is encountered quite rarely. In fact, I am aware of only one project which uses it completely consciously, as a though through choice rather than by necessity : a CD-ROM by Gerald Van Der Kaap entitled BlindRom V.0.9. (Netherlands, 1993). The CD-ROM includes a standard-looking folder named "Blind Letter." Inside the folder there are a large number of text files. You don't have to learn yet another cultural interface, search for hyperlinks hidden in images or navigate through a 3D environment. Reading these files required simply opening them in standard Macintosh SimpleText, one by one. This simple technique works very well. Rather than distracting the user from experiencing the work, the computer interface becomes part and parcel of the work. Opening these files, I felt that I was in the presence of a new literary form for a new medium, perhaps the real medium of a computer — its interface.

As the examples analyzed here illustrate, cultural interfaces try to create their own language rather than simply using general-purpose HCI. In doing so, these interfaces try to negotiate between metaphors and ways of controlling a computer developed in HCI, and the conventions of more traditional cultural forms. Indeed, neither extreme is ultimately satisfactory by itself. It is one thing to use a computer to control a weapon or to analyze statistical data, and it is another to use it to represent cultural memories, values and experiences. The interfaces developed for a computer in its functions of a calculator, control mechanism or a communication device are not necessarily suitable for a computer playing the role of a cultural machine. Conversely, if we simply mimic the existing conventions of older cultural forms such as the printed word and cinema, we will not take advantage of all the new capacities offered by a computer: its flexibility in displaying and manipulating data, interactive control by the user, the ability to run simulations, etc.

Today the language of cultural interfaces is in its early stage, as was the language of cinema a hundred years ago. We don't know what the final result will be, or even if it will ever completely stabilize. Both the printed word and cinema eventually achieved stable forms which underwent little changes for long periods of time, in part because of the material investments in their means of production and distribution. Given that computer language is implemented in software, potentially it can keep on changing forever. But there is one thing we can be sure of. We are witnessing the emergence of a new cultural meta-langauge, something which will be at least as significant as the printed word and cinema before it.

Why would artists create an interactive work?

THEORETICAL PERSPECTIVE

The History of Electronic Music as a Reflection of Structural Paradigms

Joel Chadabe

ABSTRACT

The author proposes that new and similar paradigms appear in all fields simultaneously. He posits that the two most important developments in the history of electronic music were (1) the opening up of music to all sounds and (2) the development of interactive instruments. He discusses the first development at length, suggesting ways in which it reflected general paradigm shifts in other fields. He then points to ways in which the second development may also reflect broad paradigm shifts, and concludes by speculating about the direction that such paradigmatic changes may be taking in our time.

At the time of writing this article, I am completing a book on the history of electronic music. The book narrates stories of early instruments, tape music studios, early electronic music performance groups, computer music, synthesizers and MIDI, and it speculates on the concepts involved in the development of electronic musical instruments. Except in certain fleeting moments, the book is not interdisciplinary. This article is an expansion of some of those fleeting moments.

I have been noticing for quite a while that similar major innovations in thinking seem to occur in different fields more or less simultaneously. This observation has led me to conclude that major innovations in any particular field are related to generally new perceptions of how things happen in the world and how things should happen. New structural paradigms, in other words—at least within the scope of what we know as Western cultural history—seem to develop ubiquitously. At any moment they seem to be in the air, to be whiffed by different members of the scientific and cultural vanguard. Scientists formulate new paradigms in new theories, painters show them in new imagery and composers play them in new music. In short, new structural paradigms do not flow from one field to another: it is not true, as the old adage suggests, that the sciences discover and the arts express. Ives did not read Einstein before he composed *The Unanswered Question*. New structural paradigms happen in every field at the same time.

In my view, the two most important developments in the history of electronic music were (1) the opening up of music to all sounds and (2) the development of interactive instruments, and both developments have reflected general shifts from old to new paradigms. The opening up of music to all sound reflected the shift from Newton's concept of Absolute Time to Einstein's concept of relativity as described in the Special Theory in 1905. The development of musical automata as the basis for algorithmic composition and interactive instruments reflected a paradigm shift from determinism to indeterminacy.

The Opening Up of Music to All Sound

An orrery is a tabletop model of the Newtonian solar system. Miniature planets and moons are connected to a crank via gears and mechanical arms, and as the crank is turned, the planets and moons revolve at their different speeds and in their different cycles around a stationary sun in the middle. Turning the crank is a metaphor for the passage of Newtonian time. All of the motions of the planets and moons are synchronized to it, and the speed with which time passes is defined by the speed at which the crank is turned. In his *Principia* (1687), Newton called that singular line of time Absolute Time: "Absolute, True, and Mathematical Time, of itself, and from its own nature flows. . . . All motions may be accelerated and retarded, but the True, or equable progress, of Absolute Time is liable to no change" [1].

The music composed during the Newtonian period, between 1600 and 1900, reflected the idea of a universe in which all motions were synchronized to a single line of time. In all tonal compositions, there was one line of chord progressions to which all notes, of whatever rhythmic values, were synchronized. The painting of the period also reflected Newton's concept of one line of time. Our eyes traveled along one line of perspective to the horizon, and all objects were synchronized to it.

The most important implication of Newton's universe, however, was that it was a universe scaled to human capability and common sense. It was a direct extension of what humans could perceive, understand and do with the simplest technology. It was common sense, understandable in simple human terms, that there was one line of time. How could we make appointments, for example, if our watches were not set to the same time line? It was an extension of everyday life that the painting of the era represented the world as it was seen, with objects getting smaller as they receded into the distance. It was a normal human activity to dance, march and clap to music with a discernible beat. The flow of tonal music, with its rises and falls in pitch and loudness supported by the tension and release of harmonic progressions, was understood as running parallel to the flow of human emotion: human voices also rise or fall in pitch and become louder or softer with emotional change. The music of the period from 1600 to 1900, especially as compared to the ornate music of the Renaissance, involved and affected people. It connected. And Striggio's words at the opening to Monteverdi's *Orfeo* (1607) signified the passage from the complex counterpoint of the Renaissance to the emotionally involving romanticism of the early Baroque. Striggio wrote:

> *Io la musica son, ch'ai dolci accenti*
> *so far tranquillo ogni turbato core,*
> *ed or di nobil ira ed or d'amore*
> *posso infiammar le piùgelate menti.*

Joel Chadabe (composer), P.O. Box 8748, Albany, NY 12208, U.S.A. E-mail: <jchadabe@logical.net>.

I am music, that with sweet accents
I know how to make tranquil each
turbulent heart,
and now with noble anger and now
with love
I can inflame the most frigid minds [2].

At the beginning of the twentieth century, the paradigm changed. It became clear that nature extended beyond the human scale to include items—atoms, light quanta, stars, galaxies—that were smaller or larger than those that could be seen or verified by the naked eye or common sense. In the Special Theory, Einstein went beyond all notions of human-scale perception and common sense with his descriptions of a universe in which time passed differently for every object according to the speed, relative to the speeds of other objects, with which it moved through space. The point of the Special Theory was that the faster something moved, the more slowly its time passed. Einstein's theory, in short, described a universe that was a multiplicity of parallel and unsynchronized time lines.

The idea of a universe based on a multiplicity of parallel time lines was, not surprisingly, reflected at about the same time in the arts. Analytic cubism, for example, saw the world through a prism of multiple perspectives. Any of Picasso's cubist paintings of the period would serve as an example. To choose one, consider *The Reservoir, Horta,* which was painted during a visit to Spain in the summer of 1909. This painting shows the natural angles of the rooftops and walls as distorted, so that they flow from one to the other and to the background in an overlapping, interlaced texture of planes, giving the impression that the scene is seen simultaneously from every direction. This concept was also reflected in the literary vanguard. In "Lundi Rue Christine" (1913), Guillaume Apollinaire plucked phrases from a multiplicity of parallel stories:

Trois becs de gaz allumés
La Patronne est poitrinaire
Quand tu auras fini nous jouerons une partie de jacquet
Un chef d'orchestre qui a mal à la gorge
Quand tu viendras à Tunis je te ferai fumer du kief
Ça a l'air de rimer

Three lit gas jets
The proprietress has bad lungs
When you've finished we'll play a
hand of backgammon
An orchestra conductor who has a
sore throat
When you come to Tunis I'll give you
some kef to smoke
This seems to rhyme [3]

The idea of parallel realities was also reflected in music. Debussy, Stravinsky and Ives combined bits and fragments of chords and rhythms as if they were "sampling" multiple streams of simultaneously occurring tonal activities. Upon hearing *La Soirée dans Grenade* (1903), for example, one might imagine Debussy standing in the center square of Grenada, turning his "microphone" to the left to catch a group of musicians playing a *habanera* as they walk down a hill, then turning to the right to hear a guitarist in a cafe, then turning again to catch another group approaching from another direction. One can point to the quick cuts and superimpositions in the first tableau in Stravinsky's *Petrushka* (1911), to the Petrushka chord itself—a combination of F# and C major—and to the discrete layers of music that move in dissimilar cycles in the "Wet Nurse's Dance." One could mention the superimposed processes in Ives's *The Unanswered Question* (1908), where a dialogue between an inquiring trumpet and answering woodwinds is juxtaposed with a sober, hymn-like background.

Further, the artists who perceived the world as a multiplicity of parallel processes also perceived that those processes were resident in a multiplicity of materials. Not only did they pluck moments of imagery from normal storylines, they plucked materials from their normal habitats. Indeed, from Pablo Picasso's *papiers collés* in the early 1900s to Robert Rauschenberg's combines in the mid-1900s, no matter what the specific artistic style or intent, the use of found objects in the arts became normal. This liberation of materials, wherein materials could be disassociated from their normal habitats and recombined in new settings, led to a different kind of artistic expression. No longer were artists poetic souls expressing themselves. Science, art and music became disconnected from human capability, physiology and common sense. They now conveyed a picture of the world as a collage of unrelated and juxtaposed objects.

Sounds were no longer connected to instruments, human expression or physical activity. As Gertrude Stein might have put it, a sound was a sound was a sound was a sound. Sounds became objects—*objets sonores,* as Pierre Schaeffer was to call them—for further investigation and manipulation. As synthesis and processing tools were developed, the ability to create sounds became, for many composers, the fundamental reason to engage in electronic music. Referring to the use of computers, Jonathan Harvey said, "Before the microscope, we never knew what the microworld looked like—and now, because of the tremendous precision in being able to look into sounds and work with them, the whole world of microsound has opened up and we can compose with it" [4]. Denis Smalley stated, "My musical ideas come out of the sounds themselves" [5]. Trevor Wishart observed, "You can now treat sound in the same logical way that we treated pitch before" [6]. Innumerable other composers have expressed related thoughts.

INDETERMINACY

The orrery also demonstrates Newtonian determinism in action. As we turn the crank, A, the present state of things, flows seamlessly to B, another state of things. A also flows to B in painting based on perspective, where the eye follows from one object to the next along structural lines. And A flows to B in tonal music, as one chord flows into the next. Determinism makes sense. It allows us to see our lives flowing smoothly from present to future, with the future positions of our lives following clearly from our present efforts. It allows us to believe that good is rewarded and evil punished. Determinism is comforting and satisfying, as it reinforces concepts of order, control and justice.

It also provides structure. We can turn the orrery's crank backward as easily as forward. From the vantage point of the present, we can look toward past and future with equal certainty. Time, in the context of Newtonian determinism, is symmetrical. As Norbert Wiener put it, "The music of the spheres is a palindrome" [7]. Indeed, at the level of harmonic structure, every tonal composition is a palindrome, a two-part form that leads away from a tonic chord at the beginning and back to it at the end. Those chords define the boundaries of the composition, the closure of the form at both ends, the whole. When the whole is known as the parts are being made, the parts can be made to fit. And structure, which is the division of a whole into parts, becomes possible. When structure becomes possible, the classical values of proportion and balance become possible. With an adequate sense of structure, we can see the whole of our lives—the big picture—and understand how we fit and where we belong. We can balance our activities and put our actions in proportion.

When one seeks to establish order, control and justice, words such as "chance," "randomness" and "indeterminacy" can be disturbing. When there is no continuity between past and future and no just desserts, basic values may come into doubt. When time and musical form move toward an unpredictable future, a musical composition as a whole can be known only retrospectively, after the parts have been made. Structure, proportion and balance cannot exist. It is true that indeterministic music lacks these qualities. Someone once asked John Cage, "Why doesn't your music have any structure?" and Cage replied: "My music is a process. Like the weather" [8].

In their formulation during the 1920s of what came to be known as the Copenhagen Interpretation of Quantum Theory, Niels Bohr and Werner Heisenberg presented a picture of an essential randomness in the behavior of subatomic particles. Following Heisenberg's 1926 statement of his Uncertainty Principle, which showed that there is no law that connects values for the position and momentum of a subatomic particle at one point in time with values for those properties at another point in time, Einstein objected. Underlying Einstein's often-quoted statement that God does not play dice with the world was his belief that the laws of chance represent an irrationality that is impossible in nature. In 1948, Max Born, Einstein's friend and colleague, expressed what many scientists felt at the time when he wrote: ". . . when out of his own work a synthesis of statistical and quantum principles emerged which seemed acceptable to almost all physicists he kept himself aloof and skeptical. Many of us regard this as a tragedy" [9].

However, no matter how offensive to classical values and how disturbing it may seem, indeterminacy does exist in the world; and it is interesting to note that in nature, as in music, chance often serves to produce interesting results. As molecular biologist and Nobel laureate Jacques Monod pointed out:

> Life appeared on earth: what, *before the event*, were the chances that this would occur? The present structure of the biosphere far from excludes the possibility that the decisive event occurred *only once*.. . . Our number came up in the Monte Carlo game [10].

Responding to what he observed as a human tendency to see purpose in all things, Monod continued: "Destiny is written concurrently with the event, not prior to it. Our own was not written before the emergence of the human species" [11]. He went further to point out that evolution was affected by chance: reproductive invariance is determined by information coded in DNA, yet generational change is determined by random perturbations in the DNA sequence that become, however, part of its invariant message in the succeeding generation. In other words, random change in the DNA sequence is captured in offspring and passed on to succeeding generations. In Monod's words:

> And so one may say that the same source of fortuitous perturbations, of "noise" . . . is the progenitor of evolution in the biosphere and accounts for its unrestricted liberty of creation, thanks to the replicative structure of DNA: that registry of chance, that tone-deaf conservatory where the noise is preserved along with the music [12].

The basic question, then, for a composer of electronic music is: Does a composer view a composition as an object, with its sound and structure carefully determined? If so, that composer will need to control the process of composition at every level. On the other hand, if a composer views a composition as an interactive process that can take many forms in performance depending upon who is performing it, that composer will need to accept some level of indeterminacy in leaving certain aspects of the composition open for the performer to compose. The surprises of indeterminacy are often rewarding. Of a performance of his *Untitled,* David Tudor once recalled: "It was so unpredictable, it was just wonderful" [13].

Indeterminacy is the heartbeat of the interactive system. The surprises produced by the system put its human performer in the position of a conversationalist interacting with a clever friend, giving the performer cues for further action. As Bruno Spoerri put it, referring to his interactive performance systems, "The important thing for me was to have a partner in the computer who threw balls at me, who gave me a reason to react in a certain way" [14]. George Lewis said, "I try to get the computer to do its own thing as well as follow a performer. As soon as the computer generates something independent, a performer can react to that and go with it" [15].

CONCLUSION

It is, of course, simplistic to reduce our world views to a single question of whether we prefer to plan or improvise, and it is equally simplistic to reduce our musical preferences to the question of whether we like electronic sounds. We all prefer specific types of sounds, regardless of how they are made, and we all prefer to plan and improvise in different balances at different times. New and old paradigms coexist in our minds. It is not just paradigms, but also our educational and cultural histories that determine our actions and reactions.

But it is also true that, as technology expands our powers to perceive and act, we develop new paradigms to understand what it is that we see and how we should react to it. This involves something of a chicken-and-egg process in which new paradigms lead us to new tools and new tools lead us to new paradigms. Many people, at least to some extent, abandon old paradigms when they are no longer useful. For many people, it is neither realistic nor useful to limit music to tonality and to the range of sounds that can be played by acoustic instruments.

In my view, determinism is no longer a useful paradigm. The assertion that A leads to B requires a simplification that eliminates all contributory causalities. Simply reading the daily newspapers is enough to lead one to conclude that no one cause leads to any one effect, but that everything results from an underlying complexity of causes. Even the seemingly simple cause-and-effect sequence of throwing a switch and thus turning on a light is, in fact, dependent upon an underlying support and delivery system that sometimes fails and produces blackouts. The expected, after all, is often boring.

Does indeterminacy remain a useful paradigm? Perhaps not. It was certainly useful to explain the surprises that resulted from underlying complexity. But once understood, it became an old paradigm. Newer paradigms of the 1960s dealt with issues of information processing and control, as in the cybernetic model, and controllable complexities, as in the general systems model. And even those paradigms have evolved. Given, then, an awareness of the underlying complexities in the universe and their potential for producing surprises (indeterminacy), the issues involved in processing information and controlling events (cybernetics) and the web-like multidirectional causalities within a system (general system theory), I would propose that today's prevailing paradigm is interaction. In fact, the ubiquity of the word "interactive" already indicates that interaction is fast becoming a widely accepted and, consequently, old paradigm.

Interaction means mutual influence. In environmental terms, it means that we influence changes in the environment and react to environmental events. In musical terms, it means that we influence the instrument that we play and that we are influenced by the sounds that it produces. It means that an instrument has a mind of its own, so to speak, such that it produces musical information that contains surprises. The first interactive instruments were developed around 1970, concurrently but independently by Salvatore Martirano and me. Since then, the paradigm has matured. In an article written in 1984, I referred to the interaction paradigm as "interactive composing" and concluded:

> The ultimate significance of interactive composing is that it represents a new way for composers and performers to participate in a musical activity. I offer my nontechnical perception that good things often happen—in work, in romance, and in other aspects of life—as the result of a successful interaction during opportunities presented as if by chance; to that I would add only that it seems to me reasonable that such a perception should also find expression in music [16].

The interaction paradigm is probably due for replacement. One paradigm builds on another. My guess is that, as musical control systems become increasingly complex, a strategies paradigm will be developed to deal not only with our immediate interactions with a system, but also with the implications of those interactions in terms of the achievement of a particular long-term goal. Whatever the next paradigm will be, however, it is now gradually emerging in all fields, and I, for one, am looking forward to hearing its manifestations in new music.

References and Notes

1. Cited in Max Born, *Einstein's Theory of Relativity* (New York: Dover, 1962) p. 57.

2. Alessandro Striggio, Prologue, libretto to Claudio Monteverdi, *L'Orfeo* (1607) (translation mine).

3. "Lundi Rue Christine" was first published in *Soirées de Paris* **2**, No. 19 (1913) p. 27. Quoted in William C. Seitz, *The Art of Assemblage* (New York: Museum of Modern Art, 1961) p. 15 (translation mine).

4. Jonathan Harvey, personal communication, 12 May 1994.

5. Denis Smalley, personal communication, 18 May 1994.

6. Trevor Wishart, personal communication, 17 September 1994.

7. Norbert Wiener, *Cybernetics* (Cambridge, MA: MIT Press, 1961) p. 31.

8. Cage made this remark at a public forum conducted by Lukas Foss following a production of John Cage and Lejaren Hiller's *HPSCHD* at the Brooklyn Academy of Music in 1974. The question was asked by someone in the audience. I had been artistic director of the production and was onstage as part of the forum.

9. Max Born, "Einstein's Statistical Theories," in Paul Arthur Schilpp, ed., *Albert Einstein: Philosopher-Scientist* (New York: Harper & Row, 1959) pp. 163–164.

10. Jacques Monod, *Chance and Necessity* (New York: Vintage, 1972) p. 144.

11. Monod [10] p. 145.

12. Monod [10] pp. 116–117.

13. David Tudor, personal communication, 8 September 1993.

14. Bruno Spoerri, personal communication, 21 February 1994.

15. George Lewis, personal communication, 11 December 1993.

16. Joel Chadabe, "Interactive Composing: An Overview," *Computer Music Journal* **8**, No. 1, 22–27 (1984); reprinted in Curtis Roads, ed., *The Music Machine* (Cambridge, MA: MIT Press, 1989) pp. 143–148.

Manuscript received 23 February 1996.

Interactive

Interactive is defined by the Oxford English Dictionary as "reciprocally active; acting upon or influencing each other." Although such interactivity may result in the reciprocal engagement or interaction of any two objects or beings, it was only with the addition of new technologies that the term came to be associated with primarily human–machine relations. A second definition listed by the OED presents interactive as "pertaining to or being a computer or other electronic device that allows a two-way flow of information between it and a user, responding immediately to the latter's input." A new genre of interaction has developed through participation with increasingly subjective and semiautonomous technological devices.

Television broadcasting exemplifies this mode of interactivity, or at least the beginning of its simulation. Signing off of the air for the night, the anchor says to his audience, "We'll see you tomorrow." This phrase implies a relationship with the viewer, a connection between the medium of the television and the person watching at home. In this way, it is bridging two realms: that of reality with that of televised media. In inducting the viewer into this on-screen world, "interactivity" acts as a "kind of 'suture' between ourselves and our machines." Representation comes to stand for a new reality into which the viewer becomes absorbed. Guy Debord, writing for a society newly saturated with televised and cinematic media, labeled this emerging relationship as the spectacle that results from such representation. "Where the real world changes into simple images, the simple images become real beings and effective motivations of hypnotic behavior." In more modern technologies, interactivity has evolved to consist not only of a reaching between spaces, but an immersion and affective engagement within another world or system entirely.

Interactions with these technologies allow for active user control. In her book *Virtualities: Television, Media Art, and Cyberculture*, Margaret Morse identifies interactivity as a "means of allowing the consumer/viewer to select or change the image with the help of an input device—telephone, keyboard, remote control, joystick, mouse, touch-screen, brain wave reader, et cetera." It is the ability of the user to participate in the creation or modification of a medium. Marshall McLuhan tracks the emergence of this new interaction through his explanation of 'hot' and 'cool' media. While cool media encourage the interaction of their users, "hot media do not leave so much to be filled in or completed by the audience. Hot media are, therefore, low in participation, and cool media are high in participation or completion by the audience." Hot and cool media do not necessarily have to be mechanistic, but among cool media is the telephone, one apparatus in a series of innovations that have pushed our society towards what McLuhan perceives of as a new age. "In terms of the reversal of procedures and values in the electric age, the past mechanical time was hot, and we of the TV age are cool." He describes an increased level of collaboration, or interaction, with electric media. For McLuhan, interactive media are participatory.

Increased interaction with these media introduces a new societal space, that of the virtual. A community of interactivity is one that influences individual space and time. Morse states that technologies "employ various forms of engagement to construct a *virtual relationship* between subjects in a here-and-now… The interactive user is an I or a player in discursive space and time."

This is perhaps best demonstrated through Web technology. The Internet immerses its users in an environment of abstracted space in which interactions are enacted through the click of a button. "Whether we call the noplace in which exchanges on electronic networks occur or the scene of an immersive computer graphic 'world' a virtual environment, artificial reality, or *cyberspace*, the gathering of places and sites of experience in electronic culture are increasingly situated in what amounts to *nonspace* and in which humans not only interact with human agents but also with the semiautonomous agency of the machines." The growing popularity of computer and video games, which invite viewers to interact with an entire virtual world through a character that they themselves create, is the most recent reflection of this trend.

Interactions in virtual space transcend traditional physical boundaries, and this is why the way that people interact online is often through revolutionized or liberated modes of expression. Chatrooms, e-zines, wiki pages, and networking sites such as Facebook and MySpace have produced arenas for new behavior and self-expression. Lelia Green views these types of engagement, from web collaboration to cybersex, as indicative of an entirely new cultural regulation, a new set of rules that redefine the type and scale of acceptable interaction. As interactive media become more complex and engaging, the way in which users interact with each other evolves along with their interactions with the medium itself.

Still, real societal potential is recognized in these 'no-place' interactions. It is believed by some that interaction within the virtual might be translated to interaction within reality, and thus computer-mediated relations are presented as "democracy's salvation." "Indeed," Green notes, "the Internet offers the opportunity for creative and experiential psychological interconnection with others unrivaled by traditional mass media in either the local or global context." Holmes classifies this as a "community of interactivity" that allows for interconnection and collaboration on the most global scale possible. While the regulations of the real public sphere once molded interactions of the virtual, the new modes of engagement popularized in virtual media have begun to be seen as models for reality.

Jodi Dean, however, sees only a fantasy [link] of social unity in these interactions. After all, "virtual reality in itself is a rather miserable idea: that of imitating reality, of reproducing its experience in an artificial medium." The World Wide Web is only a "deluge of screens and spectacles" that undermines the opportunity for democratic reality, a modern incarnation of the society described by Debord. Although Internet technology might be a source of democratic potential, Dean argues that it is "a mistaken notion that the Web is a public sphere." Instead, the proliferation of interactivity only leads to an obsession with this availability of computer-mediated, global information. "Enthralled by transparency" itself, no inspiring action is taken beyond the virtual level.

Nevertheless, the last two decades of art production have capitalized on this potential for interactivity promoted through virtual interface. "In observing contemporary artistic practices, we ought to talk of 'formations' rather that 'forms'. Unlike an object that is closed in on itself by the intervention of a style and a signature, present-day art shows that form only exists in the encounter and in the dynamic relationship enjoyed by an artistic proposition with other formations, artistic or otherwise." If internet interactivity represents an extension of the social arena and the potential for further democratization, then interactive art has become a means for creating and furthering both local and global relations.

Termed ‘relational aesthetics’ by the French art critic Nicolas Bourriaud and ‘dialogical art’ by Grant Kester in his book *Conversation Pieces: Community and Communication in Modern Art*, these projects “unfold through a process of performative interaction.”

Rikrit Tiravanija’s *Untitled 1992 (Free)* exemplifies the increasing interactivity of art exhibitions through its recreation of the gallery as a new site of interaction. In his work, which consisted of the makings for a provisional kitchen, Tiravanija forced participation upon the viewer by inviting visitors to cook and eat Thai noodles. Here, the medium of the artwork itself became the interactions taking place, first of the visitor with the Thai noodles, and later, the socialization with others that resulted from the unexpected sharing of this space. Tiravanija’s work, and the work of other artists who invite similar interactions with their artwork (Felix Gonzalez-Torres, Sophie Calle, Pierre Hyughe), represent a shift in participation and viewer engagement. In current artistic production, “meaning and sense are the outcome of an interaction between artist and beholder, and not an authoritarian fact. In modern art [the beholder must] make an effort to produce sense out of objects that are even lighter, ever more impalpable and even more volatile. Where the decorum of the picture used to offer a frame and a format, we must now often be content with bits and pieces. Feeling nothing means not making enough effort.” In Bourriaud’s opinion, viewer interaction has become the medium of the contemporary work of art.

These sites for interactive art and the communities they create are alternatively called platforms or stations, terms that evoke the electronic network and computer technology. “Though the means applied to this end are usually far funkier and more face-to-face than any chat room on the web,” the sociability of artwork remains closely linked to internet rhetoric.

Increasingly interactive media allow for more affective engagement with both technology and, in the case of the Internet, fellow users. While the efficacy and sincerity of such relationships continues to be debated, Bourriaud presents personal interactions as the solution to Debord’s Society of the Spectacle, which is “the opposite of dialogue.” The spectacle escapes the activity, reconsideration, and correction of men, but Bourriaud argues that it can be “analyzed and fought through the production of new types of relationships between people.” The interactivity presented by and mediated through new technologies—television, the World Wide Web, computer and video games—do not precisely mirror face-to-face relationships. Yet while the practices of one arena of interaction do not always translate directly to another, there is potential, as in all reciprocal relationships, for revision and transformation through active engagement.

Caitlin Rubin

The Chicago School of Media Theory

The Myth of Interactivity

Lev Manovich

We have only one principle still remaining from the original list: interactivity.

As with *digital*, I avoid using the word *interactive* in this book without qualifying it, for the same reason -- I find the concept to be too broad to be truly useful.

Used in relation to computer-based media, the concept of interactivity is a tautology. Modern human-computer interface (HCI) is by its very definition interactive. In contrast to earlier interfaces such as batch processing, modern HCI allows the user to control the computer in real-time by manipulating information displayed on the screen. Once an object is represented in a computer, it automatically becomes interactive. Therefore, to call computer media interactive is meaningless -- it simply means stating the most basic fact about computers.

Rather than evoking this concept by itself, in this book I use a number of other concepts, such as menu-based interactivity, salability, simulation, image- interface, and image-instrument, to describe different kinds of interactive structures and operations. The already used distinction between "closed" and "open" interactivity is just one example of this approach.

Although it is relatively easy to specify different interactive structures used in new media object, it is much more difficult to theoretically deal with user experiences of these structures. This remains to be one of the most difficult theoretical questions raised by new media. Without pretending to have a complete answer, I would like to address some aspects of this question here.

All classical, and even more so modern art, was already "interactive" in a number of ways. Ellipses in literary narration, missing details of objects in visual art and other representational "shortcuts" required the user to fill-in the missing information. Theater, painting and cinema also relied on the techniques of staging, composition and cinematography to orchestrate viewer's attention over time, requiring her to focus on different parts of the display. With sculpture and architecture, the viewer had to move her whole body to experience the spatial structure.

Modern media and art pushed each of these techniques further, putting new cognitive and physical demands on the viewer. Beginning in the 1920s new narrative techniques such as film montage forced the audiences to quickly bridge mental gaps between unrelated images. New representational style of semi- abstraction which, along with photography, became the "international style" of modern visual culture, required the viewer to reconstruct the represented objects from the bare minimum -- a contour, few patches of color, shadows cast by the objects not represented directly. Finally, in the 1960s, continuing where Futurism and Dada left of, new forms of art such as happenings, performance and installation turned art explicitly participational. This, according to some new media theorists, prepared the ground for interactive computer installations which appeared in the 1980s.

When we use the concept of "interactive media" exclusively in relation to computer-based media, there is danger that we interpret "interaction" literally, equating it with physical interaction between a user and a media object (pressing a button, choosing a link, moving the body), at the sake of psychological interaction. The psychological processes of filling-in, hypothesis forming, recall and identification, which are required for us to comprehend any text or image at all, are mistakenly identified with an objectively existing structure of interactive links.

This mistake is not new; on the contrary, it is a structural feature of history of modern media. The literal interpretation of interactivity is just the latest example of a larger modern trend to externalize of mental life, the process in which media technologies -- photography, film, VR -- have played a key role. Beginning in the nineteenth century, we witness recurrent claims by the users and theorists of new media technologies, from Francis Galton (the inventor of composite photography in the 1870s) to Hugo Munsterberg, Sergei Eisenstein and, recently, Jaron Lanier, that these technologies externalize and objectify the mind. Galton not only claimed that "the ideal faces obtained by the method of composite portraiture appear to have a great deal in common with...so-called abstract ideas" but in fact he proposed to rename abstract ideas "cumulative ideas." According to M« nsterberg, who was a Professor of Psychology at Harvard University and an author of one of the earliest theoretical treatments of cinema entitled The Film: A Psychological Study (1916), the essence of films lies in its ability to reproduce, or "objectify" various mental functions on the screen: "The photoplay obeys the laws of the mind rather than those of the outer world." In the 1920s Eisenstein was speculating about how film can be used to externalize — and control — thinking. As an experiment in this direction, he boldly conceived a screen adaptation of Marx's Capital. "The content of CAPITAL (its aim) is now formulated: to teach the worker to think dialectically," Eisenstein writes enthusiastically in April of 1928. In accordance with the principles of "Marxist dialectics" as canonized by the official Soviet philosophy, Eisenstein planned to present the viewer with the visual equivalents of thesis and anti-thesis so that the viewer can then proceed to arrive at synthesis, i.e. the correct conclusion, pre-programmed by Eisenstein.

In the 1980s, Jaron Lanier, a California guru of VR, similarly saw VR technology as capable of completely objectifying, better yet, transparently merging with mental processes. His descriptions of its capabilities did not distinguish between internal mental functions, events and processes, and externally presented images. This is how, according to Lanier, VR can take over human memory: "You can play back your memory through time and classify your memories in various ways. You'd be able to run back through the experiential places you've been in order to be able to find people, tools." Lanier also claimed that VR will lead to the age of "post-symbolic communication," communication without language or any other symbols. Indeed, why should there be any need for linguistic symbols, if everybody, rather than being locked into a "prison-house of language" (Fredric Jameson), will happily live in the ultimate nightmare of democracy -- the single mental space which is shared by everybody, and where every communicative act is always ideal (Jurgen Habermas[50]). This is Lanier's example of how post-symbolic communication will function: "you can make a cup that someone else can pick when there wasn't a cup before, without having to use a picture of the word "cup."[51] Here, as with the earlier technology of film, the fantasy of objectifying and augmenting consciousness, extending the powers of reason, goes hand in hand with the desire to see in technology a return to the primitive happy age of pre-language, pre-misunderstanding. Locked in virtual reality caves, with language taken away, we will communicate through gestures, body movements, and grimaces, like our primitive ancestors...

The recurrent claims that new media technologies externalize and objectify reasoning, and that they can be used to augment or control it, are based on the assumption of the isomorphism of mental representations and operations with external visual effects such as dissolves, composite images, and edited sequences. This assumption is shared not just by modern media inventors, artists and critics but also by modern psychologists. Modern psychological theories of the mind, from Freud to cognitive psychology, repeatedly equate mental processes with external, technologically generated visual forms. Thus Freud in The Interpretation of Dreams (1900) compared the process of condensation with one of Francis Galton's procedures which became especially famous: making family portraits by overlaying a different negative image for each member of the family and then making a single print.52 Writing in the same decade, the American psychologist Edward Titchener opened the discussion of the nature of abstract ideas in his textbook of psychology by noting that "the suggestion has been made that an abstract idea is a sort of composite photograph, a mental picture which results from the superimposition of many particular perceptions or ideas, and which therefore shows the common elements distinct and the individual elements blurred." He then proceeds to consider the pros and cons of this view. We should not wonder why Titchener, Freud and other psychologists take the comparison for granted rather than presenting it as a simple metaphor -- contemporary cognitive psychologists also do not question why their models of the mind are so similar to the computer workstations on which they are constructed. The linguist George Lakoff asserted that "natural reasoning makes use of at least some unconscious and automatic image-based processes such as superimposing images, scanning them, focusing on part of them" while the psychologist Philip Johnson-Laird proposed that logical reasoning is a matter of scanning visual models. Such notions would have been impossible before the emergence of television and computer graphics. These visual technologies made operations on images such as scanning, focusing, and superimposition seem natural.

What to make of this modern desire to externalize the mind? It can be related to the demand of modern mass society for standardization. The subjects have to be standardized, and the means by which they are standardized need to be standardized as well. Hence the objectification of internal, private mental processes, and their equation with external visual forms which can be easily manipulated, mass produced, and standardized on its own. The private and individual is translated into the public and becomes regulated.

What before was a mental process, a uniquely individual state, now became part of a public sphere. Unobservable and interior processes and representations were taken out of individual heads and put outside -- as drawings, photographs and other visual forms. Now they could be discussed in public, employed in teaching and propaganda, standardized, and mass-distributed. What was private became public. What was unique became mass-produced. What was hidden in an individual's mind became shared.

Interactive computer media perfectly fits this trend to externalize and objectify mind’s operations. The very principle of hyperlinking, which forms the basis of much of interactive media, objectifies the process of association often taken to be central to human thinking. Mental processes of reflection, problem solving, recall and association are externalized, equated with following a link, moving to a new page, choosing a new image, or a new scene. Before we would look at an image and mentally follow our own private associations to other images. Now interactive computer media asks us instead to click on an image in order to go to another image. Before we would read a sentence of a story or a line of a poem and think of other lines, images, memories. Now interactive media asks us to click on a highlighted sentences to go to another sentence. In short, we are asked to follow pre-programmed, objectively existing associations. Put diffidently, in what can be read as a new updated version of French philosopher Louis Althusser's concept of "interpellation," we are asked to mistake the structure of somebody's else mind for our own.

This is a new kind of identification appropriate for the information age of cognitive labor. The cultural technologies of an industrial society -- cinema and fashion -- asked us to identify with somebody's bodily image. The interactive media asks us to identify with somebody's else mental structure. If a cinema viewer, both male and female was lasting after and trying to emulate the body of movie star, a computer user is asked to follow the mental trajectory of a new media designer.

Healing Interactions and Interactive Digital Art
by Barbara Buckner
Afterimage, Nov-Dec, 2001

Interactive media has synthesized the traditional art forms of writing, music, painting, sculpture and the moving image. CD-ROM technology and the Web can incorporate 2D and 3D graphics, movies, animations, text and sound. An added tool to this mix has redefined the notion of artistic experience--user interaction, the user's power of choice to co-create a digital artwork within artist-defined parameters.

User interaction includes clicking, dragging and dropping and roll-overs to modify on-screen elements. Physically interacting with on-screen-elements is an empowering act, because the user becomes cause in the scenario. In a digital interactive artwork, the user engages in an artistic feedback loop to transform artistic elements and relationships-- changes in color, shape, texture, text, iconic and textual menus, animation, music, sound, voice and screen transitions. The user makes choices about these over time, creating ever new and updated information configurations and perceptions. The user builds an iterative cycle of aesthetic relationships, and through this cycle of perception and co-creation, fulfills the meaning of the work.

The power of choice in an interactive artwork is like a brush to a painter, chisel to a sculptor, pen to a writer or piano for a composer. The act of choosing allows for co-creation with the artist. I am not referring to video games or financial software, which are essentially about winning and entering/tracking of data. Interactive digital art differs from these in that it requires a series of transformations relying on the physical, contemplative and perceptual powers of the individual to unfold an aesthetic code in time and space through choice. If the artist creates a code or syntax (a group of navigational and interaction "rules") by which the user relates aesthetic entities (shape, color, sound, text) through choice, the experience can be a healing journey because the user is accepting responsibility for co-creating unity through cause and effect.

Note I say unity. A Dada-like approach to interactive art, where the user is free to click on pretty much anything and link to pretty much anywhere within the work, lacks a holistic creative vision and forces a kind of chaos on the user. The artist's role is the same as it ever was--to create clear boundaries (the aesthetic idea) realized through form and content where one finds a greater freedom and wholeness by interacting with an artistic vision, or metaphorical "play."

Someone who has knowledge of religious or spiritual mysteries is sometimes called a hierophant. This person has an ability to perceive the workings of inner forces behind physical realities. An artist-created code used to create an interactive digital artwork is like a hierophant. The hierophant can be an image, sound or text and remains in a frozen state until activated and initiated by the user. Once initiated (into the aesthetic mystery) by the user, the code moves invisible forces and qualities to modify physical onscreen elements. Each object, word or sound is in a sense sacred and can be used as a starting point for revelation of an aesthetic whole. In a digital interactive artwork the user co-creates the artwork, augmenting his or her perceptions along the way in an evolution of metaphor through the unfolding of harmonic relationships, toward unity.

An example of a hierophant could be the image of a wave crashing onto shore. Within this image are sets of forces and qualities that can modify other elements on the screen--an image of a chair and bird. If the user clicks on the crashing wave, this action modifies the chair by morphing its shape into a curved wooden bench accompanied by the sound of water. The action can also modify the bird who utters the words: "Sailors found me all wet by the ocean." User perception relates these as harmonic qualities, because each modification has "borrowed" qualities from the last cause and effect, creating a unity greater than the sum of their parts. This series of cause and effect relationships engages the user in harmonic aesthetic relationships, or a cybernetic system of metaphor. Because the harmonic relationships between objects and qualities is caused and perceived by the user, he or she experiences a healing or unifying effect through their relationship.

In a cybernetic system of metaphor (a series of aesthetic feedback loops) the user perceives unity in related sounds, text and images. The artists' perception and rules for creating the artwork, which used to be a passive or subconscious presence in traditional art forms, now becomes a conscious tool for the user to co-create. The work becomes the fulfillment of a code. This once-invisible artist's code has become a conscious syntax for co-creation, consisting of artistic elements, navigational rules and systems of interaction, Users physically unfold the art syntax and create unity in the work.

Healing is possible with an aesthetic multimedia interaction where the artist has embodied an artistic unity in the code/syntax and the user is empowered to move beyond duality (stress) to find unity among dissimilar elements. The artist's code becomes a hierophant to help an individual discover and perceive the mystery and hidden unity of the work resulting in a harmonic convergence between the user and the work.

The healing temples of the Greeks and Egyptians used color and sound harmonics to heal body, mind and emotions. Could digital art interactions using color, moving images, graphics, text and sound be designed as a twenty-first century equivalent? The personal computer station and multi-user kiosk are potential healing places. The software could even measure the reduction of stress by users to validate these programs as "healing technologies." Artists can assist others to a place of greater well-being through interactive design. An electronic syntax of light and sound elements, with user interactions guided by an artist-created code, can empower the individual via choice to create a healing cybernetic journey of aesthetic perception.