WHAT COMPUTERS CAN'T DO

A CRITIQUE OF ARTIFICIAL REASON

By Hubert L. Dreyfus

NEW YORK, EVANSTON, SAN FRANCISCO, LONDON

1817

HARPER & ROW, PUBLISHERS

7

The Role of the Body in Intelligent Behavior

Adherents of the psychological and epistemological assumptions that human behavior must be formalizable in terms of a heuristic program for a digital computer are forced to develop a theory of intelligent behavior which makes no appeal to the fact that a man has a body, since at this stage at least the computer clearly hasn't one. In thinking that the body can be dispensed with, these thinkers again follow the tradition, which from Plato to Descartes has thought of the body as getting in the way of intelligence and reason, rather than being in any way indispensable for it. If the body turns out to be indispensable for intelligent behavior, then we shall have to ask whether the body can be simulated on a heuristically programmed digital computer. If not, then the project of artificial intelligence is doomed from the start. These are the questions to which we must now turn.

Descartes, the first to conceive the possibility of robots, was also the first to suggest the essential inadequacy of a finite state machine. He remarks in the *Discourses:*

> Although such machines could do many things as well as, or perhaps even better than men, they would infallibly fail in certain others. . . . For while reason is a universal instrument which can be used in all sorts of situations, the organs of a machine have to be arranged in a particular way for each particular action. From this it follows that it is morally [i.e., practically] impossible that there

should be enough different devices in a machine to make it behave in all the occurrences of life as our reason makes us behave.[1]

Thus, although not aware of the difference between a situation and a physical state, Descartes already saw that the mind can cope with an indefinite number of situations, whereas a machine has only a limited set of states and so will eventually reveal itself by its failure to respond appropriately. This intrinsic limitation of mechanism, Descartes claims, shows the necessity of presupposing an *immaterial soul.*

This is an interesting argument, and some version of it may indeed be valid, but it gets its plausibility from the assumption that a robot can be in only a relatively small number of states. When in a modern computer the number of possible states is of the order of $10^{10^{10}}$, it is not clear just how much Descartes' objection proves. Such a machine could at least in principle respond to what would appear to be an indefinite number of situations. It would thus, on Descartes' view, be indistinguishable from a human being, destroying his argument that intelligent behavior is possible only if the mechanism behaving is somehow attached to a non-material soul. But one can raise a new objection, in some ways the exact opposite of Descartes'. A brain in a bottle or a digital computer might still not be able to respond to new sorts of situations because our ability to be in a situation might depend, not just on the flexibility of our nervous system, but rather on our ability to engage in practical activity. After some attempts to program such a machine, it might become apparent that what distinguishes persons from machines, no matter how cleverly constructed, is not a detached, universal, immaterial soul but an involved, self-moving, material body.

Indeed, it is just the bodily side of intelligent behavior which has caused the most trouble for artificial intelligence. Simon, who has been only slightly daunted by the failures of the last ten years, now feels that "machines will be capable, within *twenty* years, of doing any work that a man can do,"[2] but he admits: "Automation of a flexible central nervous system will be feasible long before automation of a comparatively flexible sensory, manipulative, or locomotive system."[3] But what if the work of the central nervous system depends on the locomotive system, or to put

it phenomenologically, what if the "higher," determinate, logical, and detached forms of intelligence are necessarily derived from and guided by global and involved "lower" forms? Then Simon's optimism, based on the three assumptions underlying artificial intelligence and traditional philosophy, would be unjustified.

The intractability of the "lower" functions has already produced a certain irony. Computer technology has been most successful in simulating the so-called higher rational functions—those which were once supposed to be uniquely human. Computers can deal brilliantly with ideal languages and abstract logical relations. It turns out that it is the sort of intelligence which we share with animals, such as pattern recognition (along with the use of language, which may indeed be uniquely human) that has resisted machine simulation.

Let us reconsider two related areas in which work in artificial intelligence has not fulfilled early expectations: game playing and pattern recognition. Thus far I have tried to account for the failure by arguing that the task in question cannot be formalized, and by isolating the nonformal form of "information processing" necessarily involved. Now I shall try to show that the nonformalizable form of "information processing" in question is possible only for embodied beings.

To make this clear we shall first have to consider human pattern recognition in more detail. With the aid of concepts borrowed from phenomenology, I shall try to show how pattern recognition requires a certain sort of indeterminate, global anticipation. This set or anticipation is characteristic of our body as a "machine" of nerves and muscles whose function can be studied by the anatomist, and also of our body as experienced by us, as our power to move and manipulate objects in the world. I shall argue that a body in both these senses cannot be reproduced by a heuristically programmed digital computer—even one on wheels which can operate manipulators, and that, therefore, by virtue of being embodied, we can perform tasks beyond the capacities of any heuristically programmed robot.

We have seen that the restricted applicability of pattern recognition programs suggests that human pattern recognition proceeds in some other way than searching through lists of traits. Indeed, phenomenologists and Gestalt psychologists have pointed out that our recognition of ordinary spatial or temporal objects does not seem to operate by checking off a list of isolable, neutral, specific characteristics at all. For example, in recognizing a melody, the notes get their values by being perceived as part of the melody, rather than the melody's being recognized in terms of independently identified notes. Likewise, in the perception of objects there are no neutral traits. The same hazy layer which I would see as dust if I thought I was confronting a wax apple might appear as moisture if I thought I was seeing one that was fresh. The significance of the details and indeed their very look is determined by my perception of the whole.

The recognition of spoken language offers the most striking demonstration of this global character of our experience. From time to time brash predictions such as Rosenblatt's have been made about mechanical secretaries into which (or at whom) one could speak, and whose programs would analyze the sounds into words and type out the results. In fact, no one knows how to begin to make such a versatile device, and further progress is unlikely, for current work has shown that the same physical constellation of sound waves is heard as quite different phonemes, depending on the expected meaning.

Oettinger has given considerable attention to the problem. His analysis of speech recognition work is worth reproducing in detail, both because this pattern recognition problem is important in itself and because this work exhibits the early success and subsequent failure to generalize which we have come to recognize as typical of artificial intelligence research.

> There was considerable initial success in building apparatus that would eke out a sequence of discrete phonemes out of the continuous speech waveform. While phonemic analysis has been dominant in that area, numerous other approaches to this decoding problem have also been followed. All have shared this initial degree of success and yet all, so far, have proved to be incapable of significant expansion beyond the recognition of the speech of a very few distinct individuals and the recognition of a very few distinct sound patterns whether they be phonemes or words or whatever. All is well as long as you are willing to have a fairly restricted universe of speakers, or sounds, or of both.
>
> Within these limitations you can play some very good tricks. There are now lots of machines, some experimental, some not so experimental, that will recognize somewhere between 20 and 100 distinct sound patterns, some of them quite elaborate. Usually the trick is something like identifying a number of features, treating these as if they were coordinates in some hyperspace, then passing planes that cordon off, if you will, different blocks of this space. If your speech event falls somewhere within one of these blocks you say that it must have been that sound and you recognize it.
>
> This game was fairly successful in the range of twenty to a hundred or so distinct things, but after that, these blocks become so small and clustered so close together that you no longer can achieve any reliable sort of separation. Everything goes to pot.[4]

This leads Oettinger to a very phenomenological observation:

> Perhaps . . . in perception as well as in conscious scholarly analysis, the phoneme comes after the fact, namely . . . it is constructed, if at all, as a *consequence* of perception not as a step in the process of perception itself.[5]

This would mean that the total meaning of a sentence (or a melody or a perceptual object) determines the value to be assigned to the individual elements.

Oettinger goes on reluctantly to draw this conclusion:

> This drives me to the unpopular and possibly unfruitful notion that maybe there is some kind of Gestalt perception going on, that here you are listening to me, and somehow the meaning of what I'm saying comes through to you all of a piece. And it is only a posteriori, and if you really give a damn, that you stop and say, "Now, here was a sentence and the words in it were of such and such type, and maybe here was a noun and here was a vowel and that vowel was this phoneme and the sentence is declarative, etc."[6]

Phenomenologists, not committed to breaking down the pattern so that it can be recognized by a digital computer, while less appalled, are no less fascinated by the gestalt character of perception. Indeed, it has been systematically studied in their account of perceptual horizons. Two forms of awareness are involved. First there is the basic figure-ground phenomenon, necessary for there to be any perception at all: whatever is prominent in our experience and engages our attention appears on a background which remains more or less indeterminate. This background, which need never have been made determinate, affects the appearance of what is determinate by letting it appear as a unified, bounded figure. In Rubin's famous "Peter-Paul Goblet" (Figure 3), "the contour which divides figure from ground 'belongs' to the figure only and changes its shape radically if a figure-ground reversal occurs."[7] Thus the figure has specific determinate characteristics, while the background can be characterized only as that-which-is-not-the figure.

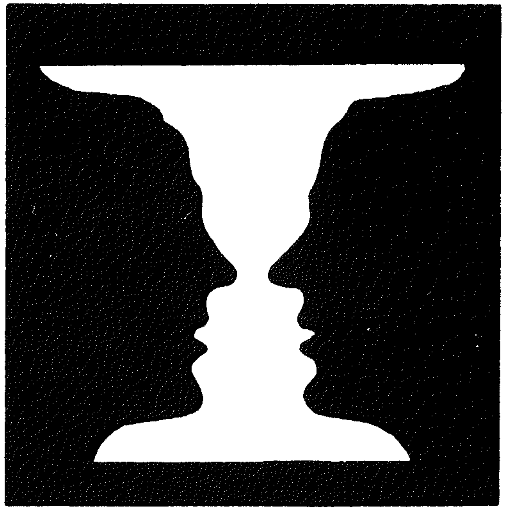

Figure 3

This indeterminacy plays a crucial role in human perception. Merleau-Ponty points out that most of what we experience must remain in the background so that something can be perceived in the foreground.

> When Gestalt theory informs us that a figure on a background is the simplest sense-datum available to us, we reply that this is not a contingent characterization of factual perception, which leaves us free, in an ideal analysis, to bring in the notion of impression. It is the very definition of the phenomenon of perception. . . . The perceptual 'something' is always in the middle of something else; it always forms part of a 'field.'[8]

It is this ground, or outer horizon as Edmund Husserl, the founder of phenomenology, called it, which in our chess example remains indeterminate and yet provides the context of the specific counting out, so that one always has a sense of the relevance of the specific move under

consideration to the rest of the game. Similarly, our sense of the overall context may organize and direct our perception of the details when we understand a sentence. For a computer, which must take up every bit of information explicitly or not at all, there could be no outer horizon. Any information to be taken into account would have to be as determinate as the figure. This leads to the unwieldy calculations which we have seen in chess programs and which Oettinger deplores in language programs.

This outer horizon, then, describes how background "information" about a conversation or a particular game is ignored without being excluded. It does not, however, describe the way the background provides information which contributes to the player zeroing in on one area of the chess board rather than another, or how our anticipation of a sentence's meaning determines our understanding of its elements as they fall into place. To understand this, we must consider a second kind of perceptual indeterminacy investigated by Husserl and Gestalt psychologists: what Husserl calls the inner horizon. The something-more-than-the-figure is, in this case, not as indeterminate as the outer horizon. When we perceive an object we are aware that it has more aspects than we are at the moment considering. Moreover, once we have experienced these further aspects, they will be experienced as copresent, as covered up by what is directly presented. Thus, in ordinary situations, we say we perceive the whole object, even its hidden aspects, because the concealed aspects directly affect our perception. We perceive a house, for example, as more than a façade—as having some sort of back—some inner horizon. We respond to this whole object first and then, as we get to know the object better, fill in the details as to inside and back. A machine with no equivalent of an inner horizon would have to process this information in the reverse order: from details to the whole. Given any aspect of an object, the machine would either pick it up on its receptors or it would not. All additional information about other aspects of the object would have to be explicitly stored in memory—in Minsky's sort of model—or counted out again when it was needed. This lack of horizons is the essential difference between an image in a movie or on a TV screen and the same scene as experienced by a human being.

When, in a film, the camera is trained on an object and moves nearer to it to give a close-up view, we can *remember* that we are being shown the ash tray or an actor's hand, we do not actually identify it. This is because the scene has no horizons.[9]

In chess and in recognizing sentences, we find the same phenomenon playing a crucial role. Our sense of the whole situation, outer horizon, and our past experience with the specific object or pattern in question, inner horizon, give us a sense of the whole and guide us in filling in the details.[10]*

This process can best be noticed when it is breaking down. If you reach for a glass of water and get milk by mistake, on taking a sip your first reaction is total disorientation. You don't taste water, but you don't taste milk either. You have a mouthful that approaches what Husserl would call pure sensuous matter or hyletic data, and naturally you want to spit it out. Or, if you find the right global meaning fast enough, you may recover in time to recognize the milk for what it is. Its other characteristics, whether it is fresh or sour, buttermilk or skimmed milk, will then fall into place.

One might well wonder how one knows enough to try "milk" rather than, say, "gasoline." Doesn't one need some neutral features to begin this process of recognition? The perceiver's apparent clairvoyance seems so paradoxical that one is tempted to embrace the computer model in spite of its difficulties. But the process seems less mysterious when we bear in mind that each new meaning is given in an outer horizon which is already organized, in this case a meal, on the basis of which we already have certain expectations. It is also important that we sometimes *do* give the wrong meaning; in these cases the data coming in make no sense at all, and we have to try a new *total* hypothesis.

A computer, which must operate on completely determinate data according to strictly defined rules, could at best be programmed to try out a series of hypotheses to see which best fit the fixed data. But this is far from the flexible interaction of underdetermined data and underdetermined expectations which seems to be characteristic of human pattern recognition.

As one might expect, the computer people, again with the support of

the philosophical tradition, and the success of physics, have rarely faced this problem. Philosophers have thought of man as a contemplative mind passively receiving data about the world and then ordering the elements. Physics has made this conception plausible on the level of the brain as a physical object. The brain does passively receive energy from the physical world and process it in terms of its present state which is a function of past energy received. If one accepts the passive view of mind and fails to distinguish the physical-processing level from the "information-processing" level, it seems self-evident that the mind, like the computer, simply receives bits of determinate data. In his introduction to the *Scientific American* issue on computers, McCarthy naïvely confuses brain and mind, energy and information, so that the passivity of the computer appears to be a self-evident model for human "information processing."

The human brain also accepts inputs of information, combines it with information stored somehow within, and returns outputs of information to its environment.[11]

Neisser is much more subtle. He too underestimates the problems posed by the role of anticipation, but his work in psychology has at least led him to see the need for "wholistic operations which form the units to which attention may then be directed,"[12] and he tries to fit this fact into his overall commitment to a digital computer model. The result is a confusion between what "global or wholistic" means in a gestalt analysis and what it would have to mean in a computer program, which is sufficiently revealing to be worth following in detail.

A general characterization of the gestalt, or global, phenomenon is: the interpretation of a part depends on the whole in which it is embedded. But this is too general. Such a definition allows Minsky, for example, to miss the whole problem. In his *Scientific American* article he speaks of Evans' analogy-solving program as being able to "recognize a 'global' aspect of the situation."[13] This turns out to mean that, on the basis of calculations made on certain local features of a figure, the program segments two superimposed figures in one way rather than another. There is nothing here to surprise or interest those concerned with the way the gestalt, or global, configuration functions in our experience.

To see the difference between the wholistic processes which interest Neisser and what Minsky calls global recognition, one needs a sharper characterization of the gestalt phenomenon. Neisser gives such a characterization in terms of a temporal gestalt, a rhythm (a favorite example of the Gestaltists):

> The parts (individual beats) get their meaning (relative position) from the whole, even though that whole does not exist at any moment of time. It exists, as one might say, in the subject's mind, as an intent . . . Gestalt. . . .[14]

The crucial feature of this gestalt interpretation, that each *part gets its meaning only in terms of the whole,* is missing in Minsky's example, as it must be, since, as we have seen, for a digital computer, each complex whole must be constructed by the logical combination of *independently defined* elements. In Minsky's example, the elements already have a precise significance (or rather two possible precise significances), and it is simply a question of deciding which interpretation is appropriate in terms of a decision based on other determinate local features of the figure.

Neisser's description of the "mind's intent" which gives the individual beats their significance, on the other hand, brings us to the center of the problem. The question is how the partially determinate anticipation, involved in game playing, pattern recognition, and intelligent behavior in general, can be simulated on a heuristically programmed digital computer so that a computer does not have to go through the enormous calculation required by an explicit internal model. More specifically for Neisser, the problem is how to reconcile his gestaltist analysis with a computer model of human performance.

Neisser thinks he has a way. In discussing linguistic performance as an example of the gestalt effect, Neisser thinks of the rules of grammar as the wholes into which the words fit as parts.

> The rules are *structural.* That is, they do not dictate what particular words are to be used, but rather how they are to be related to each other and to the sentence as a whole.[15]

But this will not work. In the case of the rhythm, the whole *determined the meaning* of each element—there is no such thing as a syncopated beat, for example, existing all by itself—but for Neisser, in the case of language, the words already have a determinate set of possible meanings; the grammar simply provides a rule for selecting a meaning and combining it with others. The elements in this case are completely determinate and can be defined independently of the rules. It is, therefore, misleading when Neisser concludes:

> A sentence is more than the sum of its parts. This is not an unfamiliar slogan. Long ago, the Gestalt psychologists used it to describe the wholistic aspects of visual perception.[16]

This confusion is already latent in Neisser's description of the anticipation involved in hearing a rhythm in the example quoted above. The description concludes: "[The anticipation] exists . . . in the subject's mind as an intent, a gestalt, *a plan, a description of a response that can be executed without further consideration.*"[17] This slide from gestalt anticipation to preset plan is an obfuscation necessitated by the computer model: A gestalt determines the meaning of the elements it organizes; a plan or a rule simply organizes independently defined elements. Moreover, just as the elements (the beats) cannot be defined independently of the gestalt, the gestalt (the rhythm) is nothing but the organization of the elements. A plan, on the other hand, can be stated as a rule or program, independently of the elements. Clearly his computer model of a formal program defined and stored separately from the independently defined bits of data which it organizes leads Neisser to betray his own gestaltist illustration. This difference is neglected in all CS models, yet it is the essence of the gestaltist insight, and accounts for the flexibility of human pattern recognition compared to that of machines.

Thus far computer programs have been unable to approach this interdependence of parts and whole. Neisser himself never sees this problem, but he unwittingly casts some new light on the important differences between mechanist and gestaltist models of psychological processes when he contrasts the digital model of neural processes postulated by the transformational linguists with the analogue model of the brain espoused by the early Gestalt psychologists.

> [The Gestaltists] were "nativists," believing that the perceptual processes were determined by necessary and innate principles rather than by learning. The proper figural organization. . . . was due to processes in the brain, which followed unvarying (and wholistic) laws of physics and chemistry. . . . The perceived world always took the "best," the "structurally simplest" form, because of the equilibrium principle that transcends any possible effects of learning or practice.[18]

Such an analogue model of brain function, in which information is integrated by equilibrium forces rather than on/off switches, was necessary if the Gestalt psychologists were to account for the role of global anticipations in structuring experience. They had been led to break with the rationalist tradition running from Descartes to Kant, which conceived of the mind as bringing independently defined innate principles (Descartes) or rules (Kant) to bear on otherwise unstructured experience. This rationalist conception (with the addition of minimal bits of determinate experience) lends itself perfectly to a computer model, but the Gestaltists saw that their principles of organization—like the equilibrium patterns formed by charged particles on curved surfaces—could not be separated from the elements they organized. Thus, even if the digital model of the brain had existed at the time, the Gestaltists would have rejected it.[19]*

Neisser does not see this. He supposes that the digital model of built-in rules, which the linguists have been led to propose, is an improvement on the analogue model proposed by the Gestaltists. Neisser's praise of the linguists' "improvement," ignoring as it does the difficulties in artificial intelligence, the latest developments in neurophysiology, and the reason the Gestaltists proposed an analogue model in the first place can only be a non sequitur:

> The Gestalt psychologists were never able to provide any satisfactory description or analysis of the structures involved in perception. The few attempts to specify "fields of force" in vision, or "ionic equilibria" in the brain, were *ad hoc* and ended in failure. In linguistics, by contrast, the study of "syntactic structures" has a long history.[20]

How the long history of syntactic structures is supposed to show that the linguists have a better model of neural processes than the Gestaltists is totally unclear. It seems to mean that at least the rules the linguists are looking for would be, if they were found, the sort of rules one could process with a digital computer which we already understand, whereas the gestaltist equilibrium principles could only be simulated on a brain-like analogue computer, which no one at present knows how to design.

This is no doubt true, but it reminds one of the story of the drunk who lost a key in the dark but looked for it under a street lamp because the light was better. It would indeed be nice to have a programmable model in linguistics, and in psychology in general, but the fact remains that modern linguists have no more detailed account of what goes on in the brain than did the Gestaltists, and, moreover, as a theory of competence, not performance, modern linguistics is not even trying to provide answers to the problem of how we produce intelligent behavior. Worse, in this case, the street lamp is not even lit. We have seen that when digital computers have been used to try to simulate linguistic *performance,* they have had remarkably little success.

The upshot of Neisser's comparison of gestalt and linguistic models of the brain, in opposition to his intent, is to call attention to a difference in brain model which exactly parallels the difference in the conception of the wholistic processes, which he also overlooks. The sort of gestalt process illustrated in Neisser's example of the rhythm which gives meaning to and is made up of its beats suggests that however the brain integrates stimuli, it does not do it like a digital computer applying independently defined heuristic rules to independently defined bits of data.

Among computer experts only Donald MacKay has seen this point. He concludes:

> It may well be that only a special-purpose 'analogue' mechanism could meet all detailed needs. . . . We on the circuit side had better be very cautious before we insist that the kind of information processing that a brain does can be replicated in a realizable circuit. Some kind of 'wet' engineering may turn out to be inevitable.[21]

If, in the light of the phenomenological and neurophysiological evidence, we accept the view that the nervous system is some sort of analogue computer operating with equilibrium fields, we must still be on guard against transferring to psychology this model of the nervous system, conceived as a brain in a bottle receiving energy from the world and sending out responses. The human perceiver must be understood in different terms than his nervous system. To have an alternative account of intelligent behavior we must describe the general and fundamental features of human activity. In the absence of a workable digital computer model, and leaving to the neurophysiologist the question of how the brain integrates incoming physical stimuli, we must again ask, How do human beings use an underdetermined, wholistic expectation to organize their experience?

Husserl has no further account beyond the assertion that we do: that "transcendental consciousness" has the "*wunderbar*" capacity for giving meanings and thus making possible the perception, recognition, and exploration of enduring objects. Like the Gestaltists, he thinks of these meanings as partially indeterminate wholes, not as explicit programs or rules. But even Husserl is not free from the traditional intellectualist view, and thus he too is vulnerable to the criticism directed at Neisser. Husserl, like Descartes and Kant, thinks of form as separable from content, of the global anticipation as separable from its sensuous feeling. Thus, his noema, or perceptual anticipation, is like a rule or program in one crucial way: it exists in the mind or transcendental consciousness independently of its application to the experience it structures.

Merleau-Ponty tries to correct Husserl's account on this point and at the same time develop a general description which supports the Gestaltists. He argues that it is the body which confers the meanings discovered by Husserl. After all, it is our body which captures a rhythm. We have a body-set to respond to the sound pattern. This body-set is not a rule in the mind which can be formulated or entertained apart from the actual activity of anticipating the beats.

Generally, in acquiring a skill—in learning to drive, dance, or pronounce a foreign language, for example—at first we must slowly, awkwardly, and consciously follow the rules. But then there comes a moment when we finally transfer control to the body. At this point we do not seem

to be simply dropping these same rigid rules into unconsciousness; rather we seem to have picked up the muscular gestalt which gives our behavior a new flexibility and smoothness. The same holds for acquiring the skill of perception. To take one of Merleau-Ponty's examples: to learn to feel silk, one must learn to move or be prepared to move one's hand in a certain way and to have certain expectations. Before we acquire the appropriate skill, we experience only confused sensations.

It is easiest to become aware of the body's role in taste, hearing, and touch, but seeing, too, is a skill that has to be learned. Focusing, getting the right perspective, picking out certain details, all involve coordinated actions and anticipations. As Piaget remarks, "Perceptual constancy seems to be the product of genuine actions, which consist of actual or potential movements of the glance or of the organs concerned. . . ."[22]

These bodily skills enable us not only to recognize objects in each single sense modality, but by virtue of the felt equivalence of our exploratory skills we can see and touch the same object. A computer to do the same thing would have to be programmed to make a specific list of the characteristics of a visually analyzed object and compare that list to an explicit list of traits recorded by moving tactical receptors over that same object. This means that there would have to be an internal model of each object in each sense modality, and that the recognition of an object seen and felt must pass through the analysis of that object in terms of common features.

My body enables me to by-pass this formal analysis. A skill, unlike a fixed response or set of responses can be brought to bear in an indefinite number of ways. When the percipient acquires a skill, he

> does not weld together individual movements and individual stimuli but acquires the power to respond with a certain type of solution to situations of a certain general form. The situations may differ widely from place to place, and the response movements may be entrusted sometimes to one operative organ, sometimes to another, both situations and responses in the various cases having in common not so much a partial identity of elements as a shared significance.[23]

Thus I can recognize the resistance of a rough surface with my hands, with my feet, or even with my gaze. My body is thus what Merleau-Ponty calls a "synergic system,"[24] "a ready-made system of equivalents and transpositions from one sense to another."[25]

Any object presented to one sense calls upon itself the concordant operation of all the others. I see a surface colour because I have a visual field, and because the arrangement of the field leads my gaze to that surface—I perceive a thing because I have a field of existence and because each phenomenon, on its appearance, attracts towards that field the whole of my body as a system of perceptual powers.[26]

A human perceiver, like a machine, needs feedback to find out if he has successfully recognized an object. But here too there is an important difference in the feedback involved. A machine can, at best, make a specific set of hypotheses and then find out if they have been confirmed or refuted by the data. The body can constantly modify its expectations in terms of a more flexible criterion: as embodied, we need not check for specific characteristics or a specific range of characteristics, but simply for whether, on the basis of our expectations, we are coping with the object. Coping need not be defined by any specific set of traits but rather by an ongoing mastery which Merleau-Ponty calls *maximum grasp.* What counts as maximum grasp varies with the goal of the agent and the resources of the situation. Thus it cannot be expressed in situation-free, purpose-free terms.

To conclude: Pattern recognition is relatively easy for digital computers if there are a few specific traits which define the pattern, but complex pattern recognition has proved intractable using these methods. Transcendental phenomenologists such as Husserl have pointed out that human beings recognize complex patterns by projecting a somewhat indeterminate whole which is progressively filled in by anticipated experiences. Existential phenomenologists such as Merleau-Ponty have related this ability to our active, organically interconnected body, set to respond to its environment in terms of a continual sense of its own functioning and goals.

Since it turns out that pattern recognition is a bodily skill basic to all intelligent behavior, the question of whether artificial intelligence is possible boils down to the question of whether there can be an artificial embodied agent. The question is philosophically interesting only if we restrict ourselves to asking if one can make such a robot by using a digital computer. (I assume there is no reason why, in principle, one could not

construct an artificial embodied agent if one used components sufficiently like those which make up a human being.)

A project to build such a *digitally* controlled robot is currently under way at M.I.T., and it is philosophically interesting to consider its program and its underlying assumptions. The project director, Minsky again, is modestly trying to make only a mechanical shoulder, arm, and hand, coordinated with a TV eye, but he proposes to make it use tools to construct things. The first simple task was to program a simplified robot arm to pick up blocks. This has indeed been accomplished and represents the early success one has learned to expect in the field. The problem which remains is, as usual, that of generalizing the present successful techniques. To bring a simple arm over to pick up a block requires locating the block in objective space, locating the arm in the same space, and then bringing the two together. This is already quite a feat. A mathematical description of the way an arm moves in objective space runs into surprising discontinuities. There are points which are contiguous in objective space which are far apart in reaching space. For example, to scratch our back we do not simply extend the position we use for scratching our ear. Living in our bodies we have built up a motor space, in which we sense these objectively contiguous points as far apart. We automatically reach for them in very different ways, and do not feel we have gone through the mathematics necessary to work out the optimal path for each specific case. For the programmer, however, who has to program the computer to calculate the movements of the mechanical arm in objective space, these discontinuities have so far proved an insurmountable obstacle. The more flexible the arm—the more degrees of freedom it has—the more difficult and time consuming such calculations become. Rumor has it that an elaborate arm with six degrees of freedom, built by Minsky by 1965, has still not even been programmed to move, let alone pick up blocks or use tools. If one adds to this the fact that, in the case of any skill which takes place in real time (such as playing Ping-Pong), all calculations must be completed in real time (before the ball arrives), the outlook is not very promising. As Feigenbaum notes in his report on the current state of robot work:

Both the MIT and Stanford University groups have worked on programs for controlling a variety of arm-hand manipulators, from the very simple to the very complex, from the anthropomorphic variety to the very non-anthropomorphic. None of the more esoteric manipulators seems to have worked out very well, though there is no published documentation of successes, failures, and reasons.[27]

In the light of these difficulties, what encourages researchers to devote their research facilities to such a project? Simply the conviction that since we are, as Minsky ingenuously puts it, "meat machines" and are able to play ping-pong, there is no reason in principle or in practice why a metal machine cannot do likewise. But before jumping to such a conclusion, the robot makers ought first to examine their underlying assumption that no essential difference exists between meat machines and metal machines, between being embodied and controlling movable manipulators. How do human beings play ping-pong, or to make the matter simpler, how do human beings use tools?

Heidegger, Merleau-Ponty, and Michael Polanyi have each devoted a great deal of thought to this question. Each discusses the important way that our experience of a tool we are using differs from our experience of an object. A blind man who runs his hand along the cane he uses to grope his way will be aware of its position and its objective characteristics such as weight, hardness, smoothness, and so forth. When he is using it, however, he is not aware of its objective position, its physical traits, nor of the varying pressure in the palm of his hand. Rather, the stick has become, like his body, a transparent access to the objects he touches with it. As Polanyi puts it:

> While we rely on a tool or a probe, these are not handled as external objects . . . they remain on our side . . . forming part of ourselves, the operating persons. We pour ourselves out into them and assimilate them as parts of our existence. We accept them existentially by dwelling in them.[28]

In this way we are able to bring the probe into contact with an object in physical space without needing to be aware of the physical location of the probe. Merleau-Ponty notes that:

> The whole operation takes place in the domain of the phenomenal; it does not run through the objective world, and only the spectator, who lends his objective representation of the living body to the active subject, can believe that . . . the hand moves in objective space.[29]

But Merleau-Ponty admits that this ability seems "magical" from the point of view of science, so we should not be surprised to find that rather than have no explanation of what people are able to do, the computer scientist embraces the assumption that people are unconsciously running with incredible speed through the enormous calculation which would be involved in programming a computer to perform a similar task. However implausible, this view gains persuasiveness from the absence of an alternative account.

To make embodiment an acceptable alternative we will have to show how one could perform physical tasks without in any way appealing to the principles of physics or geometry. Consider the act of randomly waving my hand in the air. I am not trying to place my objective hand at an objective point in space. To perform this waving I need not take into account geometry, since I am not attempting any specific achievement. Now suppose that, in this random thrashing about, I happen to touch something, and that this satisfies a need to cope with things. (More about need in Chapter 9.) I can then repeat *whatever I did*—this time *in order to* touch something—without appealing to the laws necessary to describe my movement as a physical motion. I now have a way of bringing two objects together in objective space without appealing to any principle except: "Do that again." This is presumably the way skills are built up. The important thing about skills is that, although *science* requires that the skilled performance be *described* according to rules, these rules need in no way be *involved* in producing the performance.

Human beings are further capable of remembering, refining, and reorganizing these somewhat indeterminate motor schemata. Piaget has amassed an enormous amount of evidence tracing the development of these motor skills, which he calls operations, and has come to a gestaltist conclusion:

> The specific nature of operations . . . depends on the fact that they never exist in a discontinuous state. . . . A single operation could not be an operation because the peculiarity of operations is that they form systems. Here we may well protest vigorously against logical atomism . . . a grievous hindrance to the psychology of thought.[30]*

This same analysis helps dissipate the mistaken assumptions underlying early optimism about language translation. If human beings had to apply semantic and syntactic rules and to store and access an infinity of facts in order to understand a language, they would have as much trouble as machines. The native speaker, however, is not aware of having generated multiple semantic ambiguities which he then resolved by appeal to facts any more than he is aware of having picked out complex patterns by their traits or of having gone through the calculations necessary to describe the way he brings his hand to a certain point in objective space. Perhaps language, too, is a skill acquired by innately guided thrashing around and is used in a nonrulelike way. Wittgenstein suggests this point when he notes, "In general we don't *use* language according to strict rules—it hasn't been taught us by means of strict rules either."[31]

Such a view is not behavioristic. Our ability to use language in a situation and in general the wholistic way the functional meaning organizes and structures the components of skilled acts cannot be accounted for in terms of the arbitrary association of neutral determinate elements any more than it can be analyzed in terms of their combination according to rules.

If language is understood as a motor skill, we would then assimilate language and dwell in it the way we assimilate an instrument. As Polanyi puts it,

> To use language in speech, reading and writing, is to extend our bodily equipment and become intelligent human beings. We may say that when we learn to use language, or a probe, or a tool, and thus make ourselves aware of these things as we are of our body, we *interiorise* these things and *make ourselves dwell in them.*[32]*

Again, because we are embodied, the rules necessary to give an objective analysis of our competence need in no way be involved in our performance.

The AI researcher and the transcendental phenomenologist share the assumption that there is only one way to deal with information: it must be made an object for a disembodied processor. For the transcendental phenomenologist this assumption makes the organization of our intelligent behavior unintelligible. For the AI researcher it seems to justify the assumption that intelligent behavior can be produced by passively receiving data and then running through the calculations necessary to describe the objective competence. But, as we have seen, being embodied creates a second possibility. The body contributes three functions not present, and not as yet conceived in digital computer programs: (1) the inner horizon, that is, the partially indeterminate, predelineated anticipation of partially indeterminate data (this does not mean the anticipation of some completely determinate alternatives, or the anticipation of completely unspecified alternatives, which would be the only possible digital implementation); (2) the global character of this anticipation which determines the meaning of the details it assimilates and is determined by them; (3) the transferability of this anticipation from one sense modality and one organ of action to another. All these are included in the general human ability to acquire bodily skills. Thanks to this fundamental ability an embodied agent can dwell in the world in such a way as to avoid the infinite task of formalizing everything.

This embodied sort of "information processing," in which the meaning of the whole is prior to the elements, would seem to be at work in the sort of complex pattern recognition such as speech recognition with which we began our discussion. It is also necessary, in order to account for our ability to recognize typicality, family resemblances, and similarity, where the objects recognized need have no traits in common at all. In all these cases individual features get their significance in terms of an underdetermined anticipation of the whole.

If these global forms of pattern recognition are not open to the digital computer, which, lacking a body, cannot respond as a whole, but must build up its recognition starting with determinate details, then Oettinger is justified in concluding his speech recognition paper on a pessimistic note: "If indeed we have an ability to use a global context without recourse to formalization . . . then our optimistic discrete enumerative approach is doomed. . . ."[33]

8

The Situation: Orderly Behavior Without Recourse to Rules

In discussing problem solving and language translation we have come up against the threat of a regress of rules for determining relevance and significance. Likewise, in starting a learning process, something must be known before any rules can be taught or applied. In each case we have found that if there are no facts with fixed significance, only an appeal to the context can bring this regress to a halt. We must now turn directly to a description of the situation or context in order to give a fuller account of the unique way human beings are "in-the-world," and the special function this world serves in making orderly but nonrulelike behavior possible.

To focus on this question it helps to bear in mind the opposing position. In discussing the epistemological assumption (Chapter 5) we saw that our philosophical tradition has come to assume that whatever is orderly can be formalized in terms of rules. This view has reached its most striking and dogmatic culmination in the conviction of AI workers that every form of intelligent behavior can be formalized. Minsky has even developed this dogma into a ridiculous but revealing theory of human free will. He is convinced that all regularities are rule governed. He therefore theorizes that our behavior is either completely arbitrary or it is regular and completely determined by rules. As he puts it:

". . . whenever a regularity is observed [in our behavior], its representation is transferred to the deterministic rule region."[1] Otherwise our behavior is completely arbitrary and free. The possibility that our behavior might be regular but not rule governed never even enters his mind.

We shall now try to show not only that human behavior can be regular without being governed by formalizable rules, but, further, that it has to be, because a total system of rules whose application to all possible eventualities is determined in advance makes no sense.

In our earlier discussion of problem solving we restricted ourselves to formal problems in which the subject had to manipulate unambiguous symbols according to a given set of rules, and to other context-free problems such as analogy intelligence tests. But if CS is to provide a psychological theory—and if AI programs are to count as intelligent—they must extend mechanical information processing to *all* areas of human activity, even those areas in which people confront and solve open-structured problems in the course of their everyday lives.[2]*

Open-structured problems, unlike games and tests, raise three sorts of difficulties: one must determine which facts are possibly relevant; which are actually relevant; and, among these, which are essential and which inessential. To begin with, in a given situation not all facts fall within the realm of possible relevancy. They do not even enter the situation. Thus, in the context of a game of chess, the weight of the pieces is irrelevant. It can never come into question, let alone be essential or inessential for deciding on a specific move. In general, deciding whether certain facts are relevant or irrelevant, essential or inessential, is not like taking blocks out of a pile and leaving others behind. What counts as essential depends on what counts as inessential and vice versa, and the distinction cannot be decided in advance, independently of some particular problem, or some particular stage of some particular game. Now, since facts are not relevant or irrelevant in a fixed way, but only in terms of human purposes, all facts are possibly relevant in some situation. Thus for example, if one is *manufacturing* chess sets, the weight is *possibly* relevant (although in most decisions involved in making and marketing chess sets, it will not be actually relevant, let alone essential). This situational character of relevance works both ways: In any particular situation an indefinite number of facts are possibly relevant and an indefinitely large number are irrelevant. Since a computer is not in a situation, however, it must treat *all* facts as possibly relevant at all times. This leaves AI workers with a dilemma: they are faced either with storing and accessing an infinity of facts, or with having to exclude some possibly relevant facts from the computer's range of calculations.

But even if one could restrict the universe for each particular problem to possibly relevant facts—and so far this can only be done by the programmer, not the program—the problem remains to determine what information is actually relevant. Even in a nonformal game like playing the horses—which is much more systematic than everyday open-structured problems—an unlimited, indefinitely large number of facts remain as possibly relevant. In placing a bet we can usually restrict ourselves to such facts as the horse's age, jockey, past performance, and competition. Perhaps, if restricted to these facts from the racing form, the machine could do fairly well, possibly better than an average handicapper; but there are always other factors such as whether the horse is allergic to goldenrod or whether the jockey has just had a fight with the owner, which *may* in some cases be decisive. Human handicappers are no more omniscient than machines, but they are capable of recognizing the relevance of such facts if they come across them. The artificial intelligence approach to this human ability would have to be to give the machine knowledge about veterinary medicine, how people behave when they fight their employers, and so forth. But then the problem arises of sorting through this vast storehouse of data. To which the answer is that all this information would be properly coded and tagged in the machine memory so that the machine would just have to do a scan for "horse-race betting" and get out the relevant material. But not all relevant material would have been encoded with a reference to this particular use. As Charles Taylor has pointed out in an elaboration of this example:

> The jockey might not be good to bet on today because his mother died yesterday. But when we store the information that people often do less than their best just after their near relations die, we can't be expected to tag a connection with betting on horses. This information can be relevant to an infinite set of contexts.
>
> The machine might select on the basis of the key concepts it was worrying about, horses, jockeys, jockey Smith, etc. and pick out all facts about these. But this too would give an absurdly wide scatter. Via jockey, man and horse, one would find oneself pulling out all facts about centaurs. The only way the machine could zero in on the relevant facts would be to take this broad class, or some other selected on such a broad swoop basis, and test to see whether each one had causal relevance to the outcome of the race, taking it into account if it had, and forgetting it if it hadn't.[3]*

But if the machine were to examine explicitly each possibly relevant factor as a determinate bit of information in order to determine whether to consider or ignore it, it could never complete the calculations necessary to predict the outcome of a single race. If, on the other hand, the machine systematically excluded possibly relevant factors in order to complete its calculations, then it would sometimes be incapable of performing as well as an intelligent human to whom the same information was available.

Even the appeal to a random element will not help here, since in order to take up a sample of excluded possibilities at random so that no possibility is in principle excluded, the machine would have to be provided with an explicit list of all such other possibly relevant facts or a specific set of routines for exploring all classes of possibly relevant facts, so that no facts would be in principle inaccessible. This is just what could be done in a completely defined system such as chess, where a finite number of concepts determines totally and unequivocally the set of all possible combinations in the domain; but in the real world the list of such possibly relevant facts, or even classes of possibly relevant facts, would be indefinitely large ("infinite in a pregnant sense," to use Bar-Hillel's phrase). All the everyday problems—whether in language translation, problem solving, or pattern recognition—come back to these two basic problems: (1) how to restrict the class of possibly relevant facts while preserving generality, and (2) how to choose among possibly relevant facts those which are actually relevant.

Even Minsky implicitly admits that no one knows how to cope with the amount of data which must be processed if one simply tries to store all facts:

At each moment in the course of thinking about a problem, one is involved with a large collection of statements, definitions, associations, and so on, and a network of goals. One has to deal not only with facts about objects, relations between objects, and the like, but also facts about facts, classes of facts, relations between such classes, etc. The heuristic programs that, as we shall see, so neatly demonstrate principles when applied to small models will not work efficiently when applied to large ones. Problems like looping, branching, measuring progress, and generally keeping track of what is happening will come to require a disproportional part of the computation time.[4]

Whatever it is that enables human beings to zero in on the relevant facts without definitively excluding others which might become relevant is so hard to describe that it has only recently become a clearly focused problem for philosophers. It has to do with the way man is at home in his world, has it comfortably wrapped around him, so to speak. Human beings are somehow already situated in such a way that what they need in order to cope with things is distributed around them where they need it, not packed away like a trunk full of objects, or even carefully indexed in a filing cabinet. This system of relations which makes it possible to discover objects when they are needed is our home or our world. To put this less metaphorically it is helpful to return to Charles Taylor's extension of the horse-racing example.

Much of a human being's knowledge of situations and their possibilities is know-how, that is, it cannot be exhaustively unpacked into a set of specific instructions or factual statements, but is a general capacity to generate appropriate actions and therefore, if necessary, the "instructions" underlying them. Usually we think of this kind of indefinitely unpackable form of knowledge as bound up with the know-how which underlies our actions. But the same kind of knowledge underlies what we suffer, our "passions." Thus just as I have a general grasp on what it is to walk around, use my hands, drive a car, conduct a case in court (if I'm a lawyer), etc. So I have a general grasp on what it is to be threatened, to hear good news, to be jilted by my girl friend, to be made a fool of in public.

Now the human handicapper has this general grasp of certain common human actions and passions. He has the sense of the race as a perilous enterprise which needs all the will and effort of jockey (and horse) to win. But included in this sense is the capacity to imagine or recognize an indefinite number of ways in which this will and effort could miscarry or be countered by fortune. These are not stored somewhere as separate facts in the mind or brain, they are not

"unpacked"; they are just generatable from the general grasp of the situation. Of course, the general grasp of different men may differ in scope and exactitude. If the handicapper has ever ridden horses, then he has a much firmer grasp on the activity; he can sense a lot more finely what may go wrong. But even the city-bred gangster has some general grasp of what it is to fight and strain hard to win.

But the artificial intelligence proponent may still want to protest that all this just represents an alternative method of "storage." Even if he admits that this method is not available to the machine, he might still ask how it solves the retrieval problem. How does the handicapper recognize just those odd factors which are relevant? The answer is that if we understand our grasp of the world as arising out of our dealing with it according to our different capacities, and our being touched by it according to our different concerns, then we can see that the problem of how a given concern or purpose comes to select the relevant features of our surroundings *doesn't arise. For being concerned in a certain way or having a certain purpose is not something separate from our awareness of our situation; it just is being aware of this situation in a certain light,* being aware of a situation with a certain structure. Thus being anxious for my own life because I have fallen among thugs is to sense the menace in that bulge in his pocket, to feel my vulnerability to his fist which might at any moment be swung at my face, and so on.[5]

The human world, then, is prestructured in terms of human purposes and concerns in such a way that what counts as an object or is significant about an object already is a function of, or embodies, that concern. This cannot be matched by a computer, for a computer can only deal with already determinate objects, and in trying to simulate this field of concern, the programmer can only assign to the already determinate facts further determinate facts called values, which only complicates the retrieval problem for the machine.

In *Being and Time* Heidegger gives a description of the human world in which man is at home, on the model of a constellation of implements *(Zuege),* each referring to each other, to the whole workshop and ultimately to human purposes and goals. The directional signal on a car serves as an example of a "fact" which gets its whole meaning from its pragmatic context:

The directional signal is an item of equipment which is ready-to-hand for the driver in his concern with driving, and not for him alone: those who are not travelling with him—and they in particular—also make use of it, either by giving

> way on the proper side or by stopping. This sign is ready-to-hand within-the-world in the whole equipment-context of vehicles and traffic regulations. It is equipment for indicating, and as equipment, it is constituted by reference or assignment.[6]

Wittgenstein too makes frequent references to human forms of life and concerns and to certain very general "facts of natural history" taken for granted in our use of language and in structuring our everyday activities—facts, incidentally, of a very special kind which would presumably elude the programmer trying to program all of human knowledge. As Wittgenstein says, "The aspects of things that are most important for us are hidden because of their simplicity and familiarity. (One is unable to notice something—because it is always before one's eyes.)"[7] Facts, moreover, which would be so pervasively connected with all other facts that even if they could be made explicit, they would be difficult if not impossible to classify. The basic insight dominates these discussions that the situation is organized from the start in terms of human needs and propensities which give the facts meaning, make the facts what they are, so that there is never a question of storing and sorting through an enormous list of *meaningless, isolated* data.

Samuel Todes[8]* has described in detail the field-structure of experience which is prior to the facts and implicitly determines their relevance and significance. He points out that the world is experienced as fields within fields. Bits or aspects of objects are not experienced as isolated facts but as nested in a series of contexts. And "in" has many different senses, none of them that of mere physical inclusion, which Minsky and McCarthy take as primary. Parts of objects are experienced as *in* objects which they *comprise,* objects are *in* places which they *fill,* a place is *situated in* a local environment, which itself is *in* the horizon of possible situations *in* a human world. Data, then, are far from brute; aspects of objects are not given as directly in the world but as characterizing objects in places in a local environment in space and time in the world.

We can and do zero in on significant content in the field of experience because this field is not neutral to us but is structured in terms of our interests and our capacity for getting at what is in it. Any object which we experience must appear in this field and therefore must appear in

terms of our dominant interest at that moment, and as attainable by some variant of the activity which generated the field. Since we create the field in terms of our interests, only possibly relevant facts can appear.

Relevance is thus already built in. In the horse race case, racing fits into a nested context of activities, games, sports, contests. To see an activity as a horse race is to organize it in terms of the intention to win. To return to Taylor's account:

> The handicapper is concerned to pick a winner. As a human being he has a sense of what is involved in the enterprise of winning, and his being concerned means that he is aware of a horse, jockey, etc., in a way in which dangers are salient. Hence he notices when he reads in the obituary columns that Smith's mother died yesterday (Smith being the jockey, and one he knows to be very susceptible), and for once he bets against the form. The machine would pick out Smith's mother's death, as a fact about Smith, along with all the others, such as that Smith's second cousin has been elected dogcatcher in some other city, etc., but will then have to do a check on the probable consequences of these different facts before it decides to take them into account or not in placing the bet.[9]

Thus our present concerns and past know-how always already determines what will be ignored, what will remain on the outer horizon of experience as possibly relevant, and what will be immediately taken into account as essential.

Wittgenstein constantly suggests that the analysis of a situation into facts and rules (which is where the traditional philosopher and the computer expert think they must begin) is itself only meaningful in some context and for some purpose. Thus again the elements already reflect the goals and purposes for which they were carved out. When we try to find the ultimate context-free, purpose-free elements, as we must if we are going to find the ultimate bits to feed a machine—bits that will be relevant to all possible tasks because chosen for none—we are in effect trying to free the facts in our experience of just that pragmatic organization which makes it possible to use them flexibly in coping with everyday problems.

Not that a computer model is ever really purpose-free; even a model in terms of information storage must somehow reflect the context, but such an analysis of context in terms of facts and rules is rigid and restricting. To see this, let us grant that all the properties of objects (whatever that might mean) could be made explicit in a decision tree so that each node recorded whether the object has a certain situation-independent predicate or its converse. This sort of classification structure has been programmed by Edward Feigenbaum in his EPAM model.[10]* Such a discrimination net might, in principle, represent an exhaustive, explicit, apparently situation-free characterization of an object, or even of a situation, insofar as it was considered as an object. It thus seems to provide efficient information storage, while avoiding the field/object distinction. But something crucial is omitted in the description of such an information structure: the organization of the structure itself, which plays a crucial role in the informative storage. *The information in the tree is differently stored and differently accessible depending on the order in which the discriminations are made.* As William Wynn notes in a discussion of EPAM:

> EPAM's Classification process is . . . too history-dependent and unadaptable, for the discrimination net can be grown only from the bottom down and cannot be reorganized from the top. Tests inserted in the net which later prove to be of little discriminatory power over a given stimulus set cannot be removed, nor can new tests be inserted in the upper portion of the net. Thus, once it is formed, EPAM's discrimination net is difficult to reorganize in the interest of greater retrieval efficiency. Any procedure that reorganizes the tests in the structure seriously impairs retrieval of many items held in the memory.[11]

So the order of discriminations is crucial. But in the physical world all predicates have the same priority. Only the programmer's sense of the situation determines the order in the decision tree. Through the programmer's judgment the distinction between the field and the objects in the field is introduced into the computerized model. The pragmatic context used by the programmer can indeed itself be characterized in a decision tree, but only in some order of discriminations which reflects a broader context. At each level information concerning this broader context is indeed embodied in the general structure of the tree, but at no particular node. At each level the situation is reflected in the pragmatic intuitions of the programmer governing the order of decisions; but this

fixes the facts in one order based on a particular purpose, and inevitably introduces the lack of flexibility noted by Wynn.

If, on the other hand, in the name of flexibility all pragmatic ordering could be eliminated so that an unstructured list of purified facts could be assimilated by machine—facts about the sizes and shapes of objects in the physical world and even about their possible uses, as isolable functions—then all these facts would have to be explicitly included or excluded in each calculation, and the computer would be overwhelmed by their infinity.

This is not to deny that human beings *sometimes* take up isolated data and try to discover their significance by trying to fit them into a previously accumulated store of information. Sherlock Holmes and all detectives do this as a profession; everyone does it when he is in a very unfamiliar situation. But even in these cases there must be some more general context in which we are at home. A Martian might have to proceed in a very unfamiliar context if he were on earth, but if he shared *no* human purposes his task of sorting out the relevant from the irrelevant, essential from the inessential, would be as hopeless as that of the computer.

We all know also what it is to store and use data according to rules in some restricted context. We do this, for example, when we play a game such as bridge, although even here a good bridge player stores data in terms of purpose and strategies and takes liberties with the heuristic rules. We also sometimes play out alternatives in our imagination to predict what will happen in the real game before us. But it is just because we know what it is to have to orient ourselves in a world in which we are not at home; or to follow rulelike operations like the heuristics for bidding in bridge; and how to model in our imagination events which have not yet taken place, that we know that we are not aware of doing this most of the time. The claim that we are nonetheless carrying on such operations unconsciously is either an empirical claim, for which there is no evidence, or an *a priori* claim based on the very assumption we are here calling into question.

When we are at home in the world, the meaningful objects embedded in their context of references among which we live are not a model of the world stored in our mind or brain; *they are the world itself.* This may seem plausible for the public world of general purposes, traffic regulations, and so forth. But what about *my* experience, one may ask; my private set of facts, surely that is in my mind? This seems plausible only because one is still confusing this human world with some sort of physical universe. My personal plans and my memories are inscribed in the things around me just as are the public goals of men in general. My memories are stored in the familiar look of a chair or the threatening air of a street corner where I was once hurt. My plans and fears are already built into my experience of some objects as attractive and others as to be avoided. The "data" concerning social tasks and purposes which are built into the objects and spaces around me are overlaid with these personal "data" which are no less a part of my world. After all, personal threats and attractions are no more subjective than general human purposes.

Now we can see why, even if the nervous system must be understood as a physical object—a sort of analogue computer—whose energy exchange with the world must in principle be expressible as an input/output function, it begs the question and leads to confusion to suppose that on the information-processing level the human perceiver can be understood as an analogue computer having a precise I/O function reproducible on a digital machine. The whole I/O model makes no sense here. There is no reason to suppose that the human world can be analyzed into independent elements, and even if it could, one would not know whether to consider these elements the input or the output of the human mind.

If this idea is hard to accept, it is because this phenomenological account stands in opposition to our Cartesian tradition which thinks of the physical world as impinging on our mind which then organizes it according to its previous experience and innate ideas or rules. But even Descartes is not confused in the way contemporary psychologists and artificial intelligence researchers seem to be. He contends that the world which impinges on us is *a world of pure physical motions,* while the world "in the mind" is the world of objects, instruments, and so forth. Only the relation between these two worlds is unclear. Artificial intelligence

theorists such as Minsky, however, have a cruder picture in which the world of implements does not even appear. As they see it, details of the everyday world—snapshots, as it were, of tables, chairs, etc.—are received by the mind. These fragments are then reassembled in terms of a model built of other facts the mind has stored up. The outer world, a mass of isolated facts, is interpreted in terms of the inner storehouse of other isolated, but well catalogued, facts—which somehow was built up from earlier experiences of this fragmented world—and the result is a further elaboration of this inner model. Nowhere do we find the familiar world of implements organized in terms of purposes.

Minsky has elaborated this computer-Cartesianism into an attempt at philosophy. He begins by giving a plausible description of what is in fact the role of imagination:

> If a creature can answer a question about a hypothetical experiment without actually performing it, then it has demonstrated some knowledge about the world. For, his [sic] answer to the question must be an encoded description of the behavior (inside the creature) of some submachine or "model" responding to an encoded description of the world situation described by the question.[12]

Minsky then, without explanation or justification, generalizes this plausible description of the function of imagination to all perception and knowledge:

> Questions about things in the world are answered by making statements about the behavior of corresponding structures in one's model of the world.[13]

He is thus led to introduce a formalized copy of the external world; as if besides the objects which solicit our action, we need an encyclopedia in which we can look up where we are and what we are doing:

> A man's model of the world has a distinctly bipartite structure: One part is concerned with matters of mechanical, geometrical, physical character, while the other is associated with things like goals, meanings, social matters, and the like.[14]

If all knowledge requires a model we, of course, need a model of ourselves:

> When a man is asked a general question about his own nature, he will try to give a general description of his model of himself.[15]

And, of course, for this self-description to be complete we will need a description of our model of our model of ourselves, and so forth. Minsky thinks of this self-referential regress as the source of philosophical confusions concerning mind, body, free will, and so on. He does not realize that his insistence on models has introduced the regress and that this difficulty is proof of the philosophical incoherence of his assumption that nothing is ever known directly but only in terms of models.

In general the more one thinks about this picture the harder it is to understand. There seem to be two worlds, the outer data- and the inner data-structure, neither of which is ever experienced and neither of which is the physical universe or the world of implements we normally *do* experience. There seems to be no place for the physical universe or for our world of interrelated objects, but only for a library describing the universe and human world which, according to the theory, cannot exist.

To dismiss this theory as incoherent is not to deny that physical energy bombards our physical organism and that the result is our experience of the world. It is simply to assert that the physical processing of the physical energy is not a psychological process, and does not take place in terms of sorting and storing human-sized facts about tables and chairs. Rather, the human world is the result of this energy processing and the human world does not need another mechanical repetition of the same process in order to be perceived and understood.

This point is so simple and yet so hard to grasp for those brought up in the Cartesian tradition that it may be necessary to go over the ground once again, this time returning to a specific case of this confusion. As we have seen, Neisser begins his book *Cognitive Psychology* with an exposition of what he calls "the central problem of cognition."

> There is certainly a real world of trees and people and cars and even books. . . . However, we have no direct, immediate access to the world, nor to any of its properties.[16]

Here, as we have noted in Chapter 4, the damage is already done. There is indeed a world to which we have no immediate access. We do not directly perceive the world of atoms and electromagnetic waves (if it even

makes sense to speak of perceiving them)—but the world of cars and books is just the world we *do* directly experience. In Chapter 4 we saw that at this point, Neisser has recourse to an unjustified theory that we perceive "snapshots" or sense data. His further account only compounds the confusion:

> Physically, this page is an array of small mounds of ink, lying in certain positions on the more highly reflective surface of the paper.[17]

But *physically,* what is there are atoms in motion, not paper and small mounds of ink. Paper and small mounds of ink are elements in the human world. Neisser, however, is trying to look at them in a special way, as if he were a savage, a Martian, or a computer, who didn't know what they were for. There is no reason to suppose that these strangely isolated objects are what men directly perceive (although one may perhaps approximate this experience in the very special detached attitude which comes over a cognitive psychologist sitting down to write a book). What we normally perceive is a printed page.

Again Neisser's middle-world, which is neither the world of physics nor the human world, turns out to be an artifact. No man has ever seen such an eerie world; and no physicist has any place for it in his system. Once we postulate it, however, it follows inevitably that the human world will somehow have to be reconstructed out of these fragments.

> One-sided in their perspective, shifting radically several times each second, unique and novel at every moment, the proximal stimuli bear little resemblance to either the real object that gave rise to them or to the object of experience that the perceiver will construct as a result.[18]

But this whole construction process is superfluous. It is described in terms which make sense only if we think of man as a computer receiving isolated facts from a world in which it has no purposes; programmed to use them, plus a lot of other meaningless data it has accumulated or been given, to make some sort of sense (whatever that might mean) out of what is going on around it.

There is no reason to suppose that a normal human being has this problem, although some aphasics do. A normal person experiences the objects of the world as already interrelated and full of meaning. There is no justification for the assumption that we first experience isolated facts, or snapshots of facts, or momentary views of snapshots of isolated facts, and *then* give them significance. The analytical superfluousness of such a process is what contemporary philosophers such as Heidegger and Wittgenstein are trying to point out. To put this in terms of Neisser's discussion as nearly as sense will allow, we would have to say: "The human world *is* the mind's model of the physical world." But then there is no point in saying it is "in the mind," and no point in inventing a third world—between the physical and the human world—which is an arbitrarily impoverished version of the world in which we live, out of which this world has to be built up again.

Oettinger, alone among computer experts, has seen that in the world of perception and language, where the linguist and artificial intelligence worker begins his analysis, a global meaning is always already present.

> What I want to suggest is not necessarily a novel suggestion; but it does seem to have been lost from sight, perhaps deservedly so, because, as I have pointed out, it doesn't tell one what to do next. What I suggest is that it almost seems as if the perception of meaning were primary and everything else a consequence of understanding meaning.[19]

But Oettinger does not seem to see that if one simply looks for some new sort of process, by which this global meaning is "produced," thereby reversing the current misunderstanding, one is bound to find what seems a mystery or a dead end.

> When we try to turn this around and say, "Well now, here is this stream of sound coming at you or its equivalent on a printed page, and what is it that happens to your listening to me or in reading a printed page that enables you to react to the meaning of what I say?" we seem to hit a dead end at this point.[20]

What Oettinger too fails to understand is that there is *either* a stream of sounds *or* there is meaningful discourse. The meaning is not *produced* from meaningless elements, be they marks or sounds. The stream of sounds is a problem for physics and neurophysiology, while on the level of meaningful discourse, the necessary energy processing has already taken place, and the *result* is a meaningful world for which no *new*

theory of production is required nor can be consistently conceived.

To avoid inventing problems and mysteries we must leave the physical world to the physicists and neurophysiologists, and return to our description of the human world which we immediately perceive. The problem facing contemporary philosophers is to describe the context or situation in which human beings live, without importing prejudices from the history of philosophy or the current fascination with computer models. This brings us back to the problem of regularity and rules.

Our context-guided activity in terms of which we constantly modify the relevance and significance of particular objects and facts is quite regular, but the regularity need not and cannot be completely rule governed. As in the case of ambiguity tolerance, our activity is simply as rule governed as is necessary for the task at hand—the task itself, of course, being no more precise than the rules.

Wittgenstein, like Heidegger, sees the regulation of traffic as paradigmatic:

> The regulation of traffic in the streets permits and forbids certain actions on the part of drivers and pedestrians; but it does not attempt to guide the totality of their movements by prescription. And it would be senseless to talk of an 'ideal' ordering of traffic which would do that; in the first place we should have no idea what to imagine as this ideal. If someone wants to make traffic regulations stricter on some point or other, that does not mean that he wants to approximate to such an ideal.[21]

This contextual regularity, never completely rule governed, but always as orderly as necessary, is so pervasive that it is easily overlooked. Once, however, it has been focused on as the background of problem solving, language use, and other intelligent behavior, it no longer seems necessary to suppose that all ordered behavior is rule governed. The rule-model only seems inevitable if one abstracts himself from the human situation as philosophers have been trying to do for two thousand years, and as computer experts must, given the context-free character of information processing in digital machines.

9

The Situation as a Function of Human Needs

We are at home in the world and can find our way about in it because it is *our* world produced by us as the context of our pragmatic activity. So far we have been describing this world or situation and how it enables us to zero in on significant objects in it. We have also suggested that this field of experience is structured in terms of our tasks. These are linked to goals, and these in turn correspond to the social and individual needs of those whose activity has produced the world.

What does this tell us about the possibility of AI? If the data which are to be stored and accessed are normally organized in terms of specific goals, then it would seem that the large data base problem confronting AI could be solved if one just constructed a list of objectives and their priorities—what computer workers dealing with decision-making programs call a utility function—and programmed it into the computer along with the facts.

We have seen, however, that explicit objectives do not work, even for organizing simple problem-solving programs. The difficulties of simple means-ends analysis suggest that in order for the computer to solve even well-structured problems, it is not sufficient for the machine to have an objective and to measure its progress toward this preset end. Planning requires finding the essential operations, so "pragmatic considerations," for example, the relative importance of logical operations had to be

surreptitiously supplied by the programmers themselves before the logic program could begin. We must now try to describe in more detail how this pragmatic structuring differs from means-ends analysis, ultimately asking, of course, whether this human capacity for purposive organization is in principle programmable on digital machines.

The difference between human goals and machine ends or objectives has been noted by one scientist who has himself been working on pattern recognition. Satosi Watanabe describes this difference as follows:

> For man, an evaluation is made according to a system of values which is nonspecific and quasi-emotive, while an evaluation for a robot could only be made according to a table or a specific criterion. . . . This difference is subtle but profound. [One might say] that a man has values while a machine has objectives. Certainly men too have objectives, but these are derived from a system of values and are not the final arbiter of his actions, as they would be for a robot. . . . As soon as the objective is set the machine can pursue it just as the man can. Likewise human utilitarian behavior can be easily simulated by a machine if the quantitative utility and the probability of each alternative event is fixed and given to the machine. But a machine can never get at the source from which this utility is derived.[1]

Watanabe claims that these values are essential to intelligent behavior. For one thing, as Watanabe points out, "there are infinitely many possible hypotheses that are supported by experience. Limitation of these hypotheses to a smaller subset is often done by a vaguely conceived criterion, such as the principle of simplicity, or the principle of elegance."[2] More specifically, Watanabe argues that it can be demonstrated that any two objects have the same number of predicates in common. If this does not seem to us to be the case, it is because we consider certain predicates more important than others. This decision as to what is important depends on our system of values.[3]

But why on our system of values and not on a list of objectives? How does what Watanabe calls a system of values differ from having a utility function? So far the only difference seems to be that values are vaguer. But throughout Watanabe's analysis there is no argument showing why these values are not just vague objectives which could be represented by a region on a quantitative scale. To understand this important difference, which Watanabe has noted, but not explained, one must first abandon his way of posing the problem. To speak of values already gives away the game. For values are a product of the same philosophical tradition which has laid down the conceptual basis of artificial intelligence. Although talk of values is rather new in philosophy, it represents a final stage of objectification in which the pragmatic considerations which pervade experience and determine what counts as an object are conceived of as just further characteristics of independent objects, such as their hardness or color. A value is one more property that can be added to or subtracted from an object. Once he has adopted this terminology and the philosophical position it embodies, Watanabe is unable to explain how values differ from somewhat vague properties, and thus cannot explain why he feels they cannot be programmed. To understand the fundamental difficulty Watanabe is trying to get at, we must be able to distinguish between objects, and the field or situation which makes our experience of objects possible. For what Watanabe misleadingly calls values belongs to the structure of the *field* of experience, not the objects in it.

We have seen that experience itself is organized in terms of our tasks. Like the pattern of a chess game, the world is a field in which there are areas of attraction and repulsion, paths of accessibility, regions of activity and of repose. In our own perceptual world we are all master players. Objects are already located and recognized in a general way in terms of the characteristics of the field they are in before we zero in on them and concern ourselves with their details. It is only because our interests are *not* objects in our experience that they can play this fundamental role of organizing our experience into meaningful patterns or regions.

Heidegger has described the way human concerns order experiences into places and regions:

> Equipment has its *place* or else it 'lies around': this must be distinguished in principle from just occurring at random in some spacial position. . . . The kind of place which is constituted by direction and remoteness (and closeness is only a mode of the latter) is already oriented towards a region and oriented within it. . . . Thus anything constantly ready-to-hand of which circumspective Being-in-the-World takes account beforehand has its place. The 'where' of its readiness-to-hand is put to account as a matter for concern. . . .[4]

Heidegger is also the first to have called attention to the way philosophy has from its inception been dedicated to trying to turn the concerns in terms of which we live into objects which we could contemplate and control. Socrates was dedicated to trying to make his and other people's commitments explicit so that they could be compared, evaluated, and justified. But it is a fundamental and strange characteristic of our lives that insofar as we turn our most personal concerns into objects, which we can study and choose, they no longer have a grip on us. They no longer organize a field of significant possibilities in terms of which we act but become just one more possibility we can choose or reject. Philosophers thus finally arrived at the nihilism of Nietzsche and Sartre in which personal concerns are thought of as a table of values which are arbitrarily chosen and can be equally arbitrarily abandoned or transvaluated. According to Nietzsche, "The great man is necessarily a skeptic. . . . Freedom from any kind of conviction is part of the strength of his will."[5]*

But what is missing in this picture besides a sense of being gripped by one's commitment? What difference does it make when one is trying to produce intelligent behavior that one's evaluations are based on a utility function instead of some ultimate concern? One difference, which Watanabe notes without being able to explain, is that a table of values must be specific, whereas human concerns only need to be made as specific as the situation demands. This flexibility is closely connected with the human ability to recognize the generic in terms of purposes, and to extend the use of language in a regular but nonrulelike way. Moreover, man's ultimate concern is not just to achieve some goal which is the end of a series; rather, interest in the goal is present at each moment structuring the whole of experience and guiding our activity as we constantly select what is relevant in terms of its significance to the situation at hand.[6] A machine table of objectives, on the other hand, has only an arbitrary relation to the alternatives before the machine, so that it must be explicitly appealed to at predetermined intervals to evaluate the machine's progress and direct its next choice.

Herbert Simon and Walter Reitman have seen that emotion and motivation play some role in intelligent behavior, but their way of simulating this role is to write programs where "emotions" can *interrupt* the work on one problem to introduce extraneous factors or work on some other problem.[7] They do not seem to see that emotions and concerns *accompany* and *guide* our cognitive behavior. This is again a case of not being able to see what one would not know how to program.

Heidegger tries to account for the pervasive concern organizing human experience in terms of a basic human need to understand one's being. But this analysis remains very abstract. It accounts for significance in general but not for any specific goal or specific significance. Thus Heidegger in effect assimilates all human activity to creative problem solving or artistic creation where we do not fully know what our goal was until we have achieved it. For Heidegger there can be no list of specifications which the solution must fulfill. Still, our needs are determinate enough to give things specific meaning for us, and many of our goals are quite explicit. To understand this we require a more concrete phenomenological analysis of human needs.

The philosophical and psychological tradition (with the exception of the pragmatists), however, has tried to ignore the role of these needs in intelligent behavior, and the computer model has reinforced this tendency. Thus N. S. Sutherland, Professor of Experimental Psychology at the University of Sussex, in an article "Machines and Men," writes:

> Survival and self maintenance are achieved by genetically building into the human brain a series of drives or goals. Some of the obvious ones are hunger, thirst, the sexual drive and avoidance of pain. All of these drives are parochial in the sense that one could imagine complex information processing systems exhibiting intelligent behavior but totally lacking them.[8]

We have seen, however, that our concrete bodily needs directly or indirectly give us our sense of the task at hand, in terms of which our experience is structured as significant or insignificant. These needs have a very special structure, which, while more specific than Heidegger's account, does resemble artistic creation. When we experience a need we do not at first know what it is we need. We must search to discover what allays our restlessness or discomfort. This is not found by comparing various objects and activities with some objective, determinate criterion,

but through what Todes calls our sense of gratification. This gratification is experienced as the discovery of what we needed all along, but it is a retroactive understanding and covers up the fact that we were unable to make our need determinate without first receiving that gratification. The original fulfillment of any need is, therefore, what Todes calls a creative discovery.[9]*

Thus human beings do not begin with a genetic table of needs or values which they reveal to themselves as they go along. Nor, when they are authentic, do they arbitrarily adopt values which are imposed by their environment. Rather, in discovering what they need they make more specific a general need which was there all along but was not determinate.

This is most obvious when dealing with less instinctual psychological needs. When a man falls in love he loves a particular woman, but it is not that particular woman he needed *before* he fell in love. However, after he is in love, that is after he has found that this particular relationship is gratifying, the need becomes specific as the need for that particular woman, and the man has made a creative discovery about himself. He has become the sort of person that needs that specific relationship and must view himself as having lacked and needed this relationship all along. In such a creative discovery the world reveals a new order of significance which is neither simply discovered nor arbitrarily chosen.

Sören Kierkegaard has a great deal to say about the way one's personality or self is redefined in such an experience, and how everything in a person's world gets a new level of meaning. Since such a change, by modifying a person's concerns, changes the whole field of interest in terms of which everything gets its significance, Kierkegaard speaks of these fundamental changes as changes in our sphere of existence. And because such a change cannot be predicted on the basis of our previous concerns, yet once it has taken place is so pervasive that we cannot imagine how it could have been otherwise, Kierkegaard speaks of a change of sphere of existence as a leap.[10]

This same sort of change of world can take place on a conceptual level. Then it is called a conceptual revolution. Thomas Kuhn in his book *The Structure of Scientific Revolutions* has studied this sort of transformation. As he puts it: "Insofar as their only recourse to that world is through what they see and do, we may want to say that after a revolution scientists are responding to a different world."[11]

The conceptual framework determines what counts as a fact. Thus during a revolution there are no facts to which scientists can appeal to decide which view is correct. "The data themselves [have] changed. This is the [sense] in which we may want to say that after a revolution scientists work in a different world."[12] The idea that knowledge consists of a large store of neutral data, taken for granted by Minsky, is inadequate to account for these moments of profound change. According to Kuhn, "there can be no scientifically or empirically neutral system of language or concepts."[13]

> What occurs during a scientific revolution is not fully reducible to a reinterpretation of individual and stable data. In the first place the data are not unequivocally stable. A pendulum is not a falling stone, nor is oxygen dephlogisticated air.[14]

This leads Kuhn to a rejection of the whole philosophical tradition which has culminated in the notion of reason as based on the storage and processing of "data." On the basis of his research Kuhn sees both the inadequacy of this tradition and why it nonetheless continues to seem self-evident.

> Are theories simply man-made interpretations of given data? The epistemological viewpoint that has most often guided Western philosophy for three centuries dictates an immediate and unequivocal, Yes! In the absence of a developed alternative, I find it impossible to relinquish entirely that viewpoint. Yet it no longer functions effectively, and the attempts to make it do so through the introduction of a neutral language of observations now seem to me hopeless.[15]

In suggesting an alternative view, or more exactly, in analyzing the way science actually proceeds so as to provide the elements of an alternative view, Kuhn focuses on the importance of a paradigm, that is, a specific accepted example of scientific practice, in guiding research. Here, as in the case of family resemblance studied earlier, objects are understood not in terms of general rules but rather in terms of their relation to a specific concrete case whose traits or implications cannot be completely formalized.

> [Scientists can] agree in their *identification* of a paradigm without agreeing on, or even attempting to produce, a full *interpretation* or *rationalization* of it. Lack of a standard interpretation or of an agreed reduction to rules will not prevent a paradigm from guiding research. . . . Indeed, the existence of a paradigm need not even imply that any full set of rules exist.[16]

It is just this open-ended richness of paradigms which makes them important:

> Paradigms may be prior to, more binding, and more complete than any set of rules for research that could be unequivocally abstracted from them.[17]

Without such paradigms scientists confront the world with the same bewilderment which we have suggested would necessarily confront an AI researcher trying to formalize the human form of life:

> In the absence of a paradigm . . . all of the facts that could possibly pertain to the development of a given science are likely to seem equally relevant.[18]

Indeed, without a paradigm it is not even clear what would count as a fact, since facts are produced in terms of a particular paradigm for interpreting experience. Thus finding a new paradigm is like a Kierkegaardian leap:

> Just because it is a transition between incommensurables, the transition between competing paradigms cannot be made a step at a time, forced by logic and neutral experience. Like the gestalt switch, it must occur all at once (though not necessarily in an instant) or not at all.[19]

Here it becomes clear that the idea of problem solving as simply storing and sorting through data with a specific end in view can never do justice to these fundamental conceptual changes, yet these changes determine the conceptual space in which problems can first be posed and in terms of which data get their pervasive character of relevance and significance, so that problems can be solved. The reigning conceptual framework implicitly guides research just as the perceptual field guides our perception of objects.

Finally, even more fundamental than these conceptual revolutions studied by Kuhn are cultural revolutions; for example, the beginning of Greek philosophy, as we have seen, set up a view of the nature of man and rationality on which all subsequent conceptual revolutions have rung changes. Equally radically, with the beginning of Christianity a new kind of love became possible which was not possible in Greece; heroism became suspect as a sign of pride, and goodness came to consist in the sacrifices of saints. These cultural revolutions show us, as Pascal first pointed out, that there is no sharp boundary between nature and culture—even instinctual needs can be modified and overridden in terms of paradigms—thus there is no fixed nature of man.

Man's nature is indeed so malleable that it may be on the point of changing again. If the computer paradigm becomes so strong that people begin to think of themselves as digital devices on the model of work in artificial intelligence, then, since for the reasons we have been rehearsing, machines cannot be like human beings, human beings may become progressively like machines. During the past two thousand years the importance of objectivity; the belief that actions are governed by fixed values; the notion that skills can be formalized; and in general that one can have a theory of practical activity, have gradually exerted their influence in psychology and in social science. People have begun to think of themselves as objects able to fit into the inflexible calculations of disembodied machines: machines for which the human form-of-life must be analyzed as a meaningless list of facts, rather than the flexible prerational basis of rationality. Our risk is not the advent of superintelligent computers, but of subintelligent human beings.