8

ARTIFICIAL INTELLIGENCE

When Roger Schank expressed the hope that we will be able to build a program that can learn as a child does, he was echoing words spoken by H. A. Simon over ten years earlier:

> "If GPS is a theory of how a machine can bootstrap itself into higher intelligence or how people learn language, then let it bootstrap itself, and let it learn language. This is an entirely appropriate obligation to impose. . . . Not just on behalf of myself, but on behalf of the entire group of people working in the field, I accept the obligation and hope that one of us will produce the requisite programs before too long."[1]

Both Simon and Schank have thus given expression to the deepest and most grandiose fantasy that motivates work on artificial intelligence, which is nothing less than to build a machine on the model of man, a robot that is to have its childhood, to learn language as a child does, to gain its knowledge of the world by sensing the world through its own organs, and ultimately to contemplate the whole domain of human thought. (It is worth noting, though only by the way for now, that should this dream be realized, we will have a language-understanding machine but still no theory of language understanding as such, for observing a machine "learning as a child does" does not in itself constitute an understanding of the language-acquisition process.)

Whether or not this program can be realized depends on whether man really is merely a species of the genus "information-processing system" or whether he is more than that. I shall argue that an entirely too simplistic notion of intelligence has dominated both popular and scientific thought, and that this notion is, in part, responsible for permitting artificial intelligence's perverse grand fantasy to grow. I shall argue that an organism is defined, in large part, by the problems it faces. Man faces problems no machine could possibly be made to face. Man is not a machine. I shall argue that, although man most certainly processes information, he does not necessarily process it in the way computers do. Computers and men are not species of the same genus.

Few "scientific" concepts have so thoroughly muddled the thinking of both scientists and the general public as that of the "intelligence quotient" or "I.Q." The idea that intelligence can be quantitatively measured along a simple linear scale has caused untold harm to our society in general, and to education in particular. It has spawned, for example, the huge educational-testing movement in the United States, which strongly influences the courses of the academic careers of millions of students and thus the degrees of certification they may attain. It virtually determines what "success" people may achieve in later life because, in the United States at least, opportunities to "succeed" are, by and large, open only to those who have the proper credentials, that is, university degrees, professional diplomas, and so on.

When modern educators argue that intelligence tests measure a subject's ability to do well in school, they mean little more than that these tests "predict" a subject's ability to pass academictype tests. This latter ability leads, of course, to certification and then to "success." Consequently, any correlation between the results of such tests and people's "success," as that term is understood in the society at large, must necessarily be an artifact of the testing procedure. The test itself has become a criterion for that with which it is to be correlated! "Psychologists should be ashamed of themselves for promoting a view of general intelligence that has engendered such a testing program."[2]

My concern here is that the mythology that surrounds I.Q. testing has led to the widely accepted and profoundly misleading conviction that intelligence is somehow a permanent, unalterable, and culturally independent attribute of individuals (somewhat like, say, the color of their eyes), and moreover that it may even be genetically transmittable from generation to generation.

The trouble with I.Q. testing is not that it is entirely spurious, but that it is incomplete. It measures certain intellectual abilities that large, politically dominant segments of western European societies have elevated to the very stuff of human worth and hence to the *sine qua non* of success. It is incomplete in two ways: first, in that it fails to take into account that human creativity depends not only on intellect but also crucially on an interplay between intellect and other modalities of thought, such as intuition and wisdom; second, in that it characterizes intelligence as a linearly measurable phenomenon that exists independent of any frame of reference.

Einstein taught us that the idea of motion is meaningless in and of itself, that we can sensibly speak only of an object's motion relative to some frame of reference, not of any *absolute* motion of an object. When, in speaking informally, we say that a train moved, we mean that it moved relative to some fixed point on the earth. We need not emphasize this in ordinary conversation, because the earth (or our body) is to us a kind of "default" frame of reference that is implicitly assumed and understood in most informal conversation. But a physicist speaking as a physicist cannot be so sloppy. His equations of motion must contain terms specifying the coordinate system with respect to which the motion they describe takes place.

So it is with intelligence too. Intelligence is a meaningless concept in and of itself. It requires a frame of reference, a specification of a domain of thought and action, in order to make it meaningful. The reason this necessity does not strike us when we speak of intelligence in ordinary conversation is that the required frame of reference—that is, our own cultural and social setting with its characteristic domains of thought and action—is so much with us that we implicitly assume it to be understood. But our culture and our social milieu are in fact neither universal nor absolute. It therefore behooves us, whenever we use the term "intelligence" as scientists or educators, to make explicit the domain of thought and action which renders the term intelligible.

Our own daily lives abundantly demonstrate that intelligence manifests itself only relative to specific social and cultural contexts. The most unschooled mother who cannot compose a single grammatically correct paragraph in her native language—as, indeed, many academics cannot do in theirs—constantly makes highly refined and intelligent judgments about her family. Eminent scholars confess that they don't have the kind of intelligence required to do high-school algebra. The acknowledged genius is sometimes stupid in managing his private life. Computers perform prodigious "intellectual feats," such as beating champion checker players at their own game and solving huge systems of equations, but cannot change a baby's diaper. How are these intelligences to be compared to one another? They cannot be compared.

Yet forms of the idea that intelligence is measurable along an absolute scale, hence that intelligences are comparable, have deeply penetrated current thought. This idea is responsible, at least in part, for many sterile debates about whether it is possible "in principle" to build computers more intelligent than man. Even as moderate and reasonable a psychologist as George A. Miller occasionally slips up, as when he says, "I am very optimistic about the eventual outcome of the work on machine solution of intellectual problems. Within our lifetime machines may surpass us in general intelligence."[3]

The identification of intelligence with I.Q. has severely distorted the primarily mathematical question of what computers can and cannot do into the nonsensical question of "how much" intelligence one can, again "in principle," give to a computer. And, of course, the reckless anthropomorphization of the computer now so common, especially among the artificial intelligentsia, couples easily to such simpleminded views of intelligence. This joining of an illicit metaphor to an ill-thought-out idea then breeds, and is perceived to legitimate, such perverse propositions as that, for example, a computer can be programmed to become an effective psychotherapist.

I had once hoped that it would be possible to prove that there is a limit, an upper bound, on the intelligence machines could achieve, just as Claude Shannon, the founder of modern information theory, proved that there is an upper bound on the amount of information a given information channel can transmit. Shannon proved that, for example, a specific telephone cable can carry at most a certain number of telephone conversations at any one time. However, before he could even sensibly formulate his now justly famous result, he had to have some way to quantify information. Else how could he speak of a channel's capacity to handle this "much" information but no "more"? Indeed, his design of an information measure itself constitutes an important contribution to modern science. (Given, of course, that he also founded a cogent theory within which his measure plays a decisive role.) It is now clear to me that, since we can speak of intelligence only in specific domains of thought and action, and since these domains are themselves not measurable, we can have no Shannon-like measure of intelligence and therefore no theorem of the kind I had hoped for. In plain words: we may express the wish, even the opinion, that there is a limit to the intelligence machines can attain, but we have no way of giving it precise meaning and certainly no way of proving it.

Does our inability to compute an upper bound on machine intelligence provide grounds either for the "optimistic" conclusion that "machines may surpass us in general intelligence" or for the very same "pessimistic" conclusion?* Neither. We learn instead that any argument that calls for such a conclusion, or for its denial, is itself ill-framed and therefore sterile.

These considerations shed additional light on a question alluded to in Chapter VII (p. 193), where I spoke of "objectives that are inappropriate for machines." Many people would argue that it is not

* The optimist says, "This is the best of all possible worlds!" The pessimist answers, "That's right."

reasonable to speak of machines as having objectives in the first place. But such a rhetorical quibble, if taken seriously, only begs the question, for it ignores the fact that people do in fact delegate responsibility to computers and give them objectives and purposes.

The question I am trying to pursue here is, "What human objectives and purposes may not be appropriately delegated to computers?" We can design an automatic pilot, and delegate to it the task of keeping an airplane flying on a predetermined course. That seems an appropriate thing for machines to do. It is also technically feasible to build a computer system that will interview patients applying for help at a psychiatric out-patient clinic and produce their psychiatric profiles complete with charts, graphs, and natural-language commentary. The question is not whether such a thing *can* be done, but whether it is appropriate to delegate this hitherto human function to a machine.

The artificial intelligentsia argue, as we have seen, that there is no domain of human thought over which machines cannot range. They take for granted that machines can think the sorts of thoughts a psychiatrist thinks when engaged with his patient. They argue that efficiency and cost considerations dictate that machines ought to be delegated such responsibilities. As Professor John McCarthy once put it to me during a debate, "What do judges know that we cannot tell a computer?" His answer to the question—which is really just our question again, only in different form—is, of course, "Nothing." And it is, as he then argued, perfectly appropriate for artificial intelligence to strive to build machines for making judicial decisions.

The proposition that judges and psychiatrists know nothing that we cannot tell computers follows from the much more general proposition subscribed to by the artificial intelligentsia, namely, that there is nothing at all which humans know that cannot, at least in principle, be somehow made accessible to computers.

Not all computer scientists are still so naive as to believe, as they were once charged with believing, that knowledge consists of merely some organization of "facts." The various language-understanding and vision programs, for example, store some of their knowledge in the form of assertions, i.e., axioms and theorems, and other of it in the form of processes. Indeed, in the course of planning and executing some of their complex procedures, these programs compose subprograms, that is, generate new processes, that were not explicitly supplied by human programmers. Some existing computer systems, particularly the so-called hand-eye machines, gain knowledge by directly sensing their environments. Such machines thus come to know things not only by being told them explicitly, but also by discovering them while interacting with the world. Finally, it is possible to instruct computers in certain skills, for example, how to balance a broomstick on one of its ends, by showing them how to do these things even when the instructor is himself quite incapable of verbalizing how he does the trick. The fact, then, and it *is* a fact, that humans know things which they cannot communicate in the form of spoken or written language is not by itself sufficient to establish that there is some knowledge computers cannot acquire at all.

But lest my "admission" that computers have the power to acquire knowledge in many diverse ways be taken to mean more than I intend it to mean, let me make my position very clear:

First (and least important), the ability of even the most advanced of currently existing computer systems to acquire information by means other than what Schank called "being spoon-fed" is still extremely limited. The power of existing heuristic methods for extracting knowledge even from natural-language texts directly "spoonfed" to computers rests precariously on, in Winograd's words, "the tiniest bit of relevant knowledge." It is simply absurd to believe that any currently existing computer system can come to know in any way whatever what, say, a two-year-old child knows about children's blocks.

Second, it is not obvious that all human knowledge is encodable in "information structures," however complex. A human may know, for example, just what kind of emotional impact touching another person's hand will have both on the other person and on himself. The acquisition of that knowledge is certainly not a function of the brain alone; it cannot be simply a process in which an information structure from some source in the world is transmitted to some destination in the brain. The knowledge involved is in part kinesthetic; its acquisition involves having a hand, to say the very least. There are, in other words, some things humans know by virtue of having a human body. No organism that does not have a human body can know these things in the same way humans know them. Every symbolic representation of them must lose some information that is essential for some human purposes.

Third, and the hand-touching example will do here too, there are some things people come to know only as a consequence of having been treated as human beings by other human beings. I shall say more about this in a moment.

Fourth, and finally, even the kinds of knowledge that appear superficially to be communicable from one human being to another in language alone are in fact not altogether so communicable. Claude Shannon showed that, even in abstract information theory, the "information content" of a message is not a function of the message alone but depends crucially on the state of knowledge, on the expectations, of the receiver. The message "Am arriving on 7 o'clock plane, love, Bill" has a different information content for Bill's wife, who knew he was coming home, but not on precisely what airplane, than for a girl who wasn't expecting Bill at all and who is surprised by his declaration of love.

Human language in actual use is infinitely more problematical than those aspects of it that are amenable to treatment by information theory, of course. But even the example I have cited illustrates that language involves the histories of those using it, hence the history of society, indeed, of all humanity generally. And language in human use is not merely functional in the way that computer languages are functional. It does not identify things and words only with immediate goals to be achieved or with objects to be transformed. The human use of language manifests human memory. And that is a quite different thing than the store of the computer, which has been anthropomorphized into "memory." The former gives rise to hopes and fears, for example. It is hard to see what it could mean to say that a computer hopes.

These considerations touch not only on certain technical limitations of computers, but also on the central question of what it means to be a human being and what it means to be a computer.

I accept the idea that a modern computer system is sufficiently complex and autonomous to warrant our talking about it as an organism. Given that it can both sense and affect its environment, I even grant that it can, in an extremely limited sense, be "socialized," that is, modified by its experiences with its world. I grant also that a suitably constructed robot can be made to develop a sense of itself, that it can, for example, learn to distinguish between parts of itself and objects outside of itself, that it can be made to assign a higher priority to guarding its own parts against physical damage than to similarly guarding objects external to itself, and that it can form a model of itself which could, in some sense, be considered a kind of self-consciousness. When I say therefore that I am willing to regard such a robot as an "organism," I declare my willingness to consider it a kind of animal. And I have already agreed that I see no way to put a bound on the degree of intelligence such an organism could, at least in principle, attain.

I make these stipulations, as the lawyers would call them, not because I believe that what any reasonable observer would call a socialized robot is going to be developed in the "visible future"—I do not believe that—but to avoid the unnecessary, interminable, and ultimately sterile exercise of making a catalogue of what computers will and will not be able to do, either here and now or ever. That exercise would deflect us from the primary question, namely, whether there are objectives that are not appropriately assignable to machines.

If both machines and humans are socializable, then we must ask in what way the socialization of the human must necessarily be different from that of the machine. The answer is, of course, so obvious that it makes the very asking of the question appear ludicrous, if indeed not obscene. It is a sign of the madness of our time that this issue has to be addressed at all.

Every organism is socialized by the process of dealing with problems that confront it. The very biological properties that differentiate one species from another also determine that each species will confront problems different from those faced by any other. Every species will, if only for that reason, be socialized differently. The human infant, as many observers have remarked, is born prematurely, that is, in a state of utter helplessness. Yet the infant has biological needs which, if he is to survive at all, must be satisfied by others. Indeed, many studies of orphanages have shown that more than his merely elementary physical needs must be satisfied; an infant will die if he is fed and cleaned but not, from the very beginning of his life, fondled and caressed—if, in other words, he is not treated as a human being by other human beings.[4]

A catastrophe, to use Erik Erikson's expression for it, that every human being must experience is his personal recapitulation of the biblical story of paradise. For a time the infant demands and is granted gratification of his every need, but is asked for nothing in return. Then, often after the infant has developed teeth and has bitten the breast that has fed him, the unity between him and his mother is broken. Erikson believes this universal human drama to be the ontogenetic contribution to the biblical saga of the Garden of Eden. So important is this period in the child's life that

> "a drastic loss of accustomed mother love without proper substitution at this time can lead [under otherwise aggravating conditions] to acute infantile depression or to a mild but chronic state of mourning which may give a depressive undertone to the whole remainder of life. But even under the most favorable circumstances, this stage leaves a residue of a primary sense of evil and doom and of a universal nostalgia for a lost paradise."
>
> [These early stages] "then, form in the infant the springs of the basic sense of trust and the basic sense of mistrust which remain the autogenic source of both primal hope and of doom throughout life."[5]

Thus begins the individual human's imaginative reconstruction of the world. And this world, as I said earlier, is the repository of his subjectivity, the stimulator of his consciousness, and ultimately the constructor of the apparently external forces he is to confront all his life.

> "As the child's radius of awareness, co-ordination, and responsiveness expands, he meets the educative patterns of his culture, and thus learns the basic modalities of human existence, each in personally and culturally significant ways. . . . *To get* . . . means to receive and to accept what is given. This is the first social modality

learned in life; and it sounds simpler than it is. For the groping and unstable newborn organism learns this modality only as it learns to regulate its organ systems in accordance with the way in which the maternal environment integrates its methods of child care. . . .

"The optimum total situation implied in the baby's readiness to get what is given is his mutual regulation with a mother who will permit him to develop and coordinate his means of getting as she develops and co-ordinates her means of giving. . . . The mouth and the nipple seem to be the mere centers of a general aura of warmth and mutuality which are enjoyed and responded to with relaxation not only by these focal organs, but by both total organisms. The mutuality of relaxation thus developed is of prime importance for the first experience of friendly otherness. One may say . . . that in thus *getting what is given,* and in learning to *get somebody to do* for him what he wishes to have done, the baby also develops the necessary ego groundwork *to get to be* a giver."[6]

What these words of Erikson's make clear is that the initial and crucial stages of human socialization implicate and enmesh the totality of two organisms, the child and its mother, in an inseparable mutuality of service to their deepest biological and emotional needs. And out of this problematic reunification of mother and child—problematic because it involves inevitably the trauma of separation—emerge the foundations of the human's knowledge of what it means to give and to receive, to trust and to mistrust, to be a friend and to have a friend, and to have a sense of hope and a sense of doom.

Earlier, when speaking of theories (p. 140), I said that no term of a theory can ever be fully and finally understood. We may say the same thing about words generally, especially about such words as trust and friendship and hope and their derivatives. Erikson teaches us that such words derive their meanings from universal, primal human experiences, and that any understanding of them must always be fundamentally metaphoric. This profound truth also informs us that man's entire understanding of his world, since it is mediated by his language, must always and necessarily be bounded by metaphoric descriptions. And since the child "meets the educative patterns of his culture," as Erikson says, "and thus learns the basic modalities of human existence, each in personally and culturally significant ways," each culture, indeed, each individual in a culture, understands such words and language, hence the world, in a culturally and personally idiosyncratic way.

I could go on to describe the later stages of the socialization of the individual human, the effects of schooling, marriage, imprisonment, warfare, hates and loves, the experiences of shame and guilt that vary so radically among the cultures of man, and so on. But that could be of no help to anyone who is not already convinced that any "understanding" a computer may be said to possess, hence any "intelligence" that may be attributed to it, can have only the faintest relation to human understanding and human intelligence. We, however, conclude that however much intelligence computers may attain, now or in the future, theirs must always be an intelligence *alien* to genuine human problems and concerns.

Still, the extreme or hardcore wing of the artificial intelligentsia will insist that the whole man, to again use Simon's expression, is after all an information processor, and that an information-processing theory of man must therefore be adequate to account for his behavior in its entirety. We may agree with the major premise without necessarily drawing the indicated conclusion. We have already observed that a portion of the information the human "processes" is kinesthetic, that it is "stored" in his muscles and joints. It is simply not clear that such information, and the processing associated with it, can be represented in the form of computer programs and data structures at all.

It may, of course, be argued that it is in principle possible for a computer to simulate the entire network of cells that constitutes the human body. But that would introduce a theory of information processing entirely different from any which has so far been advanced. Besides, such a simulation would result in "behavior" on such an incredibly long-time scale that no robot built on such principles could possibly interact with human beings. Finally, there appears to be no prospect whatever that mankind will know enough neurophysiology within the next several hundred years to have the intellectual basis for designing such a machine. We may therefore dismiss such arguments.

There is, however, still another assumption that information-processing modelers of man make that may be false, and whose denial severely undermines their program: that there exists one and only one class of information processes, and that every member of that class is reducible to the kind of information processes exemplified by such systems as GPS and Schank-like language-understanding formalisms. Yet every human being has the impression that he thinks at least as much by intuition, hunch, and other such informal means as he does "systematically," that is by means such as logic. Questions like "Can a computer have original ideas? Can it compose a metaphor or a symphony or a poem?" keep cropping up. It is as if the folk wisdom knows the distinction between computer thought and the kind of thought people ordinarily engage in. The artificial intelligentsia, of course, do not believe there need be any distinction. They smile and answer "unproven."

Within the last decade or so, however, neurological evidence has begun to accumulate that suggests there may be a scientific basis to the folk wisdom.[7] It has long been known that the human brain consists of two so-called hemispheres that appear, superficially at least, to be identical. These two halves, which we will call LH (Left Hemisphere) and RH (Right Hemisphere), have, however, quite distinct functions. In righthanded people—and for simplicity, we can restrict our discussion to them—the LH may be said, at least roughly, to control the right half of the body, and the RH the left half. (Actually, the connectivities are somewhat more complex, particularly between the two brain halves and the eyes, but I will not go into such details here.) Most importantly, the two halves of the brain appear to have two quite distinct modalities of thought. The LH thinks, so to speak, in an orderly, sequential, and, we might call it, logical fashion. The RH, on the other hand, appears to think in terms of holistic images. Language processing appears to be almost exclusively centered in the LH, for example, whereas the RH is deeply involved in such tasks as spatial orientation, and the production and appreciation of music.

The distinct functions of the hemispheres of the brain began to be dramatically illustrated by patients who, after suffering from extremely severe forms of epilepsy, had their two brain halves surgically separated. In normal people, the two hemispheres are connected by a part of the brain called the corpus callosum. When this is cut, no direct communication between the two halves remains possible. It was found that when a so-called split-brain patient's hands were visually hidden from him and he was given, say, a pencil in his left hand, he could not *say* what had been given to him, but he could *show* that it was a pencil by drawing a picture of it or by selecting a picture of a pencil from among pictures of many different objects. However, when the experiment was repeated, only with the right hand receiving the pencil, then he could say it was a pencil but could not produce or recognize its pictorial representation. In the first situation, the RH received the "image" of the pencil and was able to encode it into pictorial representations, but not into linguistic structures. In the second, the LH received the "image" of the pencil and was able to encode it linguistically, but not pictorially.

There is also considerable evidence, which I will not detail here, that the RH is essentially the seat of intuition, and that it thinks quite independently of the LH. One way of characterizing intuitive thought is to say that, although it is logical, the standards of evidence it uses to make judgments are very different from the standards we normally associate with logical thought. In ordinary discourse, for example, when we say that two things are the same, we mean that they are identical in almost every respect; the standard of evidence we demand to justify such a judgment is extremely demanding. But when we construct a metaphor, e.g., the overseas Chinese are the Jews of the Orient, we pronounce two things to be the same in a very different sense. Metaphors are simply not logical; when taken literally, they are patently absurd. The RH, in other words, has criteria of absurdity that are far different from those of the logical LH.

The history of man's creativity is filled with stories of artists and scientists who, after working hard and long on some difficult problem, consciously decide to "forget" it, in effect, to turn it over to their RH. After some time, often with great suddenness and totally unexpectedly, the solution to their problem announces itself to them in almost complete form. The RH appears to have been able to overcome the most difficult logical and systematic problems by, I would conjecture, relaxing the rigid standards of thought of the LH. Given the looser standards the RH employs, it was perhaps able to design thought experiments which the LH simply could not, because of its rigidity, conceive. The RH is thus able to hit upon solutions which could then, of course, be recast into strictly logical terms by the LH. We may conjecture that in children the communication channel between the two brain halves is wide open; that is, that messages pass between the two halves quite freely. That may be why children are so incredibly imaginative; e.g., for them a cigar box is an automobile one moment and a house the next. In adults, the channel has been severely narrowed—whether by education or by physiological maturational processes or by both, I cannot guess. But it is clearly more open during the dream state. I may also conjecture that psychoanalysis, quite apart from its function as psychotherapy, trains people in the use of the channel. In psychoanalysis one learns, in Theodore Reik's happy phrase, to listen with the third ear, to attend, that is, to what the unconscious is "saying." Perhaps the various meditative disciplines serve the same purpose.

These are clearly conjectures, from which we are not entitled to draw any conclusions about how either humans or computers process information. Even as a mere possibility, however, they do raise a serious question about the universality of the mode of information processing we normally associate with logical thought and with computer programs.

That the right hemisphere of the brain is, loosely speaking, the "seat of intuition" is a hypothesis in favor of which evidence appears to be accumulating. Neither philosophers nor psychologists have yet been sufficiently persuaded by the existing evidence to confidently incorporate this hypothesis into their theories of mind. But this much is firmly established: the two hemispheres of the human brain think independently of one another; they think simultaneously; and they think in modes different from one another. Furthermore, we can say something about these two distinct modes.

The great mathematician Henri Poincaré, in his celebrated essay *Mathematical Creation*,[8] wrote

> "The conscious self is narrowly limited, and as for the subliminal self we know not its limitations. . . . calculations . . . must be made

> in the . . . period of conscious work, that which follows the inspiration, that in which one verifies the results of this inspiration and deduces their consequences. The rules of these calculations are strict and complicated. They require discipline, attention, will, and therefore consciousness. In the subliminal self, on the contrary, reigns what I should call liberty, if we might give this name to the simple absence of discipline. . . . the privileged unconscious phenomena, those susceptible of becoming conscious, are those which, directly or indirectly, affect most profoundly our emotional sensibility. . . . The role of this unconscious work in mathematical invention appears to me incontestable, and traces of it would be found in other cases where it is less evident."

Of course, Poincaré, writing at the beginning of the twentieth century, knew nothing of the findings of the now-active brain researchers. And we are jumping to a conclusion when we identify what he calls the conscious and the subliminal selves with the left and right hemispheres of the brain, respectively. But our assertion here is that there are two distinct modes of human thought that operate independently and simultaneously. And that assertion Poincaré supports.

A most highly respected scientist who is now working, the psychologist Jerome Bruner, writes on this same topic from a slightly different perspective (recall that the right hand corresponds to the left hemisphere and the left hand to the right, or "intuitive," hemisphere):

> "As a right-handed psychologist, I have been diligent for fifteen years in the study of the cognitive processes: how we acquire, retain, and transform knowledge of the world in which each of us lives—a world in part 'outside' us, in part 'inside.' The tools I have used have been those of the scientific psychologist studying perception, memory, learning, thinking, and (like a child of my times) I have addressed my inquiries to the laboratory rat as well as to human beings. At times, indeed, I have adopted the role of the clinician and carried out therapy with children. . . . There have been times when, somewhat discouraged by the complexities of the psychology of knowing, I have sought to escape through neurophysiology, to discover that the neurophysiologist can help only

in the degree to which we can ask intelligent psychological questions of him.

"One thing has become increasingly clear in pursuing the nature of knowing. It is that the conventional apparatus of the psychologist—both his instruments of investigation and the conceptual tools he uses in the interpretation of his data—leaves one approach unexplored. It is an approach whose medium of exchange seems to be the metaphor paid out by the left hand. It is a way that grows happy hunches and 'lucky' guesses, that is stirred into connective activity by the poet and the necromancer looking sidewise rather than directly. Their hunches and intuitions generate a grammar of their own—searching out connections, suggesting similarities, weaving ideas loosely in a trial web. . . .

"[The psychologist] too searches widely and metaphorically for his hunches. He reads novels, looks at and even paints pictures, is struck by the power of myth, observes his fellow men intuitively and with wonder. In doing so, he acts only part-time like a proper psychologist, racking up cases against the criteria derived from hypothesis. Like his fellows, he observes the human scene with such sensibility as he can muster in the hope that his insight will be deepened. If he is lucky or if he has subtle psychological intuition, he will from time to time come up with hunches, combinatorial products of his metaphoric activity. If he is not fearful of these products of his own subjectivity, he will go so far as to tame the metaphors that have produced the hunches, tame them in the sense of shifting them from the left hand to the right hand by rendering them into notions that can be tested. It is my impression from observing myself and my colleagues that the forging of metaphoric hunch into testable hypothesis goes on all the time."[9]

That, of course, is my impression as well. Here Bruner speaks explicitly of the left hand, that is, the right hemisphere of the brain, as the artistic, the intuitive, and so on, and of the right hand, the left brain hemisphere, as the "conventional apparatus of the psychologist," and he speaks of the inadequacy of his "conceptual tools."

We learn from the testimony of hundreds of creative people, as well as from our own introspection, that the human creative act always involves the conscious interpretation of messages coming from the unconscious, the shifting of ideas from the left hand to the right, in Bruner's phrase.

The unconscious is, of course, unconscious. It is like a seething, stormy sea within us. Its waves lap on the borders of our consciousness. And what we learn from it or about it, we construct from inferences we make about the meanings of the swells and surges, the breakers and ripples that wash the fringes of our consciousness. Occasionally we wander more deeply into the surf, as when we are in that semi-hypnagogic trance that divides sleep from wakefulness. But then we experience only chaos. Our thought modalities are maximally confused. And if we rip ourselves into waking, we cannot tell, we cannot translate or transform into linguistic modalities, what we had thought.

Does not the undoubted reality of this confusion, when placed alongside all the other available evidence to which I have alluded, lend weight to the altogether plausible conjecture that the forms of information manipulated in the right hemisphere of the brain, as well as the corresponding information processes, are simply different from those of the left hemisphere? And may it not be that we can in principle come to know those strange information forms and processes only in terms that are fundamentally irrelevant to the kind of understanding we seek? When, in the distant future, we come to know in detail how the brain functions on the neurophysiological level, we will, of course, be able to give an ultimately reductionist account of the functioning of the right hemisphere. But that would not be understanding in the sense we mean here, anymore than detailed knowledge of the electrical behavior of a running computer is, or even leads to, an understanding of the program the computer is running. On the other hand, a higher-level account of the functioning of the right hemisphere may always miss its most essential features, namely, those that differentiate it from the functioning of the left hemisphere. For we are constrained by our left-hemisphere thought modalities to always interpret messages coming from the right in left-hemisphere terms.

Perhaps the LH modality of thought is GPS-like, which is to say only that perhaps it can in principle be somehow formulated (not that GPS is even a candidate for a possible formulation). Perhaps it converts a problem like

> Tom has twice as many fish as Mary has guppies. If Mary has three guppies, how many fish does Tom have?

into its own terms, for example into

$$x = 2y; y = 3,$$

and solves it using information processing and symbol-manipulation techniques characteristic of GPS-like "thought." But it is then not possible for such a mechanism to have any idea of what fishes and guppies are, or of what it can mean to be a boy named Tom, and so on. Nor can the symbolic representation of the given problem be reconverted into the original problem statement. But human problem solving, perhaps even of the apparently most routine and mechanical variety, involves both left and right modes of thought. And certainly, direct human communication crucially involves the two hemispheres.

It is much too easy, especially for computer scientists, to be hypnotized by the "fact" that linguistic utterances are representable as linear strings of symbols. From this "fact" it is easy to deduce that linguistic communication is entirely a left-hemisphere affair. But human speech also has melody, and its song communicates as well as its libretto. Music is the province of the right hemisphere, as is the appreciation of gestures. As for written communication, its function is surely, at least in large part, to stimulate and excite especially the auditory imaginations of both the writer and the reader.

We may never know whether the conjecture that a part of us thinks in terms of symbolic structures that can be only sensed but not usefully explicated is true or false. Scientists, of course, abhor hypotheses that appear not to be falsifiable. Yet it may be that, under some profound conception of truth, the hypothesis is true. Perhaps it helps to explain why we remain lifelong strangers to ourselves and to each other, why every word in our lexicon is enveloped in at least some residual mystery, and why every attempt to solve life's problems by entirely rational means always fails.

But the inference that I here wish to draw from my conjecture is that, since we cannot know that it is false any more than that it is true, we are not entitled to the hubris so bombastically exhibited by the artificial intelligentsia. Even calculating reason compels the belief that we must stand in awe of the mysterious spectacle that is the whole man—I would even add, that is the whole ant.

There was a time when physics dreamed of explaining the whole of physical reality in terms of one comprehensive formalism. Leibnitz taught that if we knew the position and velocity of every elementary particle in the universe, we could predict the universe's whole future course. But then Werner Heisenberg proved that the very instruments man must use in order to measure physical phenomena disturb those phenomena, and that it is therefore impossible in principle to know both the exact position and the velocity of even a single elementary particle. He did not thereby falsify Leibnitz's conjecture. But he did show that its major premise was unattainable. That, of course, was sufficient to shatter the Leibnitzian dream. Only a little later, Kurt Gödel exposed the shakiness of the foundations of mathematics and logic itself by proving that every interesting formal system has some statements whose truth or falsity cannot be decided by the formal means of the system itself, in other words, that mathematics must necessarily be forever incomplete. It follows from this and others of Gödel's results that "The human mind is incapable of formulating (or mechanizing) all its mathematical intuitions. I.e.: If it has succeeded in formulating some of them, this very fact yields new intuitive knowledge."[10]

Both Heisenberg's so-called uncertainty principle and Gödel's incompleteness theorem sent terrible shock-waves through the worlds of physics, mathematics, and philosophy of science. But no one stopped working. Physicists, mathematicians, and philosophers more or less gracefully accepted the undeniable truth that there are limits to how far the world can be comprehended in Leibnitzian terms alone.

Much too much has already been made of the presumed implications of Heisenberg's and Gödel's results for artificial intelligence. I do not wish to contribute to that discussion here. But there is a sense in which psychology and artificial intelligence may usefully follow the example of the new-found humility of modern mathematics and physics: they should recognize that "while the constraints and limitations of logic do not exert their force on the things of the world, they do constrain and limit what are to count as defensible descriptions and interpretations of things."[11] Were they to recognize that, they could then take the next liberating step of also recognizing that truth is not equivalent to formal provability.

The lesson I have tried to teach here is not that the human mind is subject to Heisenberg uncertainties—though it may be—and that we can therefore never wholly comprehend it in terms of the kinds of reduction to discrete phenomena Leibnitz had in mind. The lesson here is rather that the part of the human mind which communicates to us in rational and scientific terms is itself an instrument that disturbs what it observes, particularly its voiceless partner, the unconscious, between which and our conscious selves it mediates. Its constraints and limitations circumscribe what are to constitute rational—again, if you will, scientific—descriptions and interpretations of the things of the world. These descriptions can therefore never be whole, anymore than a musical score can be a whole description or interpretation of even the simplest song.

But, and this is the saving grace of which an insolent and arrogant scientism attempts to rob us, we come to know and understand not only by way of the mechanisms of the conscious. We are capable of listening with the third ear, of sensing living truth that is truth beyond any standards of provability. It is *that* kind of understanding, and the kind of intelligence that is derived from it, which I claim is beyond the abilities of computers to simulate.

We have the habit, and it is sometimes useful to us, of speaking of man, mind, intelligence, and other such universal concepts. But gradually, even slyly, our own minds become infected with what A. N. Whitehead called the fallacy of misplaced concreteness. We come to believe that these theoretical terms are ultimately interpretable as observations, that in the "visible future" we will have ingenious instruments capable of measuring the "objects" to which these terms refer. There is, however, no such thing as mind; there are only individual minds, each belonging, not to "man," but to individual human beings. I have argued that intelligence cannot be measured by ingeniously constructed meter sticks placed along a one-dimensional continuum. Intelligence can be usefully discussed only in terms of domains of thought and action. From this I derive the conclusion that it cannot be useful, to say the least, to base serious work on notions of "how much" intelligence may be given to a computer. Debates based on such ideas—e.g., "Will computers ever exceed man in intelligence?"—are doomed to sterility.

I have argued that the individual human being, like any other organism, is defined by the problems he confronts. The human is unique by virtue of the fact that he must necessarily confront problems that arise from his unique biological and emotional needs. The human individual is in a constant state of becoming. The maintenance of that state, of his humanity, indeed, of his survival, depends crucially on his seeing himself, and on his being seen by other human beings, as a human being. No other organism, and certainly no computer, can be made to confront genuine human problems in human terms. And, since the domain of human intelligence is, except for a small set of formal problems, determined by man's humanity, every other intelligence, however great, must necessarily be alien to the human domain.

I have argued that there is an aspect to the human mind, the unconscious, that cannot be explained by the information-processing primitives, the elementary information processes, which we associate with formal thinking, calculation, and systematic rationality. Yet we are constrained to use them for scientific explanation, description, and interpretation. It behooves us, therefore, to remain aware of the poverty of our explanations and of their strictly limited scope. It is wrong to assert that any scientific account of the "whole man" is possible. There are some things beyond the power of science to fully comprehend.

The concept of an intelligence alien to certain domains of thought and action is crucial for understanding what are perhaps the most important limits on artificial intelligence. But that concept applies to the way humans relate to one another as well as to machines and their relation to man. For human socialization, though it is grounded in the biological constitution common to all humans, is strongly determined by culture. And human cultures differ radically among themselves. Countless studies confirm what must be obvious to all but the most parochial observers of the human scene: "The influence of culture is universal in that in some respects a man learns to become like all men; and it is particular in that a man who is reared in one society learns to become in some respects like all men of his society and not like those of others."[12] The authors of this quotation, students of Japanese society who lived among the Japanese for many years, go on to make the following observations:

> "In normal family life in Japan there is an emphasis on interdependence and reliance on others, while in America the emphasis is on independence and self-assertion. . . . In Japan the infant is seen more as a separate biological organism who from the beginning, in order to develop, needs to be drawn into increasingly interdependent relations with others. In America, the infant is seen more as a dependent biological organism who, in order to develop, needs to be made increasingly independent of others.
>
> "The Japanese baby seems passive, and he lies quietly with occasional unhappy vocalizations, while his mother, in her care, does more lulling, carrying, and rocking of her baby. She seems to try to soothe and quiet the child, and to communicate with him physically rather than verbally. On the other hand, the American infant is more active, happily vocal, and exploring of his environment, and his mother, in her care, does more looking at and chatting to her baby. She seems to stimulate the baby to activity and to vocal response. It is as if the American mother wanted to have a vocal, active baby, and the Japanese mother wanted to have a quiet, contented baby. In terms of styles of caretaking of the mothers in the two cultures, they get what they apparently want. . . . a great deal of cultural learning has taken place by three-to-four months of age. . . . babies have learned by this time to be Japanese and American babies in relation to the expectations of their mothers concerning their behavior.
>
> "[Adult] Japanese are more 'group' oriented and interdependent in their relations with others, while Americans are more 'individ-

ual' oriented and independent. . . . Japanese are more self-effacing and passive in contrast to Americans, who appear more self-assertive and aggressive. . . . Japanese are more sensitive to, and make conscious use of, many forms of nonverbal communication in human relations through the medium of gestures and physical proximity in comparison with Americans, who predominantly use verbal communication within a context of physical separateness.

"If these distinct patterns of behavior are well on the way to being learned by three-to-four months of age, and if they continue over the life span of the person, then there are very likely to be important areas of difference in emotional response in people of one culture when compared with those in another. Such differences are not easily subject to conscious control and, largely out of awareness, they accent and color human behavior. These differences . . . can also add to bewilderment and antagonism when people try to communicate across the emotional barriers of culture."[13]

Such profound differences in early training crucially affect the entire societies involved. And they are, of course, transmitted from one generation to the next and thus perpetuated. They must necessarily also help determine what members of the two societies know about their worlds, what are to be taken as "universal" cultural norms and values, hence what in each culture is and is not to be counted as fact. They determine, for example (and this is particularly relevant to the contrast between Japanese and American social norms), what are private as opposed to public conflicts, and hence what modes of adjudication are appropriate to the defense of what human interests. The Japanese traditionally prefer to settle disputes, even those for which relief at law is statutorily available, by what Westerners would see as informal means. Actually, these means are most often themselves circumscribed by stringent ritualistic requirements that are nowhere explicitly codified but are known to every Japanese of the appropriate social class. This sort of knowledge is acquired with the mother's milk and through the whole process of socialization that is itself so intimately tied to the individual's acquisition of his mother tongue. It cannot be learned from books; it cannot be explicated in any form but life itself.

An American judge, therefore, no matter what his intelligence and fairmindedness, could not sit in a Japanese family court. His intelligence is simply alien to the problems that arise in Japanese culture. The United States Supreme Court actively recognized this while it still had jurisdiction over distant territories. For example, in the case of Diaz v. Gonzales, which was originally tried in Puerto Rico, the court refused to set aside the judgment of the court of original jurisdiction, that is, of the native court. Justice Oliver W. Holmes, writing the opinion of the Court, stated,

> "This Court has stated many times the deference due to understanding of the local courts upon matters of purely local concern. This is especially true when dealing with the decisions of a Court inheriting and brought up in a different system from that which prevails here. When we contemplate such a system from the outside it seems like a wall of stone, every part even with all the others, except so far as our own local education may lead us to see subordinations to which we are accustomed. But to one brought up within it, varying emphasis, tacit assumptions, unwritten practices, a thousand influences gained only from life, may give to the different parts wholly new values that logic and grammar never could have got from the books."[14]

Every human intelligence is thus alien to a great many domains of thought and action. There are vast areas of authentically human concern in every culture in which no member of another culture can possibly make responsible decisions. It is not that the outsider is unable to decide at all—he can always flip coins, for example—it is rather that the *basis* on which he would have to decide must be inappropriate to the context in which the decision is to be made.

What could be more obvious than the fact that, whatever intelligence a computer can muster, however it may be acquired, it must always and necessarily be absolutely alien to any and all authentic human concerns? The very asking of the question, "What does a judge (or a psychiatrist) know that we cannot tell a computer?" is a monstrous obscenity. That it has to be put into print at all, even for the purpose of exposing its morbidity, is a sign of the madness of our times.

Computers can make judicial decisions, computers can make psychiatric judgments. They can flip coins in much more sophisticated ways than can the most patient human being. The point is that they *ought* not be given such tasks. They may even be able to arrive at "correct" decisions in some cases—but always and necessarily on bases no human being should be willing to accept.

There have been many debates on "Computers and Mind." What I conclude here is that the relevant issues are neither technological nor even mathematical; they are ethical. They cannot be settled by asking questions beginning with "can." The limits of the applicability of computers are ultimately statable only in terms of oughts. What emerges as the most elementary insight is that, since we do not now have any ways of making computers wise, we ought not now to give computers tasks that demand wisdom.