

Natural Attitudes Towards Artificial Intelligence

Nathaniel S. Borenstein
Religious Studies 397 - Independent Project
December, 1979

"Every generation has a definition of man it deserves. But it seems to me that we of this generation have fared worse than we deserve. Accepting a definition is man's way of identifying himself, holding up a mirror in which to scan his own face. It is characteristic of the inner situation of contemporary man that the plausible way to identify himself is to see himself in the image of a machine. 'The human machine' is today a more acceptable description of man than the human animal. Man is simply 'a machine into which we put what we call food and produce what we call thought.' A human being is 'an ingenious assembly of portable plumbing.' The definition itself goes back to the eighteenth century. Never before, however, has it been so widely accepted as plausible. An animal stands before us as a mystery; a machine is an invention."

-A. J. Heschel [1]

"Why does man have faith in reason? Because reason has a legitimate function to fulfil, for which it is perfectly adapted; and this is to justify and illumine for man his various experiences and to give him faith and conviction in holding on to the enlarging of his consciousness."

-Sri Aurobindo Ghose [2]

"There is a greater gap between the scientist as a scientist and the same scientist as a human being than the gap between the scientists and the non-science-educated members of the human family."

-R. Buckminster Fuller [3]

With the rapid growth of the uses and capabilities of computers has arisen a branch of Computer Science known as Artificial Intelligence. Its existence goes unnoticed by a large portion of the intellectual world, yet it is, both by its assumptions and achievements, posing enormous challenges to Western forms of self-conception, or at least to the most prevalent Western interpretations of the nature of Man.

Artificial Intelligence seeks ultimately to duplicate and perhaps surpass human intelligence with computer intelligence. There are few claims that such an achievement is imminent, and no proof that it is even possible. Still, enormous progress is being made in a variety of ways, and, as a working hypothesis, the notion that all human endeavor can be simulated by a machine has met with a great deal of success.

This paper will attempt, on a non-technical level, to explore some of the crucial questions posed by such continuing successes and the inevitable concurrent demystification of Man. We shall begin by examining the human force that brings about the attempt to create a mechanical likeness of man, and the varying perspectives that it may offer or imply regarding the ultimate meaning of such an attempt.

The Desire to Transcend

"...real wisdom is likely to dwell with the higher aspiration rather than with the denial of aspiration or with the hope that limits and circumscribes itself within those narrow walls of apparent possibility which are only our intermediate houses of training." -Sri Aurobindo Ghose [4]

"In all acts he performs, man raises a claim to meaning." -A. J. Heschel [5]

"I have stressed that a mechanistic theory does not entail reducing humans to machines of a dull sort, but rather extends our understanding of what mechanism can be. Given that there are many humans who lavish immense care and attention to keep their cars in perfect condition, but ignore their children (or beat them when they can't ignore them), one might even suggest that if people were to view each other 'only' as intricate mechanisms they would be far kinder to each other! In any case, one need love another no less for appreciating something of the bodily and neural mechanisms which underlie their charm." -Michael A. Arbib [6]

Each of us has, to a greater or lesser degree, hopes and dreams that keep us ^{and} alive. The doctor who dreams of curing cancer, the politician who dreams of being President, and the laborer who dreams of going to Las Vegas, all share this common bond of dissatisfaction with an aspect of present reality and hope of a future that is somehow more desirable. Such aspiration is a manifestation of a deeply-rooted human desire, the desire to transcend. But to transcend what? Abstracted in this way, the verb cries out for an object. Often it is felt that Man wishes to transcend all that is finite, and thus to achieve some sort of communion with the Infinite. This of course presupposes

the existence of the Infinite, and not all are willing to grant this presupposition. A more cynical perspective on the deepest human aspirations is presented by the novelist Kurt Vonnegut:

"Once upon a time on Tralfamadore there were creatures who weren't anything like machines. They weren't dependable. They weren't efficient. They weren't predictable. They weren't durable. And these poor creatures were obsessed by the idea that everything that existed had to have a purpose, and that some purposes were higher than others.

"These creatures spent most of their time trying to figure out what their purpose was. And every time they found out what seemed to be a purpose of themselves, the purpose seemed so low that the creatures were filled with disgust and shame.

"And, rather than serve such a low purpose, the creatures would make a machine to serve it. This left the creatures free to serve higher purposes. But whenever they found a higher purpose, the purpose still wasn't high enough.

"So machines were made to serve higher purposes, too.

"And the machines did everything so expertly that they were finally given the job of finding out what the highest purpose of the creatures could be.

"The machines reported in all honesty that the creatures couldn't really be said to have any purpose at all.

"The creatures thereupon began slaying each other, because they hated purposeless things above all else." [7]

Machines are a favorite symbol for that aspect of humanity which we desire to transcend. Almost by definition, we desire to transcend those aspects of ourselves which resemble or may be imitated by machines. What, if anything, remains when all of this has been

transcended?

"Learning is a movement not anchored in knowledge. If it is anchored it is not a movement. The machine, the computer, is anchored. That is the basic difference between man and the machine... Our education is the gathering of a volume of knowledge, and the computer does this faster and more accurately. What need is there for such an education? The machines are going to take over most of the activities of man. When you say, as people do, that learning is the gathering of a volume of knowledge then you are denying, aren't you, the movement of life, which is relationship and behavior?"

-J. Krishnamurti [8]

By this line of thought, what remains after transcendence of the finite is relationship; it is held that something occurs within a relationship which is in no way reducible to a series of externally or subjectively discernable facts. This, like Buber's "I-Thou", then is considered the essence of being human.

Of course this belief, while comforting, remains a belief and not an argument; we are given nothing but intuition to convince us of the existence of a "relationship" which can not be duplicated by machine. Alternately, one can choose to believe that man's saving grace will be the creative 'spark' he nurtures but could never pass on fully to a machine. As John Kemeny of Dartmouth expresses it,

"Man should set the goals and tell the computer how to work toward them. It is best for man to monitor the work of the computer so that he may use his powers of intuition and evaluation to guide it in its work. In short, while 99.99 percent of the work will be done by the computer, the 1/100 percent assigned to human beings is an essential contribution to the partnership." [9]

At the heart of such speculation is the question of destiny. The human desire to transcend is essentially a faith in a destiny beyond the daily struggle for existence and comfort. In Vonnegut's story, the machine is the means by which Man may find out that such faith is illusion. To Krishnamurti, the usurping of man's functions by machines can help to clear the muddy waters that obscure Man's vision of his destiny at every moment. To others, the machine is man's destiny; the computer is the next step in the evolution of intelligence. [*] Arthur C. Clarke put it simply: "It may be that our role on this planet is not to worship God-- but to create Him." [10] In contrast to such sweeping predictions, the computer scientists' own pronouncements seem relatively restrained:

"Will the computer 'take over'? In the world of information-handling of course the computer will take over. The question is will it take over as servant or master? To this one must reply: not as servant nor as master, but as tutor, as secretary, as playmate, as research assistant. None of these in their human embodiments is a servant or a master; each is better described as a helper." [11]

But towards what we are being helped, alas, we are not told.

So we see that the very idea of Artificial Intelligence

[*] A superficial reading of Aurobindo's works (especially The Future Evolution of Man) might seem to support such a claim, but actually Aurobindo had extremely firm feelings about the limitations of reason, which will be mentioned later.

strikes a responsive chord in the man who wishes to find his destiny. Unfortunately, it seems to inspire him as easily to Vonnegut's despair and emptiness as to Krishnamurti's sense of the imminence of mystical vision. This is, perhaps, because Artificial Intelligence is not really a new idea at all, but rather a distilled and purified manifestation of an ancient longing, made newly credible by scientific discovery.

As stated before, the desire to simulate intelligence is the desire to transcend the merely human, either by surpassing it or by relegating its duties to someone else. Attempts to do this have been made since long before the alchemists struggled to turn lead into gold; the Jewish tales of the Golem are a case in point. [12] So far, we have considered this as a positive thing; the desire to transcend is seen as a noble aspiration by most religions and philosophies. But the modern methodology raises troublesome assumptions. For example, as long as a Golem was to be created through mystical incantation and the judicious use of the Name of God, it was in no way problematical to claim that such a Golem might be as much as or more than Man. Now, however, we attempt to create such a creature mechanically, as if Man himself were nothing more than a predictable collection of nuts and bolts. If the attempt is successfully, many will see it as a verification that Man is nothing more than nuts and bolts. But the

inference here is altogether unwarranted. [13] Whether or not man can be reproduced mechanically has little to do with the question of whether or not man works mechanically. Worse still, the belief that man is a machine may have dire consequences for men:

"One symptom of the attraction of the merely mechanical is the growing popularity, among some scientists and the public, of the idea that it will be possible to construct computers which are no different from man in thinking, feeling, or any other aspect of functioning. The main problem, it seems to me, is not whether such a computer-man can be constructed; it is rather why the idea is becoming so popular in a historical period when nothing seems to be more important than to transform the existing man into a more rational, harmonious, and peace-loving being. One cannot help being suspicious that often the attraction of the computer-man idea is the expression of a flight from life and from humane experience into the mechanical and purely cerebral.

"The possibility that we can build robots who are like men belongs, if anywhere, to the future. But the present already shows us men who act like robots. When the majority of men are like robots, then indeed there will be no problem in building robots who are like men." -Erich Fromm [14]

What we seem to have, then, is an interesting paradox. The desire to create an artificial intelligence is simultaneously a manifestation of man's highest and lowest urges. On the one hand it often reflects the desire to transcend, to make of this physical universe something greater than we see in it now. At the other extreme it can represent a flight from humanity into the unthinking safety of a mechanical conception of man. That these are both natural human tendencies is clear, and it helps us to understand how the science of Artificial Intelligence has

come to arise. Less clear is whether or not the goals of this science are plausible, or, more precisely, whether there are plausible goals toward which the researchers can work. In seeking such goals we shall consider first the nature of machine intelligence and then the nature of humanity. There is no consensus regarding the nature of either, but a comparison of the various views may help us to formulate a methodology for the future forging of such a consensus and for the understanding of the meaning of Artificial Intelligence.

Can machines do what humans do?

The deceptively simple question, "Can a machine ever think like a human being?" is almost certainly unanswerable, at least in such an ambiguous form. There is an obvious but fallacious line of reasoning that may lead to a negative conclusion, namely, that a human body is an absolute prerequisite for human thought. [15] This argument is obviously true in a trivial, semantic sense, but begs the larger question by simply defining human thought as that which only humans think. [16] As shall quickly become obvious, the problem of formulating the question properly may well be insurmountable.

One line of attack has been mathematical. Rather than waste their time trying to define exactly what it is humans do, many mathematicians have instead focussed on the

theoretical limitations and capabilities of idealized, abstract machines. The idea of such an approach is to formally prove that a computer-like machine can not, in theory, ever do certain things, or to prove, in some sense, that there is nothing that such a machine cannot do. Much of this type of work has been done in the last few decades, and the results have made many mathematicians optimistic about the prospects for machine intelligence. For example, one of them has seen fit to predict, from mathematical similarities he perceives in different types of knowledge, that "when a chess program has been developed capable of defeating the world champion, we shall be no more than five years away from the appearance of the 'ultra-intelligent machine', intellectually superior to man in all departments of thought." [17] This is somewhat startling in that such a program seems fairly likely in the near future, while there is much more skepticism about the imminence of an 'ultra-intelligent machine'. Another mathematician points directly towards a mathematical result that seems to imply the possibility of a computer with something akin to free will:

"...one of the basic results of the theory of computability (namely the existence of nonrecursive, recursively enumerable sets) may be interpreted as asserting the possibility of programming a given computer in such a way that it is impossible to program a computer (either a copy of the given computer or another machine) so as to determine whether or not a given item will be part of the output of the given computer." [18]

Of course, the reader will certainly be aware of the enormous gap between "nonrecursive, recursively enumerable sets" and "a free-willed computer"; this is but an extreme example of the difficulties of interpreting any mathematical results on the theory of computability. In talking about meaning, Mortimer Taube gives, perhaps inadvertently, an explanation for these difficulties:

"...any meaning, which as a proposition is the characterization of a nexus of actual entities by a selected set of eternal objects, has relevance to a nondenumerable set of possibilities which constitute a nondenumerable penumbra of meanings clustering around any given meaning. The graded relevance of non-realized possibilities to any actual proposition is what gives propositions different meanings to different people." [19]

Or, in plainer language, one may perceive and hence interpret meaning in several different ways. Thus one is never quite sure how well abstract mathematical results can be applied to real machines, especially when the capabilities of real machines are so different (and often broader) than those of the abstract models (e.g. the Turing machine). Such problems of precise definition and interpretation led Joseph Weizenbaum to conclude that, "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." [20] In short, mathematics provides us with results that are interesting but fail to satisfy our human desire to know what machines can and will achieve.

Naturally, philosophers have made their own attempts to deal with the question of machines 'doing what humans do'. Seeking formality where the mathematicians were fuzziest, many have sought to precisely define the salient features of human intelligence and hence to provide a frame in which to attempt to discover the limits, if any, of the ability of computers to reproduce it. Unfortunately, few of them seem willing to agree on anything. For example, there is a very real question of authority; who, for example, could state with authority that a machine was conscious? Sayre [21] suggests several unsatisfactory possibilities: a computer specialist could examine the controlling programs carefully and then pronounce the computer conscious or not, but what would he mean by or know about consciousness? The computer could declare itself conscious, but any computer can be told to do that now. Even when Sayre assumes that a computer could be "wired in a fashion that makes it incapable of lying," [22] he is confronted by the fact that such a computer would know no more about consciousness itself than any technologist or philosopher trying to answer the question; it would inevitably have to plead ignorance. Sayre fails to go on to note that any human trying to answer rationally the same question regarding his own consciousness would be equally at a loss. This is, in fact, the kind of impasse that results from nearly all philosophical attempts to formalize human intelligence and try to prove its

possibility either in humans or in machines. Every definition either seems to have too many exceptions, or to be too narrow, or too broad, to be useful. As another example, Sayre asserts that a necessary condition for consciousness is action [23]; by this criterion, the Buddha under the Bo tree was no more conscious than a rock. While such a statement may have a Zen-like appeal, it can also raise problems for many.

What seems clear, then, is that the intractability of the problem stems from the ambiguity of meaning itself. It is well known that almost any statement is subject to varying interpretations. Thus it is inevitable that any statement regarding the essence of the human mind will be debated hotly. What is more surprising, perhaps, is that so many seem to regard the problem as retaining the possibility of precise, unambiguous formulation and solution. Unlike the given points, lines, and planes of Euclidean geometry, or the unchanging Forms of Platonic cosmology, "meaning is clearly a relationship between message and recipient rather than a unique property of the message alone." [24] For Krishnamurti and others, we noted earlier, relationship is what remains after transcendence of the finite. It is thus illuminating to find a scientist, who has long searched for a precise definition of meaning, describing meaning itself as a relationship. This is, in fact, a widespread view among computer scientists and other students of meaning.

But relationship, Buber tells us, demands response [25], and man's response to this relativity of meaning has been varied and illuminating, especially in the context of Artificial Intelligence research. Having seen briefly the difficulties in trying to make definitive statements about the questions raised by Artificial Intelligence, we shall turn now to an exploration of some human responses to the simple asking of the questions.

Responses to the possibility of Artificial Intelligence

Obviously, the first to respond have been computer scientists themselves; their response, in many cases, has been enthusiastic pursuit of the goals of Artificial Intelligence. A few define these goals in the most extreme terms, and vehemently assert their plausibility:

"As the machine improves... we shall begin to see all the phenomena associated with the terms 'consciousness', 'intuition', and 'intelligence' itself. It is hard to say how close we are to this threshold, but once it is crossed, the world will not be the same... It is unreasonable to think that machines could become nearly as intelligent as we are and then stop, or to suppose that we will always be able to compete with them in wit and wisdom. Whether or not we could retain some sort of control of the machines, assuming that we would want to, the nature of our activities and aspirations would be changed utterly by the presence on earth of intellectually superior beings." [26]

Some enthusiastically pursue diametrically opposed expectations. "One hopes that as time goes by, the computer automation will be recognized ever increasingly for what it is: a rather moronic, but fast, storage and manipulation

device for information," says another researcher. [27] Others are commendably modest in the expression of their goals. "The primary motive of work on machine intelligence is an engineering one: we want to make a machine which is manlike in certain respects," says one. [28] Or,

"The phrase 'intelligent machine' is in fact a short statement for 'a machine which responds to information and also has some capacity for acquiring and manipulating information.' This corresponds roughly with that aspect of intelligence which is concerned with deduction of incontrovertible conclusions from specified data, leaving aside the more creative aspect of intelligence which is seen in the processes of induction and association." [29]

But this latter definition in fact leaves out many types of 'Artificial Intelligence' that have already been demonstrated. In any event, "once one has explained how a machine can be programmed to perform some process, it is hard for a human to regard the process as requiring intelligence." [30] This is the understandable reason why enthusiastic researchers are increasingly cautious in their public pronouncements; a great breakthrough in the direction of artificial intelligence is not necessarily the achievement of artificial intelligence, and hence can easily bring accusations of exaggerated claims. In general, though, the attitude of the community of computer scientists itself is very positive; Artificial Intelligence is seen as promising much and threatening little.

"...Machine Intelligence is an enterprise which may eventually offer yet one more mirror for man, in the form of a mathematical model of knowledge and reasoning. From such work we may perhaps learn a little more about our own

capacities. When one speaks of Machine Intelligence, one speaks of a collective venture to build 'knowledge machines'; but one also speaks of an unintended consequence: to fashion a mirror for the first knowledge machine of all, the machine within the skull." [31]

Of course, even as computer scientists have avidly pursued the goal of Artificial Intelligence, others have responded just as quickly with an enumeration of the pitfalls along the way. Hubert Dreyfus, a severe critic of AI research, has noted that "Before the program can function at all the total set of possible alternatives must be pre-selected by the programmer." [32] This is not merely a claim that the task is too large; it is a question of induction versus deduction, as a prominent AI researcher readily admits:

"The fact that... no means was devised whereby the program could generate new terms for itself is significant, for it is precisely at this point that current research on 'artificial intelligence' is encountering great difficulty. Deductive processes are in principle easy to mechanize. But the intellectual processes involved in induction, with their aura of 'creativity', 'originality', 'concept-formation', etc., are difficult to capture within a formal framework." [33]

That researcher went on to propose methods of solution of this 'difficulty', and Dreyfus went on to revise his own book. Without proof, one can continue to respond positively or negatively to the prospects for machine intelligence for as long as one has the strength to devise new arguments. Sri Aurobindo's words offer some light on this eternal process of rationalizing one's views, while also pointing out the weakest aspect of the computer's claim to

intellectual legitimacy:

"The truth is that upon which we are now insisting, that reason is in its nature an imperfect light with a large but still restricted mission and that once it applies itself to life and action it becomes subject to what it studies and the servant and counsellor of the forces in whose obscure and ill-understood struggle it intervenes. It can in its nature be used and has always been used to justify any idea, theory of life, system of society or government, ideal of individual or collective action to which the will of man attaches itself for the moment or through the centuries."
[34]

The key word in this passage is 'will'. Humans argue rationally (at times) in support of varying claims about Artificial Intelligence because they 'will' themselves to do so. They have chosen, for some reasons, their stands, and use reason as a tool to support these stands. If the computer is to duplicate Man, then, it may be necessary for it to develop a will to which its reason is subservient. We will return to this shortly.

We have not, as yet, discussed the responses of the lay community to the prospect of Artificial Intelligence. Naturally, much less has been written about computers by non-computer scientists than by computer scientists themselves. Still, it is easy enough to detect a few recurrent themes through conversation; ask a few friends what they think of the idea of machines more intelligent than men, and you will encounter a blend of terror, revulsion, and skepticism, and occasionally a few positive responses as well. There seems to be a basic human fear of

the computer, a fear of being surpassed, of not being needed. Abraham Heschel expressed this clearly in another context:

"Animals are content when their needs are satisfied; man insists not only on being satisfied but also on being able to satisfy, on being a need not simply on having needs. Personal needs come and go, but one anxiety remains: Am I needed? There is no human being who has not been moved by that anxiety.

"It is a most significant fact that man is not sufficient to himself, that life is not meaningful to him unless it is serving an end beyond itself, unless it is of value to someone else." [35]

Man would not be satisfied to be a useless second fiddle to superintelligent machines, even though he were well fed, clothed, and cared for as we care for our dearest pets. The human fear of the computer is a manifestation of the human need to be needed; it exists because most men do not feel needed by the universe, but only by human beings. Were it clear that man was needed in some way by the universe itself, or by a God whose existence could be depended upon, the fear of 'replacement' by the computer would dwindle or vanish. It is in the absence of such clarity that humans often respond to the computer with terror and revulsion.

Also, perhaps because of the terror, or perhaps independently of it, people respond with deep skepticism. We have already seen many expressions of such skepticism, which is often based upon an unprovable faith in the

ambiguity of reality. Speaking of the computer scientist in general, Weizenbaum said: "His knowledge is merely less ambiguous and therefore, like his computer languages, less expressive of reality." [36] That such an ambiguous perception of reality strikingly contradicts the mainstream of philosophical conceptions from Plato's Forms to Kant's noumena does not seem to enter into this response. Perhaps it is only in the face of the mechanization of all that is precise that we are coming to notice the imprecision of so much of our world and our perceptions.

In any event, it is enough to note that skepticism is a widespread response; we have already seen the near impossibility of validating such skepticism, for any task which can be well-defined can be formalized and hence can hardly serve as an example of what a computer can not do. Still, in the face of extremely widespread human intuition that there are other, ill-defined functions of the human mind that the computer has not yet begun to imitate, we ought at least to inquire into what those functions must be.

The Scope and Role of Human Intelligence

"Though the saint sees, hears, touches, smells, eats, moves, sleeps, and breathes, yet he knows the Truth, and he knows that it is not he who acts.

Though he talks, though he gives and receives, though he opens his eyes and shuts them, he still knows that his senses are merely disporting themselves among the objects of perception."

-- Bhagavad Gita [37]

"Man is undoubtedly 'wise' or 'intelligent' (the right translation of sapiens is hard to hit) but comparison with horses, dogs, dolphins, and apes seems to reveal a difference in degree rather than kind."

-- Donald Michie [38]

"However, we must not forget that in contrast to animals man is a being who not only behaves but also reflects about how he behaves."

-- A. J. Heschel [39]

If we listen to the voices of modern religious thinkers, as we must if we are to seriously consider the unprovable but nearly universal intuition that certain human functions are not computable, we find a surprising unity about one point: One of the strongest indications of the existence of 'something' beyond the strictly rational in man is the persistence of human soul-searching. Heschel phrases it simply:

"... The first answer to the question: Who is man? is that he is a being who asks questions concerning himself. It is in asking such questions that man discovers that he is a person, and it is the kind of questions he asks that reveals his condition." [40]

In the difficulty we have in conceiving of a computer asking such questions we find combined nearly all of the obstacles we perceive about the duplication of the human

mind on the computer. To ask such a question about itself, a computer would require free will, doubt, curiosity, and a very special sort of consciousness, an awareness of self. Such things may all someday prove tractable to the clever tinkering of technologists, but now raise the deepest of human doubts. Combined in the question "Who am I?", they sound resoundingly unmechanical. "The gulf between the human and the nonhuman can be grasped only in human terms."

[41]

But though we can make similar claims about a blind man's comprehension of the concept of color (e.g. it is something he is fundamentally equipped to understand because of his lack of the prerequisite physical organs), in the present case we are less able to pin down the specific deficiencies of the non-human, the machine. This is, of course, the same difficulty we have already encountered several times: were we able to 'pin down' these shortcomings, we would have described them in terms sufficiently clear to teach them to the machines. The stark truth appears to be that ultimate reality is simply not formally describable, not by computers, nor, for that matter, by the independent human intellect:

"Rationalism operates with the assumption that whatever is can also be known. It fails to distinguish between the world as given in my mind, wrapped up in concepts and categories, and the world as given to my mind as sheer being; between the self as given in my explanations of certain behavioral forms and the self as given to my mind."

[42]

With that in mind, however, we can all quickly think of several vague terms that summarize the salient aspects of true humanity; many philosophers have produced such lists with a great effort to minimize the vagueness. Aurobindo delineates four different aspects of the mind above the simple 'human intelligence' which could presumably be reproduced by machines. In ascending order, these are the 'Higher Mind' of conceptual awareness and spontaneous inherent knowledge, the 'Illumined Mind' of spiritual light and vision, the 'Intuitive Mind' of knowledge by identity (oneness with the object known), and the 'Overmind' of gnostic global knowledge. [43] While most would argue with these categories and many would deny the existence of such esoteric forms of knowledge, it is crucial to note that even skeptics can conceive of such ways of knowing well enough to dispute them. How could a computer respond to them in any way other than 'I don't know what you mean'? Erich Fromm, lamenting the apparent tendency of man to imitate the machine rather than vice versa, lists several 'humane experiences' crucial to true humanity: tenderness, compassion, empathy, love, interest, responsibility, identity, integrity, vulnerability, transcendence, and freedom. [44] Heschel adds a few of his own: preciousness, uniqueness, opportunity, nonfinality, solidarity, reciprocity, and sanctity. [45] The lists are clearly far from exhaustive, but the achievement of any one

of these qualities in an 'Artificial Intelligence' system would be a staggering and totally unexpected breakthrough. Of course, all of these qualities might well be called 'irrational', but the restriction of the goals of Artificial Intelligence to the 'rational' is far from standard. Indeed, many attempts have been made to simulate human emotions on the computer, although most have been quite transparently shallow. [46]

If then, we grant tentatively the existence of certain human traits that seem unlikely, at least, to be reproduced by computers, this still has not settled the question of intelligence. A computer need not know human anger, greed, or even love in order to surpass human intelligence. Once again, definitions are needed, but agreement is lacking. Krishnamurti says that "Intelligence is not discernment and judgement or critical evaluation. Intelligence is the seeing of what is." [47] This sounds nice, but it places all of the sciences and much of the arts outside the realm of the intelligent, and thus does not accord well with our normal notions of intelligence. Weizenbaum claims that "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." [48] This is intuitively appealing, for reason alone can not make sense of a void; it is impossible to reason about nothingness. A specified context is also inherent in

Katzen's characterization of intelligence: "Intelligent behavior, on the part of a man, a machine, or both, is the detection of the change of meaning brought about by a shift of context." [49]

Here, at last, we seem to be within reach of something tangible: We may characterize intelligence as rational actions or decisions in response to understood phenomena. In doing so we certainly do not assert this to be the basis of all human behavior or thought, although many would continue to claim that it is in fact so. However, it is a challenging and useful goal for Artificial Intelligence work. Although not always explicitly acknowledged, it is in fact often the real goal towards which current researchers are working. The lack of clarity has been largely attributable to the extreme visions of certain computer scientists and the lack of philosophically rigorous goal-delineation by many others.

In essence, what is proposed here is an empirical redefinition of intelligence as that which computers can potentially do. Thus the question 'Can computers be intelligent?' is replaced by the question 'Are there any human behaviors which are not explainable simply in terms of intelligence?' Although this may seem a semantic game at first glance, there is an important difference in orientation. The new formulation relegates the wrangling

over the theoretical limits of computation to a peripheral position. The question of whether a specific behavior can be made 'intelligible', i.e. computable, may be resolved by its achievement in the laboratory. This frees the computer scientist from the weight of constant challenges to his craft's legitimacy, while completely preserving anyone's right to doubt the computer's capability to surpass or even equal the human. No new conclusions are immediately reached with this formulation, but the term 'intelligence' is made less ambiguous and the term 'artificial intelligence' may be rendered less offensive to many. Those who proclaim the unlimited capabilities of the computer are essentially proclaiming that all human endeavor can be resolved into what we are now calling 'intelligence', i.e. discrete, logical sequences of deduction. Such people can thus have no objections to limiting the applicability of 'intelligence' to those things their machines can do. The doubters, meanwhile, may simply continue to assert the existence of that which transcends the intelligent, and to claim with Bergson that 'All that reason can do is to arrange the truths of intuition into tidy and ordered systems.' [50]

Conclusion

"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." [51]

The quest for artificial intelligence is, we have seen, a natural manifestation of the human desire to transcend the finite. Its possibility or impossibility is probably unprovable when it is stated in terms of the reproduction of all human behavior. However, if we wish to consider the possibility that human essence is composed of more than simple rationality, we can choose to describe the goals of artificial intelligence in terms of stretching the capacity of machines to their ultimate limits. By the empirical definition of intelligence, the limit of ultimate machine capacity is the limit of intelligence. Beyond that may exist much of the human soul, but the ultra-rationalist need not accept the existence of any such 'beyond' in order to accept this redefinition of intelligence. However, by defining intelligence as that which computers can do, we may prevent a great deal of hostile and fruitless argument about the legitimacy of a major and growing branch of science, Artificial Intelligence. Skeptic and science-fiction addict alike can then pursue the undoubted promise of increased mechanical competence without undue bitterness over theoretical questions about eventual capabilities. At the heart of much debate about Artificial Intelligence are

merely semantic disagreements. Perhaps, then, a redefinition of terms can promote a modicum of existential rapproachment between those who see the computer as man's evolutionary successor and those who, with Aurobindo, would see it as but a limited tool of the mind:

"The reason cannot arrive at any final truth because it can neither get to the root of things nor embrace the totality of their secrets; it deals with the finite, the separate, the limited aggregate, and has no measure for the all and the infinite. Nor can reason found a perfect life for man or a perfect society. A purely rational human life would be a life baulked and deprived of its most powerful dynamic sources; it would be a substitution of the minister for the sovereign." [52]

In this spirit, no final truth is reasonably arrived at.

Notes

(Reference is made by author's name to items in the Bibliography)

1. Heschel, pp.23-24.
2. Ghose [2], p.38.
3. Fuller, p.32.
4. Ghose [1], p.82.
5. Heschel, p. 54.
6. Arbib [2], pp.216-217.
7. Vonnegut, pp.274-275.
8. Krishnamurti, pp.133-139.
9. Kemeny, pp.18-19.
10. Arthur C. Clarke, quoted in Dorf, p.412.
11. Michie, p. 73.
12. Cf. Borenstein p.1 ff.
13. Cf. Weizenbaum pp. 132-134.
14. Fromm, p.45.
15. Dreyfus (p.235 ff.) presents this case rather well, but only as one of many arguments supporting his conclusion, which is unreserved denial of such a possibility.
16. Cf. Sayre & Crosson, pp.2-5, for a more interesting view of the meaning of the unique manner in which human intelligence functions.
17. I. J. Good, quoted in Michie, p.134.
18. Davis, p.vii.
19. Mortimer Taube, quoted in Katzan, pp.328-9.
20. Weizenbaum, p.206.
21. Sayre [1], p.7 ff.
22. Ibid., p.10.
23. Ibid., p.14 ff.
24. Mackay, p.24.
25. The idea is definitely from Buber, but I can find no specific reference for the current context.
26. Marvin Minsky, quoted in Dorf, p.405.
27. Alexsander, p. 92.
28. Michie, p.52.
29. Bell, p.1.
30. Arbib [2], p. 87.
31. Michie, p.3.
32. Dreyfus, p. 50.
33. Michie, p. 43.
34. Ghose [2], p.37.
35. Heschel, p.57.
36. Weizenbaum, p. 279.
37. Bhagavad Gita, p.51.
38. Michie, p. 2.
39. Heschel, p.9.
40. Ibid., p.28.
41. Ibid., p.22.
42. Ibid., p.30.

43. Ghose [2], pp.99-107.
44. Fromm, pp.77-90.
45. Heschel, pp.33-49.
46. It is, for example, common to have a computer express concern about a student it is tutoring, but the words are a set formula and in no way exhibit true interest or compassion. More sophisticated attempts to simulate emotions have not been spectacularly successful.
47. Krishnamurti, p.141.
48. Weizenbaum, pp.204-205.
49. Katzan, p.319
50. Paraphrased in Stace, p.195.
51. Weizenbaum, p.227.
52. Ghose [2], p.39.

Bibliography

- Alexsander, Igor, The Human Machine: A view of intelligent mechanisms; St. Saphorin, Switzerland, Georgi Publishing Company, 1978.
- Arbib, Michael A. [1], Brains, Machines, and Mathematics; New York, McGraw-Hill, Inc., 1964.
- Arbib, Michael A. [2], The Metaphorical Brain: an Introduction to Cybernetics as Artificial Intelligence and Brain Theory; New York, Wiley-Interscience, 1972.
- Bell, David A., Intelligent Machines: An Introduction to Cybernetics; New York, Blaisdell Publishing Company, 1962.
- The Bhagavad Gita, trans. Shri Purohit Swam; New York, Vintage Books, 1977.
- Borenstein, Nathaniel S., "Judaism and the Ethics of Artificial Intelligence"; unpublished, 1978.
- Buber, Martin, I and Thou; Edinburgh, T. & T. Clark, 1970.
- Dorf, Richard C., Computers and Man; San Francisco, Boyd & Fraser Publishing Company, 1977.
- Dreyfus, Hubert L., What Computers Can't Do; New York, Harper & Row, 1979.
- Fromm, Erich, The Revolution of Hope: Toward a Humanized Technology; New York, Harper & Row, 1968.
- Fuller, Richard Buckminster, Utopia Or Oblivion; Middlesex, Penguin Books, 1972.
- George, Frank H., The Robots are Coming; Sandbach, Cheshire, England, NCC Publications, 1974.
- Ghose, Sri Aurobindo [1], The Essential Aurobindo; New York, Schocken Books, 1974.
- Ghose, Sri Aurobindo [2], The Future Evolution of Man; Pondicherry, India, All India Press, 1974.
- Heschel, Abraham J., Who Is Man; Stanford, Stanford University Press, 1965.
- Hunt, Earl B., Artificial Intelligence; New York, Academic Press, Inc., 1975.

- Katzan, Harry Jr., Information Technology: The Human Use of Computers; New York, Petrocelli Books, 1974.
- Kemeny, John G., Man and the Computer; New York, Charles Scribner's Sons, 1972.
- Krishnamurti, J., The Second Penguin Krishnamurti Reader, ed. Mary Lutyens; Middlesex, Penguin Books, 1973.
- Leivick, H., "The Golem"; in The Dybbuk and other great Yiddish plays, trans. Joseph C. Landis; New York, Bantam Books, Inc., 1966.
- MacKay, Donald M., Information, Mechanism, and Meaning; Cambridge, The M.I.T. Press, 1969.
- Michie, Donald, On Machine Intelligence; Edinburgh, Edinburgh University Press, 1974.
- Raphael, Bertram, The Thinking Computer: Mind Inside Matter; San Francisco, W. H. Freeman and Company, 1976.
- Sayre, Kenneth M., Consciousness: A Philosophic Study of Minds and Machines; New York, Random House, 1969.
- Sayre, Kenneth M., and Frederick J. Crosson, eds., The Modeling of Mind: Computers and Intelligence; Notre Dame, University of Notre Dame Press, 1963.
- Stace, Walter T., The Destiny of Western Man; New York, Reynal & Hitchcock, 1942.
- Vonnegut, Kurt Jr., The Sirens of Titan; New York, Dell Publishing Co., Inc., 1959.
- Weizenbaum, Joseph, Computer Power and Human Reason; San Francisco, W. H. Freeman and Company, 1976.