HUMANISM AND ARTIFICIAL INTELLIGENCE

by Robert B. Tapp

Several factors support my imprudence in taking on this topic. In earlier days, I have taught at several liberal seminaries under the mantel of "liberal theologian" (if that is not seen here as an oxymoron). When I was writing a dissertation almost two generations ago, I learned to live with such primitive computing equipment as manual card-punchers, mechanical card-sorters, notched cards, and the like. By the time I was running a survey for the Unitarian Universalist Association in 1966, my available computers were electronic, and for the past few years my Macintosh has served me faithfully. Finally, as a result of all these varied inputs, I have been known (as Auden would have it) to "commit a social science."

At the outset, I hasten to concur with my coneague Harvey Sarles that the proper answers are Yes and No to the double questions -- are humans machines? and are machines human? I am always reminded of the Zen dialogue where the pupil says that he wants to have his mind purified and the master replies that this will be easy *if he* can just produce his mind. Until we specify what a machine *is,* we are in no position to take these questions seriously. Tennyson, for instance, somewhere between 1833 and 1850, could say "I think we are not wholly brain/Magnetic mockeries..." I Or, think of the hydraulic mechanisms that prevailed in Freud's thought, and the much more recent image of an .electrically-connected telephone exchange. None of these is today very appealing. But what of an intricate silicon chip? Or of carefully-constructed electrochemical translation points?

The prudent attitude may be the best at this juncture. While it is unwise to say "Never" in matters of science and technology, we can opine" Not Likely" to the proposal that we are on the verge of constructing analogous machines for mind/brain/body functionings in all of their complexity. Perhaps even more prudent would be the position that summary judgments are always premature, and that judgments are equally premature before the facts are in. Some of the hardest questioning comes from those philosophers who focus on machine-like minds and mind-like machines such as John Searle and Keith Gunderson. [2]

Let me move into the issues occasioned by Artificial Intelligence by reminding some background themes. Ancient Greek thinking came to be obsessed with the dangers of hubris-overweening pride. Humans must find their place in the scale of things and not overreach. The amalgam forged in the late days of the Roman empire between this heritage and Jewish themes, modulated by their Christian carriers, placed an even. lower estimate on human worth by making the "soul" the site of "being" thus leaving "body" and, by inference, "nature," in the lesser realm of "becoming." The resulting picture of a finite universe produced equilibrium for more than a millennium. Renaissance concerns with human possibilities in the arts, and sixteenth-century experiments with nature, set the stage for a new humanism. Now the stress was on the presumption that

humans *can* know nature, and *should* do so. By the end of the nineteenth century, it has become clear that our acts of knowing were by no means innocent, but that we change matters precisely by coming to understand how they proceed. Julian Huxley will give this a humanistic and biological formulation by claimimg that humans become, by this knowing, the very agents of evolution. In a slightly different sense, we became agents of cultures at the moment that we came to see such cultures in the plural. This meant that they were *in* nature, but that we could no longer assume a "natural culture" in some singular, universal, sense. But it also meant that this production of cultures by humans was of lasting significance as a kind of creation that was biologically-based(it no longer made much sense to claim it to be mind- or divine-based in any singular, universal sense). Perhaps the most useful contribution of Teilhard de Chardin will be his term"noosphere" to describe the locus of this human augment to the planet.

From 1960-63 I had the opportunity to chair a commission on "Theology and the Frontiers of Learning" for the Unitarian Universalist Association. [3] One of the assertions that made sense then, and still seems viable, was that"modern science is the way we know reality, while religion is the way we create reality." The cultural, noospheric reality that I then saw being created by religions is, of course, created by a number of other human activities as well..the arts, politics, the applications of the sciences, etc. But I think the distinction between creating and knowing is both important and useful.

Gandhi's *satyagraha* illustrates this nicely. The history of our time continues to be shaped by applications of this new technique (Gandhi would object to this sterile label for what was to him a process at once religious and political). Our scientific perspectives allow us to understand the technique more fully once it emerges. We can explore its conscious roots in Jainism, Hinduism, Jesus, Tolstoy, Thoreau. We can speculate on the less-than-conscious roots that generated this amalgam of revisions. We can analyze the cultural situations that made for its success--and by implication describe situations in which those successes might have been considerably lessened. We can trace the reactions of varied political and religious forces to this new strategy. But none of these "scientific" analyses ("knowing") would be very significant if the datum of the pratice of *satyagraha* in South Africa and India (a human "creating") were not present. As long as our concepts of world/universe remain large enough to include humans and their activities, we can say that the world can be/has been transformed by human activities. This transformation, for millenia, has included an "artifactual" world comprised of tools and symbols. More recently these have extended to transformations and storage devices such as books and digital information on disk.

These recent additions to our artifactual world, of course, are what have raised the ante on Artificial Intelligence. "Can machines think?" becomes "What thought processes can be analyzed into replicable operations?" I encourage readers here to review their own glossaries for operational words that reflect the current fruitfulness of these efforts. Start by "scanning the

memory bank, "which can be augmented by "interfacing" with an encyclopedia. Then review the "input modes" by which we can sense the world external to our consciousness, "generate and evaluate alternatives" and "output" some kind of activity. Along the way will be numerous comparisons both "digital" and "analog." Before beginning some action, we may resort to a series of "simulations" to select the preferable route. We may even discover "filters" and "suppressor circuits" that need to be circumvented. In these processes, we will continually be making use of "transforms" to alternative "symbolizations" and "languages" and "logics," which will involve "encodes" and "decodes." "Paths" may even protest being"overloaded." If we are not careful, we may wind up in an "infinite loop." The interesting thing about such an exercise as this is that almost all persons at this institute could play this same game, expanding my brief list of neologisms. And our children, who have inherited this computer-world from birth, could in many cases quickly outdo us.

Our artifacts do things for which such words have become necessary, making it much harder to deny their intelligence. At the very least, we need to use some more general term like "living"to apply to some of these machines. Geoff Simons argues quite persuasively that the four criteria of life--structural, energy -processing, information -processing, reproducing--now apply to our most advanced machines.' We could readily go beyond this and locate a number of Darwinian selection mechanisms that improve machine-species. It may be that for many persons the granting of living status to certain machines will provide a transition to considerations of the "intelligence" of these machines. There would, in any case, seem to be little doubt that any serious consideration of such questions depends upon specification of criteria. In terms of our earlier distinction between knowing and creating, there should be little difficulty in arguing that our present machines can handle many aspects of knowing'-replicating, transforming, sensing, recognizing, symbol manipulating. In almost all of these human activities that have been transferred to machines, the successes have been impressing and the reciprocal potentials of learning more about brain/mental functions from the construction of machine-analogues are even more promising. Where success is presently lacking is in the creative activities. I have already suggested *satyagraha* as one such recent human artifact. Let me now put this more formally by suggesting four areas where creativity seems dominant and, therefore, where machine-replication seems less likely.

1. Lifestyles. Unless we assume some universal stratum of human behaving, each lifestyle that has attracted some human group represents the creative act of choosing and cherishing. Within recorded history the range of such lifestyles, and the ways in which they have on occasion displaced early forms, is a major social fact. In this century diverse alternatives such as the already-mentioned *satyagraha* as well as genocidal state terror have emerged. The potential pluralism of lifestyles becomes even more evident as past' 'certainties" become untenable under scholarly scrutiny. Patriarchy will no doubt be dominant for a long time to come, but it must now be seen as ideological (Le. emerging within some particular group at some point in time) rather than "natural. "s

2. Paradigms. Further arguments for creativity are the emergence of new scientific paradigms and the differential ways that they have been received. Take the Darwin-Wallace shift to evolution. In one sense these new ideas were in the air, but in a more strictly scientific sense, it took the patient argumentation and assemblage of data by these two brilliant creators to start the paradigm shift. **In** a narrow biological sense, it took more than a generation for a new consensus to emerge. In the larger sense, there immediately began attempts to trace the "evolution" of religion, art, morals, societies, law, and almost every field of human endeavor.

The recent renewal of creationist arguments, five generations later, reminds us that paradigm shifts occur slowly and in sectors. The fact that 40 percent of the U.S. populace, along with the current resident of the White House, reject this scientific unanimity should underscore the contention that we are in a realm of creativity rather than necessity.

3~ Valuations. In focusing on states and behaviors that humans choose and cherish (i.e. values), we remain in this realm of creativity. People could always have chosen otherwise, and sometimes in fact subsequently do. Democracy and the linked issue of human rights will serve well to illustrate this. While the distant model was in ancient Athens, the eighteenth-century European and American innovators more directly inherit a series of changed evaluations of political power and of the human prospect. Kings as a class are now seen as having no special wisdom. This involved a serious rejection of many of the ecclesiastical claims supporting monarchy. If the locus of social wisdom is not in castle or church, we must either become skeptics or recognize the equal potentials of "all men" to possess sufficient wisdom to guide themselves.

I have purposely quoted the sexist designator "man" from that not-distant era. There was, of course, a series of implied footnotes to such formulations-"free," "propertied," "parish-members." The operative term here is implied. To most of the members of the U.S. Continental Congress the possibility that slaves might be participants in the commonwealth was unthinkable, and the participation of women was beyond consideration. At each stage of redefinition by expansion, the correlated meaning of "democracy" was revised.

This shift from "rights of man" to "human rights" was more than rhetorical, and resulted from the emergence of claims which were then pressed. As we move into a period when "animal rights" are seriously discussed, these ways of describing the problem will require rethinking. Animals are in no position to raise questions of their rights or to press claims for them. Nevertheless, most of us are able to understand a discussion in which these become central. The obvious point is that quite new levels of valuation can occur in the course of human history.

4. Interpretations. That the French Revolution occurred at the end of the eighteenth century has never been in dispute. What is not yet (can never be?) agreed is what that revolution meant/means. The ending of an ancient regime? The overthrow of arbitrary monarchy? The creation of revolutionary terror? The emergence of human rights for much of Europe? The creation of Bonapartism? Why such variety, especially so long after an event? The issue obviously lies much deeper than nationalist perspectives. (Michelet and Toqueville are both French).6Nor does it help to use vague terms like "subjectivity," "chance," "divine hand," which might have been useful prior to the Enlightenment. The stand any of us takes on the French Revolution is a combination of many high-level factors in ideology, valuation, lifestyle -- which I am here designating as interpretation. This linguistic-mental activity is, I believe, the most significant aspect of the nature of human nature. I have carefully avoided a simple distinction between facts and interpretations. The designation of "fact" is itself interpretive, and the using patterns to order our perceptions is equally interpretive, even when they are such obvious patterns as chronological-temporal ones. Having designated facts, the employment of them becomes a highly-creative activity. The Viet Nam War -- was it a tragedy? If so, how and for whom? Vietnamese liberals, American liberals, American hawks, French imperialists, Chinese Maoists, Cambodians of varied persuasion? If someone asks me how I feel about "Christian civilization" and my response is that "**I** would welcome it," should the exchange be understood literally or ironically?

Conclusion. The common core to enterprises such as lifestyles, paradigms, valuations, and interpretations, I have here contended, is creativity-the production of the unforeseen, the indeterminate. Somehow, this differs from products of our best, present machines which are highly determinate. The suggestion is sometimes made that so-called creativity is simply the introduction of random variation, perhaps by some such mechanism as copy-error (which could easily be programmed in present intelligent machines). One could imagine such "machine-creativity" succeeding in some contexts. For instance, the creation of new chemical compounds to control viral activity could be programmed on such a trial-and-error basis. But in that case, the criteria for the recognition of success (i.e. better control) would have to be knowable and specifiable in advance. Imagine, now, the proverbial roomful with a battery of word processors. Eventually they could/would come up with a *Hamlet-like* text, but the program specification could not "recognize" this as new and valid drama without specifications presently unimaginable. If we take the old version of the problem (Could they write the text of *Hamlet?),* we could reduce this to a problem of randomness eventually producing a pattern that compared to an already-known pattern. This would fall far short of writing the first *Hamlet,* a genuine act of creating a play that did not already exist.

This argument is not so much an attempt to limit what artificial intelligence can do as to set forth, in some non-reductionist fashion, what must be done if we are to successfully replicate human activities in their fullest. My final reminder would be that success in this kind of replicative venture would become one more human activity in need of simulation and replication.

I "In Memoriam," stanza 119.

[2] cf. John R. Searle, *Minds, Brains, and Science.* Cambridge, Univ. Press, 1984; Keith Gunderson, *Mentality and Machines.* Rev. ed. Minneapolis: Univ. of Minnesota Press, 1985.

[J] The commission was initially appointed by the American Unitarian Association, and then incorporated into the newly-merged denomination.

'Geoff Simons, *Are Computers Alive?* Boston: Burkhauser, 1983.

[5] See particularly the new study by Gerda Lerner, *The Creation of Patriarchy.* Oxford: Oxford Univ. Press, 1986.

[6] Many of these issues are brilliantly summarized in Hayden White's *Tropics of Discourse.* Baltimore: Johns Hopkins Univ. Press, 1978.