

How Can a Theological Understanding of Humanity Enrich Artificial Intelligence Work?

CHRISTOPHER B. KAISER

DEFINING THE THEOLOGICAL UNDERSTANDING OF HUMANITY

In order to answer the question posed by the title of this paper, we must first agree on what we mean by a "theological understanding of humanity." The most obvious sense of the phrase would be the understanding of human nature as defined by the councils and creeds of the Church. However, the creeds of the Church do not give us a statement of the nature of humanity sufficiently definitive to allow us to draw any direct conclusions about efforts to develop Artificial Intelligence (hereafter AI).

Creeds shared by Protestants, Roman Catholics and Orthodox Catholics tell us that we have been created (both physical body and rational soul) in the image of God, that we have somehow fallen from the grace of our original state, that our nature was assumed in its entirety (both physical body and rational soul) by the Son of God, that our bodies will be raised from the dead, and that we will be judged for things we have done in this life.

Councils recognized by the Roman Catholic Church add the more technical points that humans have one and only one soul each (against Averroists and Manichaeans) and that the soul is the form of the body (an Aristotelian term)--in other words, the soul gives each of us our distinct personality and character.¹

Both Catholic and Reformed standards generally affirm that the soul is distinct from the body and continues to exist after the death of the body (the intermediate state), but they say little about its exact relationship to the body in this life.²

Christopher B. Kaiser, Ph.D., is professor of theology at Western Theological Seminary, Holland, MI. This article is based on an invited lecture at the symposium on "The Human Dimension in Artificial Intelligence," sponsored jointly by Asbury Theological Seminary and The University of Kentucky (Lexington, KY) in April of 1988.

These statements are not definitive for our purposes for two reasons. For one thing, the times in which these articles of faith were formulated were ones in which the possibility of AI was not even being considered as a viable research program. In fact, the theologians of the early Church regarded many questions of science as hopelessly speculative--not because they were opposed to science as such, but because there was at that time no prospect of ever being able to develop the instruments needed to decide the truth of such matters. For example, Irenaeus regarded speculation about the migration of birds and the causes of meteorological phenomena like rain, thunder and lightening (all treated on speculative grounds in Aristotle's *Meteorologica*) as entirely futile.³ Basil said the same concerning discussions about what supported the earth in space.⁴ Theophilus took it as granted that humans could never produce machines that would have the capacity for motion or sensation, let alone reason.⁵

We must keep in mind that the technologies that make modern science possible are only a few hundred years old, and they place us in a context of scientific, political and moral possibilities quite different from that of the early, medieval and Reformation Church from which most of our confessional statements come. So it would make no more sense to derive implications concerning AI from Christian doctrine in a straightforward, logical manner than it would to make judgments about the theory of evolution on the basis of the book of Genesis. If we are to derive any results at all, we shall have to do so by focusing on the underlying idea of the image of God in humanity and by considering the implications that idea has had historically for the development of modern science.

A second reason for saying that Church dogma is not definitive by itself has to do with our understanding of humanity. If we know anything about human nature, it is that it is highly malleable. In contrast to other social animals like ants and bees (less so to cetaceans and the higher primates), humans are governed by their social and cultural conditions as much as by their genes. Our brains, for instance, are programmed by genetic material only in a very general way. As Jean-Pierre Changeux and others have pointed out, the number of genes in the human genome, roughly 10^5 (one hundred thousand), immense as it may seem to us, is only a small fraction of the number of synaptic connections in the human brain, roughly 10^{15} (one quadrillion).⁶ The difference is a factor of 10^{10} (ten billion)! We do not know just how the instructions for the initial formation of the brain are encoded in the human genome, but the structure of the mature brain is clearly underdetermined, genetically speaking. In other words, the formation of the brain and mind is controlled by interpersonal and environmental factors as much as it is by the genotype.

This malleability of human nature means that humans are largely what they make of themselves from one generation to the next: depending on our culture, we define ourselves through the structuring of our social relations, the degree of individual freedom in relation to group identity, the structures of our technology, moral values and religion.

Technological structures are a particularly important aspect of the cultural

formation of human nature. In the modern West, we recognize that thought and behavior of humans is rather different in a technological society than it is in a more traditional one.⁷ But the process of human self-definition has always been in a *symbiotic* relation to the development of human technology.⁸ We get hints of this as early as in the book of Genesis where the image of God is closely related to human dominion over nature--a dominion which is described in contemporary images of agriculture, the domestication of animals and metallurgy (Gen 1:28-29; 2:15-20; 4:20-22).

One of the most exciting, and sometimes frightening, aspects of the current computer revolution is the fact that we are asking questions about ourselves with renewed intensity. Are certain human traits potentialities within the laws that govern matter itself? Can we replicate ourselves through technology as well as we can through sexual reproduction? If not, which are the characteristics that cannot be reproduced in machines and why? How does all this relate to our understanding of the "image of God" within us? Does it tell us something about the relation of matter and spirit?

AI research is not just a narrow specialty that nonspecialists can ignore. While it is not as readily marketable, in the economic sense, as the "expert systems" that sometimes go by the name of AI, it confronts us anew with the central questions of our own human existence.

If and when the verdict is in on the viability of AI as a research program, we will all be in a better position to define our own nature than we are at present. Our "theological understanding of humanity" is grounded in the creeds of the Church, but it is not spelled out in explicit terms that would allow us to predetermine what can or will be. What ought to be is another question.

THE IMAGE OF GOD AND HUMAN INTELLIGENCE

So the theological understanding of humanity is not determinative in any straightforward sense: it does not have a fixed, explicit meaning, and it is subject to reinterpretation in every age. Historically, however, the theological understanding of humanity has had important and influential meaning for the development of Western culture. And, since the quest for AI is very much a product of Western technology and philosophy,⁹ this particular expression of the theological tradition is relevant to our concerns.

In fact, much of the confidence in human reason and human ingenuity that has made modern science possible was initially inspired by belief in one of the theological doctrines I mentioned earlier--belief in the creation of humans in the image of God. Traditionally this belief has had two correlates: that humans could know and understand the world God had created and that they could make moral judgments. Humans were believed to reflect the divine mind--the same mind that authored the laws of nature, and humans were believed to be accountable to the divine will--the same will that authored the moral law.¹⁰ These two ideas are the epistemic and ethical correlates of belief in the creation of humanity in the image of God. Together they constitute the historic Western theological understanding of human intelligence. This is just one tradi-

tion, but it is the tradition out of which the modern scientific quest was born. It is, therefore, the tradition to which we must refer back insofar as we want to know the meaning of the scientific enterprise as it has developed historically in the West.

THE IMAGE OF GOD AND REVOLUTIONARY SCIENCE (EPISTEMIC CORRELATE)

The idea that the divine mind is reflected in human intelligence--the epistemic correlate--goes back to the wisdom tradition of ancient Israel and to the Platonic tradition of classical Greece. Both Hebrew and Greek traditions were rooted in the mythology and cosmology of the ancient Near East. But the principle vector for transmitting these ideas to Western Europe was the patristic tradition of the Church which blended biblical and Platonic ideas and stressed the importance of mathematics for an understanding of the structures of creation.¹¹ Without such a belief there would have been no good reason to suppose that the principles of mathematics would have been applicable in any profound way to questions about nature.

Belief in the reflection of God's mind in humans meant that the world was in principle comprehensible--that is, it was comprehensible *insofar as it could be explored*.¹² Though space travel in the modern sense was unthinkable, apocalyptic writers supposed that the mechanics of the heavens would be comprehensible to someone like Enoch who was enabled by his ascent into heaven to examine them at close range.¹³ Against the skepticism of some Greek schools of philosophy, Tertullian and Augustine argued that sensory perception was basically reliable and that the world was comprehensible to the extent that our five senses allowed us to experience it.¹⁴ Even the inner recesses of the human mind were believed to be comprehensible, at least to Augustine, by virtue of God's image within it.¹⁵

Faith in the comprehensibility of the world was an essential factor in the rise of modern science. This can be seen in the endeavors of early Western scientists like the Venerable Bede (early eighth century), Adelard of Bath (twelfth century), Robert Grosseteste (thirteenth century), John Buridan (fourteenth century), Copernicus, Galileo, Kepler and Newton. Copernicus, Galileo, Kepler and Newton are generally recognized as the founders of modern science. Even though there have been so-called revolutions in twentieth-century physics--the theories of relativity and quantum mechanics--modern physics continues to be based on the method and findings of these early modern pioneers.

It is difficult to conceive of modern science developing as it has without founding figures like these, but it is just as difficult to imagine any of them working with the intensity needed to overcome the obstacles of early modern science without the kind of faith in the comprehensibility of the world that they had.

The founders of modern science were doing what Thomas Kuhn has termed "revolutionary science": that is, they were seeking answers to problems for which they had no reason to suppose (on the basis of existing science) that

answers existed.¹⁶ In theological terms, they worked by faith more than by sight--a faith sustained by their theological understanding of human nature.¹⁷ I would argue, however, that the revolutionary nature of their work was something they shared with their early medieval predecessors. Though the paradigms of modern science were not established until the sixteenth and seventeenth centuries, the attempt to reach out into the unknown, beyond what inductively established paradigms would allow, was made by natural philosophers in every century of the Middle Ages.

Scientists that have succeeded Newton have often been able to follow their example without necessarily sharing their religious faith. But those like Albert Einstein, who have pioneered radically new paradigms in scientific thought, have often accepted the comprehensibility of the world as a tenet of their faith in God. There is probably no way to establish an exact correlation, but, historically, revolutionary episodes in science appear to be ones in which faith played a significant role.

Now this epistemic correlate of belief in the image of God tells us two things about the understanding of humanity that might well enrich AI research. It tells us something about the character of the intelligence AI research may hope to reproduce, and it also tells us something about the capacity of AI researchers themselves. In other words, we can take the founders of modern science as models for the definition of what we mean by human intelligence--making the goal of reproducing it seem almost impossible--after all, AI workers have their hands full simply trying to replicate everyday common-sense behavior! Or we could take them as models for what we may expect from AI workers (they, too, are created in the image of God!) and the legitimacy of their efforts to accomplish the seemingly impossible.

SCIENTISTS SUCH AS COPERNICUS AS MODELS OF HUMAN INTELLIGENCE

The reason for taking the founders of modern science as models for our definition of the intelligence AI research aims to reproduce is that they form a bridge between classical Christian faith and modern Western science. In the lives and work of early modern scientists, the meaning of the theological understanding of human intelligence was worked out in such a way that it had implications for the scientific enterprise--implications which could not be derived in an unambiguous way from the creeds of the Church themselves. Since this is just one way of defining intelligence, or one particular aspect of human intelligence, I shall earmark it by referring to it as "Copernican intelligence," Copernicus being perhaps the most widely recognized historical expression of the revolutionary sort of intelligence I have in mind.

If we learn anything about the character of the intelligence AI hopes to replicate from these historical exemplifications of theological understanding, it seems to be this: a creature is intelligent (in the Copernican sense) only if it makes judgments concerning its own efforts to cope with reality when its very ability to cope is in doubt (and recognized to be so) and there are no induc-

tively-based paradigms for establishing the probability of success.

The mere ability of a machine to seek and to find solutions to problems--whether by heuristic programs or by semantic information processing or by scripts or by stochastic processes¹⁸--is certainly a necessary condition for intelligence, but not a sufficient one according to our definition. Why? Because the very *possibility* of arriving at a solution is never in doubt when the machine is programmed to keep trying or to quit according to some pre-determined rule--a rule based on the judgment of the AI worker who programmed the machine. In other words, problem-solving machines as we know them today can only be said to be "intelligent" or even to "solve problems" in the metaphorical sense.

For comparison, consider the way we speak of primitive prokaryotic (non-nucleated) cells as having "solved the problem" of how to produce energy more efficiently or of having "learned the trick" of photo-synthesis. Orthodox science today would readily admit that this is merely an anthropomorphism, a figure of speech. Prokaryotic cells are only able to live and reproduce themselves (by simple subdivision). The fact that some of them (the blue-green algae) evolved in such a way as to produce energy more efficiently was governed entirely by genetic mutations and environmental conditions. If intelligence came into the picture at all, it was that of the Creator, not that of the creature, however remarkable its accomplishment may have been.

At a somewhat higher level, we speak of mice "solving the problem" of finding the correct way through a maze. Again this is a figure of speech. Why? Because the propensity of mice to keep searching is due entirely to basic drives conditioned to a degree by past experience. The anticipations and skills they employ are a mere projection based on past experiences of mazes much like the one at hand. There is no possibility of a mouse considering whether, in the case of a radically new situation, there are grounds for judging whether efforts towards a solution are worthwhile.

There is evidence that early hominids like *Homo erectus* and early *Homo sapiens* (the Neanderthals) also searched for new habitats and colonized much of northern Europe and eastern Asia. Perhaps this, too, was simply a matter of basic drives conditioned by the need for survival. Or perhaps the possibility of penetrating the unknown was weighed in relation to the obvious risks involved. We don't know for sure. In any case, migration to new habitats presumably took place in small steps, none of which required a significantly greater risk than the one before it. Hence, past experience could serve as an indicator of the probability of further success.

A modern-day comparison would be the solution of crossword puzzles. Difficult as these puzzles may be for some of us, their solution does not require intelligence in the Copernican sense: we know in advance they were designed by human beings using basic patterns and language that we share on the basis of a common culture.

Or, to take a comparison closer to the topic of AI, consider the decipherment of enemy codes during World War II. Surely this was a job for "intelligence," as the term is used in the business of espionage. In fact, some of the

earliest steps in the development of computers were taken in the process of deciphering codes (Turing et al., late 1930s). But even this is not intelligence in the Copernican sense that we are using to illustrate the Western theological understanding of human nature. In military intelligence work, a cipher is always *assumed* to have a meaning. Anyone who tries to decipher it must have confidence in her or his individual ability. But the problem is known to be solvable in advance, and its solution is generally thought to be a matter of time.

In comparison to these examples, early scientific efforts by Copernicus, Kepler and Newton to understand the kinematics and dynamics of the solar system was distinctive in that there was neither compulsion nor guarantee of success. The history of Western culture could very well have gone on without these breakthroughs in science. The effort was highly contingent: it exemplified the way in which humans define themselves and their abilities--defined, in this case, in terms of their faith in the ability of the human mind to understand what God had created.

Perhaps the first indication in the fossil record of intelligence in this special, "Copernican" sense is the evidence that Neanderthals provided grave goods in the burial of their dead, particularly in the case of children.¹⁹ Perhaps there was survival value for the group in practices like these, but individuals still had to reflect on the meaning and value of practices that could be varied in an indefinite number of ways without affecting the immediate material condition of the group. In other words, they were attempting to penetrate the unknown and questions of belief became a factor in the effort to solve problems for which neither environmental conditions nor past conditioning was determinative.²⁰

TESTING A FULLY SOCIALIZED COMPUTER FOR COPERNICAN INTELLIGENCE

Under what circumstances could computers exhibit Copernican intelligence? This is still largely a matter for speculation. But a few suggestions may be offered if only to keep the ultimate goal in sight and to indicate the enormity of the task.

As philosophers like Hubert Dreyfus have argued, the way in which humans know things can not be formalized in a way that is completely independent of context.²¹ Advocates of AI like John McCarthy have recognized the same problem.²² Computers that are programmed to exercise certain skills are completely unable to cope with situations in which context may change unpredictably in such a way that the meaning of given stimuli is altered--yielding "ice cream," for example, instead of "I scream!"²³

Actually, we have a very similar difficulty with our children. We try to teach them "manners" for instance. It is proper for them to use their fingers when eating some things, but not when eating others. They may run and jump in some contexts, but not in others. There may be no way in which to formalize all the rules involved, yet our children slowly learn to make what we are willing to accept as good choices. They become socialized. Of course, some critics

argue that children are simply being limited in the scope of their imagination. Their creativity may actually diminish as their cultural baggage increases!

On the other hand, we know that children deprived of human contact in their formative years have great difficulty in behaving in socially acceptable ways. The same is true for other primates like chimpanzees.²⁴

So, even if we succeed in building computers with the hardware and built-in programs that match the capacity of a newborn child--including something like a body with five senses and two hands in order to model human interaction with the world--we must suppose that a lengthy process of socialization would be required before the "intelligence" of the machine would be formed in a way that would allow it to function in real-life situations with any degree of success.²⁵ Such a machine would have to be able to "grow" as the human brain does. Both its hardware and its software would have to develop as the machine learned of its identity and its capabilities through human contact.²⁶

Indeed, it could take many generations of humans and computers functioning together as a group for the human partners even to determine all of the hardware and programming requirements for the design of potentially intelligent machines. Design and communal relations would have to evolve together phylogenetically before an individual machine would emerge that could achieve intelligence ontogenetically. I shall refer to such a hypothetical machine as a "fully socialized computer." And, for the sake of the argument, I shall assume that the development of such a machine is possible in principle.

The question then is whether a fully socialized computer would be intelligent. At what point in the co-evolution of design and communal relations would, or could, humans recognize their computers as equally intelligent?

Insofar as the historic Western theological understanding of human nature is any guide, I would look for the day when computers began to pioneer in areas of revolutionary science.²⁷ At this juncture their human mentors would no longer be infallible guides. Computers and humans would not only share a common stock of knowledge, but they would face a common unknown as potentially equal partners. In other words, our relationship to potentially intelligent computers would be something like our relation to our graduate students at the point that we recognize them as colleagues. This is not to say that a Ph.D. is a necessary condition for intelligence. But the capability of doing independent research is assumed of any human with the necessary training as a consequence of their being intelligent. Could a fully socialized computer exhibit Copernican intelligence? Could it decide, for instance, that it was wasting its time in pursuing a particular problem of revolutionary science--that is, in cases where there is insufficient precedent to allow judging on the basis of past experience? Could it decide that its efforts would better be spent on something else? In other words, could a computer articulate a belief system about the nature of its own mind in relation to the real world (not the known world, but the real one)?

On the other hand, could our fully socialized computer choose to go on with a research project when its human colleagues had given up--not just out of

necessity or a simple projection based on past experience, but on the basis of individual judgment? Could it recognize itself as participating in a transcendent order of ideas, for instance, or as created in the image of God?²⁸ Could it exhibit greater faith than its creators?

SCIENTISTS SUCH AS COPERNICUS AS MODELS FOR AI WORKERS

Considering the founders of modern science as paradigms of the historic Western theological understanding of human nature gives us some idea of the enormity of the task of reproducing Copernican intelligence in machines. However, we must weigh this enormity against that of the capacity of human beings to understand and to invent. For this, we must take the early modern scientists as models of what we may expect from the AI workers themselves.

For example, if we consider Kepler in his own context (the early seventeenth century), we realize that he was pursuing a nearly impossible, and possibly hopeless, task in trying to find mathematical laws that would describe the data of the orbit of Mars collected by Tycho Brahe. We do not always appreciate this fact because we view the matter from the vantage point of the later Newtonian science that vindicated Kepler. We know that the orbits are simple geometric figures--ellipses with the sun at one focus. In retrospect, we can see that Kepler had a decent chance of finding a solution since the geometry of conic sections was reasonably well known.²⁹ Similarly, the non-Euclidean geometries Einstein needed to work out his general theory of relativity were available, even if not so widely known.

But neither Kepler nor Einstein, themselves, had any logical reason to suppose (on the basis of existing science) that solutions to their respective problems were at hand. They spent large amounts of time and energy pursuing projects that could have been a colossal waste of time! In fact, both Kepler and Einstein knew the agony of defeat when cherished ideas turned out to be wrong. But they persisted because they believed that abstract geometries could be applied to the real world even though most of their theorems were not derived from everyday experience.³⁰ As Kepler himself put the matter:

Those laws are within the grasp of the human mind; God wanted us to recognize them by creating us after his own image so that we could share in his own thoughts....Only fools fear that we make man godlike in doing so; for God's counsels are impenetrable, but not his material creation.³¹

If AI workers today exhibit this same kind of faith, in a curious way they reproduce the faith of Kepler and Einstein. They too exhibit what I have called Copernican intelligence.³²

Whatever the potentialities of matter may be, the potentialities of the human mind in understanding them are a matter of personal faith. After all, the conditions under which our brains were formed by natural selection millions of years ago were not ones in which the deep understanding of the laws of matter

were a significant factor. On the basis of natural selection alone there is no reason at all to suppose that our minds could ever understand the workings of the human mind, or, for that matter, that they could have understood the dynamics of the solar system or the large-scale structure of the universe. Insofar as we share the faith of Kepler and Einstein--and insofar as we share the theological understanding of human nature that lies behind it--we may also share the faith of AI workers and other scientists who pursue the quest for understanding beyond what the previous history of science would give adequate reason to suppose was likely to succeed. We may share their faith even if we have a greater sense of the enormity of their self-appointed task.

Hubert Dreyfus has also drawn parallels between AI research and early modern scientists, but, instead of pointing to Kepler and Newton as I have done, he points to the alchemical research programs of the Middle Ages and Renaissance.³³ This alternative is equally valid in my view, but it does not prove the futility of AI programs as Dreyfus implies. Historians generally recognize today that major developments of early modern chemistry--not just the chemical apparatus, as Dreyfus suggests, but the theoretical concepts and experimental results of early modern chemistry--would not have been possible without the belief structure mediated by alchemical notions. For example, Van Helmont's thesis, that each element had a distinctive "chaos" or "gas" that was given off when it was heated, was fundamental to all subsequent efforts to isolate and identify various gases.³⁴ Far from being a futile effort, alchemy played a positive, creative role in science at a time when there were no empirically-based paradigms to rely on.

We can not say whether AI will succeed in achieving the goals which it has set for itself, but, insofar as it exemplifies the human quest to understand the things God has created using the gifts God has given us, it is consistent with the theological understanding of humanity as understood historically in Western Christendom.

THE IMAGE OF GOD AND MORAL JUDGMENTS (ETHICAL CORRELATE)

For the sake of simplicity, we often differentiate between intellectual and moral questions, or between the faculties of cognitive and ethical judgment. From a theological perspective, however, the two must be related. There can be no real intelligence without consideration of justice. And, of course, the exercise of moral judgment requires a cognitive understanding of the world in which such judgments are to be made. Consequently, we must ask whether fully socialized computers would be concerned about the morality, as well as the feasibility, of their efforts. Or, in terms of the theological virtues, fully intelligent beings must exhibit love as well as faith.

Like the belief in the capacity of human mind to understand the world, the Western belief in its accountability to a moral order is rooted in the concept of God as creator and lawgiver. Christians throughout history have rightly been critical of science when it was pursued out of self-interest or even merely in the

national interest. Such, for instance, was the early Christians' critique of Greek science as they knew it.³⁵ And such was the fundamental ethic that motivated the pursuit and publication of early modern science.

For all the credit due Copernicus for his restructuring of our understanding of the solar system, it should be remembered that he originally planned to publish only some astronomical tables and rules of calculation for "common mathematicians" and was reluctant to publish the underlying theory and proof of his results. This was not out of fear of the Church, by the way, but due to a Pythagorean principle of reserving advanced forms of knowledge for personal associates and students. The fact that Copernicus did publish the full theory was largely due to the insistence of Tiedemann Giese, the bishop of Kulm. Against Copernicus, Bishop Giese argued that the Pythagorean practice of secrecy had no place in mathematical science and that the latter should be devoted to Christ and a gift to the world (for which Christ also gave his life).³⁶ The fact that the Church later placed Copernicus's work on the Index should not be allowed to obscure our indebtedness to Giese and other church officials for their positive role in science.

SCIENTISTS SUCH AS BACON AS MODELS OF INTELLIGENCE

From a theological perspective, then, true intelligence entails the consideration of how one may conform to the moral law, or in the specifically Christian sense, how one may follow the example of Jesus Christ. In order to parallel our earlier discussion of "Copernican intelligence," I shall refer to this aspect of human nature as "Baconian intelligence." It was Francis Bacon, whose insistence that science be pursued not for personal gain but for the benefit of humanity, who provided the basis for a socially-supported program of scientific research in seventeenth-century England.

As Bacon put it in *Of the Advancement of Learning*, science should not be a shop for profit or sale, but a storehouse for the glory of the Creator and the relief of the human estate.³⁷ Accordingly, the citizens of the New Atlantis, the Christian utopia that Bacon set forth as the model of the scientific community, daily prayed to God "for the illumination of [their] labors and turning them into good and holy uses."³⁸ Baconian intelligence in this sense is the ethical correlate of the theological understanding of human nature as created in the image of God.

Like the epistemic correlate, the ethical tells us two things that may enrich AI research. It tells us something about the character of the intelligence AI workers may hope to reproduce, and it tells us something about what we should expect from AI workers themselves.

TESTING A FULLY SOCIALIZED COMPUTER FOR BACONIAN INTELLIGENCE

Under what circumstances, then, could fully socialized computers (assuming again that such machines can be developed) exhibit intelligence in the Baconian sense? I would look for situations in which the need for moral judgment

characteristically arises--situations in which various alternative futures would be recognized as possibilities, and in which priorities and criteria would have to be developed to deal with the varying costs and benefits to others as well as to oneself and one's own group. Could a fully socialized computer protest against a policy that was advocated by its human associates?

Could a computer contemplate risking its own well-being in order to avoid harm or to promote the well-being of a human or of another computer--not just because it was programmed to (as Isaac Asimov has imagined in his "First Law of Robotics")³⁹ but as a matter of reasoned moral judgment? In other words, could a computer recognize itself and others as participating in a moral order? Could it conceivably exhibit even greater concern for morality than its creators?

The answer to these questions may depend on the degree to which the AI workers themselves exhibit ethical concerns in their work. After all, it is they who would function as the "parents" in relation to these new members of the community of intelligent beings. Like faith, morality is something that is learned from situations and examples. It can not be schematized or programmed. To what extent, then, will AI research be motivated by the desire to enhance the environment or alleviate human suffering? To what degree will AI workers address the issues of the social impact of the results of their work? The viability of the long-range goals of AI research may well depend on the faith and love of the AI community as much as it does on the possibilities allowed by the electronic properties of matter. And, even if AI research fails ultimately to reproduce true intelligence, it may at least force us to develop neglected aspects of our own. That is what a theological understanding of human nature can contribute to the enterprise.

Notes

1. John F. Clarkson et al., S. J., trans., *The Church Teaches: Documents of the Church in English Translation* (Herder, 1955; repr. Rockford, IL: Tan Books, 1973), pp. 142-49. The preparatory schema for Vatican I's "Dogmatic Constitution of the Principal Mysteries of the Faith" gave the further details that the soul is "immaterial, incorruptible, immortal, and gifted with intelligence and free will" and that the dual composition of human nature does not prevent the latter from being single (*ibid.*, pp. 150-151), but this statement was never adopted by the Vatican Council in formal session.
2. Among the Reformed standards are the Second Helvetic Confession, VII.6; the Heidelberg Catechism, Q. 57; the Scots Confession, XVII; and the Westminster Confession, XXXII (Philip Schaff, *Creeds of Christendom*, 3:247, 325, 459, 670-671).
3. Irenaeus, *Against Heresies*, 2.28.2.
4. Basil, *Hexameron*, 1.8-10.
5. Theophilus, *To Autolytus*, 2.4.

6. Deborah M. Barnes, "Brain Architecture: Beyond Genes," *Science* 233 (11 July 1986): 155-156.
7. Peter Berger, et al., *The Homeless Mind: Modernization and Consciousness* (New York: Random House, 1973).
8. J. David Bolter, *Turing's Man: Western Culture in the Computer Age* (Chapel Hill: University of North Carolina Press, 1984), pp. 8-14.
9. The relation of AI research to the Western philosophical tradition has been described by Hubert Dreyfus, *What Computers Can't Do* (New York: Harper & Row, revised ed., 1979), pp. 67-70, 231-234, *passim*; "Making a Mind Versus Modeling the Brain," *Dedalus* 117 (Winter, 1988): 16-18, 26-30; John Haugeland, *Artificial Intelligence: The Very Idea* (Cambridge, MA: MIT Press, 1985), pp. 15-45.
10. On the deep connection between the cosmic and moral orders in ancient Near Eastern and Old Testament thought, see Rolf Knierim, "Cosmos and History in Israel's Theology," *Horizons in Biblical Theology* 3 (1981): 96-100; Douglas A. Knight, "Cosmogony and Order in the Hebrew Tradition," *Cosmogony and Ethical Order*, ed. Robin W. Lovin and Frank E. Reynolds (Chicago: University of Chicago Press, 1985), pp. 139-140, 148-149.
11. Ernan McMullin, "Augustine of Hippo," *Dictionary of Scientific Biography* (16 vols., New York: Scribners, 1970-1980), 1:337; A. C. Crombie, *Augustine to Galileo* (2d ed., 2 vols., Cambridge, MA: Harvard University Press, 1961) 2:100, 118.
12. Biblical texts which stress the limits of human knowledge generally do so on the basis of limited access to the areas of reality under consideration: Deut 30:11-14; Job 11:7-9, 28:12-22, 38:16-38, 39:1-4; Eccl 8:17, 11:5; Jer 31:37; Sir 1:1-3, 3:21-24, 24:28.
13. Note especially 1 Enoch 17-18, 33 and the "Book of Heavenly Luminaries" (1 Enoch 72-82) which probably date from the third century B.C.E.
14. Tertullian, *Treatise on the Soul*, 18; Augustine, *City of God*, 19.18.
15. Augustine, *On the Trinity*, 14.4.6, 14.20.
16. Thomas Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 1962, 1970).
17. "Faith," as I use the term here, is not opposed to reason. It is simply a more comprehensive view of things that includes reason and gives it legitimation.
18. For a brief review of these and other AI research programs, see Dreyfus, *Mind Over Machine* (New York: Free Press, 1986), chap. 3; Marvin L. Minsky and Seymour A. Papert, *Perceptrons*, expanded ed. (Cambridge, MA: MIT Press, 1988), Prologue. For more detail, see M. Mitchell Waldrop, *Man-Made Minds* (New York: Walker, 1987), chaps 1-8.

19. There is also evidence that Neanderthals cared for the handicapped and the elderly--indications of the kind of moral awareness discussed later in this paper; see Richard E. Leakey, *The Making of Mankind* (London: Sphere Books, 1982), pp. 152-153.
20. The study of fossil homonids is one context in which moderns continue to ask basically theological questions. Who were the first humans? That is, how far back can we go and find creatures with whom we could share a social life in the pursuit of common goals? Though answers may range from the dryopithecine or ramapithecine ancestors of the great apes (15 to 20 million years ago) to Upper Paleolithic *Homo sapiens sapiens* (35,000 years ago), the quest that attracts so much attention in the media is really another way of asking the basic questions of the book of Genesis: Who were the first humans, what did they do that was different and how does that affect us? The quest for the first signs of truly human intelligence is in a way the mirror image of the quest for AI. One attempts to define our nature in terms of its origin; the other, in terms of its future capabilities. Both efforts reflect the distinctively Western interest in that which makes it possible for us to understand realms of reality beyond our immediate or past experience--intelligence in the Copernican sense.
21. Dreyfus, *What Computers Can't Do*, p. 156; *Mind Over Machine*, pp. 76-90; "Making a Mind," pp. 24-33.
22. John McCarthy, "Some Expert Systems Need Common Sense," *Annals of the New York Academy of Sciences* 426 (1984): 129-35.
23. This is known as the "frame problem": Haugeland, *Artificial Intelligence*, pp. 203-211.
24. C. M. Rogers and R. K. Davenport, "Effects of Restricted Rearing on Sexual Behavior of Chimpanzees," *Developmental Psychology* 1 (1969): 200-204.
25. Minsky and Papert, *Perceptrons*, pp. 271-272.
26. One could possibly derive the requirement of social relationships directly from the notion of the image of God itself. Genesis 1:26-27 specifically describes the image as something shared by humans in relation to each other ("male and female") and distinct from the society of other living creatures (cf. Gen 2:18-25). However, this view of the image of God has never received general acceptance among theologians: H. D. McDonald, *The Christian View of Man* (Westchester, IL: Crossway, 1981), pp. 36-37.
27. Needless to say, these requirements go far beyond what most AI workers are ready to work toward at present, but it does remain a long-term goal. See, e.g., Roger C. Shank, *The Cognitive Computer* (Reading, MA: Addison-Wesley, 1984), pp. 60-63, 117.
28. I take it that the process of human design does not rule out the possibility of the product participating in the divine image (or, at least, understanding itself as so participating) any more than natural evolution does. As Alan Turing pointed out in 1950, the stipulation that a machine could not possibly be endowed with a soul would constitute a "serious restriction of the omnipotence of the Almighty": "Computing

Machinery and Intelligence,” reprinted in *Computers and Thought*, ed. Edward Feigenbaum and Julian Feldman (New York: McGraw-Hill, 1963), p. 20.

29. In fact, computer programs have been developed that can derive Kepler’s third law from the data: Pat Langley et al., *Scientific Discovery: Computational Explorations of the Creative Processes* (Cambridge, MA: MIT Press, 1987), pp. 66-120.

30. Stanley Jaki, *Science and Creation* (Edinburgh: Scottish Academic Press, 1974), p. 268.

31. Letter to Herwart von Hohenburg, 9 April 1599, quoted in Crombie, *Augustine to Galileo*, 2:195.

32. The difference between the task of cosmologists like Kepler and Einstein and AI workers is, of course, that the former only had to build mathematical models of their subjects—the solar system and the universe, respectively. It was neither possible nor necessary for their purposes to reproduce them. AI workers, on the other hand, must reproduce the actual dynamics of their subject in order to understand it. Their work is more like that of chemists and nuclear physicists in that regard.

33. Dreyfus, *What Computers Can’t Do*, pp. 303-305.

34. Walter Pagel, “The Debt of Science and Medicine to a Devout Belief in God,” *Journal of the Transactions of the Victoria Institute* 74 (1942): 105-108; Eugene Klaaren, *The Religious Origins of Modern Science* (Grand Rapids: Eerdmans, 1977), pp. 63-71.

35. C. B. Kaiser, “The Early Christian Critique of Greek Science,” *Patristic and Byzantine Review* 1 (1982): 211-216.

36. Rheticus, *Narratio Prima*, in *Three Copernican Treatises*, trans. Edward Rosen (3d ed., New York: Octagon Books, 1971), pp. 192-195.

37. *Works of Francis Bacon*, ed. James Spedding et al. (14 vols., London, 1857-1874), 6:134.

38. Clarence J. Glacken, *Traces on the Rhodian Shore* (Berkeley: University of California Press, 1967), p. 475.

39. Isaac Asimov, *I, Robot* (New York: Ballantine Books, 1983). The question of “machine ethics” has been taken up by Mitchell Waldrop, *Man-Made Minds*, chap. 11.

