

Theological Reflections on Artificial Intelligence

Edward H. Schroeder

[ITEST Conference, St. Louis, MO, 1984. Reprinted with permission]

[Edward H. Schroeder is professor of systematic and historical theology at Christ Seminary-Seminex. An ordained clergyman, he has taught theology for 27 years both at Valparaiso (Indiana) University and at the Lutheran seminary in St. Louis. His fields of specialization are the theology and history of the Reformation, missiology, and theological ethics. He is currently active in a St. Louis based postgraduate school for laity called Crossings, a community of theological study for peoples' secular callings. Dr. Schroeder is married and the father of three children.]

"The purposes of ITEST are: (1) to act as an 'early-warning system' for the church on work being done in the scientific laboratories; (2) to translate this information into a theological/ecclesial vocabulary; (3) to identify, isolate, and respond to those scientific developments that affect Christian belief; (4) to explore the growth of Christian teaching in these matters; (5) to build a community of scientists who are dedicated both to the advancement of scientific understanding as well as to the growth of the church." (ITEST Bulletin of March, 1982)

The first four of these five statements constitute the outline of this essay. Number five, I trust, will happen when we gather in March.

I. "Early-warning System" for the Church on Work Being Done in Artificial Intelligence.

Read the attached essay by Joseph Weizenbaum, "The Computer in Your Future," The New York Review of Books (October 27, 1983), a review of The Fifth Generation: Artificial Intelligence and Japan's Computer Challenge to the World, by Edward A. Feigenbaum and Pamela McCorduck (Addison-Wesley, 1983).

Here is what my radar screen picks up from the review:

1. The authors note that sadly most intellectuals ignore the computer revolution, thereby opting for illusion, irrelevance, obsolescence. If they survive at all it will be from the charity of those in the know, the AI insiders.

2. What's happening in the world of computers? Five items. a) Computer scientists are moving into other fields, e.g., medicine; b) Computer architecture is radically changing; c) Coupled with new computer languages things go much, much faster; d) More with less, more for less; e) The Japanese are ahead of us in putting these developments together for AI machines that constitute a new "fifth" generation that can reason, make judgments, understand.

3. The authors' claims are exaggerated. Prospects are dim that these machines can do what the words "reason, judgment, understand" usually mean.

4. The authors envision the fifth generation AI world: the geriatric robot does a better job of caring than humans do, as do other human replacements, "without the prejudices (e.g., the criticism of a human physician) that often attend face-to-face interaction." Weizenbaum scorns such computer-human interfacing as a replacement for face-to-face and touch contact.

5. The totalitarian shape of the projected AI society is a Faustian bargain, a problem-solving technology that insulates us from our problems, shields us from allegedly unpleasant contact with fellow human beings, homogenizes us into Orwell's 1984. And this is urged as salvation. Every Faustian bargain must be asked: what benefit, and who gets it? The answers here are not good.

6. The authors affirm without qualification the Star-wars benefit of AI for military purposes: smart weapons, zero probability of error, sent off (as another writer put it) "to do jobs that now involve human risk." Weizenbaum marvels that this doesn't give AI people pause.

7. There is no small amount of self-promotional rhetoric in the book alongside some inaccurate (improper?) reporting on the computer patriarch von Neumann.

8. On the cardinal issue arising from the von Neumann disagreement: is everything human describable, and all human knowledge sayable? Weizenbaum says no, Feigenbaum yes. Weizenbaum faults Feigenbaum's casual arguments for his yes to this cornerstone of the controversy.

9. The frequent use of the passive voice without personal ascription of deeds and words to doers/speakers fuzzes out "responsibility; it denies the role of human will in human affairs." Here and in the previous #8 with judgments galore, the criteria remain hidden.

10. That computers process information is a truism, but that they will produce information and transfer "the burden of producing the future knowledge of the world ... from human heads to machine artifacts" is a claim for which no warrants are given. And even if they could produce information, on what grounds is that the future knowledge of the world?

11. Summa. The knowledge least well understood by the authors is that of the differences between a society centered on human beings and one centered on machines.

What's the "warning" in Weizenbaum's book review? He "warns" us that AI is already doing subtle theology, and this at the centers of scientific secularity in our country, MIT (Weizenbaum) and Stanford (Feigenbaum). The medium is silicon but the message is soteriology. But already within the AI fraternity (yes, it seems to be a male preserve),

soteriologies are in conflict. The ancient church conflict between Arius & Athanasius could hardly have been more earnest or existential (or even personal) than these debates in AI today. When ITESterS enter the AI discussion, we do not bring theology to a theological vacant scene. A form of theologizing is already going on. We might bring "better" theology, but that needs to be demonstrated.

Weizenbaum's review highlights two issues of AI's sort of theological debate: anthropology and soteriology. The first includes ethics, epistemology, the human image. Where does responsibility, accountability, the word "ought" surface in AI? What is a human being? What is "human" about human intelligence? Is there anything in the "wet carbon" that mediates human thinking which "dry silicon" cannot replicate? What's the difference between human-centered and machine-centered societies?

The anthropological debate is acknowledged among the "artificial intelligensia," to use Louis Fein's teasing term. A Stanford faculty publication (Campus Report, April 13, 1983) informs us of "500 scientists around the world working on (AI) programs and projects." In conclusion it acknowledges:

Their work is not without controversy.

Some critics, like the Dreyfus brothers, Hubert and Stuart at Berkeley, believe that machines will never "really" be able to think -- at least not the way a human thinks. That criticism is not taken seriously by scientists in the field.

Others, like Joseph Weizenbaum of MIT ... believe there are moral questions that must be answered by AI researchers. His criticism is taken more seriously, but most AI scientists think his fears are overblown.

Most seem more intent on the work itself than on what it will lead to at the moment. They are convinced the result will be a revolution, even if they are not sure just what kind of world that revolution will build.

The soteriological agenda is more subtle, though no less real despite the hype. Says John Sculley, president of Apple: "The personal computer is going to free our lives." Says Robert Jastrow, NASA futurologist, ". . . the computer will minister to our social and economic needs it will become man's salvation in a world of crushing complexity."

Just what is AI? In the Stanford faculty report cited above we have this definition:

Essentially, AI research can be divided into four areas.

1. The best known is probably robotics, the study of machines that simulate human physical activity.

2. Another is known as expert systems, computer programming that attempts to solve problems and make decisions at the level of a human expert in a particular field. This is sometimes known by the general term "knowledge engineering."

3. Another field is language understanding, which includes things like translating from one language to another, a process which requires an understanding that goes well beyond word-for-word substitution.

4. Finally there is what can only be described as epistemology, research using computers into human thought at its most philosophic, almost metaphysical, level.

"...almost metaphysical" may be Stanfordese to stay on the right side of the line between scientific and non-scientific. To say "almost theological," or just plain "theological" may be too daring. But some, like Pamela McCorduck, co-author with Feigenbaum of *The Fifth Generation*, are less hesitant.

In the concluding chapter of her work on "the history and prospects of artificial intelligence," *Machines Who Think* (1979), she concludes with a chapter on "Forging the Gods." The pun in the participle is intended, she says, creating and/or counterfeiting.

We're unquestionably in the business of forging the gods.... The enterprise is a godlike one, rightfully evoking terror in the hearts of those who think the borders between humans and gods ought to be impermeable. But the suspicion has been growing for some time that gods are a human invention. . . The invention -- the finding within -- of gods represents our reach for the transcendent at the same time we reach to understand ourselves as humans in a human world.

And so with the reflexive enterprise of artificial intelligence. We are as gods in the exercise, counterfeiting aspects of the human just as we always have, whether in theology or the arts, and for pretty much the same reasons of self-enchancement. That we might be forging gods in the other sense -- deities to rescue us from our own over-reaching -- is an idea allowing degrees of accord. Certainly artificial intelligence builders fall all along the line from dubious no to enthusiastic yes. So might we all.

Pausing for a moment to remember Isaiah's harsh words about idolatry she wonders "whether it was the creation or the worship which is objectionable," and then concludes:

Me, I breathe easy on the whole subject, being of a Hellenic, rather than Hebraic turn of mind. The accomplishments have been significant, and the promises are nearly beyond comprehension. I pause just now, before I have to call forth fortitude and exhibit courage. I pause to savor the thrill of sharing in something awesome. (p. 355-57)

- II. To Translate the Signals into Theological/Ecclesial Vocabulary, and
- III. To Identify, Isolate, and Respond.

In the 1975 ITEST Spring Conference Vice-director Bob Bertram started his paper with "Basil's complaint." The 4th-century Cappodocian theologian complained that his people in Caesarea were "busy with technology, but not theology." Bob noted that "the situation today is not unflatteringly different." In AI today the situation is somewhat different, though probably not flatteringly so. It is not flattering for our kind of Christian theology which is mostly notable by its absence. To borrow from St. Paul, there is a zeal for God of sorts, but it is not enlightened, specifically not Christicly enlightened. Amongst themselves AI people talk about "playing God," some warning Hebraically against that yea, others encouraging us all to "savor the thrill of sharing in something awesome."

In the following theses I shall try my hand at the translation and response task.

1. There is a soteriological and an anthropological controversy in the AI community.
2. The soteriologies in conflict have classical contours signalled by the terms Hellenic and Hebraic. Two "gospels" ground the two soteriologies: one technologico-gnostic, the other personalist-moral.
3. The redeemer figures in each are human, but no one human being. Extraordinary knowledge characterizes the one redeemer figure, extraordinary moral courage the other.
4. The Christian Gospel of salvation through Jesus the Christ is an "other" gospel to these two.
5. The anthropological debate is also nuanced Hellenic vs. Hebraic.
6. Two images of the human contend: an information-processing reasoning being vs. an accountable information-processing reasoning being. In New Testament rhetoric: one places the accent on gnosis, the other on epignosis (knowledge or acknowledgement). One is a consciousness-centered image, the other a conscience-centered image.
7. The importance of the human body for human being and human intelligence is in dispute. Is the body incidental or fundamental to knowing and why does the Biblical tradition come down on the pro-body side?
8. Re human life as problem-solving: to solve a problem that one "has" is one thing. To solve the problem that I "am" is out of my hands. Theologically stated: Sinners are first of all a problem for God and God's creation, and only subsequently is it also true vice versa.
9. What is the connection between suffering and knowledge? Is it: Suffering can be avoided through knowledge, or: Suffering precedes all knowledge? How to relate the latter epistemology to AI is a Christian contribution to the epistemology discussion.

10. Re human accountability in AI, we see one option minimizing it lest it be "overblown" while the other welcomes it and urges its expansion. Christian proclamation finally labels both options "insufficiently critical," the former quite obviously for its "Cain complex, the latter for its more subtle expectation that "salvation could come by the law."

11. Biblically viewed, creaturehood is limited existence. What is the particular "limit" in human creaturehood? Even if the limits of intelligence may be fluid, these are not: we cannot escape (although we may deny) our accountability, & we are really not in control of our own destinies.

12. This impacts on the debate about the calculability of everything in nature & thus everything human. By virtue of the willy-nilly creator-connectedness of every human being (call it mystery or terror, marvel or McCorduck's "awesome"), not everything human is calculable. Since God is not reducible to silicon circuitry, neither is the *imago dei*, even in her fractured state.

13. To have to reckon only with our silicon-similarity would be a relief if it would liberate us from the necessity of interfacing with the One Whose image we are. Unfortunately, the living truth is otherwise. In the long run, however, there is advantage in that adversity.

14. The good news of the Christian Gospel is framed in the metaphor of new image and new (inter)face: The crucified and risen Christ, says Paul, is the new image of God, the same God who said, "Let light shine out of darkness," (=the first *imago dei* program) who has shone in our hearts to give the light of the knowledge of the glory of God in the face of Christ (II Cor.4:4ff.).

IV. To Explore the Growth of Christian Teaching in These Matters.

If, as I have proposed, Christian theology confronts two salvation schemes with corresponding anthropologies in the AI world today, one gnostic, the other moralist, then we have some helpful precedents in the New Testament for our own exploring the growth of Christian teaching in these matters. Pneumatic gnosticism and moralizing legalism were the two most tempting "other" gospels confronting the first-century Christians. Christian theology "grew" in the first century primarily from its (dare I say it?) interface with these two movements. Drawing on those primal resources I shall make a few exploratory probes. First a reconstruction of the gnostic proposal in AI.

Technological Gnosticism

The soteriological turf needing redemption in AI gnosticism is our problem-ridden planet and the human dilemma of surviving on it with problem-ridden people as well. The current malady is rooted in our insufficient information and our inability to solve our survival problem(s). AI, a power already promising more-than-human (McCorduck says "divine") resources for problem-solving, is already on the horizon. What it takes to bring

it over the horizon is already present in the AI community. Silicon is the divine datum and AI experts know how to make it work for salvation. The redemption model is elementary. Problem: woeful ignorance in the general populace--documentably so. Solution: the super-gnosis of the super-computers of AI. Redeemer agents: the heavenly humans who by gift/intuition can manage AI to usher us out of our malady into the millenium.

Oh yes, there are, unfortunately, some hylotes ("blockheads") who never catch on, or refuse to catch on. Too bad for them. Not all can be saved. The redeemed are an elite, not necessarily a minority, but less than the whole.

A Personalist-Moralist Alternative

Weizenbaum's response to the gnostic messianism, I suppose, sounds much better to most of us. He disbelieves the gnosticism. Much of his disbelief is technical. The gnostic gurus cannot do what they claim (soon) to be able to do. Besides, the soteriological need is not for knowledge in the first place, surely not AI's kind of knowledge. There is an ignorance widespread in the land, to be sure -- maybe even stupidity. But it is moral stupidity: people making Faustian choices ignorant of the consequences. In one of his papers he lists some questions "almost never asked" by AI researchers: 1) Who is the beneficiary, who the victims? 2) What limits are proper for the application of computation to human affairs? 3) What about the almost automatic link to warfare? 4) What is the computer doing to the image human beings have of themselves? 5) What irreversible forces are we setting loose? 6) Can our kids live with the world we are herewith creating?

Weizenbaum is, I think, not proposing a decalogue that he possesses. Rather he dwells on responsibility and accountability themselves as missing items (=the soteriological defect) in the AI world. When he talks about "limit," it is not that he has a boundary line clearly in focus, not necessarily. Rather he is calling for acknowledgement of limit in the work of AI. To accept responsibility is to accept limits.

Programmers can ... gain a sense of responsibility and dignity by refusing to add to systems whose purposes and theories of operation cannot be explained to them. How else ... could programmers possibly be sure that they are not working on systems to whose ultimate purposes they could not reconcile their consciences. Clearly to accept one's responsibilities is a moral matter. It requires one, above all, to recognize and accept one's own limitations and the limitations of one's tools. Unfortunately, the temptations to do exactly the opposite are very large.

What is needed for "the construction of reliable computer systems in the social and political sphere" is not more

research in computer science, but rather a deeper theoretical understanding of the human condition. The limit, then, of the extent to which computers can help us cope with the world of human affairs is determined by what similar such limits have always been determined by: our ability to assess our situation honestly and to know ourselves.

Salvation lies not in the direction of technological gnosis, but moral epignosis; not knowledge for technological problem-solving, but acknowledgement of my responsibility for the problems I have, the problem I am, and the problems I create.

Some Seed for Growth in the Soil of AI

In the final segment of this paper I propose that we use the second lesson appointed for the first Sunday in Lent (= the weekend of our conference) as our theological grounding to explore the growth of Christian teaching in the direction of AI. It is Romans 5:12-19. Please read it searchingly before you go on here.

In his Adam-Christ typology here Paul spells out his anthropology and his soteriology. Paul's anthropological proposal is binary: Adamic anthropology and Christic anthropology. Salvation amounts to being rescued from Adamic human solidarity and translated into Christic human solidarity. His step-for-step diagnosis of Adamic humanity and his prognosis of Christic humanity looks like this:

- | | |
|--|--|
| 1. Initial Diagnosis
The one-for-all Adam is the human race that replicates sinner Adam's biography in the world. (See Genesis 3, the first reading for Lent 1, for details.) | 4. Good News for the Final Diagnosis
The one-for all human Jesus Christ's act of righteousness brings life-acquital to the race formerly in solidarity with Adam, a free gift. |
| 2. Advanced Diagnosis
Worse yet, Adamic biographies elicit death and spread death so that death reigns over all. | 5. Good News for the Advanced Diagnosis
Better yet, receivers of this free gift now do the reigning in life through the one-for-all human Jesus Christ. |
| 3. Final Diagnosis
Worse still, Adamic biographies get evaluated (krima)-- as in Genesis 3 -- and that brings katakrima, incriminated death. | 6. Good News for the Initial Diagnosis
Better still, the abundant overflow of grace ("much more" than its Adamic typos) types its recipients "righteous." That overflow brooks no limits. |

Some Lines for Exploratory Growth

1. God's evaluation and salvation are embodied events of human solidarity. Not being AI competent I can do little more than notice that its gnostic salvation proposal ignores the critical evaluation Paul proposes and is body-shy all around. Those two behaviors go together. The moral responsibility proposal does attend to the fact that the Adamic race is involved in moral judgments, but its diagnostic depth is considerably less than St. Paul's. To work on this project from just the side of a theologian promises no growth. It can only be done collaboratively with someone who "knows the AI territory."
2. If an embodied Christian soteriology could grow into winsomeness for proponents of either of the two alternatives we've been addressing, then the promise of AI and of moral commitment could be relieved of their soteriological pretensions. Although it is the chronic yen of humans to create or become salvific artifacts, McCorduck is right: there is a difference between making and worshipping what we have wrought. It takes more than insight, however, to cure idolatry. Paul's paradigm is death and resurrection. One consequence of that dying and rising is Paul's prognosis #3. That righteousness is the right use of things as people's servants, not their saviors. In more ways than one it is not right to want AI to set us free.
3. The contrast between being ruled by death and reigning in life has something helpful to say to the sticky webbing that keeps AI and weapons production together. I am not clear as to what that means. But it is the genius of the Christian gospel to open third alternatives when Adamic epistemology allows for only two options, both of which are often what Weizenbaum calls Faustian choices.
4. The foundation of the new humanity on a vindicated Suffering Servant Messiah also has growth possibilities, I would suggest, in the ethereal AI worlds. Especially in that fourth area ("almost metaphysical") where they probe the links between knowledge and experience, they might probe suffering's link to knowledge. The vastest untouched area, of course, is the suffering world of the majority human race. How AI might be turned to their service, freed of the ideology in which it is embedded, I do not know. But if AI is one of the creator's gifts to the human race, then it also belongs to the least of the siblings of the original AI (Authentic Incarnation) of the **new** human race.

Edward H. Schroeder