

Letters

■ Editor:

After discussions with other members of the medical AI community, I feel I need to lodge a mild cautionary complaint about the cover of the Summer 1989 *AI Magazine*. Those of us in medical AI have been highly sensitized to common misunderstandings of our goals, intent, and motivations. We also encounter a general lack of realistic expectations regarding the role of computers in clinical areas. Several people in the AIM community have noted with annoyance the recent cover because it further feeds the notion of computers as dehumanizing and impersonal forces in the medical scene. I think we all would have hoped for more sensitivity on this point from those involved in putting together the publication. Several of us have imagined how a slide of the cover could be misused by those who might want to misrepresent further what AIM is all about. Although I realize that the picture accompanied an article that was anticipating the impact of AI on our future society, isn't the medical example a bit sensational and hackneyed?

Ted Shortliffe

Departments of Medicine and
Computer Science
Stanford University

■ Editor:

The discussion by Randall Davis, et al. ("Expert Systems: How Far Can They Go?", *AI Magazine*, Vol. 10, Nos. 1 and 2) prompted the following reflections.

A persistent conundrum in AI has been whether AI can fulfill its scientific goal of elucidating the structure of human thought. Mr. Smith suggests the basis for a positive answer: both humans and computer employ models about the world. But there is a paradox here: this interpretation of human cognition as involving models is itself a model—an abstraction that "does violence to its subject," something that by definition is "partial." And here we find the source of the perpetual philosophical debates on

whether computer models elucidate human thought. These debates end up merely as arguments in which one model (of human intelligence) is compared with another (of what artificial intelligence is). Depending upon what properties of human and artificial intelligence are stressed we get differing, usually opposing, views of the relationship.

The problem is that the models of human intelligence that have entered into our philosophical debates often have not been rigorously and scientifically developed, and in the case of the Dreyfus model, are even of dubious theoretical value, as Davis points out. Disanalogies between human and computer "intelligence" abound: humans use intuitive judgment and common sense, humans "compile" knowledge (compilation itself is a model theoretic analogy), and humans access bits of knowledge in ways not anticipated in the original storing of that knowledge ("accessing" and "storing" knowledge are further model theoretic notions).

But the point is that if we had models of these human capabilities, we could recreate them in a computer program. A good model is a computable model. This raises two questions: (1) Is the discovery of such models an empirical inquiry (as Davis asserts)? In a sense it is. But the Winograd/Flores argument is that whatever model is created will be limited in critical respects: thus the scientific goals of AI are misguided. (2) Assuming that satisfactory models of human thought are possible, would our computer experts ever enjoy the same social trust and "understanding" that human experts do?

Failure of our expert systems today is treated as a representational problem to be addressed by knowledge engineers and programmers, whereas the limits of human expertise must be dealt with by society—and at times probably dealt with poorly (e.g. in the case of forensic psychiatry and the insanity defense). Every deployed

expert system now must be accompanied by a culture for accommodating its limited knowledge representations. Could it ever come about that we have modeled human expertise to such an extent that the limits of the computer system would no longer be a representational problem? What must take place for this to happen? Will we need to ascribe pleasure and pain to our computer experts? Must we see them as having desires? Must we be able to program a computer to make love (assuming we have the right hardware)? At what point does the need to model human consciousness stop? I find Mr. Hill's view ("The Mind at AI: Horseless Carriage to Clock," *AI Magazine*, Vol 10 No. 2, p. 36) regarding the switching of the I-you intentional stance to be more likely. But there's a problem here as well: there will be those who would cling to the "expertise" of the system for their own purposes when in fact the switch to the "dumb machine" stance would be appropriate.

As scientifically and morally intriguing as these questions are, my real problem is more immediate: What are the "systematic domains" to which the current state of the art applies in spite of its limitations? I need to understand these domains and have better tools to handle them, tools that help the knowledge-based solutions "fit into the real world," to use Mr. Smith's excellent phrase. Here is a wish-list of what is needed: standards for knowledge-based system programming (Ken Pedersen's new book, *Expert Systems Programming: Practical Techniques for Rule-Based Systems*, Wiley, New York, 1989, and the RIME methodology for XCON are initial steps); more domain specific shells, with knowledge acquisition capabilities, that address paradigmatic reasoning in the systematic domains (in more than diagnostic and process control domains); better explanation capabilities than mere rule-tracing; better natural language interface capabilities; and more graceful (controlled and informative)

SERVING THE ARTIFICIAL INTELLIGENCE COMMUNITY

Halbrecht Associates, founded in 1957, was the first search firm in the United States to specialize in Information Systems Management, Operations Research, Management Sciences, Telecommunications, Computer Sciences and Decision Support Systems. We are now also in the forefront of the Artificial Intelligence field. Because we have been so extensively involved in this emerging technology, our portfolio of people includes extensive numbers of AI professionals at all levels, from the United States and abroad.

Our client list includes companies coast to coast, large and small, including start-ups with equity opportunities. Accordingly, depending on your needs, we are in a unique position to assist you rapidly with any of your requirements.

Our current assignments range downward from several Chief Scientists, or those who can provide the technical leadership to head up major AI programs, to Senior LISP Programmer. If you would like to be considered to be part of our AI data base, please send us your credentials. This will enable us to give you first notice of appropriate opportunities. All inquiries are in strictest confidence.

Contact: Daryl Furno, Senior Associate

1200 Summer Street • Stamford, CT 06905 • 203-327-5630

For free information, circle no. 110

degradation when a shell reaches the limit of its knowledge.

Kirk D. Wilson
Senior Systems Analyst
Air Products and Chemicals, Inc
Allentown, PA 18195

■ Editor:

While clearing some shelves into a back area (and browsing a bit), I came across an older article "Knowledge Base Verification," in *AI Magazine*, Vol.8, No.2. A problem that I see arises from trying to automate such a process. (The authors keep a man in the loop.)

Suppose you want to automatically eliminate subsumed rules:

One rule: If A then C
Other rule: If A and B then C

Pure logic says eliminate the "other" rule. But pure logic is not necessarily an adequate practical criterion, especially here. In practice, of course, a knowledge base is an ever changing entity, always growing more complex as reality comes more clearly into focus.

It is entirely possible that the "other" rule was added later when it

was recognized that the first rule was incomplete, while accidentally leaving the first rule behind in the data base. If this is the case, then, as opposed to the "pure logic" approach, the first rule not the other rule should be eliminated, otherwise an automatic system is throwing away the new data and keeping the old.

Clearly what is lacking is a time stamp on the rules (or at least a record of the order of entry of the rules). With "time" in the loop, an automatic verification system can be prevented from discarding newer rules in favor of older ones.

David M. Brender
W201 Youmans Road
Wells Bridge, NY 13859

■ Editor:

The Committee on Human Factors of the National Academy of Sciences is interested in achieving greater awareness of the reports it publishes. To do this we rely on the good offices of organizations such as yours to make their membership aware of our reports. These include the following: *Research Needs for Human Factors* (1983), *Research Issues in Simulator*

Sickness (1984), *Research Needs on the Interaction Between Information Systems and Their Users* (1984), *Research and Modeling of Supervisory Control Behavior* (1984), *Designing Software to Fit Human Needs and Capabilities* (1985), *Human Factors Aspects of Simulation* (1985), *Mental Models in Human-Computer Interaction: Research Issues About What the User of Software Knows* (1987), *Human Factors in Automated and Robotic Space Systems* (1987), *Ergonomic Models of Anthropometry, Human Biomechanics, and Operator-Equipment Interfaces* (1988), *Human Factors Research and Nuclear Safety* (1988), and *Human Performance Models for Computer-Aided Engineering* (1989).

Recent reports are available free of charge on a limited basis, and all reports may be obtained at a nominal charge on a publication-on-demand basis from the National Academy Press.

Harold P. Van Cott
Committee on Human Factors
Commission on Behavioral and Social Sciences and Education
National Research Council
2101 Constitution Avenue
Washington, D.C. 20418