Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer

Hubert L. Dreyfus and Stuart E. Dreyfus (The Free Press; New York, N.Y., 1986, $16.30, hardcover)

During the past several years, we have witnessed a deluge of articles in the popular press dealing with artificial intelligence and visions of society in which machines that can think like humans may play a dominant role. To me, this is a case of *déjà vu.* Back in 1950, while I was an electrical engineering instructor at Columbia University, I wrote an article entitled "Thinking Machines—A New Field in Electrical Engineering" published in the *Columbia Engineering Quarterly.*

Alluding to headlines of that period, the opening paragraph of my article read:

> 'Psychologists Report Memory is Electrical,' 'Electronic Brain does Research,' 'Scientists Confer on Electronic Brain'—these are some of the headlines that were carried in newspapers throughout the nation during the past year. What is behind these headlines? How will 'electronic brains' or 'thinking machines' affect our way of living? What is the role played by electrical engineers in the design of these devices?

Of course, a great deal has happened since then. Computers have gained many orders of magnitude in computational capability, with massively parallel computers becoming at last a reality. We know so much more about software, networking, logic programming, automated reasoning, and much else. And yet, achieving the machine intelligence prophesied by AI pioneers like Minsky, McCarthy, Newell, and Simon has proven an elusive goal. Why? This is the principal question that Hubert and Stuart Dreyfus—members of the philosophy and industrial engineering and operations research departments at UC Berkeley—attempt to answer in *Mind Over Machine.*

Ever since he wrote his Rand report in 1965, "Alchemy and Artificial Intelligence," Hubert Dreyfus has been regarded as an enemy of AI. The artificial intelligentsia has dismissed his criticisms as those of a philosopher not knowing enough about AI to qualify as a critic. And yet, there are many influential members of the AI community—Terry Winograd among them—who concede that Dreyfus was right and they were wrong when they made unrealistic predictions about what AI would be able to accomplish. Dreyfus' favorite (and unrepentant) target in this respect is Herbert Simon, who predicted in 1965 that "machines will be capable, within 20 years, of doing any work that a man can do."

Why have so many excessively optimistic predictions been made about the future of AI? Dreyfus attributes it to what the late Y. Bar-Hillel called the "fallacy of the successful first step." In science, we have been conditioned to expect that success on a small scale can be extended to larger systems. But in AI, methods showing promise when applied to a toy problem (or a microworld) fail completely when applied to more realistic problems. Among the examples cited in *Mind Over Machine* is Winograd's natural language understanding program (SHRDLU) that created so much excitement when announced in 1972. Four years later, an MIT AI memo quoted by Dreyfus had this to say:

> AI has done well in tightly constrained domains—Winograd, for example, astonished everyone with the expertise of his blocks-world natural language system. Extending this kind of ability to larger worlds has not proved straightforward, however. . . . The time has come to treat the problems involved as central issues.

The Dreyfuses stress an AI limitation; specifically, its difficulty in dealing with commonsense knowledge. I do not agree with the authors, however, that AI is intrinsically incapable of coming to grips with commonsense reasoning. The limitation, as I see it, results from the use of first-order logic—a logical system that makes no provision for uncertainty, imprecision, and exceptions to rules. These aspects of commonsense knowledge can readily be formalized within fuzzy logic—basically a logic of approximate reasoning. In this perspective, AI based on fuzzy logic, rather than on traditional first-order logic, may be less vulnerable to the Dreyfus criticisms.

Clearly, one must differentiate between what cannot be achieved today and what is totally beyond the reach of machine intelligence now and in the future. What riles the AI community most is one of this book's central claims; namely, that machines will never come close to humans in performing cognitive tasks requiring intuition and holistic thinking. The case for this claim is drawn in the first three chapters, in which the acquisition of expertise is viewed as a process involving five steps: novice, advanced beginner, competent, proficient, and expert. The moral of the five-stage model is that there is more to intelligence than the calculative rationality that computers can mimic.

While it is difficult to argue convincingly about the ultimate limits of machine intelligence, many problems undeniably exist that will not be solved in the foreseeable future—a good example is nonstereotypical story summarization. Actually, at least several story summarization programs exist including the well-known program called Frump (developed at Yale). However, these programs merely illustrate Bar-Hillel's fallacy of the successful first step. They do have a limited ability to summarize stereotypical short stories in a narrowly prescribed domain, such as accounts of vehicular accidents. But these programs are based on methods totally incapable of extension to the types of story one may read in, say, the *New Yorker.*

Rand was a research Mecca and Stuart Dreyfus was a believer.

The inability to summarize would prevent any machine from passing the Turing test of intelligence. In this test, an interrogator is separated from a person (or machine) under interrogation, and communication is carried out on a teletype. If the interrogator could not tell whether the communication was with another person or with a machine, then the subject would be regarded as intelligent. My point is that the interrogator could easily find the answer by typing a short nonstereotypical story and asking the subject to summarize it.

Why is machine summarization much more difficult than machine translation? Primarily because summarization requires much more understanding of subject matter. To achieve such understanding, computers must deal not just with symbols but (more importantly) with real-world denotations. In fact, many machine intelligence limitations not discussed in this book link in one way or another to the absence of bridges between symbols processed by machines and the denotations of those symbols in a universe of discourse.

In his last chapter, Stuart Dreyfus

recants his faith in the effectiveness of mathematical techniques in managerial art and management science. A mathematician by training, Dreyfus was a colleague of Richard Bellman at Rand in the early 1960s and coauthored with Bellman a well-known text on dynamic programming. At the time, Rand was a research Mecca for mathematical applications to decision analysis—especially in dynamic programming, linear programming, and game theory—and Dreyfus was a believer.

Expectations for what these techniques could contribute to real-world decision making proved exaggerated, though not to the same degree as with AI. Stuart makes the point persuasively—a point still not widely accepted by theoreticians—that conventional mathematical models fail to reflect the incompleteness, imprecision, and unreliability of environments in which real-world decisions are made. In this context, Dreyfus urges a greater reliance on intuition. I would add to his suggestion the abandonment of classical quantitative frameworks and the adoption of a linguistic approach in which variable values are words rather than numbers.

The authors deserve to be commended for incisively analyzing the limitations of machine intelligence, supplying a much-needed balance between promise and reality. And yet it's certain that AI, expert systems, and robotics—despite their limitations—will play increasingly important and positive roles in our society. Machines may never be able to compose music like Beethoven, write poetry like Byron, understand speech like humans, climb trees like monkeys, or fly like birds. But they will perform many useful and complex tasks requiring high intelligence levels. Undeniably, our expectations of what machines could accomplish were—and may still be—unrealistic. But we must remember that, as Jules Verne noted at the turn of this century, exaggerated expectations drive scientific progress.

—Lotfi A. Zadeh
Electrical Engineering
and Computer Sciences Dept.
University of California
Berkeley, CA 94720.

This review appears courtesy of *Forefront*, a UC Berkeley College of Engineering publication. Last summer, *IEEE Expert* published an excerpt from the Dreyfus text ("Why Expert Systems Do Not Exhibit Expertise," Vol. 1, No. 2, pp. 86-90).