BOOK REVIEWS

Grand Hopes

The first nine chapters of this book tell the history of "artificial intelligence" (AI), drawing on interviews with 16 computer pioneers, including Marvin Minsky, Herbert Simon, and Allen Newell. Crisp, often insightful prose informs the reader about the 1956 Dartmouth workshop on thinking machines, early programs such as Logic Theorist, pathbreaking research programs at places like the Massachusetts Institute of Technology and Carnegie Mellon University, battles between paradigms and prima donnas, and the gradual emergence of the current situation, which is marked by modest commercial successes and grave scientific questions. The three concluding chapters are a meditation on the future, prophesying that eventually AI could fulfill all the grand hopes of its creators.

The author's model of AI history is a roller coaster of booms and busts: messiahs of silicon intelligence promise miracles; money and disciples flock to them; the millennium fails to materialize as prophesied; despite its solid modest accomplishments the particular technique is overshadowed by some newer fad. At MIT in the 1960s, researchers tackled carefully delimited' problems called micro worlds, notably computer manipulation of imaginary toy blocks as in Terry Winograd's SHRDLU, but they never seemed to graduate into the real world. John McCarthy's LISP programming language captivated enthusiasts who believed that symbol manipulation was the answer to their dreams, and several companies built costly machines to run it efficiently, before the inevitable rude awakening. Beginning with MYCIN, a mechanical adviser to help doctors diagnose and treat blood diseases, rule-based expert systems promised to automate professional judgment, but the more complex systems proved very hard to update and seldom matched the wisdom of human experts. Today, Crevier suggests, enthusiasts are being oversold on neural networks, a set of techniques inspired by if not actually duplicating the parallel processing employed by the human nervous system.

According to Crevier there were seven factors that encouraged pioneers to make wildly exaggerated claims about the short-term potential of AI. First, the rapid early progress of computer technology generated excessive optimism that the machines could soon accomplish almost anything. Second, AI pioneers were influenced by brands of psychology like behaviorism that asserted that human intelligence was basically very simple. Third, computers are much better than humans are at some intellectual tasks, such as rapid arithmetic, so it seemed reasonable they could easily be made to equal us at the other tasks. Fourth, because AI initially lacked the professional standing of established disciplines, those attracted to it tended to be self-promoting hustlers who were not afraid to take risks. Fifth, because it was a new subculture coalescing on the margins of existing fields, AI lacked social norms that might have restrained individual boastfulness. Sixth, unlike many new developments in science, AI could readily be explained to journalists, whose exaggerations would feed back to amplify the promoters' own hype. Finally, according to Crevier, researchers in any fast-moving field initially overestimate its short-term potential, although after disappointments they often underestimate its long-term potential.

Crevier mentions mutual influences between AI and psychology, but far more could be said about the opportunities that were lost because AI ignored much of what was going on in the conventional behavioral sciences. For example, around 1951 Minsky worked on neural networks under the influence of the psychologist B. F. Skinner, but when Skinner showed little interest Minsky quickly reversed course and became obsessed with the opposite of behaviorism, symbolic processing. The lost opportunity, I suggest, was that neural networks could have contributed much to the Skinnerian paradigm. Skinner's friend George Homans published two influential books, in 1950 and 1961, establishing a behaviorist sociology. His third great project was to have been entitled A Toy Society, showing that the Skinner-Homans axioms could explain human society. Unable to find a means for creating functioning toy societies, however, Homans was forced to abandon this project. Only near the end of his life around 1987 did he realize that neural network simulations would have been the perfect tool.

Among the pressing questions for the historian of AI is why work on neural nets languished in the 1970s and early 1980s. One theory is that the book Perceptrons by Minsky and Seymour Papert falsely convinced researchers that strict mathematical laws severely limit the potential of neural nets. An alternative view is that computers did not have sufficient power to demonstrate the value of neural nets until later. Or perhaps neural nets are just another fad that had to wait its place in line for attention. Crevier states these opposing views, but he does not attempt to assess the evidence, even though he discusses Minsky's charisma at length, and logically his boom-and-bust model demands examination of the mechanisms that either inflate or deflate the reputation of particular AI techniques.

Crevier does not consider automatic language translation to be AI, although its history fits his boom-and-bust model, and he admits there is no firm definition of his topic. In a sense, all electronic computing qualifies as artificial intelligence, especially if it incorporates pattern recognition, learning, and decision-making. At the other extreme, perhaps AI is just a term for sleight-of-hand that fakes human reactions with a machine; it is said that once you understand fully how a new computer technique works, it is no longer AI. Despite devoting many pages to neural networks, Crevier questions whether they are AI, because the traditional LISP-Minsky paradigm defined machine intelligence in terms of the manipulation of symbols.

By raw extrapolation, Crevier predicts that supercomputers will match the processing power and memory of the human brain sometime between the years 2009 and 2042. If this comes about, artificial intelligence will unmistakably announce itself a few minutes later. Then it will be possible to write a definitive history, with a clear vision of the meaning of the topic, although the author may not be a human being. Though it lacks the caution and the detail marking the work of a professional historian, Crevier's book is a stimulating introduction to perplexing episodes in the history of computers.

William Sims Bainbridge
Sociology Program,
National Science Foundation,
Washington, DC 20550