from *SIAM News*, Volume 33, Number 4

Progress in Robotics: An Accelerated Rerun Of Natural History?

Robot: Mere Machine to Transcendent Mind. *By Hans Moravec, Oxford University Press, Oxford, New York, 1999, ix + 227 pages, $25.*

BOOK REVIEW

By James Case

Hans Moravec claims that robots are destined to inherit not only the earth, but also the sun, the moon, and the stars. As a first step, robots will surpass their human progenitors in poetry, literature, and music, as well as in science, history, math, curiosity, and business acumen. Already, chess and checkers are securely in their win column. Little though there is in the annals of artificial intelligence to support such a thesis, Moravec's defense of it is worth pondering.

He begins with the observation that the American military establishment, due to (possibly bogus) intelligence suggesting that the Russians were hot on the same trail, was remarkably generous in its initial funding of AI. To prevent the opening of a "cybernetic gap," pioneers in the field received abundant access to the biggest and fastest computers extant during the 1950s—machines capable (between breakdowns) of executing upward of half a million instructions per second (MIPS). Later on, as AI's priority declined, those working in the field suffered declining access to "state-of-the-art" machines. As a result, the computing power at their disposal remained virtually static until about 1990, when ordinary PCs and workstations began to deliver first tens, then hundreds, and more recently thousands of MIPS. Only then did the computing power available to AI researchers—and with it progress in the field—begin to take off. One reason for the 30-year stagnation was the apparently mistaken belief among AI researchers themselves that a single MIPS should suffice to mimic human performance in even those tasks—such as locomotion and hand/eye coordination—for which evolution has best equipped mankind.

Moravec sees progress in robotics as an accelerated rerun of natural history. Thus, he compares the intelligence of current industrial robots, controlled as they are by already-obsolete computers capable of no more than 10 MIPS, with that of insects. Though good for spray-painting and spot-welding automobiles, assembling electronic circuit boards, carting subassemblies from place to place in factories equipped with buried signal-emitting wires, and bombing "hardened" military installations, they are stymied by more complex activities. Research machines are able to navigate well in contrived settings, both indoors and out, while commercial programs read text and transcribe speech. Moravec expects these specialized devices to be followed—at intervals of perhaps ten years—by increasingly versatile "universal" robots that have, in succeeding generations, lizard-like spatial perception, mouse-like adaptability, monkey-like imagination, and, finally, humanoid reasoning ability. After that, with opportunities to "retro-engineer" existing organisms largely exhausted, the pace of progress in AI may slacken temporarily.

Chapter 2 reviews the state of the art, comparing a robot with a baby poised for sudden growth. Little is said about the clumsy industrial robots, or about the proposed microscopic robots designed to navigate the blood stream like a fleet of tiny submarines, identifying and performing needed maintenance. Instead, the discussion features a lengthy and authoritative account of robot vehicle design—a field in which Moravec and his students at Carnegie Mellon University are acknowledged leaders—from the earliest light-seeking "tortoises" of the 1950s to Carnegie Mellon's truck-sized NAVLAB vehicles. The most recent in the Carnegie Mellon series was driven in the summer of 1995 from Washington, DC, to San Diego—at an average speed of over 100 km/hour—by a program named RALPH that remained in control of the vehicle more than 98% of the time!

Several dozen small robot-building concerns appeared and disappeared during the 1980s. A few, such as Transitions Research Company, in Connecticut, and Cybermotion, in Virginia, have survived for more than a decade by collecting modest sums from investors and selling perhaps a dozen $25,000 robots a year, mostly to research groups. A few units have been sold to actual end-users and have delivered years of effective service, patrolling warehouses, delivering meals in hospitals, and so on. On the other hand, one such unit was fired after only a month on the job when its guidance system allowed it to blunder down a staircase located in a well-marked "no robot zone." With existing techniques, high degrees of reliability are tedious and expensive to achieve. Robots able to learn routes quickly, and traverse them in safety, have proven harder than expected to design. Yet Moravec insists that such products will soon be available.

He estimates that about 100 trillion instructions per second (TRIPS), along with commensurate memory, should suffice to emulate overall human performance. To decide what constitutes commensurate memory, he reflects that the original Macintosh, with 1/2 MIPS and 1/8 megabyte of memory, was considered a very fast machine. The subsequent "fat Mac," with 1/2 MIPS and 1/2 megabyte, ran larger programs at tolerable speeds. But the 1/2-MIPS, single-megabyte Mac+ was deemed lethargic by most users, while the 4-megabyte Mac classic, the last 1/2-MIPS machine in the line, was positively slow. In time, the classic was replaced by machines compressing ten times the processing power into the same space. Since then, guided by market surveys designed to assess customers' preferred trade-offs between memory and speed, marketing departments have held personal computer designers to roughly a megabyte-per-MIPS standard.

The megabyte-per-MIPS rule, Mor-avec then argues, seems to extend beyond the world of personal computers, into nature itself. The best available evidence, he points out, seems to place the bulk of nervous system memory in the synapses between connecting neurons. an accommodate about 100 million megabytes, and therefore respond to 100 TRIPS of processing power without undue delay. Finally, he extends the megabyte-per-MIPS rule to other (animate and inanimate) information processors by means of the diagram shown in Figure 1, in which the scale is logarithmic on both axes.

In a companion figure, he plots "MIPS per thousand dollars" against time, to suggest that if Moore's law holds true only a little longer than expected—and it's due to expire in 2012—the desired 100 TRIPS should be available in about 2020. Other such figures chronicle the increases in processing power available to AI and robot programs, as well as the ratings of various chess-playing programs against their computing power. Each depicts a fairly consistent upward trend, with no apparent leveling off.

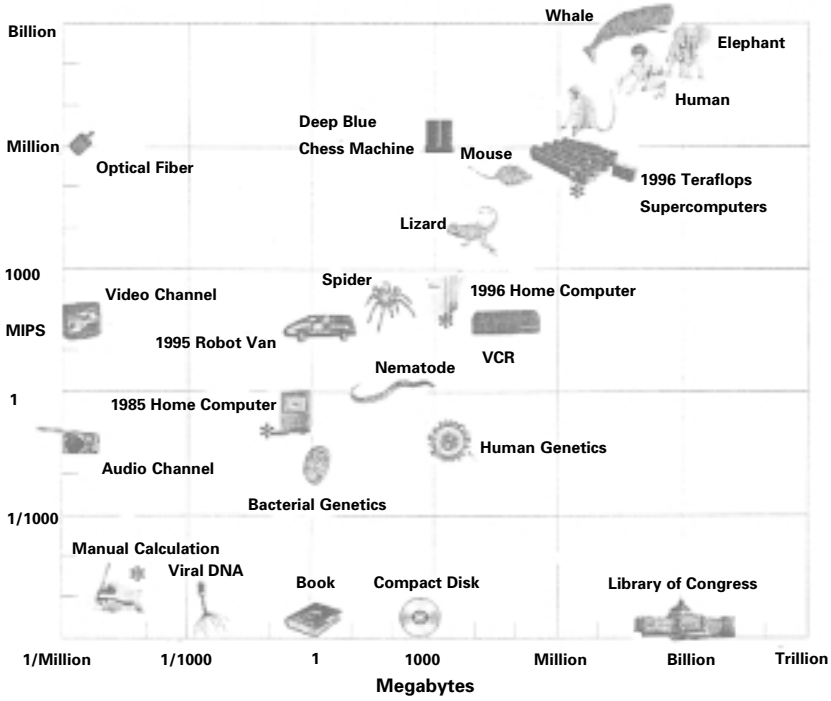

Figure 1. *Memory/speed trade-offs produced a "megabyte-per-MIPS" standard for personal computers that the author of the book under review sees extending well beyond the domain of the PC, into nature itself.*

He then lists and attempts to refute no fewer than nine separate arguments said to prove that machines, however accomplished they may become, can never really learn to *think*. Perhaps the most persuasive is Lady Lovelace's objection: Computers do only what we program them to do. At the other extreme is The Argument from Extrasensory Perception: Humans possess it, machines never can. This part of the book brings to mind the insomniac dyslexic philosopher who lay awake at night wondering if there is a Dog.

More persuasive—at least to this reviewer—were Garry Kasparov's remarks after his defeat by Big Blue, in which he described the opposing team (and Big Blue was a team) in terms usually reserved for a devious human foe. Moravec quotes similar testimony from audiences who find "classical" music composed by a computer more pleasing than the efforts of most human composers, along with the fact that a theorem-proving program operating at Argonne National Laboratory recently managed to settle a conjecture concerning Boolean algebras that had remained open for more than 50 years. What is the difference between thinking and merely seeming to think? At what point does it become pointless to deny that computers have finally learned to think? Such is the motivation behind the famous Turing test, on which Moravec's argument depends heavily.

Chapter 4 describes four generations of increasingly capable and versatile universal robots that recreate four 100-million-year stages in vertebrate brain evolution, culminating in robots with human capabilities. The stages correspond roughly to the historic enlargement of the brain stem, followed by the cerebellum, the mid-brain, and, finally, the cerebrum. The rate of development is fast-forwarded 10 million-fold. Will the end-result be conscious of its existence? Will it feel emotion?

Citing evidence that "the animating principle is not a substance, but a very particular, very complex organization," Moravec contends that robots will in time do everything humans do. He sees the requisite organization—long confined to living organisms—beginning to appear in the more sophisticated machines even now. In time, posterity will be obliged to "welcome some of them into the land of the living, however upsetting that may be to our traditional categories." He devotes ten pages to the ways in which emotions like fear, shame, joy, love, anger, pleasure, and pain will manifest themselves in machines. He concedes, however, that computers and brains will continue to have different relative strengths. Fourth-generation universal robots are likely to seem emotionally childlike, while excelling their human creators at rapid, complex reasoning. Following a pause, robots will then begin to evolve under their own power, each generation designing and manufacturing the next. How the resulting superrational superheroes will look and act is anyone's guess.

The final chapters of the book are devoted to various triumphs, such as nano-engineering, quantum computing, and time travel, that the laws of physics don't explicitly rule out and that a race of superrational superheroic robots could be expected to achieve. The text here is filled with appealing tidbits, such as the fact that Abhay Ashtekar discovered (in 1989) a change of variables that linearizes the PDEs of general relativity, thereby permitting a quantum mechanical interpretation and suggesting that space–time is composed of "loops" at the Planck scale of 10^{-33} centimeters. Connections have since been found between Ashtekar's work, the theory of knots, and "superstring" theories of the other fundamental forces.

Technology, according to Moravec, is approaching the limits of conventional matter. The best integrated circuits contain features 100 atoms wide—the limit being one atom—and switch 100 billion times a second—the limit being 100 trillion. Faster switching would sever chemical bonds. Superintelligent robots will find themselves hard put to improve their own design, at least with

conventional materials. But an improved material—perhaps the ultimate construction material—seems poised to burst onto the market. Buckminster Fullerene, the "buckyball," was discovered in 1985; a new form of molecular carbon consisting of 60 carbon atoms covalently bound, it was found to occur in nature. Related molecules, with still more carbon atoms in the shell, have been found (or synthesized) as well. One is shaped like a hollow straw, with half-buckyball caps at either end, and can in theory be of arbitrary length. Moravec notes that, in 1997, buckytubes were being grown catalytically from carbon vapor in millimeter lengths, with a high degree of efficiency. Several groups were racing to bring them to practicality.

Carbon's covalent bonds are the strongest in nature. If that strength could be carried over into tensile strength in a bulk material, making it a thousand times stronger than steel, flywheels with the energy-storage capacity of gasoline, nearly weightless rockets, impervious armor, and buildings a hundred miles tall would all become possible. It isn't entirely clear what all this has to do with robots, but it does make interesting reading. And Moravec has even more fantastic materials in mind—materials in whose atoms orbital Higgsinos replace orbital electrons in motion about nucleii composed of perhaps seventy-five protons. If such atoms could be induced to form crystals, the resulting "Higgsinium" would be a trillion times as dense as normal matter, would remain solid at temperatures of a million degrees, and would be able to support switching circuits a million times as fast!

Far more interesting, at least to this reviewer, were Moravec's thoughts on the initial stages of the coming age of robots. One need not believe that robots are destined to write sonnets and sonatas of surpassing beauty to agree with him that they will soon combine the superhuman skills and unimpeachable work habits required to operate the "lights out" unmanned factories of science fiction. Such installations will soon be churning out endless streams of high-quality, low-cost goods and services, and rendering human labor all but useless. The cheap labor that is the only comparative advantage most third-world nations seem to possess will be robbed of what little value it still retains, permitting the "workshops of the world" to be relocated yet again to the very epicenter of the markets they serve. Well within Moravec's 40-year time frame, the vast majority of human laborers seem likely to find themselves unemployable at any wage. Events have been moving in that direction for generations, and Moravec is among the few—the very few—to see the handwriting on the wall. That by itself makes his book worth reading.

Moravec likens the global workforce to a giant triangle, the broad base of which represents the low-skill, low-paying jobs occupied by most of the human race, while the narrower middle and tiny upper reaches correspond to white-collar and elite employment, respectively. He predicts that robots will enter from the bottom, forcing ever more highly trained flesh-and-blood workers upward and eventually outward through the uppermost vertex. But automation and downsizing are now having an altogether different effect. Computers and machine tools are driving a wedge between the highest- and lowest-paid lines of work, by performing tasks traditionally assigned to middle managers and skilled production workers. The jobs that remain unautomated are of two distinct types: those too sensitive or complex for existing robots to master, and those not worth building a machine to perform. Yet, as machines grow progressively smarter, human workers above and below the di-vide seem destined to be driven from the workforce, through both the top and the bottom of Moravec's employment triangle!

Moravec points out that, thanks to evolution, mankind is well equipped to enjoy the life of leisure about to be made possible by the age of robots. Stone age villagers typically worked (and in rare cases still work) only two to four hours a day—less when circumstances permit—and a few centuries of modern civilization have not significantly altered human wants and needs. He suggests that, as the robotic workforce of the future comes gradually into being, the developed nations would do well to emulate the socioeconomic organization already adopted by the oil-rich sheikdoms of the Persian Gulf. There, foreign guest-workers—paid for with petrodollars—play the role of robots. Citizens enjoy cradle-to-grave health care and free education, leading to undemanding government jobs and early retirement. Literacy and life expectancy are among the world's highest. While such societies produce both criminals and world-class achievers, the majority seem content merely to live out their unhurried lives—not unlike America's Amish—threatened only by neighboring countries whose impoverished majorities are less content with the status quo.

As Moravec observes at the outset, long-term predictions about technical developments, fantastic as they may sound when first uttered, have consistently erred on the side of conservatism. One suspects that he has avoided that pitfall with room to spare.

James Case is an independent consultant who lives in Baltimore, Maryland.