

Head and torso of a full-size Robonaut, a humanoid robot designed by the Robot Systems Technology Branch at NASA's Johnson Space Center, Houston, in a collaborative effort with DARPA. The Robonaut project seeks to develop and demonstrate a robotic system that functions as the equivalent of an extravehicular-activity astronaut.



ROBOTS, AFTER ALL

Artificial minds for mechanical bodies
capable of autonomous manual work finally
seem within reach.

COMPUTERS have invaded our everyday lives, and networked machines are finding their way into more and more of our gadgets, dwellings, clothes, even our bodies. But if pervasive computing soon handles most of our information needs, it will still not clean our floors, take out our garbage, assemble kit furniture, or do any of a thousand other essential physical tasks. The old dream of mechanical servants will remain largely unmet.

Electronics-wielding robot inventors in home, university, and industrial laboratories have tinkered with the problem since the early 20th century. While it was possible to build mechanical bodies capable of manual work, artificial minds for autonomous servants were out of reach. The problem's deceptive difficulty fooled generations of researchers attempting to solve it with computers.



BY HANS MORAVEC

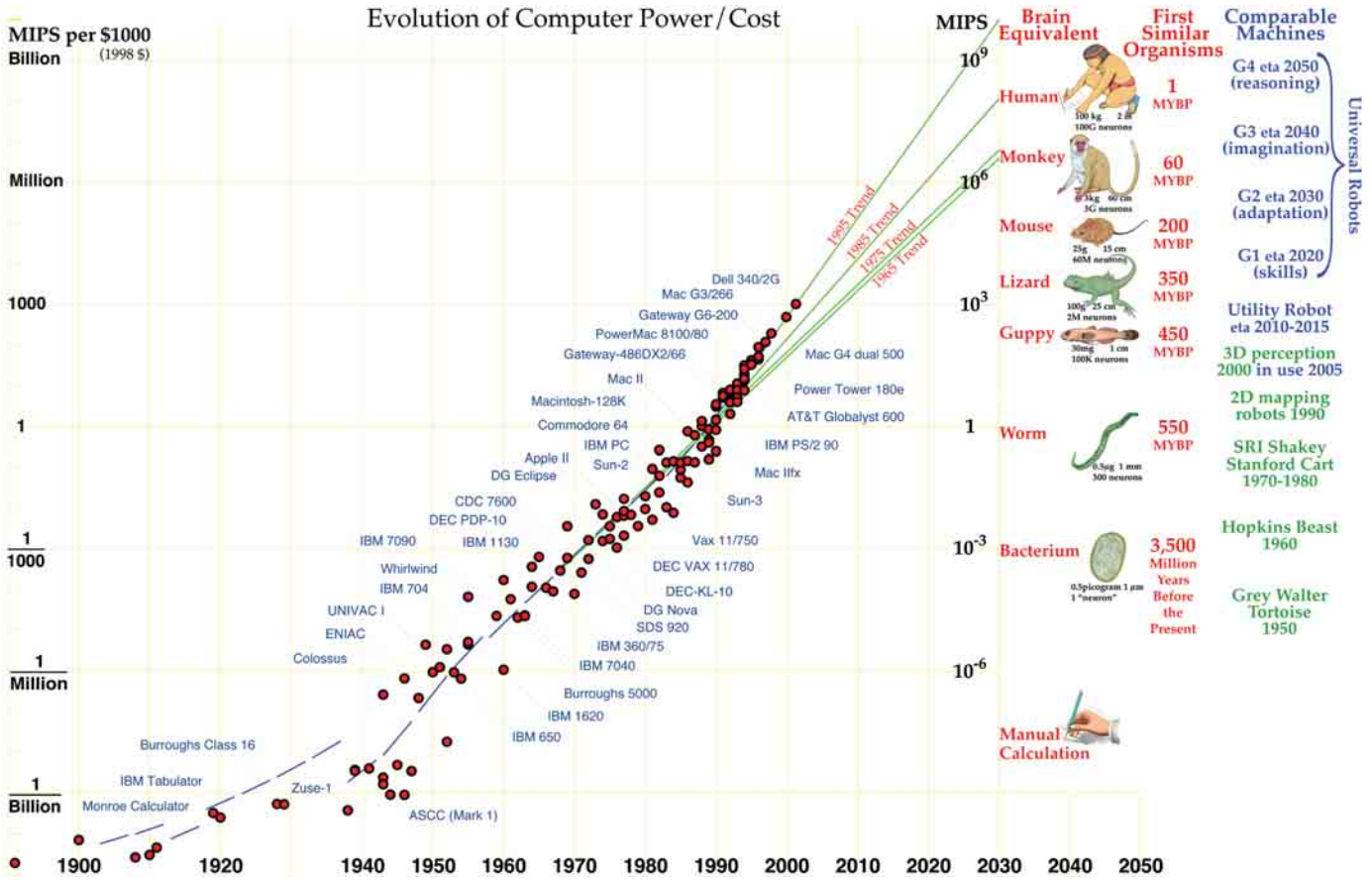


Figure 1. Mental power in four scales. The MIPS graph shows a Moore's Law-type progression for machine calculation scaled in MIPS per thousand dollars (1900 to the present). In the past decade this quantity has doubled each year as machines have become more powerful while their price simultaneously dropped. The Brain Equivalent column shows organisms whose behavior should be within reach of computers with MIPS ratings at the corresponding height, extrapolated from a retina/computer-vision comparison. The First Similar Organisms column is a timeline of the evolution of large nervous systems, showing when animals similar to those pictured first appeared in the fossil record. The Comparable Machines column attempts to locate historical and speculative future robots on this scale. Implied is that robot intelligence roughly recapitulates the evolution of human intelligence at about 10 million times the original rate. The best robots today are no smarter than the first tiny vertebrates, but their successors should cover the remaining distance to human-scale intelligence before 2050.

The earliest electronic computers in the 1950s did the work of thousands of clerks, seeming to transcend humans, as well as other machines. Yet the first reasoning and game-playing programs on those computers were a match only for individual human beginners, each performing a single narrow task. In the 1960s, computer-linked cameras and mechanical arms took hours to *unreliably* find and move a few white blocks on a black tabletop, much worse than a toddler. A modest robot industry did appear but consisted of arms and vehicles following predetermined trajectories. The situation did not improve substan-

tially for decades, disheartening waves of robotics devotees.

At long last, things are changing. Robot interactive behavior, hopelessly impossible in the 1970s and 1980s, became experimental demonstrations in the 1990s; mobile robots mapped and navigated unfamiliar office suites [1], robot vehicles drove themselves, largely unaided, across entire continents [2], and computer vision systems located textured objects and tracked and analyzed faces in real time. Programs that recognized text and speech became commercially successful. Market success extended to physical robots as the 2000s began; Sony has sold hundreds of thousands of its AIBO robot pets worldwide, despite its over-\$1,000 price tag, and several small robot vacuum cleaners, especially the \$200 iRobot Roomba, seem to be gaining customer acceptance. Not far behind, dozens of companies, established and new, are developing cleaning and transport robots using new sensors and leading-edge computers and algorithms licensed from research efforts. Emerging capabilities include mobile robots navigating ordinary places without special markers or advance maps; some systems map the surroundings in 2D or even 3D as they travel, enabling the next step—recognizing structural features and smaller objects. But why now?

Trick of Perspective

The short answer is that, after decades at about 1 MIPS (million instructions per second, each instruction representing work, like adding two 10-digit numbers), computer power available to research robots shot through 10, 100, and now 1,000 MIPS. This spurt is odd because the cost-effectiveness of computing had been rising steadily. In 1960, computers were a new and mysterious factor in the cold war, and even outlandish possibilities like artificial intelligence (AI) warranted significant investment. In the early 1960s, AI programs ran on the era's supercomputers, similar to the ones used for physical simulations by weapons physicists and meteorologists. By the 1970s, the promise of AI had faded, and the overall AI effort limped for a decade on old hardware. In contrast, weapons labs upgraded repeatedly to new supercomputers. In the 1980s, departmental computers gave way to smaller project computers, then to individual workstations and personal computers. Machine costs fell and their numbers rose, but power stayed at 1 MIPS. By 1990, the research environment was saturated with computers, and only then did further gains manifest in increased power rather than just numbers.

Mobile robot research might have blossomed sooner had the work been done on supercomputers—but pointlessly. At best, a mobile robot's computer could substitute only for a human driver, a function worth perhaps \$10 an hour. Supercomputer time costs at least \$500 an hour. Besides, the dominant opinion in the AI labs, dating from when computers did the work of thousands, was that, with the right program, 1 MIPS could encompass any human skill. The opinion remained defensible in the 1970s, as reasoning and game-playing programs performed at modest human levels.

For the few researchers in the newborn fields of computer vision and robotics, however, 1 MIPS was far from sufficient. With the best programs, individual images crammed memory, scanning them consumed seconds, and serious image analysis took hours. Human vision performed far more elaborate functions many times a second.

Hindsight enlightens. Computers calculate using as few gates and switching operations as possible. Human calculation, by contrast, is a laboriously learned, ponderous, awkward, unnatural behavior. Tens of billions of neurons in our vision and motor systems strain to analogize and process a digit a second. If our brains were rewired into 10 billion arithmetic circuits, each doing 100 calculations a second, by a mad computer designer with a future surgical tool, we'd outcompute 1 MIPS computers a million-

fold, and the illusion of computer power would be exposed. Robotics, in fact, gave us an even more apparent exposé of the limits of computing power.

Though spectacular underachievers at the wacky new stunt of longhand calculation, we are veteran overachievers at perception and navigation. Our animal ancestors across hundreds of millions of years prevailed by being frontrunners in the competition to find food, escape danger, and protect offspring. Existing robot-controlling computers are far too feeble to match this massive ultra-optimized perceptual inheritance. But by how much?

The vertebrate retina is understood well enough to be a kind of Rosetta stone roughly relating nervous tissue to computation. Besides light detectors, the retina contains edge- and motion-detecting circuitry packed into a little one-tenth-millimeter-thick, two-centimeter-across patch that reports on a million image regions in parallel about 10 times a second via the optic nerve. In robot vision, similar detections, well coded, each require the execution of a few hundred computer instructions, making the retina's 10 million detections per second worth more than 1,000 MIPS. In a risky extrapolation that must serve until something better emerges, this implies it would take about 50,000 MIPS to functionally imitate a gram of neural tissue, and almost 100-million MIPS (or 100 trillion instructions per second) to emulate the 1,500-gram human brain. By that measure, PCs today are just a match for insect nervous systems, or the 0.01 gram brain of a guppy.

Coordinated insectlike behavior in robots is probably best exhibited in the exciting field of RoboCup robot soccer (see www.robocup.org). Beginning in 1993, an international community of researchers has been organizing an effort to develop fully autonomous robots that could eventually compete in human soccer games, just as chess-playing computers compete in human chess tournaments. Their incremental development would be tested in annual machine/machine tournaments. The first "RoboCup" games were held at a 1997 AI conference in Nagoya, Japan; 40 teams entered in three competition categories: simulation, small robots, and middle-size robots. (The next size step, human scale, was reserved for the future). The robot teams (about five coffee-can-size players) were each controlled by an outside computer viewing the billiard-table-size playing field through an overhead color camera. To simplify the problem, the field was uniformly green, the ball was bright orange, and each of the players top surfaces had a unique pattern of large dots relatively easy for programs to track. The middle-size robots, approximately the dimensions of a breadbox, had cameras and com-



Figure 2. Overview and interior views of a 3D map generated by Seegrid Corp. programs using only stereoscopic camera views from a robotic vehicle's trip down the center of an L-shaped hallway.

puters onboard and played on a similarly colored but larger field.

Action was fully preprogrammed; no human intervention was allowed during play. In the first tournament, merely finding and pushing the ball was a major accomplishment (never mind the goal location), but the conference encouraged participants to share developments, and play improved in subsequent years.

Almost 400 teams signed up for RoboCup 2003, last July in Padua, Italy, and regional tournaments were introduced to cull the final tournament competitors by more than half.

In 1998, Sony provided AIBO robot dogs for a new competition category. Remarkably cute in play, AIBOs provide a standard, reliable, prebuilt hardware platform that needs only soccer software. In recent tournaments, the best teams frequently exhibit effective coordinated (goal-directed) behavior, intelligent blocking, even passing.

Though PCs today are still a daunting 100,000 times too weak, the goal of human performance is probably not impossibly far away. Computer power for a given price roughly doubled each year in the

1990s, after doubling every 18 months in the 1980s, and every two years before that. Perhaps 20 or 30 more years at the present pace would close the gap. Or, estimating the design effort ahead, the first multicelled animals with nervous systems appeared about 550 million years ago, ones with brains as capable as guppies perhaps 200 million years later. Self-contained robots covered similar ground in about 20 years. If we accept that evolutionary time roughly estimates engineering difficulty, the remaining 350 million years of our ancestry (at that pace) could be paralleled in robots in about 35 years (see Figure 1).

Better yet, sufficiently useful robots don't need full human-scale brainpower. Commercial and research experience convinces me that mental power like the pinhead-size brain of a guppy, or about 1,000 MIPS, will suffice to guide mobile utility robots reliably through unfamiliar surroundings, suiting them for jobs in hundreds of thousands of industrial environments and eventually hundreds of millions of homes. Such machines are less than a decade away, but they have been elusive so long that only a few dozen small research groups continue to pursue them.

One-Track Minds

Industrial mobile robots first appeared in 1954 when a driverless electric cart made by Barrett Elec-

tronics Corp. began pulling loads around a South Carolina grocery warehouse. Such machines, dubbed automatic guided vehicles, or AGVs, since the 1980s, originally, and still commonly, navigate by following signal-emitting wires entrenched in concrete floors. AGVs range from very small, carrying a few tens of kilograms, to very large, transporting many tons. Built for specific tasks, they are often equipped with specialized loading and unloading mechanisms like forks and lifters. In the 1980s, AGVs acquired microprocessor controllers allowing more complex behavior than afforded by simple electronic controls. New navigation techniques emerged; one uses wheel rotations to approximately track vehicle position, correcting for drift by sensing the passage of checkerboard floor tiles or magnets embedded along the path. In the 1990s, a method called laser navigation (invented and patented by several companies independently) triangulated a vehicle's position by sighting three or more retro-reflectors mounted on walls and pillars with a scanning laser on the vehicle.

About 100,000 self-guided vehicles have now found work in industry worldwide, but lighter "service robots" have yet to match even this modest success. Service robots are intended for human-service tasks like delivering mail in offices and linens and food in hospitals, floor cleaning, lawn mowing, and guard duty [4]. The most successful service robot to date is the Bell & Howell Mailmobile, designed to follow a transparent ultraviolet-fluorescent track spray-painted on office floors. About 3,000 have been sold since the late 1970s. A few dozen small AGVs from several manufacturers have been adapted to transport linens or food trays along hospital corridors. In the 1980s, several small U.S. companies were formed to exploit suddenly available microprocessors to develop small transport, floor-cleaning, and security robots that navigated by sonar, beacons, reflectors, and clever programming. But these units were expensive, often costing over \$50,000, and required expert installation. No company managed to sell more than a few dozen per decade, and all slowly expired.

Today, larger AGVs and service robots must follow carefully prearranged routes, greatly limiting their uses, usefulness, and commercial prospects. Emerging techniques, utilizing increased computer power, are poised to loosen this restriction by letting the robots do the routing, surely expanding the market. Customers will be able, unassisted, to put a robot to work where needed, enabling casual transport, floor cleaning, and other mundane tasks that cannot bear the cost, time, and uncertainty of expert human installation. Though much freer in their wanderings, these

new robots have to retain the reliability of their tracked predecessors. In my experience, customers routinely reject transport and security robots that, after a month of flawless operation, wedge themselves into corners, wander away lost, roll over employees' feet, or endanger themselves on stairs. Six months of successful operation earn a robot the equivalent of a sick day.

Sense of Space and Place

Experimental robots that chart their own routes emerged from laboratories worldwide in the mid-1990s, as microprocessors reached 100 MIPS [3]. Most built 2D maps from sonar arrays to locate and route themselves; the best were able to navigate office hallways for days between confusions. Those using sonar sensors, however, fell far short of the six-month commercial criterion. Too often, different locations in coarse 2D maps resemble one another; the same location, scanned at different heights, looks different; and small obstacles or awkward protrusions are overlooked. Greatly improving 2D mapping performance is a scanning laser sensor from German company Sick AG that scans 180 degrees in quarter-degree increments and gives reliable ranges out to several tens of meters with centimeter accuracy. Many experimental mobile robots now sport one or more Sick scanners (blue, yellow, or white, with conical scanning windows resembling compact coffee makers). Some appear to travel reliably, and commercial applications are emerging.

Meanwhile, Siemens offers a navigation package incorporating a Sick scanner for mapping and multiple sonar units for obstacle detection. The same scanners have also been incorporated into a floor-cleaning machine from Hefter Cleantech GmbH, also of Germany, that cleans the interior of an area after a human guides it around the perimeter. And an AGV from Swisslog of Switzerland, also using a Sick scanner, follows a "virtual guidepath" defined by scanner-sensed wall outlines.

Sick 2D scanning laser rangefinders provide the most effective current solution to the problem of freely navigating robots, but they're unlikely to be the final word. The maps are 2D, oblivious to hazards and opportunities above or below the scanning plane. The lasers are complex, precision electro-opto-mechanical devices that emit a powerful infrared beam, and their price is likely to fall only slowly from its current \$5,000 per unit.

I've worked more than 30 years toward practical 3D perception for robots from a variety of sensors, including cheap ones, to enable not only very reliable navigation but also such abilities as object recognition.

In the 1980s, my lab devised a way to distill large amounts of noisy sensor data into reliable maps by accumulating statistical evidence of emptiness or occupancy in each cell of a grid representing the surroundings. This approach worked well in 2D and guided many sonar-equipped robots in the following decade. 3D maps, a thousand times richer, promised to be even better but seemed computationally out of reach.

In 1992, I found economies of scale and other tricks that would reduce 3D grid costs a hundredfold and by 1996 demonstrated a test program that accumulated thousands of measurements from stereoscopic camera glimpses to map a room's volume down to centimeter-scale. With 1,000 MIPS now common in PCs, the program digests more than a glimpse per second of several stereo pairs of images (about 10,000

uct (available in about a year) will be a light-duty self-navigating cart called (for now) the SG-100 that customers install by pushing once through a facility while stopping at important locations and adding their labels to a menu. Once trained, the cart can be loaded and directed to any menu destination. It will drive to the location, stop, and wait to be unloaded, ready (with infinite patience) for the next trip. As it travels, it observes and records its surroundings in rich 3D, plans safe routes, and localizes its position relative to a map from its training tour; that map is incrementally extended and updated in subsequent trips.

Seegrid anticipates collaborations that apply these techniques to commercial floor-cleaning machines, allowing them to map and select their own cleaning routes for indicated rooms and corridors. A human custodian might supervise a flock of such semiau-

Perhaps by 2020 the process will have produced the first broadly competent 'universal robots,' the size of people but with lizardlike 10,000 MIPS minds that can be programmed for almost any simple chore.

range values), adequate for up to two feet per second (a slow walk) indoor travel.

The program was further developed from 1999 to 2003 with DARPA support, greatly improving the quality of its performance. A key addition was a learning process that tunes the sensor models through stereoscopic (and other) readings. Multiple camera images of the actual scene are projected from corresponding positions into trial 3D grids produced using particular sensor model settings; in good maps the occupied cells correspond to things in the physical scene and receive similar colors from the multiple views, so average color variance per cell is low. The learning process tunes the sensor model in the direction of decreasing average color variance.

The latest results are quite good; Figure 2 shows a 3D map constructed entirely from panoramic stereoscopic images obtained in a single traverse down the center of an L-shaped hallway. These results, as well as prior experience navigating from more sparse 3D and 2D data, convinced me that these techniques were finally ready for commercial development.

In February 2003, Scott Friedman and I founded Seegrid Corp. of Pittsburgh, PA, to do the job of developing self-navigating commercial robots (see www.ri.cmu.edu/~hpm/seegrid.html). Our first prod-

ucts are vacuum cleaners. We are also seeking to apply the approach to security robots that patrol warehouses and other large facilities detecting intrusions. We'll be able to expand these applications with routines that scan the 3D maps to recognize large features, including walls, doors, corridors, and rooms, as well as smaller objects, including furniture and people. The hardware costs for processors, cameras, and other sensors is several thousand dollars per robot in the short run, but the component costs are falling at a rate that will bring the system into the consumer price range within 10 years. Imagine small, patient, competent robot vacuum cleaners automatically learning their way around a home, exploring unoccupied rooms, and cleaning when the humans are away, recharging their batteries and emptying their dust containers at a docking station.

Any Simple Chore

Commercial success will provoke competition and accelerate investment in manufacturing, engineering, and research. Vacuuming robots should beget smarter cleaning robots with dusting, scrubbing, and picking-up arms, followed by larger multifunction utility robots with stronger, more dexterous arms and better sensors. Programs will be written to make such machines pick up clutter, store, retrieve, and deliver

things, take inventory, guard homes, open doors, mow lawns, and play games. New applications will expand the market and spur further advancement whenever the robots fall short in acuity, precision, strength, reach, dexterity, skill, or processing power. Capability, numbers sold, cost effectiveness, and engineering and manufacturing quality will improve in a mutually reinforcing spiral. Perhaps by 2020 the process will have produced the first broadly competent “universal robots,” the size of people but with lizardlike 10,000 MIPS minds that can be programmed for almost any simple chore [2].

Like competent but instinct-ruled reptiles, first-generation universal robots will handle only those contingencies explicitly covered in their currently running application programs. Unable to adapt to changing circumstances, they are likely to perform inefficiently or not at all. Still, so much physical work awaits them in businesses, streets, fields, and homes that robotics could begin to overtake pure information technology as a commercial enterprise.

A second generation of universal robot with a mouselike 300,000 MIPS, which could be available within about 30 years, will be able to adapt, unlike the first generation, and even be trainable. Besides application programs, these robots would host a suite of software “conditioning modules” generating positive and negative reinforcement signals in predefined circumstances. Application programs would have alternatives for every step small and large (grip under/over hand, work in-/outdoors). As jobs are repeated, alternatives that had resulted in positive reinforcement would be favored, those with negative outcomes shunned. With a well-designed conditioning suite, including, say, positive for doing a job quickly, keeping the batteries charged, and negative for breaking or hitting something, a second-generation robot will slowly learn to improve its performance.

A Kind of Consciousness

A monkeylike 10 million MIPS will permit a third generation of robots, which could be available within about 40 years, as indicated in Figure 1, to learn quickly from mental rehearsals in simulations modeling physical, cultural, and psychological factors. Physical properties include shape, weight, strength, texture, and appearance of things, along with what people do with them. Cultural aspects include a thing’s name, value, location, and purpose. Psychological factors, applied to humans and other robots, include goals, beliefs, feelings, and preferences.

Developing these simulators will be a huge undertaking spread across many organizations and involve thousands of programmers and experience-gathering

robots. A simulation would track external events, constantly tuning its models to keep them faithful to reality. It should let a robot learn a skill through imitation and afford a kind of consciousness. Asked why candles are on a table, a third-generation robot might consult its simulation of house, owner, and self to honestly reply that it put them there because its owner likes candlelit dinners, and it likes to please its owner. Further queries would elicit more detail about a simple inner mental life concerned only with concrete situations and the people in its work area.

Fourth-generation universal robots with a humanlike 300-million MIPS, possibly available within 50 years, will be able to abstract and generalize. The earliest AI programs 40 years ago reasoned abstractly almost as well as people, albeit in very narrow domains, and many existing expert systems outperform us. But the symbols these programs manipulate are meaningless unless interpreted by humans. For instance, a medical diagnosis program needs a human practitioner to observe and enter a patient’s symptoms and implement a recommended therapy. Not so a third-generation robot, whose simulator provides a two-way conduit between symbolic descriptions and physical reality. Fourth-generation machines can be expected to result from melding powerful reasoning programs to third-generation machines. They may be able to reason about everyday actions by referring to their simulators in the same way that IBM researcher Herbert Gelernter’s 1959 geometry theorem prover examined analytic-geometry “diagrams” to check special-case examples before trying to prove general geometric statements.

Properly educated, the resulting robots are likely to be intellectually formidable. **C**

REFERENCES

1. Coradeschi, S., Tadokoro, S., Noske, R., and Birk, A., Eds. *RoboCup 2001: Robot Soccer World Cup V (Lecture Notes in Artificial Intelligence)*. Springer Verlag, Berlin, 2002.
2. Moravec, H. *Robot: Mere Machine to Transcendent Mind*. Oxford University Press, New York, 1999.
3. Murphy, R. *An Introduction to AI Robotics (Intelligent Robotics and Autonomous Agents)*. MIT Press, Cambridge, MA, 2000.
4. Schraft, R. and Schnierer, G. *Service Robots: Products, Scenarios, Visions*. AK Peters, Ltd., Natick, MA, 2000.

HANS MORAVEC (hpm@cs.cmu.edu) is a research professor in the Robotics Institute at Carnegie Mellon University, Pittsburgh, PA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
