THE UNIVERSAL ROBOT

Hans Moravec
Carnegie Mellon University
Pittsburgh, Pennsylvania

Abstract

Our artifacts are getting smarter, and a loose parallel with the evolution of animal intelligence suggests one future course for them. Computerless industrial machinery exhibits the behavioral flexibility of single-celled organisms. Today's best computer-controlled robots are like the simpler invertebrates. A thousand-fold increase in computer power in the next decade should make possible machines with reptile-like sensory and motor competence. Properly configured, such robots could do in the physical world what personal computers now do in the world of data—act on our behalf as literal-minded slaves. Growing computer power over the next half-century will allow this reptile stage will be surpassed, in stages producing robots that learn like mammals, model their world like primates and eventually reason like humans. Depending on your point of view, humanity will then have produced a worthy successor, or transcended some of its inherited limitations and so transformed itself into something quite new

Introduction: State of the Art

Instincts which predispose the nature and quantity of work we enjoy probably evolved during the 100,000 years our ancestors lived as hunter-gatherers. Less than 10,000 years ago the agricultural revolution made life more stable, and richer in goods and information. But, paradoxically, it requires more human labor to support an agricultural society than a primitive one, and the work is of a different, "unnatural" kind, out of step with the old instincts. The effort to avoid this work has resulted in domestication of animals, slavery and the industrial revolution. But many jobs must still be done by hand, engendering for hundreds of years the fantasy of an intelligent but soulless being that can tirelessly dispatch the drudgery. Only in this century have electronic sensors and computers given machines the ability to sense their world and to think about it, and so offered a way to fulfill the wish.

As in fables, the unexpected side effects of robot slaves are likely to dominate the resulting story. Most significantly, these perfect slaves will continue to develop, and will not long remain soulless. As they increase in competence they will have occasion to make more and more autonomous decisions, and so will slowly develop a volition and purposes of their own. At the same time they will become indispensable. Our minds were evolved to store the skills and memories of a stone-age life, not the enormous complexity that has developed in the last ten thousand years. We've kept up, after a fashion, through a series of social inventions—social stratification and division of labor, memory aids like poetry and schooling, written records stored outside the body, and recently machines that can do some of our thinking entirely without us. The portion of absolutely essential human activity that takes place outside of human bodies and minds has been steadily increasing. Hard working intelligent machines may complete the trend.

Serious attempts to build thinking machines began after the second world war. One line of research, called Cybernetics, used simple electronic circuitry to mimic small nervous systems, and produced machines that could learn to recognize simple patterns, and turtle-like robots that found their way to lighted recharging hutches [Wiener 1961]. An entirely different approach, named Artificial Intelligence (AI), attempted to duplicate rational human thought in the large computers that appeared after the war. By 1965, these computers ran programs that proved theorems in logic and geometry, solved calculus problems and played good games of checkers [Feigenbaum 1963].

In the early 1970s, AI research groups at MIT (the Massachusetts Institute of Technology) and Stanford University attached television cameras and robot arms to their computers, so their "thinking" programs could begin to collect their information directly from the real world.

What a shock! While the pure reasoning programs did their jobs about as well and about as fast as college freshmen, the best robot control programs took hours to find and pick up a few blocks on a table. Often these robots failed completely, giving a performance much worse than a six month old child. This disparity between programs that reason and programs that perceive and act in the real world holds to this day. In recent years Carnegie Mellon University produced two desk-sized computers that can play chess at grandmaster level, within the top 100 players in the world, when given their moves on a keyboard. But present-day robotics could produce only a complex and unreliable machine for finding and moving normal chess pieces.

In hindsight it seems that, in an absolute sense, reasoning is much easier than perceiving and acting—a position not hard to rationalize in evolutionary terms. The survival of human beings (and their ancestors) has depended for hundreds of millions of years on seeing and moving in the physical world, and in that competition large parts of their brains have become efficiently organized for the task. But we didn't appreciate this monumental skill because it is shared by every human being and most animals—it is commonplace. On the other hand, rational thinking, as in chess, is a newly acquired skill, perhaps less than one hundred thousand years old. The parts of our brain devoted to it are not well organized, and, in an absolute sense, we're not very good at it. But until recently we had no competition to show us up.

By comparing the edge and motion detecting circuitry in the four layers of nerve cells in the retina, the best understood major circuit in the human nervous system, with similar processes developed for "computer vision" systems that allow robots in research and industry to see, I've estimated that it would take a billion computations per second (the power of a world-leading Cray 2 supercomputer) to produce the same results at the same speed as a human retina. By extrapolation, to emulate a whole brain takes ten trillion arithmetic operations per second, or ten thousand Crays worth [Moravec 1988]. This is for operations our nervous system do extremely efficiently and well.

Arithmetic provides an example at the other extreme. In 1989 a new computer was tested for a few months with a program that computed the number pi to more than one billion decimal places. By contrast, the largest unaided manual computation of pi was 707 digits by William Shanks in 1873. It took him several years, and because of a mistake every digit past the 527th was wrong! In arithmetic, today's average computers are one million times more powerful than human beings. In very narrow areas of rational thought (like playing chess or proving theorems) they are about the same. And in perception and control of movement in the complex real world, and related areas of common-sense knowledge and intuitive and visual problem solving, today's average computers are a million times less capable.

The deficit is evident even in pure problem solving AI programs. To this day, AI programs exhibit no shred of common sense—a medical diagnosis program, for instance, may prescribe an antibiotic when presented a broken bicycle because it lacks a model of people, diseases or bicycles. Yet these programs, on existing computers, would be overwhelmed were they to be bloated with the details of everyday life, since each new fact can interact with the others in an astronomical "combinatorial explosion." [A ten year project called Cyc at the Microelectronics and Computer Consortium in Austin Texas is attempting to build just such a common-sense data base. They estimate the final result will contain over one hundred million logic sentences about everyday objects and actions [Lenat 1989].]

Machines have a lot of catching up to do. On the other hand, for most of the century, machine calculation has been improving a thousandfold every twenty years, and there are basic developments in research labs that can sustain this for at least several decades more. In less than fifty years computer hardware should be powerful enough to match, and exceed, even the well-developed parts of human intelligence. But what about the software that would be required to give these powerful machines the ability to perceive, intuit and think as well as humans? The

Cybernetic approach that attempts to directly imitate nervous systems is very slow, partly because examining a working brain in detail is a very tedious process. New instruments may change that in future. The AI approach has successfully imitated some aspects of rational thought, but that seems to be only about one millionth of the problem. I feel that the fastest progress on the hardest problems will come from a third approach, the newer field of robotics, the construction of systems that must see and move in the physical world. Robotics research is imitating the evolution of animal minds, adding capabilities to machines a few at a time, so that the resulting sequence of machine behaviors resembles the capabilities of animals with increasingly complex nervous systems. This effort to build intelligence from the bottom up is helped by biological peeks at the "back of the book"—at the neuronal, structural, and behavioral features of animals and humans.

The best robots today are controlled by computers which are just powerful enough to simulate the nervous system of an insect, cost as much as houses, and so find only a few profitable niches in society (among them, spray painting and spot welding cars and assembling electronics). But those few applications are encouraging research that is slowly providing a base for a huge future growth. Robot evolution in the direction of full intelligence will greatly accelerate, I believe, in about a decade when the mass-produced general purpose, universal robot becomes possible. These machines will do in the physical world what personal computers do in the world of data—act on our behalf as literal-minded slaves.

The Dumb Robot (ca. 2000-2010)

To be useful in many tasks, the first generation of universal robots should navigate efficiently over flat ground and reliably and safely over rough terrain and stairs, be able to manipulate most objects, and to find them in the nearby world. There are beginnings of solutions today. In the 1980s Hitachi of Japan developed a mobility system of five steerable wheels, each on its own telescoping stalk that allows it to accommodate to rises and dips in uneven terrain, and to climb stairs, by raising one wheel at a time while standing stably on the other four. My laboratory at Carnegie Mellon University in Pittsburgh has developed a navigation method that enables a robot equipped with sonar range measuring devices and television cameras to build probabilistic maps of its surroundings to determine its location and plan routes [Moravec 1987]. An elegant three-fingered mechanical hand at the Massachusetts Institute of Technology can hold and orient bolts and eggs and manipulate a string in a humanlike fashion [Mason 1985]. A system called 3DPO from SRI International in Menlo Park, California can find a desired part in a jumble seen by special range-finding camera [Bolles 1984]. The slow operation of these systems suggests one other element needed for the universal robot, namely a computer about one one thousand times as powerful as those found on desks and in robots today. Such machines, able to do one billion computations per second, would provide robots approximately the brain power of a reptile, and the personality of a washing machine.

Universal robots will find their first uses in factories, where they will be cheaper and more versatile than the older generation of robots they replace. Eventually they will become cheap enough for some households, extending the reach of personal computers from a few tasks in the data world to many in the physical world.

As with computers, many applications of the robots will surprise their inventors. Some will do light mechanical assembly, clean bathrooms, assemble and cook gourmet meals from fresh ingredients, do tuneups on a certain year and make of cars, hook patterned rugs, weed a lawn, run robot races, do detailed earthmoving and stonework, investigate bomb threats, deliver to and fetch from warehoused inventories, and much more. Each application will require its own original software (very complex by today's computer program standards), and some may also need optional hardware attachments for the robot such as special tools and chemical sensors.

Learning (2010-2020)

Useful though they will be, the first generation of universal robots will be rigid slaves to

37

simple programs. If the machine bangs its elbow while chopping beef in your kitchen making Stroganoff, you will have to find another place for the robot to do its work, or beg the software manufacturer for a fix. Second generation robots with more powerful computers will be able to host a more flexible kind of program able to adjust itself by a kind of conditioned learning. First generation programs will consist primarily of sequences of the type "Do step A, then B, then C...." The programs for the second generation will read "Do step A1 or A2 or A3 ... then B1 or B2 or B3 ... then C1 or C2 or C3 ...." In the Beef Stroganoff example, A1 might be to chop with the right hand of the robot, while A2 is to use the left hand. Each alternative in the program has a "weight," a number that indicates the desirability of using it rather than one of the other branches. The machine also contains a "pain" system, a series of programs that look out for problems, such as collisions, and respond by reducing the weights of recently invoked branches, and a "pleasure" system that increases the relevant weights when good conditions, such as well charged batteries or a task efficiently completed, are detected. As the robot bangs its elbow repeatedly in your kitchen, it gradually learns to use its other hand (as well as adapting to its surroundings in a thousand other ways). A program with many alternatives at each step, whose pain and pleasure systems are arranged to produces a pleasure signal on hearing the word "good" and a pain message on hearing "bad" could be slowly trained to do new tasks, like a small mammal. A particular suite of pain- and pleasure-producing programs interacting with a robot's individual environment would subtly shape its behavior and give it a distinct character.

Imagery (2020-2030)

Adaptive robots will find jobs everywhere, and the hardware and software industry that supports them could become the largest on earth. But teaching them new tasks, whether by writing programs or through punishment and reward, will be very tedious. This deficiency will lead to a portentous innovation, a software world-modeler (requiring another big increase in computer power), that allows the robot to simulate its immediate surroundings and its own actions within them, and thus to think about its tasks before acting. Before making Beef Stroganoff in your kitchen, the new robot would simulate the task many times. Each time its simulated elbow bangs the simulated cabinet, the software would update the learning weights just as if the collision had physically happened. After many such mental run-throughs the robot would be well trained, so that when it finally cooks for real, it does it correctly. The simulation can be used in many other ways. After a job, the robot can run though its previous actions, and try variations on them to improve future performance. A robot might even be configured to invent some of its own programs by means of a simpler program that can detect how nearly a sequence of robot actions achieves a desired task . This training program would, in repeated simulations, provide the "good" and "bad" indications needed to condition a general learning program like the one of the previous section.

It will take a large community of patient researchers to build good simulators. A robot entering a new room must include vast amounts of not directly perceived prior knowledge in its simulation, such as the expected shapes and probable contents of kitchen counters and the effect of (and force needed for) turning faucet knobs. It needs instinctive motor-perceptual knowledge about the world that took millions of years of evolution to install in us, that tells us instinctively when a height is dangerous, how hard to throw a stone, or if the animal facing us is a threat . Robots that incorporate it may be as smart as monkeys.

Reasoning (2030-2040)

In the decades while the "bottom-up" evolution of robots is transferring the perceptual and motor faculties of human beings into machinery, the conventional Artificial Intelligence industry will be perfecting the mechanization of reasoning. Since today's programs already match human beings in some areas, those of 40 years from now, running on computers a million times as fast as today's, should be quite superhuman. Today's reasoning programs work from small amounts of clear and correct information prepared by human beings. Data from robot sensors such as cameras

is much too voluminous and too noisy for them to use. But a good robot simulator will contain neatly organized data about the robot and its world. For instance, if a knife is on a countertop, or if the robot is holding a cup. A robot with simulator can be married to a reasoning program to produce a machine with most of the abilities of a human being. The combination will create beings that in some ways resemble us, but in others are like nothing the world has seen before.

First Generation Technicalities

Both industrial robot manipulators and the research effort to build "smart" robots are twenty five years old. Universal robots will require at least another decade of development, but some of their elements can be guessed from the experience so far. One consideration is weight. Mobile robots built to work in human sized spaces today weigh too many hundreds of pounds. This dangerously large mass has three major components: batteries, actuators and structure. Lead-acid batteries able to drive a mobile robot for a day contribute about one third of the weight. But nickel-cadmium aircraft batteries weigh half as much, and newer lithium batteries can be half again as light. Electric motors are efficient and precisely controllable, but standard motors are heavy and require equally heavy reducing gears. Ultrastrong permanent magnets can halve the weight and generate high torque without gears. Robot structure has been primarily aluminum. Its weight contribution can be cut by a factor of four by substituting composite materials containing superstrength fibers of graphite, aramid or the new material Spectra. These innovations could be combined to make a robot with roughly the size, weight, strength and endurance of a human.

The first generation robot will probably move on wheels. Legged robots have advantages on complicated terrain, but they consume too much power. A simple wheeled robot would be confined to areas of flat ground, but if each wheel had a controlled suspension with about a meter of travel, the robot could slowly lift its wheels as needed to negotiate rough ground and stairs. The manipulation system will consist of two or more arms ending in dexterous manipulators. There are several designs in the research labs today, but the most elegant is probably that of the so-called Stanford-JPL hand (mentioned above, now found at MIT), which has three fingers each with three controlled joints.

The robot's travels would be greatly aided if it could continuously pinpoint its location, perhaps by noting the delay from a handful of small synchronized transmitters distributed in its environment. This approach is used in some terrestrial and satellite navigation systems. The robot will also require a sense of its immediate surroundings, to find doors, detect obstacles and track objects in its workspace. Research laboratories, including my own, have experimented with techniques that do this with data from television cameras, scanning lasers, sonar transducers, infrared proximity sensors and contact sensors. A more precise sensory system will be needed to find particular work objects in clutter. The most successful methods to date start with three dimensional data from special cameras and laser arrangements that directly measure distance as well as lateral position. The robot will thus probably contain a wide angle sensor for general spatial awareness, and a precise, narrow angle, three dimensional imaging system to find particular objects it will grasp.

Research experience to date suggests that to navigate, visually locate objects, and plan and control arm motions, the first universal robots will require a billion operations per second of computer power . The 1980s have witnessed a number of well publicized fads that claim to be solutions to the artificial intelligence or robot control problem. Expert systems, the Prolog logical inference language, neural nets, fuzzy logic and massive parallelism have all had their spot in the limelight. The common element that I note in these pronouncements is the sudden enthusiasm of group of researchers experienced in some area of computer science for applying their methods to the robotics problems of perceiving and acting in the physical world. Invariably each approach produces some simple showcase demonstrations, then bogs down on real problems. This pattern is no surprise to those with a background in the twenty five year research robotics effort.

Making a machine to see, hear or act reliably in the raw physical world is much, much more difficult than naive intuition leads us to believe. The programs that work relatively successfully in

these areas, in industrial vision systems, robot arm controllers and speech understanders, for example, invariably use a variety of massive numerical computations involving statistics, vector algebra, analytic geometry and other kinds of mathematics. These run effectively on conventional computers, and can be accelerated by array processors (widely available add-ons to conventional machines which rapidly perform operations on long streams of numbers) and by use of modest amounts of parallelism. The mind of the first generation universal robot will almost certainly reside in quite conventional computers, perhaps ten processors each able to perform 100 million operations per second, helped out by a modest amount of specialized computing hardware that preprocesses the data from the laser eyes and other sensors, and that operates the lowest level of mobility and manipulation systems.

Mind Children (2050+)

The fourth robot generation and its successors, with human perceptual and motor abilities and superior reasoning powers, could replace human beings in every essential task. In principle, our society could continue to operate increasingly well without us, with machines running the companies and doing the research as well as performing the productive work. Since machines can be designed to work well in outer space, production could move to the greater resources of the solar system, leaving behind a nature preserve subsidized from space. Meek humans would inherit the earth, but rapidly evolving machines would expand into the rest of the universe.

This development can be viewed as a very natural one. Human beings have two forms of heredity, one the traditional biological kind, passed on strands of DNA, the other cultural, passed from mind to mind by example, language, books and recently machines. At present the two are inextricably linked, but the cultural part is evolving very rapidly, and gradually assuming functions once the province of our biology. In terms of information content, our cultural side is already by far the larger part of us. The fully intelligent robot marks the point where our cultural side can exist on its own, free of biological limits. Intelligent machines, which are evolving among us, learning our skills, sharing our goals, and being shaped by our values, can be viewed as our children, the children of our minds. With them our biological heritage is not lost. It will be safely stored in libraries at least; however its importance will be greatly diminished.

What about life back on the preserve? For some of us the thought of being grandly upstaged by our artificial progeny will be disappointing, and life may seem pointless if we are fated to spend it staring stupidly at our ultra-intelligent progeny as they try to describe their ever more spectacular discoveries in baby-talk that we can understand. Is there any way individual humans might join the adventure?

You've just been wheeled into the operating room. A robot brain surgeon is in attendance, a computer waits nearby. Your skull, but not your brain, is anesthetized. You are fully conscious. The robot surgeon opens your brain case and places a hand on the brain's surface. This unusual hand bristles with microscopic machinery, and a cable connects it to the computer at your side. Instruments in the hand scan the first few millimeters of brain surface. These measurements, and a comprehensive understanding of human neural architecture, allow the surgeon to write a program that models the behavior of the uppermost layer of the scanned brain tissue. This program is installed in a small portion of the waiting computer and activated. Electrodes in the hand supply the simulation with the appropriate inputs from your brain, and can inject signals from the simulation. You and the surgeon compare the signals it produces with the original ones. They flash by very fast, but any discrepancies are highlighted on a display screen. The surgeon fine-tunes the simulation until the correspondence is nearly perfect. As soon as you are satisfied, the simulation output is activated. The brain layer is now impotent—it receives inputs and reacts as before but its output is ignored. Microscopic manipulators on the hand's surface excise this superfluous tissue and pass them to an aspirator, where they are drawn away.

The surgeon's hand sinks a fraction of a millimeter deeper into your brain, instantly compensating its measurements and signals for the changed position. The process is repeated for the next layer, and soon a second simulation resides in the computer, communicating with the first

and with the remaining brain tissue. Layer after layer the brain is simulated, then excavated. Eventually your skull is empty, and the surgeon's hand rests deep in your brainstem. Though you have not lost consciousness, or even your train of thought, your mind has been removed from the brain and transferred to a machine. In a final, disorienting step the surgeon lifts its hand. Your suddenly abandoned body dies. For a moment you experience only quiet and dark. Then, once again, you can open your eyes. Your perspective has shifted. The computer simulation has been disconnected from the cable leading to the surgeon's hand and reconnected to a shiny new body of the style, color, and material of your choice. Your metamorphosis is complete.

Your new mind has a control labeled "speed." It had been set at 1, to keep the simulations synchronized with the old brain, but now you change it to 10,000, allowing you to communicate, react, and think ten thousand times faster. You now seem to have hours to respond to situations that previously seemed instantaneous. You have time, during the fall of a dropped object, to research the advantages and disadvantages of trying to catch it, perhaps to solve its differential equations of motion. When your old biological friends speak with you, their sentences take hours—you have plenty of time to think about the conversations, but they try your patience. Boredom is a mental alarm that keeps you from wasting your time in profitless activity, but if it acts too soon or too aggressively it limits your attention span, and thus your intelligence. With help from the machines, you change your mind-program to retard the onset of boredom. Having done that, you will find yourself comfortably working on long problems with sidetracks upon sidetracks. In fact, your thoughts routinely become so involved that you need an increase in your memory. These are but the first of many changes. Soon your friends complain that you have become more like the machines than the biological human you once were. That's life.

References

Norbert Wiener, Cybernetics, or Control and Communication in the Animal and the Machine (second edition), MIT Press, Cambridge, Massachusetts, 1961.

Edward Feigenbaum and Julian Feldman, eds., Computers and Thought, McGraw-Hill Inc., New York, 1963.

Robert Bolles, Patrice Horaud and Marsha Jo Hannah, 3DPO: A Three-Dimensional Part Orientation System, in Robotics Research : The First International Symposium, Michael Brady and Richard Paul, eds., MIT Press, Cambridge, Massachusetts, 1984, pp 413:424.

Matt Mason and Kenneth Salisbury, Robot Hands and the Mechanics of Manipulation, MIT Press, Cambridge, Massachusetts, 1985.[

Hans Moravec, Sensor Fusion in Certainty Grids for Mobile Robots, AI Magazine v9#2, Summer 1988, pp 61-77.

Hans Moravec, Mind Children: The Future of Robot and Human Intelligence, Harvard University Press, Cambridge, Massachusetts, 1988.

Douglas Lenat and Rajiv Guha, Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project, Addison-Wesley Publishing Co., Reading, Massachusetts, 1989.