

## Pigs in Cyberspace

Hans Moravec  
 Robotics Institute  
 Carnegie Mellon University  
 Pittsburgh, PA 15213

Exploration and colonization of the universe awaits, but Earth-adapted biological humans are ill-equipped to respond to the challenge. Machines have gone farther and seen more, limited though they presently are by insect-like behavioral inflexibility. As they become smarter over the coming decades, space will be theirs. Organizations of robots of ever increasing intelligence and sensory and motor ability will expand and transform what they occupy, working with matter, space and time. As they grow, a smaller and smaller fraction of their territory will be undeveloped frontier. Competitive success will depend more and more on using already available matter and space in ever more refined and useful forms. The process, analogous to the miniaturization that makes today's computers a trillion times more powerful than the mechanical calculators of the past, will gradually transform all activity from grossly physical homesteading of raw nature, to minimum-energy quantum transactions of computation. The final frontier will be urbanized, ultimately into an arena where every bit of activity is a meaningful computation: the inhabited portion of the universe will be transformed into a cyberspace.

Because it will use resources more efficiently, a mature cyberspace of the distant future will be effectively much bigger than the present physical universe. While only an infinitesimal fraction of existing matter and space is doing interesting work, in a well developed cyberspace every bit will be part of a relevant computation or storing a useful datum. Over time, more compact and faster ways of using space and matter will be invented, and used to restructure the cyberspace, effectively increasing the amount of computational spacetime per unit of physical spacetime.

Computational speed-ups will affect the subjective experience of entities in the cyberspace in a paradoxical way. At first glimpse, there is no subjective effect, because everything, inside and outside the individual, speeds up equally. But, more subtly, speed-up produces an expansion of the cyber universe, because, as thought accelerates, more subjective time passes during the fixed (probably lightspeed) physical transit time of a message between a given pair of locations—so those fixed locations seem to grow farther apart. Also, as information storage is made continually more efficient through both denser utilization of matter and more efficient encodings, there will be increasingly more cyber-stuff between any two points. The effect may somewhat resemble the continuous-creation process in the old steady-state theory of the physical universe of Hoyle, Bondi and Gold, where hydrogen atoms appear just fast enough throughout the expanding cosmos to maintain a constant density.

A quantum-mechanical entropy calculation by Bekenstein suggests that the ultimate amount of information that can be stored given the mass and volume of a hydrogen atom is about a megabyte. But let's be conservative, and imagine that at some point in the future only "conventional" physics is in play, but every few atoms stores a useful bit. There are about  $10^{56}$  atoms in the solar system. I estimate that a human brain-equivalent can be encoded in less than  $10^{15}$  bits. If a body and surrounding environment takes a thousand times more storage in addition, a human, with immediate environment, might consume  $10^{18}$  bits. An AI with equivalent intelligence could probably get by with less, since it does without the "body-simulation" life support needed to keep a body-oriented human mind sane. So a city of a million human-scale inhabitants might be efficiently stored in  $10^{24}$  bits. If the atoms of the solar system were cleverly rearranged so every 100 could represent a bit, then a single solar system could hold  $10^{30}$  cities—far more than the number ( $10^{22}$ ) of stars in the visible universe! Multiply that by  $10^{11}$  stars in a galaxy, and one gets  $10^{41}$  cities per galaxy. The visible universe, with  $10^{11}$  galaxies, would

then have room for  $10^{51}$  cities—except that by the time intelligence has expanded that far, more efficient ways of using spacetime and encoding data would surely have been discovered, increasing the number much further.

### *Mind without Body?*

Start with the concepts of telepresence and virtual reality. You wear a harness that, with optical, acoustical, mechanical and chemical devices controls all that you sense, and measures all of your actions. Its machinery presents pictures to your eyes, sounds to your ears, pressures and temperatures to your skin, forces to your muscles and even smells and tastes for the remaining senses. Telepresence results when the inputs and outputs of this harness connect to a distant machine that looks like a humanoid robot. The images from the robot's two camera eyes appear on your "eyeglass" viewscreens, and you hear through its ears, feel through its skin and smell through its chemical sensors. When you move your head or body, the robot moves in exact synchrony. When you reach for an object seen in the viewscreens, the robot reaches for the object, and when it makes contact, your muscles and skin feel the resulting weight, shape, texture and temperature. For most practical purposes you inhabit the robot's body—your sense of consciousness has migrated to the robot's location, in a true "out of body" experience.

Virtual reality retains the harness, but replaces the remote robot with a computer simulation of a body and its surroundings. When connected to a virtual reality, the location you seem to inhabit does not exist in the usual physical sense, rather you are in a kind of computer-generated dream. If the computer has access to data from the outside world, the simulation may contain some "real" items, for instance representations of other people connected via their own harnesses, or even views of the outside world, perhaps through simulated windows.

One might imagine a hybrid system where a virtual "central station" is surrounded by portals that open on to views of multiple real locations. While in the station one inhabits a simulated body, but when one steps through a portal, the harness link is seamlessly switched from the simulation to a telepresence robot waiting at that location.

The technical challenges limit the availability, "fidelity" and affordability of telepresence and virtual reality systems today—in fact, they exist only in a few highly experimental demonstrations. But progress is being made, and it's possible to anticipate a time, a few decades hence, when people spend more time in remote and virtual realities than in their immediate surroundings, just as today most of us spend more time in artificial indoor surroundings than in the great outdoors. The remote bodies we will inhabit can be stronger, faster and have better senses than our "home" body. In fact, as our home body ages and weakens, we might compensate by turning up some kind of "volume control." Eventually, we might wish to bypass our atrophied muscles and dimmed senses altogether, if neurobiology learns enough to connect our sensory and motor nerves directly to electronic interfaces. Then all the harness hardware could be discarded as obsolete, along with our sense organs and muscles, and indeed most of our body. There would be no "home" experiences to return to, but our remote and virtual existences would be better than ever.

The picture has us become a "brain in a vat," sustained by life-support machinery, and connected by wonderful electronic links, at will, to a series of "rented" artificial bodies at remote locations, or to simulated bodies in artificial realities. But the brain is a biological machine not designed to function forever, even in an optimal physical environment. As it begins to malfunction, might we not choose to use the same advanced neurological electronics that make possible our links to the external world, to replace the gray matter as it begins to fail? Bit by bit our brain is replaced by electronic equivalents, which work at least as well, leaving our personality and thoughts clearer than ever. Eventually everything has been replaced by manufactured parts. No physical vestige of our original body or brain remains, but our thoughts and awareness

continue. We will call this process, and other approaches with the same end result, the downloading of a human mind into a machine. After downloading, our personality is a pattern impressed on electronic hardware, and we may then find ways to move our minds to other similar hardware, just as a computer program and its data can be copied from processor to processor. So not only can our sense of awareness shift from place to place at the speed of communication, but the very components of our minds may ride on the same data channels. We might find our selves distributed over many locations, one piece of our mind here, another piece there, and our sense of awareness at yet another place. Time becomes more flexible—when our mind resides in very fast hardware, one second of real time may provide a subjective year of thinking time, while a thousand years of real time spent on a passive storage medium may seem like no time at all. Can we then consider ourselves to be a mind without a body? Not quite.

A human totally deprived of bodily senses does not do well. After twelve hours in a sensory deprivation tank (where one floats in a body-temperature saline solution that produces almost no skin sensation, in total darkness and silence, with taste and smell and the sensations of breathing minimized) a subject will begin to hallucinate, as the mind, somewhat like a television tuned to a nonexistent channel, turns up the amplification, desperately looking for a signal, becoming ever less discriminating in the theories it offers to make sense of the random sensory hiss it receives. Even the most extreme telepresence and virtual reality scenarios we have presented avoid complete bodylessness by always providing the mind with a consistent sensory (and motor) image, obtained from an actual remote robot body, or from a computer simulation. In those scenarios, a person may sometimes exist without a physical body, but never without the illusion of having one.

But in our computers there are already many entities that resemble truly bodiless minds. A typical computer chess program knows nothing about physical chess pieces or chessboards, or about the staring eyes of its opponent or the bright lights of a tournament. Nor does it work with an internal simulation of those physical attributes. It reasons instead with a very efficient and compact mathematical representation of chess positions and moves. For the benefit of human players this internal representation is sometimes translated to a recognizable graphic on a computer screen, but such images mean nothing to the program that actually chooses the chess moves. For all practical purposes, the chess program's thoughts and sensations—its consciousness—is pure chess, with no taint of the physical, or any other, world. Much more than a human mind with a simulated body stored in a computer, a chess program is a mind without a body.

So now, imagine a future world where programs that do chess, mathematics, physics, engineering, art, business or whatever, have grown up to become at least as clever as the human mind. Imagine also the most of the inhabited universe has been converted to a computer network—a cyberspace—where such programs live, side by side with downloaded human minds and accompanying simulated human bodies. Suppose that all these entities make their living in something of a free market way, trading the products of their labor for the essentials of life—in this world memory space and computing cycles. Some entities do the equivalent of manual work, converting undeveloped parts of the universe into cyberspace, or improving the performance of existing patches, thus creating new wealth. Others work on physics or engineering problems whose solutions give the developers new and better ways to construct computing capacity. Some create programs that can become part of one's mental capacity. They trade their discoveries and inventions for more working space and time. There are entities that specialize as agents, collecting commissions in return for locating opportunities and negotiating deals for their clients. Others act as banks, storing and redistributing resources, buying and selling computing space, time and information. Some we might class as artists, creating structures that don't obviously result in physical resources, but which, for idiosyncratic reasons, are deemed valuable by some customers, and are traded at prices that fluctuate for subjective reasons. Some entities in the cyberworld will fail to produce enough value to support their requirements for existence—these eventually shrink and disappear, or merge with other ventures. Others will succeed and grow. The closest present

day parallel is the growth, evolution, fragmentation and consolidation of corporations, whose options are shaped primarily by their economic performance.

A human would likely fare poorly in such a cyberspace. Unlike the streamlined artificial intelligences that zip about, making discoveries and deals, reconfiguring themselves to efficiently handle the data that constitutes their interactions, a human mind would lumber about in a massively inappropriate body simulation, analogous to someone in a deep diving suit plodding along among a troupe of acrobatic dolphins. Every interaction with the data world would first have to be analogized as some recognizable quasi-physical entity: other programs might be presented as animals, plants or demons, data items as books or treasure chests, accounting entries as coins or gold. Maintaining such fictions increases the cost of doing business, as does operating the mind machinery that reduces the physical simulations into mental abstractions in the downloaded human mind. Though a few humans may find a niche exploiting their baroque construction to produce human-flavored art, more may feel a great economic incentive to streamline their interface to the cyberspace.

The streamlining could begin with the elimination of the body-simulation along with the portions of the downloaded mind dedicated to interpreting sense-data. These would be and replaced with simpler integrated programs that produced approximately the same net effect in one's consciousness. One would still view the cyberworld in terms of location, color, smell, faces, and so on, but only those details we actually notice would be represented. We would still be at a disadvantage compared with the true artificial intelligences, who interact with the cyberspace in ways optimized for their tasks. We might then be tempted to replace some of our innermost mental processes with more cyberspace-appropriate programs purchased from the AIs, and so, bit by bit, transform ourselves into something much like them. Ultimately our thinking procedures could be totally liberated from any traces of our original body, indeed of any body. But the bodiless mind that results, wonderful though it may be in its clarity of thought and breadth of understanding, could in no sense be considered any longer human.

So, one way or another, the immensities of cyberspace will be teeming with very unhuman disembodied superminds, engaged in affairs of the future that are to human concerns as ours are to those of bacteria. But, once in a long while, humans do think of bacteria, even particular individual bacteria seen in particular microscopes. Similarly, a cyberbeing may occasionally bring to mind a human event of the distant past. If a sufficiently powerful mind makes a sufficiently large effort, such recall could occur with great detail—call it high fidelity. With enough fidelity, the situation of a remembered person, along with all the minutiae of her body, her thoughts, and feelings would be perfectly recreated in a kind of mental simulation: a cyberspace within a cyberspace where the person would be as alive as anywhere. Sometimes the recall might be historically accurate, in other circumstances it could be artistically enhanced: it depends on the purposes of the cybermind. An evolving cyberspace becomes effectively ever more capacious and long lasting, and so can support ever more minds of ever greater power. If these minds spend only an infinitesimal fraction of their energy contemplating the human past, their sheer power should ensure that eventually our entire history is replayed many times in many places, and in many variations. The very moment we are now experiencing may actually be (almost certainly is) such a distributed mental event, and most likely is a complete fabrication that never happened physically. Alas, there is no way to sort it out from our perspective: we can only wallow in the scenery.