



The Coming Merging of Mind and Machine by Ray Kurzweil

Ray Kurzweil predicts a future with direct brain-to-computer access and conscious machines. From Scientific American.

Originally published September 1, 1999 in Scientific American. Published on KurzweilAI.net February 22, 2001.

Sometime early in the next century, the intelligence of machines will exceed that of humans. Within several decades, machines will exhibit the full range of human intellect, emotions and skills, ranging from musical and other creative aptitudes to physical movement. They will claim to have feelings and, unlike today's virtual personalities, will be very convincing when they tell us so. By 2019 a \$1,000 computer will at least match the processing power of the human brain. By 2029 the software for intelligence will have been largely mastered, and the average personal computer will be equivalent to 1,000 brains.

Within three decades, the author maintains, neural implants will be available that interface directly to our brain cells. The implants would enhance sensory experiences and improve our memory and thinking.

Once computers achieve a level of intelligence comparable to that of humans, they will necessarily soar past it. For example, if I learn French, I can't readily download that learning to you. The reason is that for us, learning involves successions of stunningly complex patterns of interconnections among brain cells (neurons) and among the concentrations of biochemicals, known as neurotransmitters, that enable impulses to travel from neuron to neuron. We have no way of quickly downloading these patterns. But quick downloading will allow our nonbiological creations to share immediately what they learn with billions of other machines. Ultimately, nonbiological entities will master not only the sum total of their own knowledge but all of ours as well.

As this happens, there will no longer be a clear distinction between human and machine. We are already putting computers--neural implants--directly into people's brains to counteract Parkinson's disease and tremors from multiple sclerosis. We have cochlear implants that restore hearing. A retinal implant is being developed in the U.S. that is intended to provide at least some visual perception for some blind individuals, basically by replacing certain visual-processing circuits of the brain. Recently scientists from Emory University implanted a chip in the brain of a paralyzed stroke victim that allows him to use his brainpower to move a cursor across a computer screen.

In the 2020s neural implants will improve our sensory experiences, memory and thinking. By 2030, instead of just phoning a friend, you will be able to meet in, say, a virtual Mozambican game preserve that will

seem compellingly real. You will be able to have any type of experience--business, social, sexual--with anyone, real or simulated, regardless of physical proximity.

How Life and Technology Evolve

To gain insight into the kinds of forecasts I have just made, it is important to recognize that technology is advancing exponentially. An exponential process starts slowly, but eventually its pace increases extremely rapidly. (A fuller documentation of my argument is contained in my new book, *The Age of Spiritual Machines*.)

The evolution of biological life and the evolution of technology have both followed the same pattern: they take a long time to get going, but advances build on one another and progress erupts at an increasingly furious pace. We are entering that explosive part of the technological evolution curve right now.

Consider: It took billions of years for Earth to form. It took two billion more for life to begin and almost as long for molecules to organize into the first multicellular plants and animals about 700 million years ago. The pace of evolution quickened as mammals inherited Earth some 65 million years ago. With the emergence of primates, evolutionary progress was measured in mere millions of years, leading to *Homo sapiens* perhaps 500,000 years ago.

The evolution of technology has been a continuation of the evolutionary process that gave rise to us--the technology-creating species--in the first place. It took tens of thousands of years for our ancestors to figure out that sharpening both sides of a stone created useful tools. Then, earlier in this millennium, the time required for a major paradigm shift in technology had shrunk to hundreds of years.

The pace continued to accelerate during the 19th century, during which technological progress was equal to that of the 10 centuries that came before it. Advancement in the first two decades of the 20th century matched that of the entire 19th century. Today significant technological transformations take just a few years; for example, the World Wide Web, already a ubiquitous form of communication and commerce, did not exist just nine years ago.

Computing technology is experiencing the same exponential growth. Over the past several decades, a key factor in this expansion has been described by Moore's Law. Gordon Moore, a co-founder of Intel, noted in the mid-1960s that technologists had been doubling the density of transistors on integrated circuits every 12 months. This meant computers were periodically doubling both in capacity and in speed per unit cost. In the mid-1970s Moore revised his observation of the doubling time to a more accurate estimate of about 24 months, and that trend has persisted through the 1990s.

After decades of devoted service, Moore's Law will have run its course around 2019. By that time, transistor features will be just a few atoms in width. But new computer architectures will continue the exponential growth of computing. For example, computing cubes are already being designed that will provide thousands of layers of circuits, not just one as in today's computer chips. Other technologies that promise orders-of-magnitude increases in computing density include nanotube circuits built from carbon atoms, optical computing, crystalline computing and molecular computing.

We can readily see the march of computing by plotting the speed (in instructions per second) per \$1,000 (in constant dollars) of 49 famous calculating machines spanning the 20th century [see graph below]. The graph is a study in exponential growth: computer speed per unit cost doubled every three years between 1910 and 1950 and every two years between 1950 and 1966 and is now doubling every year. It took 90 years to achieve the first \$1,000 computer capable of executing one million instructions per second (MIPS). Now we add an additional MIPS to a \$1,000 computer every day.

Why Returns Accelerate

Why do we see exponential progress occurring in biological life, technology and computing? It is the result of a fundamental attribute of any evolutionary process, a phenomenon I call the Law of Accelerating Returns. As order exponentially increases (which reflects the essence of evolution), the time between salient events grows shorter. Advancement speeds up. The returns--the valuable products of the process--accelerate at a nonlinear rate. The escalating growth in the price performance of computing is one important example of such accelerating returns.

A frequent criticism of predictions is that they rely on an unjustified extrapolation of current trends, without considering the forces that may alter those trends. But an evolutionary process accelerates because it builds on past achievements, including improvements in its own means for further evolution. The resources it needs to continue exponential growth are its own increasing order and the chaos in the environment in which the evolutionary process takes place, which provides the options for further diversity. These two resources are essentially without limit.

The Law of Accelerating Returns shows that by 2019 a \$1,000 personal computer will have the processing power of the human brain--20 million billion calculations per second. Neuroscientists came up with this figure by taking an estimation of the number of neurons in the brain, 100 billion, and multiplying it by 1,000 connections per neuron and 200 calculations per second per connection. By 2055, \$1,000 worth of computing will equal the processing power of all human brains on Earth (of course, I may be off by a year or two).

The accelerating rate of progress in computing is demonstrated by this graph, which shows the amount of computing speed that \$1,000 (in constant dollars) would buy, plotted as a function of time. Computer power per unit cost is now doubling every year.

Programming Intelligence

That's the prediction for processing power, which is a necessary but not sufficient condition for achieving human-level intelligence in machines. Of greater importance is the software of intelligence.

One approach to creating this software is to painstakingly program the rules of complex processes. We are getting good at this task in certain cases; the CYC (as in "encyclopedia") system designed by Douglas B. Lenat of Cycorp has more than one million rules that describe the intricacies of human common sense, and it is being applied to Internet search engines so that they return smarter answers to our queries.

Another approach is "complexity theory" (also known as chaos theory) computing, in which self-organizing algorithms gradually learn patterns of

information in a manner analogous to human learning. One such method, neural nets, is based on simplified mathematical models of mammalian neurons. Another method, called genetic (or evolutionary) algorithms, is based on allowing intelligent solutions to develop gradually in a simulated process of evolution.

Ultimately, however, we will learn to program intelligence by copying the best intelligent entity we can get our hands on: the human brain itself. We will reverse-engineer the human brain, and fortunately for us it's not even copyrighted!

The most immediate way to reach this goal is by destructive scanning: take a brain frozen just before it was about to expire and examine one very thin slice at a time to reveal every neuron, interneuronal connection and concentration of neurotransmitters across each gap between neurons (these gaps are called synapses). One condemned killer has already allowed his brain and body to be scanned, and all 15 billion bytes of him can be accessed on the National Library of Medicine's Web site. The resolution of these scans is not nearly high enough for our purposes, but the data at least enable us to start thinking about these issues.

We also have noninvasive scanning techniques, including high-resolution magnetic resonance imaging (MRI) and others. Their increasing resolution and speed will eventually enable us to resolve the connections between neurons. The rapid improvement is again a result of the Law of Accelerating Returns, because massive computation is the main element in higher-resolution imaging.

Another approach would be to send microscopic robots (or "nanobots") into the bloodstream and program them to explore every capillary, monitoring the brain's connections and neurotransmitter concentrations.

Fantastic Voyage

Although sophisticated robots that small are still several decades away at least, their utility for probing the innermost recesses of our bodies would be far-reaching. They would communicate wirelessly with one another and report their findings to other computers. The result would be a noninvasive scan of the brain taken from within.

Most of the technologies required for this scenario already exist, though not in the microscopic size required. Miniaturizing them to the tiny sizes needed, however, would reflect the essence of the Law of Accelerating Returns. For example, the translators on an integrated circuit have been shrinking by a factor of approximately 5.6 in each linear dimension every 10 years.

The capabilities of these embedded nanobots would not be limited to passive roles such as monitoring. Eventually they could be built to communicate directly with the neuronal circuits in our brains, enhancing or extending our mental capabilities. We already have electronic devices that can communicate with neurons by detecting their activity and either triggering nearby neurons to fire or suppressing them from firing. The embedded nanobots will be capable of reprogramming neural connections to provide virtual-reality experiences and to enhance our pattern recognition and other cognitive faculties.

To decode and understand the brain's information-processing methods (which, incidentally, combine both digital and analog methods), it is not necessary to see every connection, because there is a great deal of

redundancy within each region. We are already applying insights from early stages of this reverse-engineering process. For example, in speech recognition, we have already decoded and copied the brain's early stages of sound processing.

Perhaps more interesting than this scanning-the-brain-to-understand-it approach would be scanning the brain for the purpose of downloading it. We would map the locations, interconnections, and contents of all the neurons, synapses and neurotransmitter concentrations. The entire organization, including the brain's memory, would then be re-created on a digital-analog computer.

To do this, we would need to understand local brain processes, and progress is already under way. Theodore W. Berger and his co-workers at the University of Southern California have built integrated circuits that precisely match the processing characteristics of substantial clusters of neurons. Carver A. Mead and his colleagues at the California Institute of Technology have built a variety of integrated circuits that emulate the digital-analog characteristics of mammalian neural circuits.

Developing complete maps of the human brain is not as daunting as it may sound. The Human Genome Project seemed impractical when it was first proposed. At the rate at which it was possible to scan genetic codes 12 years ago, it would have taken thousands of years to complete the genome. But in accordance with the Law of Accelerating Returns, the ability to sequence DNA has been accelerating. The latest estimates are that the entire human genome will be completed in just a few years.

By the third decade of the 21st century, we will be in a position to create complete, detailed maps of the computationally relevant features of the human brain and to re-create these designs in advanced neural computers. We will provide a variety of bodies for our machines, too, from virtual bodies in virtual reality to bodies comprising swarms of nanobots. In fact, humanoid robots that ambulate and have lifelike facial expressions are already being developed at several laboratories in Tokyo.

Will It Be Conscious?

Such possibilities prompt a host of intriguing issues and questions. Suppose we scan someone's brain and reinstate the resulting "mind file" into a suitable computing medium. Will the entity that emerges from such an operation be conscious? This being would appear to others to have very much the same personality, history and memory. For some, that is enough to define consciousness. For others, such as physicist and author James Trefil, no logical reconstruction can attain human consciousness, although Trefil concedes that computers may become conscious in some new way.

At what point do we consider an entity to be conscious, to be self-aware, to have free will? How do we distinguish a process that is conscious from one that just acts as if it is conscious? If the entity is very convincing when it says, "I'm lonely, please keep me company," does that settle the issue?

If you ask the "person" in the machine, it will strenuously claim to be the original person. If we scan, let's say, me and reinstate that information into a neural computer, the person who emerges will think he is (and has been) me (or at least he will act that way). He will say, "I grew up in Queens, New York, went to college at M.I.T., stayed in the Boston area, walked into a scanner there and woke up in the machine here. Hey, this

technology really works." But wait, is this really me? For one thing, old Ray (that's me) still exists in my carbon-cell-based brain.

Will the new entity be capable of spiritual experiences? Because its brain processes are effectively identical, its behavior will be comparable to that of the person it is based on. So it will certainly claim to have the full range of emotional and spiritual experiences that a person claims to have.

No objective test can absolutely determine consciousness. We cannot objectively measure subjective experience (this has to do with the very nature of the concepts "objective" and "subjective"). We can measure only correlates of it, such as behavior. The new entities will appear to be conscious, and whether or not they actually are will not affect their behavior. Just as we debate today the consciousness of nonhuman entities such as animals, we will surely debate the potential consciousness of nonbiological intelligent entities. From a practical perspective, we will accept their claims. They'll get mad if we don't.

Before the next century is over, the Law of Accelerating Returns tells us, Earth's technology-creating species-us-will merge with our own technology. And when that happens, we might ask: What is the difference between a human brain enhanced a millionfold by neural implants and a nonbiological intelligence based on the reverse-engineering of the human brain that is subsequently enhanced and expanded?

The engine of evolution used its innovation from one period (humans) to create the next (intelligent machines). The subsequent milestone will be for the machines to create their own next generation without human intervention.

An evolutionary process accelerates because it builds on its own means for further evolution. Humans have beaten evolution. We are creating intelligent entities in considerably less time than it took the evolutionary process that created us. Human intelligence--a product of evolution--has transcended it. So, too, the intelligence that we are now creating in computers will soon exceed the intelligence of its creators.

The Coming Merger of Mind and Machine reproduced with permission. Copyright (C) Scientific American, Inc. All rights reserved.

Original article at Sciam.com