J. Andrew Ross

Will Robots See Humans as Dinosaurs?

Prehistory

The *JCS* report on the ASSC10 conference at St Anne's College, Oxford, in June 2006, by Claude Pasquini (2006) was delightfully evocative of a delightful conference. But I was startled to read this dramatic paragraph:

'Just like these guys' — Andrew Ross pointed to the giant skeletons of the dinosaurs — 'are now the remnants of our distant past, we humans will be the remnants of the distant past of robots which will replace us humans. From the point of view of those future machines the world as we know it will be their primeval soup and humans will have been some biological dirt one could dispense with.' (p. 86)

These are not my exact words, of course, but Pasquini has captured the general sense (or nonsense) of my proclamation quite nicely. The question is whether we should take this sort of utterance seriously. Is there any good reason for believing such ideas? I think there is, and with your indulgence I shall now try to explain.

Robots

The future is hard to predict, but the robots we now build will almost certainly soon be seen as primitive machines. Consider how we now see museum specimens of the machines built only a few decades ago,

Correspondence: Andy Ross   *me@andyross.net*

which our grandparents regarded as the first forerunners of spaceships that would colonise the solar system and so on. Every generation has its utopian visions, it seems, and ours is no exception. Our collective vision, suggested by numerous science-fiction movies as well as by technology futurologist Ray Kurzweil (2005), is apparently that robots in some form will overshadow us not only in high-tech factories but also more generally.

The basis for this idea is a generalisation of Moore's law that has some initial plausibility. Moore's law, to remind you, is the claim that the electronics industry will continue to double the speed and component count per chip of integrated circuits every eighteen months or so for many years. Formulated by Gordon Moore (1965), this 'law' still reflects the facts quite well, and is expected to do so for another few doublings at least. This kind of exponential progress is visible in the initial histories of many new technical developments, such as in the performance of cars or aircraft, although the doubling time of the microelectronics graph is shorter than that of any previous such development. Kurzweil expects similarly fast exponential growth to be evident soon in three new fields: nanotechnology, genetic engineering and robotics.

The basis for the growth of robotics is the increasing power of the microprocessors that serve as the 'brains' of robots. As progress in nanotechnology allows continuing miniaturisation of the components on those chips, robot designers will be able to build more and more intelligence into their products. Human brains have some hundred billion neurons in them, and within a decade our biggest computers will have some hundred billion logic gates in them, so in principle we shall be able to simulate, understand and reconstitute in hardware the higher cognitive functions of human beings.

Lest this be seen as trivialising the task of understanding our higher cognitive functions, consider how parallel developments elsewhere may converge to accelerate progress. The Blue Brain project led by Henry Markram at the Brain Mind Institute in Lausanne (Markram 2006) will use an IBM Blue Gene supercomputer to create the most detailed and accurate simulation to date of neural processes in the mammalian neocortex. The neocortex is a folded layer of tissue that consists essentially of a large number of pyramidal columns that each contain some ten thousand neurons. The rat cortex contains a few thousand of these columns and the human cortex contains a few million of them. The project will simulate one such column. If we understand in detail how the columns work, we shall have understood in principle how the cortex as a whole works. The IBM Blue Gene

contains over eight thousand processors, so the Blue Brain simulation will devote the processing power of a few hundred million transistors to the simulation of each neuron in a column. The simulation will be tuned to reflect exact results obtained from ongoing laboratory work on rat brains.

One can argue that the Blue Brain project is a 'brute force' approach to understanding the brain, but it is not the only approach currently under way. To cite just one example among many, studies involving functional magnetic resonance imaging can realistically hope to achieve in the near future a spatial resolution on the order of a single cortical column and a temporal resolution on the order of an action potential or at least a post synaptic potential (Bandettini *et al.* 2005). These and other studies should increase our understanding of the brain in the next few years by at least as much as the work done in the last decade of the twentieth century, the 'Decade of the Brain', which advanced our understanding by orders of magnitude.

With such understanding, and with nanotech supercomputers the size of a grapefruit not too far away, we can expect to be able to build robots that equal us in terms of raw intellect in the foreseeable future. The hardest question here is how to program such brutes to work intelligently and safely with people. The art of computer programming as it now exists offers little comfort. But new ideas from evolutionary programming based on genetic algorithms and approaches based in outline on how humans learn such skills as how to walk and talk and how to understand what they see and hear show more promise. We can hope that such ideas will enable us to build robots that coexist harmoniously and productively alongside humans.

The other side of the story is motivation, and there the message is clear. Numerous well funded groups worldwide are building robotic systems for a wide variety of purposes, from domestic tasks such as vacuum cleaning through routine tasks such as driving vehicles over natural terrain to more telegenic tasks such as playing football at tournament level. Purely military applications bring more funding to the table. Pentagon projects for unmanned combat air vehicles involve massive robotics investment and bring readily defensible benefits in terms of delivering ordnance precisely on target without risking pilots. More benignly, medical robots slaved to a surgeon working through a remote interface can already perform operations beyond the ability of unaided humans. In all these ways, we are welcoming robots into our lives as fast as we can.

Humans

For all its faults, this is a golden age in the history of *Homo sapiens*. We have achieved unprecedented knowledge of ourselves and our past, as well as our natural environment in the physical universe, and realised just how insignificant we are in the cosmic perspective. Yet despite this humbling knowledge, we are still able to see ourselves as lords of creation, in the sense that we are not yet aware of any creatures anywhere whose powers essentially exceed our own.

There are two main ways in which this could change. Either we could discover extraterrestrial planets hosting advanced civilizations that make their presence known to us in ways we cannot or dare not imagine. Or we could build robots that build more robots and so on to build beings whose powers transcend our own.

Given the astronomical distances between earthlike planets and the discouraging effect of that distance on any communication between them, as well as the fact of our cosmic isolation to date, it seems quite likely that our demise as lords of creation will be our own doing. We can either endure an ignominious demise, as we blow each other up or drown in our own toxic waste and get replaced by a master race of cockroaches, or we can enjoy a more decorous demise, like that of parents who give up all they have to their children. More likely, some of us, those who build and control the robots, will retire decorously to pampered communities that will be in effect human zoos, while all the less privileged will ultimately be forced to go down fighting. But now, before that parting of the ways, each of us individually has the choice of either resisting every incursion of the machines or relaxing and learning to see them as our collective progeny.

Apart from all that, there is a different and more likely way in which humans as we now know them will probably soon cease to be the lords of creation, namely that we shall learn to rebuild and improve ourselves. This is the promise of genetic engineering. Our enhanced successors may call themselves *Homo superior* and look back on us as simian old saps who were lucky to have survived as long as we did. Their superior enhancements may take them a lot further than the present state of genetic engineering would suggest. Genetic engineering is still in what we might call the tinkering phase, where methods are inefficient and our detailed understanding of gene functions and interactions is limited, but this is changing fast. In parallel with advances in nanotechnology and robotics, and building on many of the same enabling developments in science and society, progress in genetic engineering, or rather synthetic biology, will accelerate

exponentially, too, until we can essentially rebuild our genomes to realise any traits we can competently design.

The key to understanding the opportunities that await us here is to remember that DNA molecules together with RNA and ribosomes function as nanoscale Turing machines (Shapiro & Benenson 2006) and can be reprogrammed as freely as any computer. All we need is a suitable set of nanomanipulation tools and some reasonable idea of what we want to do. Soon we shall be building synthetic organisms to do such things as recycle waste materials or environmental toxins. Soon after, we shall build organisms to seek and destroy tumors or rejuvenate telomeres. But these are the baby steps.

Imagine what this capability could do for people who want to 'improve' themselves. Long legs, superenhanced immune systems, arbitrarily extended lifetimes, genius intellects, cosmic consciousness, pimple-free faces and so on will be just the beginning. Whatever turns someone on will be available at a price, whether others like it or not. People will inject themselves with nanobots that seek out and upgrade their DNA to keep up with the latest fashions. In more organized societies, mass production of superman clones will be sure to follow. Soon after that, disenchanted by the random variety of species that nature has evolved, our superior cousins will doubtless embark upon some serious extermination programmes.

However, the greatest benefit of all this genetic reprogramming will be that *Homo superior* will achieve convergence with machines. Instead of staring at images on screens, we shall be able to shoot the bit streams that made up the images directly to visual cortex via carbon nanotube electrodes plugged into individual neurons. Then we shall refashion the bit streams to reduce the computational load on those superior brains and thus enhance their capacity to cerebrate on the refashioned contents. And instead of controlling machines via levers and buttons, we shall rely on the nanotube electrodes to relay our thoughts to the machines. We shall drive rovers on Mars or flap our own robot wings as easily as we focus our thoughts.

The same technology will enable enhanced people to commune with each other directly, brain to brain, and, if it turns them on, to enjoy about the same transmission bandwidth between their brains as we have between our cerebral hemispheres via the *corpus callosum*, and thus effectively remove interpersonal barriers between each other. If used widely for daily intercourse, this will achieve for the superior beings a kind of collective consciousness. When globally networked machines become integrated as intimately as this with enhanced

humans, the entire regimented and engineered ecosystem of life on Earth will become for all practical purposes a single global organism.

This is still science fiction. But the elements are compelling, for me at least, and the juggernaut of science and technology, as well as the global economic machine that supports it, is clearly moving in this direction.

Dinosaurs

The convergence of humans with machines is likely to happen both incrementally, with no sudden discontinuity, and faster than we might now expect. By analogy, consider how far we have come already, and how fast. Ten thousand years ago, people considered things like rain-proof houses or warm clothing to be luxuries. One thousand years ago, false teeth and eyeglasses were still rarities. A hundred years ago, people still walked a lot and died of things like the common cold. Ten years ago, people still communicated everyday information on paper. Now most of us in the Western world can hardly imagine life without air conditioning, cars, televisions, drugs for all our ailments, schools that teach evolution and big bang cosmology, online living via email and the web, and so on.

Everything in recent history suggests that even with the crises and catastrophes that will doubtless plague us in future, the exponentially accelerating intimacy of our embrace of the products of our own technology will continue. Our successors in the not-so-distant future will probably find the prospect of life without their latest technology to be unthinkable. Indeed if they mess up their natural environment fast enough, the sort of unaugmented lifestyles that some of us can still enjoy will become simply impossible.

We look back a hundred million years to see the dinosaurs. Given the exponential increases in the rates of evolution of new forms, either of biological species or of technological products, in the terrestrial exosphere, we need only look forward a few centuries to imagine how humans as we now know them may seem to our superior successors to be more like dinosaurs than fellow beings.

Let us imagine, then, that the enhanced humans who live in symbiosis with the machines a few centuries hence are genetically augmented to resist engineered viruses pumped into air and water to keep down the feral humans who still lurk in desert caves and jungles. These genetically enhanced symbionts communicate wirelessly via nanotech implants in their skulls, like the Borg drones in the inspired *Star Trek* stories, and do so in a global web with enough bandwidth to

sustain a collective consciousness. This is worth elaborating more carefully for the discerning readers of the *JCS*.

Consciousness has two sides. We are conscious of an external world and we are conscious of ourselves. Nowadays, with the help of science and modern media, we know fairly well how most of our fellows see and understand the external world, and our remaining scruples about their inscrutability tend to fall back on such things as the ineffability of the colour red or the difficulty of sharing the feelings of people from radically different cultures. As for our selves, we know that they are in part constructions of limited validity, made and remade to suit our social purposes, and based on feelings arising from a body that has its own unique constitution and history but is similar in all essentials to every other human body. If we could know each other's thoughts as easily as our own left and right hemispheres know each other's thoughts, which I believe will soon be within the bounds of technical possibility, then nothing would prevent us from claiming quite reasonably to share a collective consciousness. Nothing in our present understanding of consciousness prevents such sharing in principle, and plenty in the transpersonal tradition suggests that sharing in this way could be both benign and enlightening.

To be more specific about this enlightened community of symbionts, imagine that they live in intelligent pods that are connected into a body I shall tendentiously call the Global Online Dominion, which executes governance functions for the entire integrated and online global economic machine. Whenever the symbionts emerge from their pods, they do so in robotic exoskeletons that they plug into via nanotube implants to allow direct neural control. Further, these robotic suits interface seamlessly in turn with cars and aircraft and so on, to provide what they regard as the normal implementation of physical empowerment, as opposed to the virtual empowerment they enjoy as online agents when they are physically in their pods.

These posthuman organisms regard their robot suits as parts of themselves. We can well imagine they love their robosuits and obsess about the latest accessories for them. They think of their shrunken and repurposed biological bodies as mere cores. They see their own DNA coding rather as we see Windows, as mere functional code that is subject to overnight automatic update from a genome bank in the Global Online Dominion. For practical purposes these posthumans are inseparable from their pods and their suits. We need not pursue the fantasy further, since such ideas have been well mined by science fiction authors.

The point of this extended fantasy is that such enhanced and superior members of a global collective organism, inseparable as they are from their pods and their robosuits, will no longer think of themselves as human at all. They are more likely to see themselves as robots with soul. Their posthuman cores will serve as mere gateways to the online collective consciousness veiling what for them is the mystic union of their souls. In fact they may see themselves as angelic beings who each day become temporarily incarnate as robotic cyborgs to maintain their physical world.

Now to the crunch. These angels will sometimes contemplate the prehistory of the global organism they naturally call GOD. Perhaps they will visit museums and view dioramas representing the earlier subspecies of humans, accurately depicted in representative scenes, for example on a campsite by a beach, surrounded by old cars, messy dogs, filthy toilets, flyblown garbage, smoky barbecues, melting ice cream and so on. They will look at the people and see flabby bellies, blotchy skins, brutish hair, clawlike nails, rotten teeth and the like. Perhaps they will imagine the contents of those primitive minds, filled with images of hard-core sex and junk food, football and fast cars, scanty clothes and celebrity gossip, and think … *dinosaurs!*

References

Bandettini, P. A., N. Petridou, and J. Bodurka (2005), 'Direct detection of neuronal activity with MRI: fantasy, possibility, or reality?', *Applied Magnetic Resonance*, 29(1). http://fim.nimh.nih.gov/people/pab/P48

Kurzweil, Ray (2005), *The Singularity is Near* (London: Duckworth). http://www.kurzweilai.net

Markram, Henry (2006), http://bluebrainproject.epfl.ch

Moore, Gordon (1965), http://www.intel.com/technology/silicon/mooreslaw/

Pasquini, Claude (2006), 'A (mostly) sunny and sober anniversary', *Journal of Consciousness Studies*, **13** (6), pp. 79–100.

Shapiro, Ehud, and Yaakov Benenson (2006), Bringing DNA computers to life, *Scientific American*, 294(5), 32–39. http://www.sciam.com