“Why robots will never rule the world” a Café Scientifique event

Chris Malcolm, a retired lecturer from the School of Informatics, at the University of Edinburgh gave an absorbing, clear talk littered with colourful and simple analogies, debating why he thinks robots will never rule the world to a packed, sceptical Filmhouse café on Monday the 18th of February 2008.

In 1872 Samuel Butler anonymously published *Erewhon*, a satire on Victorian society and the first novel to instil a fear of machines developing consciousness through Darwinian selection. It was ridiculed at the time, with many suggesting that these ideas reduced Darwin’s theories to an absurdity. Today, the increasing ability of robots has made it a serious topic of debate.

Eminent researchers and robotocists Professors Moravec and Warwick both believe that computers will exceed the power or intelligence of the human brain. Moravec believes this to be the next stage of evolution through our “mind children”. Whilst Warwick doesn’t believe that we will be so willing to hand over dominance of the planet to robots. More importantly, we will be helpless to stop them. Ray Kurzweil, inventor and futurist, takes a different approach "We'll have intelligent nanobots go into our brains through the capillaries and interact directly with our biological neurons," thus merging humans with computers.

The speaker eloquently explained the power of Moore’s well known first law that predicts how computing power will double every 18 months. This law, which has held true from the ‘60s till today, implies that no matter how much we have underestimated the power of our brain, either individually or in a group, computing power will unavoidably overtake our brain power. Furthermore this will happen sooner rather than later; it might only be a few decades before the computer I am typing on now becomes cleverer than me.

But there are physical and technological limitations to computing power. Microchips can only be so small and we have already reached nanometer dimensions for silicon microchips and will soon come up against the limit of microchip miniaturisation. However, humanity has time and again jumped seamlessly from one computing technology to another. Will tomorrow’s computer be quantum, optical or DNA-based?

Another problem is expense: Moore’s second law states that expense will also exponentially increase as transistor size decreases. Smaller transistors are required in order to “pack the same punch” within a computer. Therefore, computer component factories are becoming ever-more expensive to build. A simple statistic can place this problem into perspective: sales of windows-compatible PCs have made $3,100 billion since 1981 and most of that was in the last 5 years; on its own, Microsoft managed to rake in $51.12 billion in 2007.

So if money or technological limitations can’t hold back the rise of the robots, what can? Malcolm emphasised that the previous arguments for robots taking over the world assume that intelligence is equal to computing power. However, intelligence is much more than the raw ability to make calculations – computers can already beat us

at chess. Intelligence requires the ability to learn, to understand and to interpret. Artificial intelligence is already highly sophisticated, yet its improvement requires research and research is a slow process.

Malcolm wittily followed this argument with a story about a robotocist and a physicist. The robotocist dared the physicist that she'd be able to create a robot the physicist would be unable to destroy. A year later the robotocist showed the physicist a small furry robot which looked far from invincible, handing her a hammer to smash it with. But when the physicist lifted the hammer, the robot promptly flipped onto its back and squeaked piteously. Distressed, the physicist was unable to destroy it. The robot was simply programmed to react in such a manner. It did not feel actual fear. However, the human associated the robots behaviour with hard wired instinct that would make it very difficult to hurt the robot. This analogy cleverly shows that robots would simply be an imitation of humans and not actual humans possessing human desires. Therefore, even though humans naturally want to lord over everything, why should a robot?

The key difference is that humans have been crafted by 3.7 billion years of evolution. The process of evolution is a simple concept with an incredible level of complexity behind it that computers have not gone through. So where would they gain the desire to be, like humans, voraciously self-replicating, lust after power, out-compete others for resources and, above all else, survive? It might be possible to program all these qualities into a robot, but why would we want to? Furthermore, great leaps in research would be required to reach the level of technology and software development needed to instil such attributes into a robot.

The discussion and Q&A that followed also touched on topics ranging from what is consciousness? Can we test it and can we program it? What does being clever mean; is a chess player cleverer than a footballer? (Interestingly the footballer came top in this one). The event closed at 10.30, however, some of the audience remained with the speaker, continuing the discussion of the fascinating topics raised at this Café Scientifique event.

Edward Duca
February 2008

Further reading:

http://www.dai.ed.ac.uk/homes/cam/ - Chris Malcolm's homepage
http://en.wikipedia.org/wiki/Technological_singularity - the "intelligence explosion".
http://en.wikipedia.org/wiki/Erewhon - Wikipedia summary of Victorian sci-fi novel Erewhon
http://www.c-i-a.com/pr0806.htm - Almanac-Press release: PC sales statistics
http://news.bbc.co.uk/1/hi/technology/7085019.stm - BBC guide to future computing
http://news.bbc.co.uk/1/hi/world/americas/7248875.stm - BBC article *Machines 'to match man by 2029'*