The Computer and the Brain, and the Potential of the Other 95%

Tai Sing Lee
Carnegie Mellon University
July 15, 2015.

By some accounts, the formation of the Computer Science Department 50 years ago was in large part motivated by Herb Simon and colleagues' agenda of using computers to simulate and understand human intelligence and decision-making. The founding fathers of the field of computer science in fact cared deeply about understanding the brain. Alan Turing and a number of mathematicians shared beers with biologists like Horace Barlow in their informal dining Ratio Club in Cambridge to discuss the principles of cybernetics and the neural basis of intelligence. Computer science and neuroscience had been more intimately coupled than most practitioners of either field might realize nowadays. The last paper John Von Neumann wrote, on his deathbed, was his contemplation on the connection between *the Computer and the Brain*.

Computer science and neuroscience have made tremendous strides in the last 50 years. Thanks to the visionaries who founded and lead CSD and later the School of Computer Science, Carnegie Mellon has been at the forefront of many innovations in artificial intelligence. Allen Newell and Herb Simon's Logic Theorist in 1955 was the first AI program. It proved 38 of the 52 well-known theorems and even discovered a more elegant new proof on its own! Hans Berliner's BKG Backgammon defeated the reigning world champion at the time in 1979. Feng-Hsiung Hsu's Deep Thoughts achieved grand-master level in 1988 when he was a third year Ph.D. student here. Later his Deep Blue team at IBM defeated the reigning world chess champion Gary Kasparov in 1997. Kai-Fu Lee, Xuedong Huang and Raj Reddy's CMU Sphinx was the first successful continuous-speech, speaker-independent speech recognition system. Achievements like Chuck Thorp and Dean Pomerleau's autonomous vehicle driving itself across the country in 1995, Red Whitaker's Red team winning the DARPA's Urban Challenge in 2007, Manuela Veloso winning a number of Robot soccer world cups continued to make AI history.

On the other hand, as the fields of computer science and neuroscience have grown explosively, they have also grown apart. Only a very tiny fraction of the researchers in either field now actually care about the connection and relationship between the brain and the computer. It is no longer a topic of mainstream discourse in either field. While the field of artificial intelligence is close to solving many practical problems such as self-driving car, object and speech recognition, progress in understanding the neural basis and computational principles of human intelligence has been remarkably slow. Many of the recent remarkable accomplishment of machine learning and neural networks are due more to big data and a dramatic increase in computing power, rather than new theoretical insight or improved understanding of the computational principles of intelligence. Geoff Hinton's deep learning or convolution neural network beat the state-of-the-art computer vision algorithms in object recognition using the neuron model, developed by McCulloch and Pitt 80 years ago, and the back-propagation algorithms developed 40 years ago. While convolution neural networks came back to dominate computer vision and speech recognition for the time being, it has provided no additional fundamental insights to neural computation and human intelligence that we didn't already know back in the 80s.

It is true that a revolution is coming in neuroscience. With the rapid development of new recording and manipulation technologies in neuroscience, there is a rapid growth in demand for advanced computing and machine learning in neural data analysis and modeling. With the Human Brain project in Europe, and the Brain Initiative programs in the United States and China, an international brain race is raging to unravel the secrets of the brain, believing that better knowledge of the brain will one day allow us to build better and more powerful computers that can be as flexible, adaptive, creative, imaginative and intelligent as humans, beating humans in our own games. John von Neumann has described a "singularity" in human history as a point in time when accelerating progress in technology lead to an intelligence explosion or emergence of super-intelligence beyond which human society will no longer be the same. Ray Kurzweil has argued that singularity is fast approaching as computers, robots and the internets are becoming capable of recursive self-improvement, redesigning and programming themselves. Elon Musk has warned against summoning the artificial intelligence demon which ultimately can pose the most serious existential threat to the human race, as we know it.

How close are we to this "singularity"? I can't know for sure. But I am not as optimistic or pessimistic as our modern-day technological heroes and prophets. Probably not in 50 years, I would think. For one thing, our understanding of the neural basis of intelligence remains rather limited. We are still far away from understanding the most basic mechanisms and circuits underlying learning, memories, reasoning and decision making, not to mention creativity and imagination. The neural networks for producing all the marvelous state-of-the-art results in object and speech recognition nowadays are essentially a feed-forward cascade of McCulloch-Pitt neurons doing a sophisticated kind of regression. Neurons in the brain, those in the visual system for example, are listening to signals coming from the input path with only 5% of their synapses. What it could mean is that the neural networks nowadays are utilizing only 5% of the machinery or potential of the brain. A better understanding of what the other 95% synapses do, I believe, would likely help make computing machines more flexible, adaptive, imaginative, creative, predictive and introspective in the future. Already computers can fly airplanes and drive cars, play chess and video games better than most of us. In 50 years, computers and robots will likely pass the Turing test in the imitation game. They could be quite autonomous in thinking and behaviors that we probably will take for grant that they might have some form of consciousness, in the sense that they have an sufficient internal representation of themselves for introspection and deliberation, for selecting actions based on predictions of their consequences. Were we able to understand the other 95% of the brain's secrets, for constructing our internal models of the world to reason and to guide our action, we will be one step closer to making robots and computers that are more human like, that can grow like a baby and learn like a kid. However, robots are still an imitation of human, created after our own images, and at best approximate our intelligence.

Ironically, the internet might evolve and develop into a form of new intelligence and consciousness without intentional "intelligent" design. This form of intelligence could actually be more similar to the working of our brain and mind than current artificial intelligence and neural networks. Connected by billions of humans, machines, sensors and actuators as well as the vast human knowledge, it will have a mind of its own, breathing and pulsing like a super-organism with humans and machines like bees in a beehive, very much like the spider-like web French philosopher Denis Diderot envisioned to describe how our minds work in D'Alembert's dream back in 1769. It might be sooner that the internet, rather than AI robots, will develop a mind can rival that of ours in intelligence. Such an intelligent "life" form might have the great potential for evil, as is already embodied in the dark web. It might also have great capacity for

good. It is therefore important for us to reconvene the Ratio Club, where computer scientists and neuroscientists will have beers together -- invite the sociologists, psychologists and philosophers as well, to contemplate intelligence and the mind, natural and artificial, so that we can answer the challenges of the emerging new super-organism of collective human intelligence and machine intelligence connected in a web, and to harness this Prometheus fire to shape human destiny in a positive way.

Related readings:

1. Diderot, D. (1769) *D'Alembert's Dream* (*Le Rêve D'Alembert*).
2. von Neumann, J. (1958) *The Computer and the Brain*. Yale University Press.
3. Simon, H. (1969) The Sciences of the Artificial. MIT Press.
4. Kurzweil, R. (2005) *The Singularity is Near*. New York: Viking Books.
5. Douglas, R.J. and Martin, K.A.C. (2012) Behavioral Architecture of the Cortical Sheet, *Current Biology*, vol 22 (24), 1033-1038.
6. LeCun, Y., Bengio, Y. and Hinton, G.E. (2015) Deep Learning. *Nature*, vol 521, 436-444.
7. Lee, T.S. (2015) The Visual System's Internal Model of the World. *Proceedings of the IEEE*, vol 103 (8), 1359-1378.