What to Think About Machines That Think

Today's Leading Thinkers on the Age of Machine Intelligence

EDITED BY JOHN BROCKMAN

HARPER PERENNIAL

NEW YORK • LONDON • TORONTO • SYDNEY • NEW DELHI • AUCKLAND

AI WILL MAKE YOU SMARTER

TERRENCE J. SEJNOWSKI
Computational neuroscientist; Francis Crick Professor, Salk Institute for Biological Studies; coauthor (with Steven R. Quartz), *Liars, Lovers, and Heroes: What the New Brain Science Reveals About How We Become Who We Are*

Deep learning is today's hot topic in machine learning. Neural-network learning algorithms were developed in the 1980s, but computers were slow back then and could simulate only a few hundred model neurons, with one layer of "hidden units" between the input and output layers. Learning from examples is an appealing alternative to rule-based AI, which is highly labor-intensive. With more layers of hidden units between the inputs and outputs, more abstract features can be learned from the training data. Brains have billions of neurons in cortical hierarchies ten layers deep. The big question back then was how much the performance of neural networks could improve with the size and depth of the network. What was needed was not only much more computer power but also a lot more data to train the network.

After thirty years of research, a million-times improvement in computer power, and vast data sets from the Internet, we now know the answer to this question: Neural networks scaled up to twelve layers deep, with billions of connections, are outperforming the best algorithms in computer vision for object recognition and have revolutionized speech recognition. It's rare for any algorithm to scale this well, which suggests that they may soon be able to solve even more difficult problems. Recent breakthroughs have been made that allow the application of deep learning to natural-language processing. Deep recurrent networks with short-term memory were trained to translate English sentences into French sentences at high levels of performance. Other deep-learning networks could create English captions for the content of images with surprising and sometimes amusing acumen.

Supervised learning using deep networks is a step forward, but still far from achieving general intelligence. The functions they perform are analogous to some capabilities of the cerebral cortex, which has also been scaled up by evolution, but to solve complex cognitive problems the cortex interacts with many other brain regions.

In 1992, Gerald Tesauro at IBM, using reinforcement learning, trained a neural network to play backgammon at a world-champion level. The network played itself, and the only feedback it received was which side won the game. Brains use reinforcement learning to make sequences of decisions toward achieving goals, such as finding food under uncertain conditions. Recently, DeepMind, a company acquired by Google in 2014, used deep reinforcement learning to play seven classic Atari games. The only inputs to the learning system were the pixels on the video screen and the score, the same inputs humans use. The program for several of the games could play better than expert humans.

What effect will these advances have on us in the near future? We're not particularly good at predicting the impact of a new invention, and it often takes time to find its niche, but we already have one example to help us understand how this could unfold. When Deep Blue beat Garry Kasparov, the world chess champion, in 1997, did human chess players give up trying to compete with machines? Quite the contrary:

Humans have used chess programs to improve their game, and as a consequence the level of play in the world has improved.

Humans aren't the fastest or the strongest species, but we're the best learners. Humans invented formal schools where children labor for years to master reading, writing, and arithmetic and to learn more specialized skills. Students learn best when an adult teacher interacts with them one-on-one, tailoring lessons for that student. However, education is labor-intensive. Few can afford individual instruction, and the assembly-line classroom system found in most schools today is a poor substitute. Computer programs can keep track of a student's performance, and some provide corrective feedback for common errors. But each brain is different, and there's no substitute for a human teacher who has a long-term relationship with the student. Is it possible to create an artificial mentor for each student? We already have recommender systems on the Internet that tell us, "If you liked X, you might also like Y," based on data of many others with similar patterns of preference.

Someday the mind of each student may be tracked from childhood by a personalized deep-learning system. To achieve this level of understanding of a human mind is beyond the capabilities of current technology, but there are already efforts at Facebook to use their vast social database of friends, photos, and likes to create a Theory of Mind for every person on the planet.

So my prediction is that as more and more cognitive appliances, like chess-playing programs and recommender systems are devised, humans will become smarter and more capable.