The Opinion Pages | OP-ED CONTRIBUTORS

Where Computers Defeat Humans, and Where They Can't

By ANDREW McAFEE and ERIK BRYNJOLFSSON   MARCH 16, 2016

ALPHAGO, the artificial intelligence system built by the Google subsidiary DeepMind, has just defeated the human champion, Lee Se-dol, four games to one in the tournament of the strategy game of Go. Why does this matter? After all, computers surpassed humans in chess in 1997, when IBM's Deep Blue beat Garry Kasparov. So why is AlphaGo's victory significant?

Like chess, Go is a hugely complex strategy game in which chance and luck play no role. Two players take turns placing white or black stones on a 19-by-19 grid; when stones are surrounded on all four sides by those of the other color they are removed from the board, and the player with more stones remaining at the game's end wins.

Unlike the case with chess, however, no human can explain how to play Go at the highest levels. The top players, it turns out, can't fully access their own knowledge about how they're able to perform so well. This self-ignorance is common to many human abilities, from driving a car in traffic to recognizing a face. This strange state of affairs was beautifully summarized by the philosopher and scientist Michael Polanyi, who said, "We know more than we can tell." It's a phenomenon that has come to be known as "Polanyi's Paradox."

Polanyi's Paradox hasn't prevented us from using computers to accomplish complicated tasks, like processing payrolls, optimizing flight schedules, routing telephone calls and calculating taxes. But as anyone who's written a traditional computer program can tell you, automating these activities has required painstaking precision to explain exactly what the computer is supposed to do.

This approach to programming computers is severely limited; it can't be used in the many domains, like Go, where we know more than we can tell, or other tasks like recognizing common objects in photos, translating between human languages and diagnosing diseases — all tasks where the rules-based approach to programming has failed badly over the years.

Deep Blue achieved its superhuman performance almost by sheer computing power: It was fed millions of examples of chess games so it could sift among the possibilities to determine the optimal move. The problem is that there are many more possible Go games than there are atoms in the universe, so even the fastest computers can't simulate a meaningful fraction of them. To make matters worse, it's usually far from clear which possible moves to even start exploring.

What changed? The AlphaGo victories vividly illustrate the power of a new approach in which instead of trying to program smart strategies into a computer, we instead build systems that can learn winning strategies almost entirely on their own, by seeing examples of successes and failures.

Since these systems don't rely on human knowledge about the task at hand, they're not limited by the fact that we know more than we can tell.

AlphaGo does use simulations and traditional search algorithms to help it decide on some moves, but its real breakthrough is its ability to overcome Polanyi's Paradox. It did this by figuring out winning strategies for itself, both by example and from experience. The examples came from huge libraries of Go matches between top players amassed over the game's 2,500-year history. To understand the strategies that led to victory in these games, the system made use of an approach known as deep learning, which

has demonstrated remarkable abilities to tease out patterns and understand what's important in large pools of information.

Learning in our brains is a process of forming and strengthening connections among neurons. Deep learning systems take an analogous approach, so much so that they used to be called "neural nets." They set up billions of nodes and connections in software, use "training sets" of examples to strengthen connections among stimuli (a Go game in process) and responses (the next move), then expose the system to a new stimulus and see what its response is. AlphaGo also played millions of games against itself, using another technique called reinforcement learning to remember the moves and strategies that worked well.

Deep learning and reinforcement learning have both been around for a while, but until recently it was not at all clear how powerful they were, and how far they could be extended. In fact, it's still not, but applications are improving at a gallop, with no end in sight. And the applications are broad, including speech recognition, credit card fraud detection, and radiology and pathology. Machines can now recognize faces and drive cars, two of the examples that Polanyi himself noted as areas where we know more than we can tell.

We still have a long way to go, but the implications are profound. As when James Watt introduced his steam engine 240 years ago, technology-fueled changes will ripple throughout our economy in the years ahead, but there is no guarantee that everyone will benefit equally. Understanding and addressing the societal challenges brought on by rapid technological progress remain tasks that no machine can do for us.

Andrew McAfee is a principal research scientist at M.I.T., where Erik Brynjolfsson is a professor of management. They are the co-founders of the M.I.T. Initiative on the Digital Economy and the authors of "The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies."

Follow The New York Times Opinion section on Facebook and Twitter, and sign up for the Opinion Today newsletter.