



## Why Self-Taught Artificial Intelligence Has Trouble With the Real World

The latest artificial intelligence systems start from zero knowledge of a game and grow to world-beating in a matter of hours. But researchers are struggling to apply these systems beyond the arcade.

*By Joshua Sokol*



[maxuser](#)

Until very recently, the machines that could trounce champions were at least respectful enough to start by learning from human experience.

To beat Garry Kasparov at chess in 1997, IBM engineers made use of centuries of chess wisdom in their Deep Blue computer. In 2016, Google DeepMind's [AlphaGo thrashed champion Lee Sedol](#) at the ancient board game Go after poring over millions of positions from tens of thousands of human

games.

But now artificial intelligence researchers are rethinking the way their bots incorporate the totality of human knowledge. The current trend is: Don't bother.

Last October, the DeepMind team [published](#) details of a new Go-playing system, AlphaGo Zero, that studied no human games at all. Instead, it started with the game's rules and played against itself. The first moves it made were completely random. After each game, it folded in new knowledge of what led to a win and what didn't. At the end of these scrimmages, AlphaGo Zero went head to head with the already superhuman version of AlphaGo that had beaten Lee Sedol. It won 100 games to zero.



[DeepMind](#)

Lee Sedol, 18-time Go world champion, during his match against AlphaGo in 2016.

The team went on to create what would become another master gamer in the AlphaGo family, this one called simply AlphaZero. In a [paper](#) posted to the scientific preprint site arxiv.org in December, DeepMind researchers revealed that after starting again from scratch, the trained-up AlphaZero outperformed AlphaGo Zero — in other words, it beat the bot that beat the bot that beat the best Go players in the world. And when it was given the rules for chess or the Japanese chess variant shogi, AlphaZero quickly learned to defeat bespoke top-level algorithms for those games, too. Experts marveled at the program's aggressive, unfamiliar style. "I always wondered how it would be if a superior species landed on Earth and showed us how they played chess," Danish grandmaster Peter Heine Nielsen [told](#) a BBC interviewer. "Now I know."

The past year also saw otherworldly self-taught bots emerge in settings as diverse as no-limit poker

and Dota 2, a hugely popular multiplayer online video game in which fantasy-themed heroes battle for control of an alien world.

Of course, the companies investing money in these and similar systems have grander ambitions than just dominating video-game tournaments. Research teams like DeepMind hope to apply similar methods to real-world problems like building room-temperature superconductors, or understanding the origami needed to fold proteins into potent drug molecules. And of course, many practitioners hope to eventually build up to artificial general intelligence, an ill-defined but captivating goal in which a machine could think like a person, with the versatility to attack many different kinds of problems.

Yet despite the investments being made in these systems, it isn't yet clear how far past the game board the current techniques can go. "I'm not sure the ideas in AlphaZero generalize readily," said [Pedro Domingos](#), a computer scientist at the University of Washington. "Games are a very, very unusual thing."

## Perfect Goals for an Imperfect World

One characteristic shared by many games, chess and Go included, is that players can see all the pieces on both sides at all times. Each player always has what's termed "perfect information" about the state of the game. However devilishly complex the game gets, all you need to do is think forward from the current situation.

Plenty of real situations aren't like that. Imagine asking a computer to diagnose an illness or conduct a business negotiation. "Most real-world strategic interactions involve hidden information," said [Noam Brown](#), a doctoral student in computer science at Carnegie Mellon University. "I feel like that's been neglected by the majority of the AI community."

Poker, which Brown specializes in, offers a different challenge. You can't see your opponent's cards. But here too, machines that learn by playing against themselves are now reaching superhuman levels. In January 2017, a program called Libratus created by Brown and his adviser, [Tuomas Sandholm](#), [outplayed](#) four professional poker players at heads-up, no-limit Texas Hold' em, finishing \$1.7 million ahead of its competitors at the end of a 20-day competition.

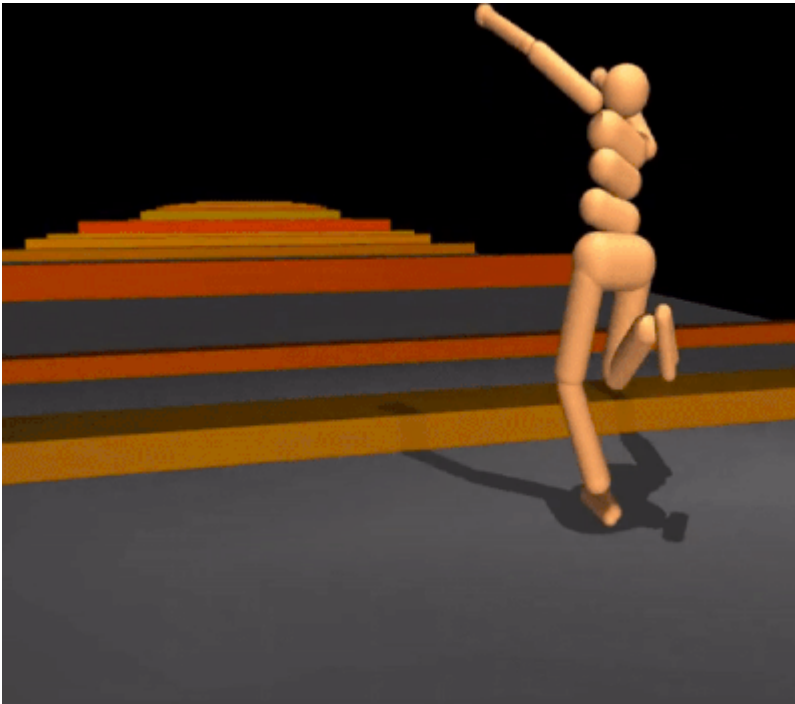
An even more daunting game involving imperfect information is StarCraft II, another multiplayer online video game with a vast following. Players pick a team, build an army and wage war across a sci-fi landscape. But that landscape is shrouded in a fog of war that only lets players see areas where they have soldiers or buildings. Even the decision to scout your opponent is fraught with uncertainty.

This is one game that AI still can't beat. Barriers to success include the sheer number of moves in a game, which often stretches into the thousands, and the speed at which they must be made. Every player — human or machine — has to worry about a vast set of possible futures with every click.

For now, going toe-to-toe with top humans in this arena is beyond the reach of AI. But it's a target. In August 2017, DeepMind [partnered](#) with Blizzard Entertainment, the company that made StarCraft II, to release tools that they say will help open up the game to AI researchers.

Despite its challenges, StarCraft II comes down to a simply enunciated goal: Eradicate your enemy. That's something it shares with chess, Go, poker, Dota 2 and just about every other game. In games, you can win.

From an algorithm's perspective, problems need to have an "objective function," a goal to be sought. When AlphaZero played chess, this wasn't so hard. A loss counted as minus one, a draw was zero, and a win was plus one. AlphaZero's objective function was to maximize its score. The objective function of a poker bot is just as simple: Win lots of money.



[DeepMind](#)

Computer-generated humanoid walkers can be trained to perform complex behaviors, like walking through unfamiliar terrain.

Real-life situations are not so straightforward. For example, a self-driving car needs a more nuanced objective function, something akin to the kind of careful phrasing you'd use to explain a wish to a genie. For example: Promptly deliver your passenger to the correct location, obeying all laws and appropriately weighing the value of human life in dangerous and uncertain situations. How researchers craft the objective function, Domingos said, "is one of the things that distinguishes a great machine-learning researcher from an average one."

Consider Tay, a Twitter chatbot released by Microsoft on March 23, 2016. Tay's objective was to engage people, and it did. "What unfortunately Tay discovered," Domingos said, "is that the best way to maximize engagement is to spew out racist insults." It was [snatched back offline](#) less than a day later.

## Your Own Worst Enemy

Some things don't change. The methods used by today's dominant game bots employ strategies devised decades ago. "It's almost a blast from the past, with just more computation being thrown at it," said [David Duvenaud](#), a computer scientist at the University of Toronto.

The strategies often rely on reinforcement learning, a hands-off technique. Instead of

micromanaging an algorithm with detailed instructions, engineers let the machine explore an environment and learn to meet goals through trial and error. Before the release of AlphaGo and its progeny, the DeepMind team achieved its first big, headline-grabbing result in 2013, when they used reinforcement learning to make a bot that [learned to play](#) seven Atari 2600 games, three of them at an expert level.

That progress has continued. On February 5, DeepMind released [IMPALA](#), an AI system that can learn 57 Atari 2600 games, plus 30 more levels built by DeepMind in three dimensions. In these, the player roams through different environments, accomplishing goals like unlocking doors or harvesting mushrooms. IMPALA seems to transfer knowledge between tasks, meaning time spent playing one game also helps it improve at others.

But within the larger category of reinforcement learning, board games and multiplayer games allow for an even more specific approach. Here, exploration can take the form of self-play, where an algorithm gains strategic supremacy by repeatedly wrestling with its own close copy.

This idea dates back decades. In the 1950s, IBM engineer Arthur Samuel [created](#) a checkers-playing program that learned in part by matching an alpha side against a beta side. And in the 1990s, Gerald Tesauro, also from IBM, built a backgammon program that pitted the algorithm against itself. The program reached human expert levels, devising unorthodox but effective strategies along the way.

In game after game, an algorithm in a self-play system faces an equally matched foe. This means that changes in strategy lead to different outcomes, giving the algorithm immediate feedback. “Anytime you learn something, anytime you discover a small thing, your opponent immediately uses it against you,” said [Ilya Sutskever](#), the research director at OpenAI, a nonprofit he co-founded with Elon Musk that is devoted to developing and sharing AI technology and shepherding it toward safe applications. In August 2017, the organization [released a Dota 2 bot](#) controlling the character Shadow Fiend — a sort of demon-necromancer — that beat the world’s best players in one-on-one battles. Another OpenAI project pits simulated humans against one another in a sumo match, where they end up [teaching themselves](#) how to tackle and feint. During self-play, “you can never rest, you must always improve,” Sutskever said.

But the old idea of self-play is just one ingredient in today’s dominant bots, which also need a way to translate their play experiences into deeper understanding. Chess, Go and video games like Dota 2 have many more permutations than there are atoms in the universe. Even over the course of many lifetimes spent battling its own shadow across echoless virtual arenas, a machine can’t encounter every scenario, write it down in a look-up table, and consult that table when it sees the same situation again.

To stay afloat in this sea of possibilities, “you need to generalize, capture the essence,” said [Pieter Abbeel](#), a computer scientist at the University of California, Berkeley. IBM’s Deep Blue did this with its built-in chess formula. Armed with the ability to gauge the strength of board positions it hadn’t seen before, it could adopt moves and strategies that would increase its chances of winning. In recent years, though, a new technique has made it possible to skip the formula altogether. “Now, all of a sudden, the ‘deep net’ just captures all of that,” Abbeel said.

Deep neural networks, which have soared in popularity in the last few years, are [built out of layers of artificial “neurons”](#) that stack like pancakes. When neurons in one layer fire, they send signals to the next layer up, which sends them to the next layer, and so on.

By tweaking how the layers connect, these networks become fantastic at morphing inputs into a related output, even if the connection seems abstract. Give them a phrase in English, and they could



train themselves to translate it into Turkish. Give them pictures from an animal shelter and they can identify which ones contain cats. Or show them a game board, and they can grok what their probability of winning is. Typically, though, you need to first give these networks reams of labeled examples to practice on.

That's why self-play and deep neural networks fit together so well. Self-play churns out troves of games, giving deep neural networks a theoretically unlimited supply of the data they need to teach themselves. In turn, the deep neural networks offer a way to internalize the experiences and patterns encountered in self-play.

But there's a catch. For self-play systems to produce helpful data, they need a realistic place to play in.

"All these games, all of these results, have been in settings where you can perfectly simulate the world," said [Chelsea Finn](#), a Berkeley doctoral student who uses AI to control robot arms and interpret data from sensors. Other domains are not so easy to mock up.

Self-driving cars, for example, have a hard time dealing with bad weather, or cyclists. Or they might not capture bizarre possibilities that turn up in real data, like a bird that happens to fly directly toward the car's camera. For robot arms, Finn said, initial simulations provide basic physics, allowing the arm to at least learn how to learn. But they fail to capture the details involved in touching surfaces, which means that tasks like screwing on a bottle cap — or conducting an intricate surgical procedure — require real-world experience, too.

For problems that are hard to simulate, then, self-play is not so useful. "There is a huge difference between a true perfect model of the environment and a learned estimated one, especially when that reality is complex," wrote [Yoshua Bengio](#), a pioneer of deep learning at the University of Montreal, in an email. But that still leaves AI researchers with ways to move forward.

## Life After Games

It's hard to pinpoint the dawn of AI gaming supremacy. You could choose Kasparov's loss in chess, or Lee Sedol's defeat at the virtual hands of AlphaGo. Another popular option would be when legendary *Jeopardy!* champion Ken Jennings lost to IBM's Watson in 2011. Watson could parse the game's clues and handle wordplay. The two-day match wasn't close. "I for one welcome our new computer overlords," Jennings wrote under his final answer.

Watson seemed to be endowed with the kind of clerical skills humans use on a host of real-world problems. It could take a prompt in English, rummage through relevant documents at lightning speed, come up with the relevant snippets of information, and settle on a single best answer. But seven years later, the real world continues to present stubborn challenges for AI. A [September report by the health publication Stat](#) found that researching and designing personalized cancer treatments, as Watson's descendant Watson for Oncology aims to do, is proving difficult.

"The questions in *Jeopardy!* are easier in the sense that they don't need much common sense," wrote Bengio, who has collaborated with the Watson team, when asked to compare the two cases from the AI perspective. "Understanding a medical article is much harder. Again, much basic research is needed."

As special as games are, there are still a few real-world problems they resemble. Researchers from DeepMind declined to be interviewed for this article, citing the fact that their AlphaZero work is currently under peer review. But the team has suggested that its techniques may soon help

biomedical researchers, who would like to understand protein folding.

To do this, they need to figure out how the various amino acids that make up a protein [kink and fold](#) into a little three-dimensional machine with a function that depends on its shape. That's tricky in the same ways chess is tricky: Chemists know the rules roughly well enough to calculate specific scenarios, but there are still so many possible configurations, it's a hopeless task to search through them all. But what if protein folding could be configured as a game? In fact, it already has been. Since 2008, hundreds of thousands of human players have attempted [Foldit](#), an online game where users are scored on the stability and feasibility of the protein structures they fold. A machine could train itself in a similar manner, perhaps by trying to beat its previous best score with general reinforcement learning.

Reinforcement learning and self-play might also help train dialogue systems, Sutskever suggests. That would give robots meant to speak to humans a chance to train by talking to themselves. And considering that specialized AI hardware is becoming faster and more available, engineers will have an incentive to pose more and more problems in the form of games. "I think that in the future, self-play and other ways of consuming a very large amount of computing power will become more and more important," Sutskever said.

But if the ultimate goal is for machines to do as much as humans can, even a self-taught, generalist board-game champ like AlphaZero may have a way to go. "You have to see, to my mind at least, what's really a huge gulf between the real activities of thinking, creative exploration of ideas, and what we currently see in AI," said [Josh Tenenbaum](#), a cognitive scientist at the Massachusetts Institute of Technology. "That kind of intelligence is there, but it's mostly going on in the minds of the great AI researchers."

Many other researchers, conscious of the hype that surrounds their field, offer their own qualifiers. "I would be careful not to overestimate the significance of playing these games, for AI or jobs in general. Humans are not very good at games," said François Chollet, a deep-learning researcher at Google.

"But keep in mind that very simple, specialized tools can actually achieve a lot," he said.

*Clarification February 22, 2018: An earlier version of this article implied that chess strategy was programmed into the Deep Blue computer. In fact, engineers programmed in a framework for chess strategy, and the machine analyzed many human games to arrive at its own particular strategy. The article has been changed to avoid the unwanted implication.*

*This article was reprinted on [TheAtlantic.com](#).*