Move over AlphaGo: AlphaZero taught itself to play three different games

Enlarge / Starting from random play and knowing just the game rules, AlphaZero defeated a world champion program in the games of Go, chess, and shogi (Japanese chess).
DeepMind Technologies, Ltd.

reader comments

43 with 30 posters participating

Share this story

Google's DeepMind—the group that brought you the champion game-playing AIs AlphaGo and AlphaGoZero—is back with a new, improved, and more-generalized version. Dubbed AlphaZero, this program taught itself to play three different board games (chess, *Go*, and shogi, a Japanese form of *chess*) in just three days, with no human intervention.

A paper describing the achievement was just published in *Science*. "Starting from totally random play, AlphaZero gradually learns what good play looks like and forms its own evaluations about the game," said Demis Hassabis, CEO and co-founder of DeepMind. "In that sense, it is free from the constraints of the way humans think about the game."

Chess has long been an ideal testing ground for game-playing computers and the development of AI. The very first chess computer program was written in the 1950s at Los Alamos National Laboratory, and in the late 1960s, Richard D. Greenblatt's Mac Hack IV program was the first to play in a

human chess tournament—and to win against a human in tournament play. Many other computer chess programs followed, each a little better than the last, until IBM's Deep Blue computer defeated chess grand master Garry Kasparov in May 1997.

As Kasparov points out in an accompanying editorial in *Science,* these days your average smartphone chess playing app is far more powerful than Deep Blue. So AI researchers turned their attention in recent years to creating programs that can master the game of Go, a hugely popular board game in East Asia that dates back more than 2,500 years. It's a surprisingly complicated game, much more difficult than chess, despite only involving two players with a fairly simple set of ground rules. That makes it an ideal testing ground for AI.

Enlarge / Starting from random play, knowing just the basic rules, AlphaZero defeated a world champion program in the games of Go, chess, and shogi.
DeepMind Technologies, Ltd.

AlphaZero is a direct descendent of DeepMind's AlphaGo, which made headlines worldwide in 2016 by defeating Lee Sedol, the reigning (human) world champion in Go. Not content to rest on

its laurels, AlphaGo got a major upgrade last year, becoming capable of teaching itself winning strategies with no need for human intervention. By playing itself over and over again, AlphaGo Zero (AGZ) trained itself to play Go from scratch in just three days and soundly defeated the original AlphaGo 100 games to 0. The only input it received was the basic rules of the game.

The secret ingredient: "reinforcement learning," in which playing itself for millions of games allows the program to learn from experience. This works because AGZ is rewarded for the most useful actions (i.e., devising winning strategies). The AI does this by considering the most probable next moves and calculating the probability of winning for each of them. AGZ could do this in 0.4 seconds using just one network. (The original AlphaGo used two separate neural networks: one determined next moves, while the other calculated the probabilities.) AGZ only needed to play 4.9 million games to master Go, compared to 30 million games for its predecessor.

> *"Instead of processing human instructions and knowledge at tremendous speed, AlphaZero generates its own knowledge."*

AGZ was designed specifically to play Go. AlphaZero generalizes this reinforced-learning approach to three different games: Go, chess, and shogi, a Japanese version of chess. According to an accompanying perspective penned by Deep Blue team member Murray Campbell, this latest version combines deep reinforcement learning (many layers of neural networks) with a general-purpose Monte Carlo tree search method.

"AlphaZero learned to play each of the three board games very quickly by applying a large amount of processing power, 5,000 tensor processing units (TPUs), equivalent to a very large supercomputer," Campbell wrote.

"Instead of processing human instructions and knowledge at tremendous speed, as all previous chess machines, AlphaZero

generates its own knowledge," said Kasparov. "It does this in just a few hours, and its results have surpassed any known human or machine." Hassabis, who has long been passionate about chess, says that the program has also developed its own new dynamic style of play—a style Kasparov sees as much like his own.

There are some caveats. Like its immediate predecessor, AlphaZero's basic algorithm really only works for problems where there are a countable number of actions one can take. It also requires a strong model of its environment, i.e., the rules of the game. In other words, Go is not the real world: it's a simplified, highly constrained version of the world, which means it is far more predictable.

Enlarge / AlphaZero searches only a small fraction of the positions considered by traditional chess engines.
DeepMind Technologies, Ltd.

"[AlphaZero] is not going to put chess coaches out of business just yet," Kasparov writes. "But the knowledge it generates is information we can all learn from." David Silver, lead researcher of the AlphaZero project, has high hopes for future applications of that knowledge. "My dream is to see the same kind of system applied not just to board games but to all

kinds of real-world applications, [such as] drug design, material design, or biotech," he said.

Poker is one contender for future AIs to beat. It's essentially a game of partial information—a challenge for any existing AI. As Campbell notes, there have been some programs capable of mastering heads-up, no-limit Texas Hold 'Em, when only two players are left in a tournament. But most poker games involve eight to 10 players per table. An even bigger challenge would be multi-player video games, such as *Starcraft II* or *Dota 2*. "They are partially observable and have very large state spaces and action sets, creating problems for Alpha-Zero like reinforcement learning approaches," he writes.

One thing seems clear: chess and Go are no longer the gold standard for testing the capabilities of AIs. "This work has, in effect, closed a multi-decade chapter in AI research," Campbell writes. "AI researchers need to look to a new generation of games to provide the next set of challenges."

DOI: *Science*, 2018. 10.1126/science.aar6404 (About DOIs).

https://arstechnica.com/?p=1422731