Review of “Mastering the game of Go with deep neural networks and tree search”
By Brennon York
On 2017/02/19

A brief summary of the paper's goals or techniques introduced (if any).

AlphaGo sets out to develop an artificial intelligence capable of defeating any human player in the game of Go. The game of Go has previously been thought of as the hardest game to develop an artificial intelligence for given its b^d numbers being incomprehensible (b ~= 250, d ~= 150). Chess, previously beaten by Deep Blue, has a b^d where b ~= 35 and d ~= 80 as a comparison. In this paper the authors introduce two main new techniques to the field, specifically approximate value comprehension and sampling through a policy of probability distributions over the set of possible moves. They former is a form of tree search pruning that removes vast subsections of trees based on predictions at that current state of the game. The latter are a set of policies trained as convolutional neural networks in supervised and reinforced fashion. The first and second stage of the pipeline (the supervised and reinforcement policy networks) are tied together with the position evaluation through a Monte Carlo tree search (MCTS) algorithm. Even before the tied the first, second, and third stage together, they achieved an 85% win rate against the best open source program, Pachi, where the previous best was only 11%. They trained the first stage, the supervised learner, as a 13-layer network on 30 million positions from the KGS Go server. The second stage, the reinforcement learner, was trained by playing against the current policy network and randomly selected previous iterations of the network. They then played the second stage against the first, with a second stage success rate of 80%. For the final stage they trained it on a generated set of 30 million random game points rather than the KGS dataset as, when they did that, the system incorrectly memorized specific game outcomes that couldn’t generalize (overfitting). Tying the system together gave the authors a 99.8% success rate when playing a tournament against all other known open source and commercial Go implementations. They went further to build a distributed version of the system which was then able to achieve 100% success against any other computer player.

A brief summary of the paper's results (if any).

AlphaGo has, for the first time, developed an artificial intelligence capable of besting the top human players in the classic game of Go. Not only is this considered one of artificial intelligence’s “grand challenges” they furthered the general field by fundamentally changing how an artificial intelligence evaluates its state on a game board or otherwise. They demonstrated that, in “thousands of times fewer positions” than Deep Blue evaluated, they were able to show a 100% success rate against all other forms of computer-based Go players and a 5-0 official match against the top Go human player, Fan Hui. Further they were able to create a novel evaluation policy that, rather than search more, search more intelligently based on previous games played by experts and evaluating more effectively with a value network thought to be much closer to what humans naturally do. AlphaGo demonstrates a leap in artificial intelligence research as it shows how policy networks trained on previous gameplay and deep convolutional neural networks can overcome the traditional evaluation methods around pure tree search.