UDACITY - ARTIFICIAL INTELLIGENCE NANODEGREE

Adversarial Game Playing Project

by George Pampalis, January 2018

Research Review

"Mastering the game of Go with deep neural networks and tree search"

Silver *et al.* in their paper of 2016 explain their approach of creating an intelligent program (AlphaGo) that is able to compete against the best human players in the game of Go. Specifically, they combined two different approaches (deep neural networks and tree search) in order to achieve an expert game-play level.

Even though Go is a game with a great depth and breadth (each an order of magnitude larger than chess), the authors were able to consistently beat the European champion by using supervised and reinforcement learning to achieve a policy leading to winning games, as well as reinforcement learning to predict the value of each position without the need of branching down to the leaf nodes of the tree.

Firstly, the authors trained a policy network to predict expert moves using a 13-layer convolutional network. Stochastic gradient ascent was applied to maximize the probabilities of selecting a good action a for a given state s. This step resulted in a 55% accuracy of predicting expert moves.

Next, the policy network was improved upon by applying policy gradient reinforcement learning. The new policy was initialized with the previous one and was trained using random previous iterations of itself, in order to prevent overfitting, and the reward values for each terminal time step is updated using stochastic gradient ascent. As a result, the reinforced policy won 80% of the games against the expert policy network of the previous step.

Finally, the value of each position was predicted using a similar reinforced learning network, combined with stochastic gradient descent in order to find the minimum difference between the mean-squared error and the predicted value, as well as a data-set of self-played games with millions of uncorrelated distinct positions to prevent overfitting during the training process. This step resulted in similar accuracy with the Monte Carlo rollouts method of other state-of-the-art approaches, but using 1/15,000 of the computational power.

By combining all of the above steps, AlphaGo was able to defeat almost always all of the existing Go programs, while also showing great results even with handicapped opponents. However, its most astonishing feat was defeating the European champion Fan Hui in 2015, during the matches of which thousands of fewer positions were evaluated when compared to Deep Blue and its 1997 match against Gary Kasparov.