

Mastering the game of Go with Deep Neural networks and tree search

Summary

AlphaGo was developed by Deepmind to improve the efficiency of tackling the board game Go using deep neural networks and tree search. The search algorithm used Monte Carlo with value and policy networks. Due to the size of the search space being so complex, it appeared near infeasible.

Deepmind introduced the use of value functions based on truncated the Monte Carlo search algorithm which terminate rollouts before the end of the game and use a value function in place of the terminal reward. AlphaGo also evaluates full rollouts with truncated rollouts. They challenged the problem of position evaluation using value networks. To control the search strategy, the algorithm preferred actions with high action value using the value network. To increase computing time, they also exploited the use of symmetry of the board

The team first employed a supervised learning technique, to build upon existing research using labelled data from the KGS Go server. This network was only able to achieve the potential of a strong amateur. The team then applied reinforcement learning techniques to learn from self-play instances and achieved a more advanced game-playing agent.

Network Architecture

The network consisted of a 13 layer deep policy network using an input image size of 23x23 with 2 to 12 hidden layers, a 3x3 kernel size with a stride of 1 and using a rectifier nonlinearity and a fully connected layers with an output layer using single tanh. Using this network, they were able to identify actions with high action value (best move).

The authors introduced a rollout policy to determine the probability of how the program should search through the tree branches. While this increase the speed of the search, it was not as accurate. It was used as a look ahead approach to predict the most likely outcome of the next move.

Results

As a result of this research, a human professional Go player was defeated by a computer program for the first time in history. AlphaGo achieved a 99.8% winning rate against other Go programs and won 5 to 0 against the human Go European champion. This achievement which was once thought infeasible and one of the challenging decision-making tasks for a computer program.