Research Review on *AlphaGo* by the DeepMind Team

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

The game of Go is a classic game which has two adversarial players taking turns to place a piece on a grid of 19 x 19 lines and the player who surrounds more territory wins. It is considered as the most challenging game among the other classic games such as chess and checkers because the average depth of the search tree is about 150 and the average branching factor is about 250, and therefore, the exhaustive search with size of 150^{250} is nowhere close to possible.

However, two general principles could be used to reduce the search size. The first principle is to reduce the depth of the search tree by position evaluation where a good static evaluation on a node could replace the value backed up by its subtree. The second principle is to reduce the average breadth of the tree by a policy where only part of the moves is selected to expand further.

The DeepMind team at Google builds the AlphaGo program which implements the idea of the above-mentioned principles. They first train a supervised learning (SL) policy network using deep convolutional networks where multiple layers are used based on different features. They used the 30 million positions from the KGS Go server as the training data to generate the policy network which it can produce the probability distribution of moves on any state. Reinforcement learning (RL) is then used to improve the performance of the SL policy network. The last stage is to train pipeline to get the value evaluation of positions by running simulations of two agents using the RL policy networks playing against each other.

AlphaGo then combines the policy and value networks together with the Monte Carlo tree search (MCTS) algorithm which selects best actions by searching ahead.

The final version of standalone AlphaGo uses 40 search threads, 48 CPU, and 8 GPUs, while a distributed version employs multiple severs with 40 search threads, 1202 CPUs, and 176 GPUs.

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

AlphaGo is the first computer program that ever defeated a human professional player in a full-size game of Go and achieved 99.8% winning percentage against other Go programs. It evaluates thousands of times of fewer states than Deep Blue did in the chess match against Kasparov by selecting more intelligent moves based on the policy network, and evaluate more accurately on the moves using the value network.