Research Review of DeepMind's AlphaGo Paper

Why Go is a challenge?

The ancient game of Go has long been viewed as the most challenging of classic games for artificial intelligence owing to its enormous search space and the difficulty of evaluating board positions and moves. The number of possible states is estimated to be 250^{150} which is roughly equal to 10^{761}. Because these it is prohibitively difficult to use traditional AI methods such as alpha-beta pruning, tree traversal and heuristic search.

New techniques

The previously strongest machine players at Go were based on Monte Carlo Tree Search (MCTS) techniques, which basically try to reduce the complexity of the search tree via sampling/making simulations.

AlphaGo's algorithm uses MCTS guided by two deep neural network technologies `Policy Network` and `Value Network`.

The `Policy Network` provides the probability distribution of moves. It learns to predict the most likely moves to be played. This is first trained by Supervised Learning using 30 million positions of data from KGS Go Server. To improve this further, a policy gradient reinforcement learning techniques are used later.

The `Value Network` is used for evaluating moves during the game play, using MCTS. It is trained from games for which the outcome is already known, by using the strongest policy they have trained.

Evaluation

By introducing these new techniques, AlphaGoo achieved a 99.8% winning rate against other Go programs, and defeated the human European Go champion even though other previous Go programs were amateur level.