Research Review
– Mastering the game of Go with deep neural networks and tree search

The article introduces a new approach to develop computer GO that involving neural networks "value networks" and "policy networks" as well as a new search algorithm that combines Monte Carlo simulation with value and policy networks. By applying these new upgrades, AlphaGO achieved a 99.8% wining rate against other GO programs and defeated the human European Go champion by 5:0.

The neural networks have been used to reduce the effective depth and breadth of the search tree by evaluating positions using a value network and sampling actions using a policy network. A brief summary of the specific steps as following:

1. Train a supervised learning (SL) policy network directly from expert human moves, in order to provide fast, efficient learning updates with immediate feedback and high-quality gradients.
2. Improve the policy network by training reinforcement learning (RL) and optimizing the outcome of games of self-play.
3. Train a value network that predicts the winner of games played by the RL policy network against itself.

Monte Carlo tree search (MCTS) is an algorithm which uses Monte Carlo rollouts to estimate the value of each state in a search tree. AlphaGo's new searching algorithm combines the policy and value networks in a MCTS algorithm that selects actions by look-ahead search. Because that evaluating policy and value networks requires several orders of magnitude more computation than traditional search heuristics, AlphaGo uses an asynchronous multi-threaded search that executes simulations on CPUs, and computes policy and value networks in parallel on GPUs.

To evaluate AlphaGO, an internal tournament wans run among AlphaGO and several other GO programs that also based on high-performance MCTS algorithms. The results of the trounament suggest that single-machine AlphaGo is many dan ranks stronger than any previous Go program, winning 494 out of 495 games (99.8%) against other Go programs. Moreover, the distributed version AlphaGo was evaluated against Fan Hui, a professional 2 dan, and winner of the 2013 – 2015 European Go championships. AlphaGo won the match 5 games to 0.