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

Goals

To win the GO game or the other GO program, this research introduces a new approach that uses “value network” and “policy network” to play the game. The value network was trained by reinforcement learning to evaluate board position; the policy network was trained by supervise learning by expert moves and reinforcement learning by self-play to choose the best move. Also, while previous research only used shallow policies or value function to reduce the calculation, this research use Monte Carlo tree search to prune the search tree and keep the high probability actions only.

Results

The trained program AlphaGo achieved a 99.8% winning rate while playing with other GO program, and also defeated the human European GO champion. This is the first time that a computer program defeated a human expert GO player.