Summary of the Paper: Mastering the game of Go with deep neural networks and tree search

The Paper's Goals

The paper aims to create an artificial go player named AlphaGo. It utilized the combination of supervised learning of policy networks, reinforcement learning of policy networks and reinforcement learning of value networks as a whole, to evaluate the playing strength of AlphaGo. With adjusted regulations of playing, this paper executed the program on real world games, hoping to achieve good winning records.

The Paper's Techniques

The main techniques of this paper can be divided into four parts: supervised learning of policy networks, reinforcement learning of policy networks, reinforcement learning of value networks, and evaluating the playing strength of AlphaGo.

For supervised learning of policy networks, this paper utilized SL policy network between convolutional layers with weights sigma, and rectifier nonlinearities. It also used softmax layer as probability distribution outputs.

For reinforcement learning of policy networks, this paper utilized RL policy network to train this system, and try to figure out which point is possibly the best at given status. It mainly executed tremendous simulations with Monte Carlo search program, and chose the best policy according to MSE comparison.

For reinforcement learning of value networks, it figures out a perfect value function estimating outcome prediction and game position.

After establishing these three techniques, it combines them and search the optimal position and strategy with policy and value networks.

The Paper's Result

By combining tree search with policy and value network, AlphaGo has finally reached a professional level in Go, and wined almost every Go game online or in real competitions.