

Summary: AlphaGo with deep neural networks and tree search

The game of Go is very challenging for artificial intelligence because of the very large search space. Go has a number of legal moves (breath (b)) of about 250 per turn and its game length (depth (d)) is about 150. This gives b^d sequences of moves and make searching the whole game tree unfeasible. The second difficulty is the correct and fast evaluation of the board positions and the best moves to choose.

This new approach builds on existing knowledge but expands the functionality of the Go playing agent through the use of value networks combined with Monte Carlo Tree Search.

A neural network pipeline is used to train a value network. This is done by using a mixture of supervised learning and reinforcement learning on human expert moves and self play.

The first stage of the training pipeline is built on prior work on predicting expert moves in the game of Go using supervised learning. The second stage of the training pipeline aims at improving the policy network by policy gradient reinforcement learning.

AlphaGo combines the policy and value networks in an MCTS algorithm that selects actions by lookahead search.

This combination of different techniques combined together gives AlphaGo a significant advantage over existing Go-playing programs.

The strength of AlphaGo is validated through tournaments against other Go-playing programs. The effect of using more CPUs and GPUs on the agent's performance is tested and documented in detail. As expected more computational power gives better performance of the agent. The improvements of AlphaGo are enormous and it was possible to test the power of AlphaGo in playing against a professional human Go player. AlphaGo could win the tournament against a human player. This milestone in the advance of AI was expected to take place a decade later.

This gives a bright outlook on the future of AI as the results provide hope that human-level performance can now be achieved in other seemingly intractable artificial intelligence domains.