

Research review

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

Summary

The paper “Mastering the game of Go with deep neural networks and tree search” introduces a new approach to use computer to play Go. Go is a classical game of artificial intelligence due to the very large space to evaluate positions and moves. The new approach is based on deep convolutional neural networks, the network are trained with supervised learning from human expert and reinforcement learning from games of self play. The goal achieved by the paper is a new search algorithm that combine neural network evaluations and Monte Carlo simulation. This search algorithm defeated the human European Go champions by 5 games to 0.

The Go game has a large search space where and exhaustive search of the optimal value function for the next position it is infeasible. In general the search space can be reduced by two principles: position evaluation and sampling actions from a policy. Both principles produces only weak amateur level play Go.

In order to reduce the search space in the paper is presented the use of a deep convolutional neural networks. The depth (game length) and breadth (number of legal moves per position) of the search tree are reduced using a value network, and sampling actions using a policy network.

The training of the neural network is did in two phases. First is a supervised learning (SL) of policy network, directly from expert human moves. Second is a reinforcement learning (RL) improving the SL by optimising the final result of games of self-play.

AlphaGo combine the policy and value networks in a Monte Carlo search tree (MCTS) algorithm that selects the next move to do. To efficiently exploit MCTS with deep neural network the AlphaGo uses CPUs and GPUs. The simulation are executed in an asynchronous multi-thread search on CPUs. The policy and value networks are computed in parallel GPUs. The final version of AlphaGo uses 40 search threads, 48 CPUs and 8 GPUs.

Result

The result of the paper is a Go program that plays at the level of the strongest human player. The program is based on neural network trained from gameplay through general-purpose supervised and reinforcement learning method avoiding the construction of handcrafted evaluation function as for instance Deep Blue.

AlphaGo has reached a professional level in Go providing a solution of an intractable search space problem. AlphaGo exploits a new search algorithm that combines neural network evaluations with Monte Carlo rollouts producing a high performance tree search engine. This search algorithm could be applied successfully with advances in several other domains; for example, general game-playing, classical planning, scheduling and constraint satisfactions.