

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

Goals:

The go game, unlike other game, has a very large search space so that the normal search algorithm and prune method doesn't work well. This paper gives us a brief introduction of how AlphaGo works and provides the algorithm to tackle with this situation.

Techniques:

AlphaGo use the Monte Carlo tree search combined with supervise/reinforcement learning techniques to deal with the search tree spaces. The traditional Minimax algorithm with alpha-beta pruning can't deal with the breadth and depth in go, while the Monte Carlo method can return a score about how well the game looks like by sampling and aim at the best result. The machine learning algorithms are used to train the policy network and value network which are two networks. Policy network is used for playing piece and value network is used for evaluating the situations. These two networks were combined into the Monte Carlo tree search as heuristics. Policy network is trained by with chess manual with supervised learning techniques and refined by reinforcement learning algorithm. Then another reinforcement learning algorithm is applied to build the value network. After that, the whole pipeline was built.

Result:

The whole game agent performs very well in go game and beat the world championship. Also, the calculating time through the game is far less than the deep blue versus Kasparov. Besides the game playing, AlphaGo gives us the methods and confident to deal with large search space which is a great evolution in AI field.