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

Structure summery

The paper mainly introduced a new approach to conquer the most challenge classic game for artificial intelligence, Go. In the past, Monte Carlo tree search was the most popular strategy for building game agent. Due to its linear combination of input feature and limited evaluation function, the most powerful Go play agent can only achieve strong amateur level. A new search algorithm is introduced in this paper which combines new algorithm with Monte Carlo tree search to build a policy neural network, the following content briefly explains how it works.

Researchers in this project, basically borrowed the principle of Monte Carlo tree search and modified evaluation stage using deep neural network.

During rollout stage, firstly, a supervised learning policy network (SL) has been trained first which processes input directly from expert players and provide fast update and high-quality gradient. Such network mainly switches between convolutional layer and rectifier nonlinearities, and the soft-max function at last provide a probability distribution. To maximize the prediction, stochastic gradient ascent has been utilized as well. In the meantime, a fast rollout policy has been built to predict human player's moves as well.

To further improve the performance, a reinforcement learning policy network (RL) has been implemented which inherit the weight from SL and play against random selected nodes from previous level, then use reward function to find the best next move.

In the last stage, the value network is to provide a value according to the position state of the game. The utilization of reinforcement learning network is result from the strong correlation of successive positions.

Result

The result from tournament of AlphaGo is outstanding. With 494 wins out of 495 games against other agents. With handicap, AlphaGo can win 77%, 86%, 99% against Crazy Stone, Zen and Pachi. All those results reveals the power of combination of deep neural networks and Monte Carlo tree search.