Research Review

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

The paper presents the method that AlphaGo system used to beat human play in Go. As we learned in the lectures, to effectively build a game-playing agent, we need build efficient search algorithm and evaluation algorithm.

One of the main challenge to build an agent playing Go is the huge search space. The breadth of Go is about 250 with depth is about 150, much larger than the game like chess. Several approaches were proposed to further reduce the search space. And the most successful method is called Monte Carlo tree search. The main idea of MCTS is expanding the search tree use random sampling of the search space. By random sampled child node, the search algorithm will play the game to the end. And based on the state of the end game., the agent will update the information in the nodes. By repeating doing this, a limited search network will build and the agents will choose the node with maximal visit counts.

To further boost performance of the search, the author builds the deep neural network to guide the sampling. The board of the game will be converted to the matrix of features and then input into the neural networks. Two networks, policy network and value network, were trained to provide the evaluation of the position and the board status.

To train the model, the author first use 30 million positions from online Go server as the target to finish a supervised learning. After that, the AlphaGo use reinforcement learning to play against the variation of itself. The two steps training finally make AlphaGo can compete with human professional players.

Compared to the previous AI agent of Go, AlphaGo is much more powerful. AlphaGo wins 494 out of 495 games against other AI agents. Even giving the opponent free move, AlphaGo still wins most of the games. Using Elo rating, we can quantitively compare the AI agent. And AlphaGo is about 500 to 1000 stronger than previous strongest AI (equal to 2 to 4 dan). The researcher also compare the different combination of the techniques that AlphaGo uses to show the importance of the different methods.

I believed such strategies can be used for other complex games and to build a successful agent we need always focus on the search algorithm and evaluation algorithm.