# Part 2: Research Review

Name : Kwangshin Oh
Udacity Email Address : kwangshin@gmail.com

Subject : AlphaGo by the DeepMind Team.
Title : Mastering the game of Go with deep neural networks and tree search

Along with the development of Artificial Intelligence, the game is the first and best target to evaluate the output of Artificial Intelligence research and product. The simple game is easy to build a rule and provide the definite result : win or lose. So it was not a big challenge to win the simple game with small game board. The computer software quickly calculate all the possible scenarios using best algorithm, so it becomes the untouchable area by human being. However the traditional Go game looks like the impregnable fortress to Artificial Intelligence due to the complexity of game and the difficulty of evaluating the possible moving positions to decide next position within limited time periods.

To win a board game, there is no secret strategy. As a human being does it, just choose the next position having the higher probability of winning. It is the way that the Artificial Intelligence wins the simple board game against human being : Go all the possible positions and choose the next position having the winning route. However if the time is limited, then give the score for each possible positions based on the winning probability after visiting as much as places within given time period.

Researchers have been trying to improve the evaluation algorithm of current board position and the selecting algorithm of next moving position. However the traditional Go game looks like the final stage to be conquered by Artificial Intelligence. Along with the advanced Machine Learning algorithm, the AlphaGo team introduces new approach and apply those idea as new search algorithm after trained by supervised learning and reinforcement learning.

From this paper, the AlphaGo team introduces 2 new approaches: value networks and policy networks. The value networks will be used to evaluate current board positions, and the policy networks will be used to select next moves. Those networks will be built and trained by an innovative combination of supervised learning with data from games done by human expert and reinforcement learning with data from games done by self-play. As a final step, both value networks and policy networks would be combined into most notable Monte Carlo Tree Search (known as MCTS) heuristic search algorithm.

As a result of introducing combined search algorithm with value networks as well as policy networks, the AlphaGo program achieved marvelous 99.8% winning rate against other Go programs. Moreover, it is the surprising news that the AlphaGo won the battle with the human European Go champion with unbelievable game score 5 versus 0. It is a remarkable first step that a computer program wins a victory in the full-sized Go game against human professional player.