

Research Review By Ng Fang Kiang

Udacity – Artificial Intelligence Nanodegree (Class of 2017)

Paper :AlphaGo (<https://goo.gl/U5ypHp>)

The purpose of this research was to show how artificial intelligence mastering the game of Go with deep neural networks and tree search exceed human expertise. AlphaGo was invented by DeepMind team for solving the most challenging of classic games, called Go! Comparing the 9x9 Chess board, the game of Go was very much complex than chess in the complexity of games. The games of Go was created with 19x19 grid board with 10360 of game tree complexity, while chess on a 9x9 board has a game-tree complexity of 10123. Due to the sheer number of moves, the traditional search tree methods over all possible position did not have a chance in Go. Therefore, AlphaGo combines advanced tree search with deep neural networks which called Monte Carlo tree search(MCTS).

AlphaGo Paper Techniques introduced

- Using ‘value networks’ to evaluate board positions and ‘policy networks’ to select moves.
- A new search algorithm that combines Monte Carlo simulation with value and policy networks.(Achieved a 99.8% winning rate against other Go programs, and defeated the human European Go Champion by 5 games to 0)

The AlphaGo team have combines supervised learning and reinforcement learning to train their neural networks. First, the DeepMind team train their neural networks with the supervised learning of policy networks for predicting opponent moves in games in a data set of positions. Then, the team improves AlphaGo win rate with implement reinforcement learning of policy networks that rely on the positions from the self-play data set. Self-play data was played an important role reinforcement learning algorithms, it helps to train the AlphaGo getting better from its past experiences. The process is keep repeated rewarding winning games and punishment for losing game. To evaluate on position, they using reinforcement learning of value networks helps them to optimal under perfect play.

AlphaGo has achieved the professional level and won the best human player of Go, Sedol Lee with 4:1 result in March 2016. After that, AlphaGo won against another best human player from over the world with 60:0 result on 5 January 2017. Continuously, AlphaGo won the No.1 ranked human player in the world, Kajie with 3:0 result in May 2017. As the result show, the mixture of the output of the value network and the result of a self-play simulation of the fast policy network did a pretty good job. AlphaGo has finally achieved the human-professional level.