

Research Review: AlphaGo

For my review I have chosen paper by Deepmind - "[Mastering the game of Go with deep neural networks and tree search](#)"

The game of Go has long been viewed as the most challenging of classic games for artificial intelligence because of its enormous search space and the difficulty of evaluating board positions and moves. DeepMind's AlphaGo made waves when it became the first AI to beat a top human Go player in March of 2016.

As a zero-sum game of perfect information, the game of Go has an optimal value associated to every board position and this optimal value can be solved by traversing a search tree containing approximately b^d sequence of moves where b is the game's breadth (number of legal moves per position) and d is its depth (game length). The challenge, then is that such exhaustive search would be infeasible for relatively complex games such as chess ($b \approx 35$, $d \approx 80$) or Go ($b \approx 250$, $d \approx 150$).

There are two general principles to address this problem and reduce the effective search space. 1) Reduce depth of search by finding an approximate value function. This approach has led to strong performance in chess, checkers, and othello, but was considered to be difficult for Go. 2) Reduce breadth of search by simulating likely future actions. This approach has resulted in strong performance in backgammon and Scrabble, and weak amateur level play in Go.

AlphaGo is a finely tuned combination of the two approaches in its implementation of (1) 'value networks' to evaluate board positions and 'policy networks' to select moves and (2) a search algorithm that combines Monte Carlo simulation with value and policy networks. AlphaGo is made up of a number of relatively standard techniques: behavior cloning (supervised learning on human demonstration data), reinforcement learning (REINFORCE), value functions, and Monte Carlo Tree Search (MCTS).

Udacity Artificial Intelligence Nanodegree

Results:

The AlphaGo implementation uses a parallel asynchronous non-locking combined CPU and GPU architecture which scales.

AlphaGo sports a 99.8% win rate against other go programs and most dramatically defeated the current world champion of go, Lee Sedol 4 - 1. Against Fan Hui, the european champion of go, AlphaGo evaluated thousands of times less moves than Deep Blue did in chess against Kasparov. This demonstrates that the policy network can choose more intelligent moves while evaluating them more accurately using the value network.

[A new paper](#) was released in October 2017 detailing a new neural net [AlphaGo Zero](#) that does not need humans to show it how to play Go. Not only does it outperform all previous Go players, human or machine, it does so after only three days of training time.

References

Silver, Huang et al. "[Mastering the Game of Go with deep neural networks and tree search](#)." *Nature* 529 (2016).