Research Review: “Mastering the game of Go with deep neural networks and tree search” Silver et al. *Nature,* 28 January 2016.

Goals and Techniques

The DeepMind team sought to solve a game-playing problem previously thought to be intractable. The game Go has proved to be one of the most challenging board games for artificial intelligence due to the huge search space (b ~=250, d~=150, compare this to chess with b~=35, d~=80).

Tree search methods, such as the ones used by chess playing bots like DeepBlue, can approach the optimal value function. However, these methods used in Go’s search space can quickly lead to combinatorial explosion. Instead, the DeepMind team mind team used policy networks, a form of search that calculates a probability distribution without branching. To create effective policy networks, DeepMind used convolutional neural nets to analysis large datasets of board positions. The deep neural net was trained using both supervised and reinforcement learning so that AlphaGo could both learn from past master players and optimize play towards the goal of winning.

During the reinforcement learning phase AlphaGo played games against a random pool of opponents and received a reward for winning and penalty for losing. These rewards were back propagated using stochastic gradient ascent to maximize a winning outcome.

Results

To evaluate the strength of AlphaGo, DeepMind ran internal tournaments against other Go programs. In a fair game, AlphaGo consistently won the majority of games against other Go agents (around 99.8%). To better evaluate potential performance against professional human players, AlphaGo played handicapped games against the other AI agents. Using a single machine, AlphaGo won 77% of games against the next strongest computer program. Distributed AlphaGo proved to be more effective, winning 100% of games against other programs.

This results against computer opponents transferred well to play against professional human players. In a match against the European champion, AlphaGo won 5 games to 0. This was an unprecedented result. AlphaGo did more than defeat a professional Go player. It also proved that novel techniques can conquer what are thought to be intractable problems. In particular, using a combination of policy networks and deep learning techniques, there is potential for AI agents to approach human performance in many other domains.