AlphaGo Paper Summary

Goal

AlphaGo aims to outperform all existing Go programs and play full-sized game of Go at human professional level.

Technique

The major challenges of Go, compared with other perfection information games that have been “solved” by AI (e.g. Chess), are its enormous search space and the difficulty of evaluating the game state (i.e. board positions) and game moves. The state-of-the-art Monte Carlo search approach that simulates thousands of random games of self-play has achieved human amateur level.

AlphaGo introduces deep neural networks to tackle the above-mentioned challenges of Go. Specifically, two types of neural networks are implemented - value networks and policy networks. Value networks are responsible for evaluating board positions, while policy networks specialize in selecting moves. Together they help reduce the infamous gigantic search space of Go and produce more accurate evaluation of the game state. The training of these neural networks start with supervised learning from human expert games, and later move on to reinforcement learning from countless games of self-play.

The deep neural networks alone, without any search into future moves, is able to play Go at a level that is on par with the best computer Go programs.

Additionally, the team behind AlphaGo also combines Monte Carlo tree search with value network and policy network that selects actions by lookahead search.

Result

At the time of publishing, AlphaGo consistently beat other Go programs at a 99.8% winning rate, and defeated the European Go champion Fan Hui by 5 to 0. In March 2016, AlphaGo won a five-game match with Lee Sedol, the first time a 9-dan professional player losing to a computer Go program.