Research review: AlphaGo by the Deepmind Team

Laurent George

February 20, 2017

1   A brief summary of the paper's goals or techniques introduced

The presented approach combines machine learning and tree search techniques (Monte Carlo Tree Search) to play the game of Go. This approach does not rely on a handcrafted evaluation function: it is based on deep neural networks learning.

Two kinds of neural networks are used to reduce the depth and breadth of the search tree: policy and value networks. The policy networks aim at selecting intelligently which positions to explore in the search tree. These policy networks are first tained to predict which moves a human expert would choose. A dataset of 30 million positions from 160,000 games played by expert human players is used during the learning stage. A reinforcement learning approach is then used to improve this neural network. Concerning the value network, its aim is to predict precisely the winner of games played by the learned policy using board positions. A regression is used to train the value network.

The policy and value networks are finaly combined in a Monte Carlo Tree Search algorithm in order to selects moves by lookahead search. The AlphaGo program is an implementation of this algorithm. A single machine and a distributed version of the program have been developed.

2   A summary of the paper's results

The Alpha Go version which used only the policy networks to select the move, using no search at all, reached 85% of game against Pachi (the strongest opensource program for Go, which uses search). This highlights the power of deep neural network and reinforcement learning in this context.

The AlphaGo program (combining learning and search) running on a single machine achieved a 99.8% winning rate (494 out of 495 games) against other Go programm. When playing with handicap (i.e 4 free moves for the opponent), Alpha-Go won more than 77% of the games against the three best Go programs.

The distributed version of the AlphaGo was significantly stronger winning 77% of the games against single-machine AlphaGo and 100% of games against other programm. The distributed version was also evaluated against a human professional Go player. AlphaGo won the match. This was the first time that a computer Go program ever defeated a human professional player in the full game of Go.