

Research Review: AlphaGo

Christopher Brian Currin

Introduction

Silver *et al.* (2016) aimed to build a game-playing agent that could consistently win at the game of Go against any player. Because of the complexity of Go (with high breadth and depth), novel and advanced methods were goals of the development towards a winning agent.

Methodology

AlphaGo combines deep convolutional neural networks and Monte Carlo tree search (MCTS) to effectively build a winning game-playing agent.

Two principle types of deep convolutional neural networks were used: policy and value networks. The value network evaluated the board (positions of the players) to output a value indicating the likelihood of winning from that position. The policy network evaluated the best possible move from the current set of legal moves.

The training of the networks involved 1) supervised learning from (good) human playing data to classify and predict the next move for the policy network (to maximise prediction of moves), 2) reinforcement learning from playing the policy networks against iterations of themselves (to maximise winning probability), and consequently 3) regression of predicted outcomes to build the value network.

MCTS involved selecting moves of maximum value (moving to a better state according to 'action values' plus the trained policy network's probabilities), with consequential moves being evaluated as part of the selection. Post-simulation the game is then evaluated (using a fast rollout policy) according to who won and action values are updated for the next simulation.

Results

As part of the authors' evaluation of AlphaGo's ability, the pitted AlphaGo against other established Go-playing agents and one of the top players in the world; with consistent success. Evaluation of the optimal AlphaGo system itself involved comparing the abilities of the system with different combinations of its constituents (MCTS, policy network, value network), as well as performance on different architectures (number of GPUs and threads) for scalability.

The AlphaGo team claims to have solved one of artificial intelligence's "grand challenges" – that of creating a game-playing agent that can beat the best human players. In their comparison to previous famous AI's, mainly that of DeepBlue, they surmise that AlphaGo plays in a more similar (but advanced) manner compared to humans: *viz.* choosing moves more intelligently (with greater precision on evaluating said move).