

Review of "Mastering the Game of Go with Deep Neural Networks and Tree Search"

Michelle L. Gill

Overview

In Silver, *et. al.*, *Nature*, 2016, 529:484-489, the DeepMind team reported the development of a game playing agent for the game Go, called AlphaGo. Go is a particularly challenging target for artificial intelligence research, relative to games like checkers, chess, and Othello, because Go moves encompass a much larger search space, and the evaluation of both board position and future moves is more challenging.

Technical Innovations

The chief innovations enabling this development were a pair of convolutional neural networks, referred to as the policy and value networks, respectively. The policy network suggests next moves, effectively reducing the breadth of the search space for future moves. The value network reduces the required search depth of possible future moves by predicting the outcome.

The policy networks were trained using a catalog of known, expert human moves. The predictions from this model were then improved using reinforcement learning (RL) via self play, whereby the current policy network was matched against a previous iteration. This model was then tested against Pachi, the strongest open source Go program at the time. In these matches, 85% of games were won by AlphaGo's RL policy network.

The value networks, which evaluate the current position during game play, were also trained by reinforcement learning. In this case, completed games were used for the predictions, which initially resulted in overfitting. To counteract this issue, a new dataset containing positions sampled from separate games was generated.

These two neural networks were combined via Monte Carlo tree search (MCTS), which determines moves via lookahead search. The MCTS implementation in AlphaGo is asynchronous and multithreaded, utilizing 40 threads, 48 CPUs, and 8 GPUs. The authors also published a distributed version of AlphaGo that performs slightly better, albeit with greater hardware requirements.

Results

AlphaGo was tested against a variety of computational Go players, both open source and proprietary. Against these computational counterparts, AlphaGo was many dan ranks stronger, even when a handicap of four stones was given to the opponents. Overall, AlphaGo won an astounding 494 of 495 games against other computational Go players. Finally, AlphaGo played the European Go champion, Fan Hui. In this series, AlphaGo won all five matches.