

Research Review - Karan Pinto - AIND Isolation

Mastering the game of Go with deep neural networks and tree search

Exploring all possible outcomes of player-opponents moves, may come across as one of the optimum ways to decipher the best move made by a game agent. However, the game of GO, with a staggering $361!$ explorable possibilities, renders this method of traversing all possibilities in the game tree, inefficient and almost impossible within a realistic time constraint. The neural network approach, enables the agent to overcome this limitation.

The AlphaGo design consists of the Policy and Value Network respectively. The policy network, expresses the probability of positioning in a co-ordinate system given a specific game state, helping select moves. The value network, reflects the best overall game state by evaluating it, post the completion of the positioning. The following steps are taken to ensure perfect policy and value networks:

The policy network was trained for classifications by using a KGS data set that contained 29.4 million positions. The training was done using Deep Convolutional Neural Network (DCNN) and Monte Carlo Tree Search (MCTS) technologies. Researchers also used reinforced learning/self-playing games for policy network optimisation. Supervised learning- regression was then utilised to derive the value network (15 DCNN layers) from the policy network. These two networks enable AlphaGo to obtain the best move given the current game state.

In conclusion, this neural network architecture, trained by combining supervised (from human expert games) and reinforcement learning (from games of self-play), achieves an Elo rating of 2890 and 3140, beating mainstream Go programs like CrazyStone (1929) and even professional Go players.

This search algorithm that combines Monte Carlo simulation with value and policy networks, enables our AlphaGo program to achieve a 99.8% winning rate against other Go programs, and also successfully defeated the human European Go champion by 5 games to 0. It should be noted that this is the first time that a computer program has defeated a human professional player in the full-sized game of Go, a feat previously thought to be at least a decade away.
