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

By

Asarudheen

For Build an Adversarial Search Agent Project

Udacity.

Go is an abstract strategy board game for two players, in which the aim is to surround more territory than the opponent. Go is considered as one of the complex game for computers to win when compared to other games, because Go game has much larger branching factor which makes it more difficult to solve using traditional AI techniques such as MIN-MAX, Alpha-beta, etc…

What makes the AlphaGo to outperform?

The techniques used in AlphaGo took it to win against the human player.

Below are the techniques used in AlphaGo

- It uses 'value networks' to evaluate board positions and 'policy networks' to select moves.
- The neural networks are trained by supervised and reinforcement methods.
    - Stage 1: The policy network is trained on randomly sampled state-action pairs from 30 million position from KGS Go server.
    - Stage 2: The main aim of this stage is to improve the policy network by policy gradient reinforcement learning.
    - Stage 3: The weight of the value network is trained by the regression on state-outcome pairs (s, z), using stochastic gradient descent to minimize the mean squared error (MSE) between the predicted value and the corresponding outcome.
- Alpha Go uses Monte Carlo tree search.
    - AlphaGo combines the policy and value networks in an MCTS algorithm that selects actions by look ahead search. Each edge is evaluated in two ways: using the value network and by running a rollout to the end of the game with the fast rollout policy, then computing the winner.

Different versions of AlphaGo:

The final version of AlphaGo used 40 search threads, 48 CPUs, and 8 GPUs. The distributed version of AlphaGo that exploited multiple machines, 40 search threads, 1,202 CPUs and 176 GPUs. The Methods section provides full details of asynchronous and distributed MCTS.

Summary: Evaluating the playing strength of AlphaGo

In final tournament AlphaGo won 494 out of 495 games (99.8%) against other Go programs. AlphaGo played against Fan Hui, a professional 2, and the winner of the 2013, 2014 and 2015 European Go championships and AlphaGo won the match 5 games to 0. This is the first time that a computer Go program has defeated a human professional player, without handicap, in the full game of Go.