RESEARCH REVIEW

Mastering the game of Go with Deep Neural networks and tree search

Mastering Go is a very difficult problem. Unlike Chess or other games, **the branching factor of Go games is around 250** and the **depth length is around 150** making an exhaustive search impossible. Hence, while AI could have master Chess with a mixed of Minimax with alpha-Beta pruning, with database of opening moves and ending games and other techniques to maximize the heuristic and reduce the tree search, those kinds of techniques are impracticable for the game of GO. Before the publication of this paper, the best AI was based on the Monte Carlo tree search which is a technique that uses probabilities of wining as the weight for each branch. However, this technique isn't sufficient to master the game of GO.

The DeepMind team manages to avoid searching in a complex tree search by using a convolutional network to represent the position on the board. They then train a supervised learning algorithm from expert human move and used reinforcement learning to optimize the final outcome. Finally, they used Monte Carlo tree search to choose the best choice to make (the one that have the higher probability to lead to a wining state). AlphaGo uses several stages to build a strong AI:

Firstly, they used a 13-layer convolutional neural network that they trained over 30 million positions. They were able to gain a significant gain as the network predicted expert moves with and **accuracy of 57%** (before the release of this paper the accuracy was up to 44,4 %).

They then entered the second phase of the training pipeline where they improved the policy network by using reinforcement learning. At this point the Reinforcement Learning policy network won 80% of games against the Supervised Learning policy created at the first stage. They also play against Patchi and won 85% of the times where previous state of the art Supervised Convolutional networks won 11% of times.

The final stage of the training pipeline focuses on position evaluation. The goal is to accurately predict the outcome of the game from a certain position of the game. At this point they were able to approach the accuracy of Monte Carlo rollouts but **using 15 000 times less computation**.

Finally they combine the policy and value networks in a Monte Carlo Tree search to evaluate the best choice to make at each state of the game.

Finally, they come up with an AI that masters the game of GO. The results are unequivocal. **AlphaGo won 494 out of 495 games against other Go program**. With 4 handicap stones, AlphaGo manage to win respectively 77%, 86% and 99% of the times against Crazy Stone, Zen and Patchi[1]. The distributed version of alphaGo won 77% of games against single-machine Alphao and 100% of its games against other programs. Finally AlphaGo win 5 games to 0 against Fan Hui, the winner of the 2013, 2014, 2015 European Go championships.

The same principles applied to solve Chess was impracticable to use for an AI to master the game of GO. Other ways was explored and bring a huge improvement in the game of GO by using MCTS. Yet it wasn't sufficient to master the game of GO, DeepMind team come up with the idea to combine MCTS with policy and value network to finally build an AI that masters GO.

[1] *Patchi, Zen and Crazy Stone are other names of AI programs that can play GO*