REVIEW SUMMARY OF THE PAPER "MASTERING THE GAME OF GO WITH DEEP NEURAL NETWORKS AND TREE SEARCH"

Veeresh Taranalli
veeresht@gmail.com

1 INTRODUCTION AND GOALS OF THE PAPER

In this review, we present a concise summary of the seminal paper "Mastering the game of Go with deep neural networks and tree search" that appeared in the journal Nature in early 2016. This paper presents various improvements to the state-of-the-art Go playing artificial intelligence (AI) agents, with the help of which the proposed agent termed "AlphaGo" was able to defeat the human European Go champion by 5 games to 0. The proposed techniques are based on the use of deep neural networks which made the "AlphaGo" program feasible in practice.

Go is a 2-player strategy board game of perfect information similar to other games such as Chess and Checkers. Games of perfect information can be solved using an exhaustive tree search algorithm that explores the future of every possible move at a particular stage in the game and chooses the best possible move assuming the opponent would do the same. However, the complexity of such exhaustive search makes it infeasible for even small board games with very few possibilities. Prior to this work, the game of Go had been one of the grand challenges in the field of AI simply because of its large size of search space. Past AI research has given us various algorithms for game playing that are efficient in practice, notably the alpha-beta pruning with iterative deepening minimax search in which the search stops at a particular depth and all game states at that depth are evaluated using an approximate value function (also termed heuristic evaluation function) that assign a particular value $\nu(s)$ for a game state s that reflects the likelihood of a win from that state. Another approach is to use Monte-Carlo tree search (MCTS) where not all actions possible at a particular state are considered, but instead actions are sampled according to a policy $p(a|s)$, and by playing a large number of games, the true value function of a state $\nu(s)$ and the optimal action policy at a particular state $p(a|s)$ are estimated.

This paper proposes the use of a supervised learning (SL) policy neural network (p_σ) that learns to predict the best possible move at a state of the game. The input to the neural network consists of the board position image of size 19×19 pixels and also some derived features that are simple functions of the board position, as described in detail in the paper. Such a neural network is trained using a database of human expert moves. Next the paper also uses a reinforcement learning policy network (p_ρ) that optimizes/directs the SL policy network to achieve the final goal of winning rather than prediction accuracy of current move which is not important. Finally the paper also proposes a value function neural network (ν_θ) that predicts the winner of games played by the RL policy network against itself. The proposed AlphaGo program combines all these techniques with the MCTS technique to achieve the record-breaking feat of beating a human Go champion. The use of deep neural networks for SL policy learning and value function evaluation contribute to the novelty of this work.

2 KEY RESULTS OF THE PAPER

The AlphaGo progam was evaluated against other Go playing programs that are based on high-performance MCTS algorithms. Two versions of AlphaGo were considered, one that used a single machine (40 search threads, 48 CPUs and 8 GPUs) and another that was distributed across multiple machines (40 search threads, 1202 CPUs and 176 GPUs). The single machine AlphaGo version lost just 1 game out of 495 games played against other competitive computer programs. The distributed version won 77% of all its games against single machine AlphaGo and all its games against the other computer programs. Apart from these, the AlphaGo program was also able to win a tournament, 5 games to 0, played across multiple days with Fan Hui, a professional go player and a 3-time European Go champion. A computer beating a human consistently at the game of Go was the feat achieved for the first time in history of the game of Go.