Name : Ahmed Yamout
Email : a.yamout@nu.edu.eg

AlphaGo

AlphaGo Goals and techniques:

The AlphaGo paper explains how the agent uses deep learning to assign score values to gamestates in their gametree using a neural network that weights the game positions, then determines which game branch to play using a policy neural network that uses heuristic functions.

AlphaGo uses supervised learning ,that is the machine learning networks are given input and the desired equivalent output to it is provided as well, Then the neural networks adapt to obtain the same outputs from the input given. In a later stage however, AlphaGo started using Reinforced machine learning, where the AlphaGo would play against an instance of itself , and by iterating the method on a huge amount of matches it developed the neural networks that have won against the top Go player.

AlphaGo uses Monte Carlo Tree Search , which is a search algorithm to evaluate (using the four policy neural networks) the weight of each gamestate in the game tree and select the ones more weighted towards the program win.

The first network used in assesing gamestates is the supervised learning neural network which is a 13-layer deep convolutional neural network trained on 30 million Go game positions. Its purpose is to predict the next move from a certain gamestate.

The next neural network depends on reinfoced machine learning. The neural network started with the same structure as the supervised neural network, but rather than being fed input and equivalent optimal output, two isntances of the neural network play against each other and adapted to produce the optimal winning plays and situations. Their had been 1.2 million reinforced learning matches to produce this powerful neural network.

The third neural network is called the fast roll-out policy network, it is similar to the supervised learning network however is altered. It provides more a faster expectation for the next move , but is not as accurate as the supervised learning CNN.

The final neural network is the Value network , it assigns the probability of each game state of leading the agent to win, it has been trained to assign probability of win to gamestates by using over 1.5 billion matches of reinforced learning and 30 million games that have been played against humans.

AlphaGo results:

AlphaGo significantly outperforms any AI AlphaGo playing agents that existed from before, it has beaten the world's top player as well.AlphaGo has been implemented using several hardware settings and tested to obtain the "Elo rate" which is the rate used to determine how good a player is in Go game. The highest Elo rating of AlphaGo was obtained by using running (in distributed mode) 1920 cpus and 280 gpus and using 64 search threads; this resulted in an Elo rating of 3168.