Stephane Chong / chong.stephane@gmail.com

AlphaGo Paper Summary

Goal: the goal of the program AlphaGo is to combine deep neural networks with Monte Carlo Tree Search (MCTS) techniques to develop a Go game-playing agent capable to beat all previously known Go game-playing agent, ultimately capable to beat a human professional player in the full-sized game of Go, a feat previously thought to be a decade away.

Prior work: The strongest current Go programs are based on MCTS, yet limited on shallow policies and value functions based on linear combination of features.

New techniques: AlphaGo uses of deep neural networks and modern computation power to make it possible to develop more complex policies and value functions to perform an improved MCTS

- AlphaGo uses Value network to evaluate board positions and a Policy network to select moves based on their probability of winning. These networks are 13-layers neural networks trained by a combination of supervised learning from human experts moves and reinforcement learnings from self-play. The value and policy networks are trained using available large data sets on expert games and available Go game-playing agents. Additionally a faster but less accurate “fast rollout” policy is trained for later use in the tree search due to its fast action selection.
- A search algorithm is then implemented using MCTS, where the best action is determined after 4 steps: selection, expansion, evaluation and backup.
    - Selection is performed by maximising the action value plus a bonus proportional to the prior probability given by the policy network for the edge
    - Leaf nodes may be expanded, with new nodes being processed once by the policy network and the output being stored as prior probabilities for the action
    - The end of the simulation is evaluated using a combination of the value network and a “fast rollout” policy to the end of the game
    - Action values are then updated back-up to the root position to reflect the mean value of all evaluation of their subtree, and the most visited is chosen

Results:

AlphaGo achieved 99.8% winning rate against other Go players and defeated the human European Go champion 5 games to 0.

Reference: https://storage.googleapis.com/deepmind-media/alphago/AlphaGoNaturePaper.pdf