Summary of DeepMind's AlphaGo paper - Artificial Intelligence Nanodegree, Udacity

12 February 2017

Silver, Huang et. al show that the game of Go – an ancient Chinese zero sum board game where two players compete to surround as much territory as possible using black and white stones placed on a 19 by 19 grid board – can be mastered by a game-playing agent based on deep neural networks. In particular, the authors show that leveraging human expert Go knowledge to inform a supervised deep neural network and using reinforcement learning to play many games against itself, this agent is able to play at the same state-of-the-art level as Monte Carlo tree-search agents without any look-ahead. Combining these two deep networks, a 'value'-based network and a 'policy'-network, the agent designed by Silver, Huang et. al achieves a win rate of 99.8% against other computerized Go game-playing agents, and was able to beat the European Go champion by 5 games to 0.

Silver, Huang et. al begin by framing the problem of leveraging traditional tree-search algorithms to find optimal solutions to Go, despite the fact that it is a game of perfect information and that there exists an optimal value function that determines the outcome of every game. This issue stems from the average branching factor and tree depth of Go, which is about 250 and 150, respectively – more than the number of atoms in the Universe.

The paper continues to describe that the use of convolutional neural networks, and passing the game state as a 19 by 19 game board image as input, and using supervised training through expert human moves to inform the network on the appropriate label of classification for the next move, which helps achieve efficient learning updates with immediate feedback. Next, the authors train a reinforcement learning deep neural network that improves the first network by optimizing the final outcome of games of self-play. Finally, Silver, Huang et. al train a neural network that tries to predict the winner from games played by the reinforcement learning network.

Using deep neural networks, and combining several of them, the authors show that it is possible to create a game-playing agent that plays at the same level as the strongest human Go players by leveraging the optimized policy network to select moves, rather than brute-force searching a large part of the search tree of available moves for every possible game state.