**Udacity AI Nanodegree: Research Review of Alpha-Go**

**Advanced Game Playing Agent Module Project**

**Omkar Vedpathak, April 2017 cohort**

The goal of the Alpha-Go research project was to try to exceed the performance of previous Go game playing agents, with the stated goal of beating expert human champions of Go. Another goal was to play in an intuitive manner, similar to humans, and exhibit creativity. The ultimate aim of the initiative is to apply the knowledge and algorithms developed to other fields where there are clear combinatorial explosions e.g. drug design and material design.

The Alpha-Go research paper introduced a means to generate an evaluation of any board state, called a ‘Value Network’, which had not previously generated for the game of Go before. This ‘Value Network’ generates a score from 0 to 1 of the probability of a given board state that the board state would lead to winning game for the Alpha-Go Agent. This allowed the Alpha-Go agent, at any particular state of the tree, to evaluate whether a given set of moves to produce a board state would be likely to lead to a winning state or not.

The Alpha-Go paper combined this ‘Value Network’ with a ‘Policy Network’, which is a probability map produced from performing Reinforcement Learning using Deep Neural Nets from millions of amateur games of Go as to what likely moves human players would make for a given board. This Value network allowed the Alpha-Go agent to reduce the branching factor significantly from around 200 to around 10, since the Value Network provides a means of generating the best n moves for a given board state. The Policy Network allows the agent to select moves.

The Policy Network and Value Network were used in combination with a modified Monte-Carlo tree searching (MTCS) algorithm, to significantly improve the performance of MTCS algorithm to produce optimal moves. The use of the Monte-Carlo network was also seen to be unconventional as more recent trends in AI research opted for purely Neural Net and Deep Learning based methods.

The research outlined in the paper showed that Deep Mind achieved and exceeded some of its goals, namely:

* Beating a professional Alpha-Go player
* Having a 99.8% winning rate against other Alpha-Go programs
* Beating the Alpha-Go world champion, and Go ‘genius’ Lee Sedol
* Displaying ‘creativity’ – the Alpha-Go agent played a lot of creative and unconventional moves, and using a far less ‘brute-force’ approach vs Deep Blue (1000s fewer positions were evaluated than Deep Blue, despite the search space being an order of magnitude greater than Chess)