Artificial Intelligence Nanodegree Program: Research review by Luis Palacios

AlphaGo by the DeepMind Team
Mastering the game of Go with deep neural networks and tree search

The specific goal of the research is to master the game of Go with **deep neural networks** and **tree search**. The program that was created for this challenge is AlphaGo. The DeepMind team's goals was to find and apply a novel approach that is successful in games that have enormous search spaces. How enormous exactly is not defined, but the DeepMind team refers to the search space of Go as being "infeasible". Just to illustrate a bit more considering b^d possible sequences of moves, where b is the game's breadth (number of legal moves per position) and d is its depth (game length). for chess this is (b≈35, d≈80) and for Go (b≈250, d≈150). The massiveness of Go was previously thought to make the game unconquerable by 2 computers until later in the future. The DeepMind team introduce several novel techniques (or combinations of techniques) for example:

- Using 'value networks' to evaluate board positions and 'policy networks' to select moves
- Training these deep neural networks using a combination of supervised learning from human expert games, and reinforcement learning from games of self-play
- A new search algorithm that combines Monte Carlo simulation with the 'value' and 'policy' networks

As Results of these efforts Fan Hui and more recently in 2016 Lee Sedol both human professional Go players defeated by a computer program (AlphaGo) for the first time in history and decades in advance of public expectation. Fan described the program as "very strong and stable, it seems like a wall. ... I know AlphaGo is a computer, but if no one told me, maybe I would think the player was a little strange, but a very strong player, a real person" Looking back at Deep Blue that mainly relied on brute computational force to evaluate millions of positions, AlphaGo relied on neural networks and reinforcement learning which more closely resemble human decision-making. DeepMind research revealed that the final version of AlphaGo used 40 search threads, 48 CPUs, and 8 GPUs and a distributed version with 40 search threads, 1,202 CPUs, and 176 GPUs was also implemented, the program's competitiveness in terms of Elo rating exhibited diminishing returns.

So The DeepMind team by facing "a challenging decision-making task, an intractable search space" provided "an optimal solution so complex it appears infeasible to directly approximate using a policy or value function." They conquer a problem that previously had been regarded as a hard problem in machine learning that was expected to be out of reach for the technology of the time thus they made the scientific community start discussing preparations for the possible future impact of machines with general purpose intelligence

"The end of an era... board games are more or less done and it's time to move on."
Deep Blue's Murray Campbell on AlphaGo's victory against Lee Sedol