Review of "Mastering the game of Go with deep neural networks and tree search"[1]

Grant Bartel

The game of Go is an ancient Chinese game that requires from its players a keen sense of intellect and strategy. For a long time it was thought to be in a distant future when computers would one day have the ability to challenge the most skilled players of this complex game. However, recently the computer program AlphaGo, using a mix of techniques well-known in the artificial intelligence community, was able to beat the world's most expert players. The following is a review of the article "Mastering the game of Go with deep neural networks and tree search" by Google DeepMind on how AlphaGo and its developers were able to achieve such a feat along with their outcomes.

The major goal of creating AlphaGo was to be able to defeat the most advanced players at the game of Go. To achieve this, they used a combination of convolutional neural networks (CNNs) and search trees. While it is common to use search trees to play zero-sum games such as chess, checkers, and Go, it's not always feasible to solely depend on them. When taking into consideration all of the possible game states, moves, and outcomes, some games prove to have a computationally infeasible space to evaluate. As deep neural networks have become the goto solutions for such huge, nonlinear problems, it was only a matter of time they would be used to master the most difficult games.

There are two phenomena the developers of AlphaGo wanted the program to learn: expert Go moves (knowledge) and when best to use them (skill). To do this, they built multiple CNNs that could take images of Go game states (i.e., boards) and produce predictions of next best moves as well as the final outcome based on those moves. To predict the next best moves, they trained a *policy network* via supervised learning based on games played by experts. After training this CNN extensively, they then used its final weights as a "jumping off point" for two other virtually identical CNN architectures, which would use reinforcement learning to produce the final *policy network* and *value network*.

Reinforcement learning differs from supervised learning in the sense that optimal input/output pairs are almost never used and the underlying network must learn over time the best path to take through a combination of exploring uncharted territory and exploiting current knowledge. Using this learning technique and the initial CNN weights of the initial *policy network*, the final CNNs were trained by actually playing games of Go; against humans, other computers, and itself. Through the new *policy network* and *value network*, AlphaGo developed the knowledge and skill to come up with expert moves and how those moves might affect the outcome of the game. However, since there are an incomprehensible amount of possible next moves and outcomes based on a given Go game state, a way to search these possibilities in a smart way was required.

To search through these possible scenarios when playing games, it's common to use search trees. Basically, at a current state of the game there are a number of possible moves and outcomes which lead to many more moves and outcomes. By creating a search tree that uses the previously trained CNNs while traversing through the tree, AlphaGo was able to determine which moves were more likely to lead to a win. At every game turn granted to AlphaGo, within given time constraints it would traverse this tree as intelligently as possible while running its CNNs to decide its next move.

Something like AlphaGo was not expected to be realized until at least a decade in the future. However, via deep convolutional neural networks, supervised learning, reinforcement learning, search trees, a lot of data, and a great team, the developers of AlphaGo achieved in creating a program that could beat the best Go players in the world. Depending on the program's underlying hardware, the levels of game playing expertise can grow very large. At the time of the article's publication, AlphaGo's greatest feat had been defeating Fan Hui (a Go world champion) in a formal five-round match, all of which Alpha won. Currently AlphaGo is the world's raining Go champion and has left all competitors in its wake.

[1] Mastering the game of Go with deep neural networks and tree search