Research Review: AlphaGo

Patricio Beltran Molas

February 2018

The research paper chosen for this report was "Mastering the game of go with deep neural networks and tree search".

Go has always been an interesting game for the AI community because of the unbelievably huge game tree it produces. Its branching factor comes close to 250 and its depth more than 150. This gives us almost 3.48×10^{17} possible situations. Traversing this tree to endgame would probably take longer than the time planet Earth has existed. Because of this, until recently, the possibility of a computer program beating a human at professional-level Go was thought to be decades away.

The team behind AlphaGo accomplished this by taking advantage of recent advancements in deep convolutional neural networks research. They designed a revolutionary training multi-stage pipeline that leveraged such advancements and tackled training with a variety of approaches. This, in combination with high performance searching algorithm Monte Carlo, resulted in huge advancements in the field of game playing.

At the highest level, two networks were used: a policy network that was intended to reduce the breadth of the tree, by choosing suitable paths, and a value network, that computed the value of nodes to try and reduce the depth of the tree.

During the first stage, a 13-layer policy network was trained using 30 million positions from previous human played games. This led to almost a 10% increase in correct predictions over the current state-of-the-art research.

Then, a second policy network was trained, this time using self play. The goal of this network was to reinforce the learning of the first network. This would result in an increased prediction of correct moves that would drive the program towards winning the game. The result of this stage was an 85% win over the current state-of-the-art program.

The last stage of the pipeline was the training of the value network. The goal of this network was to effectively evaluate positions to choose the right path.

The combination of MonteCarlo search and the expertly-trained CNNs yielded great results that became visible when playing against several other pro-level programs, losing just 1 out of 495 games. But the most impressive result was beating the European Go champion Fan Hui, 5 - 0 in formal matches.