Summary of Mastering the game of Go with deep neural networks and tree search paper

Ruben Rodriguez-Fernandez

May 15, 2017

Since Garri lost to Deep Blue in the 1990s, artificial intelligence pioneers have been looking for new challenges, and it was not until 2016 that a new challenge was overcome, this time playing the Go game.

Larger games such as Chess and Go has been object of study due to their huge search space, which is roughly b^d where $(b \approx 35, d \approx 80)$, $(b \approx 250, d \approx 150)$ for Chess and Go respecively, cannot be addressed using exhaustive search techniques such as Minimax. Altough there are several techniques that are able to beat professional human players in chess, the large difference between its search space size make then infeasible for Go. Thus, new techniques should be created in order to reduce in some way the search space.

AlphaGo's team use a set of techniques which allows them to overcome the problems presented so far and beat the world champion Lee Sedol. The techniques used by AlphaGo, that are organized in pipelines, are described in the following paragraph.

In the first stage of the pipeline, they build a *Supervised Learning policy network* using data from Professional Players. Then, in the second state of the pipeline, a *Policy gradient reinforcement learning* is built with the same structure and weights as the SL policy network. After training the Policy gradient RL using randomly selected previous iterations of the policy network, the RL policy network won 80% of the matches against the SL policy network. The last stage of the pipeline focuses on *position evaluation*, as the $v^*(s)$ function which determines the outcome of the game in any state is infeasible, a approximation is built using the RL policy network.

Finally, AlphaGo combines the policy along the value networks in a Monte Carlo Three Search (MCTS) algorithm which determines the actions by lookahead search.

In conclusion, the AlphaGo team developed a new architecture based on Deep Neural Networks and Three Search that was not only able to play Go at levels of professional players but also have applications in domains where the search space is large such as game-playing, classical planning and scheduling.