ARTICLE SUMMARY

By Vinicius Ribeiro Machado da Silva

Mastering the Game of Go with Deep Neural Networks and Three Search

doi:10.1038/nature16961

Goals & Developments

The state of the art to solve games was not applicable to Go game due to its complexity. The need to evaluate the outcome of the game for every board position would lead to an intractable number of evaluations. The search space needed to become smaller, which is guaranteed by reducing its depth and the breadth.

The authors suggests that the depth of the search was truncated at a certain level and the remaining subtree was represented by an approximate function. The breadth would be reduced by sampling moves from a probability policy distribution of possible moves “a” in position “s”, where the Monte Carlo approach provides an effective position evaluation by averaging over the rollouts sampled (Monte Carlo search tree - MCST).

Beyond the MCST, the new search algorithm uses deep convolutional neural network to reduce the breadth and depth of the search tree, evaluating positions using a value network and sampling actions from a policy network. The strategy to train such neural network was adopted as the pipeline presented as follows:

a) Train the policy networks on a supervised learning way to predict human expert moves;
b) Train a reinforcement learning policy network that improves the supervised learning policy network by optimizing the final outcome of games of self-play;
c) Train a value network that predicts the winner of the games by the reinforcement learning policy network against itself;

Results Summary

The authors presented novel ways to reduce depth and breadth in the search space and a novel training of the deep neural network used, combing a pipeline of supervised and reinforcement learning.The developed algorithm was capable of playing at the same level of the strongest human players. Using the search algorithm, the program achieved a 99.8% winning rate against other Go algorithms.

The AlphaGo algorithm had a much better performance than Deep Blue, the algorithm used to beat Kasparov in chess game, reducing the amount of moves evaluation in a smarter way with the new techniques developed.

With the succeeded developments, they implied in advancements in other domains, such as classical planning, scheduling and constraint satisfaction, as examples.