

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

This research focus on providing a new aproach to Artificial Inteligente on playing Go Games. T

The Go Game is considered to be the most challenging classic game for artificial intelligence, because of its humongous search space.

Techniques used:

The technique introduced uses search, but not on the game space as is, but in a pre-built “value tree” using Monte Carlo Tree Search.

First, there was built a neural network to predict the next move based on actual state, called “Policy Network”. The game would choose, based on actual state, a new state, based on examples of a professional players database. Then, this network was refined using Reinforcement Learning, playing with itself and rewarding winning moves for this player.

Another network was built, called “value network”, to predict the winner, trained with reinforcement learning analysing states of two policy networks playing with each other. To minimize overfitting, there was generated a sample from diferent games at diferent positions, otherwise the network would simply memorize the sequence of previous played games and would not generalize.

Then, the final algorithm combines the two networks using Monte Carlo Tree Search (partially random) with total exploration, but selecting only most promising nodes. To encourage exploration, a visit counter penalizes already explored edges.

Even with all the pretrained value and policy guides, searching in Go is still huge space, with the overhead of evaluating neural networks, several times of magnitude more computing than traditional heuristic. Then, there was developed an asynchronous, distributed, multi-threaded search.

Results:

The results of the research are impressive. The algorithm beated previous Go algorithms almost in all matches made, with . And, for the first time, a machine coul beat a professional human Go Champion, the European three times champion Fan Hui, a fact that was achieved ten years before if was expected.

(In fact, with some improvements made with iterations with humans players, including Fan Hui, focusing on discover the algorithm weak points, AlphaGo beated the world champion a year later, but this was latter than the research publication, so it is not on this text)