# Review on *AlphaGo* paper

Weixiong (Victor) Zheng, Sep. 04, 2017

## 1. Goals and techniques of the *AlphaGo:*

1) Goals:

There's one main goal of the AlphaGo: to defeat professional human Go player. Go has long been considered as the most challenging classical games for artificial intelligence (AI) due to the enormous space for searching and position/move evaluation. Technically, researchers wanted to demonstrate the feasibility of combining machine learning with AI to hugely advance the agent ability on playing the game of Go.

2) Issue to address:

Due to the curse of the high dimensionality, pure search-based AI agent suffers from horizon effect. All the possibilities that the agent can search within limited time is still a little portion of all the possibilities of existing moves. As no intuition is introduced into the search, it's quite possible AI agent moves like an amateur leaving opponent great chances of winning. So basically, the core problem is how to introduce such "intuition" like human does.

3) Techniques:

Generally speaking, the technique is to combine searching algorithm with the "intuitions" provided by machine learning algorithm to boost the searching efficiency. In total there are three stages of machine learning operations plus a new hybrid searching method combining Monte Carlo search and neural networks

a) Machine learning techniques:

There are three stages of machine learning algorithms to provide such intuitions. The first stage is a supervised learning policy network is trained to select moves based on training received from playing with experts; the second stage is that a reinforcement policy network is engaged to further improve the move selection with self-play using previous iterations; the third stage is to use reinforcement learning of value networks to predict the expected outcome at a specific position.

Even without a lookahead search algorithm, the neural nets play Go at a level of programs based upon state-of-art Monte Carlo tree search (MCTS).

b) Searching algorithm:

Besides neural networks, *AlphaGo* still introduces a newly developed algorithm combining MCTS with policy networks to select actions. Differing from conventional MCTS, *AlphaGo* linearly combines Monte Carlo rollout with value and policy networks to produce a series of leaf evaluations in order to generate an action value. Once search is done, the move most visited from the root position will be chosen. With this newly developed searching algorithm, *AlphaGo* plays at a winning rate of 99.8% against other Go programs.

## 2. Summary of the results:

Up to the time of publishing the paper, by playing games with 5s of computations per move, *AlphaGo*'s had achieved:

1) Single-machine *AlphaGo* won 99.8% in fair games against other state-of-art MCTS based Go AI programs

2) Single-machine *AlphaGo* won from 77% to 99% of the games against other state-of-art MCTS based Go AI programs with four handicap stones (i.e. four free moves for opponent programs).

3) Distributed *AlphaGo* won European champion Fan Hui (5:0). What's interesting is that Fan Hui even preferred some moves by *AlphaGo* over his own moves by intuition.