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

Go is considered one of the most complicated classical games for Artificial intelligent with 250^{150} possible move sequences, large then chess which has only 35^{80} possible move sequences.

The aim of the pepper is to reduce depth and breadth of search using deep and reinforcement learning, to create agent which can out perform state of the art Go program which use Monte Carlo tree search.

The authors reduced depth and breadth of search tree by using convolutional neural network to analyse 19 x 19 board image to construct representation of the positions, and two neutral network. Value network to evaluate the positions and policy network to evaluate sampling actions.

The authors trained the neural networks using pipe line which consist of the following stages of machine learning:

- Supervised learning to train 13-layer policy network (SL) from 30 million positions from the KGS Go Server taken from directly from human expert which provide fast learning with immediate feedback.
- Reinforcement learning to train policy network which used to improve (SL) policy network, by optimising the final outcome of game of self play.
- Final stage is to train value network to predict the winner of the games played by (RL) policy network against it self.

The approach used effectively combine RL policy and value network with Monte Carlo tree search, which utilise random sampling of tree with evaluation of the game tree branch.

To evaluate the AlphaGo the authors ran internal tournament among variants of AlphaGo and others open source Go programmer and state of the art commercial Go programs.

The result of the tournament show's single Alpha Go is stronger then many existing Go program, winning 494 games out of 495 (99.8%).

The final challenge for AlphaGo is game agains Fan Hui professional 2 dan and winner of the 2013, 2014 and 2015 European championships, where AlohaGo was the winner.