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

A Summary Of Techniques Introduced

AlphaGo had introduced a new approach that uses 'value networks' to evaluate board positions and 'policy networks' to select moves and introduce a new search algorithm that combines Monte Carlo simulation with value and policy networks.

Value Networks (VN)

The value network provides an estimate of the value of the current state of the game. The input to the value network is the whole game board, and the output is a single number, representing the probability of a win.

Policy Networks (PN)

The policy networks provide guidance regarding which the next best action to choose. It takes the game board 19 x 19 image as input, then computed and output a 19 x 19 x 2 image that represents the "next best" board position for both white and black.
The first stage of PN trained by supervised learning using 30 million positions from the KGS data set. The second stage of policy network is trained using policy gradient reinforcement learning.

Monte Carlo Tree Search (MCTS) Simulation

AlphaGo using the MCTS algorithm, which searches game trees by mixing in truncated Monte Carlo rollouts, they don't consider all possible moves in a given state but chose the one with the highest probability according to a policy that is trained to predict human expert moves and leading to a win.

A Summary Of The Key Results

AlphaGo was evaluated by ran an internal tournament among variants of AlphaGo and several other Go programs, all of them based on high-performance MCTS algorithms.
The results of the tournament AlphaGo achieved a 99.8% winning rate against other Go programs, and the distributed version of AlphaGo was significantly stronger, winning 77% of games against single-machine AlphaGo and 100% of its games against other programs.
The AlphaGo defeated Fan Hui the European Championship for 2013, 2014 and 2015 by 5 to 0. This is the first time that a computer program has defeated a human professional player in the full-sized game of Go, a feat previously thought to be at least a decade away.