Sachin Solkhan
AIND - Isolation

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

The paper talks about AI approaches to create the AlphaGo computer program used to play game of Go that achieved 99.8% winning rate. This winning rate is significant given the complexity of this game to computationally evaluate through a large search space given the large permutations and combinations.

AI approaches encompasses using neural networks that are trained by combination of a) supervised learning from human expert games and b) reinforcement learning from self plays. These were combined with Monte Carlo simulation to create a new search algorithm. It uses value networks to evaluate board positions and policy networks to select moves. It helped in creating highly effective move selection and position evaluation functions. That helped in efficient and intelligent evaluation of through the large and complex search space.

This approach led AlphaGo to work 5-0 against human European Go champion in 2015, a feat that was believed to be a decade away.

The approach involves the following stages -
1. Supervised learning by training of a 13-layer neural network called the SL policy network that was trained with 30 MM positions from KGS Go Server.
   It was improved upon in next 2 stages.
2. Improving the policy network by policy gradient reinforcement learning.
3. The final stage of the training pipeline focussed on position evaluation, estimating a value function that predicts the outcome from position s of games played by using policy for both players
4. Combining the policy and value networks in an Monte Carlo Tree Search (MCTS) algorithm that selected actions by lookahead search.
    To efficiently combine MCTS with deep neural networks, it uses an asynchronous multi-threaded search that executes simulations on CPUs, and computes policy and value networks in parallel on GPUs