Silver - "Mastering the Game of {Go} with Deep Neural Networks and Tree Search"

Goals

- AlphaGo is a computer program that was created by Google Deepmind with the purpose of defeating a human professional player in the full-sized game of Go for the first time (this feat was thought to be at least a decade away).
- AlphaGo's neural networks are trained directly from game-play purely through general-purpose supervised and reinforcement learning methods.
- The motivation for developing this general approach to learning the game of Go was to help enable advancements towards achieving human-level performance in other seemingly intractable artificial intelligence domains.
- Just like Go, these domains tend to involve an intractable search space with an optimal solution so complex it appears infeasible to directly approximate using a policy or value function.

Techniques

- Rather than using a handcrafted evaluation function required by IBM's chess playing program - Deep Blue, AlphaGo was able to evaluate thousands of times fewer positions by selecting those positions more intelligently, using the policy network, and evaluating them more precisely, using the value network. This is perhaps closer to how humans play.
  - AlphaGo uses **value networks** to evaluate board positions and **policy networks** to select moves.
  - These deep neural networks are trained by a novel combination of supervised learning from human expert games, and **reinforcement learning** from games of self-play.
  - Without any lookahead search, the neural networks play Go at the level of state-of-the-art **Monte-Carlo** tree search programs that simulate thousands of random games of self-play.
  - AlphaGo also introduce a new search algorithm that combines Monte-Carlo simulation with value and policy networks.

Results

- Using this search algorithm, AlphaGo achieved a 99.8% winning rate against other Go programs, and defeated the European Go champion, 5 - 0.