Research: AlphaGo **by the DeepMind Team**

Summary:
AlphaGo is an AI computer player developed by DeepMind. It played with professional players and won. Algorithms used include policy network and value network. Essentially, AlphaGo learns from human players and develops improved solutions, thus becoming a better player.

Techniques:

1. **Monte Carlo Tree Search (MCTS)**

- Another approach to searching the game tree.
- Run many game simulations. The more simulations are executed, the more accurate the numbers are, the higher likelihood that an optimal play will be converged to.

2. **Policy Network**

- Provides guidance regarding which action to choose, given the current state of game. Output is the probability value for each possible legal move. Actions with higher probability values correspond to actions that have a higher chance of leading to a win.
- Deep reinforcement learning
    - Improved policy networks by letting the system play against each other using the outcome of the games as a training signal.

3. **Value Network**

- Provides an estimate of the value of the current state of game such as the probability of the white player to win the game given the current state. The input to the value network is the whole game board and the output is a single number representing the probability of a win.
- Trained on 30 million game positions obtained while policy network played against itself.
- Makes random predictions at the beginning of the game and becomes better at predicting the final game outcome when more moves have been played.

Result of AlphaGo:

- High computational power
- Won professional players

Downside of AlphaGo:

- Imitation of human players
- Relies on large dataset
- System is trained in phases - neural networks and tree search