# Research Review - AlphaGo
Qi Shao

Introduction

AlphaGo is a narrow AI, computer program developed by Alphabet Inc.'s Google DeepMind in London to play the board game Go.AlphaGo's algorithm uses a Monte Carlo tree search to find its moves based on knowledge previously "learned" by machine learning, specifically by an artificial neural network (a deep learning method) by extensive training, both from human and computer play.

Architecture
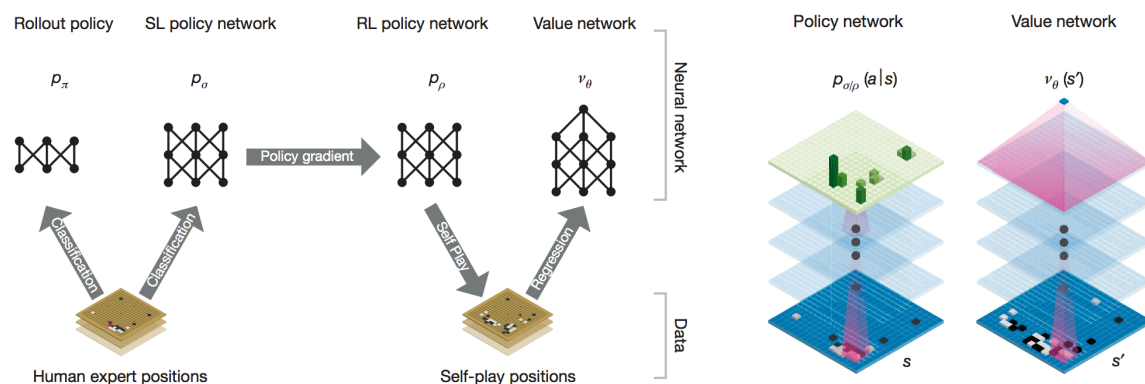
From Article: **Mastering the game of Go with deep neural networks and tree search**

The neural networks is trained through a pipeline consisting of different stages of machine learning:

1. p_π a policy network trained on random rollout
2. p_σ a supervised learning (SL) policy network directly from expert human moves, which provides fast, efficient learning updates with immediate feedback and high-quality gradients
3. p_ρ a reinforcement learning (RL) policy network that improves the SL policy network by optimizing the final outcome of games of self-play and adjusts the policy towards the correct goal of winning games, rather than maximizing predictive accuracy
4. v_θ a value network trained by regression that predicts the winner of games played by the RL policy network against itself

Policy Network takes representation of board positions as input, passes it through many convolutional layers with **σ** or **ρ**, and outputs a probability distribution (probability map over the board) **$p_\sigma(a|s)$ or $p_\rho(a|s)$** over legal moves **a**.

Value Network does similar work with **θ**, but outputs a scalar value **$v_\theta(s')$** that predicts expected outcome in position **s'**.

Evaluations

AlphaGo outperformed several other Go programs such as Crazy Stone, Zen, Pachi, Fuego, and GnuGo that are also based on MCTS algorithms in an internal tournament. What's more, it won 5-0 against Fan Hui, a professional 2 dan.