Paper Review: "Mastering the game of Go with deep neural networks and tree search" by the DeepMind team.

AlphaGo utilizes Monte Carlo Tree Search (MCTS) using a policy to give values to different board configurations. It also employs a pipeline of three Machine Learning stages. First, it was trained with Supervised Learning using data from human games. It then improves upon itself using Reinforcement Learning, by playing Go games with itself. Finally, a value network was trained that predicts the winner of games played by the RL policy network against itself. AlphaGo then efficiently combines the policy and value network with MCTS.

Most notable results include:
- AlphaGo won 99.8% (out of 495 games) against other Go-playing programs.
- AlphaGo was more efficient than Deep Blue. In particular, during the match against Fan Hui, it searched some 1000x fewer positions than Deep Blue vs. Kasparov.
- AlphaGo' evaluation functions were trained through SL and RL rather than being handcrafted as in the case for Deep Blue.