

Research Review of “Mastering the game of Go with deep neural networks and tree search”

Andrew Cai

Paper’s goal:

The paper introduces a new approach to computer Go that use “value networks” to evaluate board positions and “policy networks” to select moves. The goal of the paper is to master the game of go and create a program named “AlphaGo”.

To reach the goal, the authors of the paper trains these deep neural networks with a novel combination of supervised learning from human expert game and reinforcement learning from games of self-play. The program uses search algorithm that combines Monte Carlo simulation with value and policy networks.

The SL policy networks have 13 layers and was trained from 30 millions positions from the KGS go Server. The RL policy networks are initialized to SL policy networks and trained by randomly selecting previous iteration of the policy network. Alpha Go then combines the policy and value networks in an MCTS algorithm that selects actions by lookahead search.

Paper’s result:

The distributed version of AlphaGo defeated a professional 2 dan, and the winner of European Go championships by 5 to 0. This is the first time that a computer Go program has defeated a human professional player, without handicap, in the full game of Go, a feat that was previously believed to be at least a decade away.

The single-machine AlphaGo is many dan ranks stronger than any previous Go programs, winning 494 out of 495 games against other Go programs. To provide a greater challenge to AlphaGo, it won 77%, 86% and 99% of handicap games against Crazy Stone, Zen and Pachi.