Research Paper March 21, 2017

Olya Flores, AIND Search

Summary of Mastering the Game of Go with deep neural networks and tree search

(https://storage.googleapis.com/deepmind-media/alphago/AlphaGoNaturePaper.pdf)

Traditional approaches to AI for playing games of Go included applying Monte Carlo search trees to reduce the breadth of the game tree, identifying probabilistically most advantageous sub-trees and then traversing a limited depth of those trees. However, by itself Monte Carlo algorithms were only able to deliver average-level Go gameplay (although they have been very effective in scrabble, chess and others).

The authors of the paper developed a novel approach where they trained convolutional neural networks to analyze Go game positions and evaluate them, and then they combined the neural net with a Monte Carlo implementation. This combination yielded impressive results, beating 99% of computer Go players as well as defeating a top human player 5-0.

The authors mention several techniques and approaches of interest. For example, rather than translating the Go board into a digital representation, they instead chose to represent it as an image, and use existing deep learning and image processing techniques to evaluate board states (using convolutional network layers to represent the player's position). Also, the authors separated policy from value calculations, training multiple networks: 1) a supervised learning policy network, trained from human expert player moves (also with a fast version of the network); 2) reinforcement learning policy network, trained on self-play games and 3) value network that predicts the outcomes of games played by network 2. All of these networks are then integrated in evaluating Monte Carlo Search Trees. The tree is traversed by simulation in lookahead search, and each time moves are picked according to their probability, but with a negative weight assigned for repeated visits (so you don't get stuck visiting the same nodes again and again).

The authors mention that, in addition to achieving 5-0 gameplay results, they also find that AlphaGo evaluates many fewer positions than algorithms like DeepBlue do in playing their games. It's able

to do so because it selects positions more intelligently, which brings it closer to the way humans tend to play these games.