Jan-Benedikt Jagusch

Master the game of Go with deep neural networks and tree search

Paper Review

Introduced Techniques

In games like Go, where breadth and depth of the game tree make it unfeasible to compute the outcome of all terminal leaves, it is critical to reduce search space. In such scenarios, you can apply two general principles: reducing the breadth or pruning the depth. Silver et al. introduce novel approaches for both variants, using convolutional deep learning, where the network input corresponds to a grid representation of the game state.

First, the breadth of the game tree can be reduced by only taking a sample of possible actions, determined by a policy. This is achieved by training a neural network with soft-max layer that returns the probability distribution of all subsequent game states. The network was initially trained in a supervised learning manner, where it learned to reproduce human exert moves. After achieving satisfying performance, the network was enhanced over a reinforcement learning process, where it focused on winning games, rather than copying moves.

You can prune the game tree's depth at a defined cutoff level, by replacing the subtree with an approximate evaluation function. Instead of designing such evaluation function manually, Silver et al. developed an artificial neural network that was trained to minimize the error between the prediction and the corresponding outcome.

Results

The resulting agent AlphaGo combined policy and value neural networks. It was tested against several other Go programs, where it won 99.8% of all games. Next, the program competed with Fan Hui, the European Go Champion of 2015. AlphaGo won the competition 5:0, with four resignations from Hui after the first match.