

The game of Go has long been viewed as the most challenging of classic games for artificial intelligence. The paper's goal is to create a powerful game agent with artificial intelligence.

The two key points of the problem is how to evaluating board positions and moves as well as how to select moves.

The paper uses 'value networks' to evaluate board positions and 'policy networks' to select moves. Both of them are deep neural networks trained by a combination of supervised learning from human expert games, and reinforcement learning from games of self-play.

First, build on prior work on predicting expert moves in the game of Go using supervised learning. The policy network is trained on randomly sampled state-action pairs (s, a) using stochastic gradient ascent to maximize the likelihood of the human move a selected in state s . In the paper, they trained a 13-layer policy network.

Second, using policy gradient reinforcement learning to improve the policy network.

Final stage of the training is to evaluate the positions. The paper approximate the value function using a value network with weights. This neural network has a similar architecture to the policy network, but the outputs are single prediction instead of a probability distribution. The weights of the value network is trained by regression on state-outcome pairs (s, z) , using stochastic gradient descent to minimize the MSE between predicted value and the corresponding outcome z .

The AlphaGo combines the policy and value networks in an MCTS algorithm that selects actions by search. Each edge (s, a) of the search tree stores an action value $Q(s, a)$, visit count $N(s, a)$ and prior probability $P(s, a)$. The tree is traversed by simulation, starting from the root state.

The result:

The AlphaGo is many dan ranks stronger than any previous program, winning 494 out of 495 games against other Go programs in the tournament. And defeated the human European Go champion by 5 games to 0. This is the first time that a computer program has defeated a human professional player in the full-sized game of Go.