Research Review

By

Ajay U. Makode

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

Goals:

The ultimate goal of the DeepMind Team was to find noval techniques that are successful for game with huge search space. The degree of massiveness is not defined as such but we can compare the search space with chess. A game of Go using a 19x19 board with branching factor of 250, a state complexity of 10^{170} and a game tree complexity of 10^{250}, while chess has much smaller values. Searching in such a huge search space in not feasible. To reduce the search space this paper briefly describes the two methods:

1. The depth of the search may be reduce by position evaluation
2. The breadth of the search tree may be reduced by sampling actions from a policy p(a-s) that is a probability distribution over possible moves a in position s.

The DeepMind team used artificial Intelligence to help solve the problems. The board position for each game is represented by a 19x19 pixels image and it is fed to convolution network to construct the position representation. The search space is reduced using a 'value network' for position evaluation and 'policy network' for moves sampling.

Policy and Value Network's training:

1. A policy network is trained using supervised learning directly on the human experts move. This provides the fast efficient learning updates with immediate feedback and high quality gradients.
2. The policy network is further trained using reinforcement learning which optimizes on the previous policy network.
3. Finally reinforcement learning value network is trained which predicts the winner of games player by reinforcement policy network.

Results:

AlphaGo defeated a human professional Go player for the first time in the history and decades in advance of public expectations. AlphaGo had 99.8% winning rate against other computer Go programs. DeepMind's research also revealed the level of computational power required to conquer such a task.

To challenge AlphaGo, we also played games with four handicap stones(that is, free moves for the opponent). AlphaGo won 77%,86% and 99% of handicap games against Crazy Stone, Zen and Pachi respectively. AlphaGo further competed against Fan Hui the European Championship winner for 2013, 2014 and 2015 and AlphaGo won the tournament as 5 games to 0.

Finally, We can conclude from the results of the tournament that AplhaGo is performing far better than Go players.