Research Review: Mastering the game of Go with deep neural networks and tree search[1]

By: Marie-Anne Valiquette

Goal

The ancient Chinese game of GO has simple rules: two players take turns to place black or white stone on a 19x19 board trying to capture the opponent's stones or surround empty space to make points. This game is considered one of the most challenging for artificial intelligence due to its massive search space, the board positions and moves. The game has a branching factor of 250, which means that a player is faces with a choice of approximately 250 possible moves and a search tree of 250^{150} .[2]

Methods

The paper, by DeepMind team, explains the technical details behind AlphaGo which is a computer agent that was able to beat the world's champion by using:

- 'Value networks' as an advanced heuristic function to evaluate board positions
- 'Policy networks' as an advanced version of alpha-beta pruning or minimax algorithm to select moves
- Training these 2 deep neural networks using supervised learning from human expert games
- Training these 2 deep neural networks using reinforcement learning from games of self-play
- An alternative approach to searching the game tree, Monte Carlo Tree Search, is used. Many game simulations run at the same time, each of these simulations starts at the current game state and stops when the game is won by one of the two players.
- A combination of Monte Carlo simulation and the 2 deep neural networks, 'value and policy networks', to create a new search algorithm

To execute the simulations AlphaGo uses 40 search threads, 28 CPUs, and 8 GPUs. The DeepMind team also implemented a distributed version of AlphaGo that exploited multiple machines, 40 threads, 1 202 CPUs and 176 GPUs.[1]

Conclusion

In conclusion, for the first time in history a human professional Go player was defeat by a computer agent called AlphaGo. The computer agent relies on the combination of a tree search procedure and convolutional networks that guides the tree search procedure. Also, AlphaGo uses reinforcement learning by playing against itself instead of using hand-craft heuristics. Finally, the ways that the DeepMind team took advantage of the development of deep neural network, formatted the training data and trained the system was impressive and gives access to other field of applications.

---

[1] Mastering the game of Go with deep learning neural networks and tree search
https://storage.googleapis.com/deepmind-media/alphago/AlphaGoNaturePaper.pdf

[2] Google DeepMind's AlphaGo: How it works
https://www.tastehit.com/blog/google-deepmind-alphago-how-it-works/