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

Joao Paulo Fernandes Ventura

SUMMARY

In perfect information games (ex. chess, Go, isolation), it is possible to execute a recursive tree search function *v(s)* to determinate the game outcome for a given of the board state *s*.

As previously mentioned during the course, being b the sequence of all possible moves for a given board state, and D the game length, the search tree will contain b^d possible moves.

Eventhough chess piece moves are many in quantity and complexity, there are 35^{80} possible moves. However with much simpler rules an pieces, Go reaches astonishing 250^{150} (i.e., more atoms than exist in the whole universe). Thus exhaustive search is infeasible.

Although it was possible beat master chess players using *v*(s),* where the function result was replaced by approximations (allowing reasonable computation time) it was believe that such approach was not possible to Go.

The study breakthrough was create an estimate function *v*(s)* combining Monte Carlo Tree Search (MCTS) with deep convolutional neural networks (created with supervisioned training) to reduce the tree search depth.

The deep convolutional neural network (DCNN) training consisted in three major steps:

- Supervised learning of policy networks
- Reinforcement learning of policy networks
- Reinforcement learning of value networks

RESULTS

AlphaGo was first compared to several commercial programs such as Crazy Stone. Zen, Pachi and Fuego, which are based on strong MCTS algorithms with a 5s timeout per move.

Without a handicap, AlphaGo won 99.8% of all maches, while won 77%, 86%, and 99% of handicap game against Crazy Stone, Zen and Pachi, respectively.

However it is know that the best AI players were not match to professional Go players, so AlphaGo was also tested with Fa Hui (a professional 2 dan, and the winner of the 2013, 2014 and 2015 European Go championships).

Over 5–9 October 2015 AlphaGo an Fan Hui competed in a formal five-game match. AlphaGo won the match 5 games to 0.

As Deep Blue vs Kasparov [2] was an historical achievement in 1994, AlphaGo was the first system to defeat a human professional player, without handicap, in the full game of Go.

Even more important is the fact that scientists strongly believed that it would be only possible with the computational power of a decade in the future.

This is a strong evidence of how much powerful deep learning is and how astonishing results may be obtained with supervised training with an adequate training data set.

Now It is also believed that deep learning will be capable of solve other challeging decision-making task, an intractable search space, and optimal solution, spawning even harder AI challenges such as StarCraft [3] and DOTA 2 [4].

REFERENCES

[1] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel & D. Hassabis 2016. Mastering the game of Go with deep neural networks and tree search. Nature International Journal of Science. http://dx.doi.org/10.1038/nature16961

[2] Murray Campbell, A. Joseph Hoane, Jr., and Feng-hsiung Hsu. 2002. Deep Blue. Artif. Intell. 134, 1-2 (January 2002), 57-83. DOI=http://dx.doi.org/10.1016/S0004-3702(01)00129-1

[3] DeepMind and Blizzard open StarCraft II as an AI research environment, https://deepmind.com/blog/deepmind-and-blizzard-open-starcraft-ii-ai-research-environment/

[4] Open AI reveals self-play information after successful Dota 2 test, https://www.teslarati.com/openai-self-play-dota-2-musk/