

AlphaGo Zero

The AI That Has Nothing to Learn From Humans

(*The Atlantic*)

Bertrand Louargant

Udacity AI Nanodegree

<https://deepmind.com/blog/alphago-zero-learning-scratch/>

1. Forewords

Although the subject initially proposed was about alphago, the recent publication of Deep Mind concerning AlphaGo Zero, its successor, cannot be ignored, especially in light of the meaning of “Zero” in it’s name ...

That’s why this research will concentrate on Alpha Go zero and try to explain why this “Zero” is so important.

2. Introduction

AlphaGo Zero is the latest development of Google DeepMind’s Go AI player. While AlphaGo, it’s predecessor, was able to beat the world’s best Go player - something thought to be impossible just a few years ago - this new version is way ahead of it’s venerable ancestor not only because it was able to beat it to a final score of **100** to **0** but also by the ressources it needed to do so.

3. Final Score: 100 to 0

Let’s go back a little to the time when Alpha Go beated the - now - famous Lee Sedol, world champion of Go. At that time, AlphaGo showed moves that nobody has seen before which is pretty impressive given the fact that the game of Go is nearly four thousand years old.

Moreover, the game of Go offer such a large search space that it is completely possible that a less capable AI is lucky enough to win over a better AI.

So, is it possible that AlphaGo Zero was able invent new bizare moves ?

That’s seems to be the case, a fact somehow backed by the scarcity of ressources needed compared to it predecessor.

4. Always Less, Less of Everything.

The second thing that strikes about AlphaGo Zero is that it needs so much less resources, it needs less of just everything.

- **Less computation power**

With only 4 Google TPUs needed, this is less than 1/10th of the 48 TPUs needed by the previous version. This alone is a serious hint that something big is happening here.

Note that with 4 Nvidia Volta GPUs, like the DGX-1 Workstation, one has enough computation power to replicate the experiment. If this is not cheap it's an order of magnitude cheaper than a super computer and also way more accessible.

- **Less training time**

After only three days of training it was able to defeat AlphaGo "Lee", the version that won against Lee Sedol in March 2016. After 21 days, it reaches the AlphaGo "Master" level, a version of AlphaGo that defeated the world's best players and world number one Ke Jie. After 40 days, it outperformed any other version of AlphaGo.

- **Less training data**

Usually, the way to improve a neural network is either a bigger database or to generate more synthetic data by doing all sorts of data manipulations. However for AlphaGo Zero, DeepMind used what they call "Self-play reinforcement learning".

More on this in the next section.

5. The AI That Has Nothing to Learn From Humans

Previous versions of AlphaGo were heavily relying on human games to learn of to play Go. But AlphaGo Zero just skips this step and directly start by playing against itself.

If the first games are completely random, in just 3 hours it learns how to play like a human beginner and 16 hours later, it has acquired the base of more advanced strategies. After 70 hours of training it plays at super-human level. To achieve this DeepMind used a novel form of reinforcement learning where AlphaGo Zero is its own teacher.

With each iteration of self-play, the system learns to become a stronger player by exploiting the results of the Monte Carlo Tree Search (MCTS) results of the previous game winner. The recombination of the updated neural network and the search algorithm results in a better version of itself.

6. Last Words

By creating an AI capable of learning from scratch without prior human knowledge of a game as complex as the game of Go, DeepMind has demonstrated the power of its new approach to reinforcement learning and bode well for this branch of Deep Learning and for Artificial Intelligence in general.