A Critical Analysis of AlphaGo

Electrical and Electronic Engineering
(general stream)
Shi Yuhan(3035540736)
14/2/2019

AlphaGo is a computer program that plays the board game Go. It was developed by Alphabet Inc.'s Google DeepMind in London.[1] It was the first computer program to defeat professional human chess players and the first to defeat the world champion of Go. The main principle of Alpha Go is deep learning.

In October 2015, AlphaGo Fan defeated the French National Go Team Head Coach Fan Hui, European Go champion, five to zero. This was the first time that a computer program has defeated a human professional player in the full-sized game of Go. In March 2016, AlphaGo Lee competed with Lee Sedol and won by 4:1. Lee, a 9 dan pro, is the world champion and had won 18 international titles at that time. At the end of 2016 and the beginning of 2017, the program used “Master” as its registered account on Chinese chess websites to compete with dozens of top Chinese, Japanese and Korean Go players. It won in 60 consecutive games without a defeat. In May 2017, AlphaGo Master beat Ke Jie who was the world champion at that time by 3:0 at The Future of Go Summit in Wuzhen. It is generally acknowledged that Alpha Go has exceeded the top level of professional Go players.

On October 18, 2017, DeepMind team released the strongest version of Alpha Go, named AlphaGo Zero. After just three days of self-training, Alpha Go Zero beat Alpha Go Lee with a record of 100:0. After 40 days of self-training, Alpha Go Zero defeated the Alpha Go Master version.

The main principle of AlphaGo is "Deep learning”. "Deep learning" refers to a multi-layered artificial neural network and a method of training it. The input of the neural network is features extracted from a large amount of sample data. Each layer of the neural network assigns different weights to the features. After training by the multi-layer neural network, another data set is finally generated as an output by the nonlinear activation method. The output of the network will be evaluated and the results will be fed back to the network to adjust the parameters for better results. Deep learning discovers attribute categories by combining low-level features to form more abstract high-levels, thereby discovering distributed feature representations of data.

Alpha Go combine Monte Carlo Search Tree with Policy and Value networks. These deep neural networks are trained by a novel combination of supervised learning from human expert games, and reinforcement learning from games of self-play. The Monte Carlo Tree Search connects these parts together to form a complete system.

First, it uses a lot of human expert samples to train a Supervised Learning policy network to get the probability density distribution of the human expert movements in current status, so as to predict the next maximum probability action. Similar to a prior work, it can reduce the complexity of the system quickly. At the same time, using the Rollout policy which is a small network used to quickly analog entire games to train a fast policy. Second, it trains a Reinforcement Learning policy network by optimizing the output of the network to improve the supervised network. The output of the network is adjusted to the gain (+1) or the debt (-1) of the target (winning the game) instead of improving the accuracy of the prediction. Third, using a Reinforcement Learning network, it trains a value network. Let Alpha Go play against itself to predict the winners in current status. Finally, during the repetitive training, the system will adjust the parameters to get the optimal next move.[2] Since this processor has a lot of random elements, it is impossible for people to know exactly how the network "thinks", but more training can make it evolve better.

In previous versions, Alpha Go learned a lot from human expert games. It combines Convolutional Neural Network and Monte Carlo Search Tree by using two networks. One of them is policy network which is used to predict the next move. The other one is value network which is used to evaluate the winner in the current state. A new version of Alpha Go names Alpha Zero. AlphaGo Zero is trained completely by Reinforcement learning, from the beginning of the random game without any supervision or human expert data. It uses only the black and white sub-boards as input features (the previous AlphaGo has many features built artificially). What’s more it uses only one neural network, not a separate policy network and value network. The input of the network is the current layout. The output is the current state's equity and the location and probability of the next action.[3]

Google has been trying to use artificial intelligence in a variety of different fields to achieve General AI. Real General AI is a system that demonstrates the ability to reason, understand, and complete complex tasks at human level. Google’s vision of General AI is to evolve calculation from mobile first to artificial intelligence first (AI first). Nowadays, each of us carries a computing device with us. In the future, this device may become an assistant to our work and life. It has the wisdom of human beings to help us get more information and solve more problems.

The chess or other kinds of games has their own rules to follow. It is easier to learn. But in the real world, the reality environment is much more complicated than the game. The limitation of general artificial intelligence is that it can't automatically summarize and generate new solutions from the surrounding environment. Although Alpha Go system developed by Deep Mind seems to be powerful. But there is still a long way to go to implement general artificial intelligence.

[1] En.wikipedia.org. (2019). AlphaGo. [online] Available at: https://en.wikipedia.org/wiki/AlphaGo [Accessed 14 Feb. 2019].

[2] Silver, D., Huang, A., Maddison, C., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T. and Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. Nature, 529(7587), pp.484-489.

[3] Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T. and Hassabis, D. (2017). Mastering the game of Go without human knowledge. Nature, 550(7676), pp.354-359.