The game of Go was a long time challenging for Aritificial Intelligence. The AI overcomed the chess at 1977. But even roughly 40 years path after winning the chess world champion the AI still can't overcome the Go. Because there was a big problem. The problem was the search space was too big. Chess had about 35 breadth and 80 depth at tree search but Go has about 250 breath and 150 depth. So the exhaustive search was infeasible. Deepmind team solve this problem by truncating the breath and death both. So the Go AI AlphaGo can search the tree more efficiently and effectively.

First, they need to truncate the breath. When the Go is in state S, we need to choose the position A. But we can't choose every single position. We need to limit the position and there need to be a reasonable reason why we cut off the others and choose some of the position in concern. Deepmind team limit the position where seems the human expert might move. With 30 million positions data in KGS Go Server they trained CNN(Convolutional Neural Network) with SL(Supervised Learning). This CNN called SL policy network. And CNN called rollout policy using a linear softmax of small pattern features that trained faster but less accurate was also made.

After that they made another CNN that has identical structure to the SL policy network. And by self-play they used RL(Reinforcement Learning). RL policy network won more than 80% of games against the SL policy network.

Second, they truncate the depth with the CNN called Value network. Value network evaluate 'Will the game will win or lose'. So we can estimate 'who will win the game when the state is S and we choose position A' and this means we don't need to go through the end of the game to get the results. So we truncate the depth of the search tree. Value network is based on RL policy network architecture, but outputs a single prediction instead of a probability distribution. As same as policy network Value network also used RL in self-play. This RL was not just for performance it was also to prevent overfitting.

With this two solutions the search space reduced feasible to search.

And the last they used MCTS(Monte Carlo Tree Search) algorithm that selects actions by lookahead search. Policy network show the legal move and Action value Q plus u(P) that depends on a stored prior probability P for that edge choose where to go. And when traverse enough depth, it evaluate the node with value network and fast rollout policy. After evaluation they back up and update the Q value and the visit count for each node. This process repeated many times. Once the search is complete, the algorithm choose the most visited move from the root position.

With this new algorithms AlphaGo achieved a 99.8% winning rate against other Go programs, and defeated the human European Go champion by 5 games to 0. Also distributed version of Alpha go, it defeated the human World Go champion LEE SEDOL by 5 games to 1. Beating human professional player in the full-sized game of Go previously thought to be at least a decade away but it was not. Exponential growth in AI field debunk the previous predictions. Amazing future is coming.