ALPHAGO

The game of Go, the most challenging classic game for an AI to beat expert player. AlphaGo archived 99.8% winning rate against other Go programs, and defeated the human European Go champion by 5 games to 0.This is the first time, a computer program has defeated a human professional player in the full-sized game of Go.

AlphaGo team used recent machine learning technique to train the game agent. They pass the board position as 19x19 image and use convolutional layers to construct a representation of the position. They use neural networks to reduce the effective depth and breadth of the search tree: evaluating positions using a value network, and sampling actions using a policy network.

They begin by training a supervised learning (SL) policy network directly from expert human moves. This provides fast, efficient learning updates with immediate feedback and high-quality gradients. Next, they train a reinforcement learning (RL) policy network that improves the SL policy network by optimizing the final outcome of games of self-play. This adjusts the policy towards the correct goal of winning games, rather than maximizing predictive accuracy. Finally, they train a value network $v\theta$ that predicts the winner of games played by the RL policy network against itself.

To evaluate AlphaGo, they ran an internal tournament among variants of AlphaGo and several other Go programs, including the strongest commercial programs Crazy Stone and Zen, and the strongest open source programs Pachi and Fuego. They included the open source program GnuGo, a Go program using state-of-the-art search methods that preceded MCTS. All programs were allowed 5s of computation time per move.

Finally, they evaluated the distributed version of AlphaGo against Fan Hui, a professional 2 dan, and the winner of the 2013, 2014 and 2015 European Go championships. Over 5–9 October 2015 AlphaGo and Fan Hui competed in a formal five-game match. AlphaGo won the match 5 games to 0.

The AlphaGo team have developed a Go program, based on a combination of deep neural networks and tree search, that plays at the level of the strongest human players, thereby achieving one of artificial intelligence's "grand challenges"31–33. They have developed, for the first time, effective move selection and position evaluation functions for Go, based on deep neural networks that are trained by a novel combination of supervised and reinforcement learning. They have introduced a new search algorithm that successfully combines neural network evaluations with Monte Carlo rollouts. AlphaGo integrates these components together, at scale, in a high-performance tree search engine.

During the match against Fan Hui, AlphaGo evaluated thousands of times fewer positions than Deep Blue did in its chess match against Kasparov; compensating by selecting those positions

more intelligently, using the policy network, and evaluating them more precisely, using the value network—an approach that is perhaps closer to how humans play. Furthermore, while Deep Blue relied on a handcrafted evaluation function, the neural networks of AlphaGo are trained directly from gameplay purely through general-purpose supervised and reinforcement learning methods.

By combining tree search with policy and value networks, AlphaGo has finally reached a professional level in Go, providing hope that human-level performance can now be achieved in other seemingly intractable artificial intelligence domains.