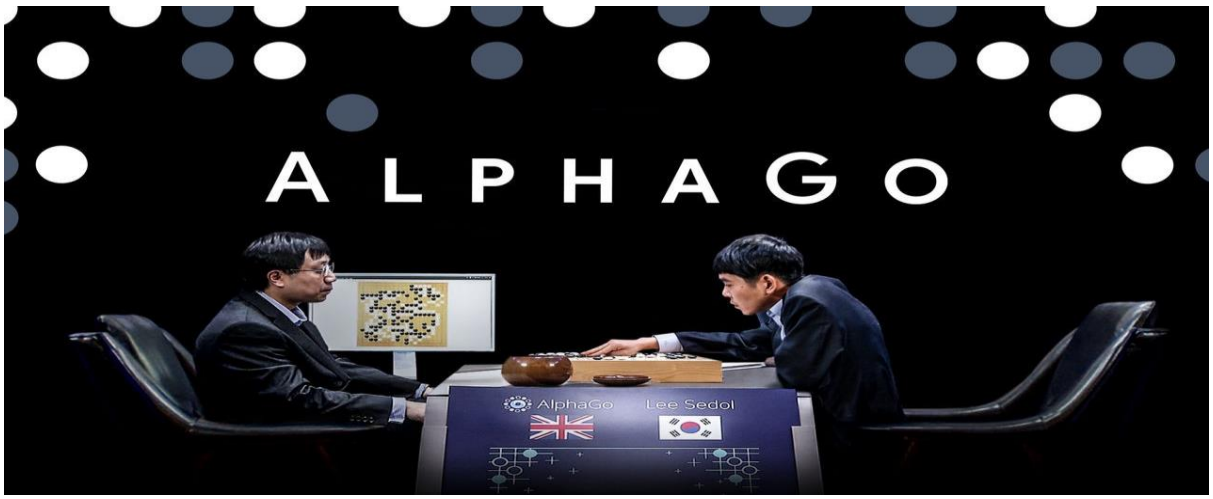
Task 9.1: Advanced intelligent system

Go is an adversarial game with the objective of surrounding a larger total area of the board with one's stones than the opponent. As the game progresses, the players position stones on the board to map out formations and potential territories. Contests between opposing formations are often extremely complex and may result in the expansion, reduction, or wholesale capture and loss of formation stones.

Go is played on a square grid of lines, typically 19 intersections to a side. Other common board size is 13*13 and 9*9 boards. player uses a set of stones typically coloured black and white. The objective of the game is to surround more territory than your opponent. There are four main rules to adhere to. The first rule is to set the flow of game i.e. in a standard match black begins, each player takes a turn to place a single stone onto an intersection. Stones that are placed on the board stay on the respective intersection and not moved unless captured by opponent. The second rule deals with how stones are captured, the empty intersections next to the stone are called liberties. A single stone in the middle of the board has four liberties, a stone on the side has three liberties and stone on the corner has two liberties. Stones that have all liberties occupied by stones of opposite colour are removed from the board and taken by opponent as captures. Connected stones of the same colour share the liberties and if captured are captured as a group. A group that is under thee threat of being captured or in other words having only one liberty. The third rule is co rule accordingly a player may not immediately replicate a previous board position. It leads to what is known as ko fight is one of the most difficult part of the game. The fourth rule deals with how the game ends and how it is scored. Instead of playing a stone a player can pass which is done when the player feels there is no more available worthwhile rules. The game ends when each player pass in success. The stone are immediately removed as captures. Territories are calculated for each player by the number of intersections surrounded by their stone. Point counts as territory when they only reach stones of singe players colour. Each player score is also reduced by the number of points equal to the number of stone that was captured by their opponent. It is to be noted scoring system vary based on the method of rule set but doesn't alter the outcome of the game. And the player with higher resulting score wins the game.

AlphaGo is the first computer program to defeat a professional human Go player, the first to defeat a Go world champion, and is arguably the strongest Go player in history.



DeepMind created AlphaGo, a computer program that combines advanced search tree with deep neural networks. These neural networks take a description of the Go board as an input and process it through several different network layers containing millions of neuron-like connections. One neural network, the “policy network”, selects the next move to play. The other neural network, the “value network”, predicts the winner of the game. AlphaGo was introduced to numerous amateur games to help it develop an understanding of reasonable human play. Then AlphaGo played against different versions of itself thousands of times, each time learning from its mistakes. Over time, AlphaGo improved and became increasingly stronger and better at learning and decision-making. This process is known as reinforcement learning.

AlphaGo architecture for the game of Go starts by passing in the board position as a 19×19 image and use convolutional layers to construct a representation of the position. Then use these 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. It then trains the neural networks using a pipeline consisting of several stages of machine learning. Then it begins by training on a supervise learning (SL) policy network directly from expert human moves which provides fast, efficient learning updates with immediate feedback and high-quality gradients. It also trains a fast policy that can rapidly sample actions during rollouts. Next, it trains on a reinforcement learning (RL) policy network that improves the SL policy network by optimizing the outcome of games of self-play. This adjusts the policy towards the correct goal of winning games, rather than maximizing predictive accuracy. Finally, its trained on a value network that predicts the winner of games played by the RL policy network against itself. The program AlphaGo efficiently combines the policy and value networks with MCTS (Monte Carlo Tree Search). To efficiently integrate large neural networks into AlphaGo, it implemented an asynchronous policy and value MCTS algorithm (APV-MCTS).

AlphaGo went on to defeat Go world champions in different global arenas and arguably became the greatest Go player of all time.

“I thought AlphaGo was based on probability calculation and that it was merely a machine. But when I saw this move, I changed my mind. Surely, AlphaGo is creative” (*LEE SEDOL WINNER OF 18 WORLD GO TITLES*)

References:

I. 2020, RETRIEVED 23 MAY 2020, <HTTPS://YOUTU.BE/UVTRWWZUYBO>.
II. 2020, STORAGE.GOOGLEAPIS.COM, RETRIEVED 23 MAY 2020, <HTTPS://STORAGE.GOOGLEAPIS.COM/DEEPMIND-MEDIA/ALPHAGO/ALPHAGONATUREPAPER.PDF>.
III. "GO (GAME)" 2020, RETRIEVED 23 MAY 2020, <HTTPS://EN.WIKIPEDIA.ORG/WIKI/GO_(GAME)>.
IV. 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 & HASSABIS, D 2016, "MASTERING THE GAME OF GO WITH DEEP NEURAL NETWORKS AND TREE SEARCH", NATURE, VOL. 529, NO. 7587, PP. 484-489.
V. 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 & HASSABIS, D 2017, "MASTERING THE GAME OF GO WITHOUT HUMAN KNOWLEDGE", NATURE, VOL. 550, NO. 7676, PP. 354-359.