AI agents are designed to play games by exploring the possible states for the next move in a way to maximize agent's own utility from the move and minimize the opponent's utility. The complexity of search problem is determined based on the branching factor i.e. the number of possible immediate next moves and the depth of the search graph i.e. levels to look ahead in the game with each level alternating between agent and its opponent. AI agent developed as part of the Isolation game is an implementation of Game Theory. AI agent developed for Isolation game looks ahead a certain depth of game using MiniMax algorithm and Iterative deepening. In the Minimax algorithm the agent predicts the next move by considering its own and opponent possible moves for next d level of depth. In the iterative deepening the algorithm continue the exploring depth of search graph till the pre-set timer runs out. The complexity of an AI agent increases exponentially even with a linear increase of branching factor and depth. Therefore games such as Chess, Go remain computing challenge and pushes the frontier of artificial intelligence.

The game of Go is incredibly complex because the number of possible states which increases the search space to identify the best next move for an agent. The researchers at Google developed an AI agent for Go that combined deep neural network and tree search capabilities to intelligently select the most efficient next move. The deep neural networks were trained by a combination of supervised learnings from human experts and reinforcement learnings from the self-play. Prior to Alpha Go implementation, AI agents for the game of GO were based on Monte Carlo tree search (MCTS). MCTS technique uses simulations to predict the value of each state in a search tree.

Alpha GO implementation combines deep neural networks and tree search for optimal move selection and position evaluations functions for GO. Alpha Go training architecture applied many layers of neurons in a convolutional neural networks by representing the board position as a 19 X 19 image. Neural network was used to reduce the effective depth and breadth of the search tree. Convolutional neural networks work by many layers with each layer focusing on a localized representation of the problem space. The first stage of Alpha Go training architecture was built upon prior work of predicting expert moves using supervised learnings (SL). The SL policy network alternates between convolutional layers and rectifier nonlinearities. A softmax function outputs the probability distribution of all the legal moves. This 13-layer policy network called SL policy network was trained from 30 million positions. The second stage of training pipeline aims at improving the policy network by policy gradient reinforcement learning (RL). RL policy network won more than 80% of the games against the SL policy network. The final stage of training pipeline focuses on position evaluation, estimating a value function that predicts the outcome from position of games played by using policy for both players.

AlphaGo combines the policy and value networks in an MCTS algorithm that selects action by lookahead search. The search is traversed using a simulation starting from the root state. At each time step of each simulation, an action is selected from state so as to maximize action value plus a bonus that is proportional to the prior probability but decays with repeated visits to encourage exploration. At the end of each simulation, the action values and visit counts of all traversed edges are updated. Once the search is complete, the algorithm chooses the most visited move from the root position.

An enormous computing power is required for Alpha Go agent compared to the traditional search heuristics. To efficiently combine MCTS with deep neural networks, Alpha Go uses an asynchronous multi-threaded search that executes simulations on CPU and computes policy and value network on GPUs.

Alpha Go uses Alpha Go program defeated a professional human player by 5-0. This is the first time a computer program defeated a human player beginning a new era for the field of artificial intelligence.