Summary of AlphaGo by the DeepMind Team: I am summaring the AlphaGo, a computer program which is developed by Google's DeepMind. AlphaGo has defeated Lee Sedol – the second highest-ranking professional Go player, in a five game competition and AlphaGo won with score 4-1.

Summary of Techniques:
Monte Carlo Tree Search (MCTS): MCTS is an alternative simulation based search algorithm to minimax algorithm to searching the game tree. With games like Go, it is very challenging to apply minimax search algorithm, as there 10^{761} possible games are possible. Each simulation starts at the current game state and stops when the game is won by one of the two players. At first, the simulations are completely random: actions are chosen randomly at each state, for both players. At each simulation, some values are stored, such as how often each node has been visited, and how often this has led to a win. These numbers guide the later simulations in selecting actions. With more and more simulations, it selects winning moves and converges with optimal games.

Deep convolutional networks: Deep convolutional networks are a sub-type of neural networks, which are mimicking the structure of Brain neurons, and there has been lot of success in processing of image data. These networks take image data as input.
hu
Deep reinforcement learning: The Deep reinforcement learning is the combination of deep learning and reinforcement which will able to create artificial agents to achieve human-level performance. Reinforcement learning is that agents can learn for themselves to achieve successful strategies that lead to the greatest long-term rewards. This paradigm of learning by trial-and-error, solely from rewards or punishments, is known as reinforcement learning (RL).

The AlphaGo has three convolutional neural networks are trained which are two policy networks and one value network, in which both types of networks take image as input.

Value network: The value network provides an estimate of the value of the current state of the game which is the probability of a player to win the game, given the current state. The input to the value network is the whole game board, and the output is a single number, which is the probability of a winning the game.

Policy network: The policy networks provide guidance regarding which action to choose, given the current state of the game. The output is a probability value for each possible legal move and output of the network is the board. Actions/steps with higher probability values correspond to actions/steps that have a higher chance of winning the game.

Summary of results:

AlphaGo has evaluated by conducting tournaments with several other other Go programs like Crazy Stone, Zen, Pachi and Fuego and result suggest that AlphaGo is many dan ranks stronger than any of the listed Go programs.

The most commonly used system to compare the strength of Go players in Elo rating system. The player with higher rating will be defined based on his chance of winning games which indicates player with higher rating is most likely win games.

AlphaGo with distributed version reported with 3140 Elo ratings and all ratings are reported as below.