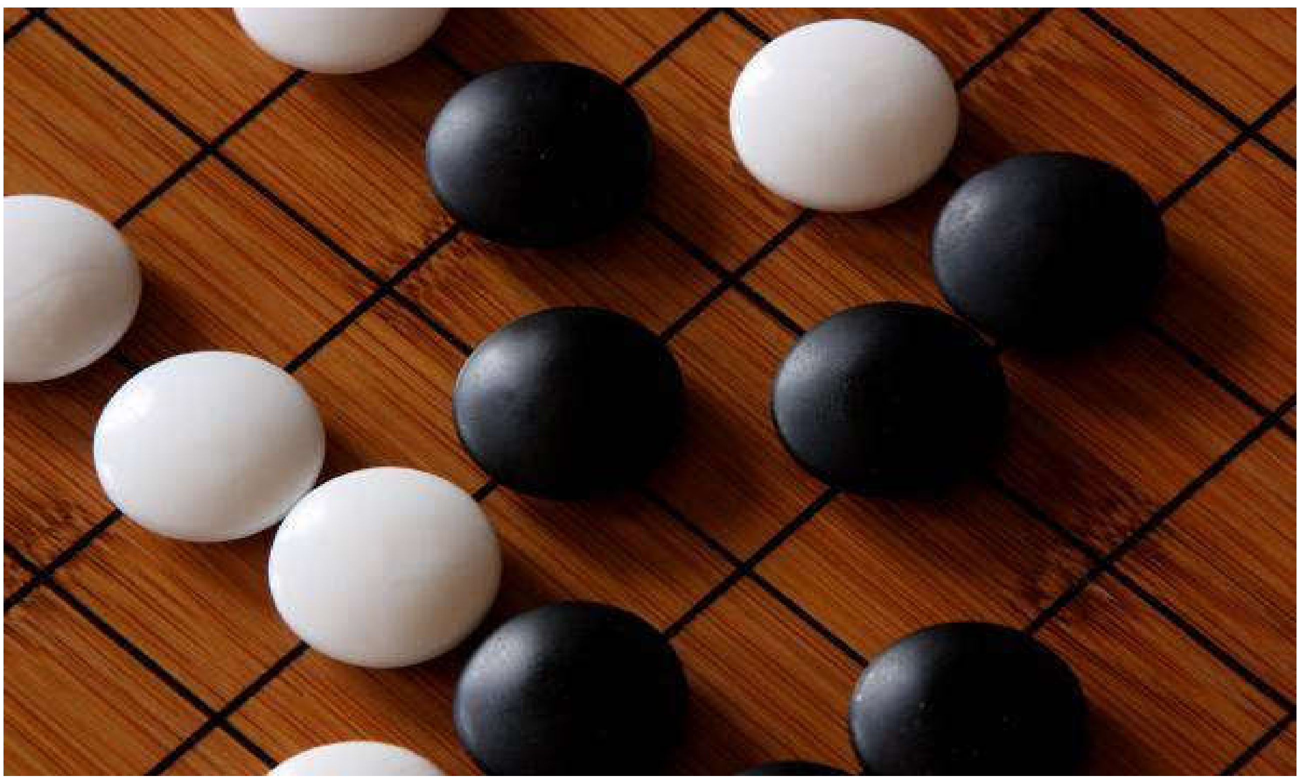
RESEARCH REVIEW

Mastering the game of Go with Deep Neural Networks and Tree Search



Submitted by Vinod Krishna

As part of ***Artificial Intelligence Nano Degree, Udacity***

TABLE OF CONTENTS

WHY GAME OF GO IS MOST CHALLENGING GAME FOR AI

Game Go is a perfect information game; as with every perfect information game; an optimal value functions exist which determines the outcome of the game, for every board position or state, under perfect play by all players. Traditional AI methods, constructs a search tree over all possible board positions. The challenge is, the enormity of search space, evaluation of board positions and moves with traditional AI methods. For a quick reference, the search space has 10^{170} possible board positions – more than the number of atoms in the known universe10^{80}.

ALPHAGO AND ITS ACHEIVEMENTS

AlphaGo's greatest achievement is to play the game in an approach that is perhaps closer to how humans play. AlphaGo has reached a full professional level in Go game play. This achievement comes with a hope that human-level performance can now be achieved in other artificial intelligence domains. Moreover, to achieve such levels of game play by utilizing standard general-purpose supervised and reinforcement learning methods to train the neural networks of AlphaGo, purely from gameplay; provides added hope in solving intractable AI problems.

Some highlights:

- *AlphaGo achieved a 99.8% winning rate against other Go Programs.*
- *AlphaGo defeated human European Go Champion Fan Hui (5-0)*
- *AlphaGo defeated human World Go Champion the legendary Lee Sedol*

NOVEL TECHNIQUES

- *Use of self-play to train the policy and value networks.*
- *A Search algorithm that combines policy and value network evaluations with Monte Carlo rollouts.*

OVERVIEW OF THE IMPLEMENTATION

The AlphaGo uses a pipelined stages of machine learning to train its neural networks. The First stage is training a policy network using Supervised Learning [SL] based on expert human moves. The Second stage is the Reinforced Learning [RL] deep neural network that improves SL policy network based on self-play. This adjusts the policy towards winning games. Finally a Value deep neural network that predicts the winner of the games played by the RL policy network against itself. Monte Carlo Tree search algorithm ties both policy and value networks together.