AlphaGo - research review

Goals

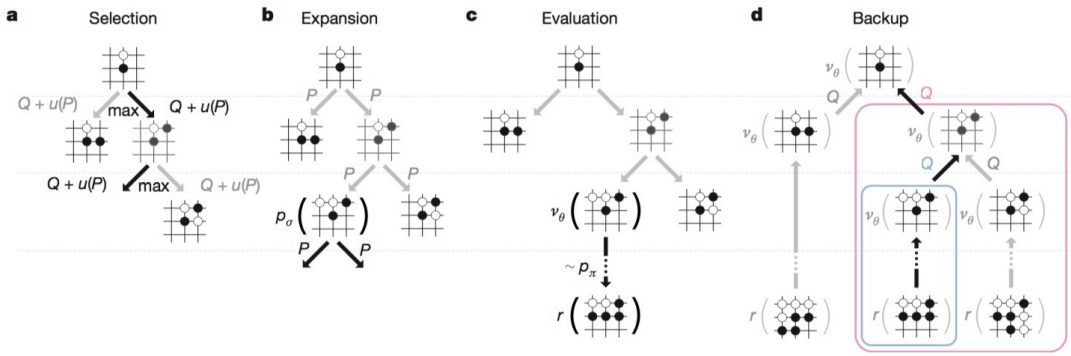
The goal was to improve the current state of the art in IA, using the 'most challenging of classic games', Go. Achieving the feat of beating the human European Go champion by 5 games to 0, a feat though to be still a decade into the future.

The innovation introduced was based on a paradigm shift from solo Monte Carle tree search (MCTS), technique used currently in the strongest Go programms to a combination with deep convolutional neural network in order to reduce the effective depth and breath of the search tree.

The step of using a deep learning approach was based on:

- 1.a Train a simple, low effective yet extremely fast "rollout policy" network.
- 1.b Train a complex, more accurate " policy network" (44,4%, double than the 'rollout' one) thus, slower.
- 2.a Improve the "policy network" by the use of reinforcement learning.
- 3.a Reiforcement learning for the "value networks".

Once the networks where trained, they were combined using a MCTS algorithm.



Results

They remarked that the SL policy network performed better than the stronger RL policy, being this effect provoked in theory by the SL being trained on human moves (humans tend to chose a beam of promising moves) rather than the RL network, which focused on the best move available.

To asses the strenght of AlphaGo, an internal tournament using several modifications of AlphaGo and other Go programms. Results suggest the single-machine AlhpaGo is many dans (Go mastery scale), above previous Go programms.

AlphaGo is an example of a challenging decision-making task, an intractable search space, and an optimal solution so complex it appears infeasible to directly approximate using a policy or value function.