**Review of Mastering the game of go**

**Summary**

The game of Go is the most challenging domains in terms of human intellect, and widely viewed as a grand challenge for artificial intelligence. It require a precise and sophisticated lookahead in vast search spaces. AlphaGo (Fan)was the first g]program to achieve superhuman performance in Go, which used two deep neural networks: a policy network that outputs move probabilities and a value network that outputs a position learning to accurately predict human expert moves, and was subsequently refined by policy-gradient the winner of games played by the policy network against itself.

AlphaGo Zero differs from AlphaGo Fan and AlphaGo Lee in several important aspects.

**Contributions**

① AlphaGo Zero trained solely by self-play reinforcement learning, starting from random play, without any supervision or use of human data.

② It uses only the black and white stones from the board as input features.

③ It uses a single neural network, rather than separate policy and value networks.

④ They have deployed effective move selection and position evaluation functions for Go, based on deep neural networks that are trained by a novel combination of supervised and reinforcement learning.

⑤ Introduced a new search algorithm that successfully combines neural network evaluations with Monte Carlo rollouts.

⑥ Integrates these components together, at scale, in a high-performance tree search engine.

**Comments**

AlphaGo evaluated thousands of times fewer positions than Deep Blue did in its chess match against Kasparov; compensating by selecting those positions more intelligently, using the policy network, and evaluating them more precisely, using the value network.

Their results comprehensively demonstrate that a pure reinforcement learning approach is fully feasible, even in the most challenging of domains: it is possible to train to superhuman level, without human examples or guidance, given no knowledge of the domain beyond basic rules. AlphaGo Zero defeated the strongest previous versions of AlphaGo, which were trained from human data using handcrafted features.