Reading Report: *Mastering the Game of Go without Human Knowledge*

Summary

In the paper *Mastering the Game of Go without Human Knowledge*, the authors introduce AlphaGo Zero, a novel Go-playing system that surpasses all previous versions of AlphaGo. Unlike earlier iterations, AlphaGo Zero discards all human expert knowledge and learns entirely from self-play using reinforcement learning. The key innovation lies in using a single deep neural network to simultaneously predict the move probabilities (policy) and the game winner (value), coupled with Monte Carlo Tree Search (MCTS) to guide exploration. This approach enables AlphaGo Zero to achieve superhuman performance in a matter of days, outperforming previous versions trained with human data. The system demonstrates that it is possible to reach a high level of strategic understanding without relying on human priors, challenging the traditional paradigm of incorporating expert data into AI systems.

Questions

Q1. To what extent can the self-play reinforcement learning method used by AlphaGo Zero be transferred to real-world problems?
Unlike Go, many real-world tasks lack clearly defined rules or a reliable environment simulator. Can a self-play framework be adapted for such domains, or are there inherent limitations?

Q2. Does AlphaGo Zero's training paradigm indicate a shift away from supervised learning toward a more generalized AI framework?
AlphaGo Zero achieves higher performance without any human data. Should we expect future AI systems to rely less on human knowledge, or are there scenarios where supervision remains indispensable?