University of South Wales Prifysgol De Cymru

Understanding Mind of the Machine [AlphaZero's Chess Mastery]

Comparing and Contrasting — Man Vs. Machine

Student Name: Salma Javid
Student ID: 30107961

♔ Introduction — The Case of Chess ♕

The Chess game has been a grand challenge to humans since the dawn of computing age with early pioneers including Babbage, Turing, Shannon and von Neumann all trying their hand at designing chess programs. It subsequently became the grand challenge for the AI generation researchers. Indeed, Chess is one of the most and longest-studied domains in the history of AI.

After decades of research it had enormous progress. Finally a landmark was achieved in 1997 when the IBM program **Deep Blue** defeated the human world champion Garry Kasparov. Computer chess programs continued to progress steadily and two decades later in 2017 AlphaZero, developed by **Google's DeepMind**, was introduced.

AlphaZero achieved superhuman performance and became the strongest player in the history of games, winning games not just against world champion chess engines, but also against top human players.

♘ How does AphaZero play?

Algorithm of AlphaZero: To learn each game, an untrained neural network plays millions of games against itself via reinforcement learning. RL is a type of machine learning technique that enables an agent to learn in an interactive environment by trial and error using feedback from its own actions and experiences. At first, it plays randomly, but over time the system learns from wins, losses, and draws to adjust the parameters of the neural network, making it more likely to choose advantageous moves in the future.

AlphaZero uses a general purpose **Monte Carlo tree search (MCTS)** algorithm to select the most promising moves in games. For each move, AlphaZero searches only a small fraction of the positions considered by traditional chess engines.

Search: AlphaZero Vs. World Champion Chess Engines: AlphaZero searches just 80 thousand positions per second in chess, compared to 70 million for Stockfish and 35 million for Elmo.

♘ Wondering what's Zero in AlphaZero?

Zero refers to having ZERO "human" knowledge of playing chess. AlphaZero was not fed with any opening books, endgame tablebases or domain-specific heuristics of ways to win. It wasn't provided with any history of the human plays, zero examples were given, and absolutely zero data on guidance was given.

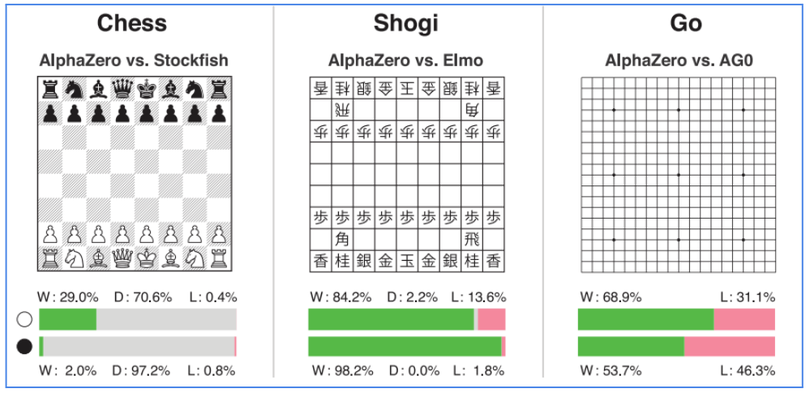

Figure 1: Tournament evaluation of AlphaZero in chess, shogi, and Go in matches against Stockfish, Elmo & AlphaGo Zero. In the top bar, AlphaZero plays white; in the bottom bar AlphaZero plays black. Each bar shows the results from AlphaZero's perspective: win ('W', green), draw ('D', grey), loss ('L', red).

♖ World Champion's View on AlphaZero

"I can't disguise my satisfaction that it plays with a very dynamic style, much like my own! The implications go far beyond my beloved chessboard. Not only do these self-taught expert machines perform incredibly well, but we can actually learn from the new knowledge they produce." **Garry Kasparov**

"It's like discovering the secret notebooks of some great players from the past. Impressively, it manages to impose its style of play across a very wide range of positions and openings," **says Matthew Sadler**, who observes that it plays in a deliberate style from its first move with a very human sense of consistent purpose.

"AlphaZero could be a powerful teaching tool for the whole community." **Natasha Regan.**

♔ Anatomy: World Champion Chess Engines Vs. AlphaZero

Anatomy of World Champion Chess Engine: Traditional world computer chess champions (viz., Stockfish & Deep Blue) rely on domain knowledge, thousands of rules, a vast array of specialized components, knowledge, a high performance alpha-beta search-tree using a large number of clever heuristics & domain-specific adaptations and strategies handcrafted by strong human players that accounts for every eventuality in the game.

Anatomy of AlphaZero = All Components Removed. AlphaZero replaces all traditional game playing programs with a deep neural network, a general-purpose *tabula rasa* self-play reinforcement learning algorithm and a general purpose tree Monte-Carlo Tree search algorithm that knows nothing about the game beyond the basic rules.

AlphaZero's Neural Network at Play AlphaZero uses a deep neural network (p, v) = fθ(s) with parameters θ. This neural network fθ(s) takes board position ***s*** as input and outputs a vector of move probabilities ***p*** with components pa = Pr(a|s) for each action a, and a scalar value v estimating expected outcome ***z*** from position s, v ≈ E[z|s]. AlphaZero learns these move probabilities and value estimates entirely from self play; these are then used to guide its search.

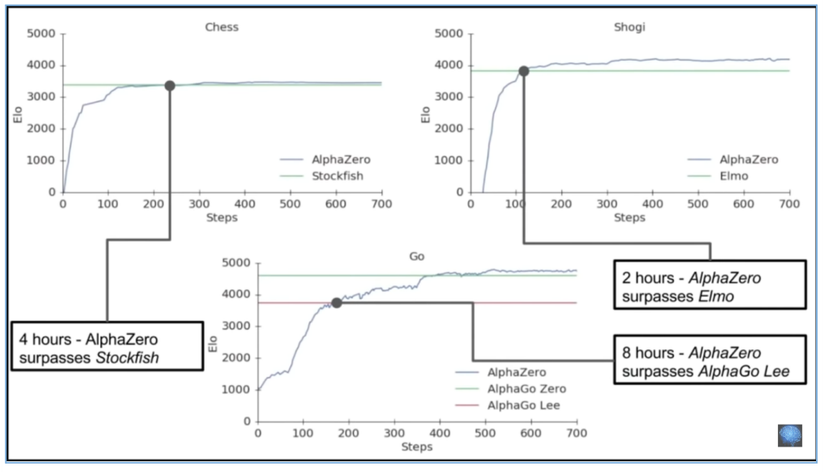

Figure 2: Performance of AlphaZero in 700,000 steps
A. In chess, it beat the 2016 world-champion Stockfish in 4 hours (300,000 steps).
B. In shogi, it beat 2017 CSA world-champion Elmo in 2 hours (110,000 steps).
C. In Go, AlphaZero beat AlphaGo after 8 hrs (74,000 steps).

♖ Critical Evaluation: AlphaZero Vs. Human Play

1. Learning Process:
Humans: Learn through a combination of study, practice and experience. They acquire knowledge through books, coaches, databases and playing against others to improve their openings, tactics and strategies.

AlphaZero: Learns entirely through self-play and reinforcement learning. It discovers its own strategies through millions of games played against itself, continuously improving its performance by adjusting its neural network. Alpha Zero's play is not static; it continues to learn and improve over time, making it a dynamic and evolving chess player. It does not rely on human input or databases.

2. Style of Play:
Humans: Tend to follow established play and well-known opening theories; have memorized specific sequences of moves to gain an early advantage or reach familiar positions. Players focus on exploiting weaknesses or surprising opponents with unconventional openings. They exhibit a wide range of playing styles, from aggressive and tactical to positional and strategic. Creativity and intuition play a significant role in human play. Being human, our style may also depend on mood based on opponents, tournament situations, or specific positions including making conscious decisions to switch strategies during a game.

AlphaZero's style of play — relentless aggression. It is free from all feelings and emotions. It displays a unique style that is more intuitive, creative, innovative, diverse and unconventional yet strategic moves. Not bound by traditional thinking, it has surprised human players with flexible and innovative strategies. It doesn't follow any predetermined patterns and does not adhere to any established principles or well-known lines. It explores a wide range of novel and creative play. This ability to learn each game afresh, unconstrained by the norms of human play, results in a distinctive, unorthodox, yet creative and dynamic playing style.

3. Tactical and Strategic Understanding:
Humans: Excel in tactical awareness, pattern recognition and calculation with feeling, insight and intuition. They're also skilled at finding tactical opportunities during a game and prioritize long-term plans & maneuvering.
AlphaZero: Demonstrates a high level of tactical acumen. Its evaluation function is learned from scratch, so it evaluates positions differently than humans and has been known to prioritize long-term strategic considerations over short-term tactical gains. It sacrifices material for long-term positional advantages, showcasing a unique strategic approach.

Conclusion

AlphaZero brings us a step closer to realizing a longstanding ambition of AI — to create programs that can learn for themselves using first principles. These intelligent systems (flexibile with the ability to generalize to new situations) will have the potential to address a diverse array of real-world problems, opening up a world of possibilities for AI across innumerable domains.

References

- Silver, D. et al. (2018) 'A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play'. Science 362(6419), pp. 1140-1144 at https://www.science.org/doi/abs/10.1126/science.aar6404
- Silver, D. et al. (2017) 'Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm' at: http://arxiv.org/abs/1712.01815
- AlphaZero: Shedding light on chess, shogi & Go (2018) at: https://deepmind.google/blog/alphazero-shedding-new-light-on-chess-shogi-and-go/