Measuring the Strength of AlphaGo(Zero)

Authors: Kien Nguyen, Khiem Nguyen, Matthew Bolding
Advisors: Dr. Liran Ma, Dr. Ze-Li Dou

TCU College of Science & Engineering
reSEaRCh
Science and Engineering Research Center

Overview

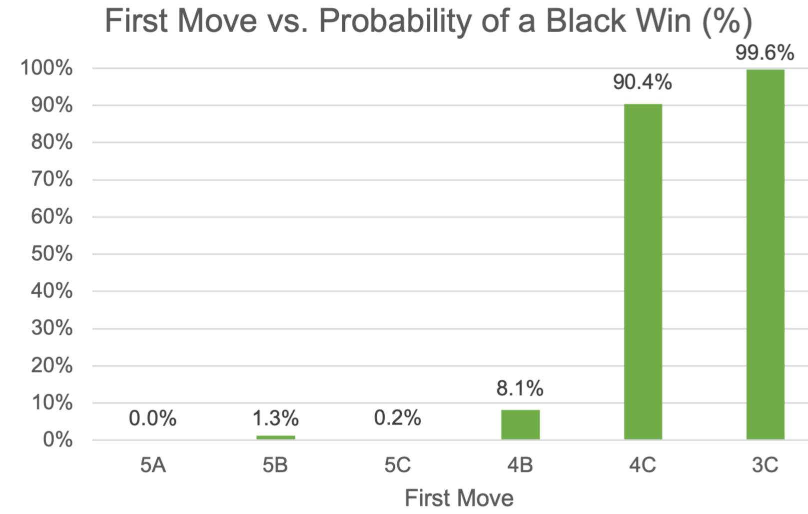
SciCom Let's Talk Science

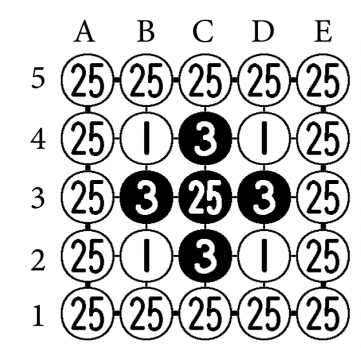
- Continuing from our first poster, "Exposing AlphaGo(Zero)'s Weaknesses", we now present our initial findings in measuring the strength of an AlphaGo(Zero) agent.
- Important concepts to recall include:
  - **Go**: a zero-sum adversarial board game for 2 players where Black plays first;
  - **Goal**: secure as much board space as possible;
  - **Komi**: the compensation added to White's final score to compensate for going second.
- We expect to see that the agent plays too aggressively when losing and too consecutively when winning.

Proof of Progress

- In parallel to a figure by van der Werf *et. al.*, we found that the first move often dictates the remainder of the game; whichever color plays in the center is likely to win.

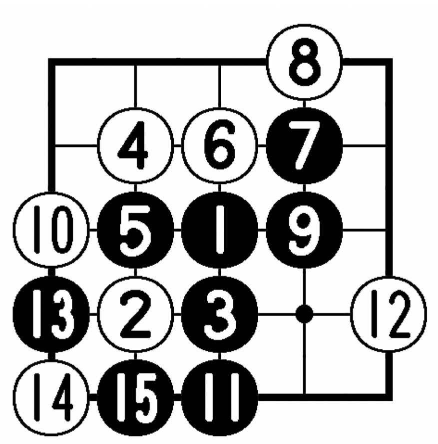
First Move vs. Probability of a Black Win (%)





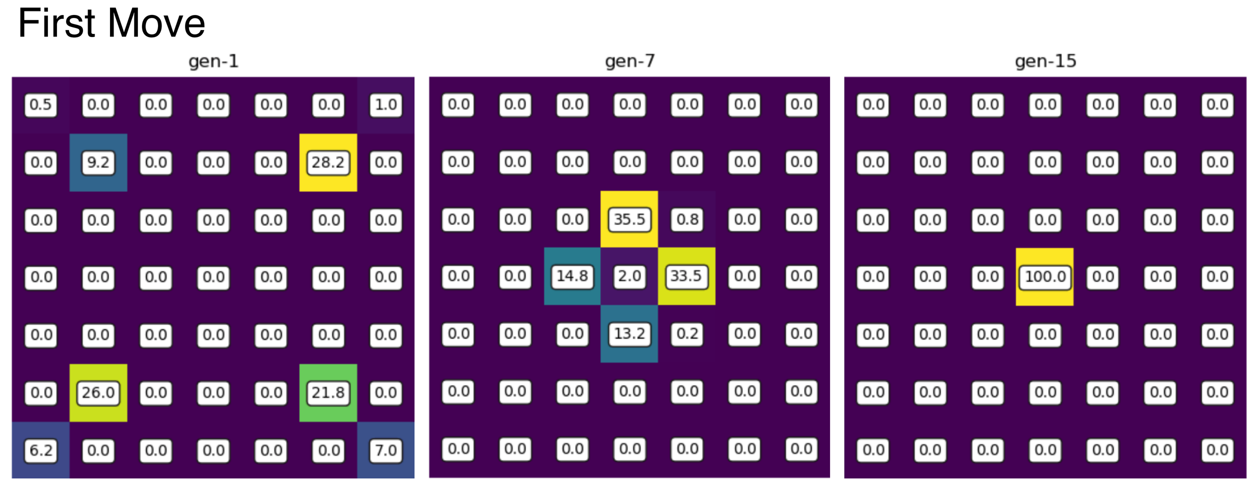
Opening Moves

- Our agent is able to find, as verified by "Solving Go On Small Boards," an optimal line of play. At the eleventh move, our agent plays an equally optimal move — here, Black wins either way.

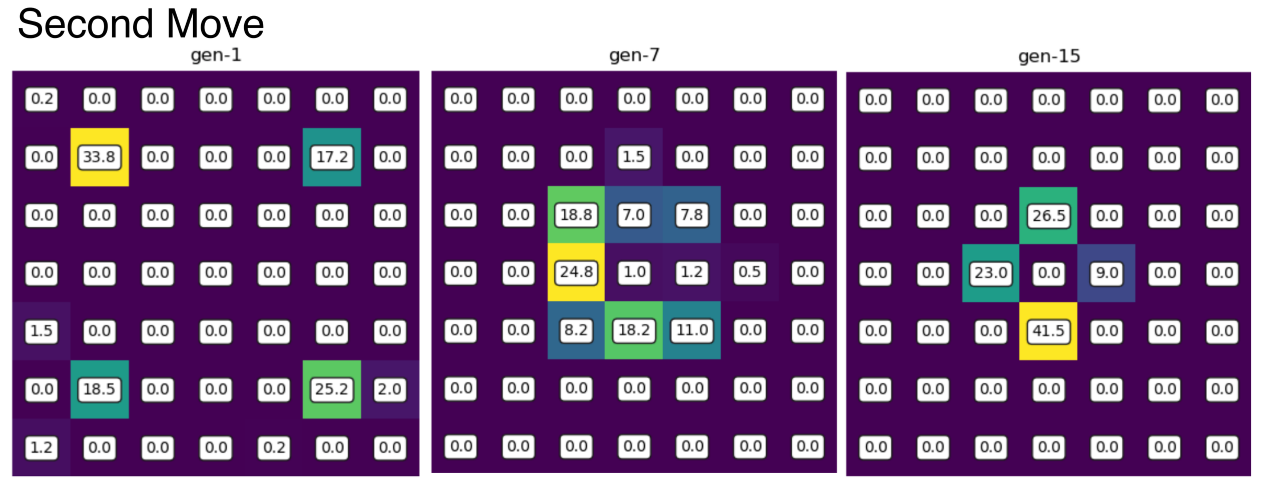


- On a 7×7 board, early moves trend to the center as the number of generations grow; first move approaches the center square, while the next move approaches cardinally adjacent squares.

First Move

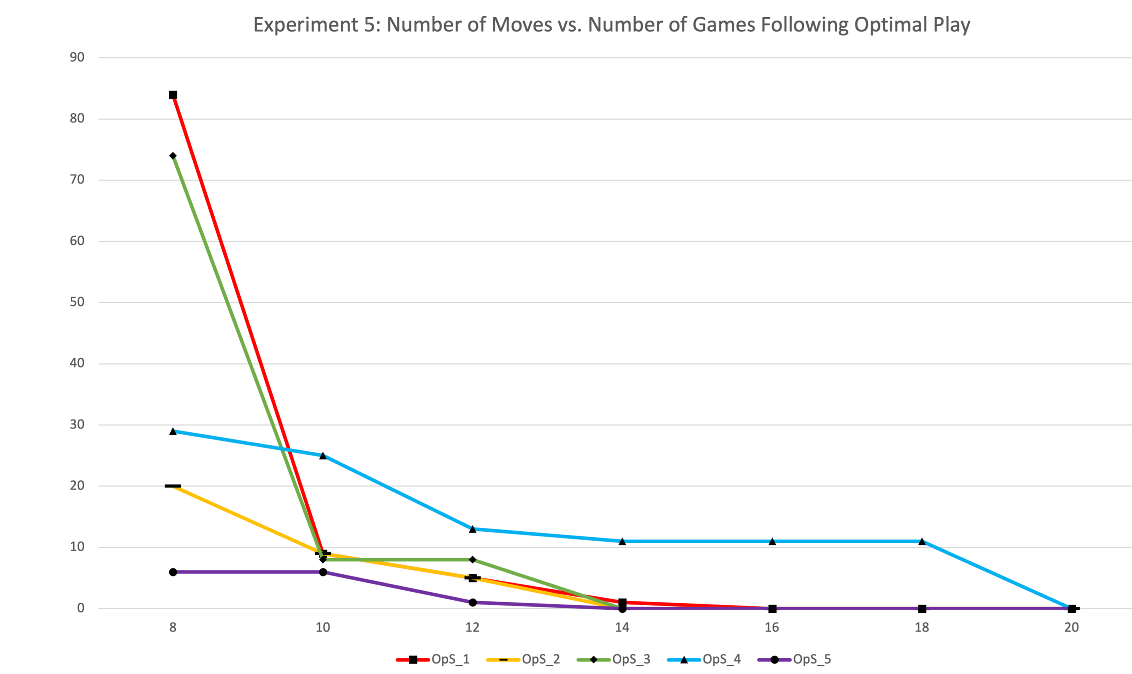


Second Move



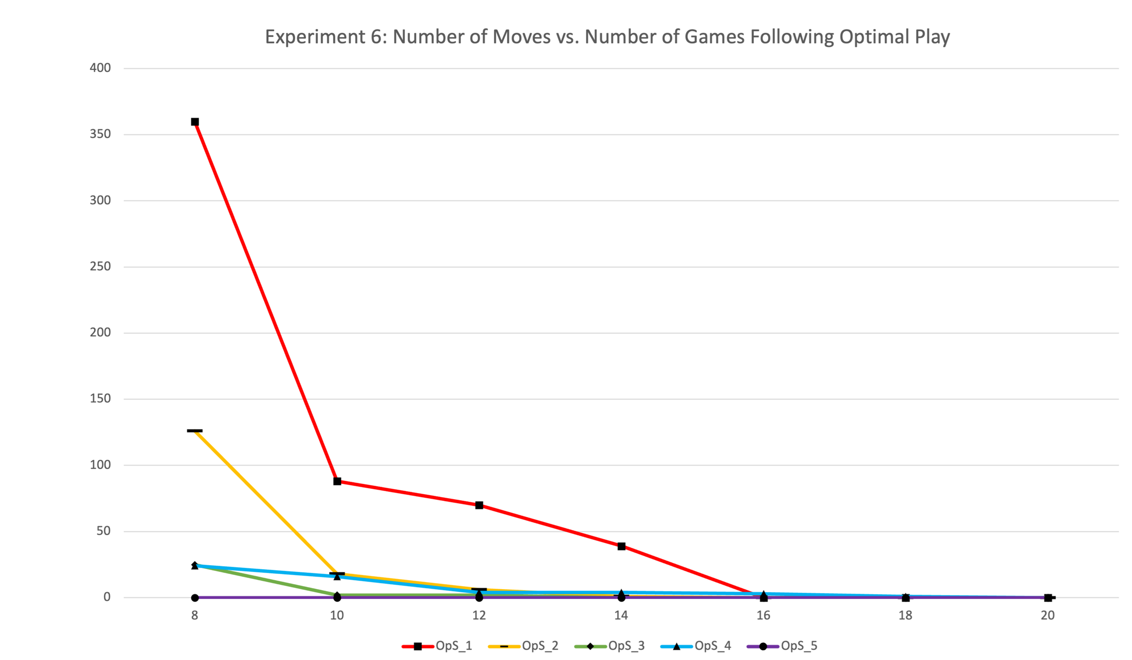
Optimal Solutions

- Our experiments confirm that AlphaGo(Zero) type agents do not play optimally. Instead, they play to maximize win rate.



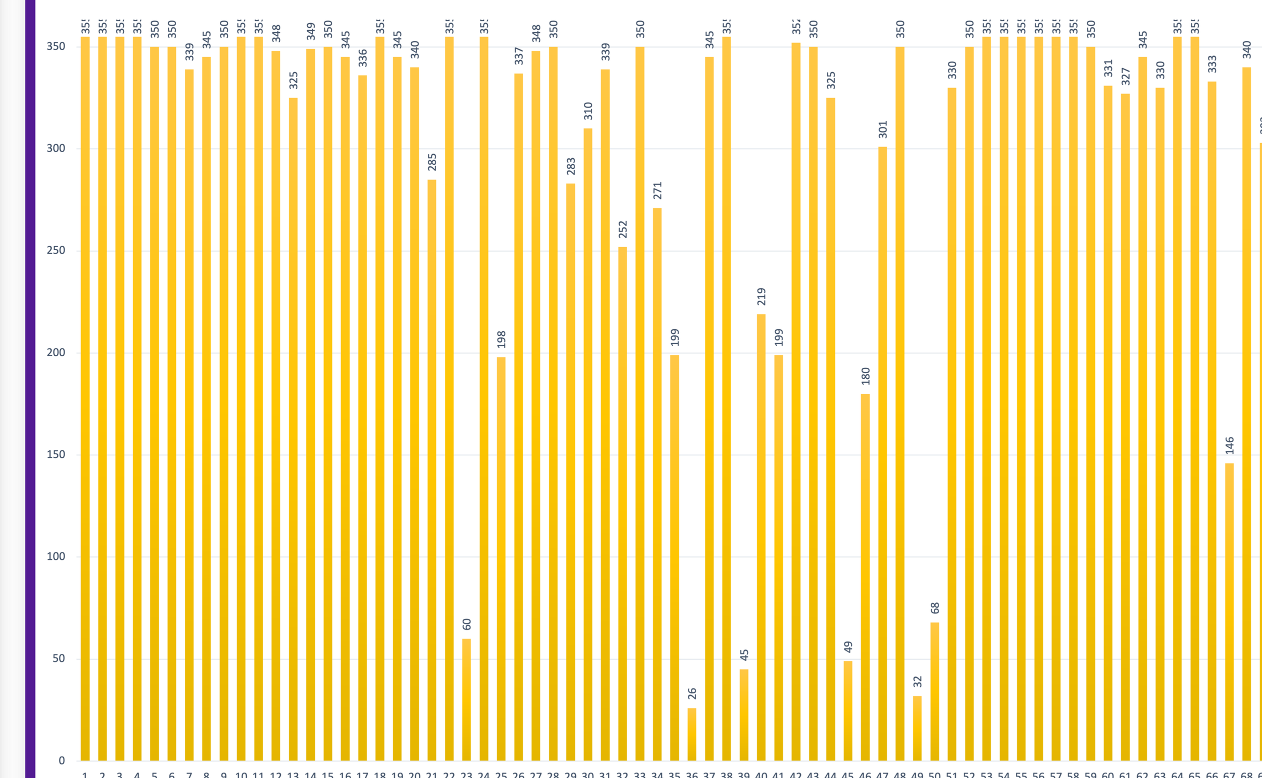
Observe that after only twenty moves, none of the agents played a game following optimal play.

- Based on solving puzzles whose solutions are known, we aim to analyze the extent to which the agent does or does not follow optimal lines of play.
- With a larger sample of games, we see that further into a game, the agent strays further from optimal lines of play. This is indicative of the deficiency in the algorithm itself, as we exposed in our work.



Puzzle Solving

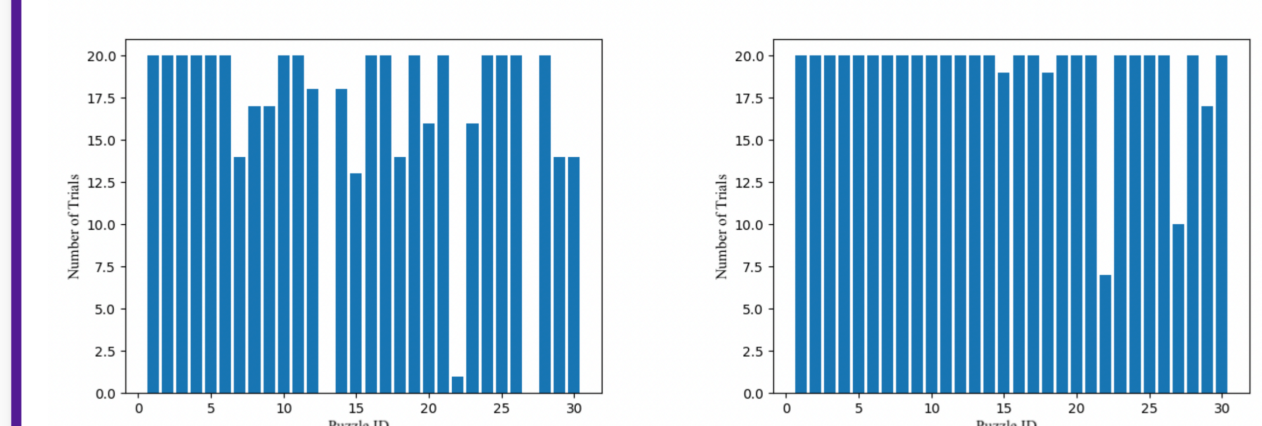
- We provided set game positions (puzzles) to the agent to determine the percentage of accuracy between our agent's moves and the optimal ones.



Puzzle Number vs. Number of Generations Solving the Puzzle

Targeted Training

- When the agent is unable to play optimally, we attempted to enhance the agent's performance by modifying the agent's *tabula rasa* approach to learning by inserting set game positions into the training process without providing solutions.



(a) Self-play-trained agent. (b) Target-trained agent.

Conclusion and Future Work

- On the 5x5 board, we successfully found optimal plays.
- We demonstrated on the 7x7 board that we could enhance the performance of the AlphaGo(Zero) agent by targeted training — inserting specially formed puzzles into the training process enhances its performance.
- Design a more precise and robust evaluation system.
- Find methods to enhance the performance of our playing agent.

Technologies Used

TensorFlow — Deep Learning framework
Python — Write training script
C++ — Implement Go game

References

- Z. Ahmad, N. Sturtevant, and M. Bowling, "Measuring the Solution Strength of Learning Agents in Adversarial Perfect Information Games," 2020. [Online]. Available: http://aaai-rlg.mlanctot.info/
- van der Werf, Eric C. D., van der Herik, H. Jaap, and Uiterwijk, Jos W. H. M., "Solving Go On Small Boards," ICGA Journal, vol. 26, no. 2, pp. 92–107, October 2003.
- Ze-Li Dou, Liran Ma, Khiem Nguyen, and Kien X. Nguyen, "Paradox of AlphaZero: Strategic vs. Optimal Plays," IPCCC, 2020.

Acknowledgements

The team would like to thank the following individuals:

- **Dr. Liran Ma:** for guidance throughout moments of uncertainty, technical assistance and advice, and facilitating substantive conversation.
- **Dr. Ze-Li Dou:** for providing excellent insight into Go, handcrafting numerous puzzles for the agent to play through, and providing solutions to such puzzles.