# CMPUT 396: Epilogue

AlphaGo Zero

Let's go back to that question that David Silver was asked about AlphaGo: is it possible that starting with human data biased the results of AlphaGo? Is it possible to do the same thing *without* starting from human data?

At the time, David Silver said he didn't know. But, they worked on it and the result is AlphaGo Zero.

AlphaGo Zero is *tabula rasa*, or blank slate, so they started without any human data, leaving out the initial supervised learning step that AlphaGo had, and learned completely from self-play.

Remember game 4 against Lee Sedol where AG's win rate dropped at move 78. Lee Sedol played a move that was ranked very low by human professional players. Did learning from the human pros bias its results here? If it had learned on its own and not from humans would it have made this error?

How can we do the same thing but without human game records?

Starting from a blank slate might take too long to learn, so how do we speed it up? Added a powerful search feature in their learning to speed it up.

Some differences from previous versions:
- Only uses black and white stones from board as input (previous versions included a small number of hand-engineered features)
- Uses one neural network instead of two (AG used a policy network to select the next move to play and a value network to predict the winner of the game from each position, these are combined in AGZ for more efficient training and evaluation).
- Does not use rollouts (the fast random games to predict who will win from current board position), instead relies on high quality neural networks to evaluate positions.

These differences improved performance and made it more general.

After 3 days of self-play training, AlphaGo Zero beat the Lee Sedol version of AlphaGo 100-0.

Computing resources used by each version (note this is what it played on, not what it trained on):
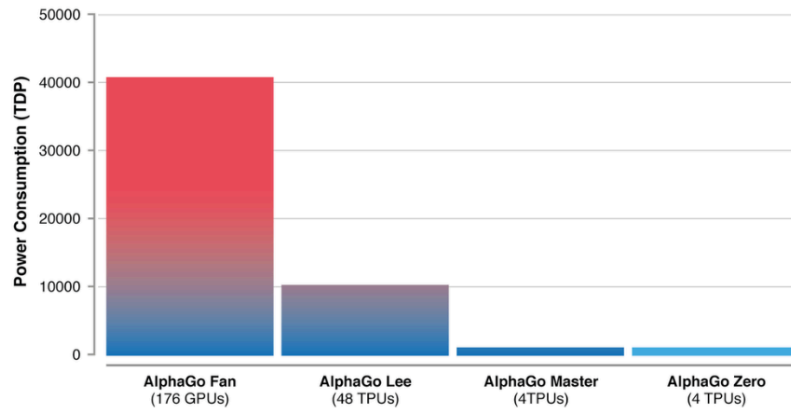


TPU: new product of Google, like GPU but tailored to their kinds of problems.

**Some lingering questions:**
1) How many machines (and what hardware) did DeepMind train AlphaGo Zero on?
2) DeepMind explains many details of the AlphaGo Zero architecture, but parameter tuning is critical in neural net training and performance. Did AlphaGo Zero use (perhaps to start) AlphaGo neural net parameter values? If yes, is it fair to say that AGZ learns *tabula rasa*?

## AlphaZero

AlphaZero is an even more general version that can achieve tabula rasa, superhuman performance given only the rules of the game. In 24 hours it was able to achieve superhuman level of play in chess, shogi (Japanese chess), and Go, defeating a world champion program in each case.

## Conclusion

The DeepMind story is one of all the pieces falling together at the right time and the right place.

Not only did the right people come together, but also the right ideas came along at the same time (MCTS, DCNN) to help them achieve this amazing feat.