

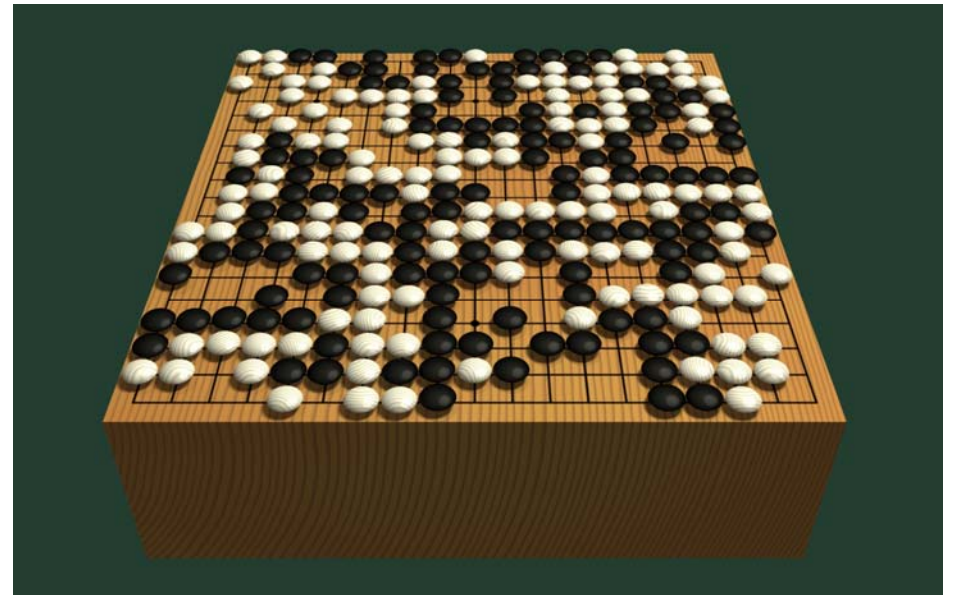
Mastering the game of Go without human knowledge

Paper by D.Silver *et al.* (2017)
Seminar by Jack Kleinsman

jkle533 113166519

The game of Go

- $\sim 2 \times 10^{170}$ legal game states
- Thousands of years of history
- Over 20 million active players



Machine Learning Methods

Supervised Learning

*AlphaGo Fan, AlphaGo Lee,
Alpha Go Master*

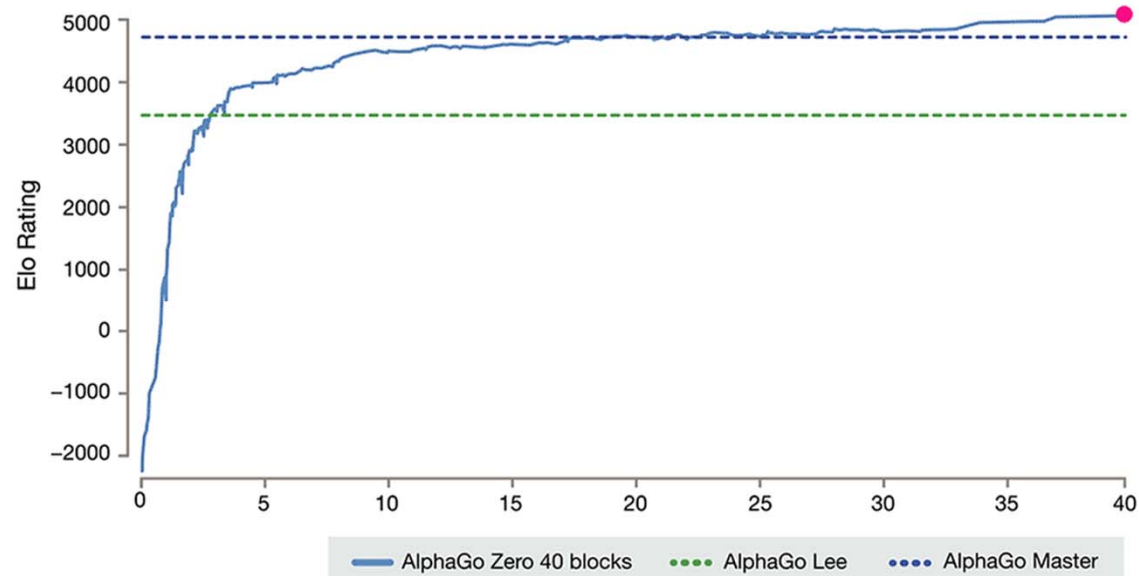
- Learns from labelled data
- Requires a large dataset
- Susceptible to overfitting

Reinforcement Learning

AlphaGo Zero

Learns from repeated self-play
Only requires knowledge of the
games rules

AlphaGo Zero's performance



jkle533 113166519

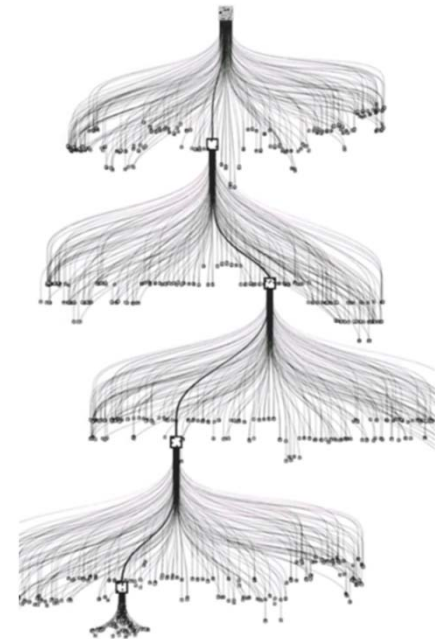
Knowledge learnt by AlphaGo Zero

- The fundamental human knowledge of Go
- Discovered new techniques

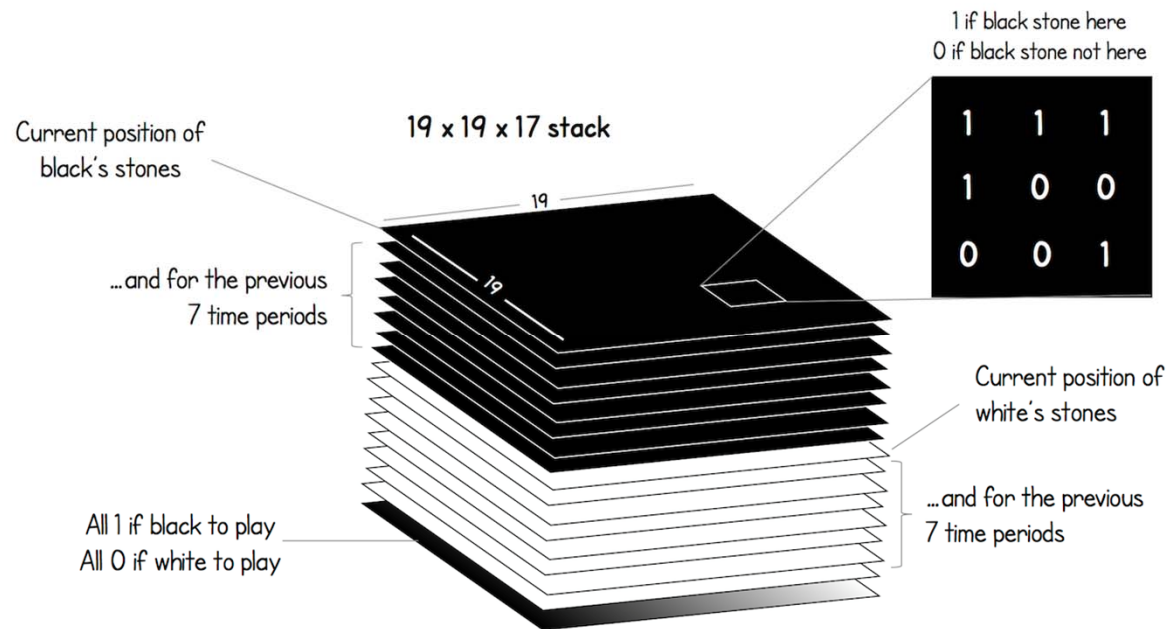
AlphaGo Zero surpassed thousands of years of Go knowledge in just 40 days.

AlphaGo Zero's architecture

- Input is the current game state
- One convolutional layer
- 40 residual layers
- Value and policy heads



What AlphaGo Zero sees...



jkle533 113166519

Convolutional layer

- The input game state has 6,137 dimensions
- It suffers the from 'curse of dimensionality'
- The convolutional layer reduces the dimensions of the data while preserving the important information

Convolutional layer

1 _{x1}	1 _{x0}	1 _{x1}	
0 _{x0}	1 _{x1}	1 _{x0}	
0 _{x1}	0 _{x0}	1 _{x1}	

Image

4		

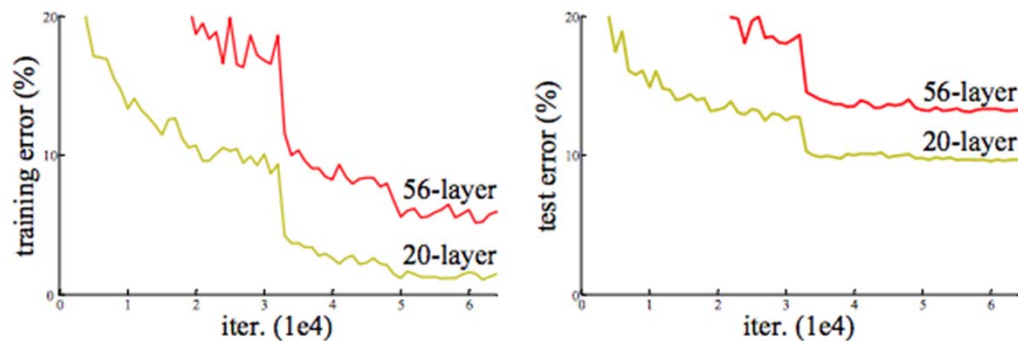
Convolved
Feature

gif from www.towardsdatascience.com

jkle533 113166519

Residual layers

- After passing a threshold, the performance of traditional CNN's decrease as the number of layers increase.



jkle533 113166519

Residual layers

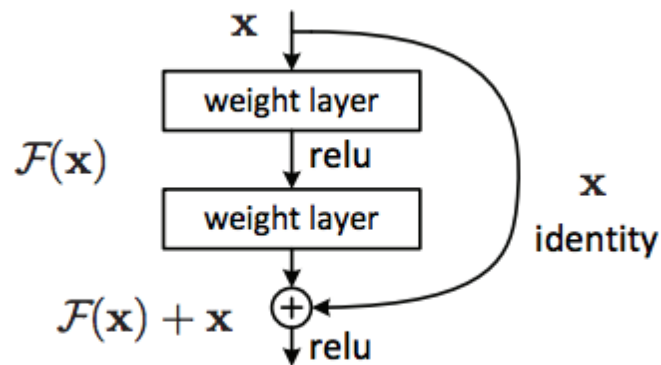


Figure 2. Residual learning: a building block.

Limitations

- AlphaGo Zero requires full knowledge of the rules
- Simulating problems is computationally expensive
- There must be a way to calculate how correct a solution is

Conclusion

- Computers have achieved superhuman performance in Go
- A reinforcement learning based algorithm surpassed the supervised learning algorithms
- Reinforcement learning was able to learn strategies both existing and new
- A generalized version of the AlphaGo Zero algorithm, AlphaZero, has achieved similar results in chess and shogi

References

- D. Silver *et al.* “Mastering the game of Go without human knowledge”, *Nature*, vol. 550, pp. 354-359, Oct. 2017. doi:<https://doi.org/10.1038/nature24270>.
- D. Silver *et al.* “Mastering the game of Go with deep neural networks and tree search”. *Nature*, vol. 529, pp. 484-489, Jan. 2016. doi:<https://doi.org/10.1038/nature16961>.
- S. Saha. “A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way”. [towardsdatascience.com](https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53), <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53> (accessed Aug. 23, 2020).
- C. Shorten. “Introduction to ResNets”. [towardsdatascience.com](https://towardsdatascience.com/introduction-to-resnets-c0a830a288a4), <https://towardsdatascience.com/introduction-to-resnets-c0a830a288a4> (accessed Aug. 23, 2020).

References

- R. Sutton and A. Barto. “Reinforcement Learning: An Introduction”, Second edition, in progress. London, England: MIT Press, 2015.
- J. Tromp and G. Farnebäck. “Combinatorics of Go”. Jan. 2016. tromp.github.io/go/gostate.pdf (accessed Aug. 23 2020).
- American Go Association, “A Brief History of Go”. usgo.org, <https://www.usgo.org/brief-history-go> (accessed Aug. 23 2020).