



Pattern Recognition in Chess

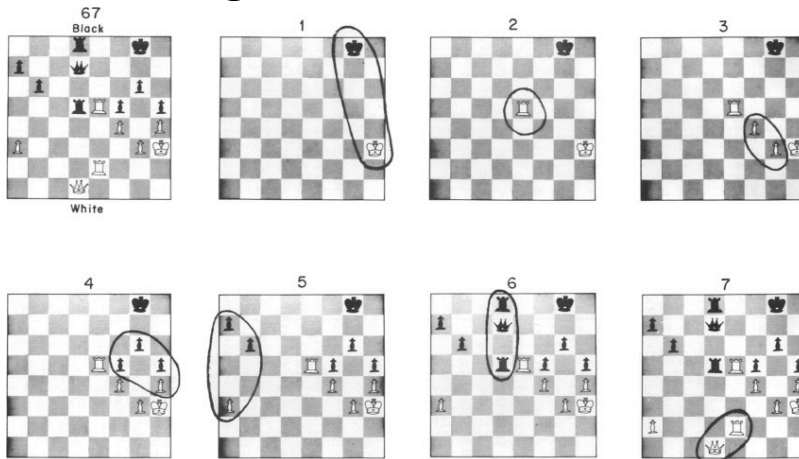
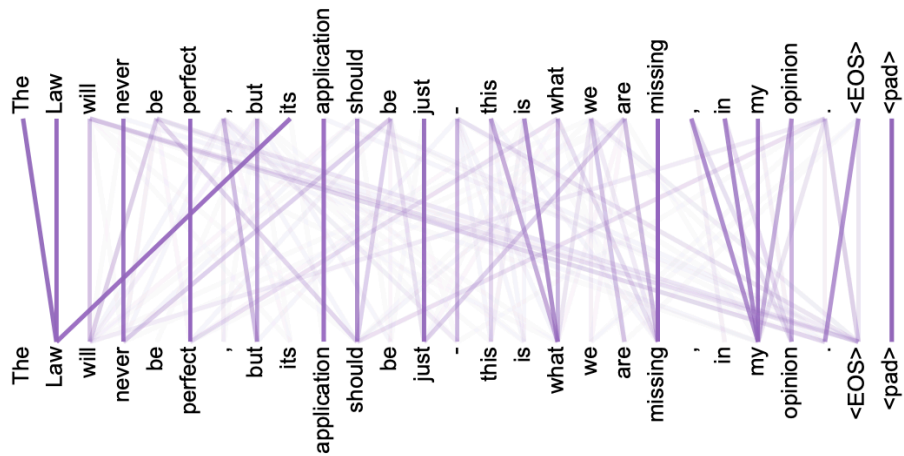


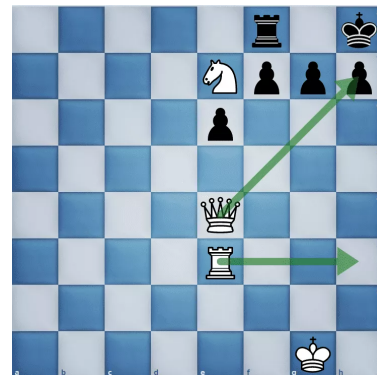
Image from: The mind's eye in chess, by Chase and Simon

Patterns in NLP Systems



Pattern Recognition: Rules vs. NN

The knight goes to e7
 where it controls g8 and g6;
 the queen is sacrificed on h7
 to open the h-file,
 and a rook makes use of the
 newly opened h-file to
 deliver checkmate.



Source: <https://www.chessable.com/blog/chess-pattern-recognition/#:~:text=Chess%20pattern%20recognition%20allows%20you%20to%20use%20Lolli's%20Mate%20on,is%20b7%20instead%20of%20g7.>

NNs vs Rule-based Systems

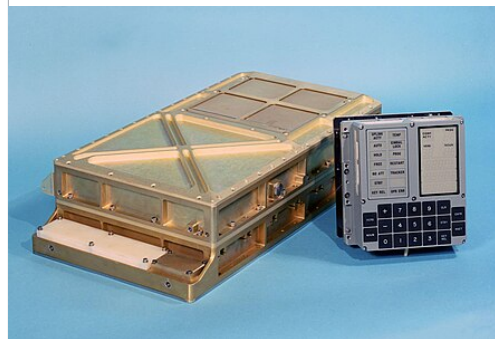
- In the late 60's researchers knew that NNs were the better technology for AI.
- Selfridge: "The most important learning process of all is still untouched. No current program can generate test features on its own. The effectiveness of all of them is forever restricted by the ingenuity or arbitrariness of their programmers."¹⁾
- Minsky, one of the most influential person in AI: "Today, machines solve problems mainly according to the principles we build into them. Before long, we may learn how to set them to work upon the very special problem of improving their own capacity to solve problems."²⁾

1) Oliver Selfridge and U. Neisser. Pattern Recognition by Machine. In: Computers and Thought. Feigenbaum, E. A. and Feldman, J. editors. McGraw-Hill. 1963.

2) Marvin Minsky. Machines are More Than They Seem. Science Journal 4(10). pages 3-43. October 1968.

Computing Power of Apollo 11

- Around 40,000 instructions per second
- 15-bit word length + 1-bit parity
- 2048 words RAM (magnetic-core memory)
- 36,864 words ROM (core rope memory)



Source: https://en.wikipedia.org/wiki/Apollo_Guidance_Computer

Computing Power to Train ChatGPT

- GPT-3 required $3.14e23$ flops of computing power for its initial training.
- Assuming GPT-4 is about 10 times bigger than GPT-3 with 175 billion parameters, it would require about $3.14e24$ flops of computing power for its training.¹⁾
- At least **10,000 high-end NVIDIA GPUs**
- ChatGPT-3 has **175 billion** parameters and requires 45 terabytes of data for training.²⁾

1) [https://www.quora.com/How-much-computing-power-and-cost-is-involved-in-running-ChatGPT-4#:~:text=As%20you%20can%20see%2C%20you,or%20A%20series%20from%20Nvidia\).&text=How%20much%20CPU%20power%20does,power%20for%20its%20initial%20training](https://www.quora.com/How-much-computing-power-and-cost-is-involved-in-running-ChatGPT-4#:~:text=As%20you%20can%20see%2C%20you,or%20A%20series%20from%20Nvidia).&text=How%20much%20CPU%20power%20does,power%20for%20its%20initial%20training)

2) <https://www.linkedin.com/pulse/secrets-chatgpts-ai-training-look-high-tech-hardware-behind-kandel/>

Learning by Self-Play

- AlphaGoZero
 - Trained by self-play reinforcement learning:
 - Initially random play
 - Without any supervision
 - Without human training/expertise
 - Just the rules of go
 - Learned patterns by itself
 - Discovered a remarkable level of Go knowledge.
 - Is believed to have learned a strategy that is qualitatively different to human play.

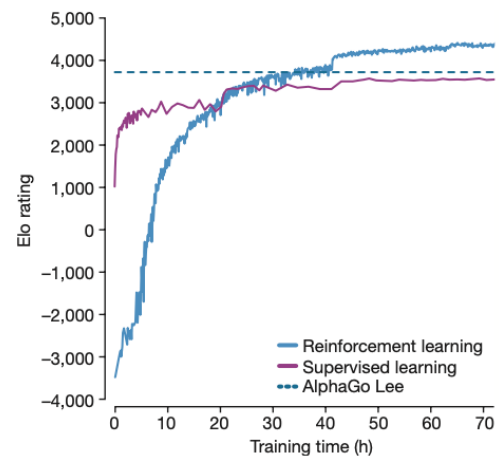


Figure from: Silver et al. Mastering the game of Go without human knowledge. NATURE, VOL 550, 19 October 2017