Alpha Zero
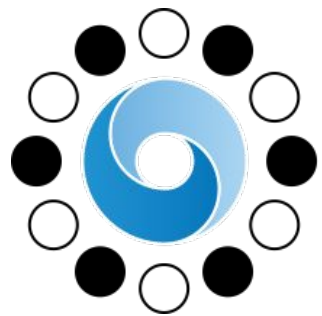
Введение
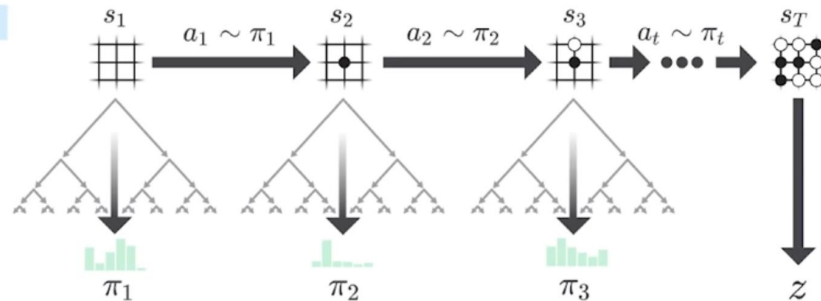
Введение
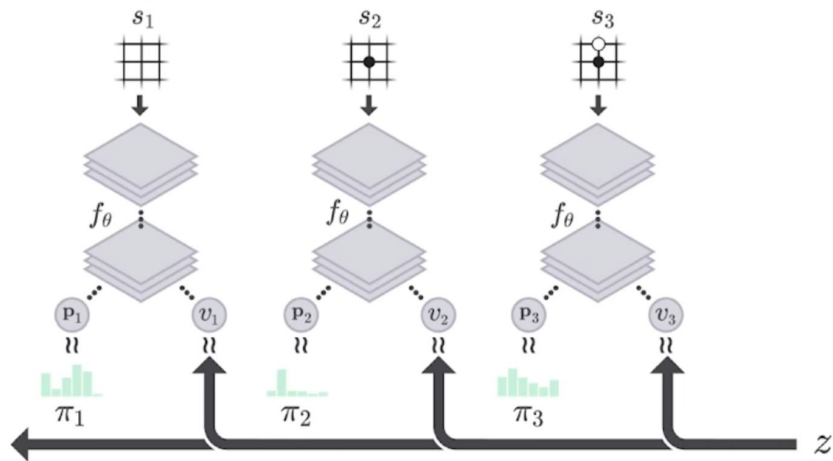

    Далее был разработан новый вариант алгоритма *AlphaGo Zero*, примененный к шахматам.

Alpha Zero для шахмат
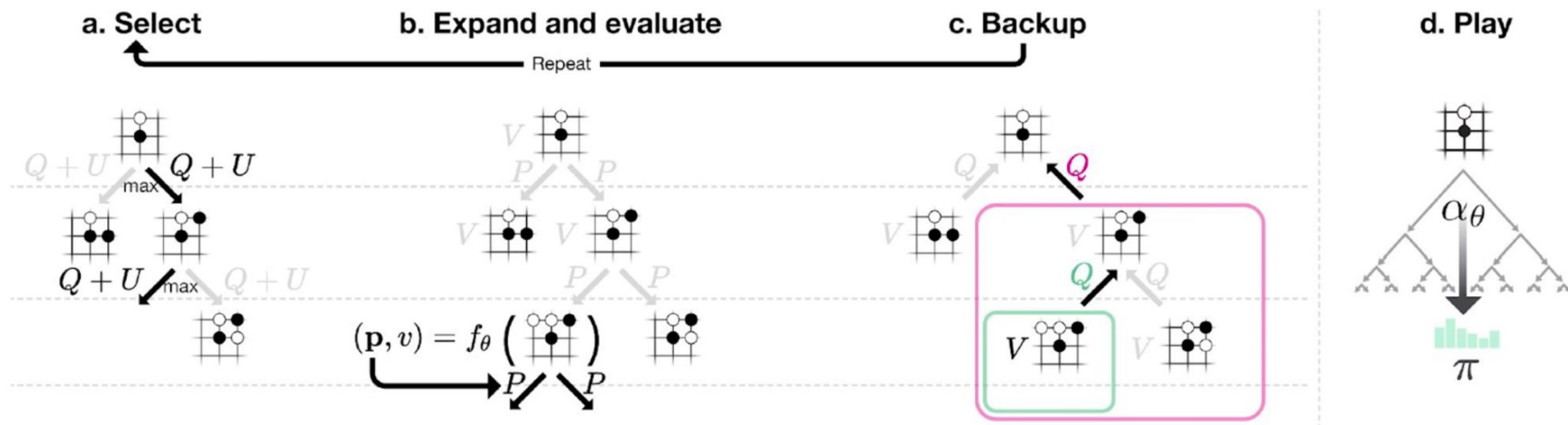
Alpha Zero для шахмат

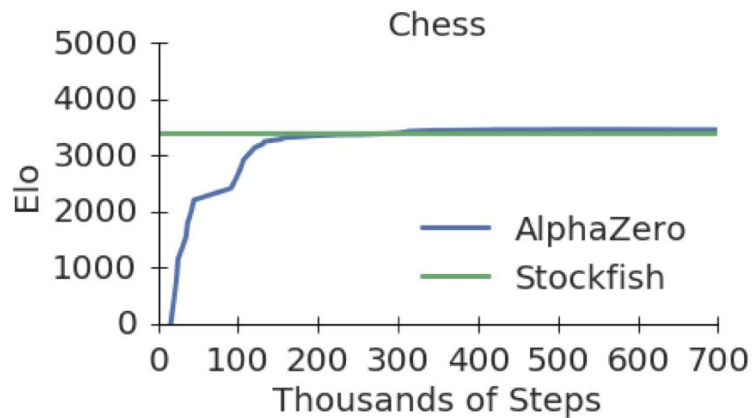
$$l = (z - v)^2 - \boldsymbol{\pi}^\top \log \mathbf{p} + c||\theta||^2$$

Alpha Zero vs Stockfish

Alpha Zero vs Stockfish

Game	White	Black	Win	Draw	Loss
Chess	*AlphaZero*	*Stockfish*	25	25	0
	Stockfish	*AlphaZero*	3	47	0
Shogi	*AlphaZero*	*Elmo*	43	2	5
	Elmo	*AlphaZero*	47	0	3
Go	*AlphaZero*	*AG0 3-day*	31	–	19
	AG0 3-day	*AlphaZero*	29	–	21

Alpha Zero vs Stockfish

Program	Chess	Shogi	Go
AlphaZero	80k	40k	16k
Stockfish	70,000k		
Elmo		35,000k	

Источники:

- https://arxiv.org/pdf/1712.01815.pdf
- http://incompleteideas.net/book/RLbook2020.pdf
- https://deepmind.com/research?filters=%7B%22tags%22:%5B%22Reinforcement%20learning%22%5D%7D
- https://www.nature.com/articles/s41586-020-03051-4.epdf?sharing_token=kTk-xTZpQOF8Ym8nTQK6EdRgN0jAjWel9jnR3ZoTv0PMSWGj38iNIyNOw_ooNp2BvzZ4nIcedo7GEXD7UmLqb0M_V_fop31mMY9VBBLNmGbm0K9jETKkZnJ9SgJ8Rwhp3ySvLuTcUr888puIYbngQ0fiMf45ZGDAQ7fUI66-u7Y%3D