A Brief Story About AlphaGo Zero

Speaker: Lu Jia
Date: 2017.10.27

History

- DeepMind Technologies Limited

2010   Founded in London

2014   Acquired by Google

             Part of the Alphabet group

History

- AlphaGo Fan

2015.10
Defeated Fan Hui
2016 .1
Made headlines
on nature

From wiki

From deepmind.com

History

- AlphaGo Lee

2016.3
Defeated Lee Sedol

From deepmind.com

History

- AlphaGo Master

2017.5
Defeated KeJie

From deepmind.com

History

AlphaGo Fan   AlphaGo Lee   AlphaGo Master   AlphaGo Zero
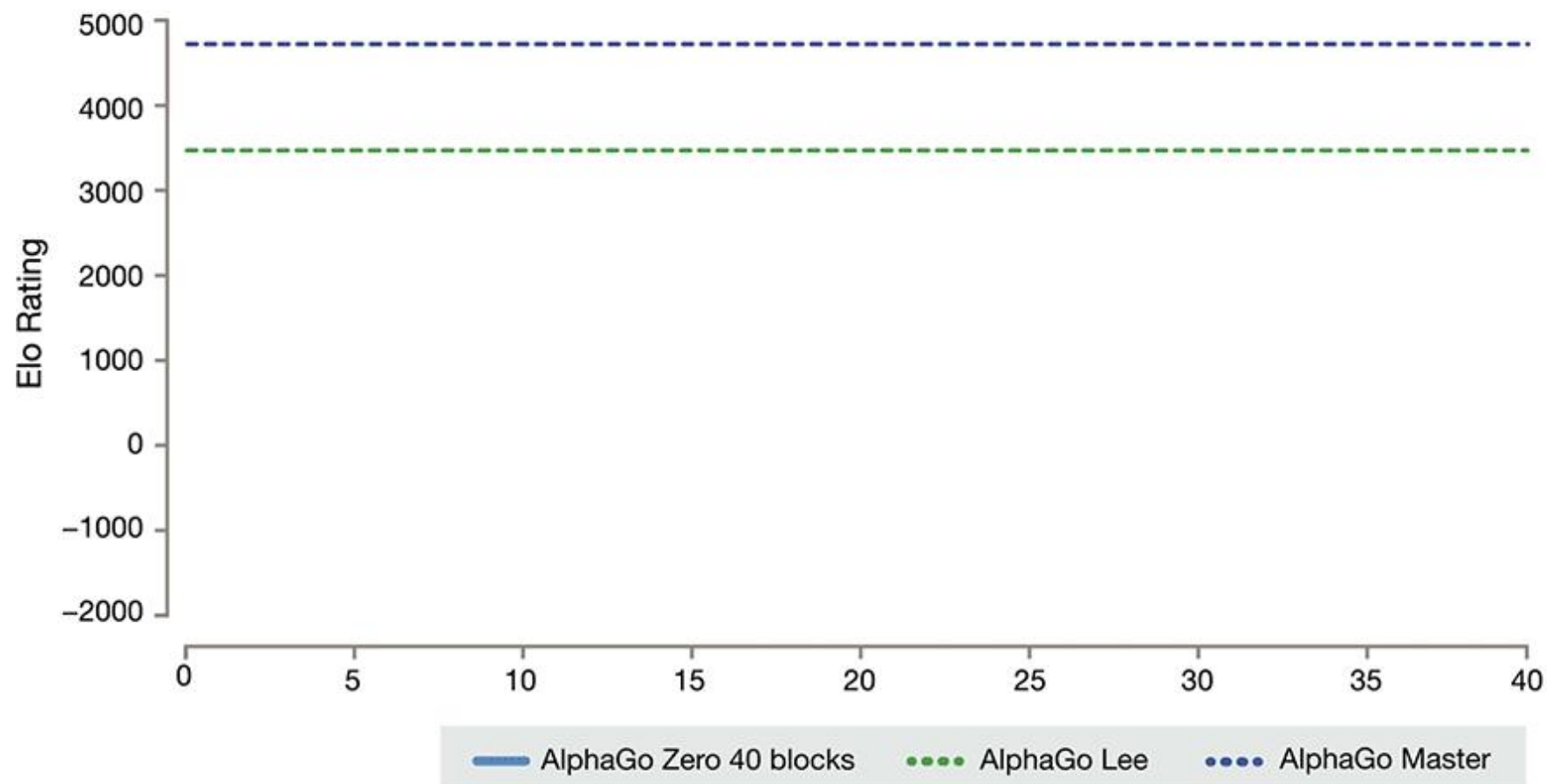
2015.11   2016.3   2016.12 ~ 2017.5   2017.10

From http://mp.weixin.qq.com/s/1S0CW4HxvftffhUZVya20g

AlphaGo Zero
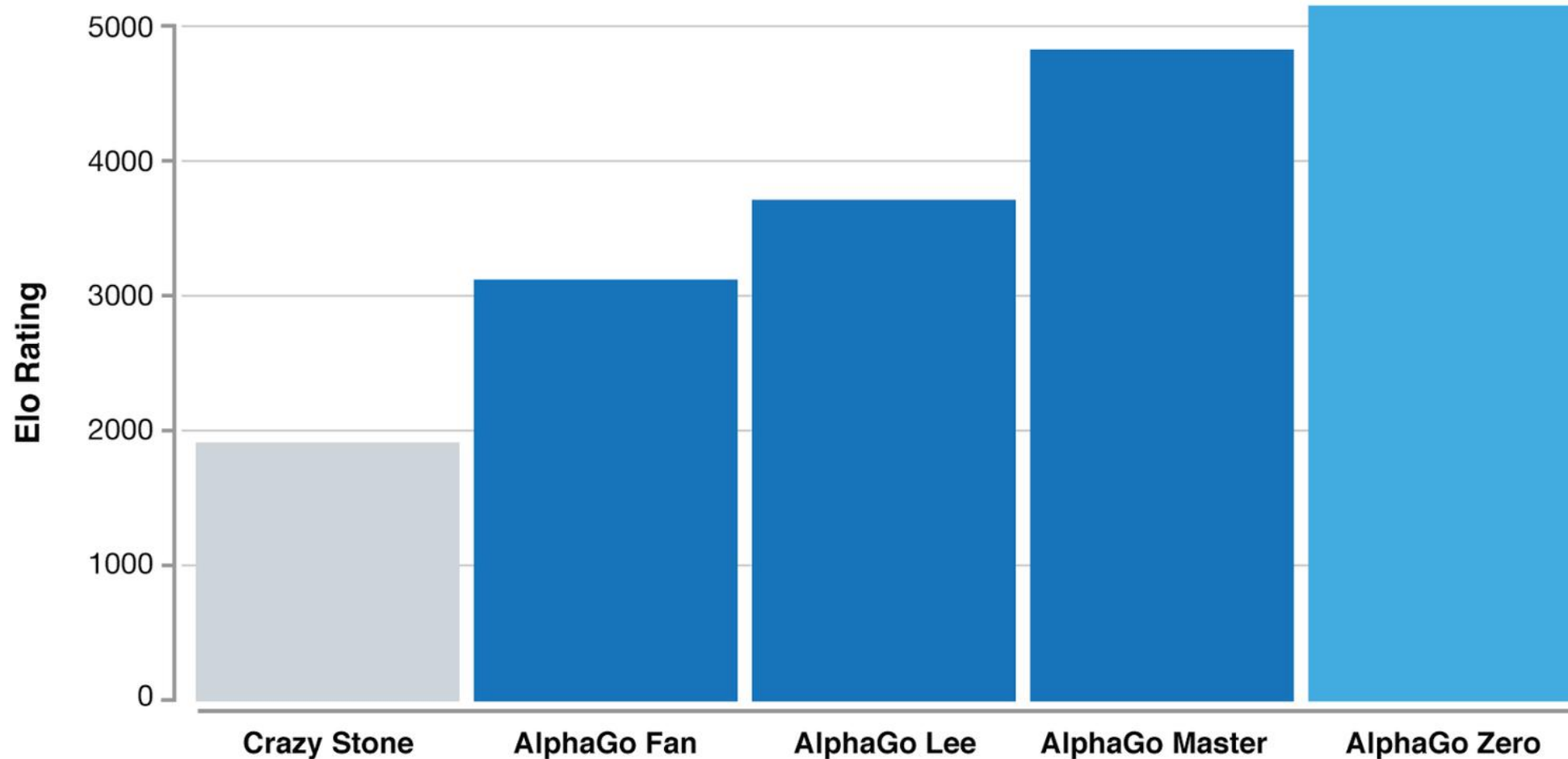

Training time graphic
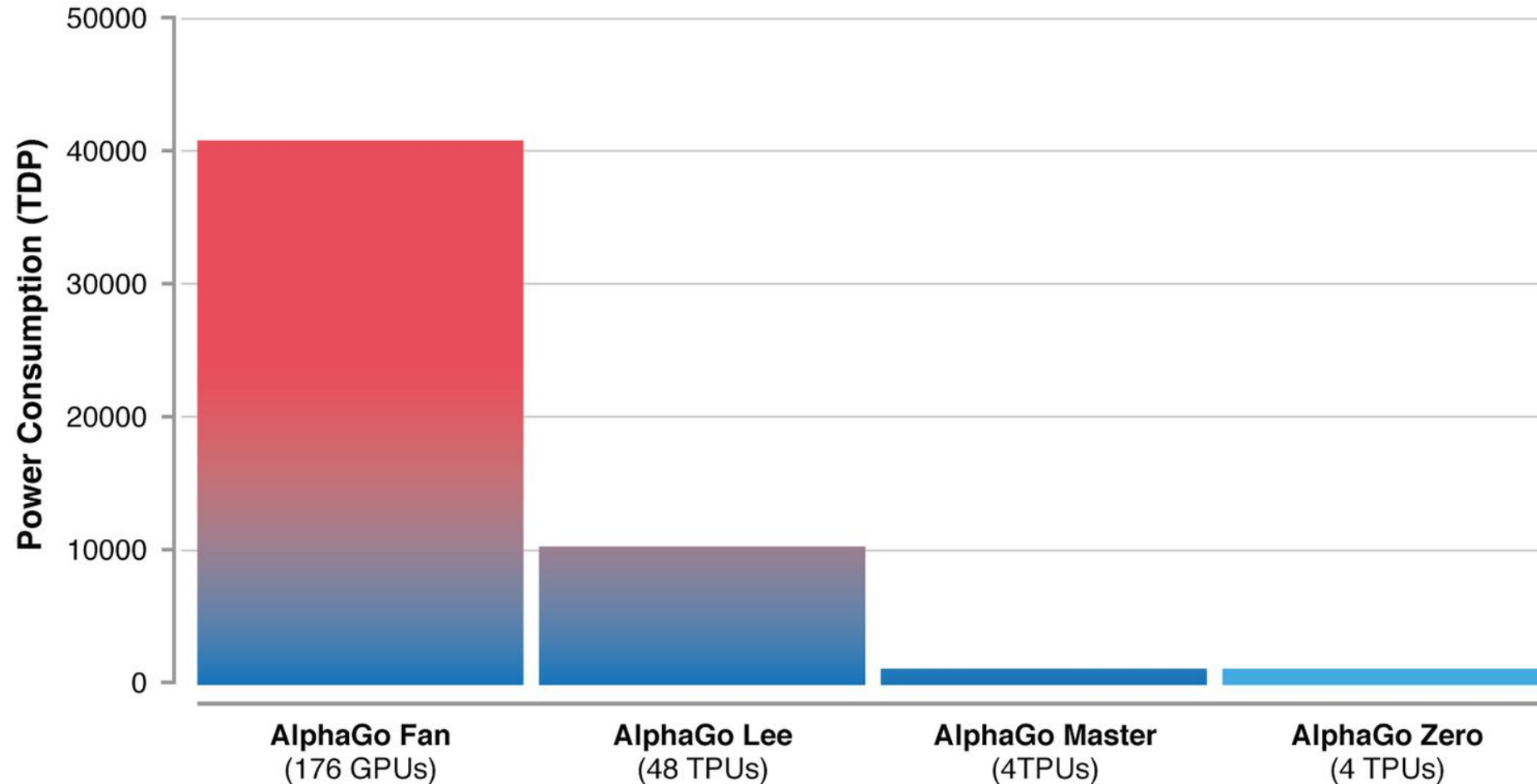*From https://deepmind.com/blog/alphago-zero-learning-scratch/*

AlphaGo Zero

Elo ratings - a measure of the relative skill levels of players in competitive games such as Go - show how AlphaGo has become progressively stronger during its development
From https://deepmind.com/blog/alphago-zero-learning-scratch/

AlphaGo Zero

AlphaGo has become progressively more efficient thanks to hardware gains and more recently algorithmic advance
From https://deepmind.com/blog/alphago-zero-learning-scratch/

Differences

| Previous versions	AlphaGo Zero
trained from human data	without human data beyond rules
policy network and value network	one deep neural network
supervised learning	reinforcement learning

Reinforcement Learning Algorithm

——Why is AlphaGo Zero so powerful?

- A deep neutral network f_θ

- Monte Carlo tree search (MCTS)

Deep Neural Network f_θ

- $(p,v) = f_\theta(s)$

b Neural network training
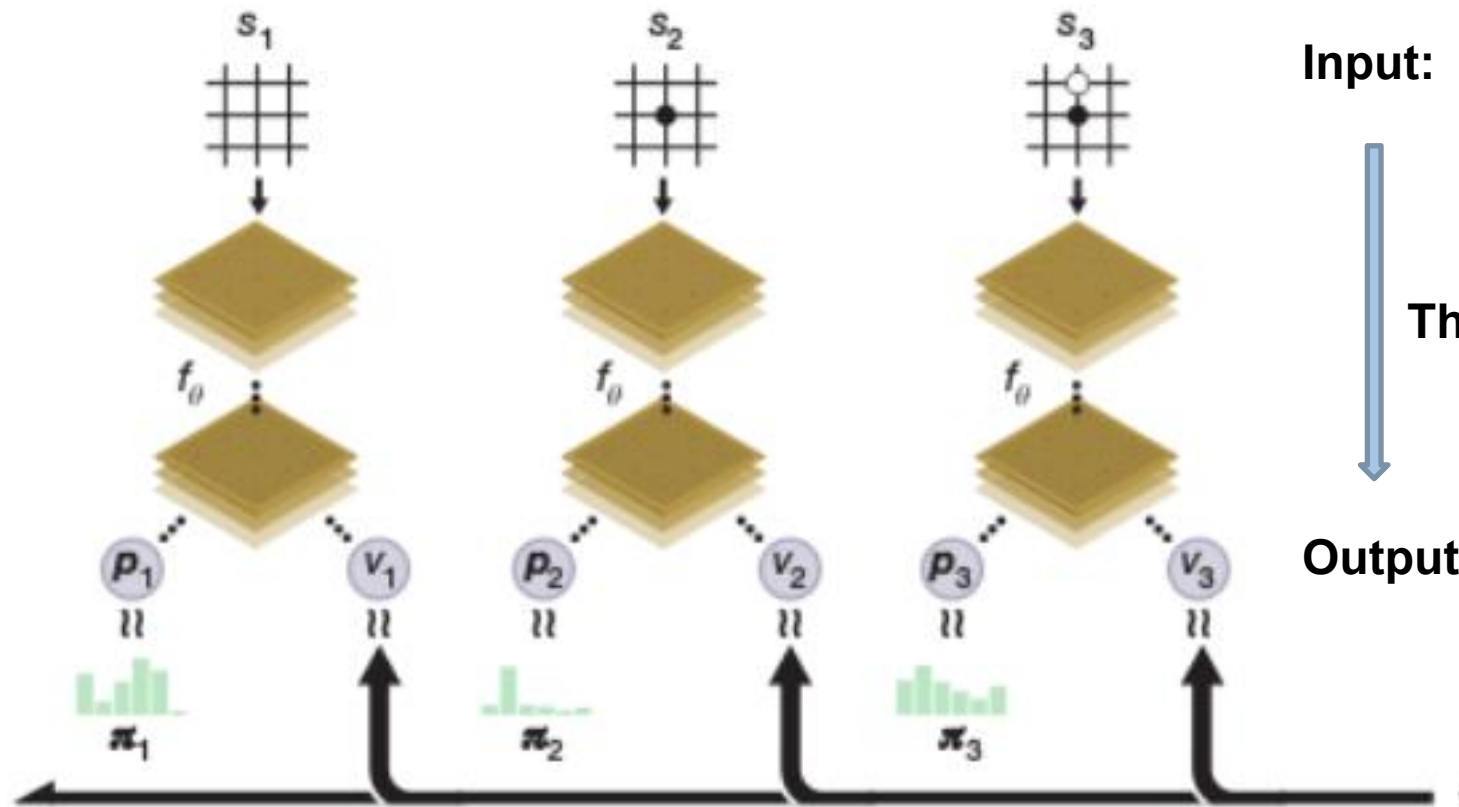
Input: **s**: position and history

Through: **θ**: parameters

Output

p: vector of move posibilities of selecting move a
$pa = Pr (a \mid s)$
v: probility of curent player wining from s

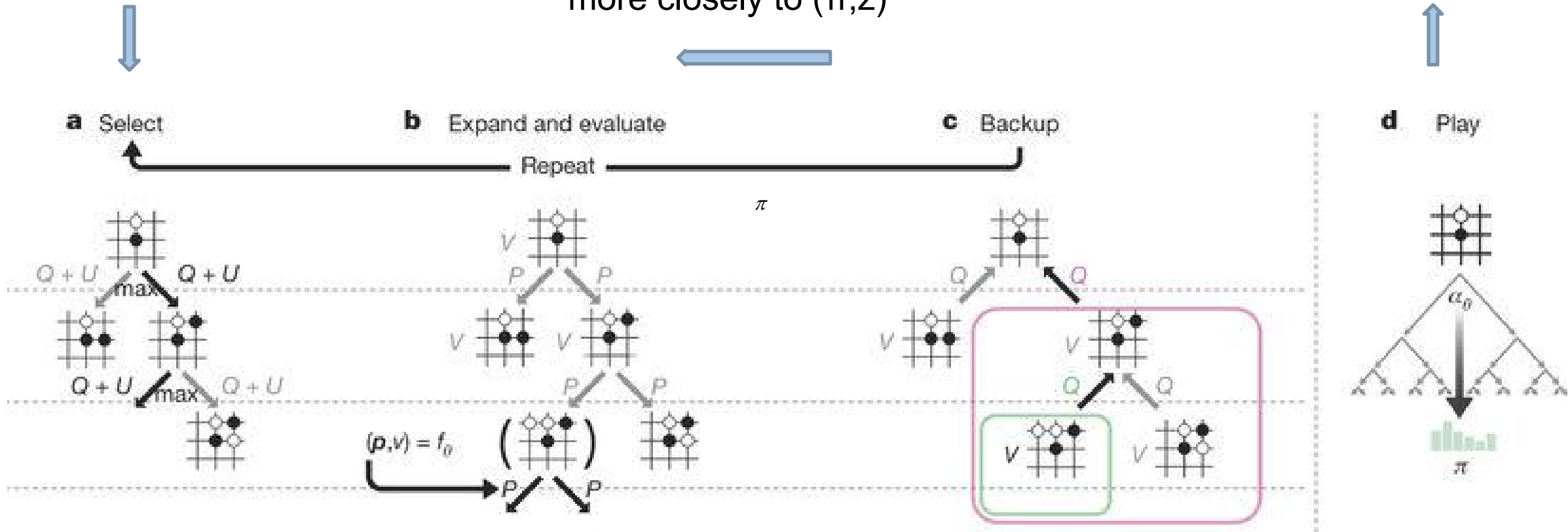Figure 1 | Self-play reinforcement learning in AlphaGo Zero. a, The

From David Silver et al Nature 550, 354–359 (19 October 2017)

Monte Carlo tree search (MCTS)

π : probability of moving

z : self-play winner

Depending on f_θ

reset (p,v) to make them more closely to (π,z)

Steps of Monte Carlo Tree Search
From https://www.nature.com/nature/journal/v550/n7676/full/nature24270.html

Self-play pipeline

s : position   a : move

π : probability of moving   z : self-play winner
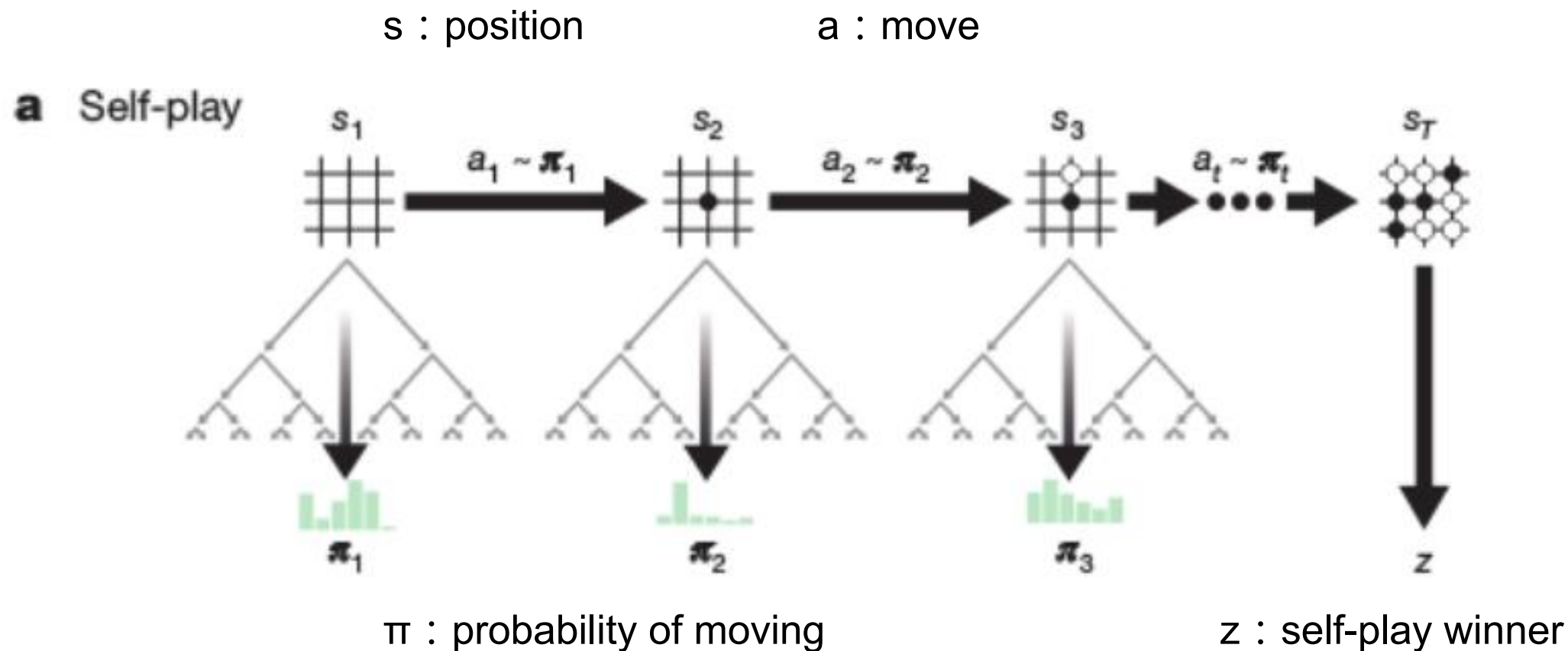
From David Silver et al Nature 550, 354–359 (19 October 2017)

Applications

- Protein folding

- Reducing energy comsumption

- Searching for new materials

Machine learning in astronomy

- Classification of Galaxies

- Redshift estimation

- Gamma-ray separation

- Determine stability of exoplanet

- Reference:
https://www.zhihu.com/question/34224788?sort=created   top5 answers

References

- https://deepmind.com/research/alphago

- https://en.wikipedia.org

- https://www.nature.com

- https://deepmind.com/research/alphago/alphago-vs-alphago-self-play-games/

- https://online-go.com/learn-to-play-go/