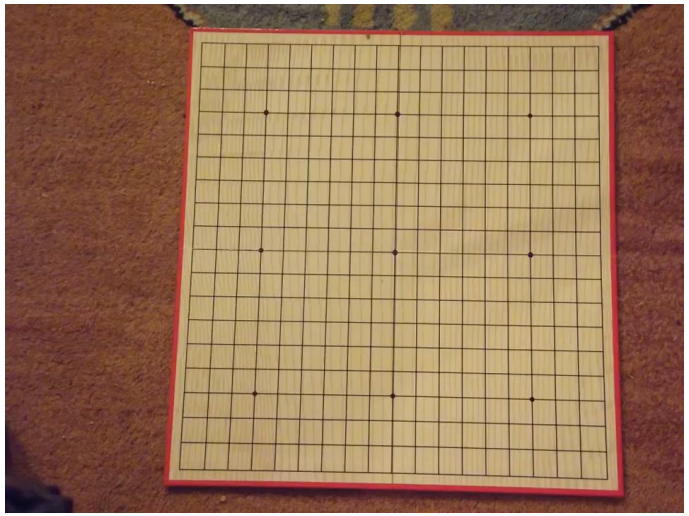
AlphaGo

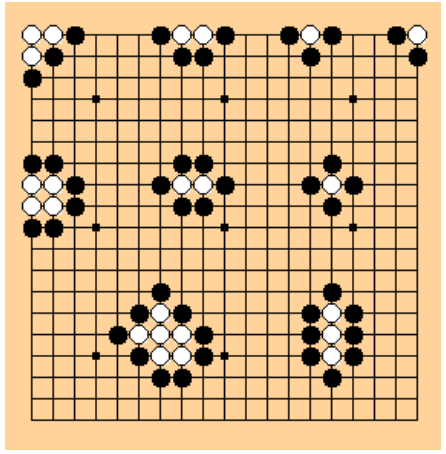
“Mastering the game of Go with deep neural networks and tree search”

What is the game of Go?

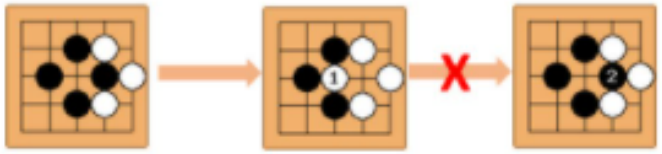
- **Objective:** Surround a larger area of the board with your own stone than your opponent.



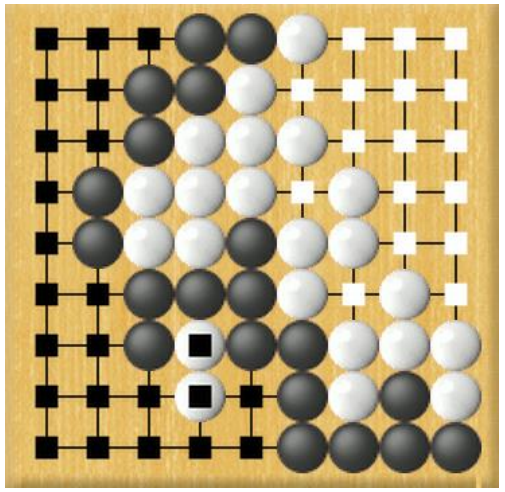
- **Rules:**
  1. **Flow of the game** – black moves first
  2. **Capture Rule** – capturing stones of your opponent by surrounding all liberties
  3. **Eternity/Ko Rule** – player is not allowed to make a move that returns the game to the previous state of the game
  4. **Scoring** – helps in evaluating points at the end of the game.



Capture Rule

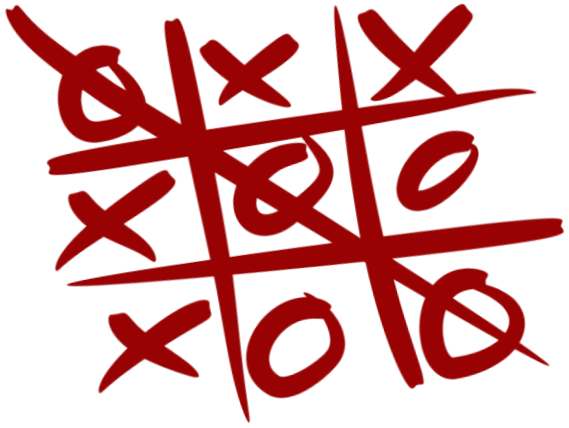


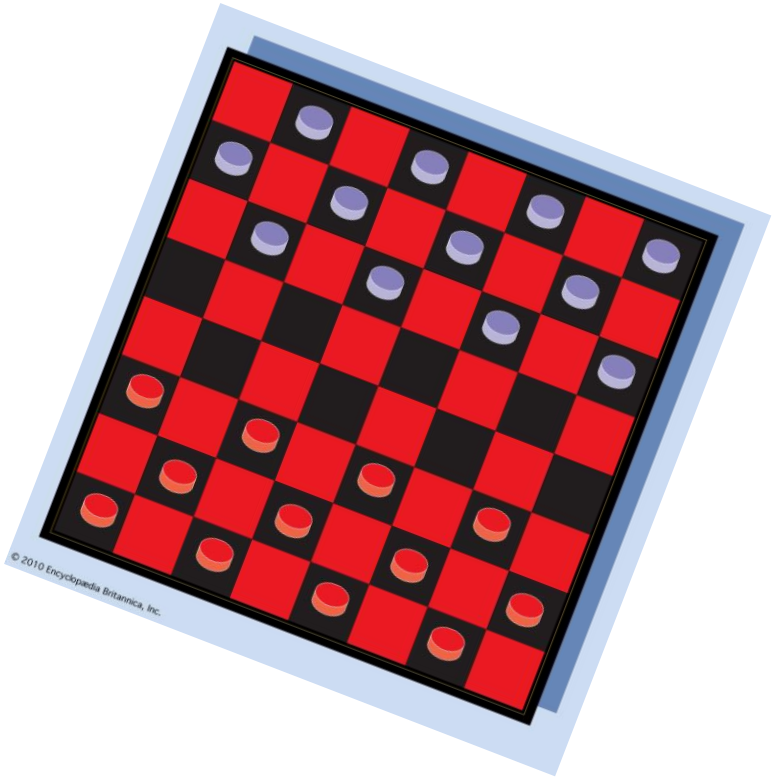
Ko Rule



Scoring Rule

Can you think of similar games as *"Go"*?





What do these games have in Common?

- **Two-Player**
  - Two players involved
- **Sequential**
  - Players take turns to move (take actions)
  - Act alternately
- **Zero-sum**
  - Utility values of the two players at the end of the game have equal absolute value and opposite sign
- **Perfect Information**
  - Each player can fully observe what actions other agents have taken

Problem Setting

To build AIs that can play any of these **“two-player sequential zero-sum perfect-information games.”**

Planning and its role in AI

- **Fundamental Question:**

  What is the optimal sequence of actions to achieve a goal?

- **Simple case: planning**

  The rules of the game(model) are given

- **Computer Go: *drosophila* of this kind of planning algorithm**

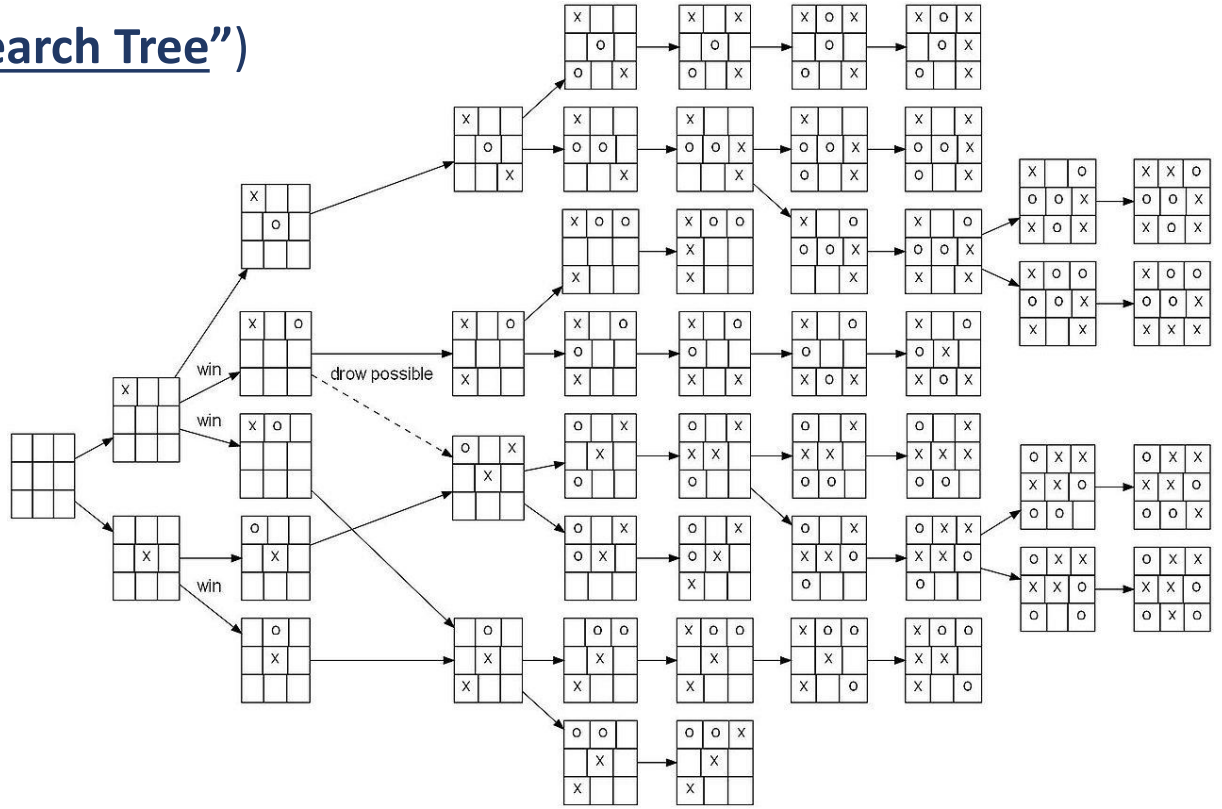
  Simple rules/model

  Can be simulated very efficiently

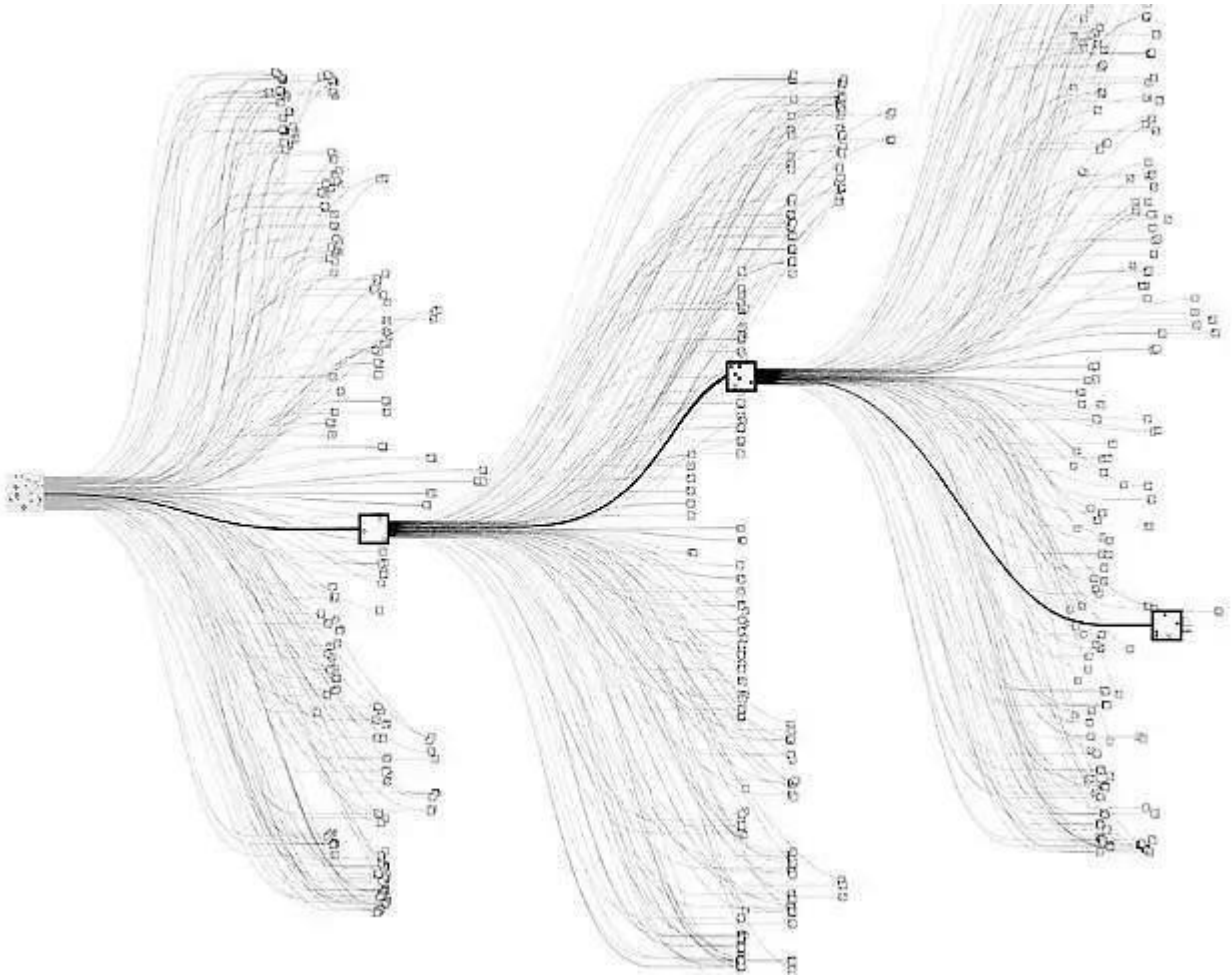
  Objective measure of progress

Define a two-player sequential zero-sum perfect-information games as a "**Search Problem**"

(can be visualized as a "**Search Tree**")



Game Tree of Go



Search Space
?????

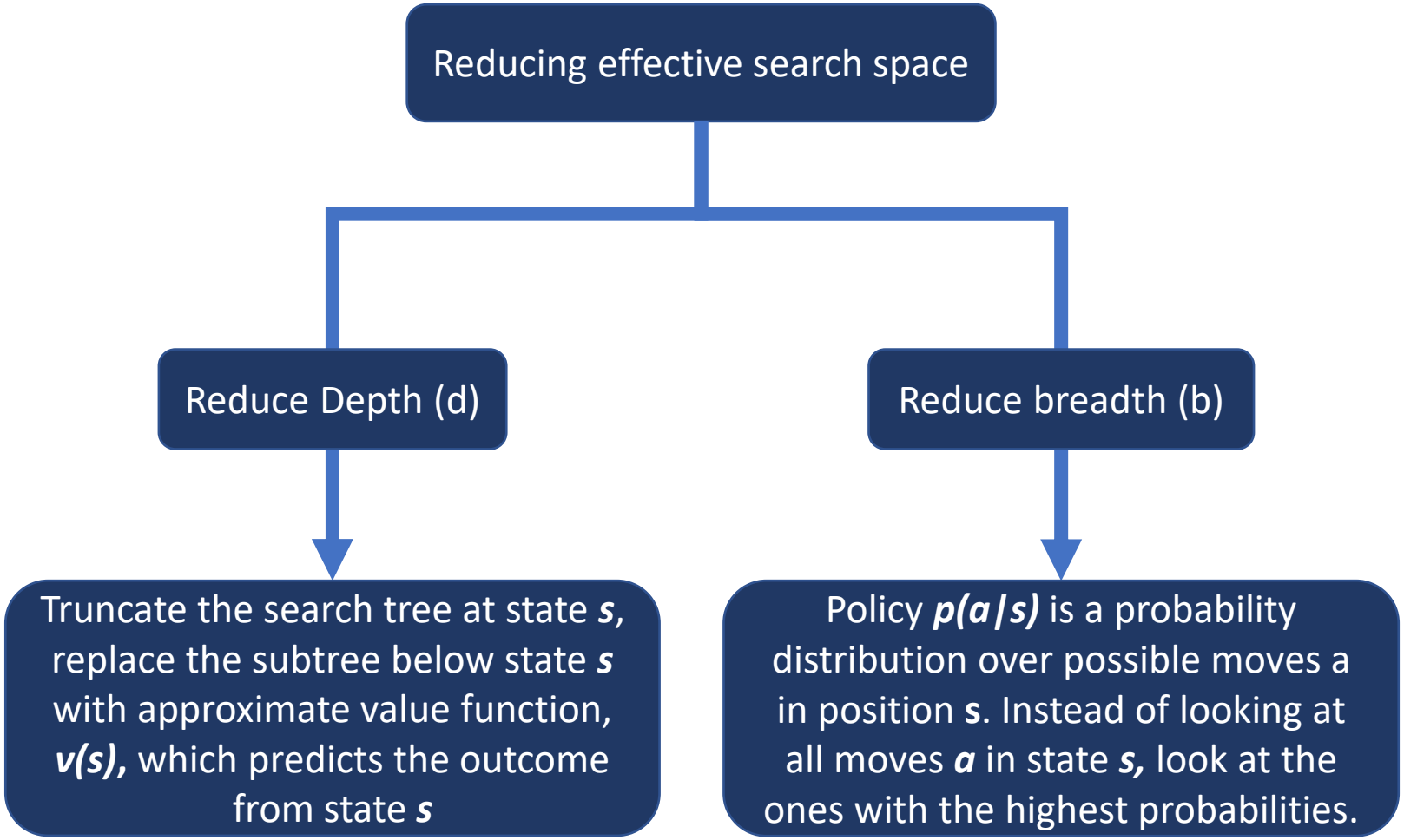
Presented by: Shubham Kamble
Supervisor: Friedrich Solowjow

Search space

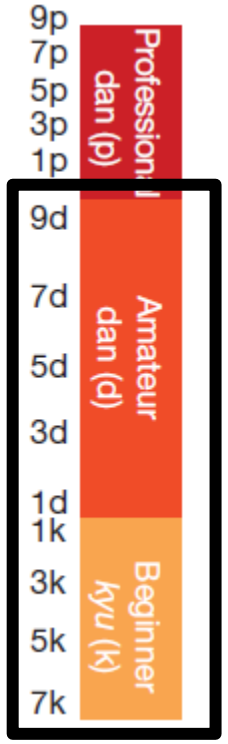
	b = breadth	d = depth	Configurations = b^d
Tic-Tac-Toe	9	9	9^9
Chess	≈35	≈80	35^{80}
Go	≈250	≈150	250^{150}

Go possesses more configurations than the number of atoms in the Universe !!

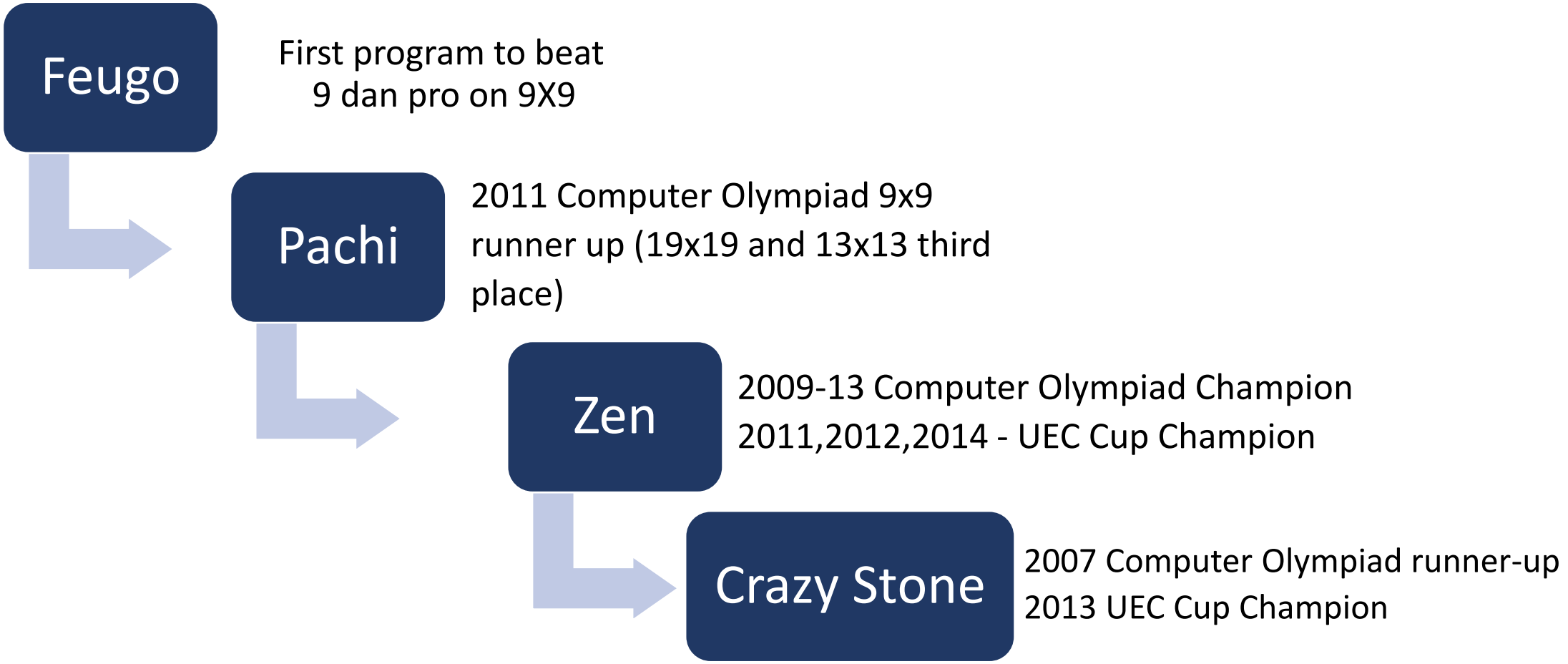
Early Awakening



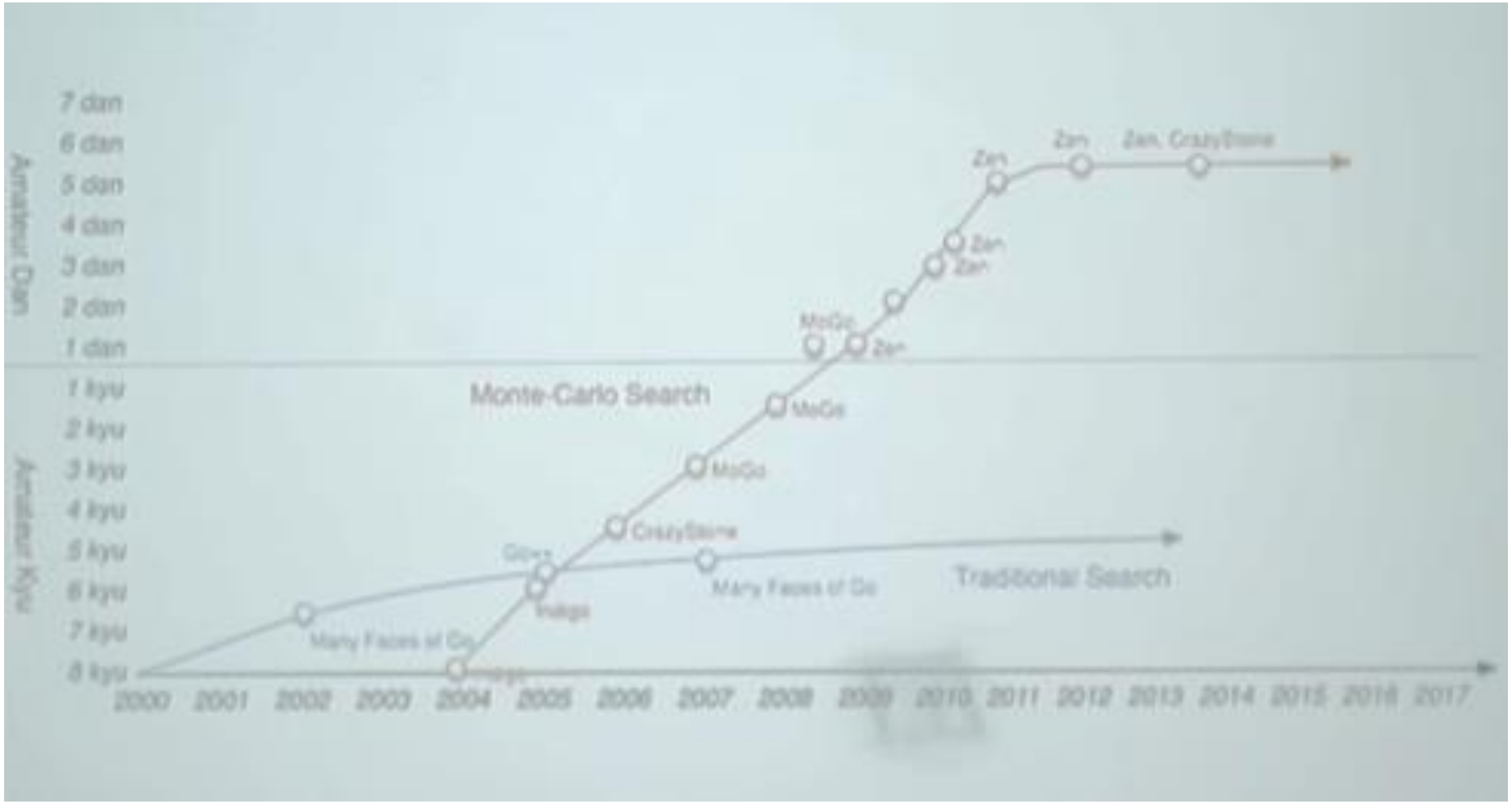
Evolving state-of-the-art approaches (1/2)



Existing program based on “Monte Carlo Tree Search”

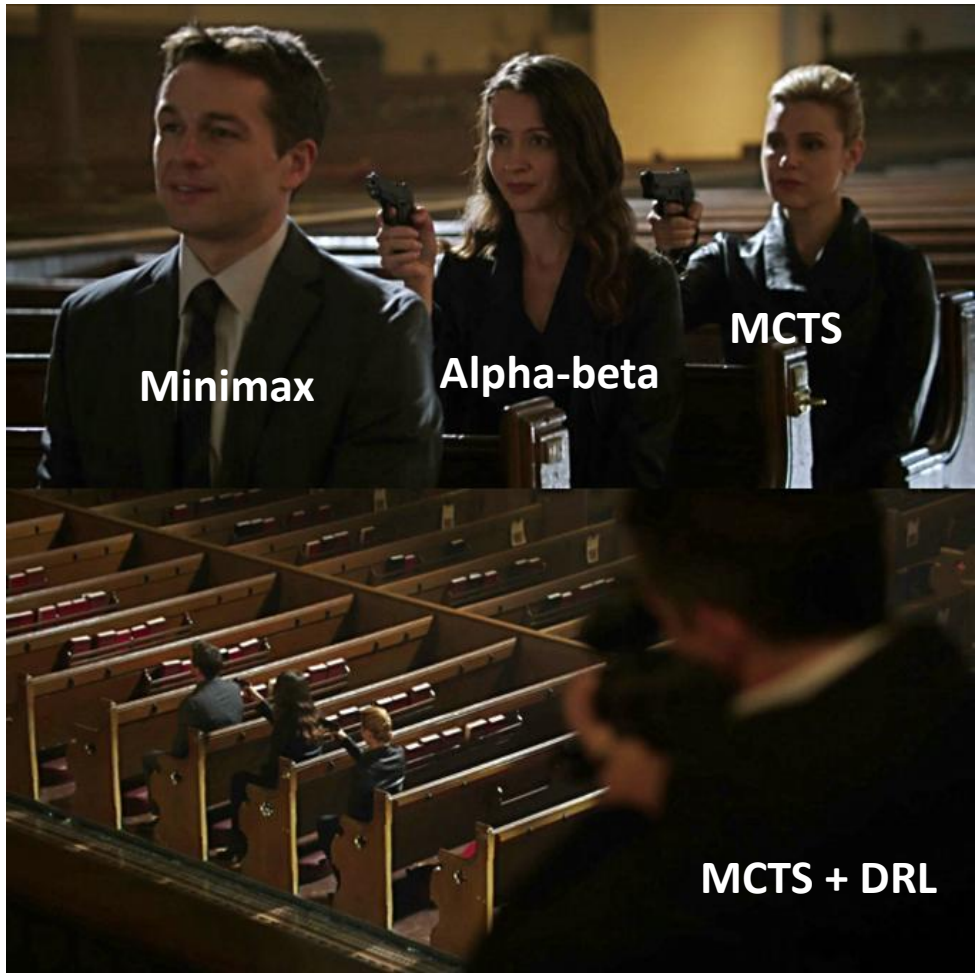


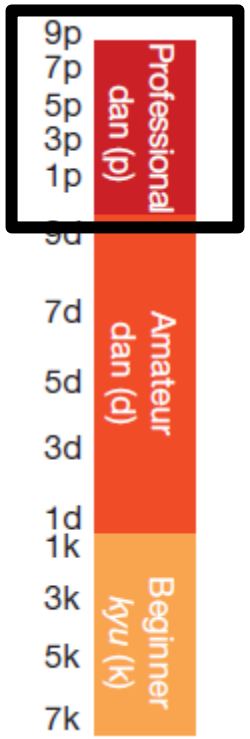
Progress in Computer Go (2000 – 2015)



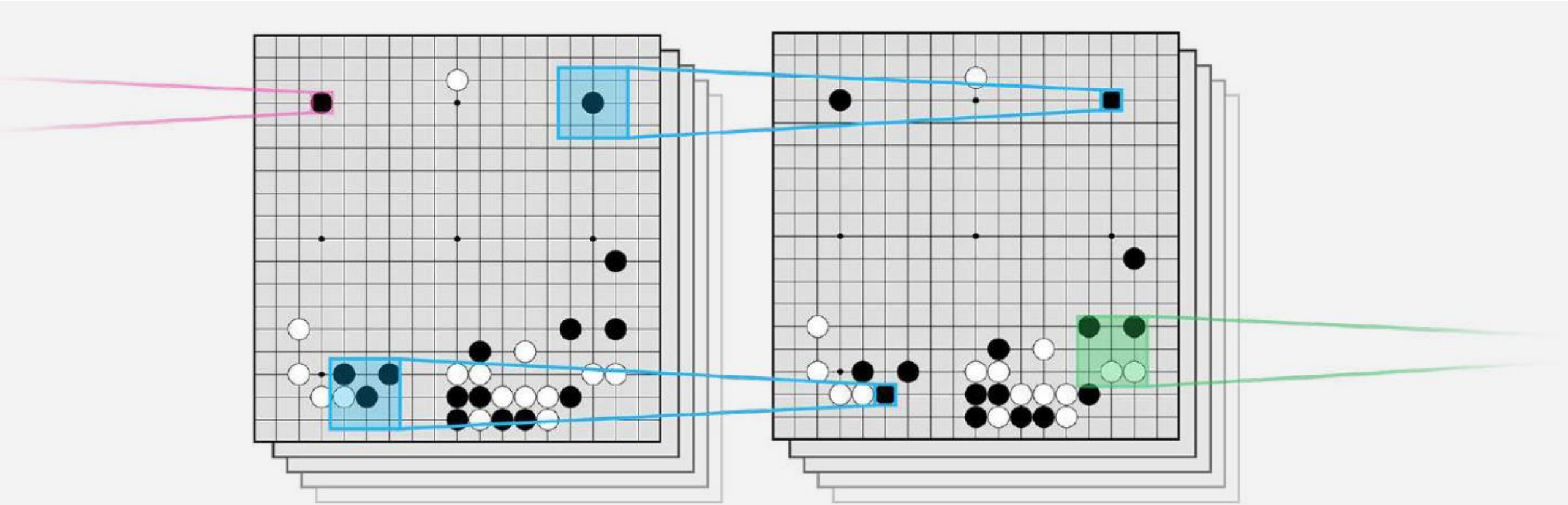
Presented by: Shubham Kamble
Supervisor: Friedrich Solowjow

Evolving state-of-the-art approaches (2/2)

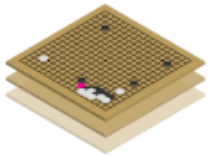




Convolutional Neural Networks



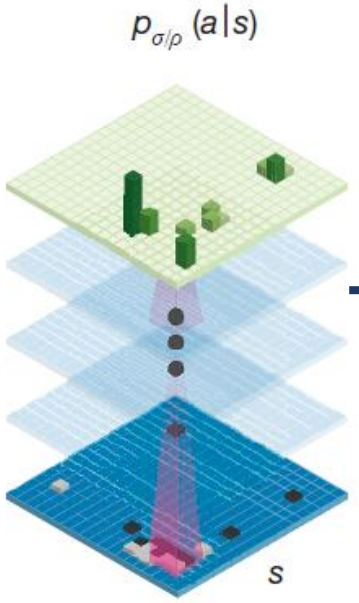
AlphaGo offline training pipeline



Human Expert Plays

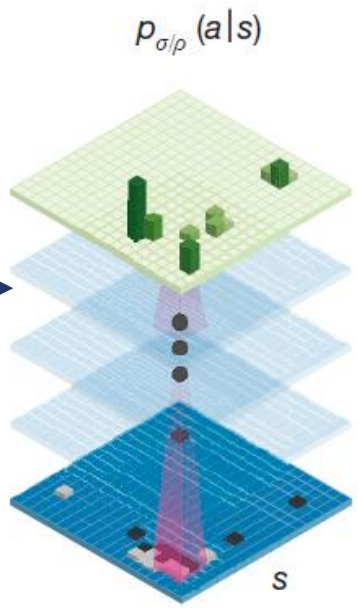
KGS Server

Supervised Learning



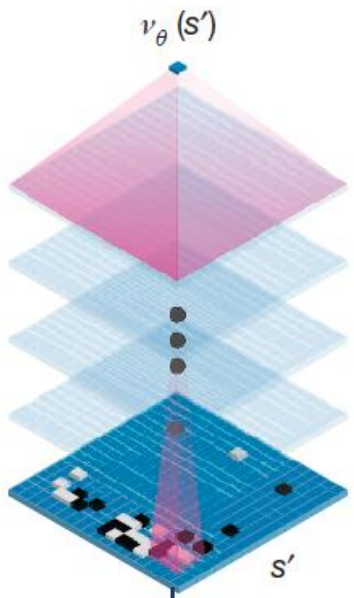
SL policy network

Policy gradient



RL policy network

Reinforcement Learning



Self play

$$\Delta\sigma \propto \frac{\partial \log p_\sigma(a|s)}{\partial\sigma}$$

Maximize Likelihood

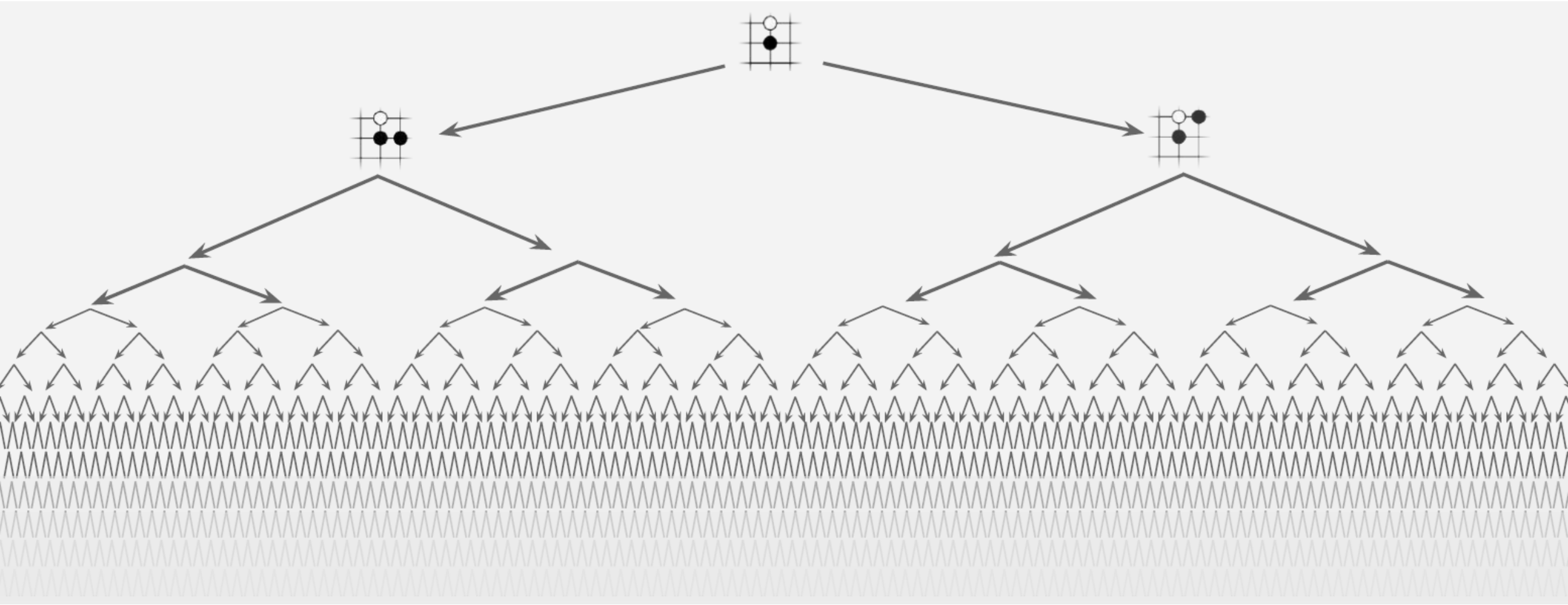
$$\Delta\rho \propto \frac{\partial \log p_\rho(a_t|s_t)}{\partial\rho} z_t$$

Maximize wins *z*

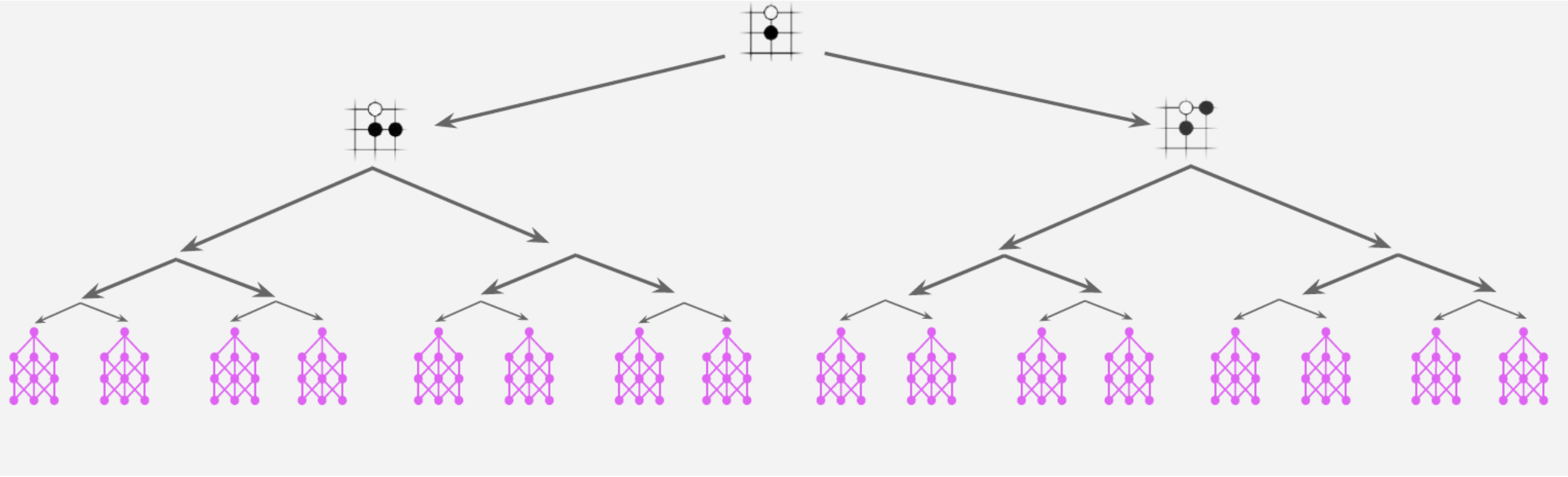
$$\Delta\theta \propto \frac{\partial v_\theta(s)}{\partial\theta}(z - v_\theta(s))$$

Minimize MSE between *z* and *v(s)*

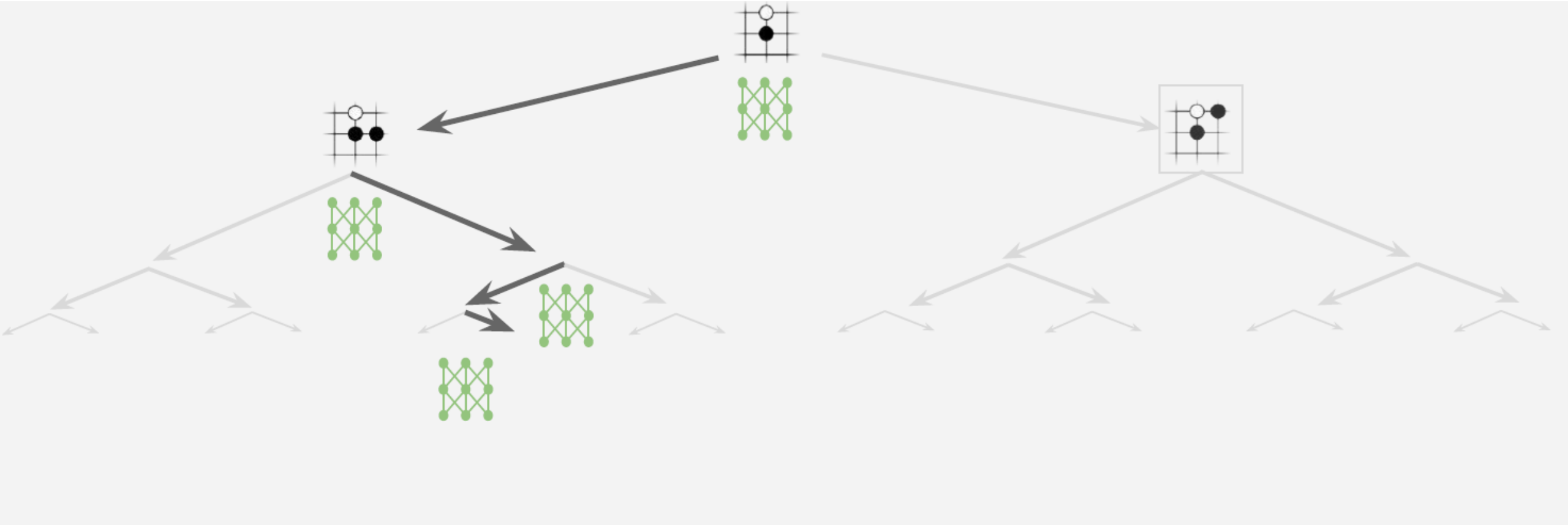
Back to how to reduce this search space?



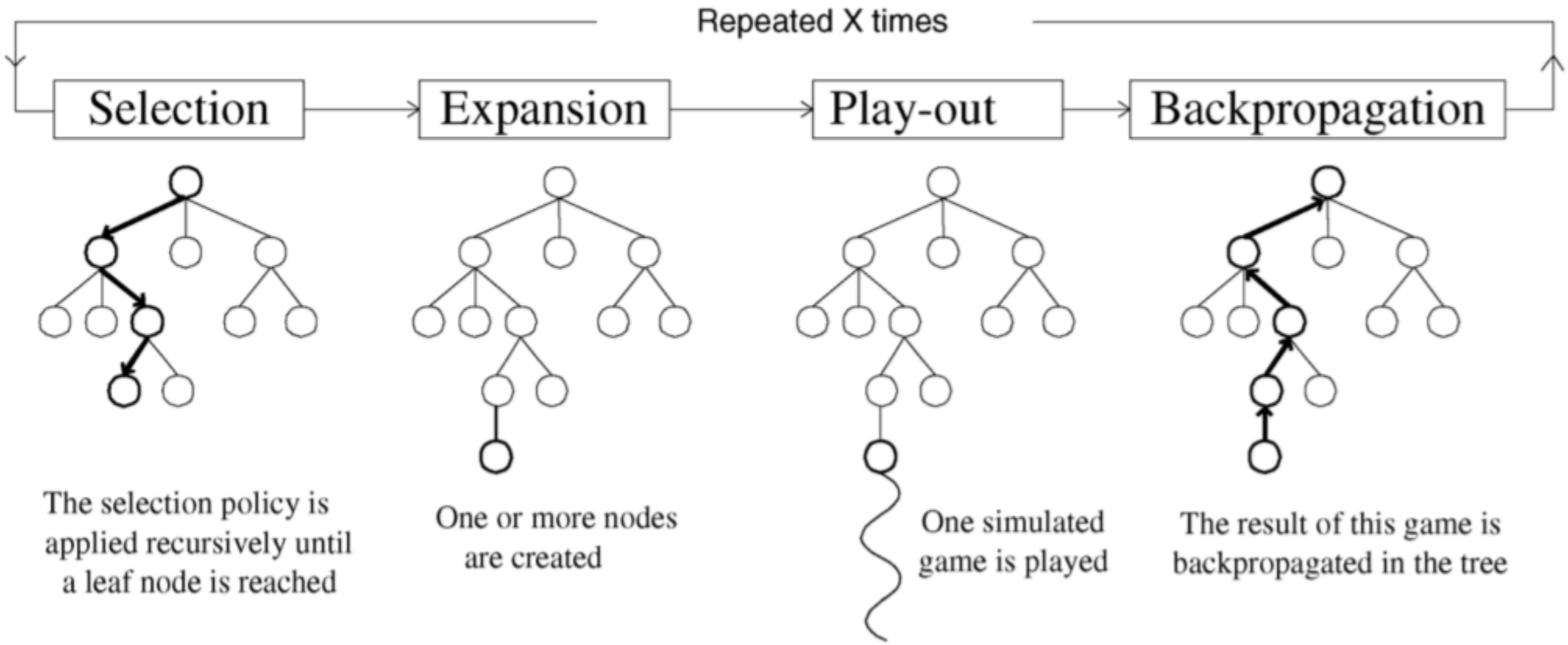
Reducing Depth with Value Network



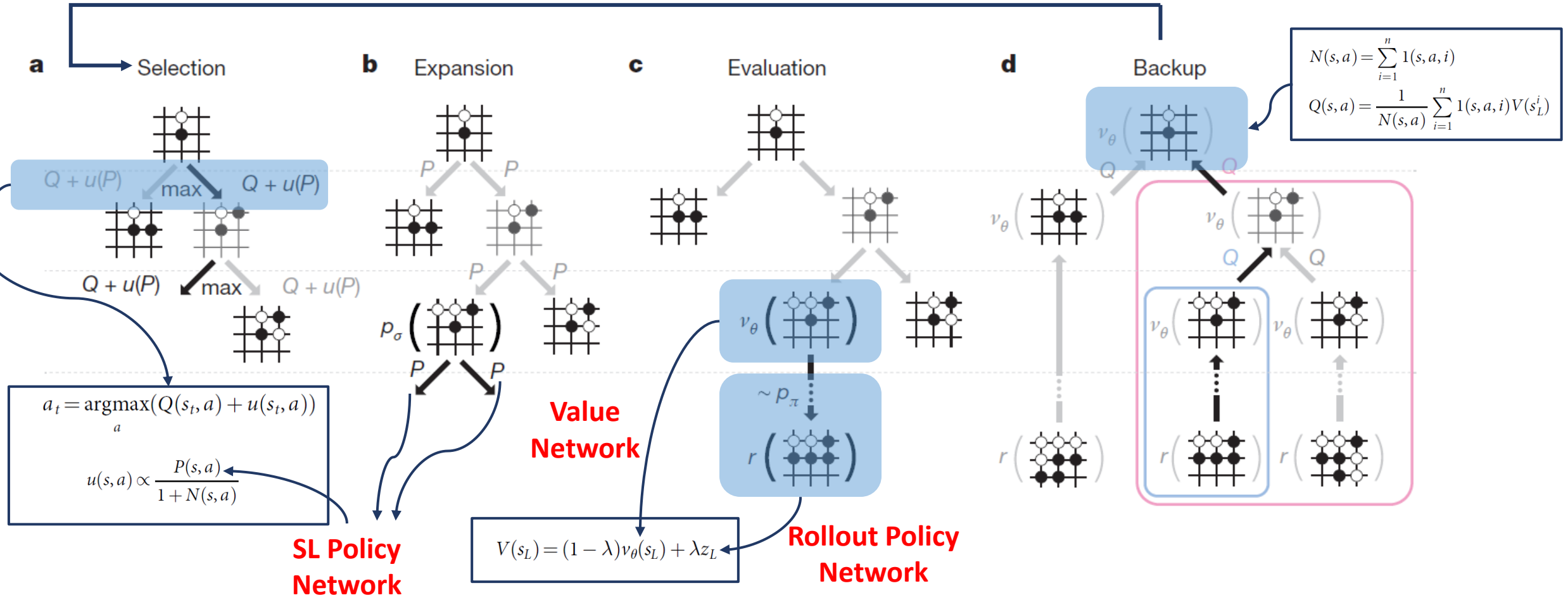
Reducing breadth with Policy Network



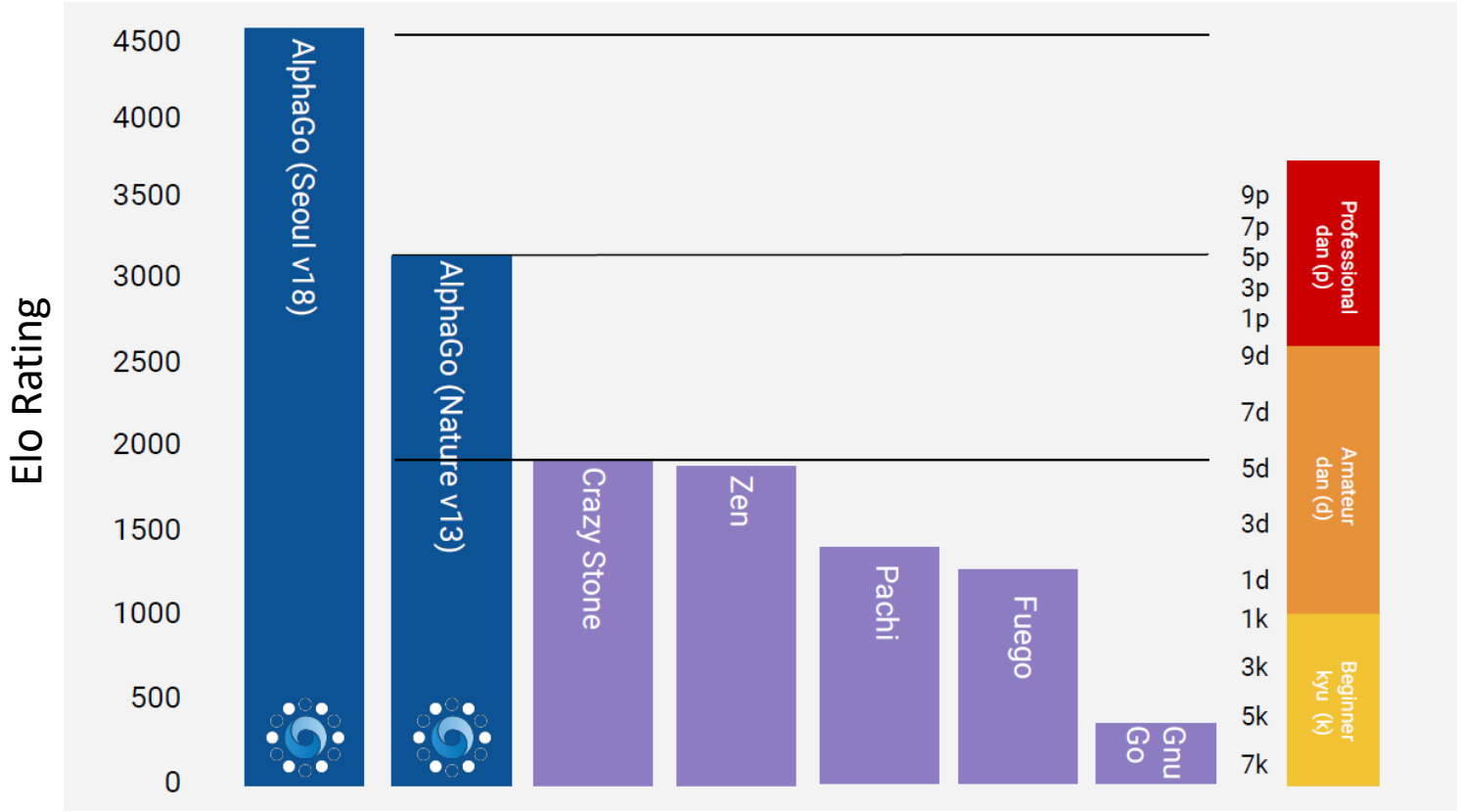
Basic Monte Carlo Tree Search



AlphaGo Monte Carlo Tree Search



Evaluating AlphaGo against computers



AlphaGo against humans

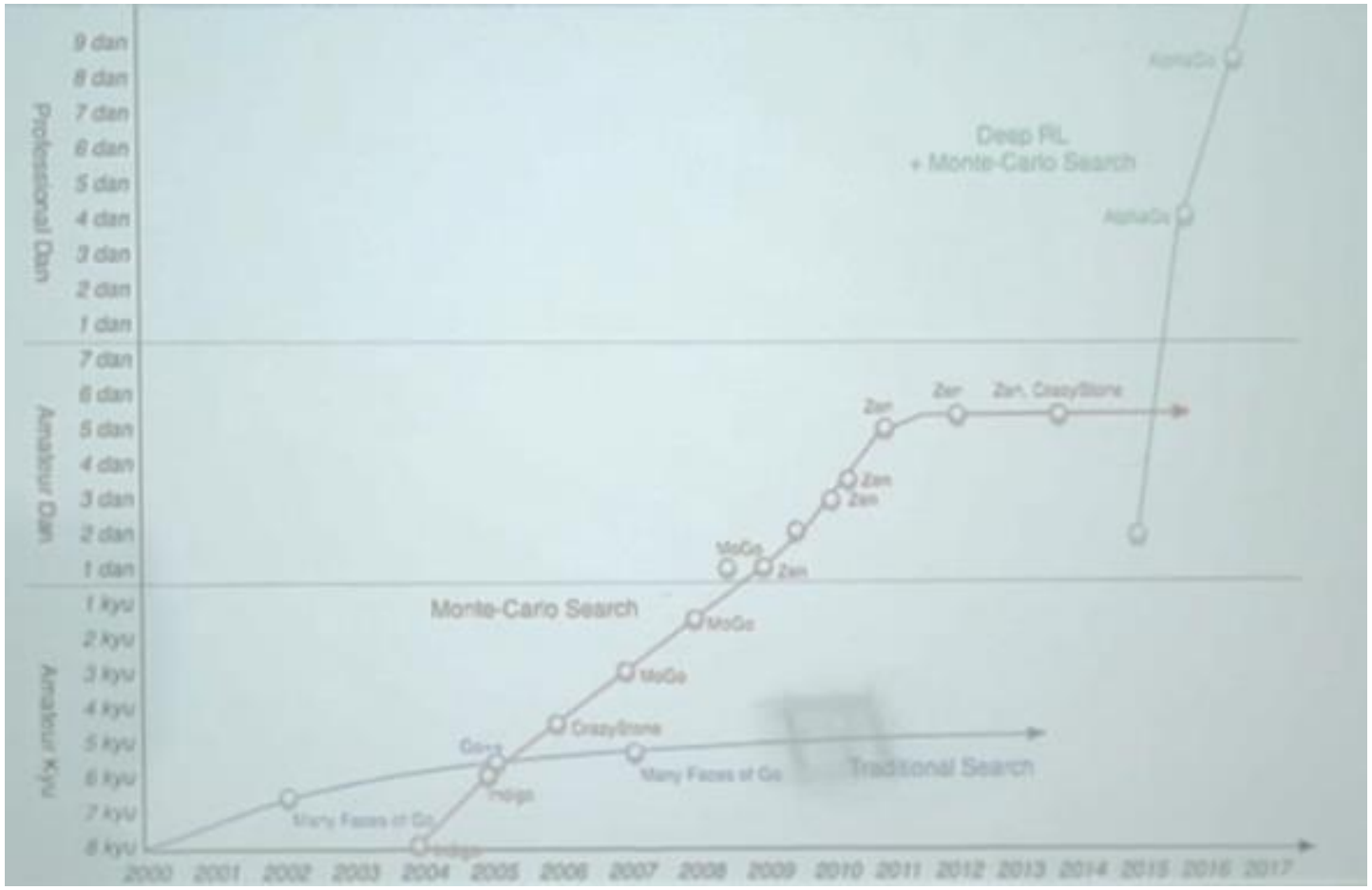
AlphaGo from DeepMind 2015: defeats a European Go Champion (Fan Hui, 2-dan professional) on a 19x19 board without handicap. It used both human and machine training.

Deep Mind's AlphaGo 2016: Defeats former world champion Lee Sedol by 4-1.

DeepMind's AlphaGo Master 2017:

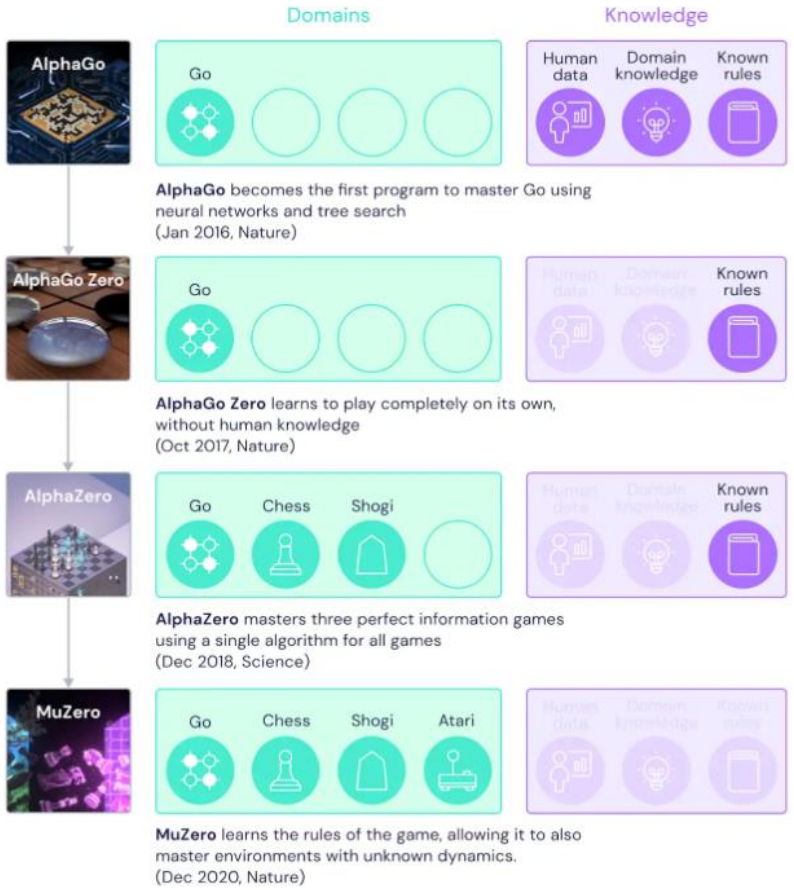
Defeats current world champion Ke Jie (also beat a 5-person team). Commentators noted that Ke appeared to borrow moves from AlphaGo 2016. But Ke noted that "AlphaGo is improving too fast" and is a different player from last year.

Progress in Computer Go (2000 – 2017)



Presented by: Shubham Kamble
Supervisor: Friedrich Solowjow

Why is AlphaGo considered the state-of-the-art ?



References

- Silver, D., Huang, A., Maddison, C. *et al.* Mastering the game of Go with deep neural networks and tree search. *Nature* **529,** 484–489 (2016).
- Schrittwieser, J., Antonoglou, I., Hubert, T. *et al.* Mastering Atari, Go, chess and shogi by planning with a learned model. *Nature* **588,** 604–609 (2020).
- Gelly, S. & Silver, D. Combining online and offline learning in UCT. In *17th International Conference on Machine Learning*, 273–280 (2007).
- Maddison, C. J., Huang, A., Sutskever, I. & Silver, D. Move evaluation in Go using deep convolutional neural networks. *3rd International Conference on Learning Representations* (2015).
- Sutton, R. & Barto, A. *Reinforcement Learning: an Introduction* (MIT Press, 1998).
- Gelly, S. *et al.* The grand challenge of computer Go: Monte Carlo tree search and extensions. *Commun. ACM* **55,** 106–113 (2012).
- KGS. Rating system math. http://www.gokgs.com/help/rmath.html.
- AlphaGo - The Movie | Full award-winning documentary - https://www.youtube.com/watch?v=WXuK6gekU1Y&ab_channel=DeepMind
- ICML 2017: Test of Time Award (Sylvain Gelly & David Silver) - https://www.youtube.com/watch?v=Bm7zah_LrmE&t=448s&ab_channel=DeepMind
- Tesla's AI Day video - https://youtube.com/watch?v=j0z4FweCy4M

Questions ???