Trustworthy AI Autonomy

M5-2 Trustworthy RL-Interpretability

Ding Zhao

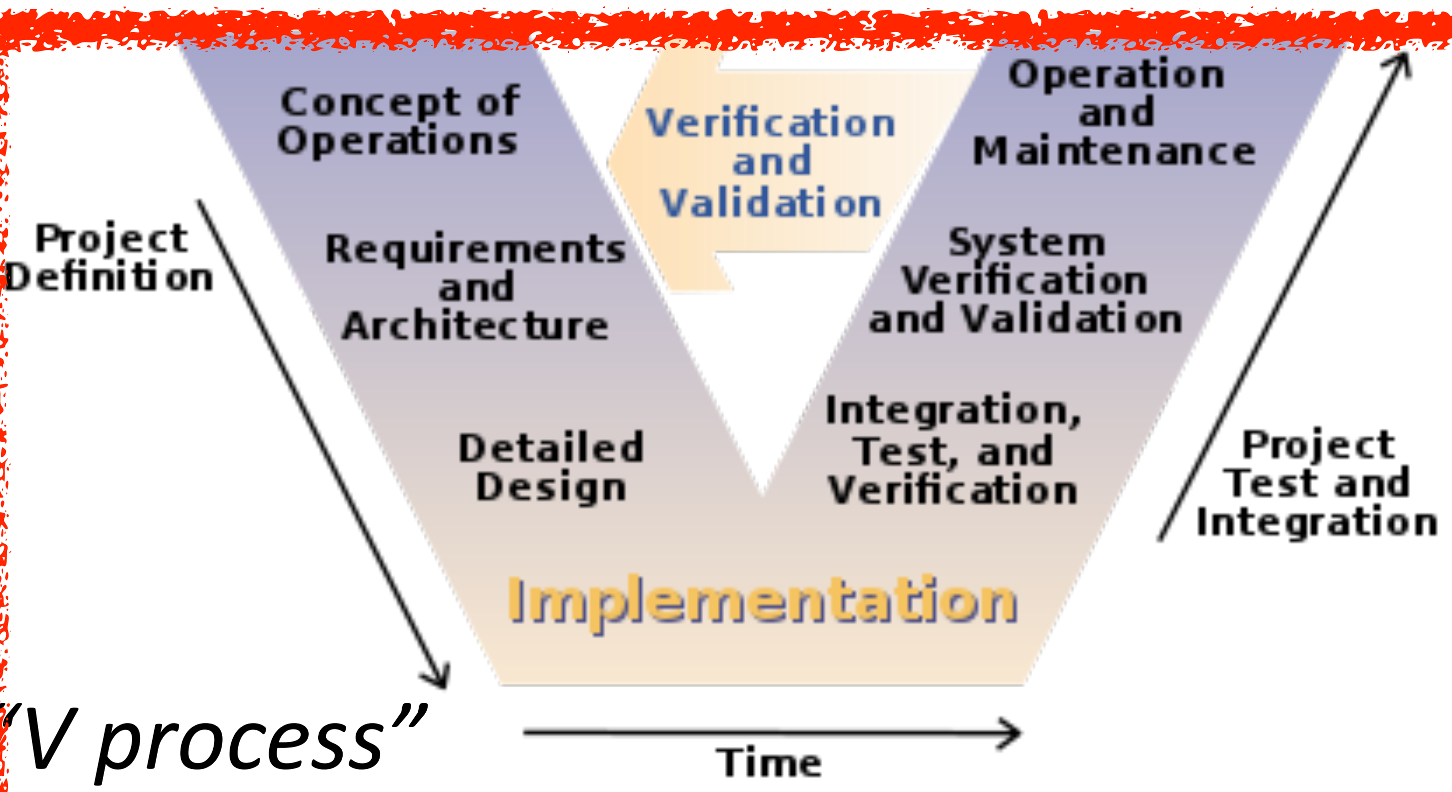
Assistant Professor

Carnegie Mellon University

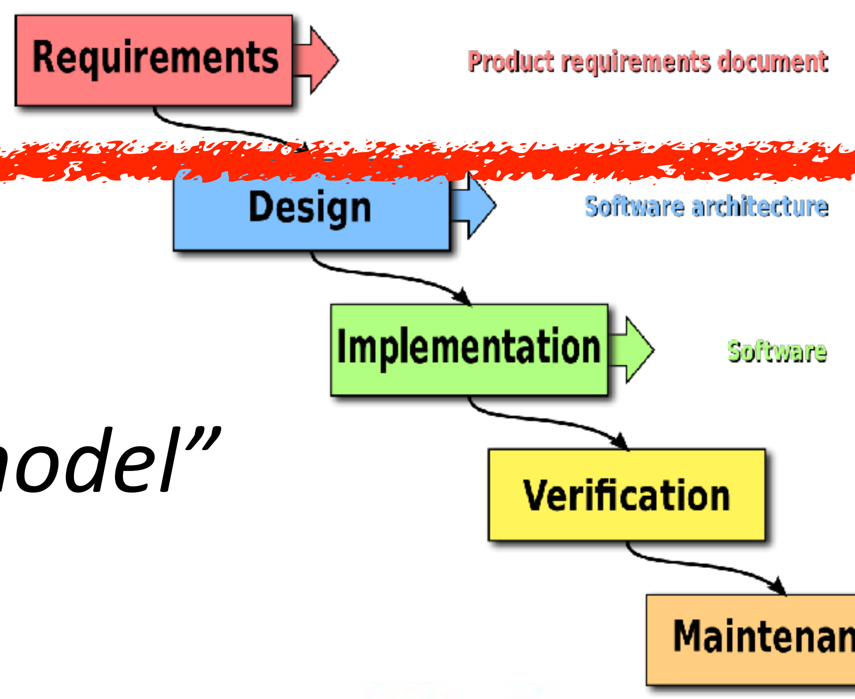
Carnegie Mellon University

Safe AI Lab @CMU

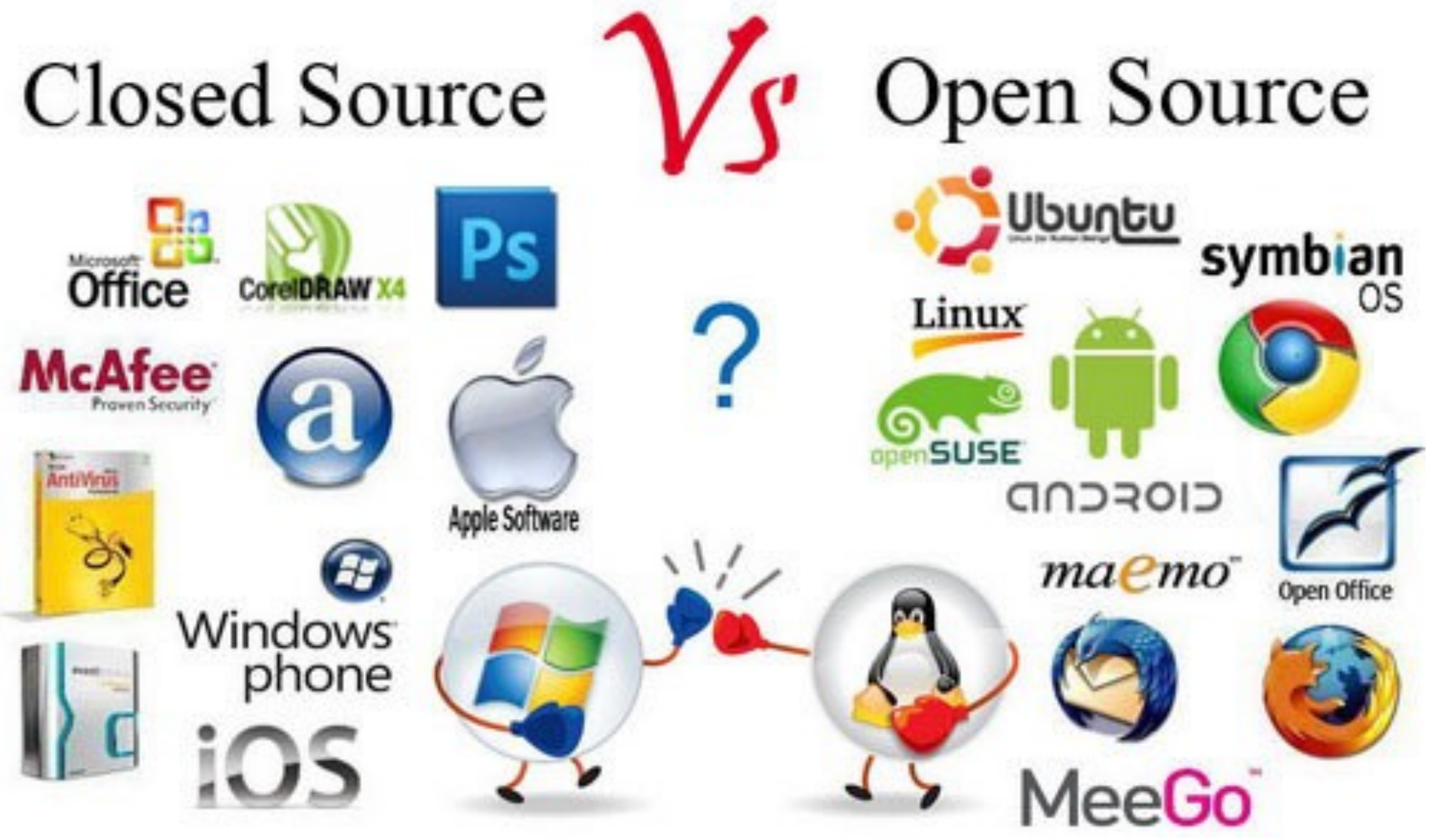
We are on the cusp to revolute the way to make machines



"V process"
Wikipedia



"Waterfall model"



Connected

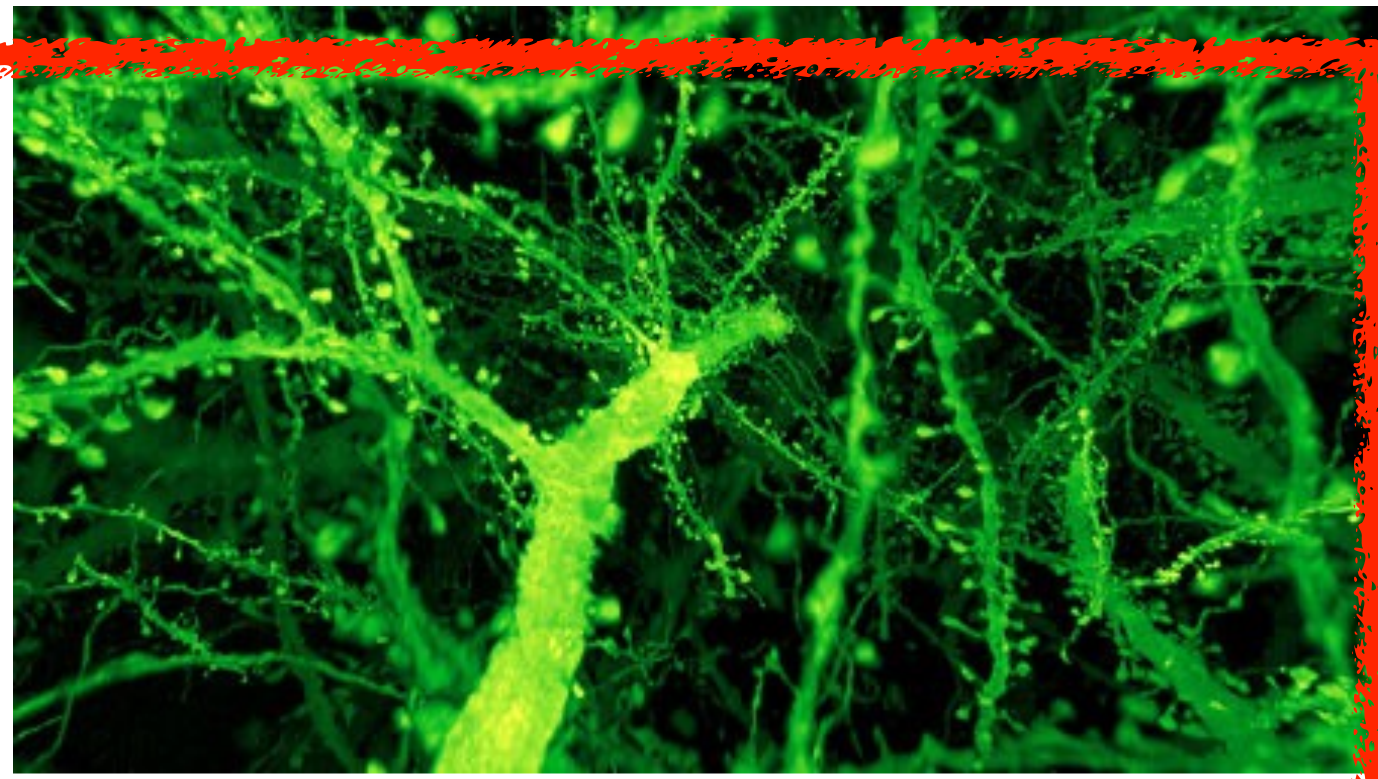
By complex structures

Evolving

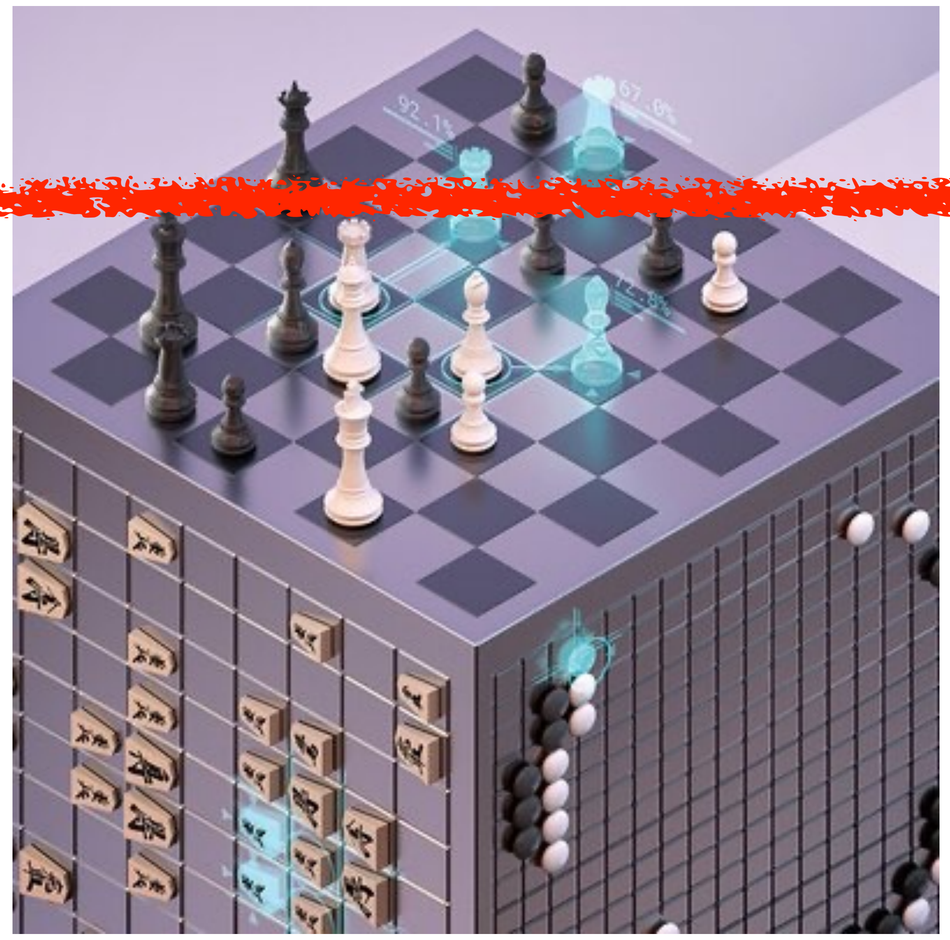
In a self-supervised way

Sharing

With blackboxes and uncertainty



Neural Network
[Science, 2019]



Reinforcement Learning
[Science, 2018]

Open Code/data
[Science, 2017]

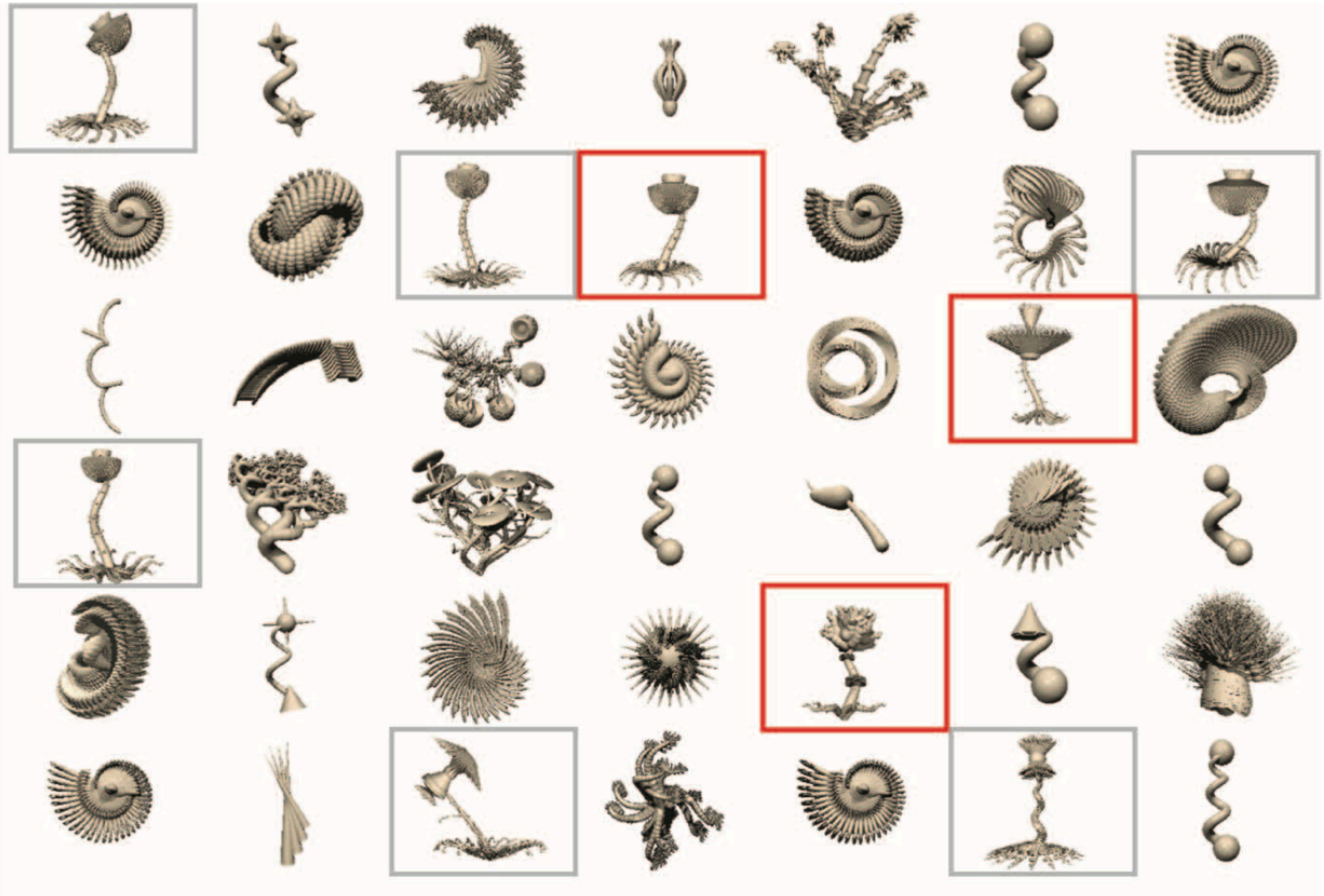
Contents

- Hierarchical AI structures
- Trees
  - Decision trees
  - Random tree/forests
  - Monte Carlo Tree search, Alpha Go
- Hierarchical RL
  - Manager-worker
  - Option/Semi-MDP
- Hierarchical structures in Meta learning
  - Neural Processes

Generalizations of concepts

- Human children learning names for object concepts routinely make strong generalizations from just a few examples. The same processes of rapid generalization can be studied in adults learning names for novel objects created with computer graphics.

- Given these alien objects and three examples (boxed in red) of “tufas” (a word in the alien language), which other objects are tufas? Almost everyone selects just the objects boxed in gray.



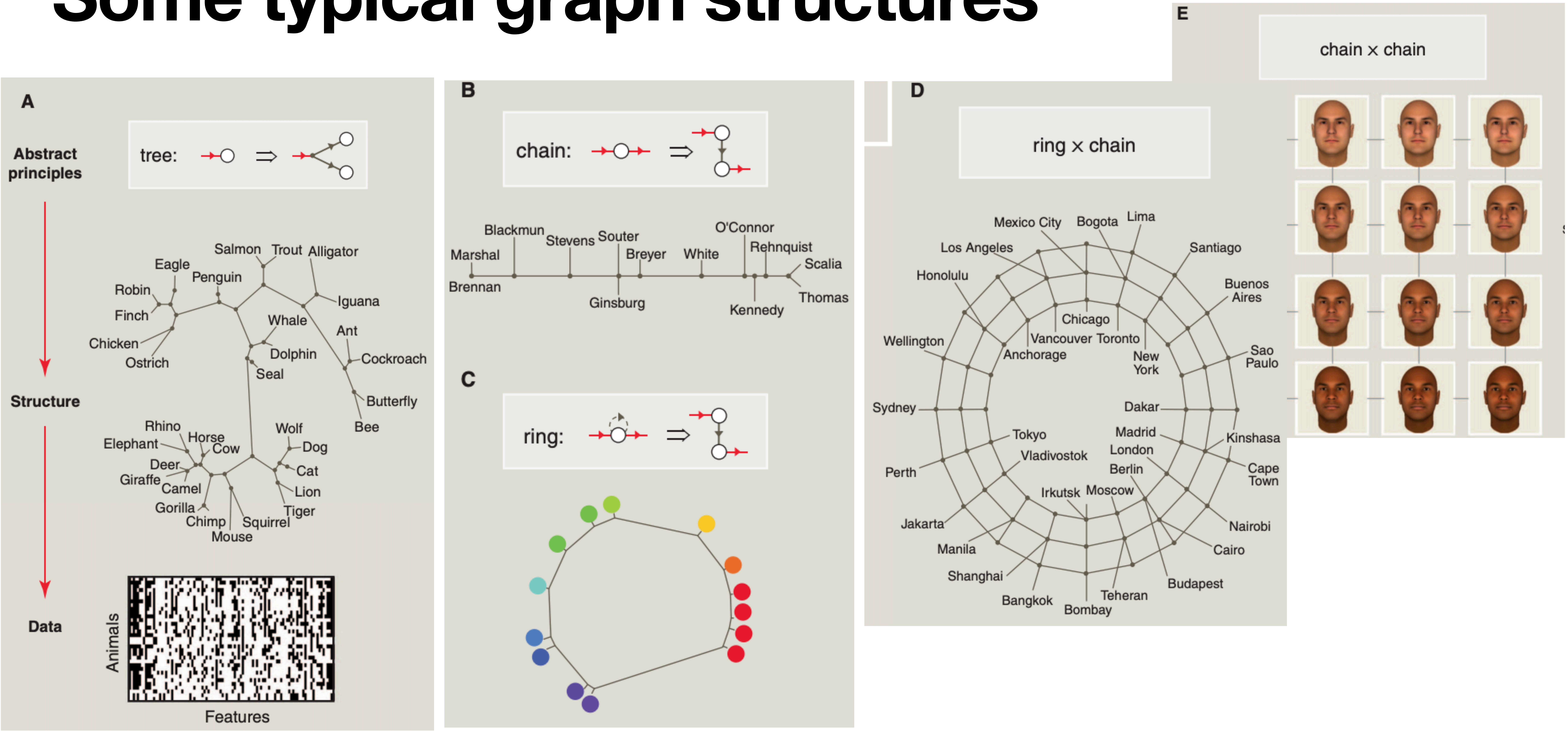
Tenenbaum, Joshua B., et al. "How to grow a mind: Statistics, structure, and abstraction." *science* 331.6022 (2011): 1279-1285.

Generalizations of concepts

- Learning names for categories can be modeled as (Bayesian) inference over a tree-structured domain representation.
- Objects are placed at the leaves of the tree, and hypotheses about categories that words could label correspond to different branches.
- Branches at different depths pick out hypotheses at different levels of generality.

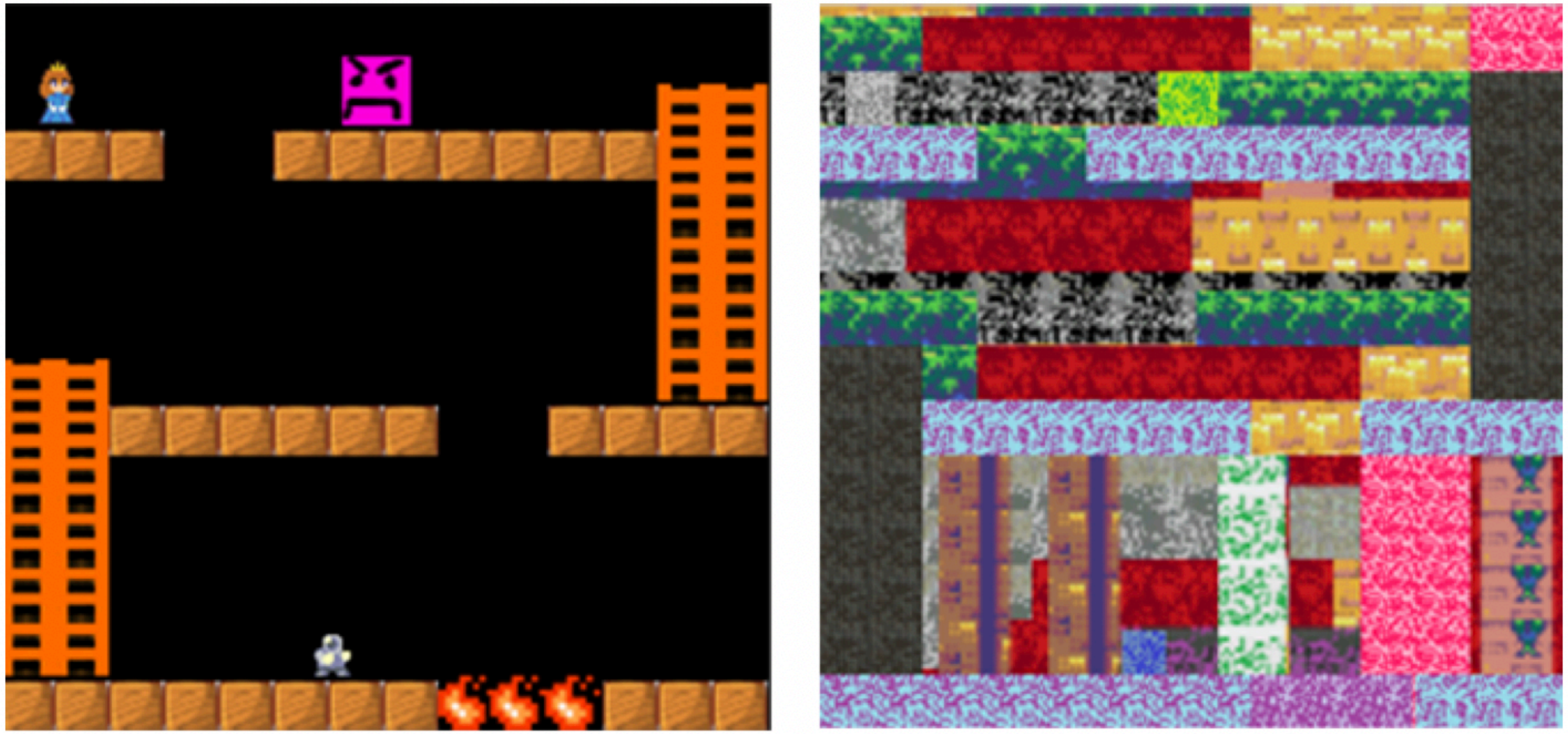
Tenenbaum, Joshua B., et al. "How to grow a mind: Statistics, structure, and abstraction." *science* 331.6022 (2011): 1279-1285.

Some typical graph structures



Tenenbaum, Joshua B., et al. "How to grow a mind: Statistics, structure, and abstraction." *science* 331.6022 (2011): 1279-1285.

Human may not always be able to generalize

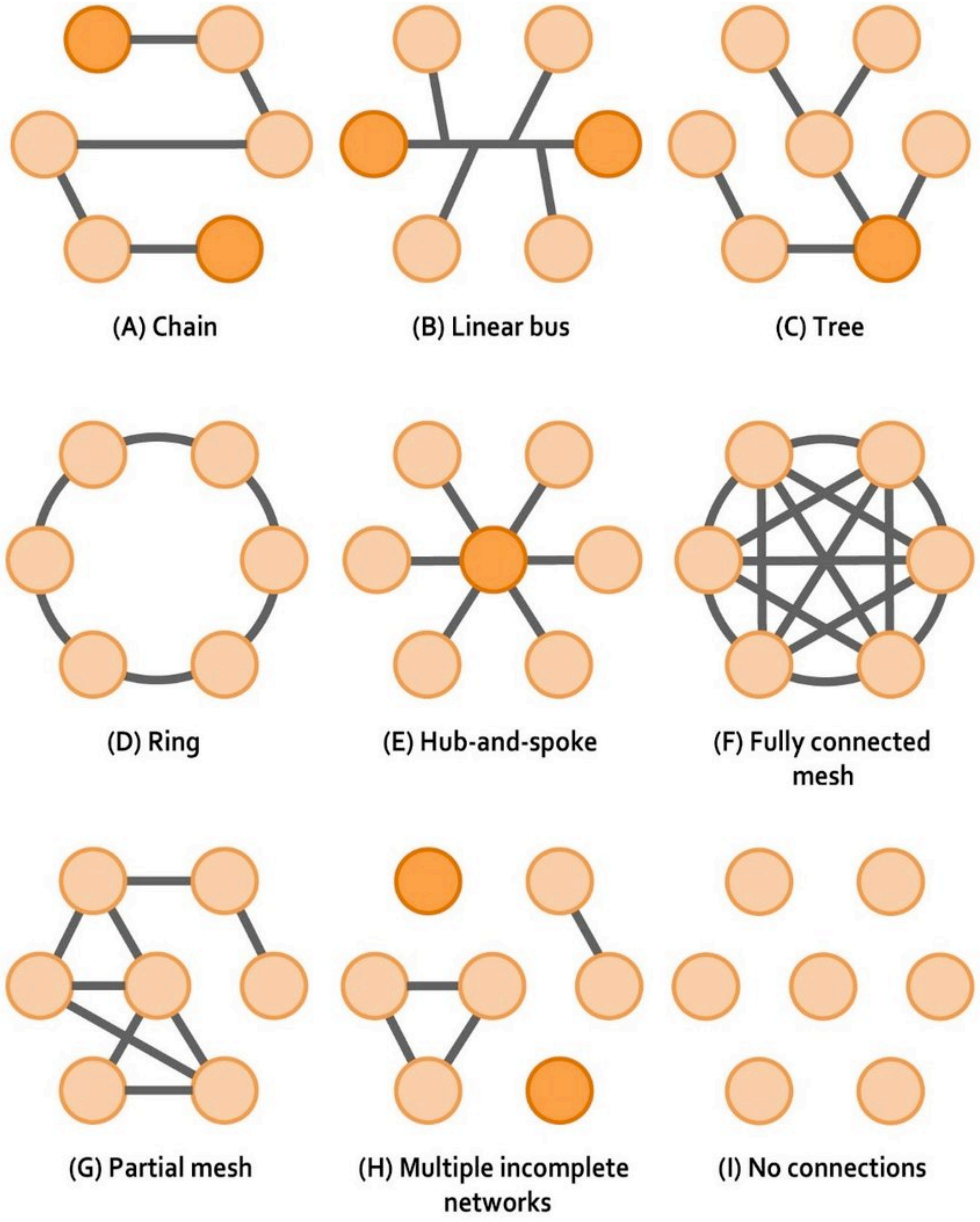


(a) Original Game (b) Modified Game

Tenenbaum, Joshua B., et al. "How to grow a mind: Statistics, structure, and abstraction." *science* 331.6022 (2011): 1279-1285.
Frank, Michael C., Edward Vul, and Scott P. Johnson. 2009. "Development of Infants' Attention to Faces during the First Year." *Cognition* 110 (2): 160–70.
https://www.medicalnewstoday.com/articles/baby-sticking-tongue-out

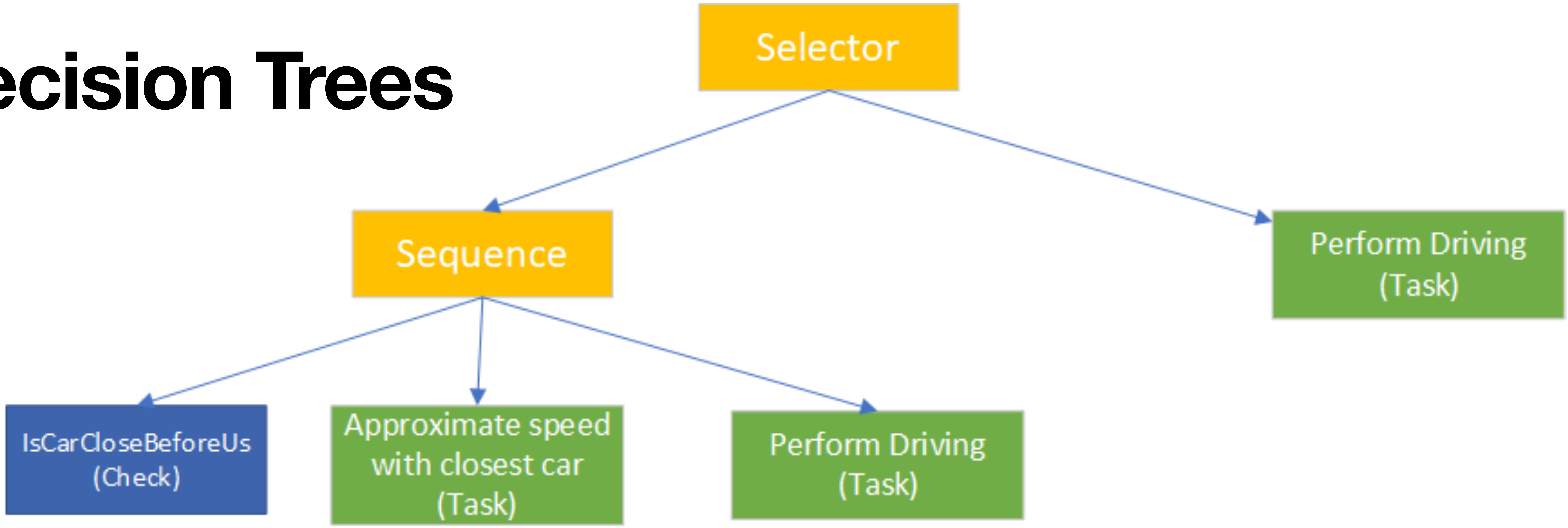
Graph grammars to describe knowledge

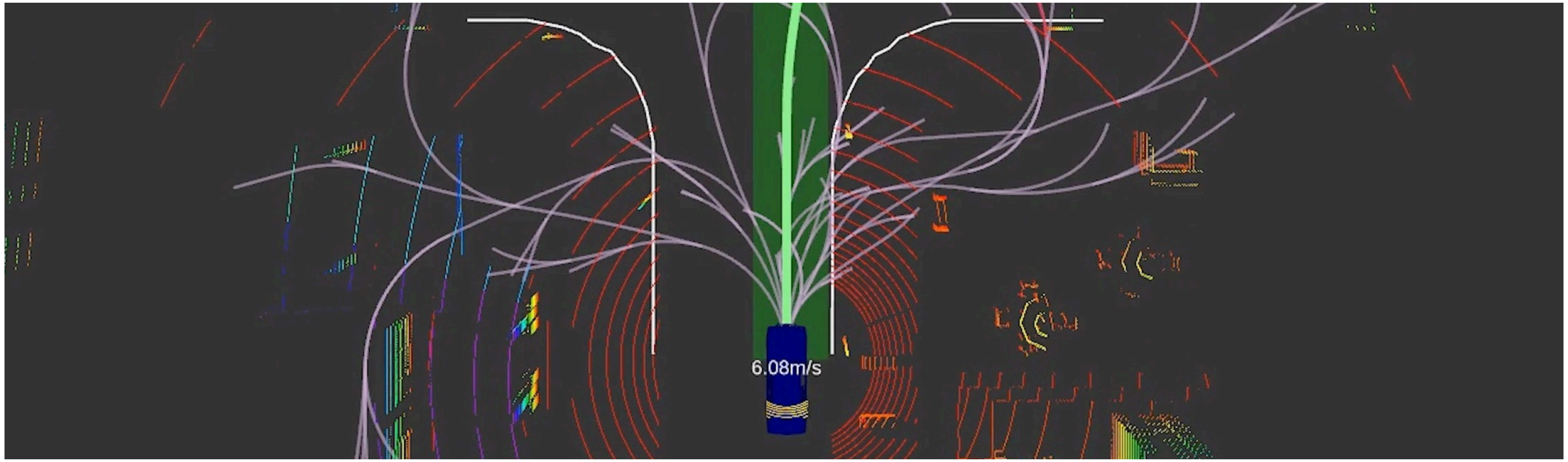
- A graph G is a set of nodes (vertices) connected by directed/undirected edges.
- This is a very flexible data structure
  - If there are no edges, then it becomes a set.
  - A tree is an undirected graph in which any two vertices are connected by exactly one path.
  - A forest is an undirected graph in which any two vertices are connected by at most one path.



Gladden, Matthew E. Neuroprosthetic Supersystems Architecture: Considerations for the Design and Management of Neurocybernetically Augmented Organizations. Synthypnion Academic, 2017.

Decision Trees





https://medium.com/behavior-trees-for-path-planning-autonomous/behavior-trees-for-path-planning-autonomous-driving-8db1575fec2c
https://spectrum.ieee.org/cars-that-think/transportation/self-driving/realtime-robotics-motion-planning-chip-autonomous-cars

Decision Trees vs Random Forests

- Issues of decision trees: overfitting
- Random forests could avoid this by
    - training with a random subset of data (bootstrapping)
    - Randomly select a subset of attributes
    - Take an aggregation of results

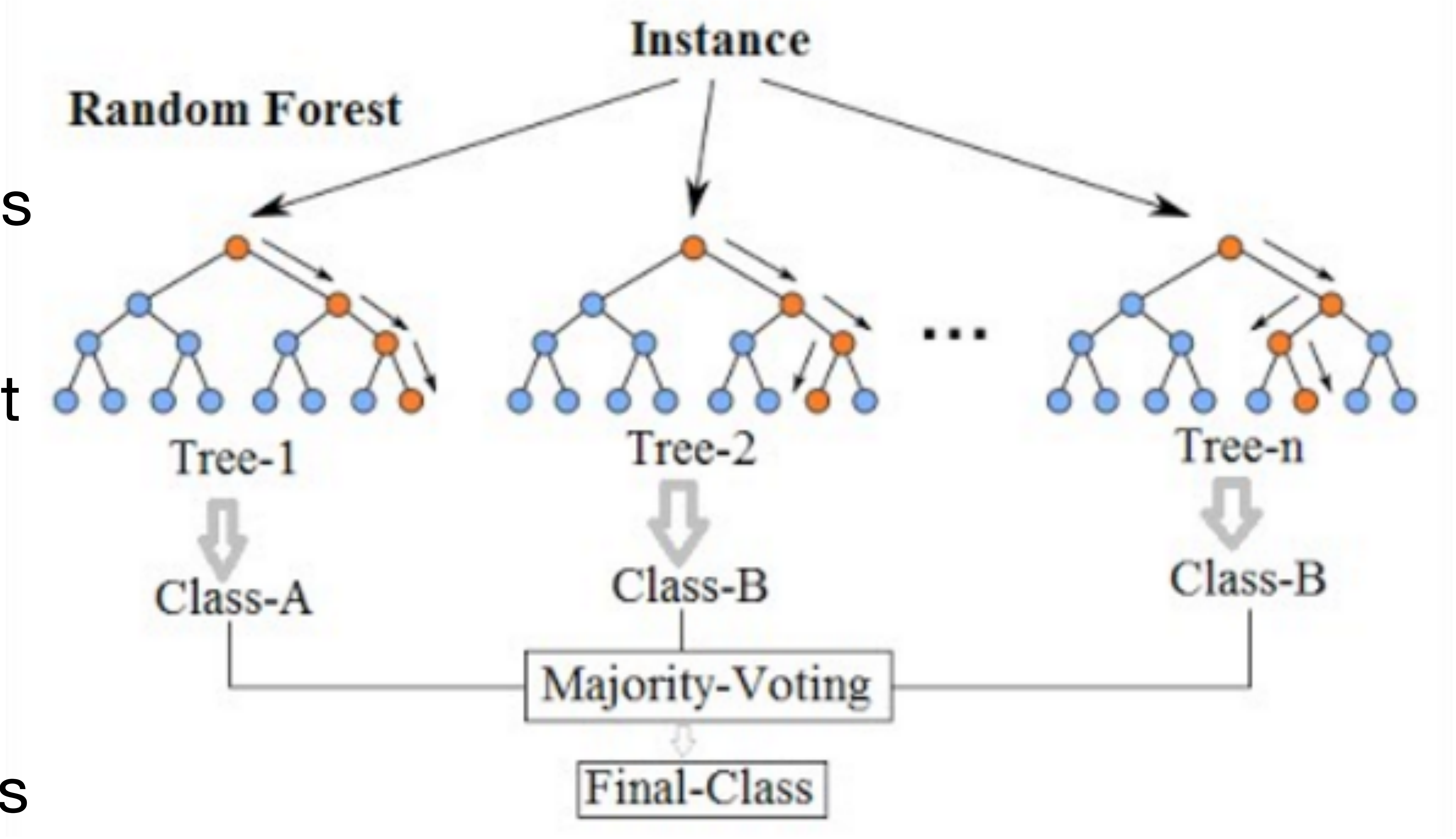
Chest Pain	Good Blood Circ.	Blocked Arteries	Weight	Heart Disease
No	No	No	125	No
Yes	Yes	Yes	180	Yes
Yes	Yes	No	210	No
Yes	No	Yes	167	Yes

* **B**ootstrapping the data using the **ag**gregation to make a decision is called **bagging**

https://williamkoehrsen.medium.com/random-forest-simple-explanation-377895a60d2d
https://youtu.be/J4Wdy0Wc_xQ

Decision Trees vs Random Forests

- Issues of decision trees: overfitting
- Random forests could avoid this by
  - training with a random subset of data (bootstrapping)
  - Randomly select a subset of attributes
  - Take an aggregation of results

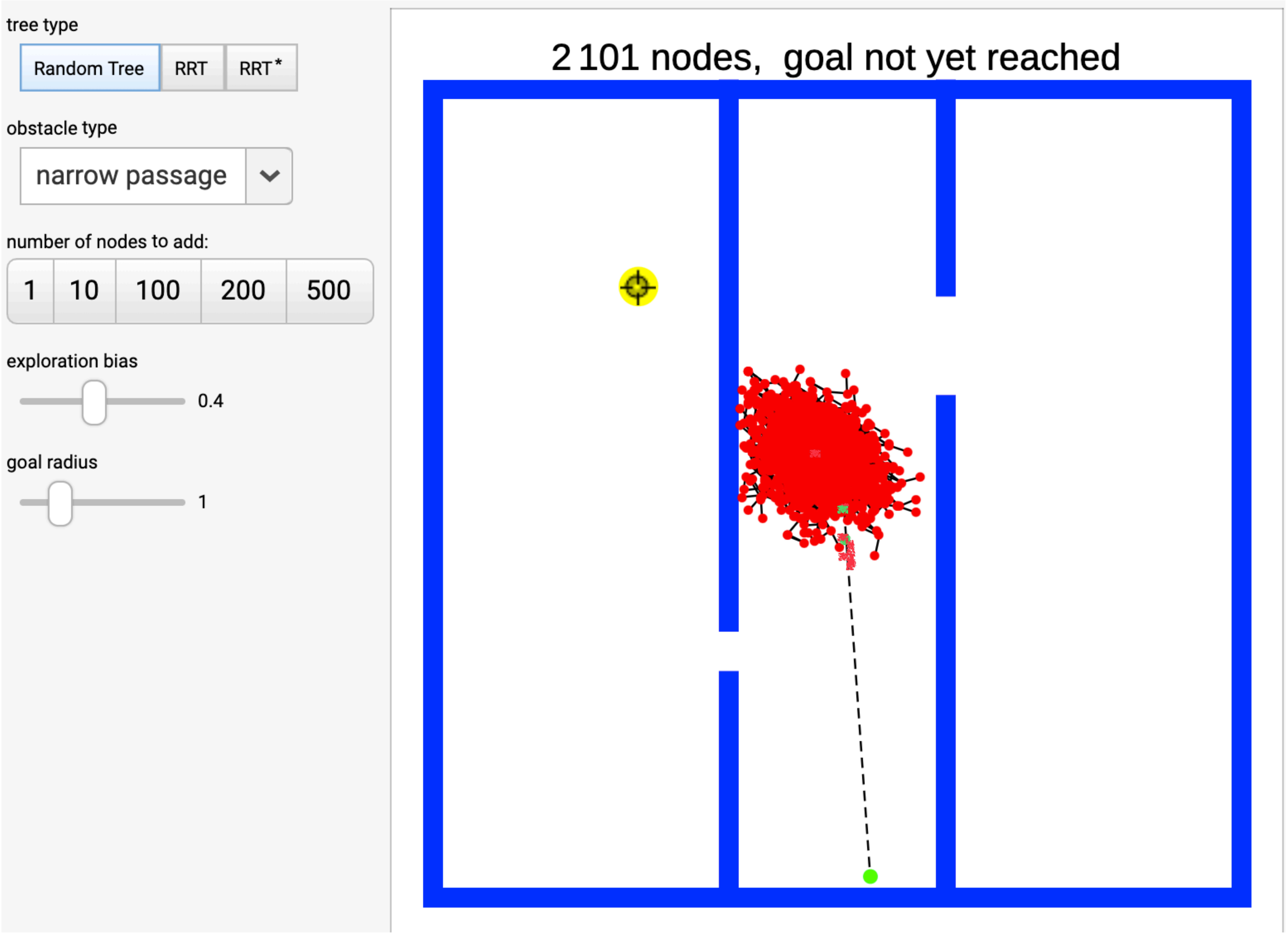


* **B**ootstrapping the data using the **ag**gregation to make a decision is called **bagging**

https://williamkoehrsen.medium.com/random-forest-simple-explanation-377895a60d2d

How to grow a tree to search: Random Tree

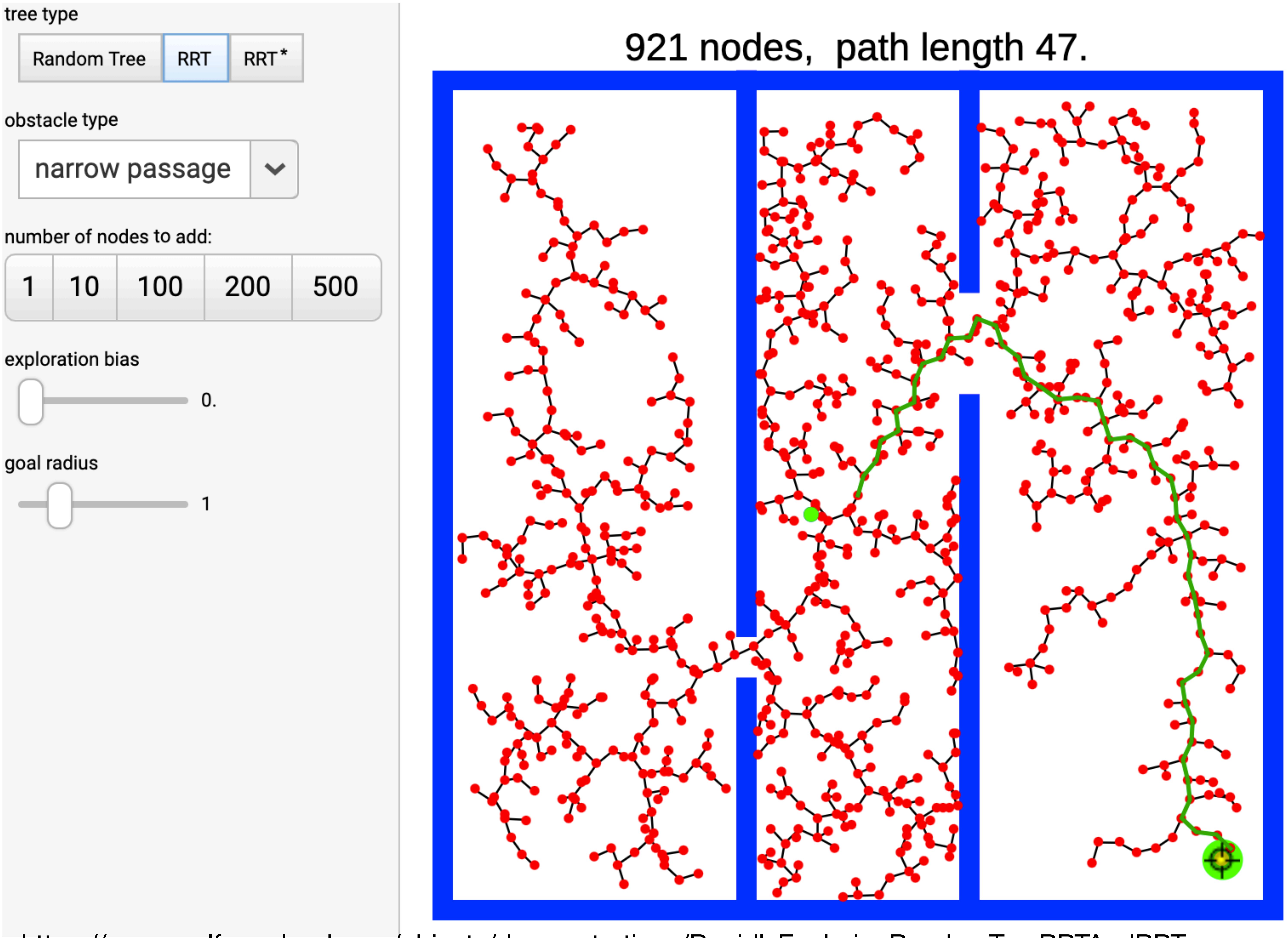
- A Random Tree selects a node at random from the tree and adds an edge in a random direction.



https://www.wolframcloud.com/objects/demonstrations/RapidlyExploringRandomTreeRRTAndRRT-source.nb

Rapid Random Tree (RRT)

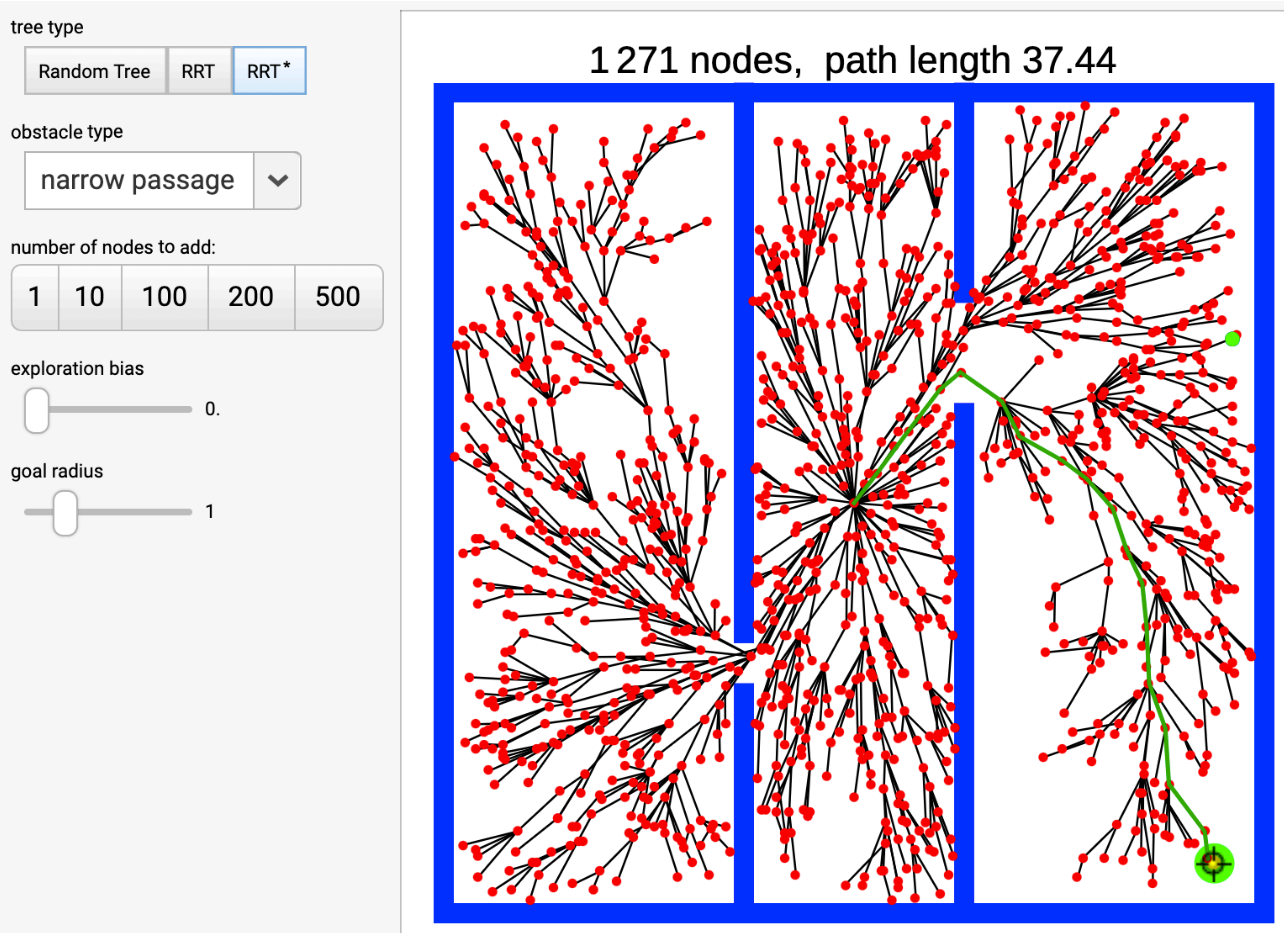
- A RRT first selects a random goal. point, then tries to add an edge from the closest node in the tree toward the goal point.



https://www.wolframcloud.com/objects/demonstrations/RapidlyExploringRandomTreeRRTAndRRT-source.nb

Rapid Random Tree Star (RRT*)

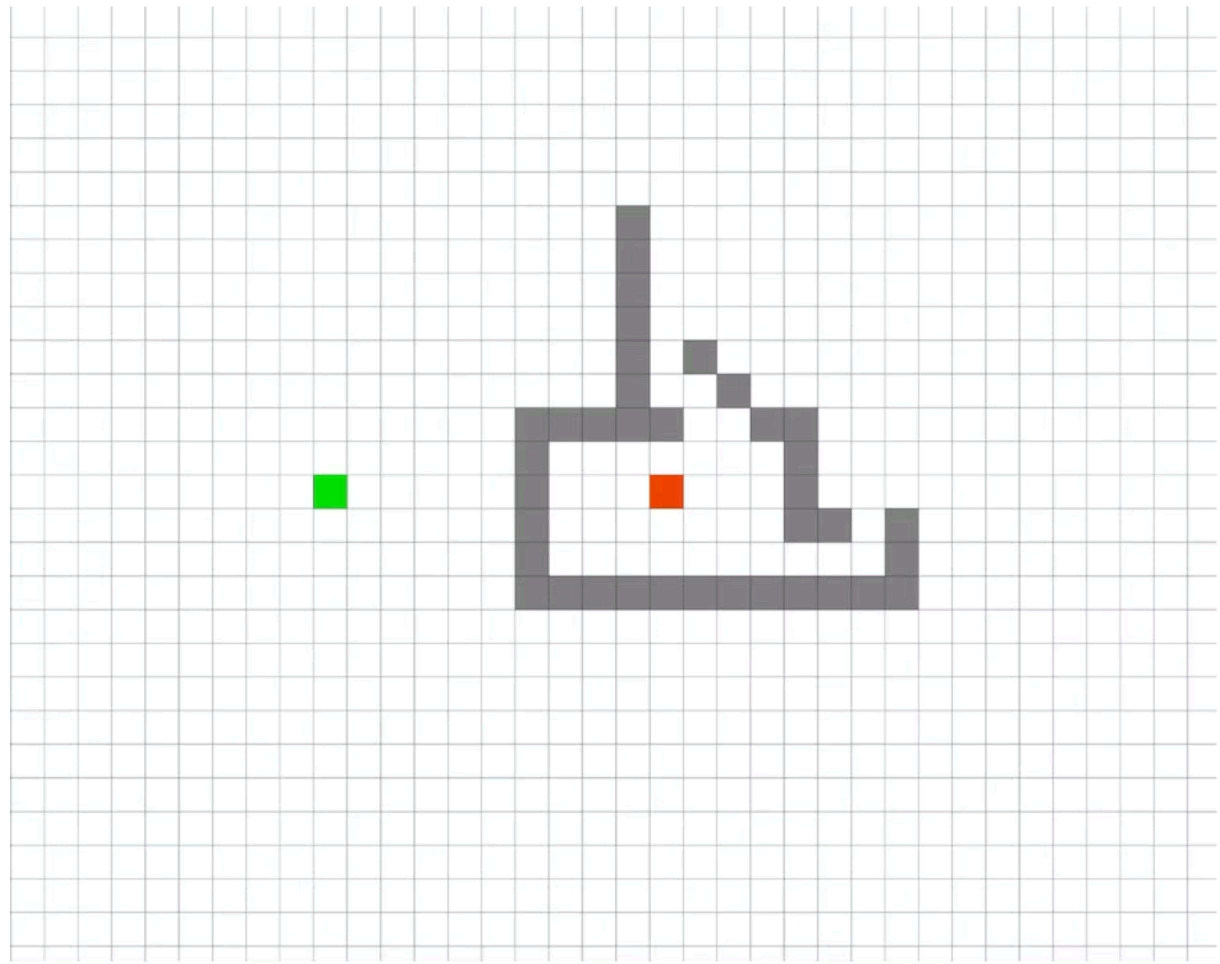
- RRT* improves this by rewiring the tree to form shortest paths.

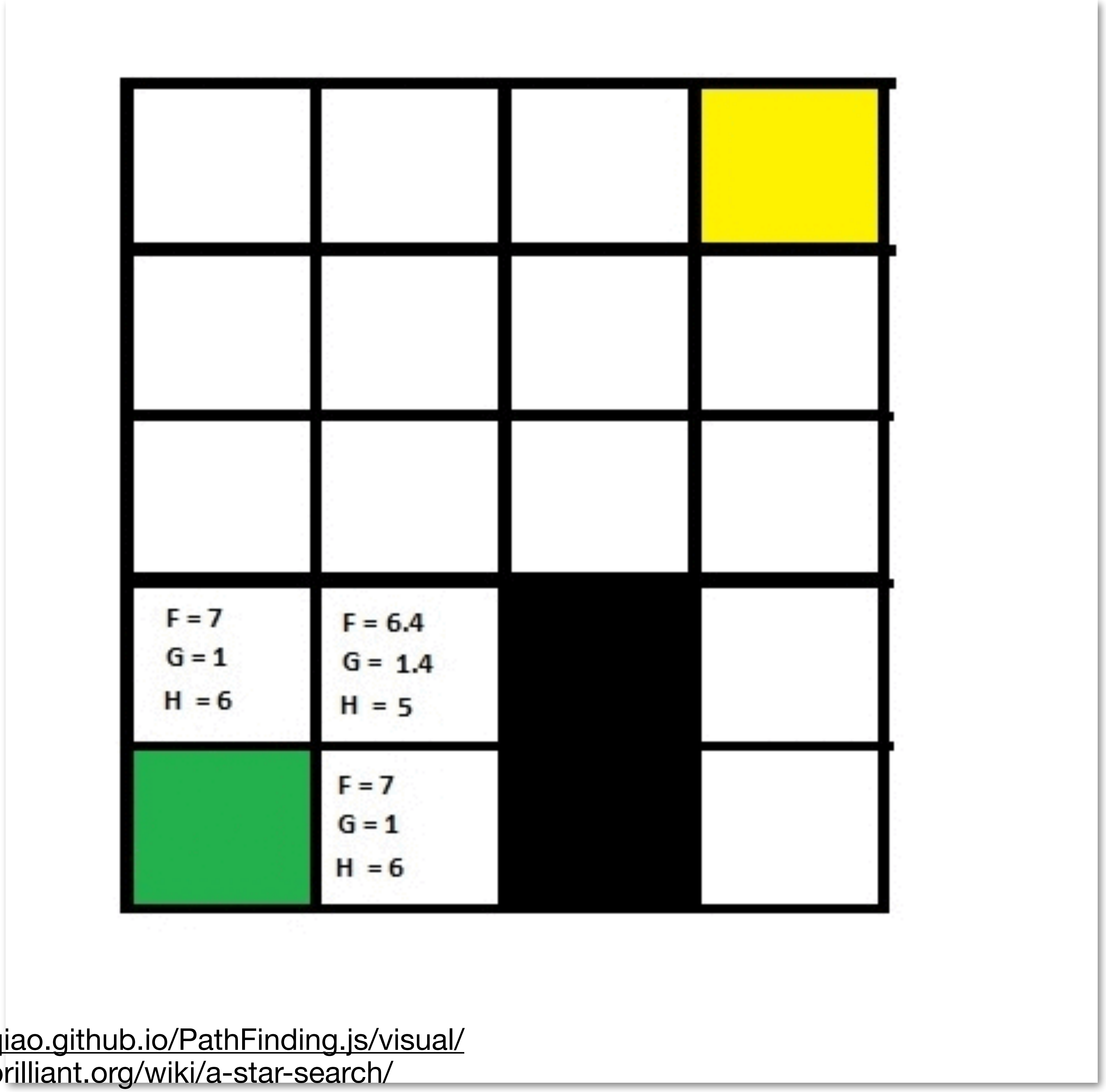


https://www.wolframcloud.com/objects/demonstrations/RapidlyExploringRandomTreeRRTAndRRT-source.nb

A*: Random tree search with heuristic

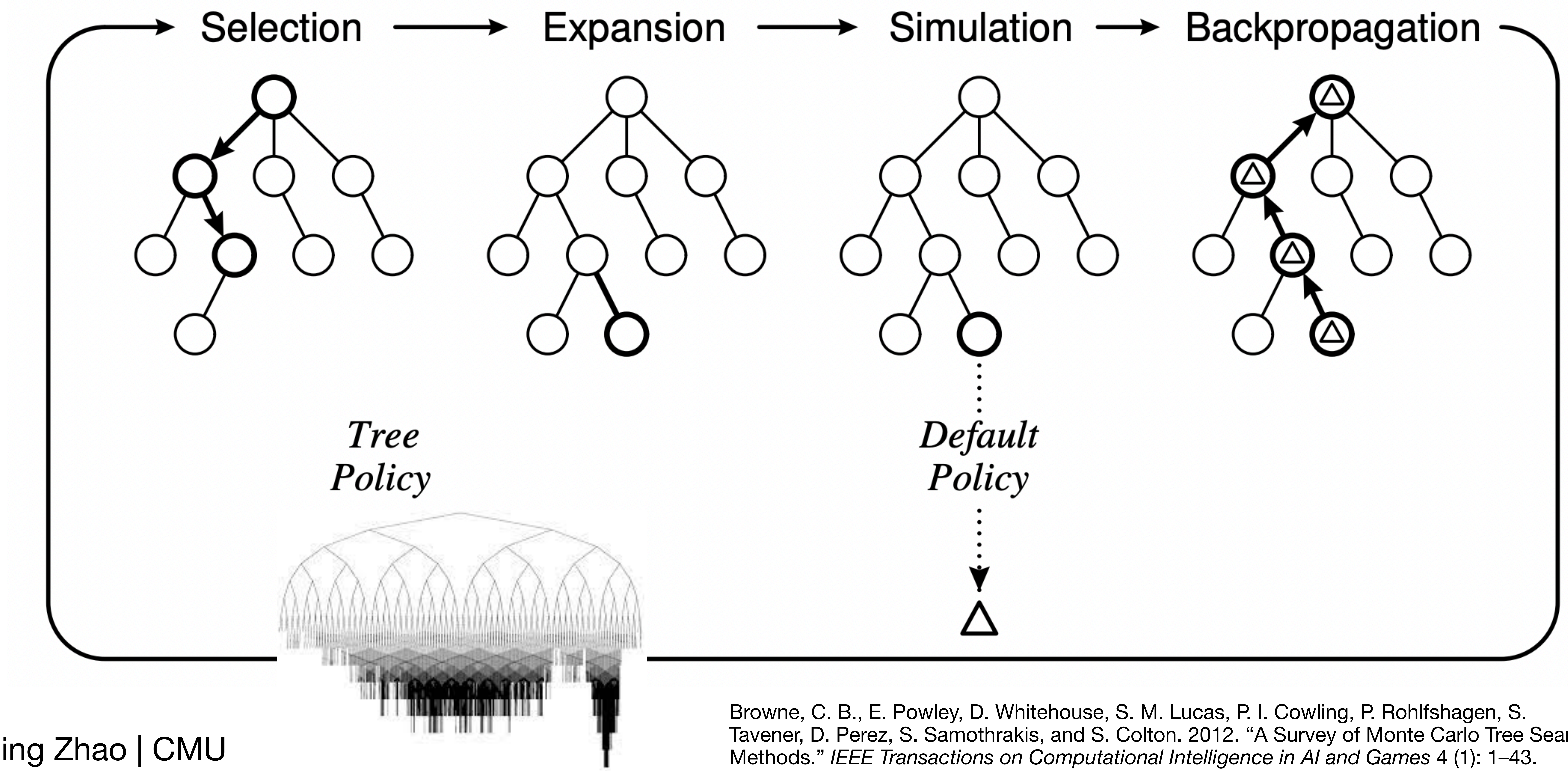
- $f(n) = g(n) + h(n)$
  - Where $f(n) =$ total estimated cost of path through node n
  - $g(n) =$ cost so far to reach node n
  - $h(n) =$ estimated cost from n to goal. This is the heuristic part of the cost function.





https://qiao.github.io/PathFinding.js/visual/
https://brilliant.org/wiki/a-star-search/#:~:text=A*%20(pronounced%20as%20%22A,or%20points%2C%20on%20the%20graph.&text=With%20A*%2C%20a%20robot%20would,diagram%20on%20the%20right%20below.

Monte Carlo Tree Search



Browne, C. B., E. Powley, D. Whitehouse, S. M. Lucas, P. I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis, and S. Colton. 2012. “A Survey of Monte Carlo Tree Search Methods.” *IEEE Transactions on Computational Intelligence in AI and Games* 4 (1): 1–43.

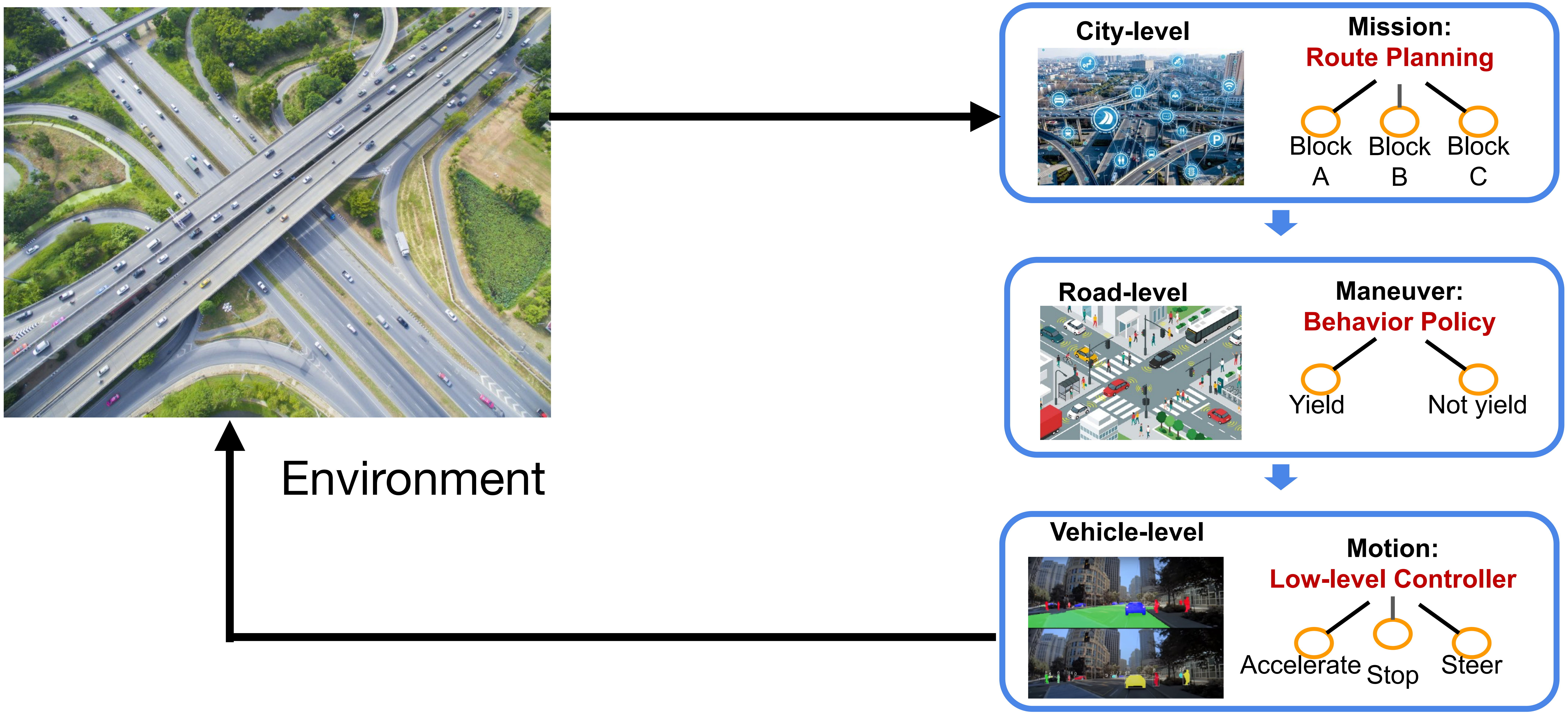
Monte Carlo Tree Search case study: Alpha Go

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).

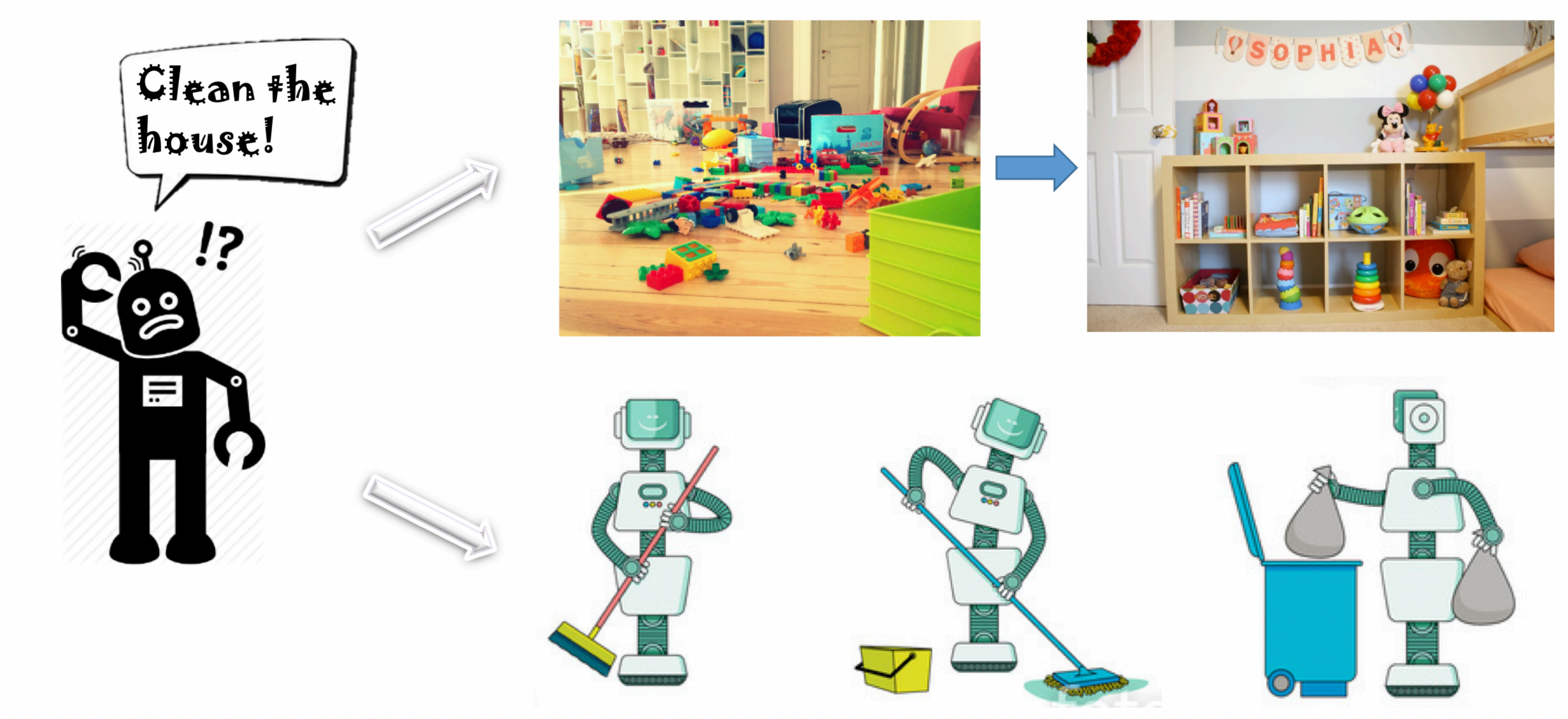
Contents

- Hierarchical AI structures
- Trees
  - Decision trees
  - Random tree/forests
  - Monte Carlo Tree search, Alpha Go
- Hierarchical RL
  - Manager-worker
  - Option/Semi-MDP
- Hierarchical structures in Meta learning
  - Neural Processes

Hierarchical Decision-Making (Autonomous Vehicles)



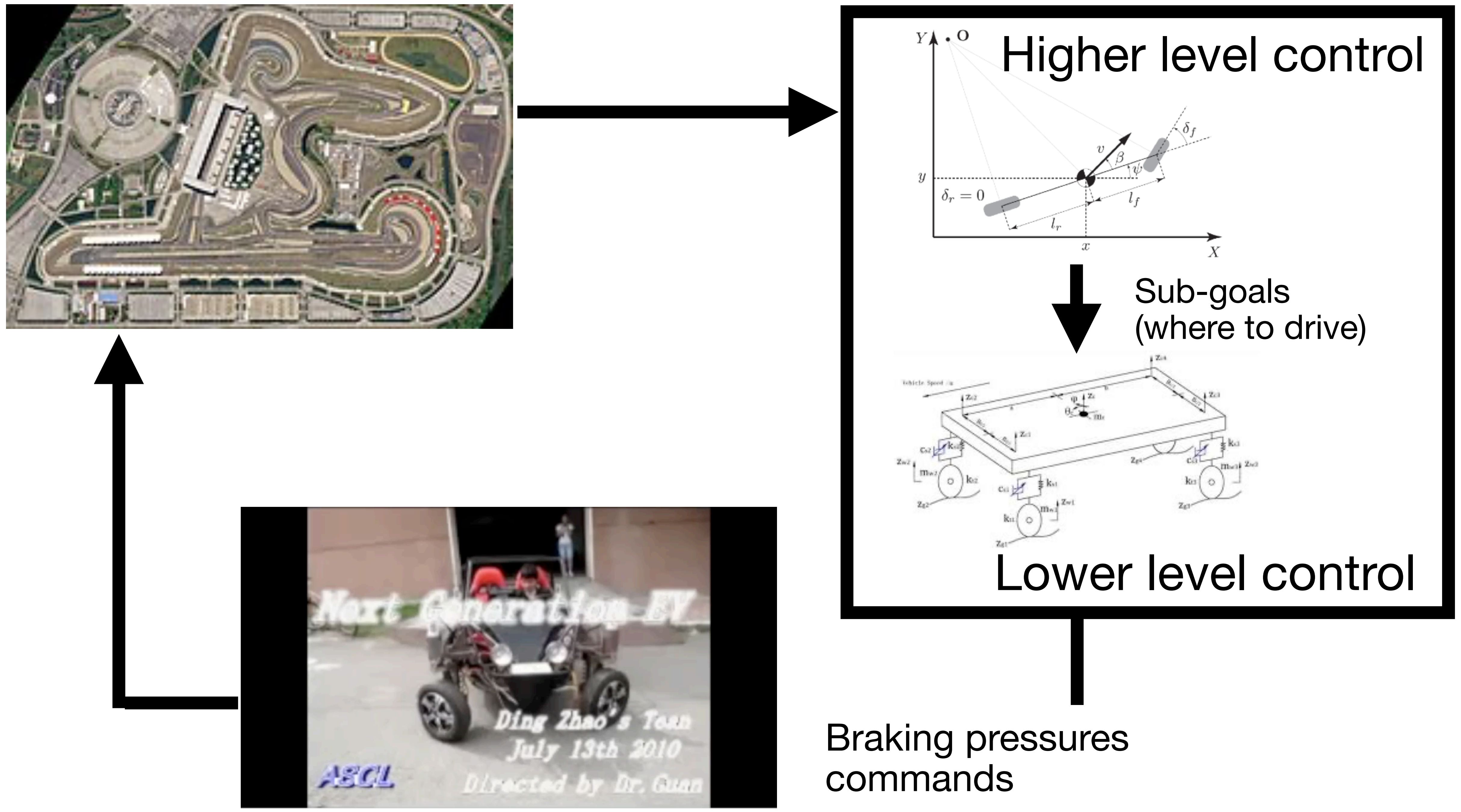
Hierarchical Decision-Making (home robots)



https://phillipi.github.io/6.882/2020/notes/The%20problem%20of%20learning%20for%20long%20horizons/FeudalRL.pdf

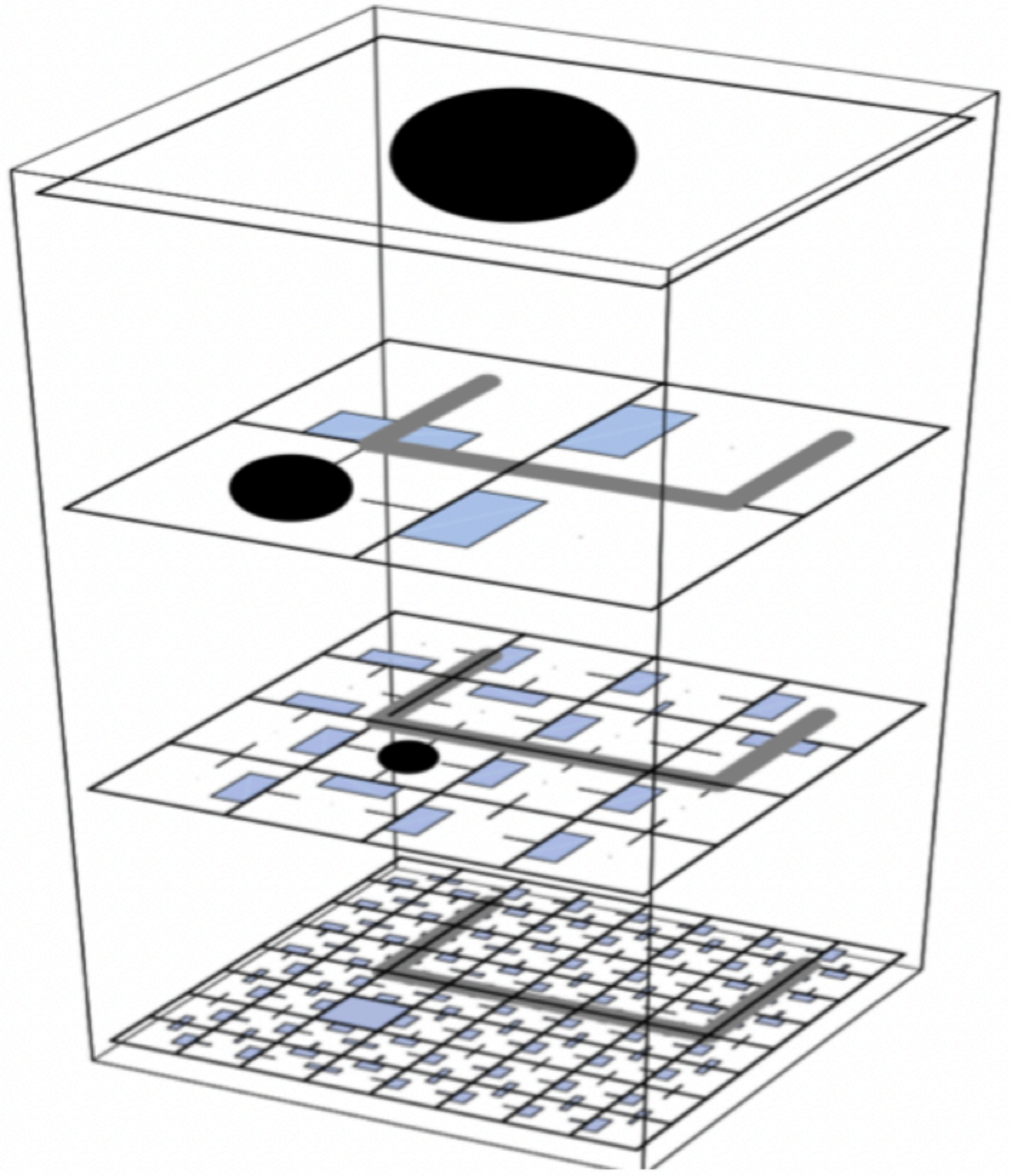
Hierarchical Reinforcement Learning

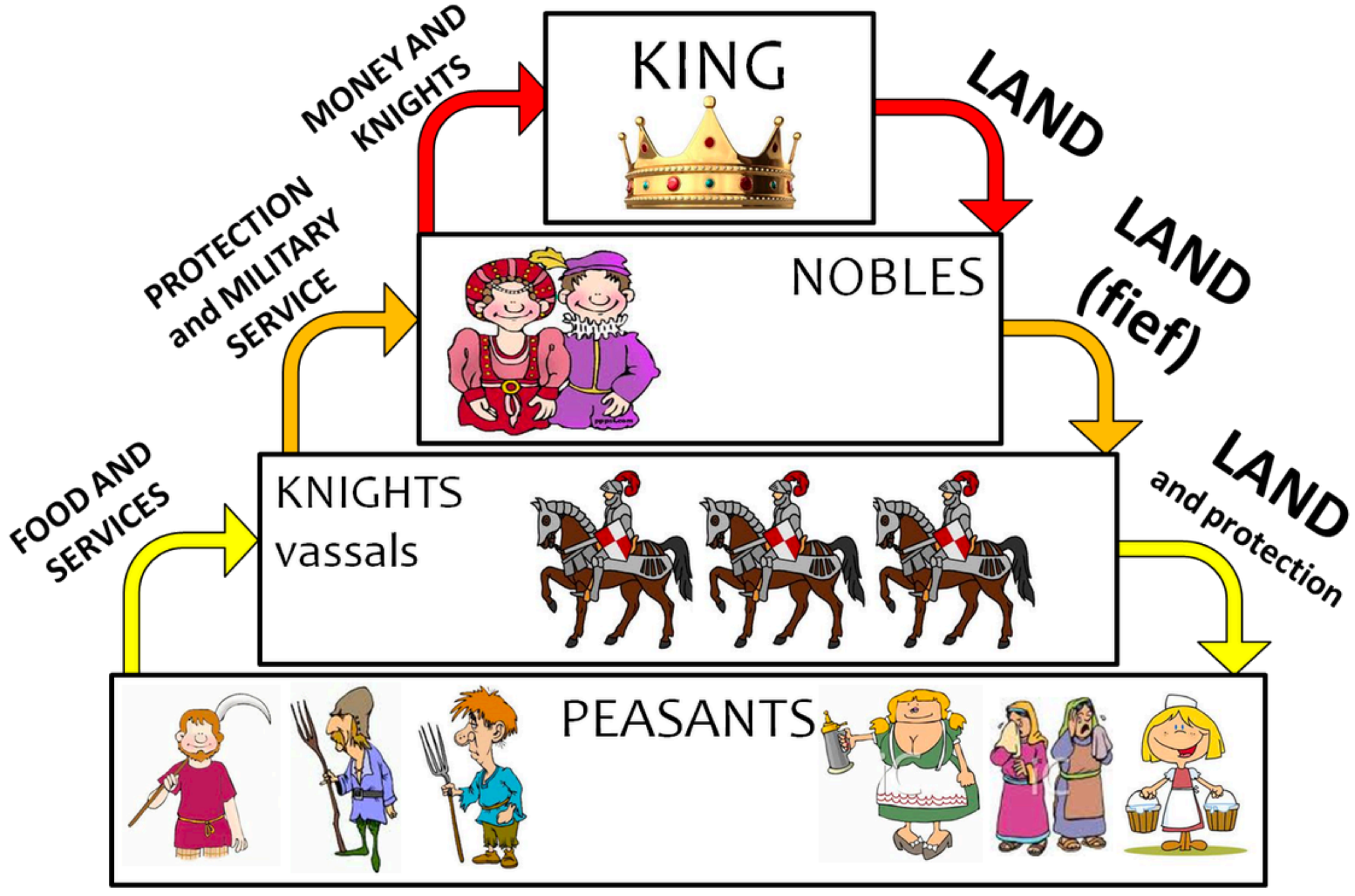
- Benefits
  - Efficiency/Scalability
  - Transfer/reusability of skills
  - Explainability/maintenance
- Different hierarchical frameworks
  - Manager-submanager: manager sets subgoals and rewards for sub-managers
    - Feudal RL; FeUdal Networks (FUNs)
  - Option: no explicit subgoals learn and discover options
    - Option-Critic; Meta Learning Shared Hierarchies (MLSH)



Feudal Reinforcement Learning

- Good concept
- Was not widely used

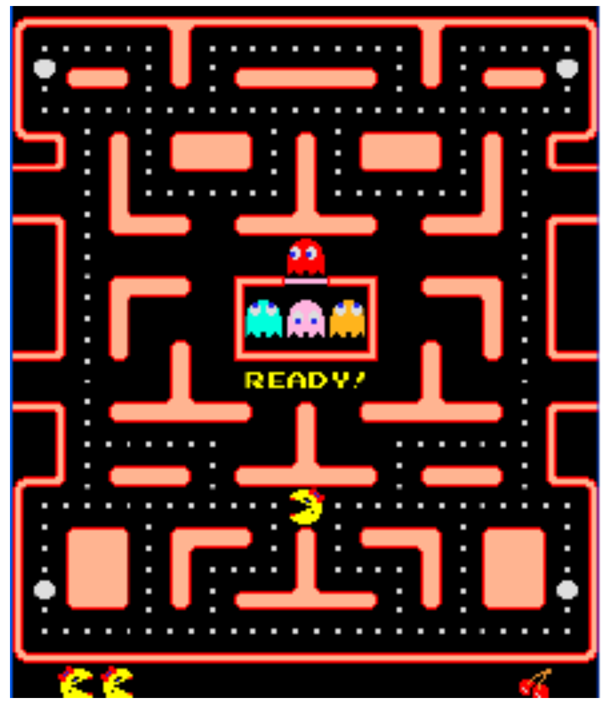


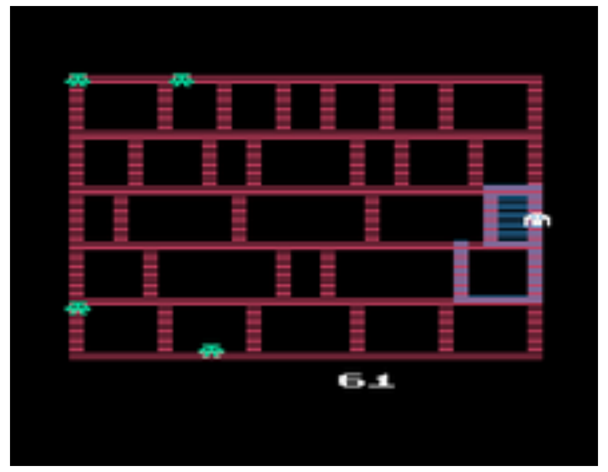


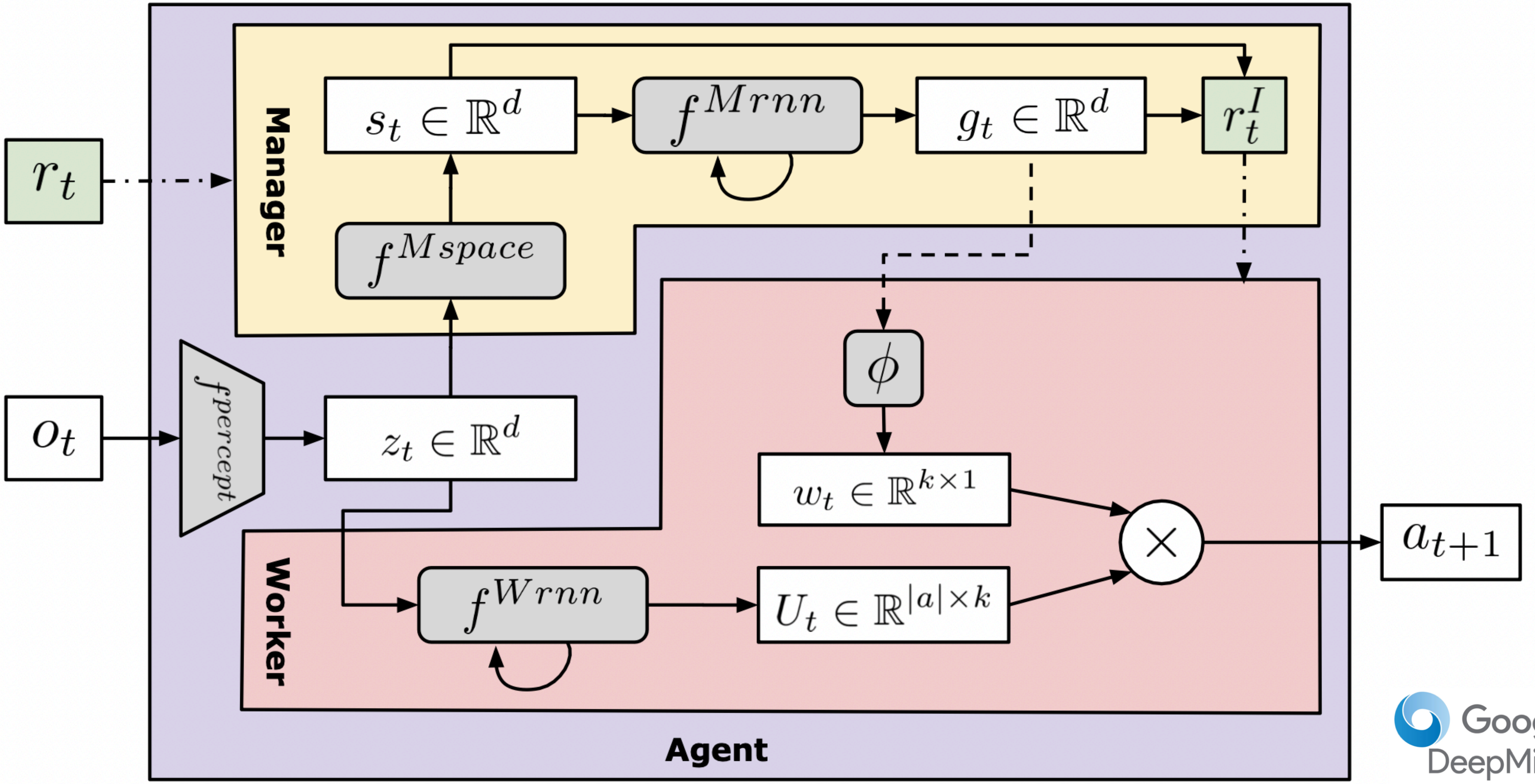
Feudal Pyramid of Power

Dayan, Peter, and Geoffrey E. Hinton. 1993. “Feudal Reinforcement Learning.” In *Advances in Neural Information Processing Systems*
https://www.pinterest.com/pin/109493834677707526/

FeUdal Networks (FUNs)



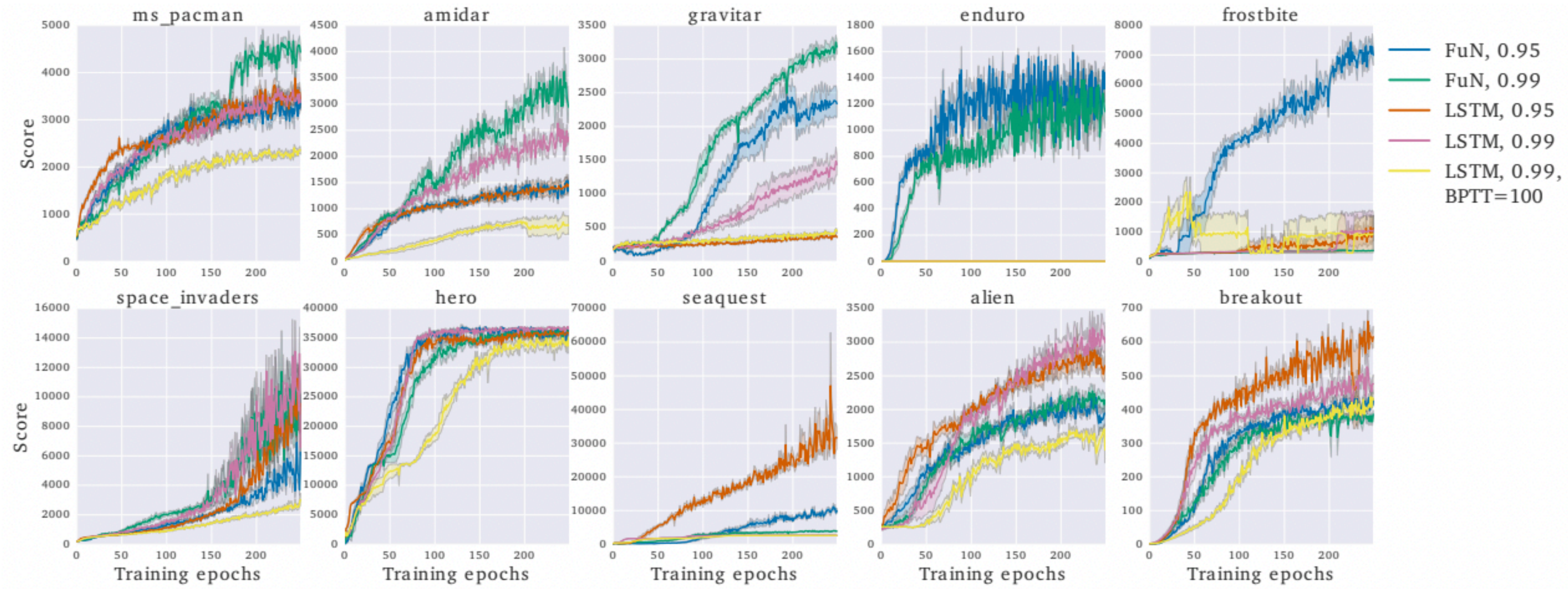




Google DeepMind

Vezhnevets, A. S., S. Osindero, and T. Schaul. "Feudal Networks for Hierarchical Reinforcement Learning." *ICML, 2017*

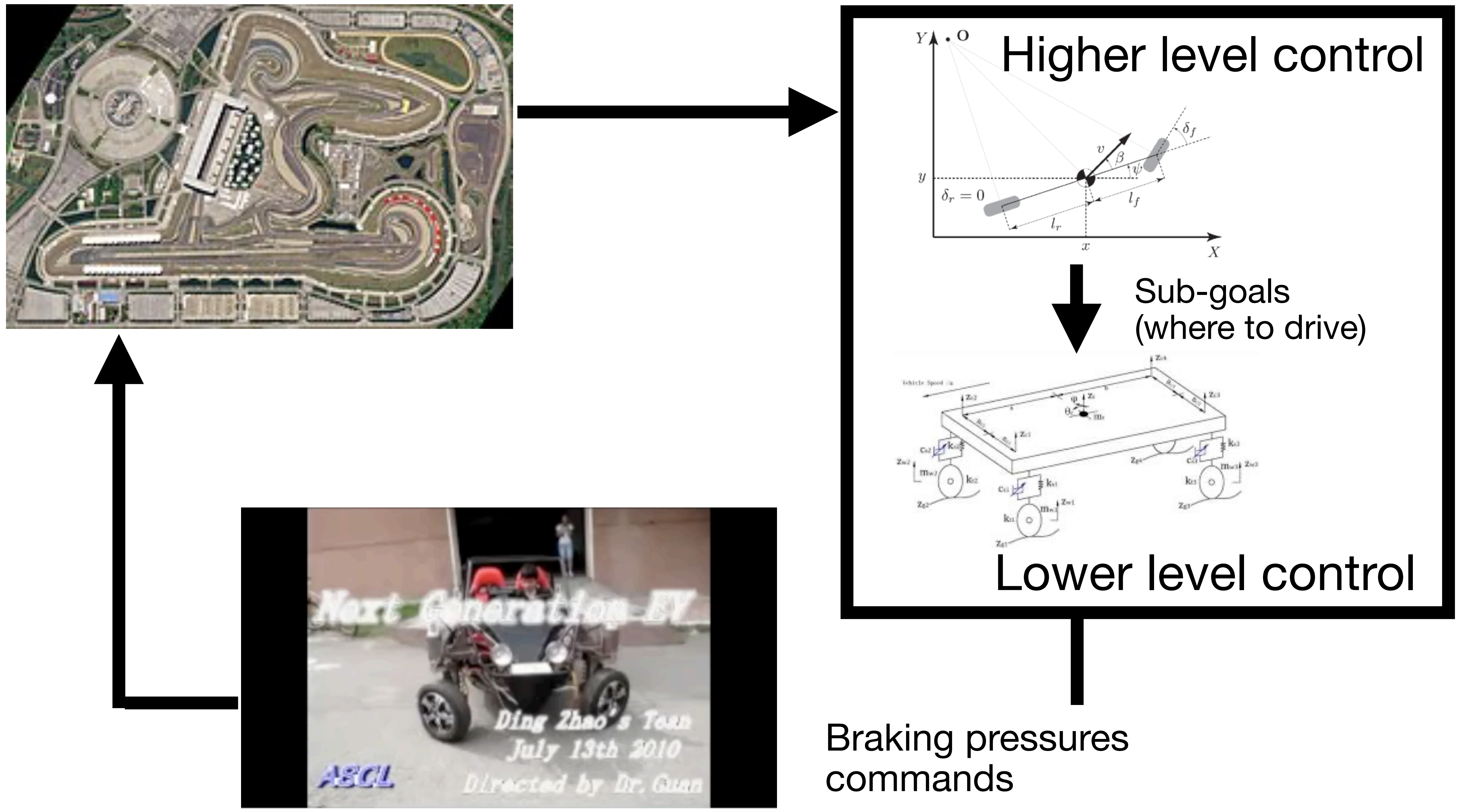
FeUdal Networks (FUNs) Empirical Results



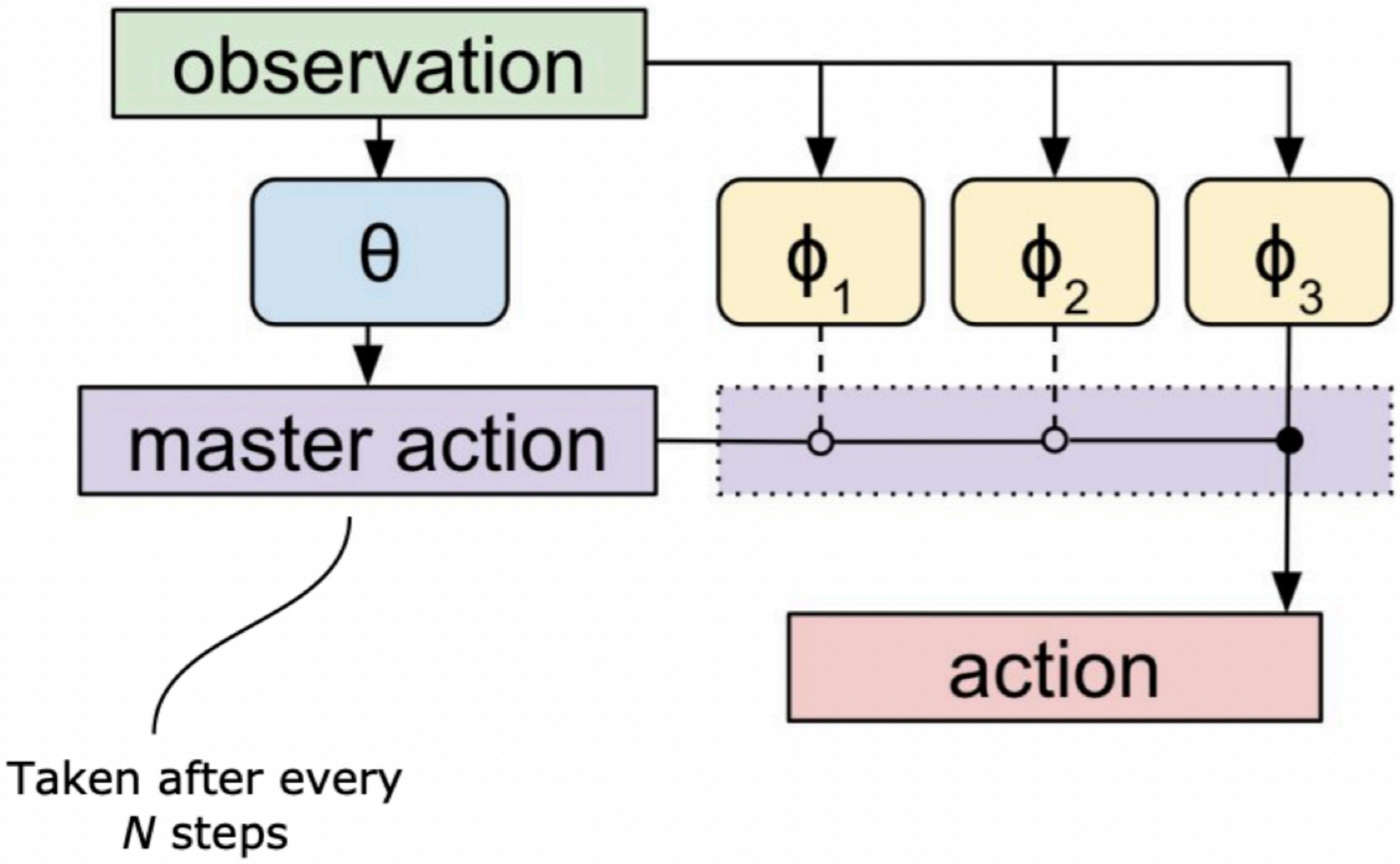
Vezhnevets, A. S., S. Osindero, and T. Schaul. 2017. "Feudal Networks for Hierarchical Reinforcement Learning." *Machine Learning*. http://proceedings.mlr.press/v70/vezhnevets17a.html.

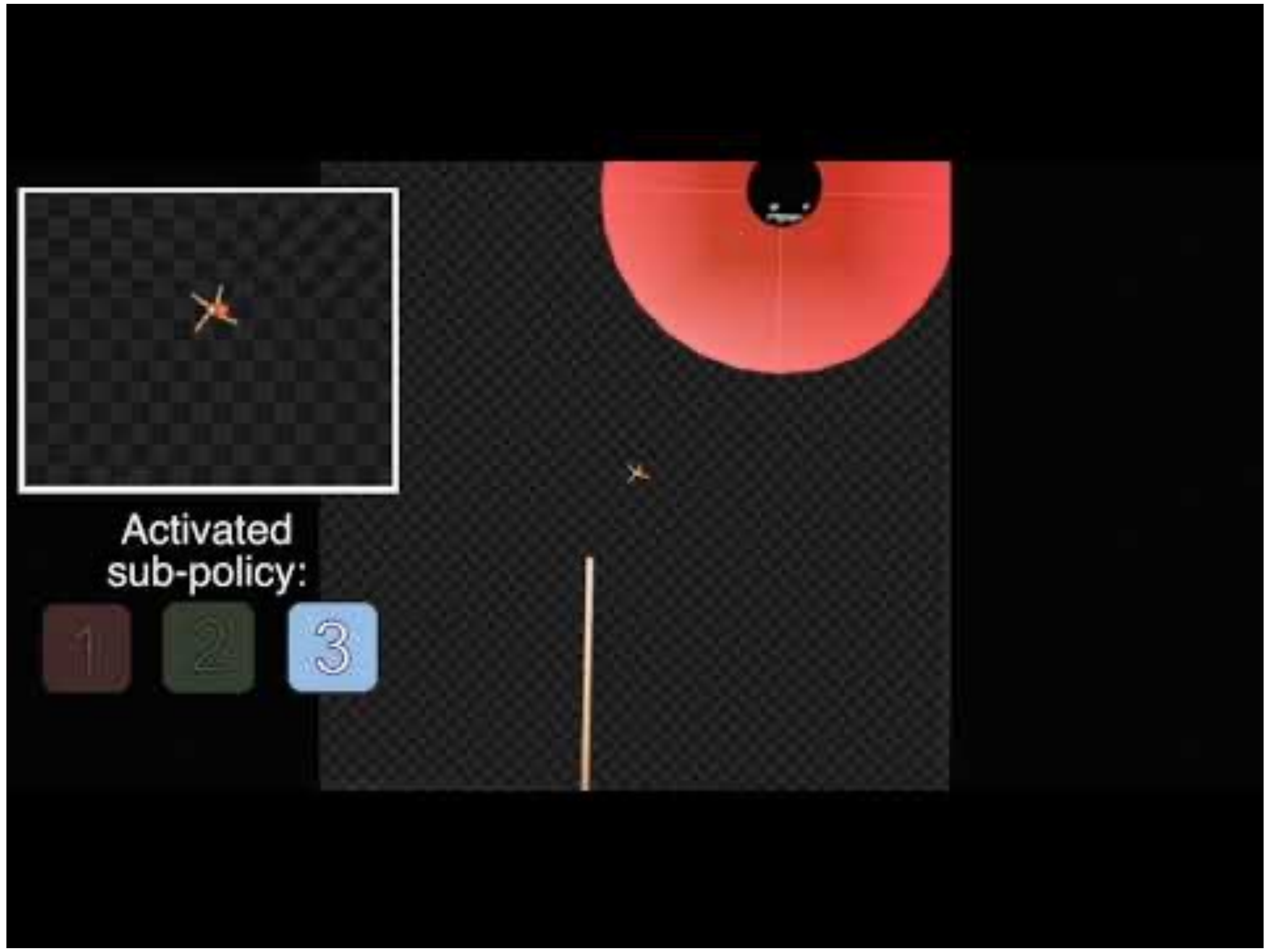
Hierarchical Reinforcement Learning

- Benefits
  - Efficiency/Scalability
  - Transfer/reusability of skills
  - Explainability/maintenance
- Different hierarchical frameworks
  - Manager-submanager: manager sets subgoals and rewards for sub-managers
    - Feudal RL; FeUdal Networks (FUNs)
  - **Option: no explicit subgoals learn and discover options**
    - Option-Critic; Meta Learning Shared Hierarchies (MLSH)



Meta Learning Shared Hierarchies

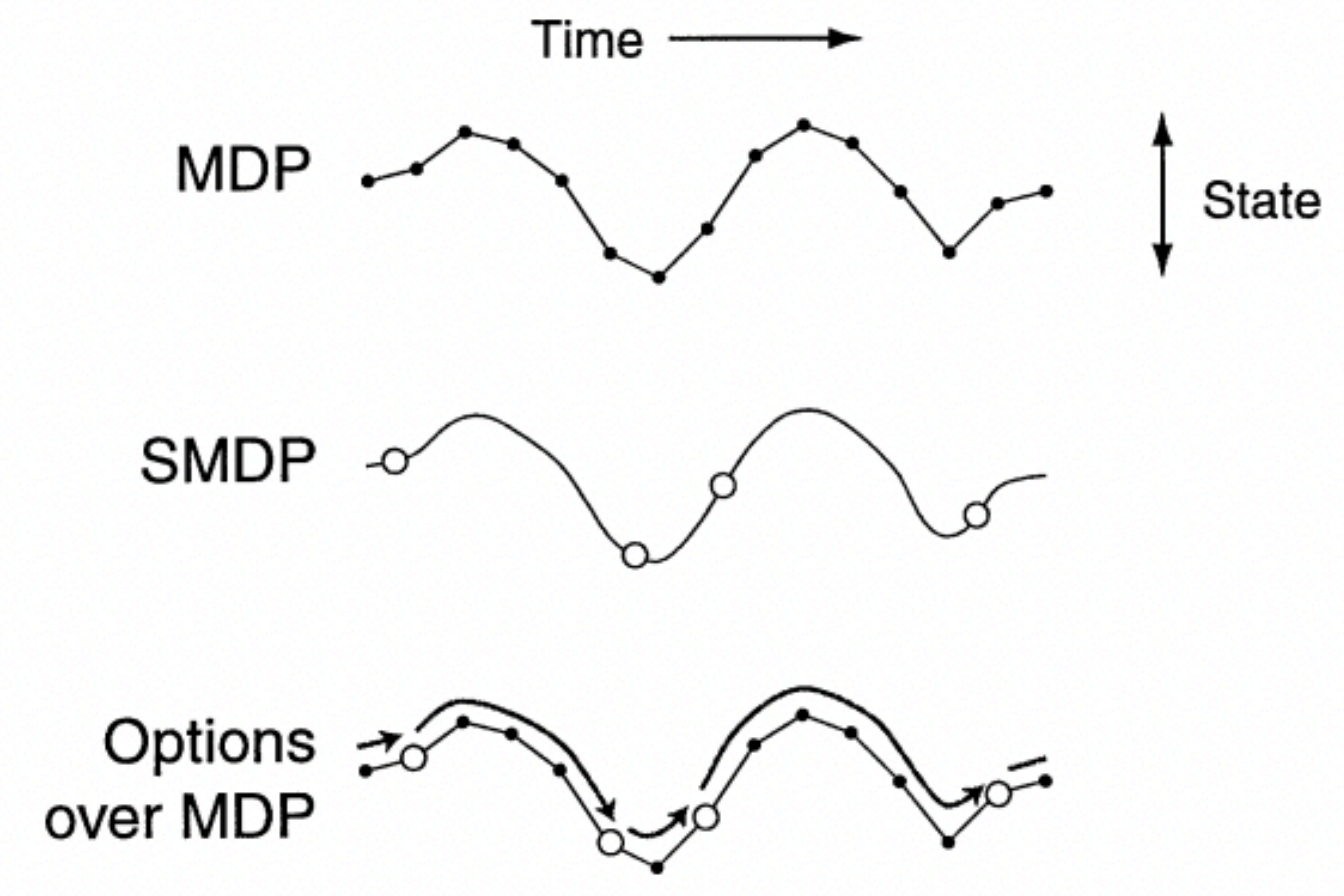




Frans, Kevin, Jonathan Ho, Xi Chen, Pieter Abbeel, and John Schulman. 2017. “Meta Learning Shared Hierarchies.” ICLR 2018

Options: Temporal abstraction in RL

- MDP + Options = Semi-MDP
- Semi-Markovian
  - Transition probability:
  - $p\left(s', \tau \mid s\right) = p\left(s' \mid s\right) p(\tau \mid s)$
  - where τ indicates the time to transition

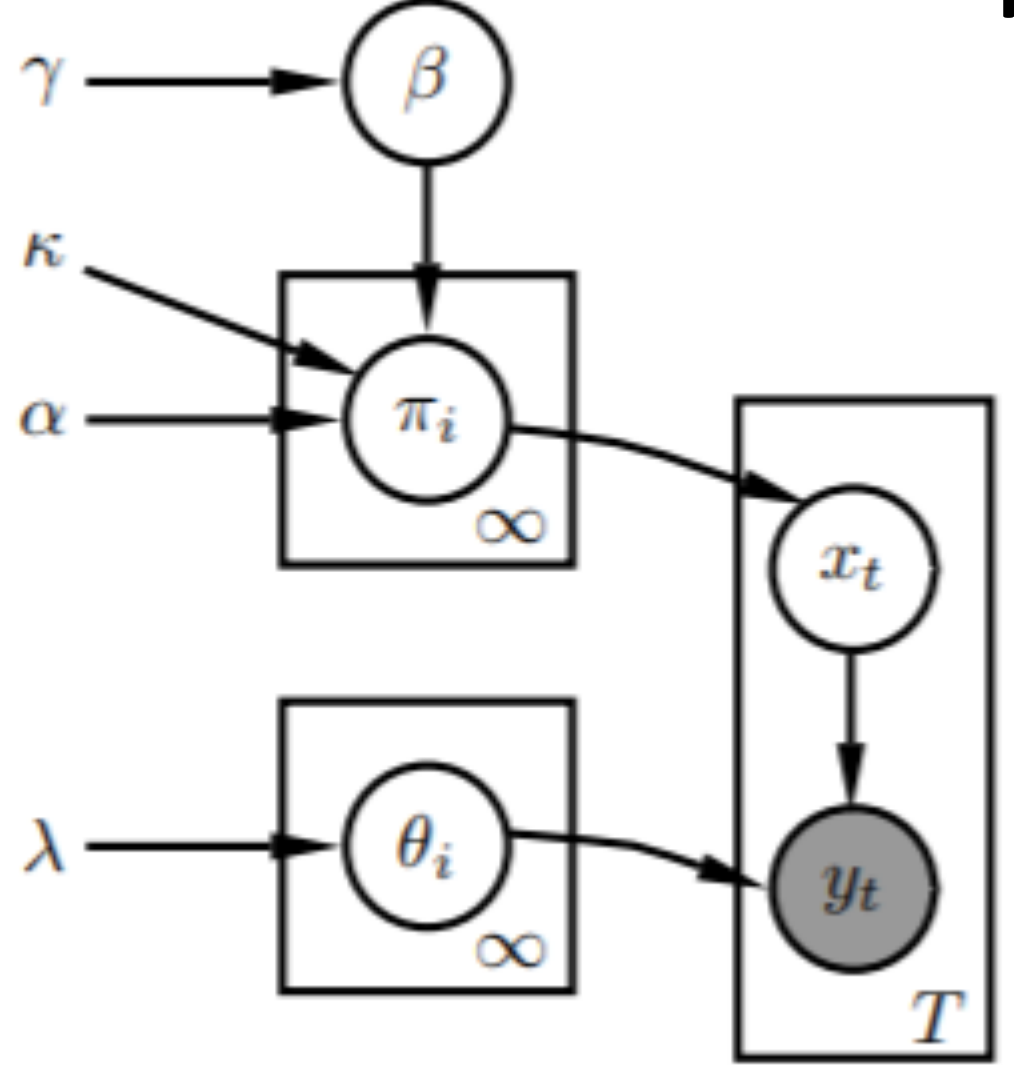


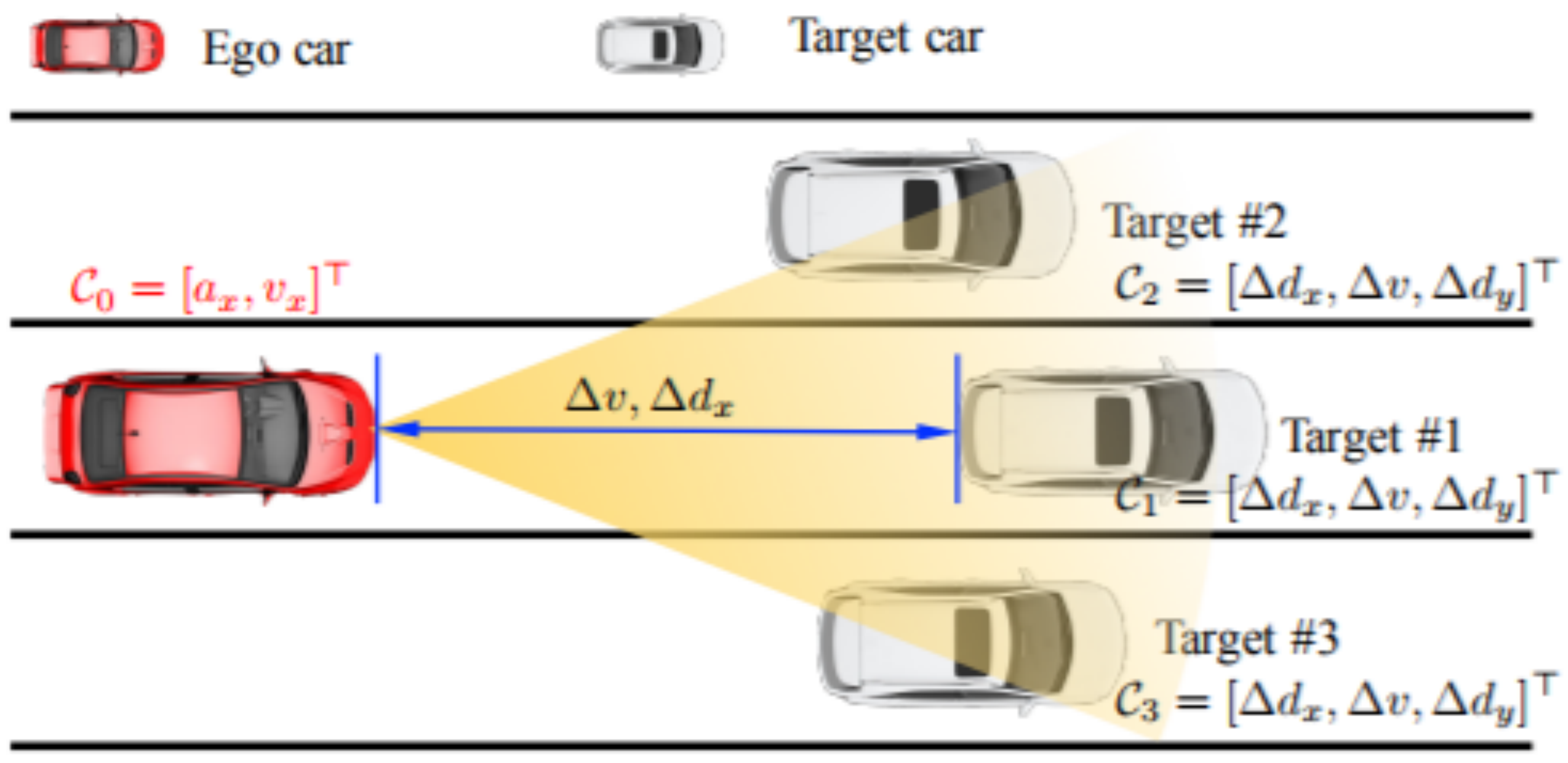
Sutton, Richard S., Doina Precup, and Satinder Singh. 1999. "Between MDPs and Semi-MDPs: A Framework for Temporal Abstraction in Reinforcement Learning." *Artificial Intelligence* 112 (1): 181–211.
CS885, University of Waterloo, Pascal Poupart

Example 1: Traffic Primitives

Traffic primitive is referred to the representation of fundamental building blocks of the traffic environment in spatiotemporal space.

Toyota (**PI**) "Extracting **Traffic Primitives** from Millions of Naturalistic Driving Encounters -- A Synthesized Method based on Nonparametric Bayesian and Deep Unsupervised Learning"





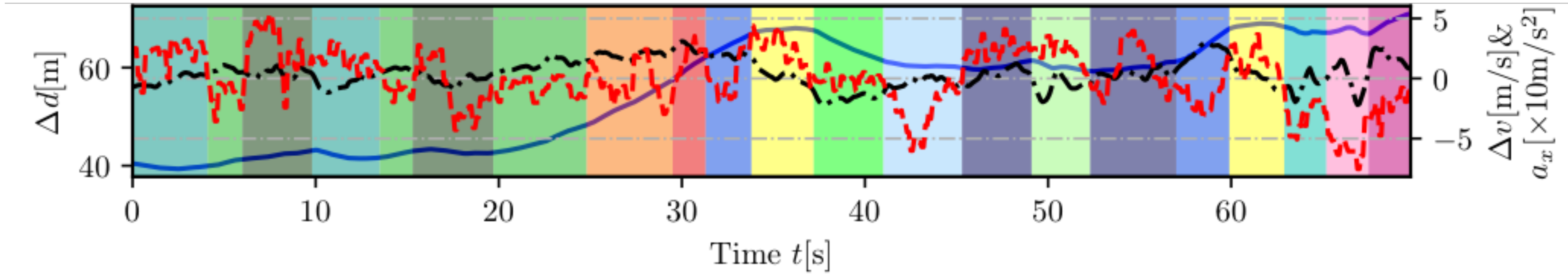
Previous methods:

- **Subjectively-selected scenarios**

Traffic Primitive:

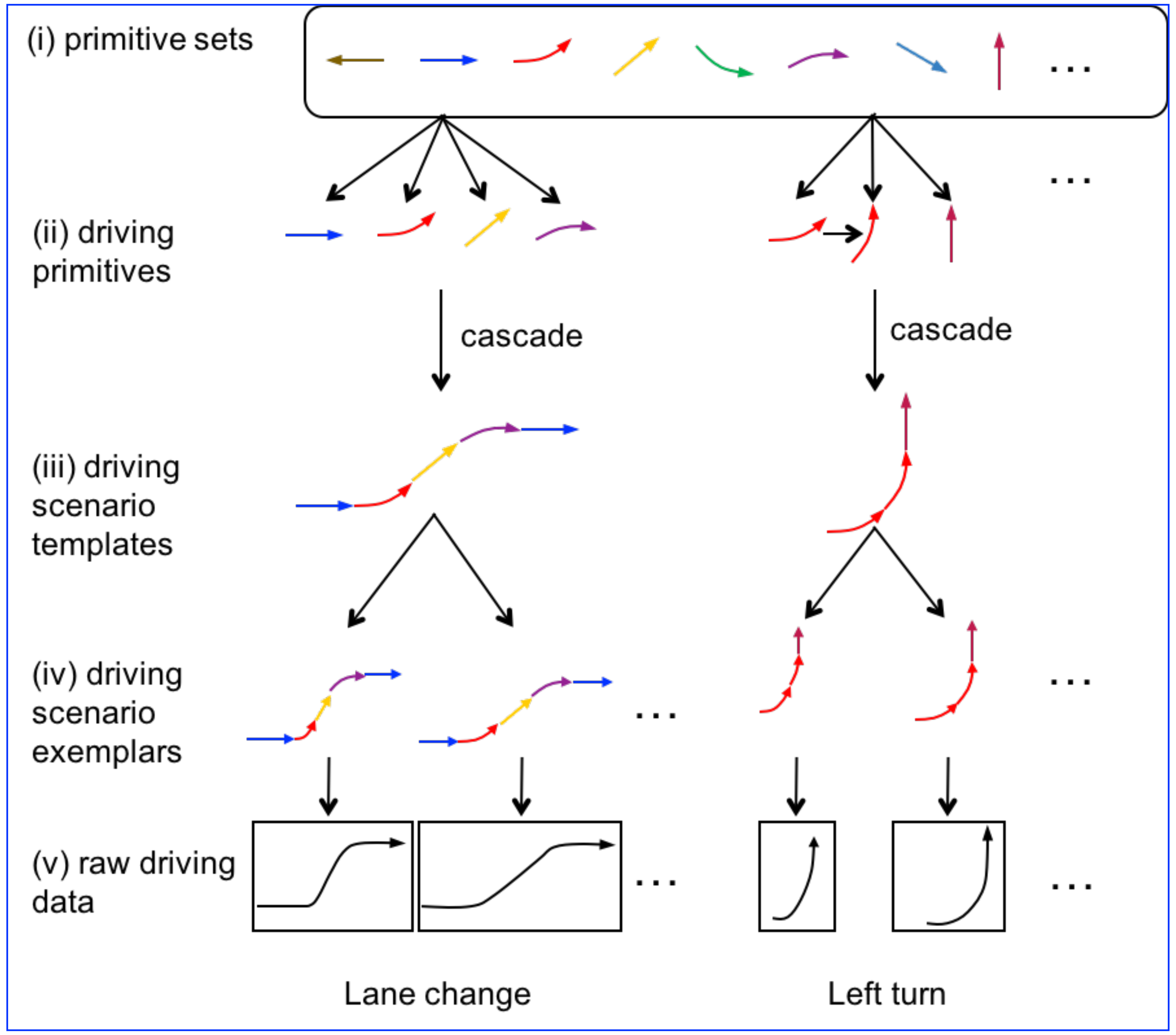
- Segment/cluster similar traffic scenes automatically using unsupervised learning

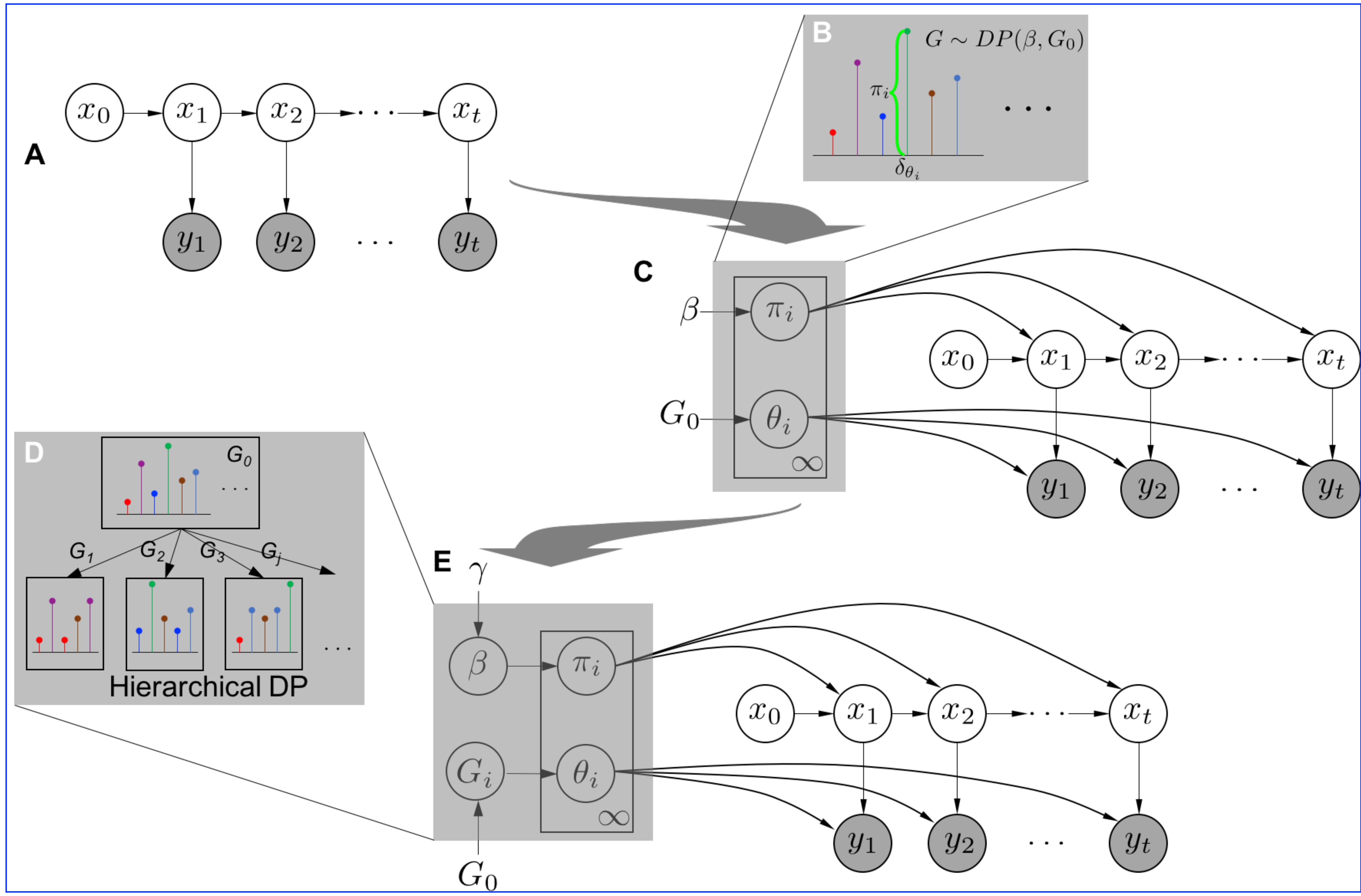
- Objectively-selected scenarios



[Wang, Zhao, 'Extracting Traffic Primitives Directly from Naturalistically Logged Data for Self-Driving Applications, ICRA, 2018]

Primitive extraction & analysis





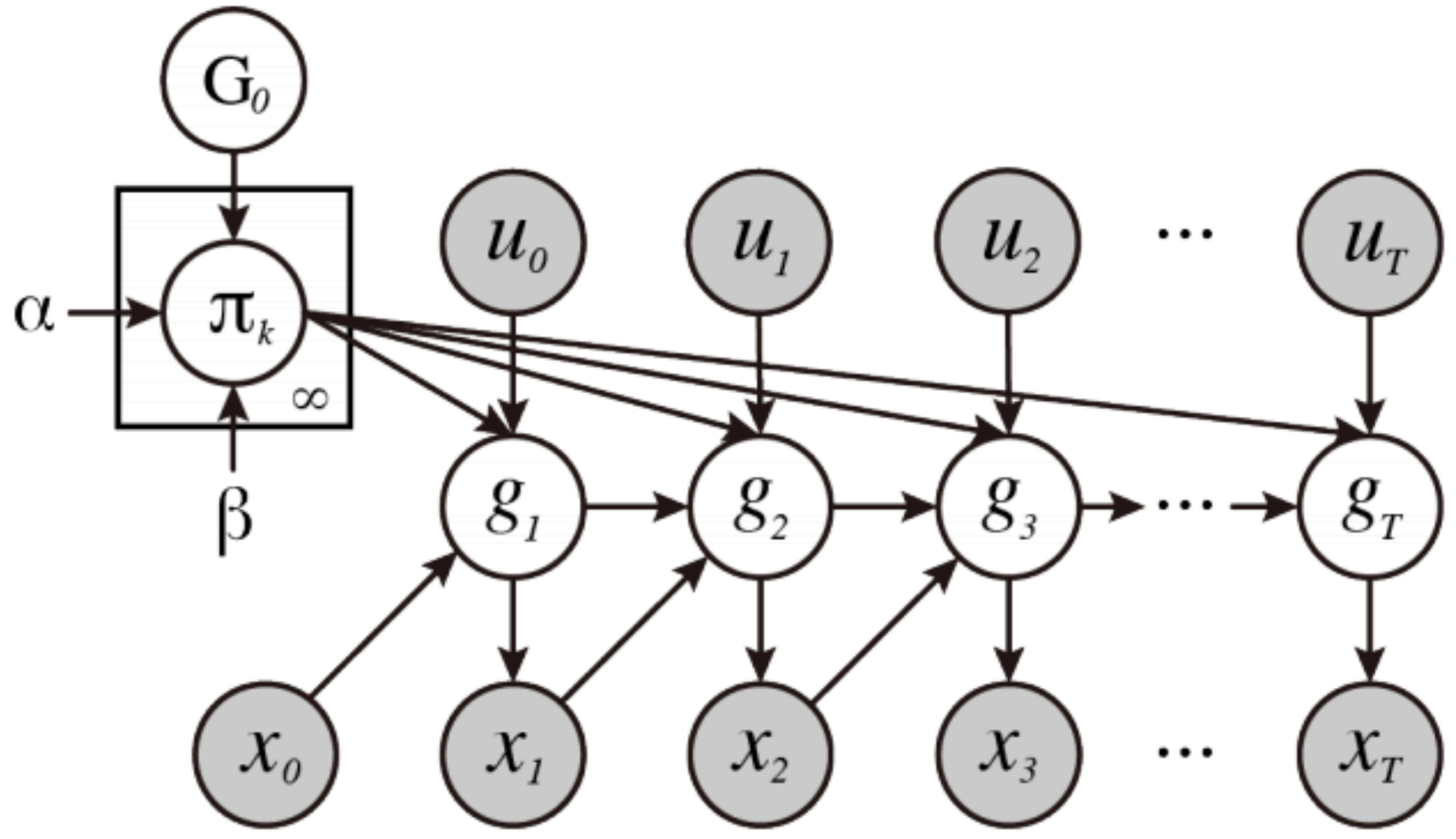
Extracting driving primitives

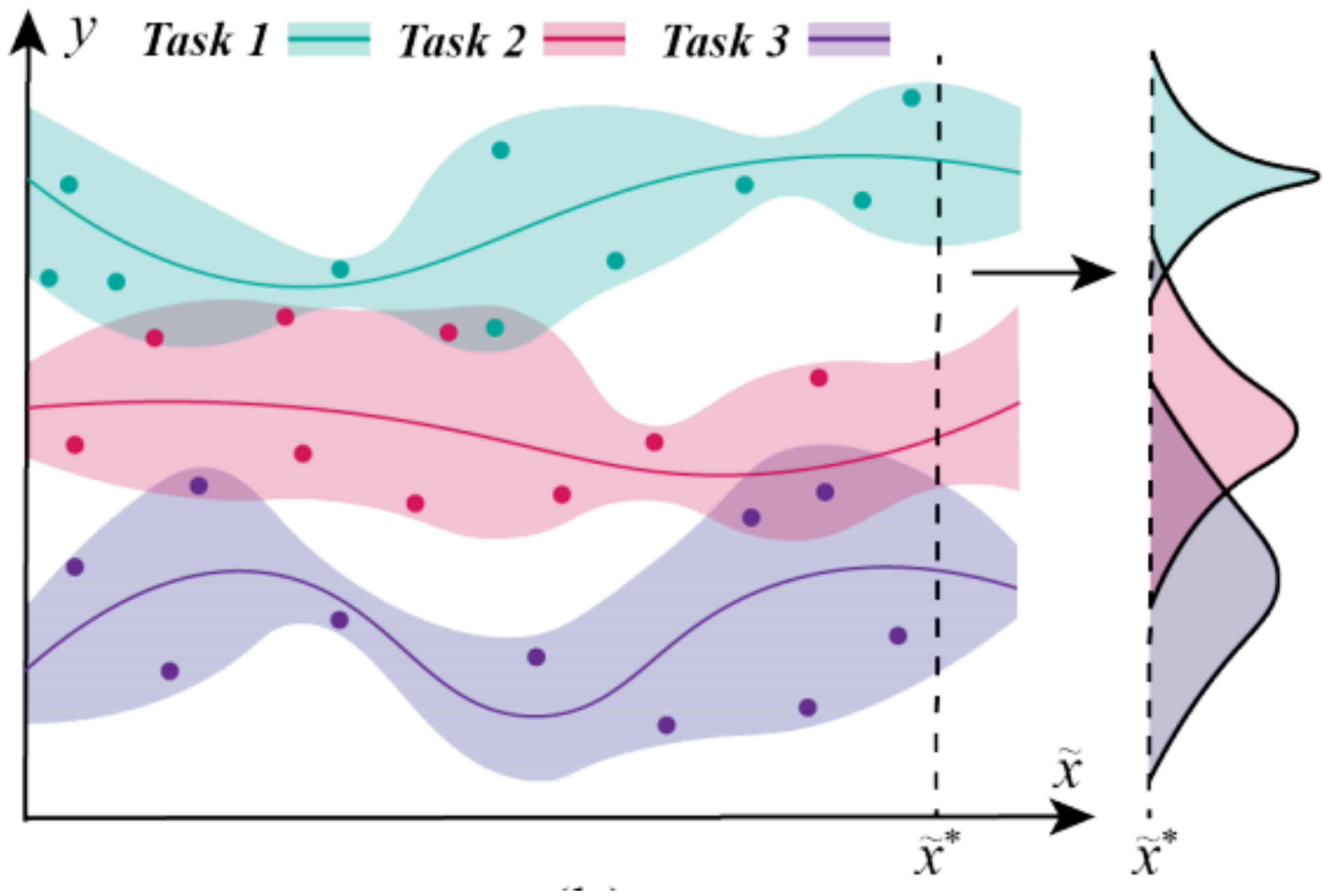
Nonparametric Bayesian learning (HDP-HMM)

[Wang, Zhang, Zhao,'Understanding V2V Driving Scenarios through Traffic Primitives', IEEE ITS, 2020]

Example 2: DPGP-MBRL

- Use Model-based RL with an infinite mixture of Gaussian Processes as the learned dynamics model.





- Do not require pre-training <- GP.
- Handle substantially different tasks <- mixture model.
- Online setting with streaming data <- streaming variational inference

Contents

- Hierarchical AI structures
- Trees
  - Decision trees
  - Random tree/forests
  - Monte Carlo Tree search, Alpha Go
- Hierarchical RL
  - Manager-worker
  - Option/Semi-MDP
- Hierarchical structures in Meta learning
  - Neural Processes

Meta learning

- Meta-learning (aka “learning to learn”) consists in training a model on various different tasks so that it can solve new learning tasks more efficiently using only a small number of training samples.

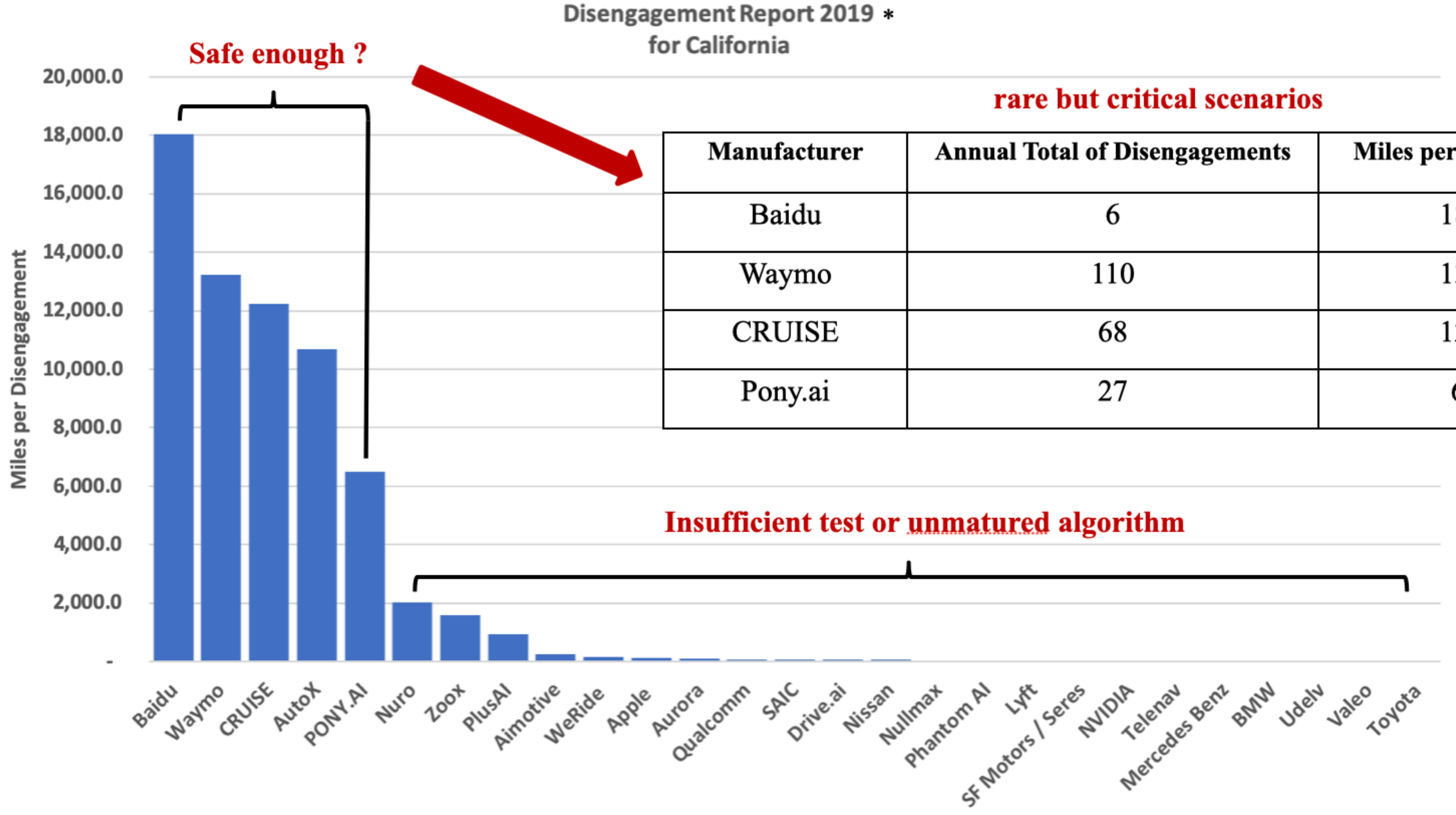
- Meta-RL is meta-learning on reinforcement learning tasks. After trained over a distribution of tasks, the agent is able to solve a new task by developing a new RL algorithm with its internal activity dynamics.

Koller, Torsten, et al. "Learning-based model predictive control for safe exploration." *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018.

Why Meta reinforcement learning

- Task imbalanced environment is prevalent for safety critical applications

Why do we care about generative models

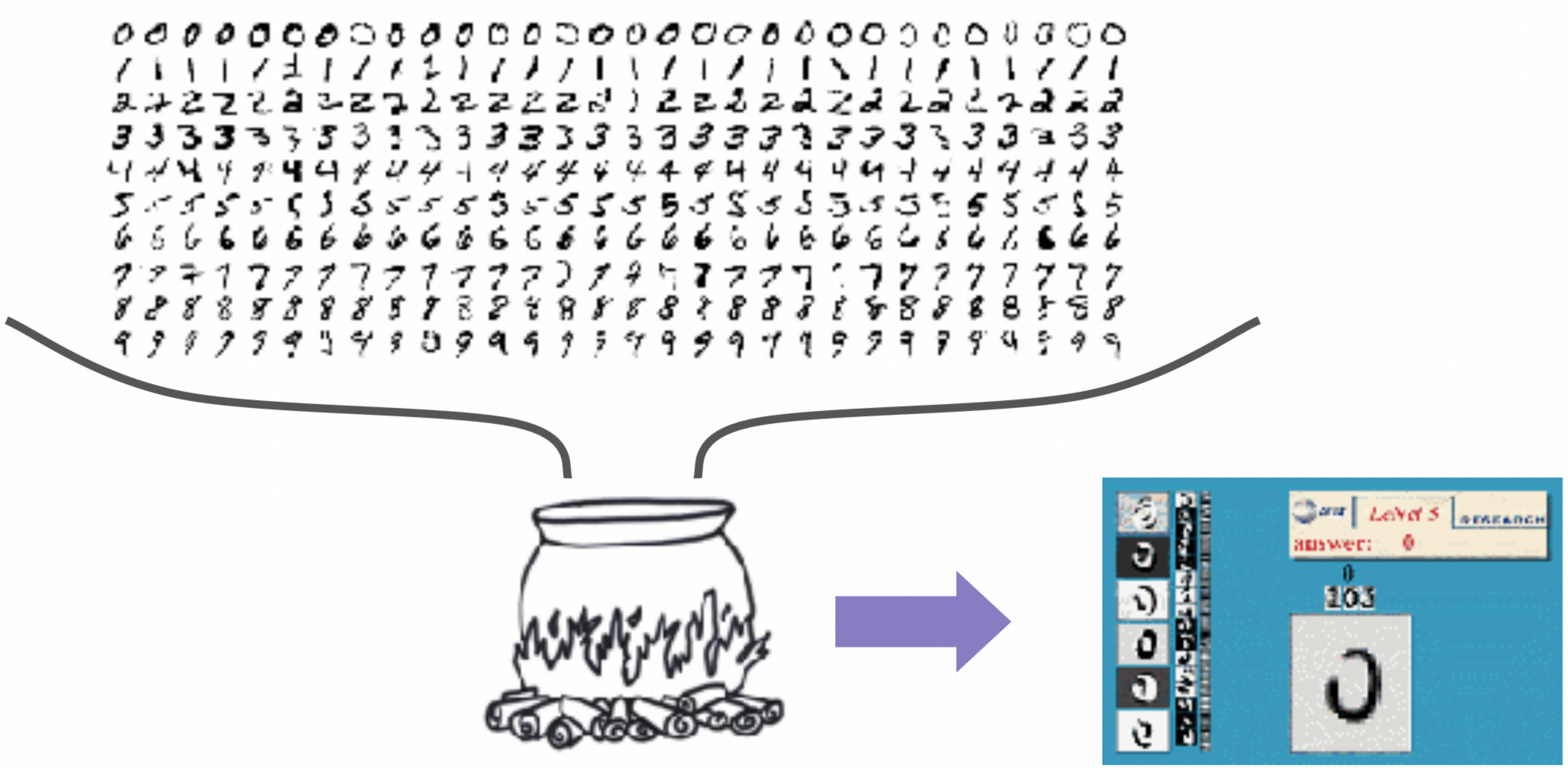


Manufacturer	Annual Total of Disengagements	Miles per Disengagement
Baidu	6	18,050.0
Waymo	110	13,219.4
CRUISE	68	12,221.2
Pony.ai	27	6,475.8

** Data source: California Department of Motor Vehicle 2019 disengagement report*

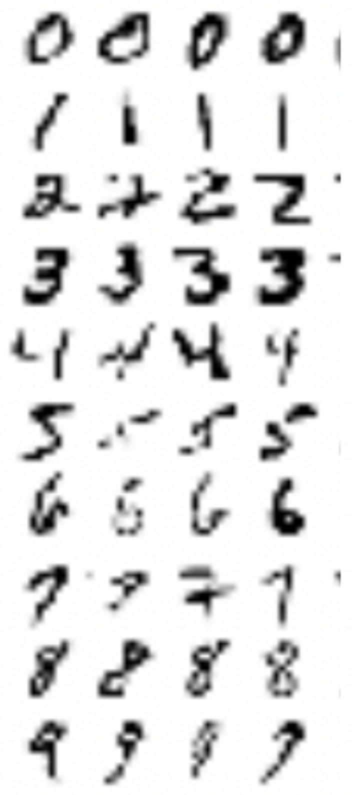
Ding Zhao | CMU

Machine learning with big data



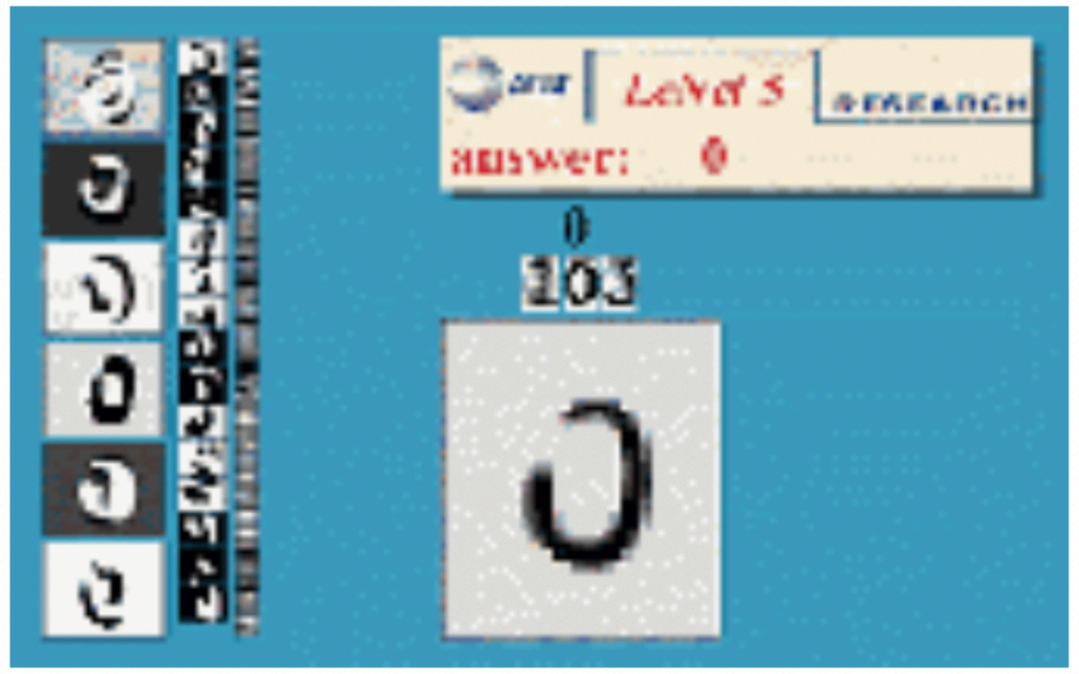
Meta-Learning— an idiosyncratic tutorial, Yee Whye Teh

Machine learning with small/imbalanced dataset

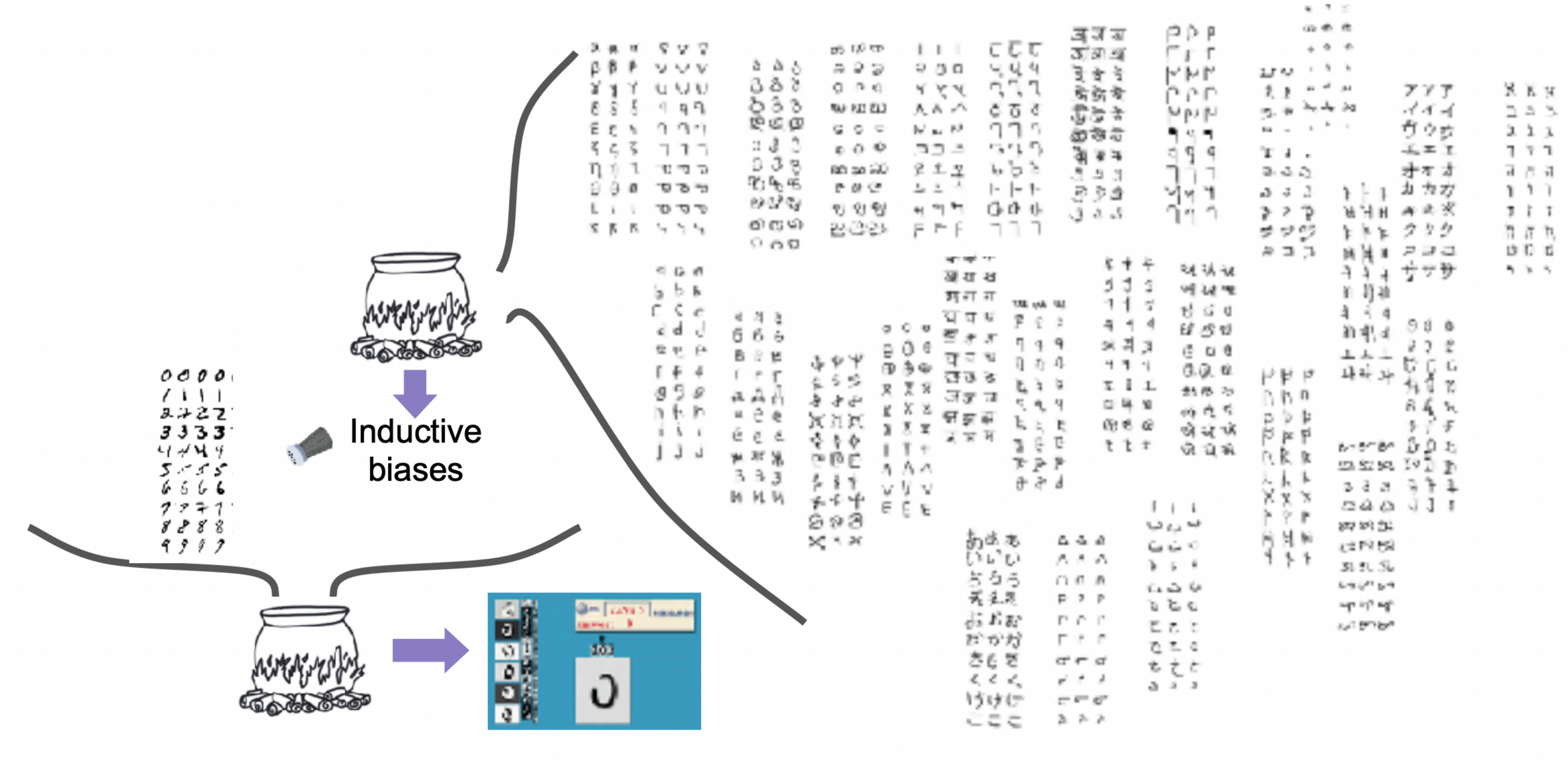


Inductive biases

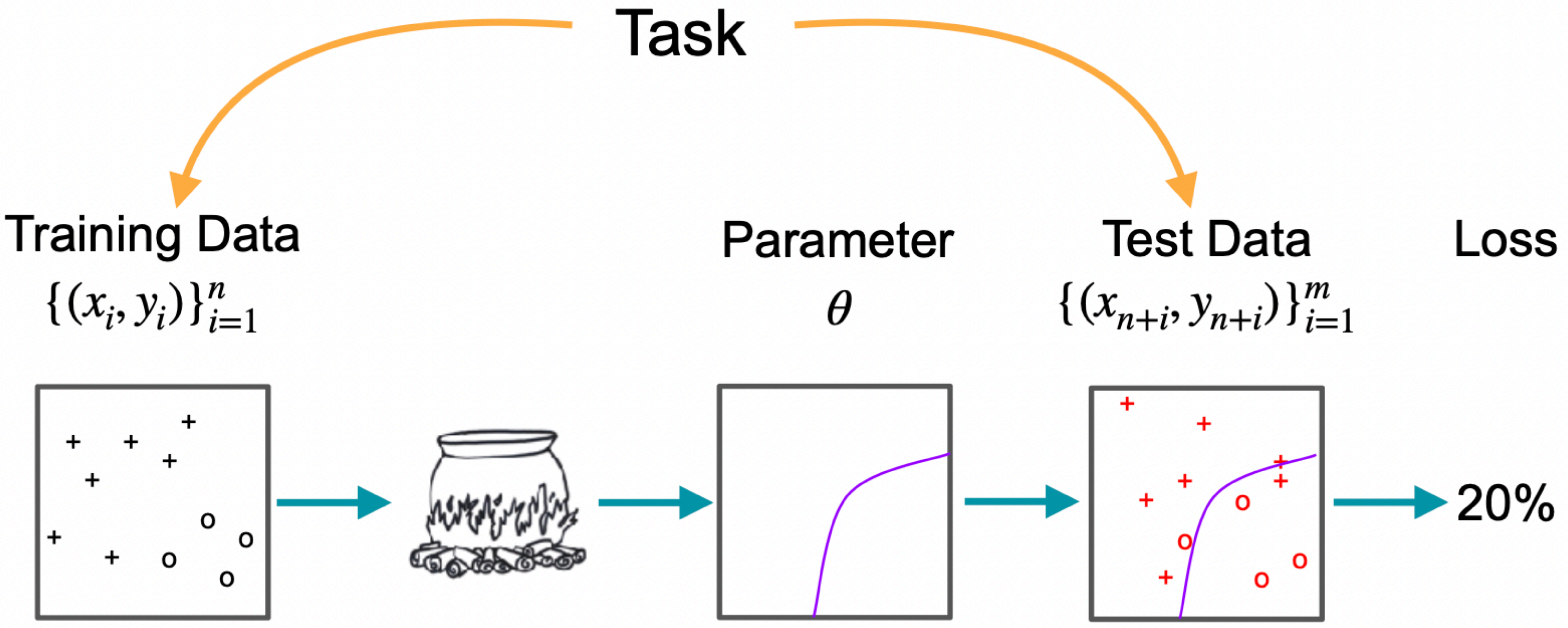
- Preprocessing
- Feature engineering
- Data augmentation
- Model architectures
- Regularisation



Meta-learning, learning-to-learn

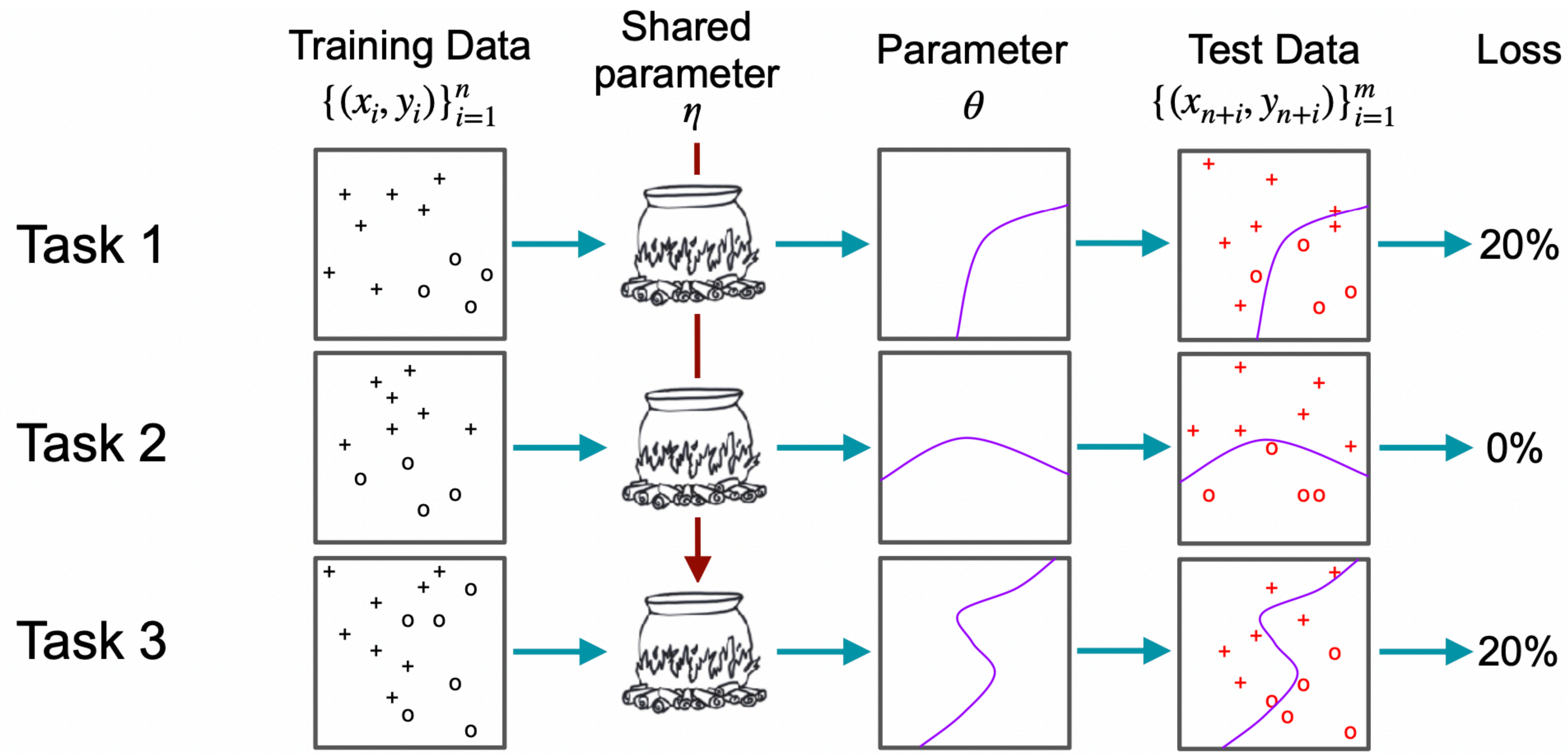


Single-task learning

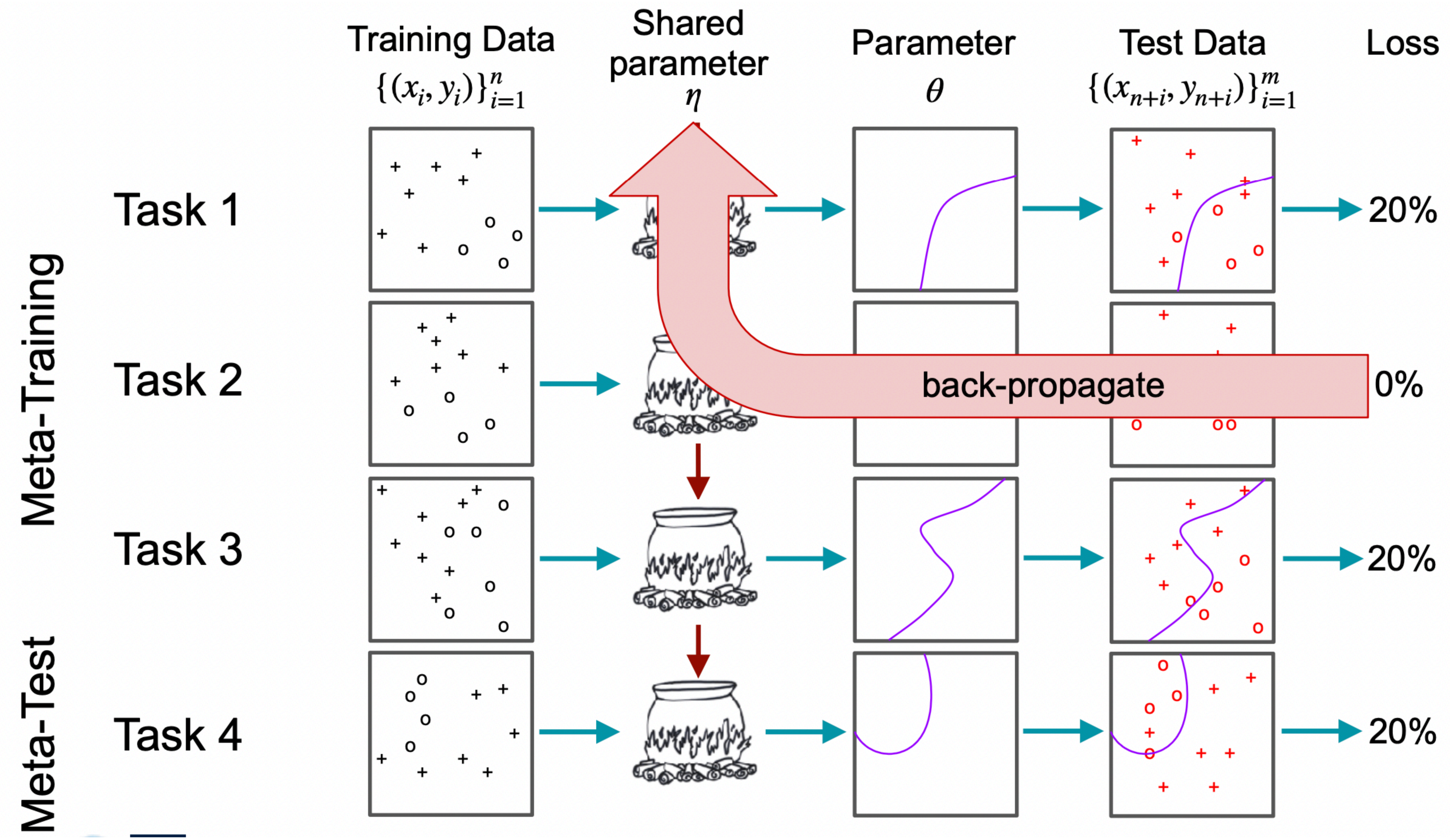


Meta-Learning— an idiosyncratic tutorial, Yee Whye Teh

Multi-task learning

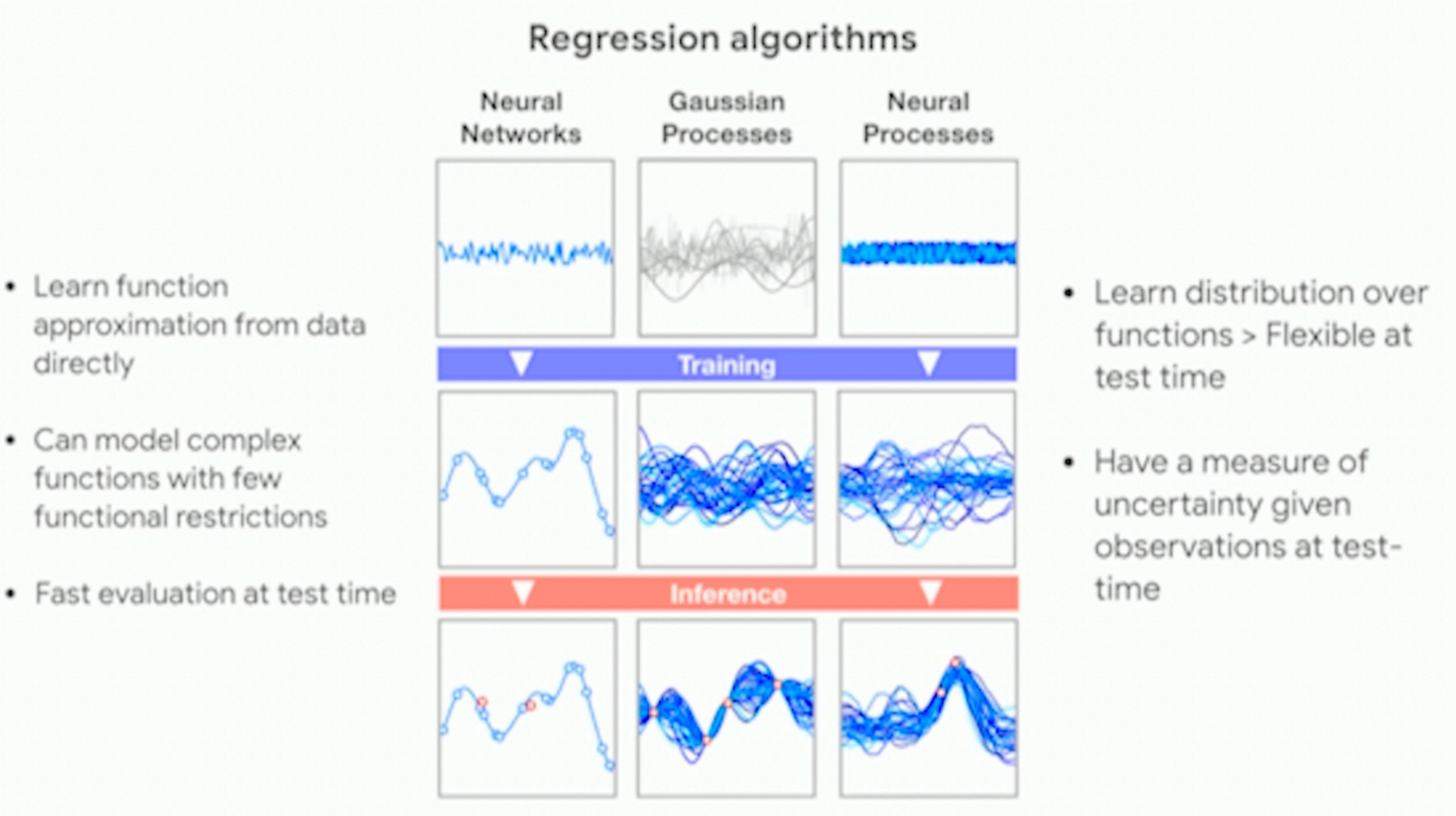


Meta-Learning



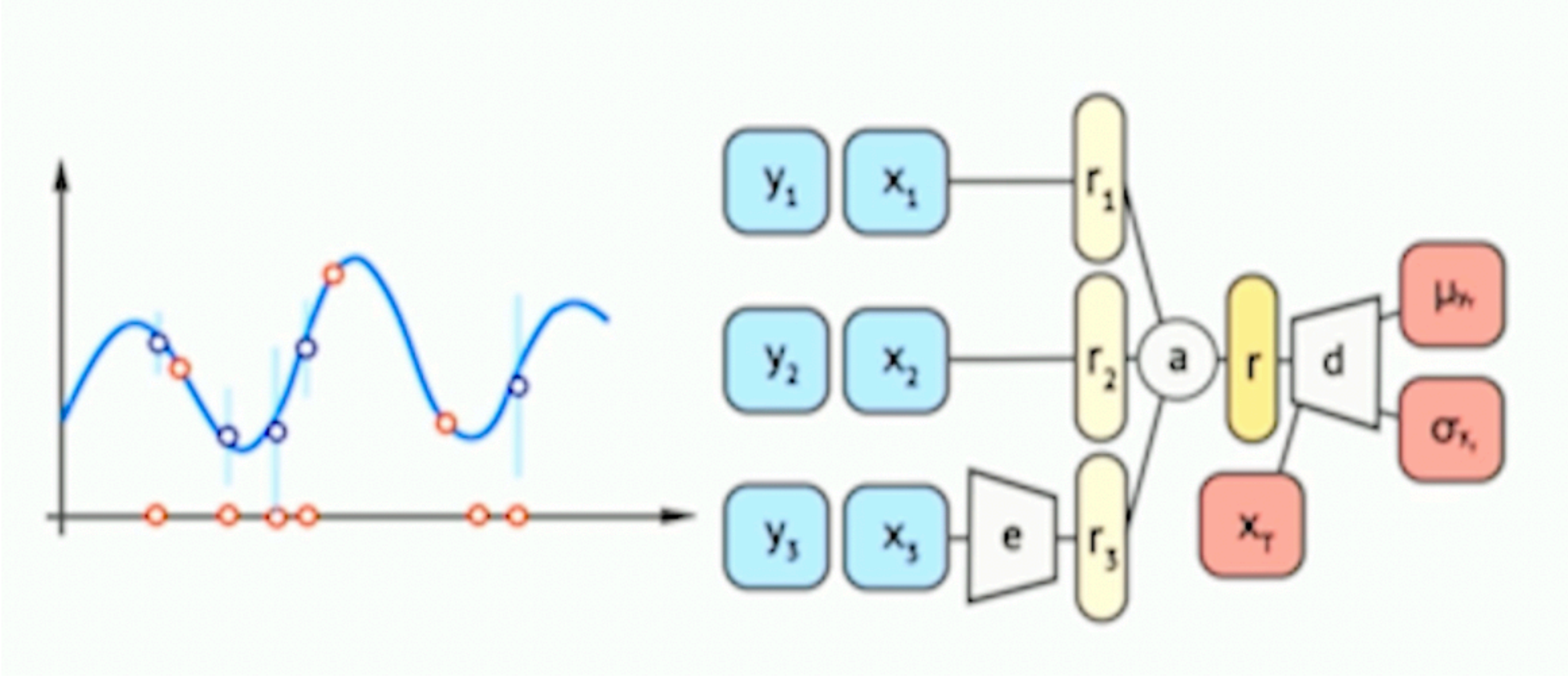
Meta-Learning— an idiosyncratic tutorial, Yee Whye Teh

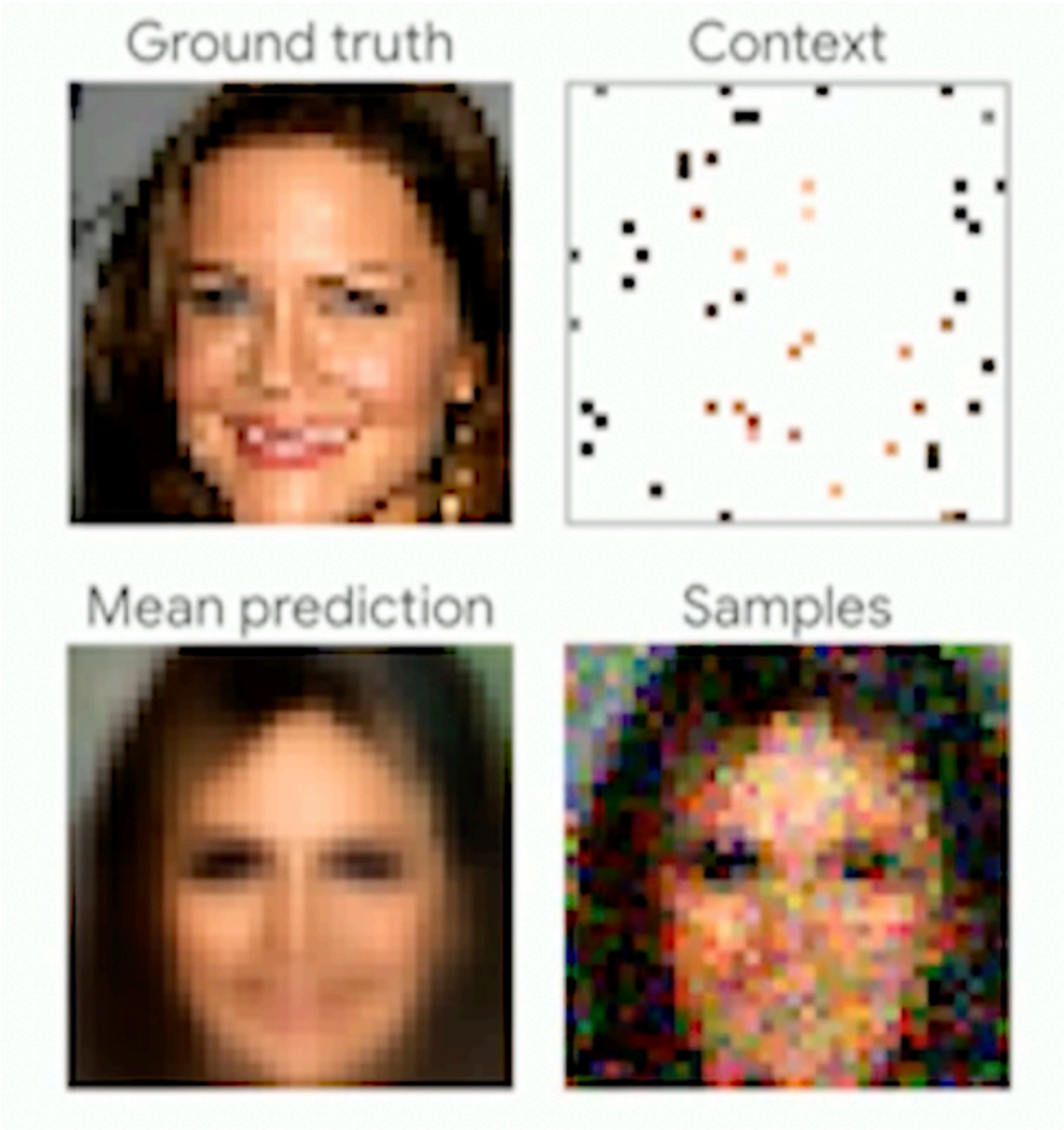
Neural Processes



https://vimeo.com/312299226

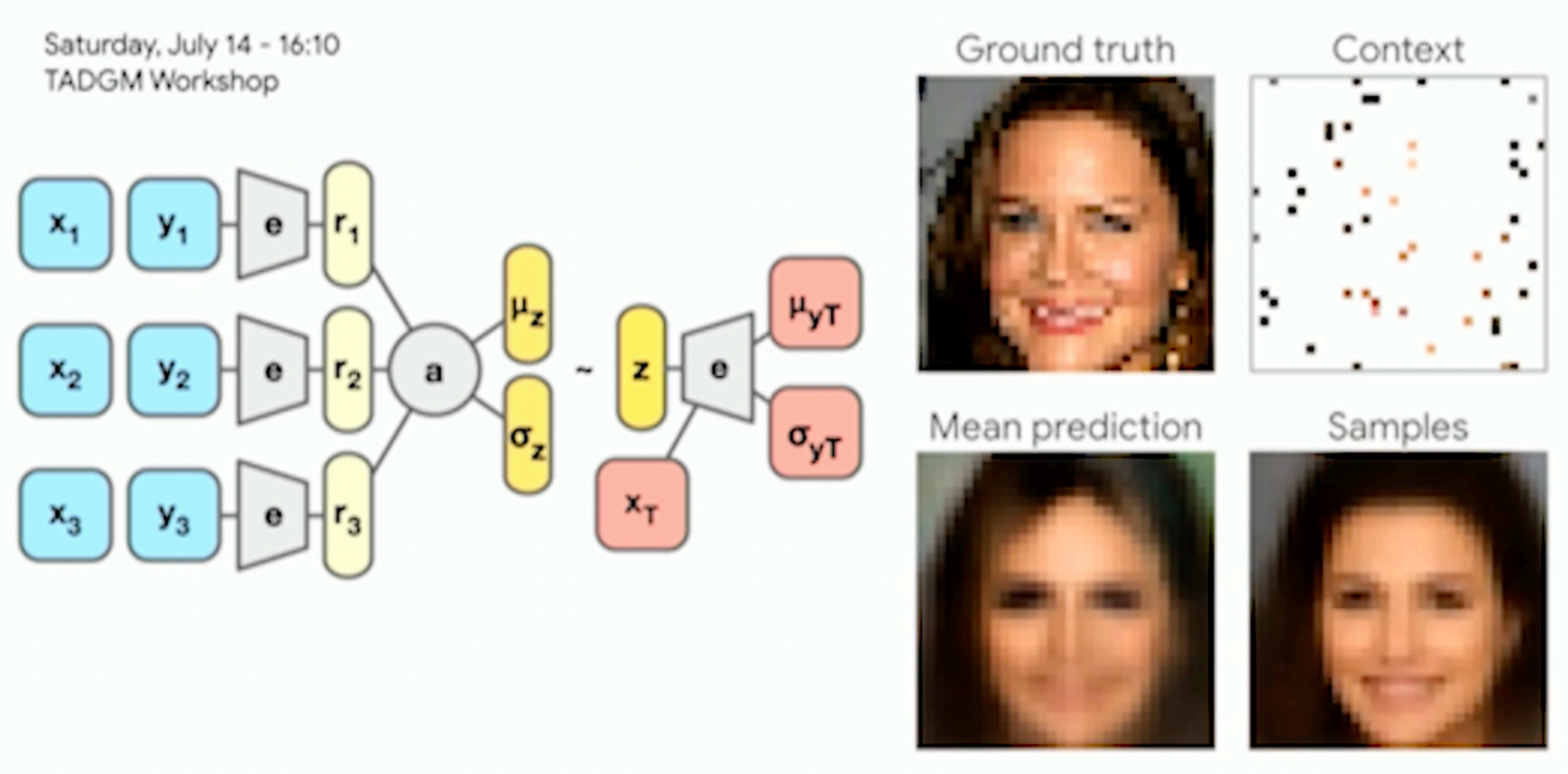
Neural Processes





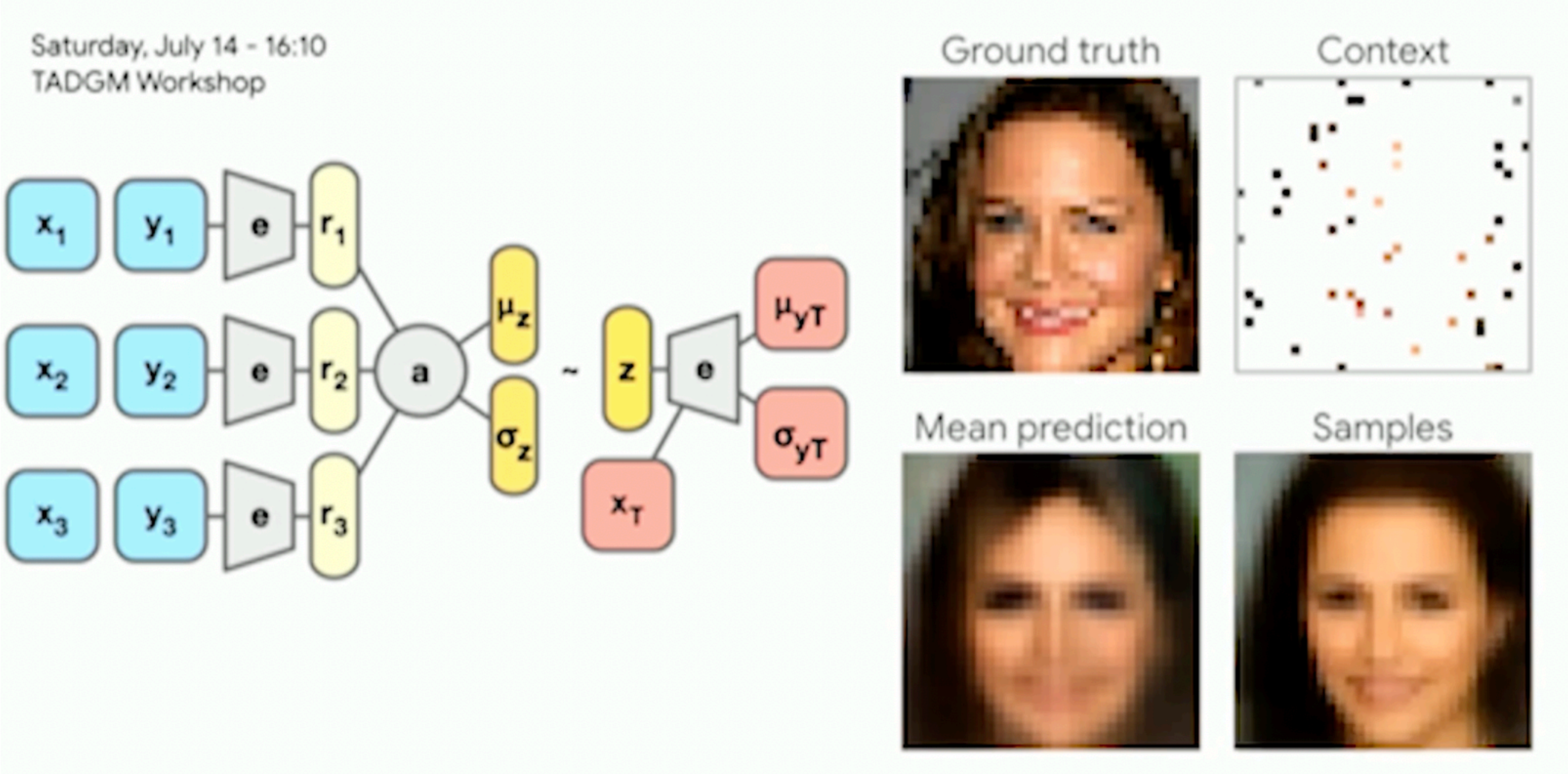
https://vimeo.com/312299226

Generate Coherent Samples



https://vimeo.com/312299226

Generate Coherent Samples



https://vimeo.com/312299226

Skepticism on Hierarchical RL

- “Surprisingly, we find that most of the empirical benefit of hierarchy in our considered settings can be attributed to improved exploration.”
- “These proposed exploration methods enable non-hierarchical RL agents to achieve performance competitive with state-of-the-art HRL. Although our analysis is empirical and thus our conclusions are limited to the tasks we consider, we believe that our findings are important to the field of HRL.”
- “Our findings reveal that only a subset of the claimed benefits of hierarchy are achievable by current state-of-the-art methods, even on tasks that were previously believed to be approachable only by HRL methods.”

Nachum, Ofir, Haoran Tang, Xingyu Lu, Shixiang Gu, Honglak Lee, and Sergey Levine. 2019. “Why Does Hierarchy (Sometimes) Work So Well in Reinforcement Learning?” *arXiv [cs.LG]*. arXiv. http://arxiv.org/abs/1909.10618.

Summary

- Hierarchical AI structures
- Trees
    - Decision trees
    - Random tree/forests
    - Monte Carlo Tree search, Alpha Go
- Hierarchical RL
    - Manager-worker
    - Option/Semi-MDP
- Hierarchical structures in Meta learning
    - Neural Processes

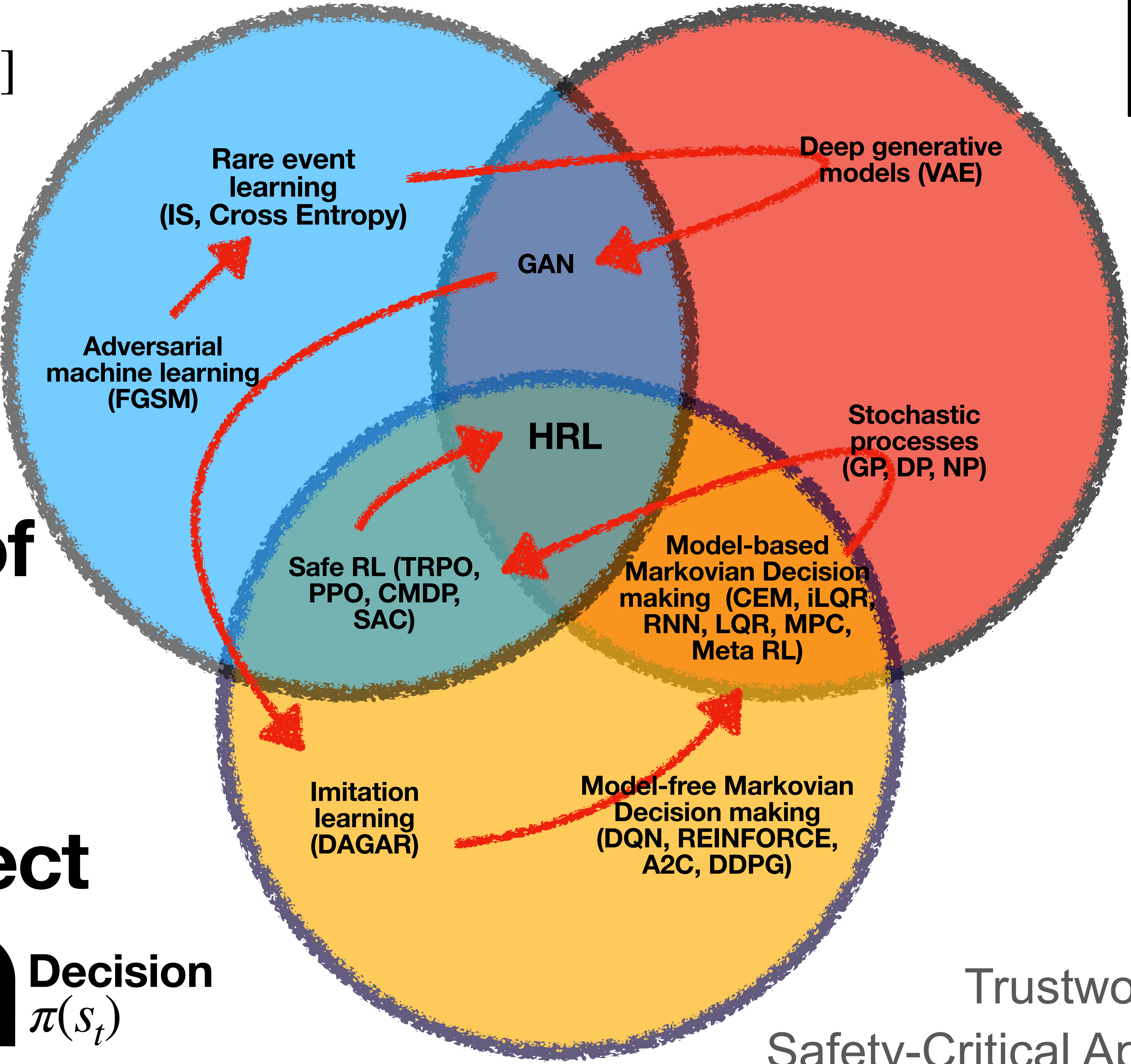
Additional reading materials

- High level intro to HDP
  Tenenbaum, Joshua B., et al. "How to grow a mind: Statistics, structure, and abstraction." *science* 331.6022 (2011): 1279-1285.

- Frans, Kevin, Jonathan Ho, Xi Chen, Pieter Abbeel, and John Schulman. 2017. “Meta Learning Shared Hierarchies.” ICLR 2018

Evaluation
$\mathbb{E}_{p,\pi}[\Sigma_t r(s_t, a_t)]$

Modeling
$p(s_{t+1} \mid s_t, a_t)$



Summary of TAIAT: 12 lectures +25 papers +Final project

Decision
$\pi(s_t)$

Trustworthy AI for Safety-Critical Applications