CS 6756: Learning for Robot Decision Making

Sanjiban Choudhury

PORTAL

Cornell Bowers CIS
Computer Science

WHAT A TIME TO BE ALIVE!

Exciting time for Artificial Intelligence

Deep Q Networks

Playing Atari with Deep Reinforcement Learning

Volodymyr Mnih Koray Kavukcuoglu David Silver Alex Graves Ioannis Antonoglou
Daan Wierstra Martin Riedmiller
DeepMind Technologies

AlphaGo

Transformers

GPT-4

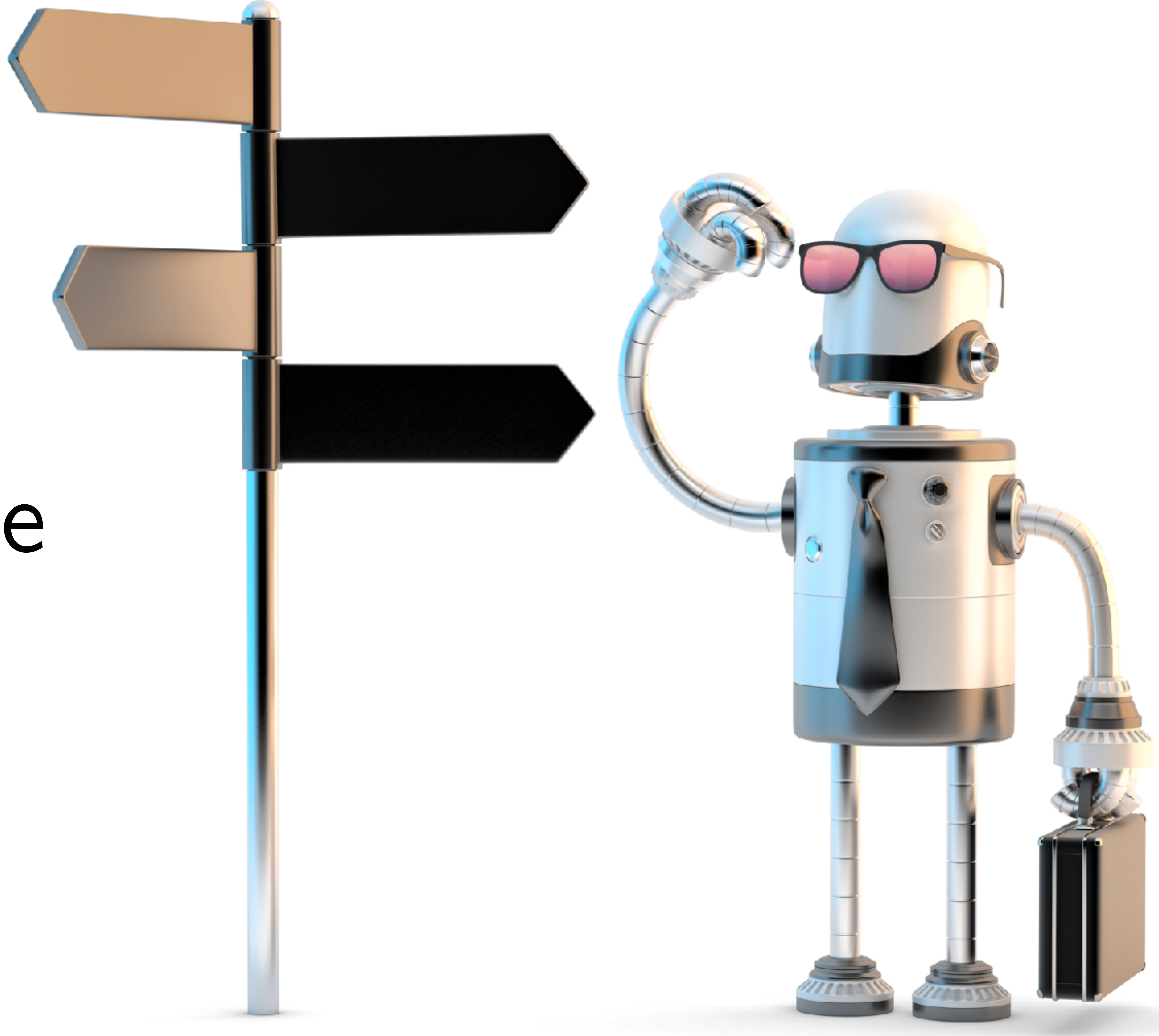
DALL-E

An image generated by DALL-E 2 based on the text prompt "Teddy bears working on new AI research underwater with 1990s technology"

stability ai

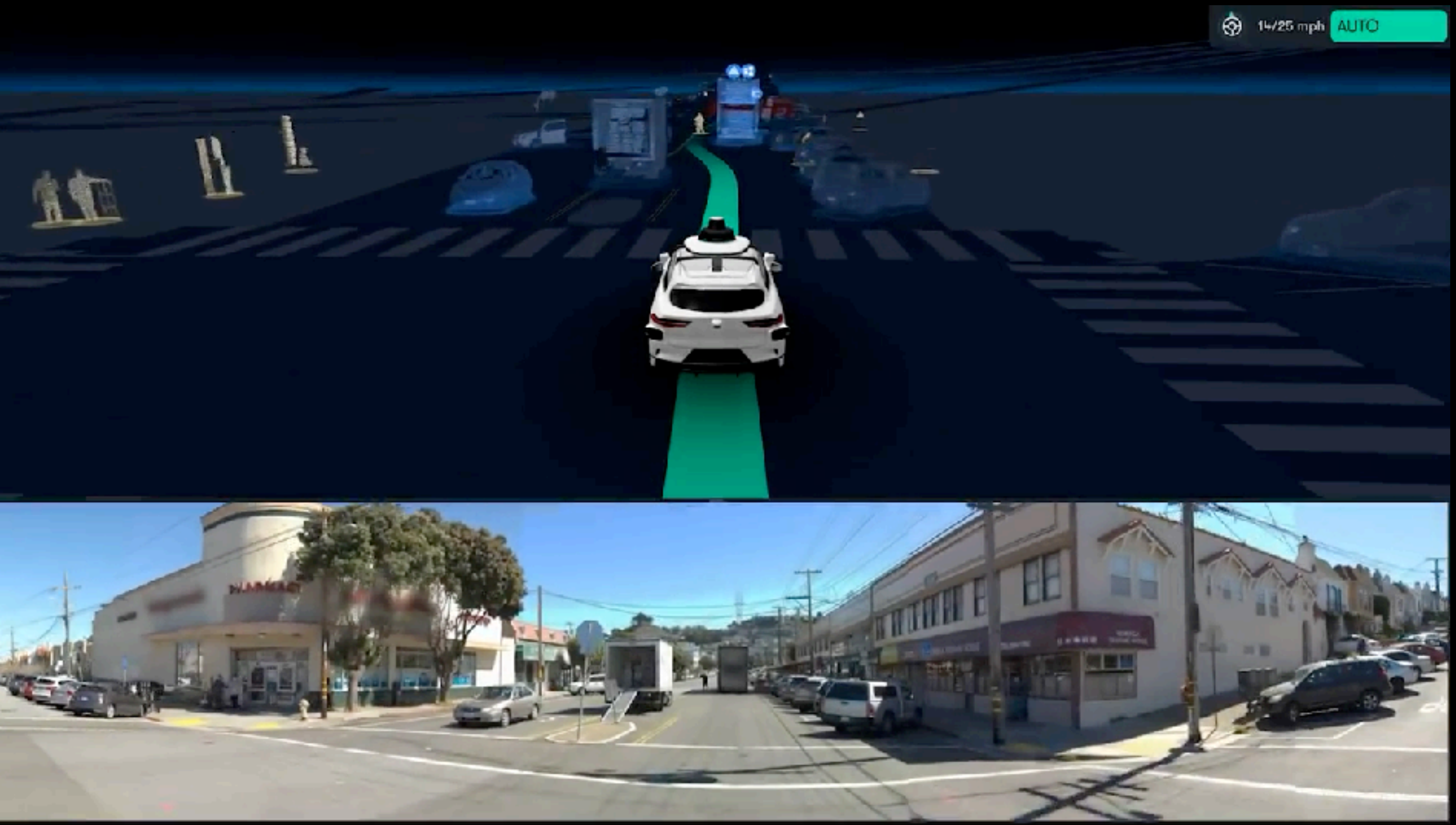
2013 2016 Today

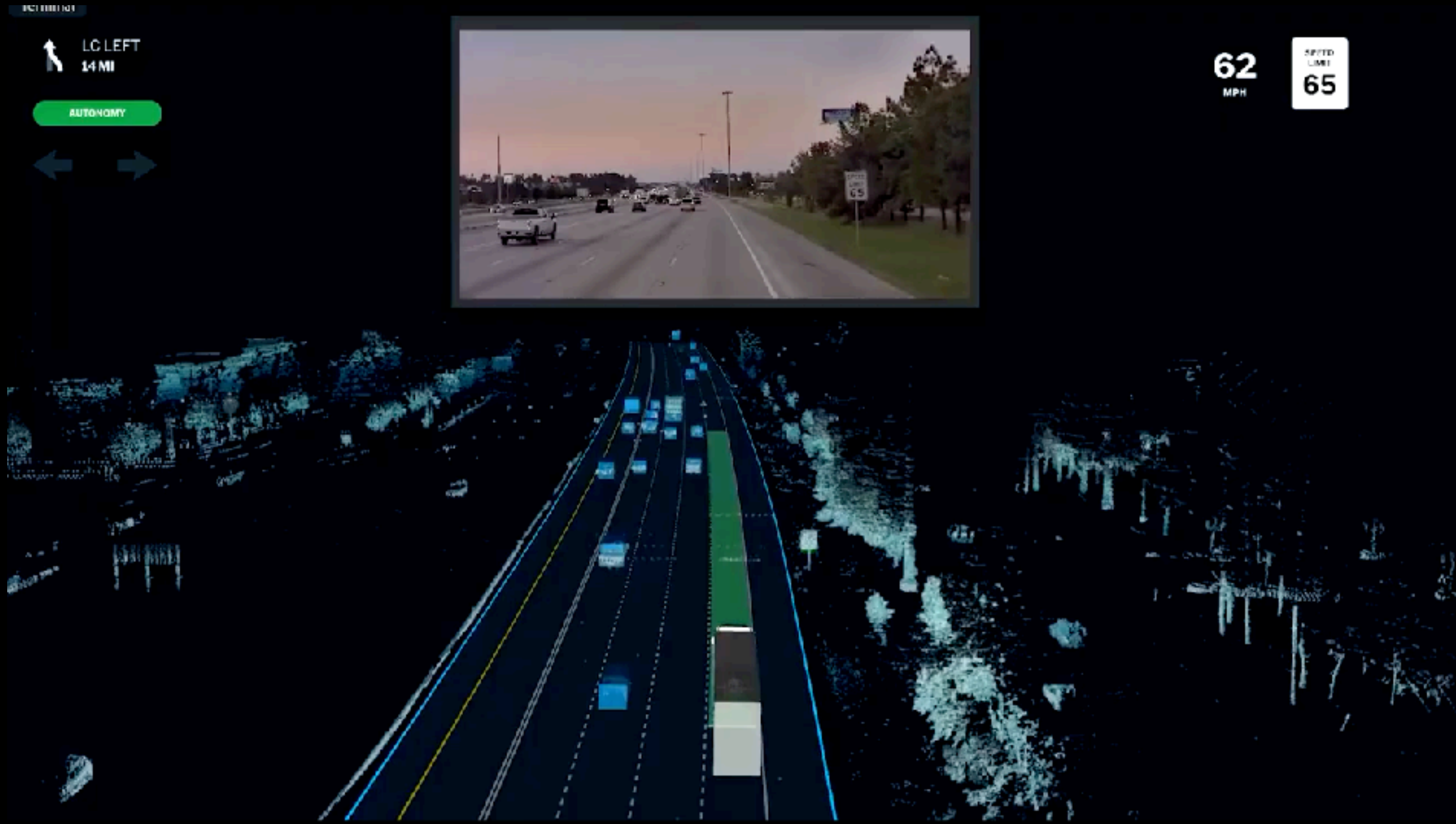
Where are the robots?



Robots are not far behind!

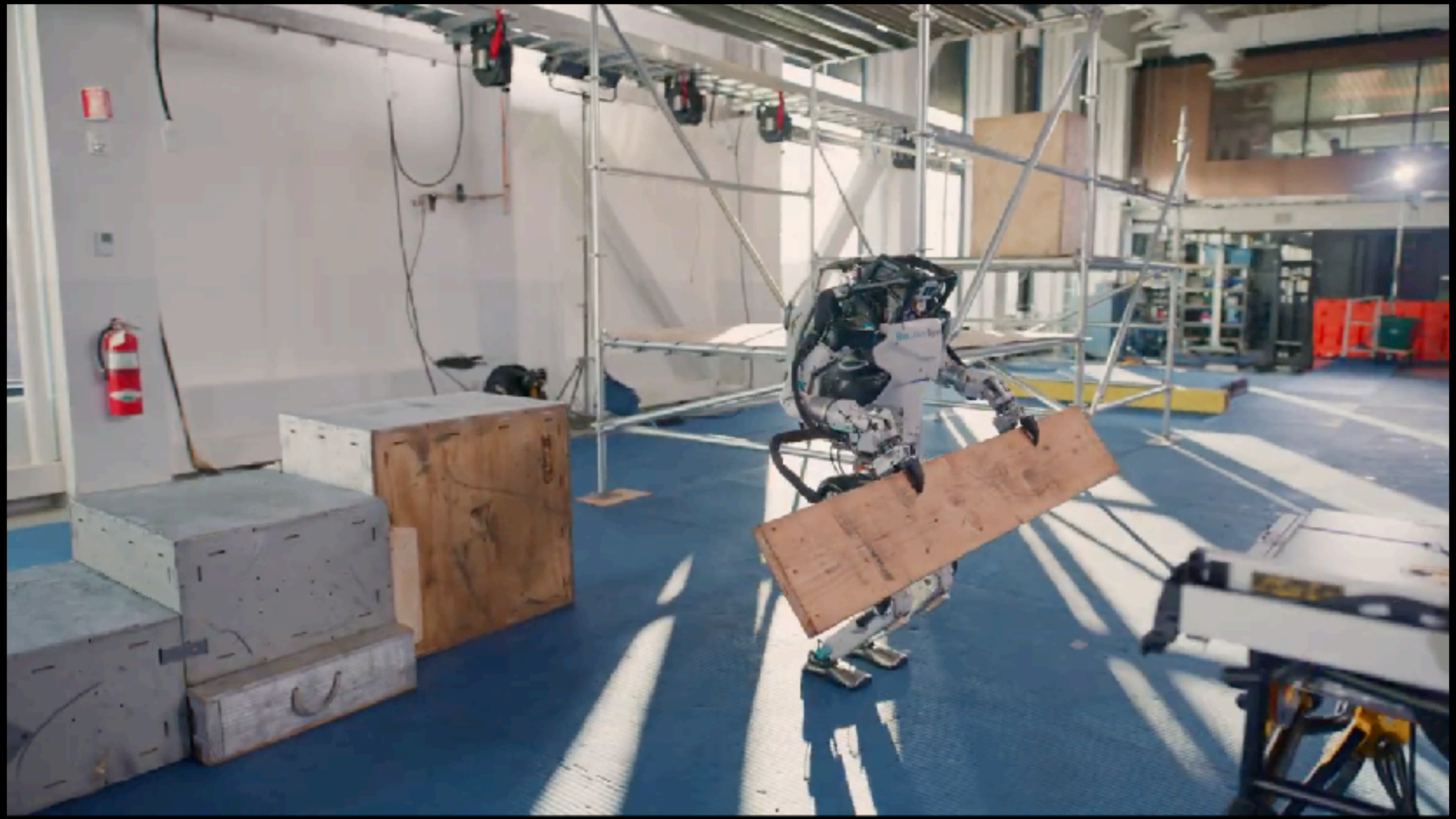
Robots are not far behind!





Self-driving companies going driverless ...

Robots are not far behind!



Boston Dynamics are starting to sell their robots ...

Robots are not far behind!

Drones are getting more reliable ...

But ...

... robots are not in millions of homes yet.

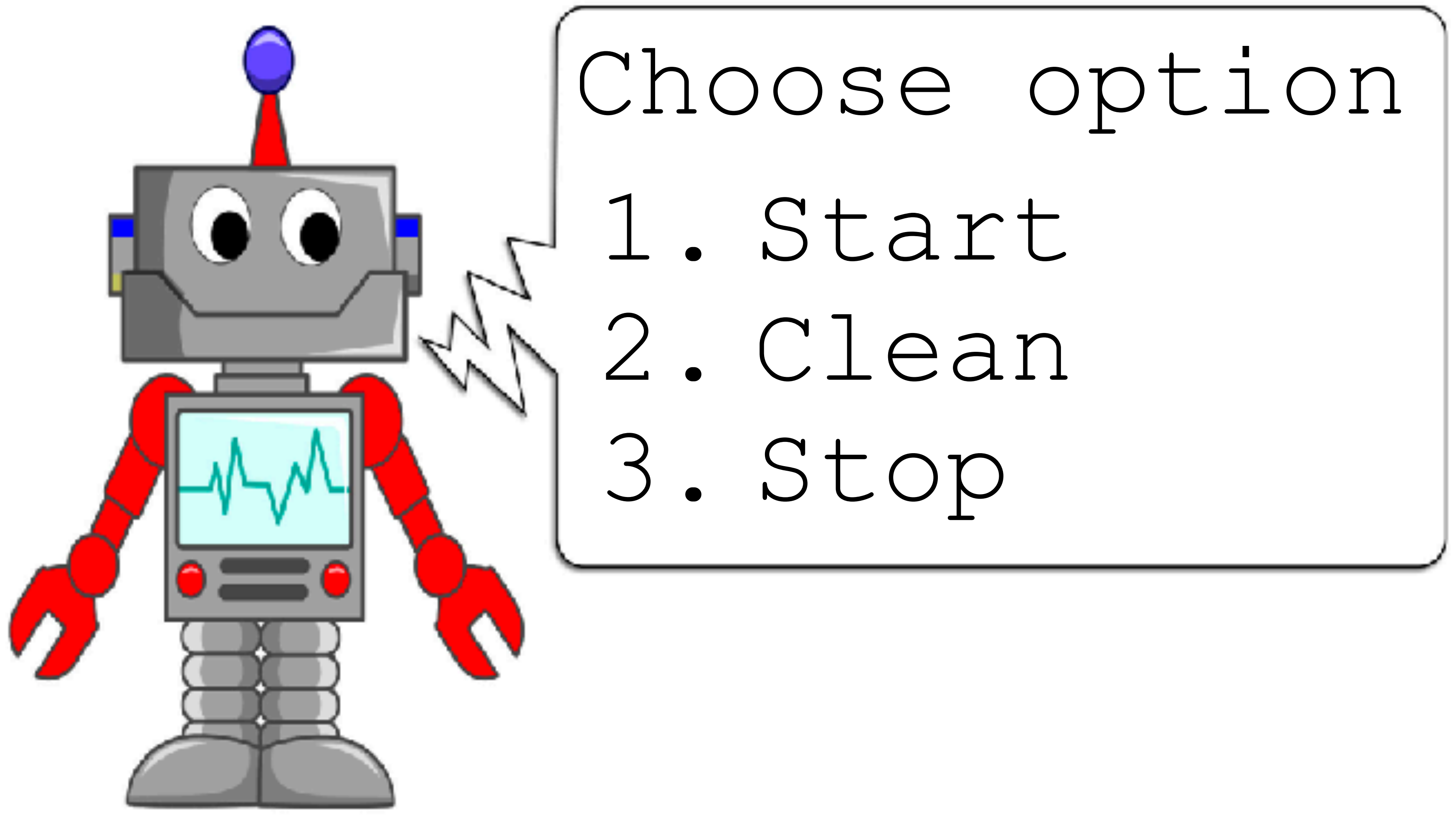
Why?

Why are robots not in millions of homes yet?

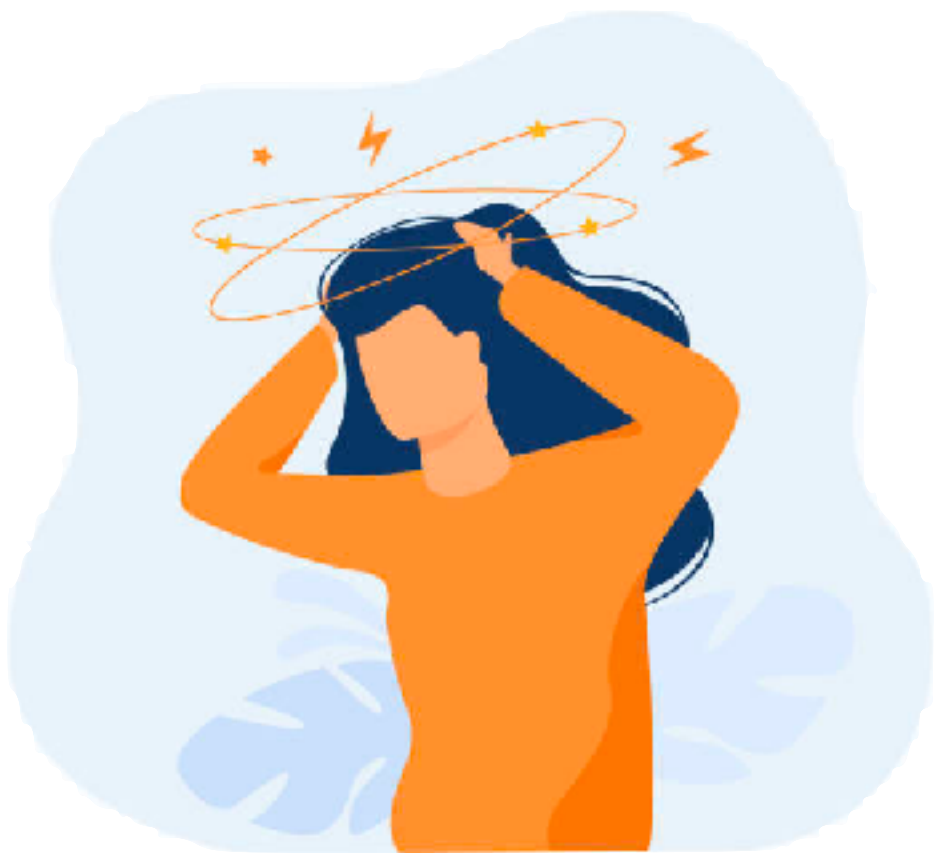
PollEv.com/sc2582

The way we program robots today is ... rigid!

Engineers hand-craft behaviors



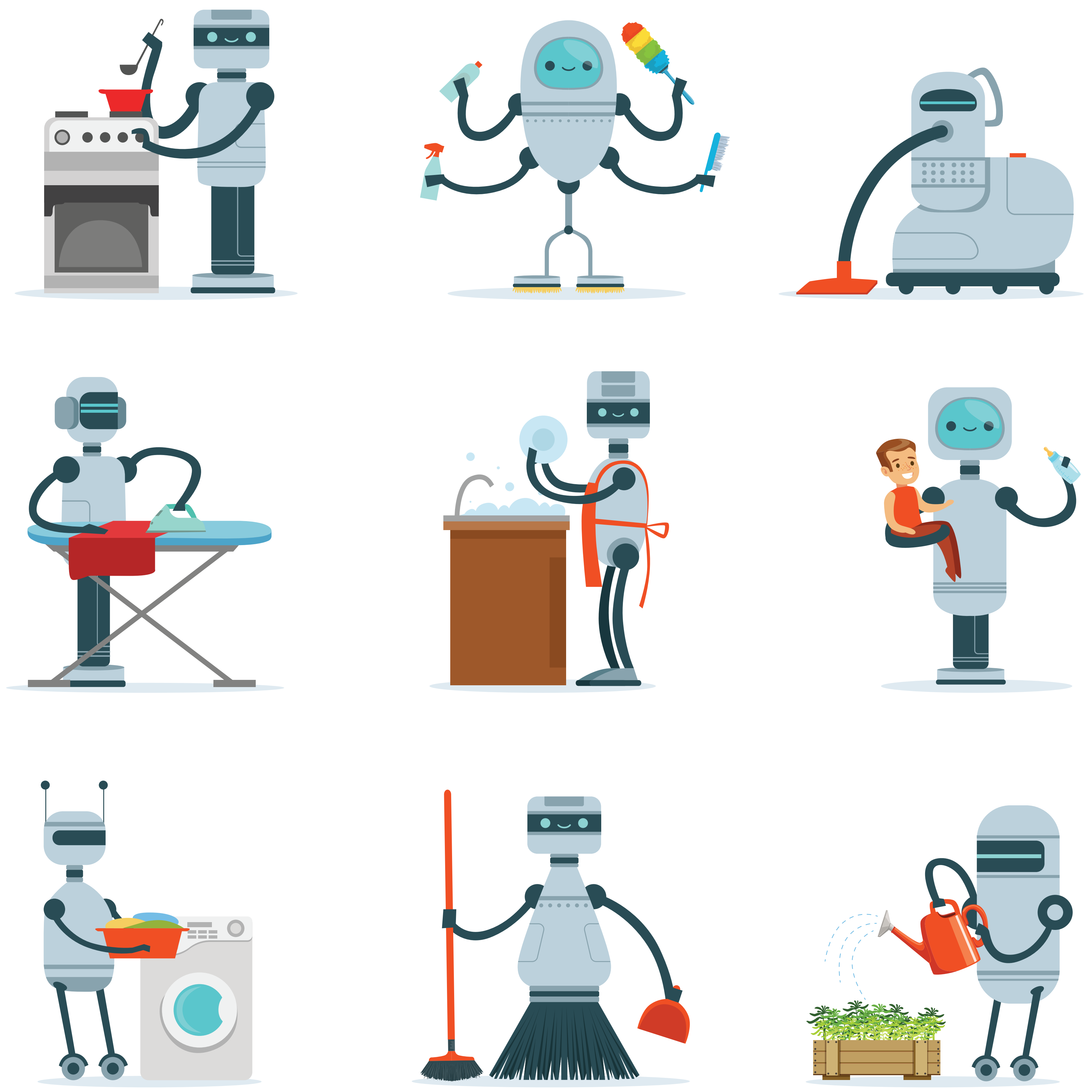
Ship robot



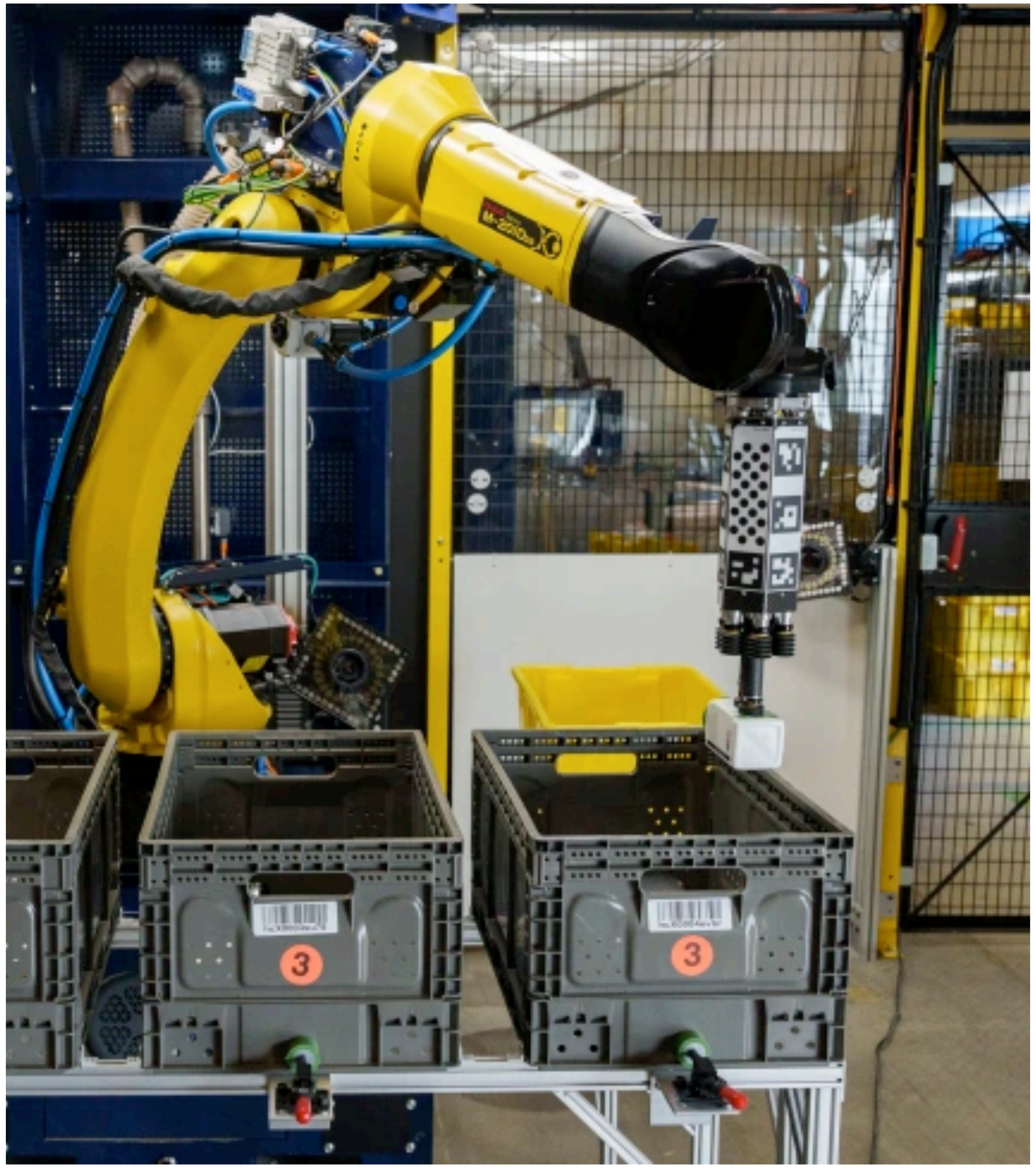
Frustrate users!

Not flexible enough to be used by everyday users for everyday tasks

This restricts robots to a CLOSED world



The Dream

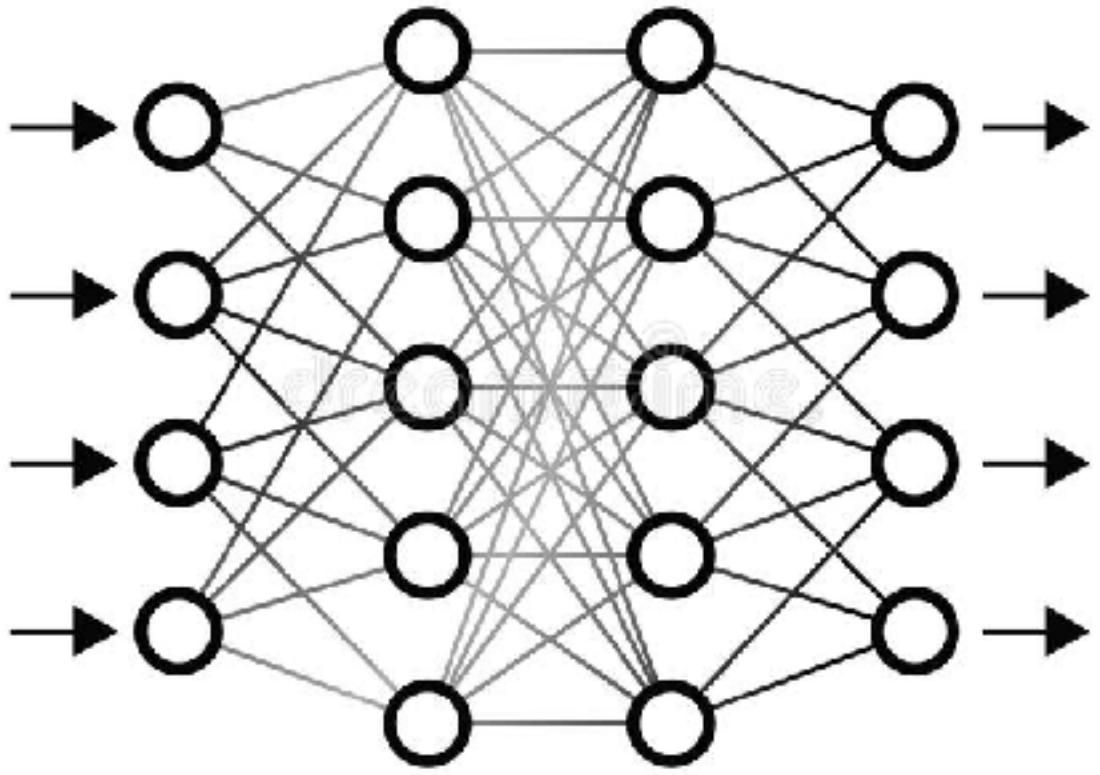


Reality

How can we get robots out of the factory into the OPEN WORLD?

Robotics 2.0: Scale and improve with data

Formulate as a learning problem

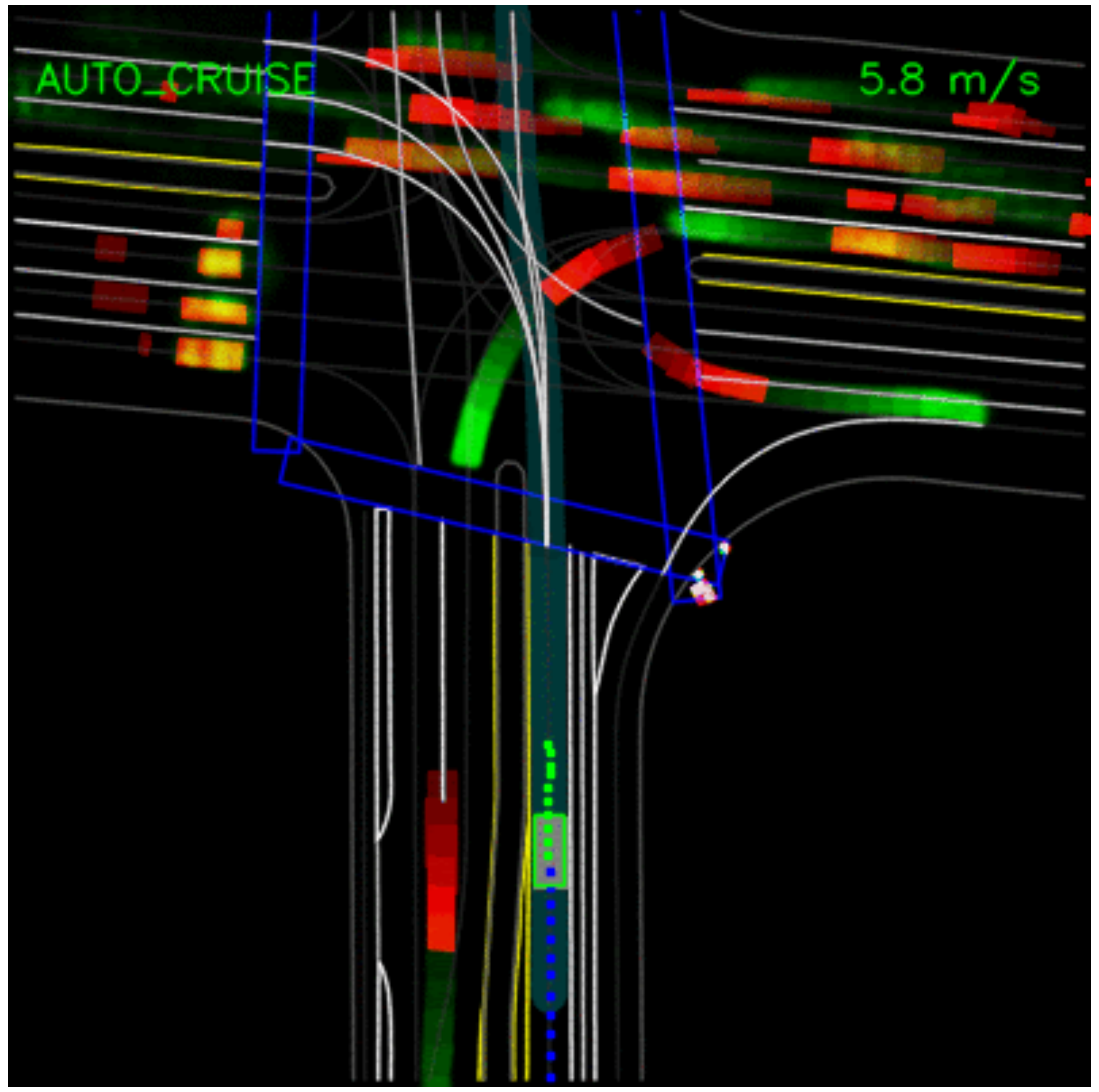


Invest in good ML pipelines

Robotics 2.0: Scale and improve with data

Formulate as a learning problem

Invest in good ML pipelines



Self-driving led the way!

What is special about learning for robot decision making?

Standard learning

$$\min_{\theta} \; \mathbb{E}_{x,y} \, \ell(y, \theta(x))$$

x is a sequence of inputs, y is a sequence of outputs, θ is a model

In decision making:

x is the sequence of observations

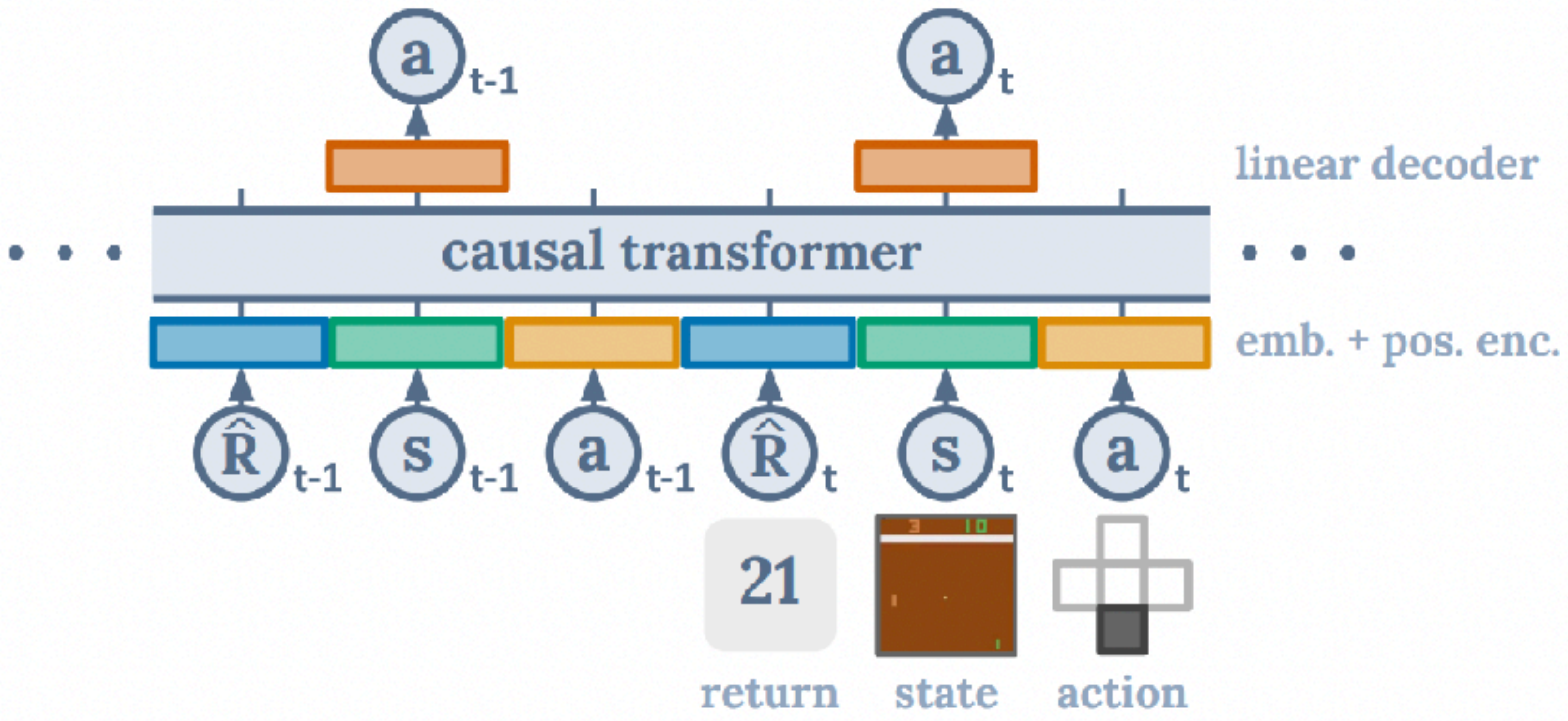
y is the sequence of decisions (plan)

What is special about learning for robot decision making?

$$\min_{\theta} \mathbb{E}_{x,y}\, \ell(y, \theta(x))$$

x is a sequence of inputs, y is a sequence of outputs, θ is a model

Transformers are pretty standard choice for the model



What is special about learning for robot decision making?

$$\min_{\theta} \mathbb{E}_{x,y} \ell(y, \theta(x))$$

x is a sequence of inputs, y is a sequence of outputs, θ is a model

Problem 1: What's special about the data?

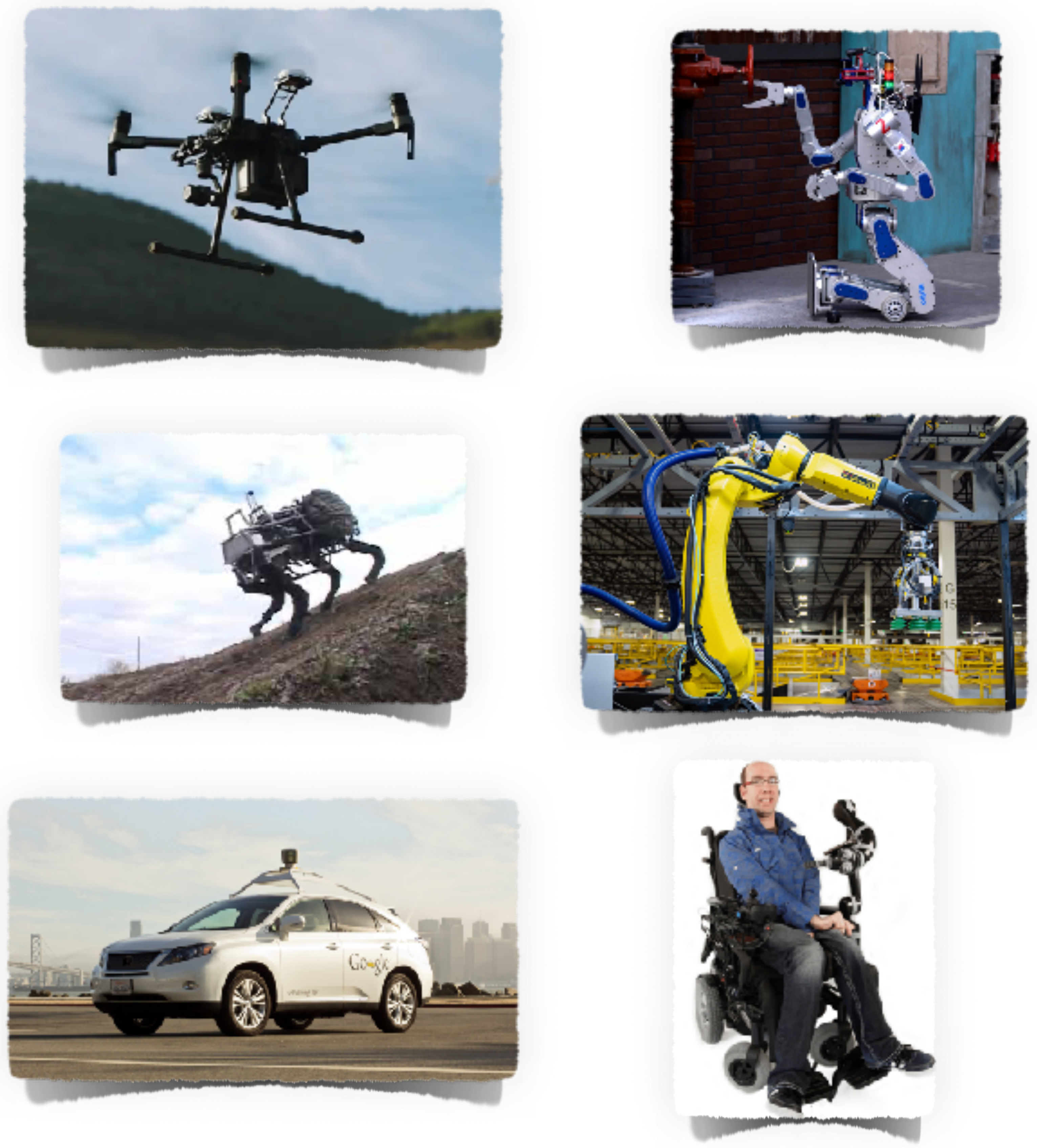
What is special about learning for robot decision making?

$$\min_{\theta} \mathbb{E}_{x,y} \, \ell(y, \theta(x))$$

x is a sequence of inputs, y is a sequence of outputs, θ is a model

Problem 2: What's special about the loss?

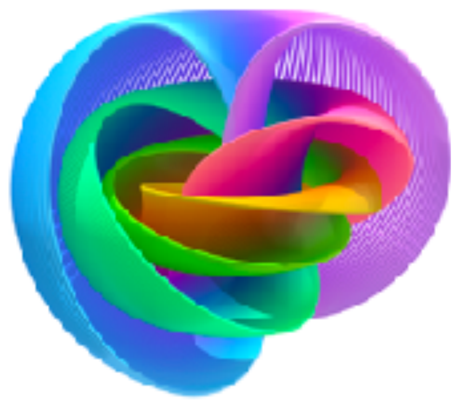
WHY this course?



Take *any*
robot application

Formulate as a Markov Decision Problem (MDP)

Analyze and Solve MDPs
(unified framework + algorithmic toolkit)



Develop a unified framework
(that ties old and new ideas)

Belonging

The Crew

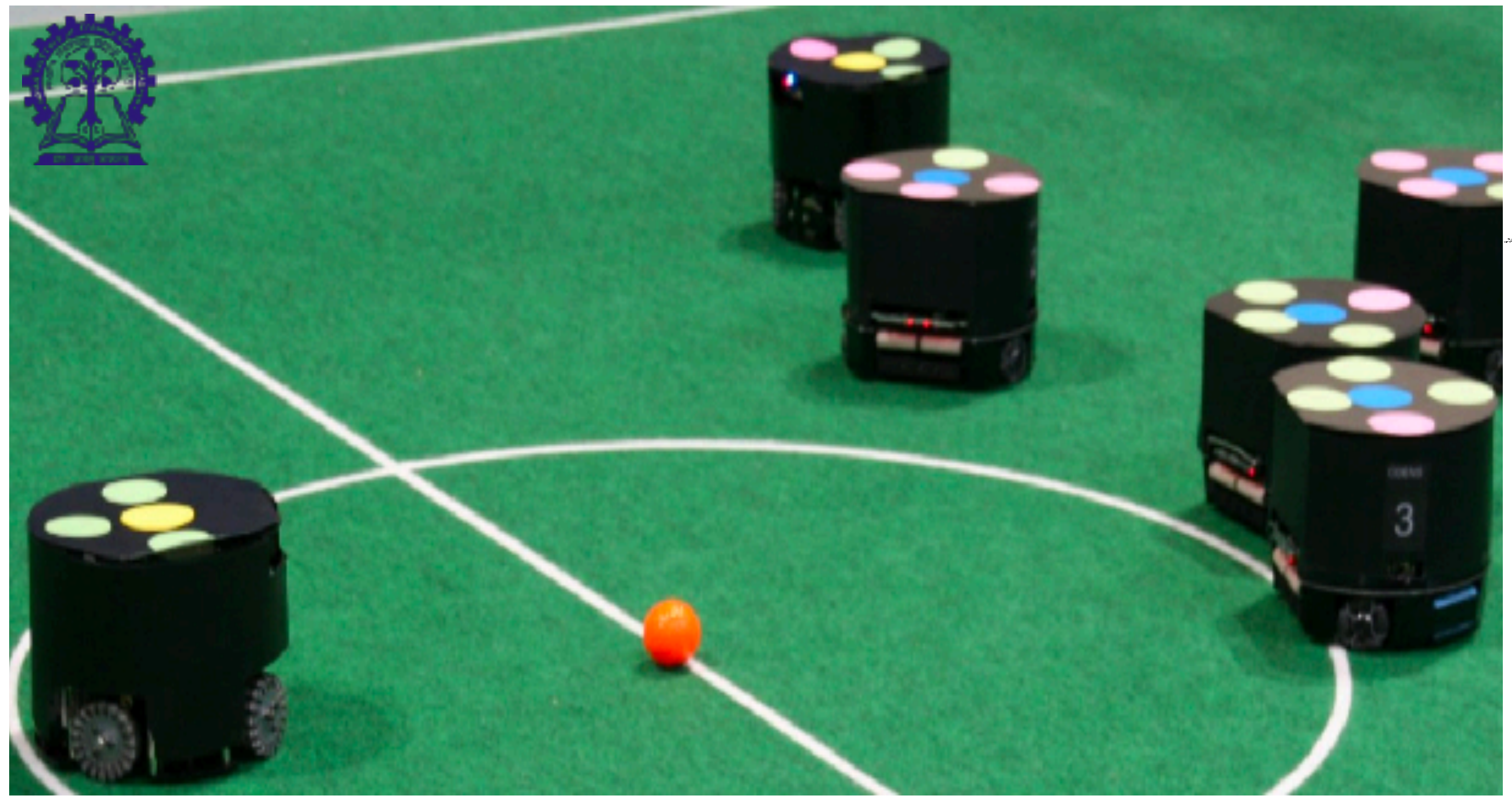
Build robots that can learn from humans!

"Sanjiban"

He / Him

Office hours:
Tues 11:30 – 1:30pm
Gates 413B

Undergrad



PhD

PostDoc

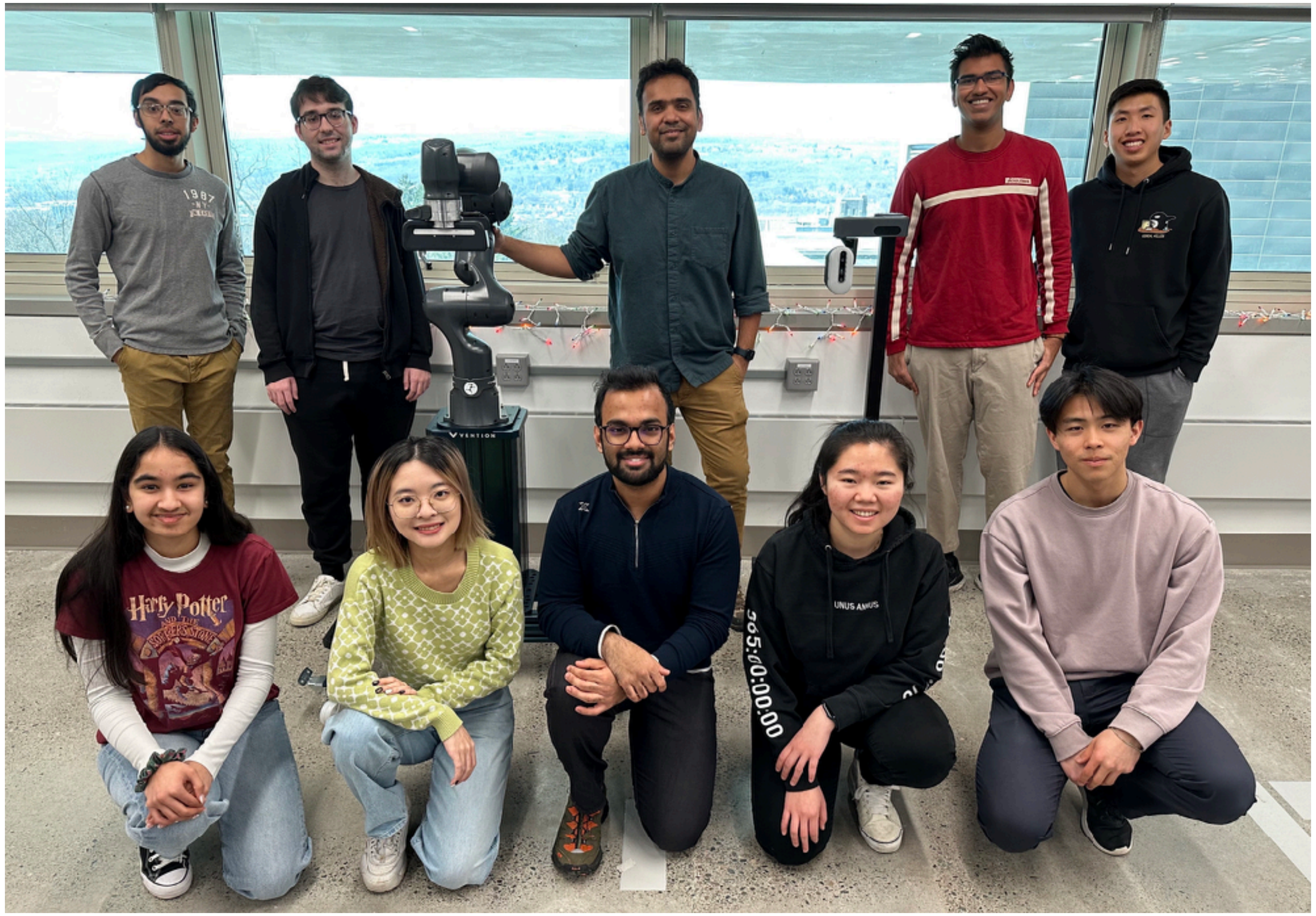
Research Engineer

- Kushal Kedia, PhD Student
- Research Interests: Forecasting, Imitation Learning
- Fun Fact about me: I love collecting merchandise for my favourite sports team, Chelsea! Let’s chat about soccer :)

Office hours:
Thursday 12:30 – 2:30pm,
Rhodes 402

We are PoRTaL

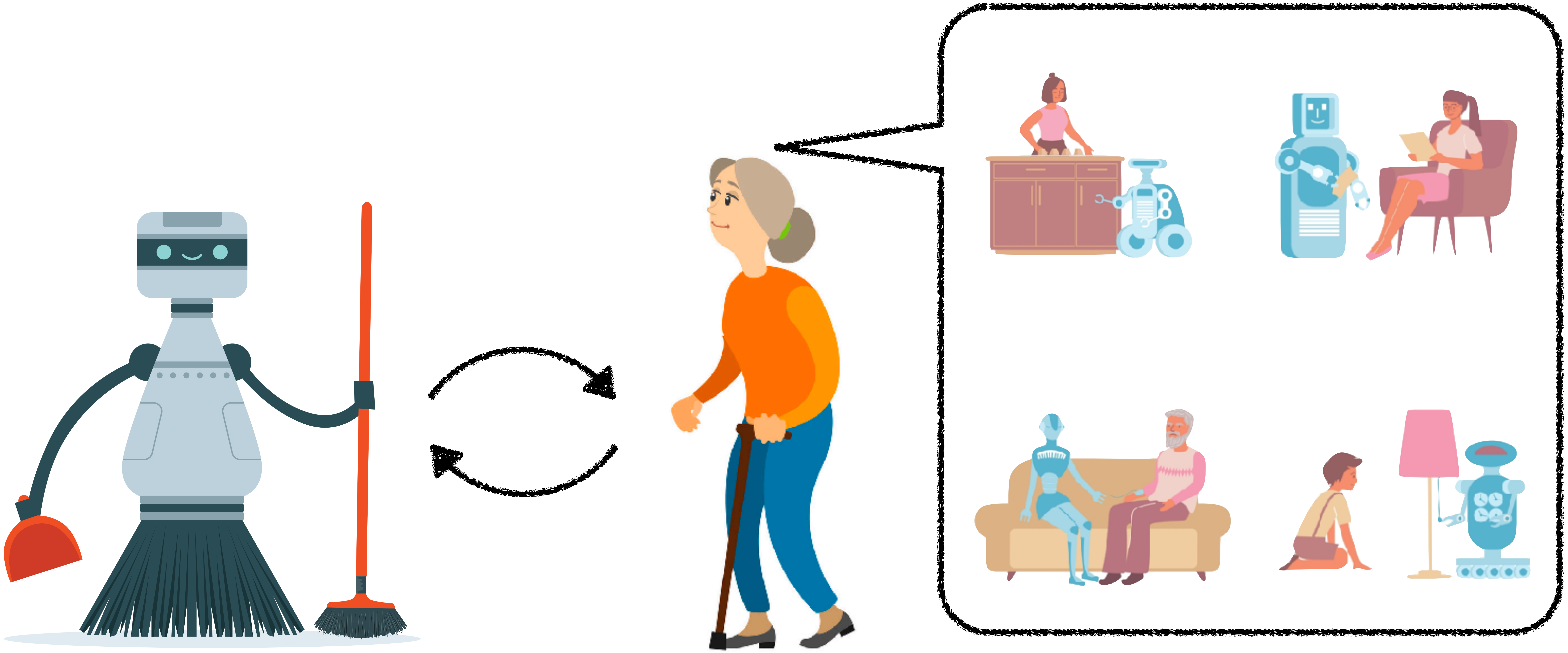
(People and Robots, Teaching and Learning)



https://portal.cs.cornell.edu/

PORTAL

Everyday Robots for Everyday Users



HAL & DORA

Helping Out In the Kitchen

HAL

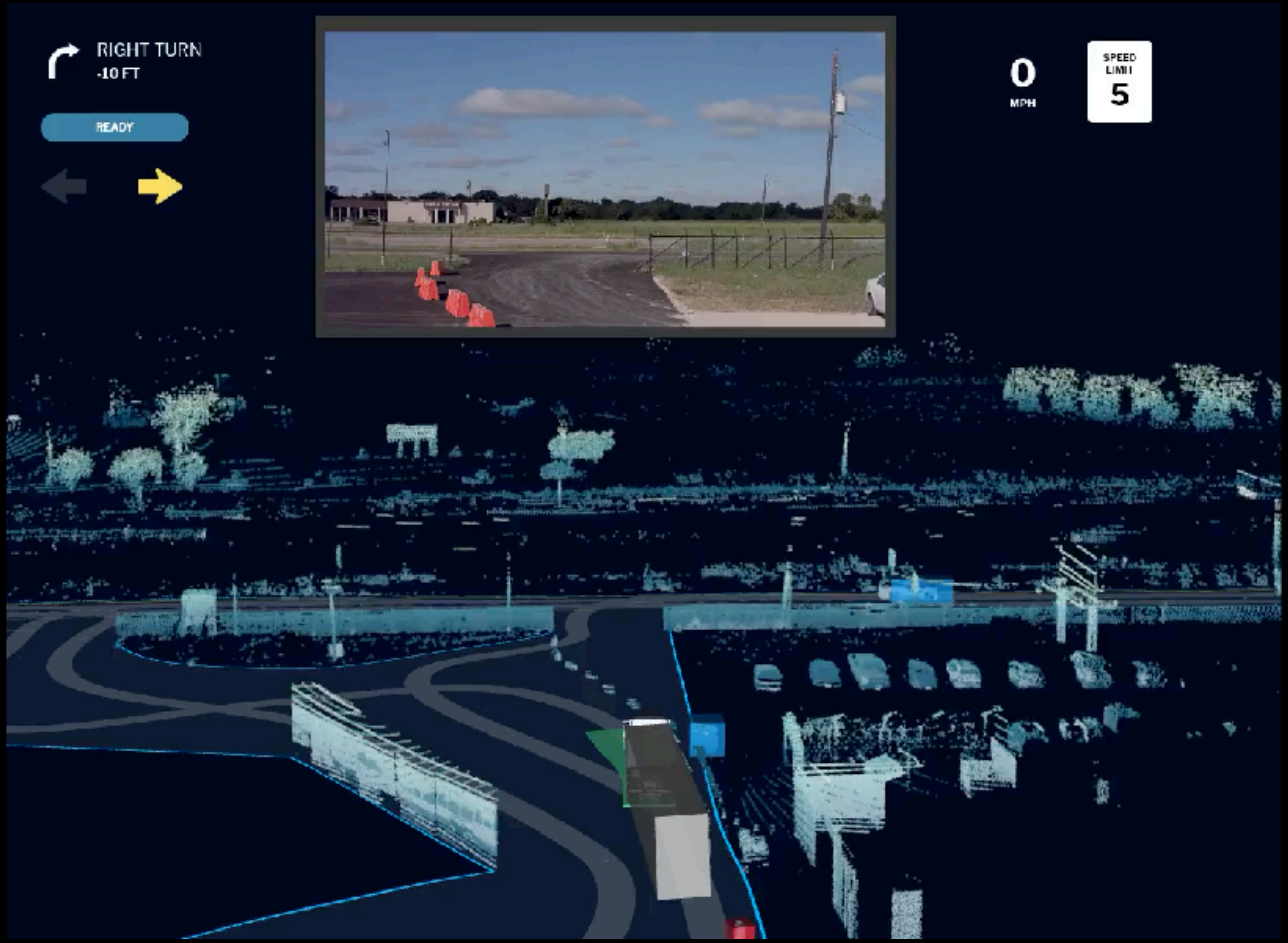
DORA

PORTAL

Let's get started!

How should robots learn to make good decisions?

Grounded in two “personal” applications



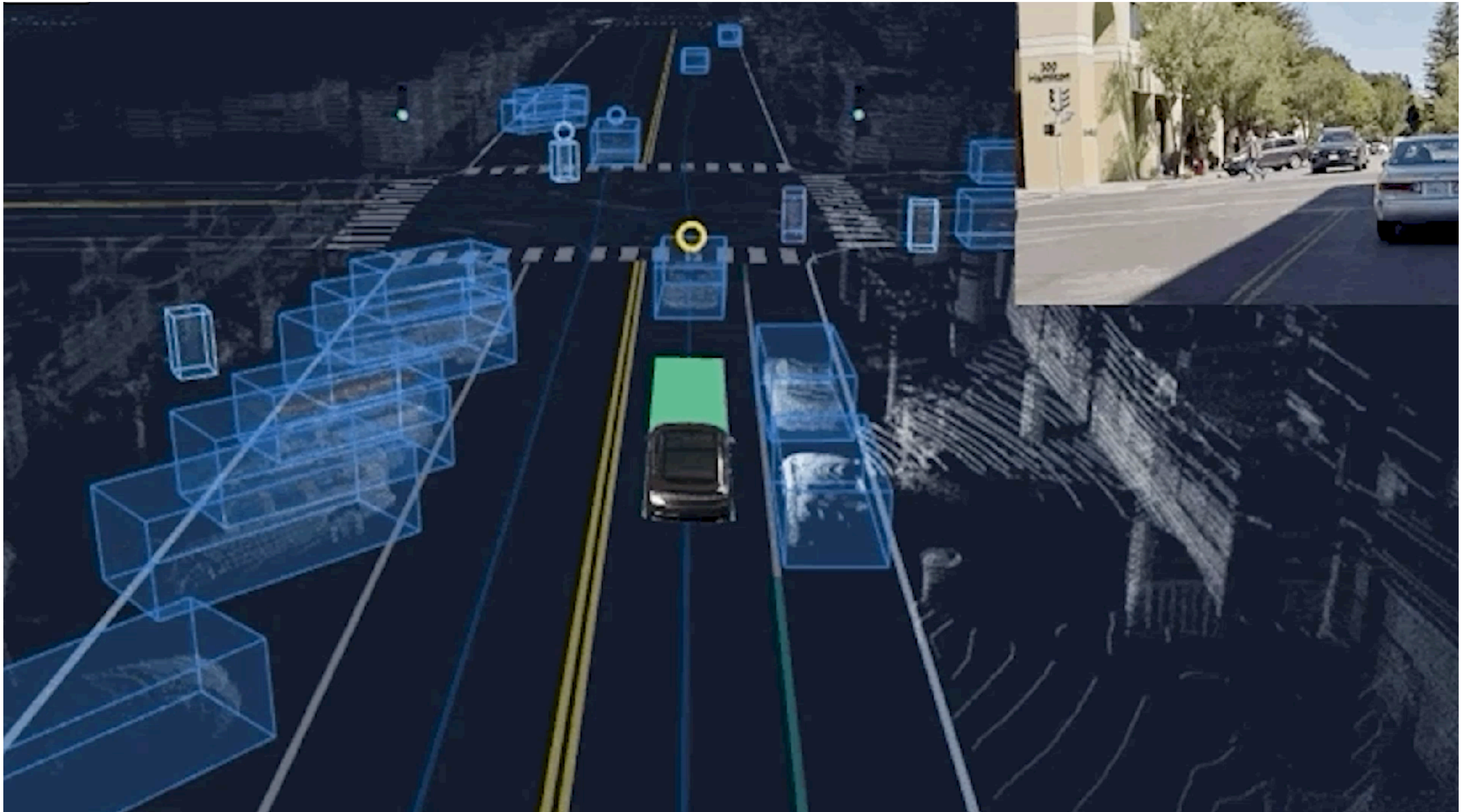
Self-driving

Home Robots

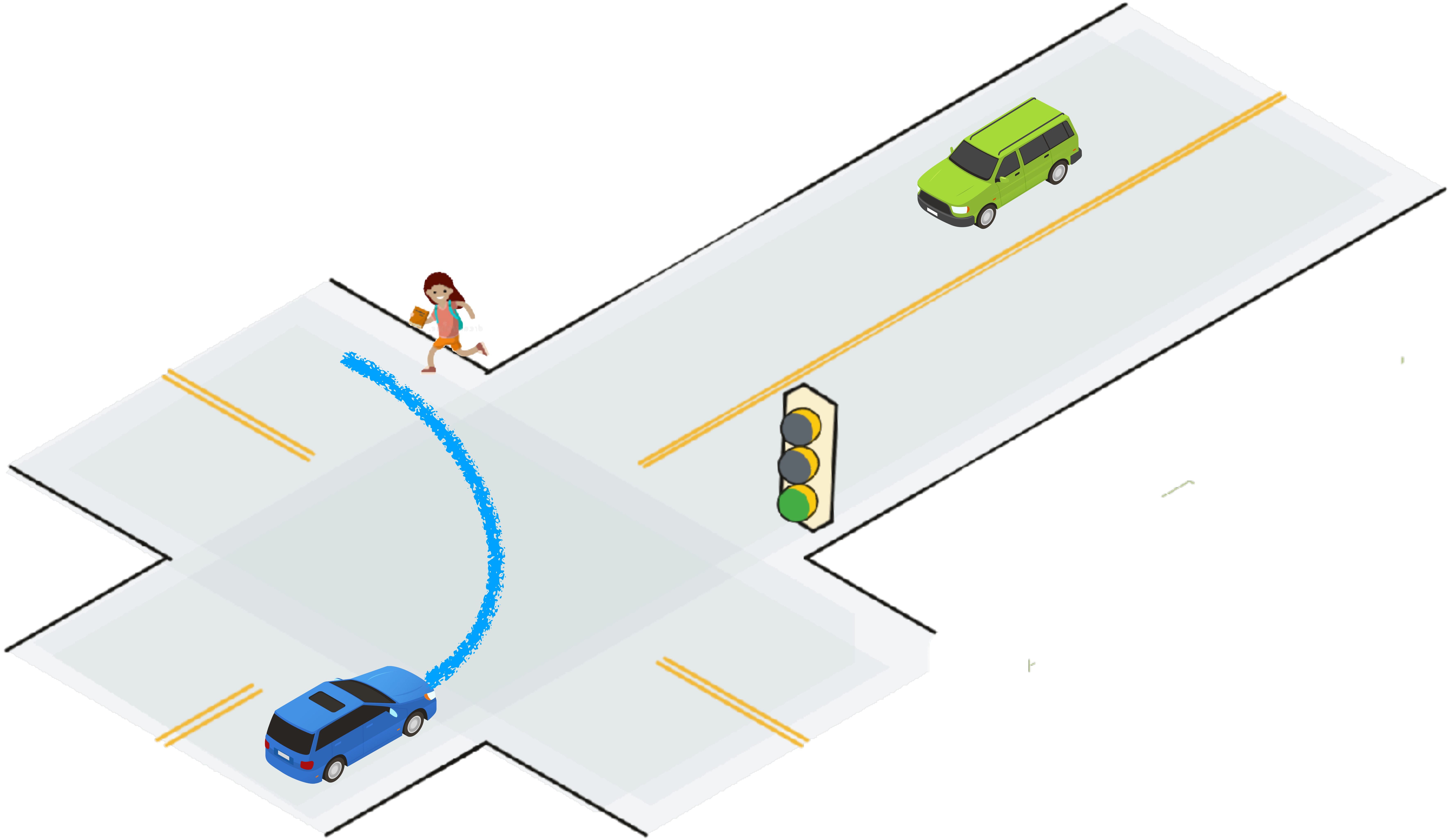
Self-Driving

Activity!

Brainstorm: What is "good" behavior in a left turn?



Brainstorm: What is "good" behavior in a left turn?

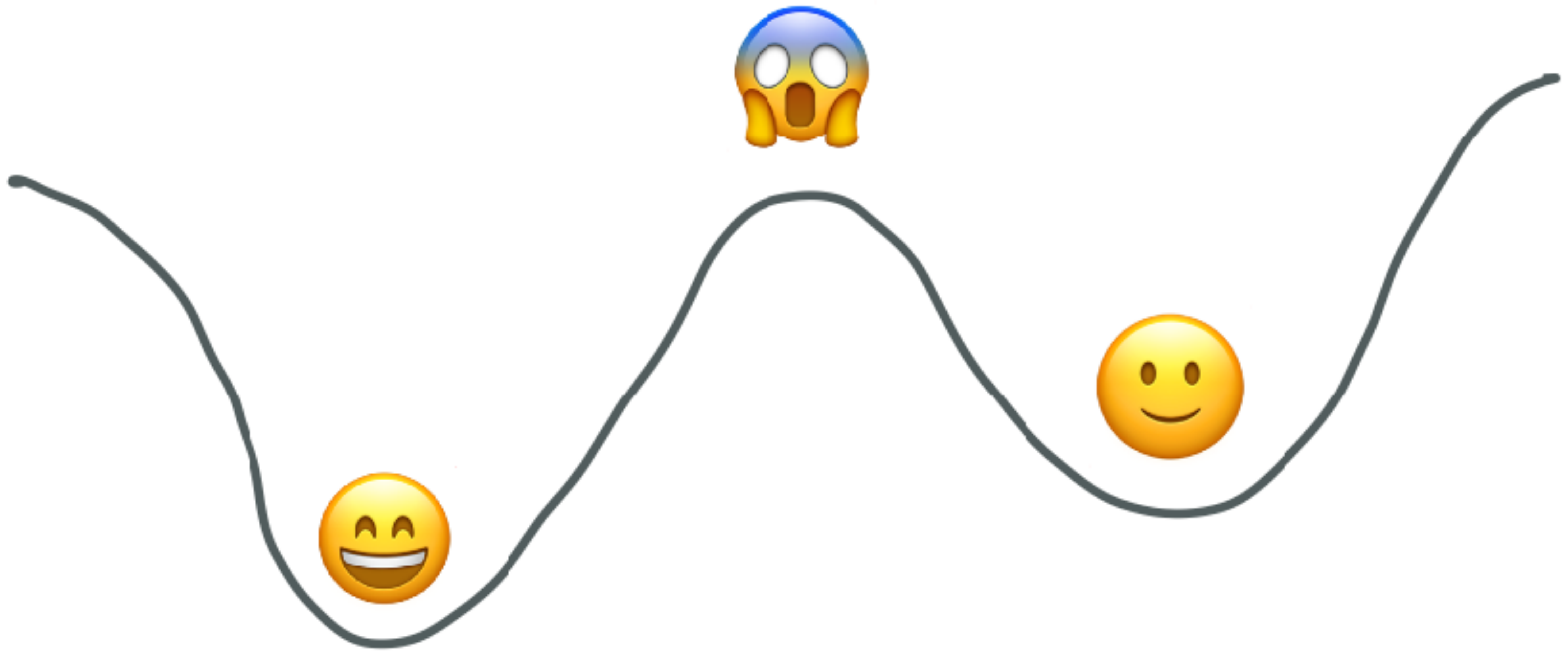


How should robots learn to make good decisions?

Three fundamental questions

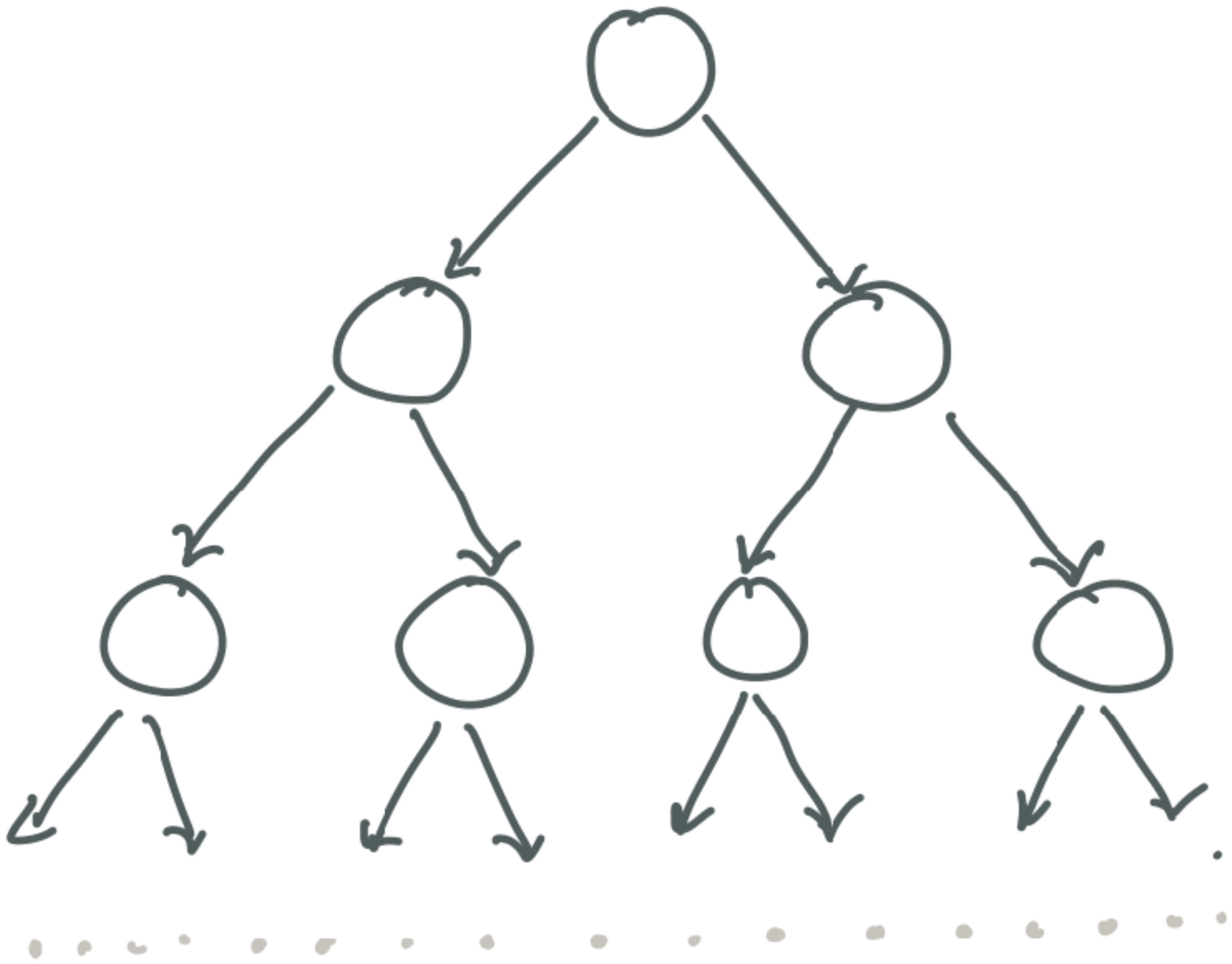
Values

What are good / bad states?



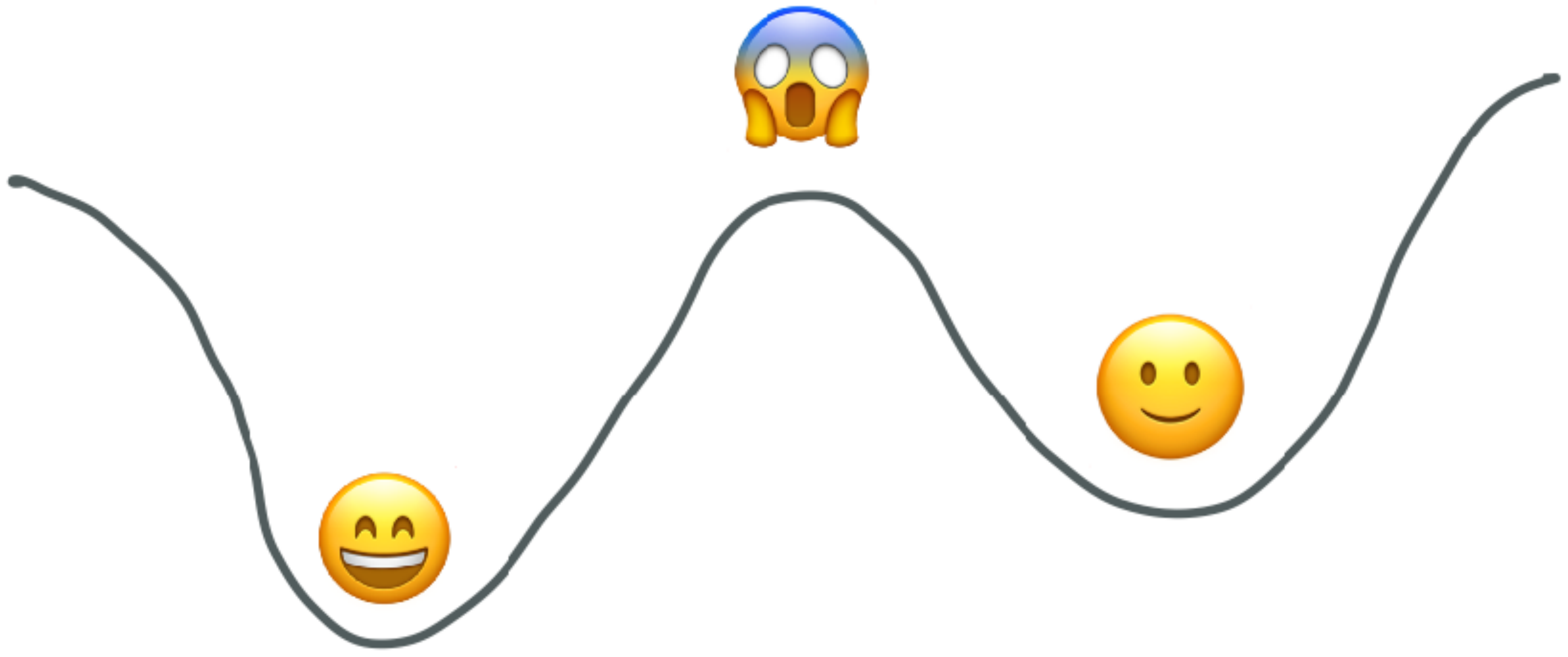
Models

How do decisions affect states?



Values

What are good / bad states?



Optimization

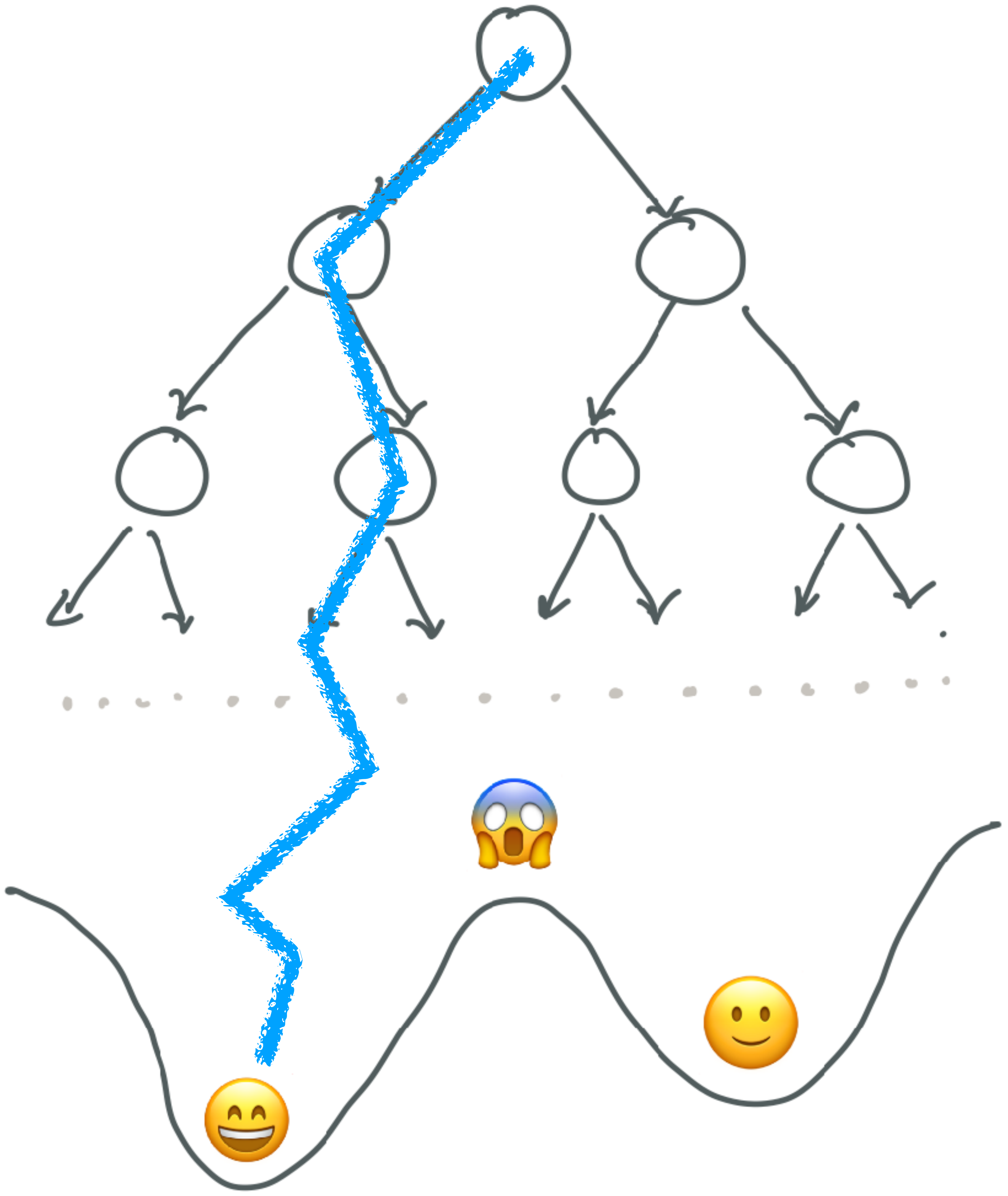
How do we efficiently find
the optimal sequence of decisions?

Models

How do decisions affect
states?

Values

What are good / bad
states?



Optimization

How do we efficiently find the optimal sequence of decisions?

Models

How do decisions affect states?

Values

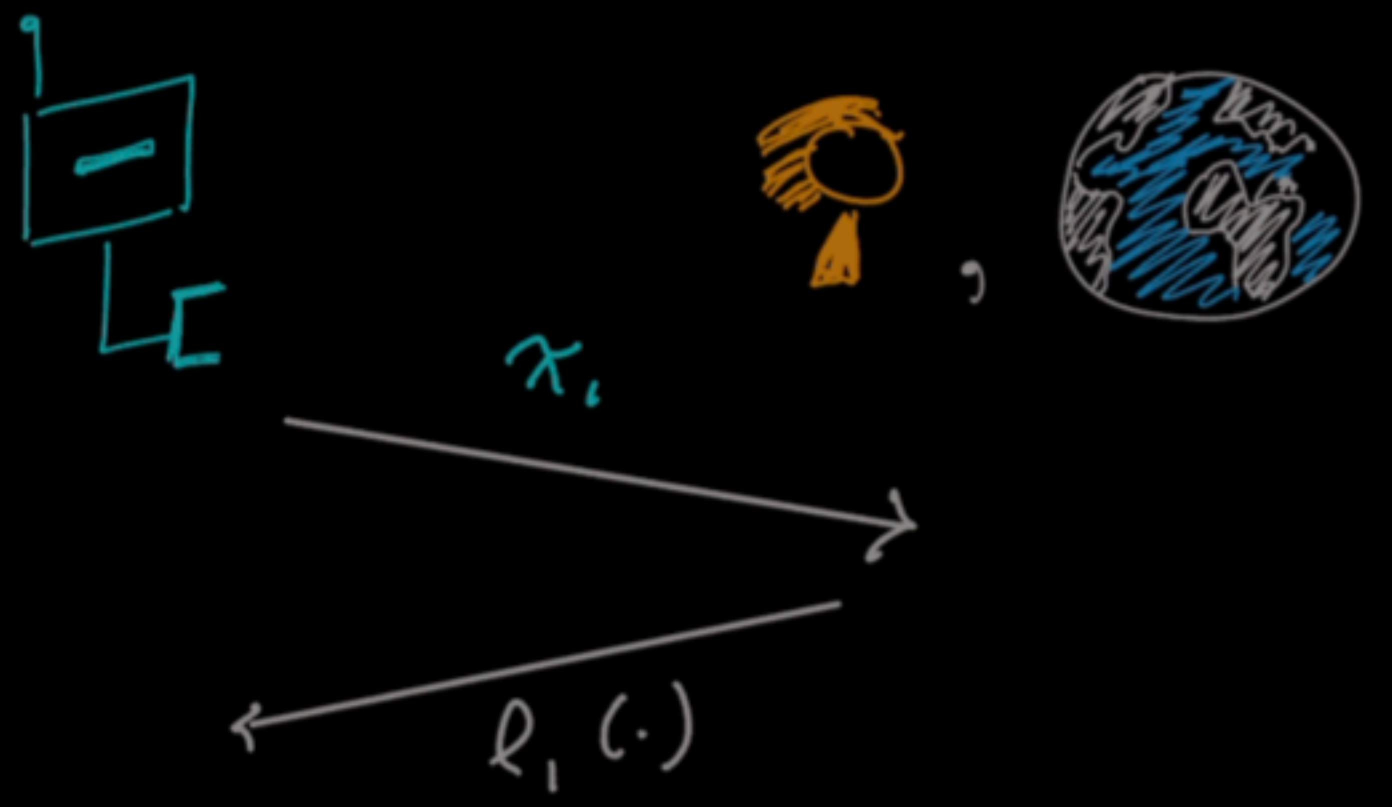
What are good / bad states?

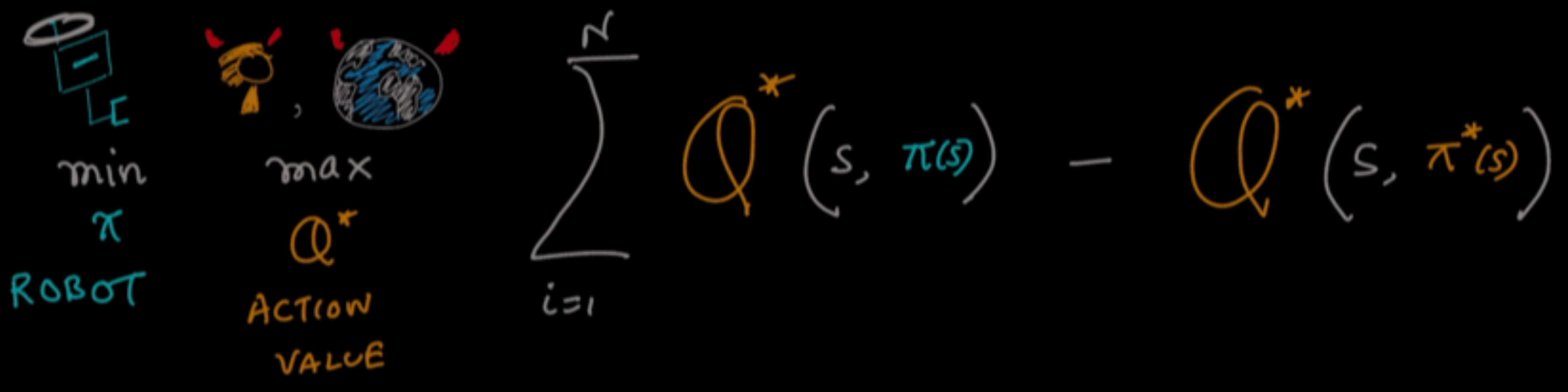
Learning

5 Levels

of

Robot Learning





Optimization

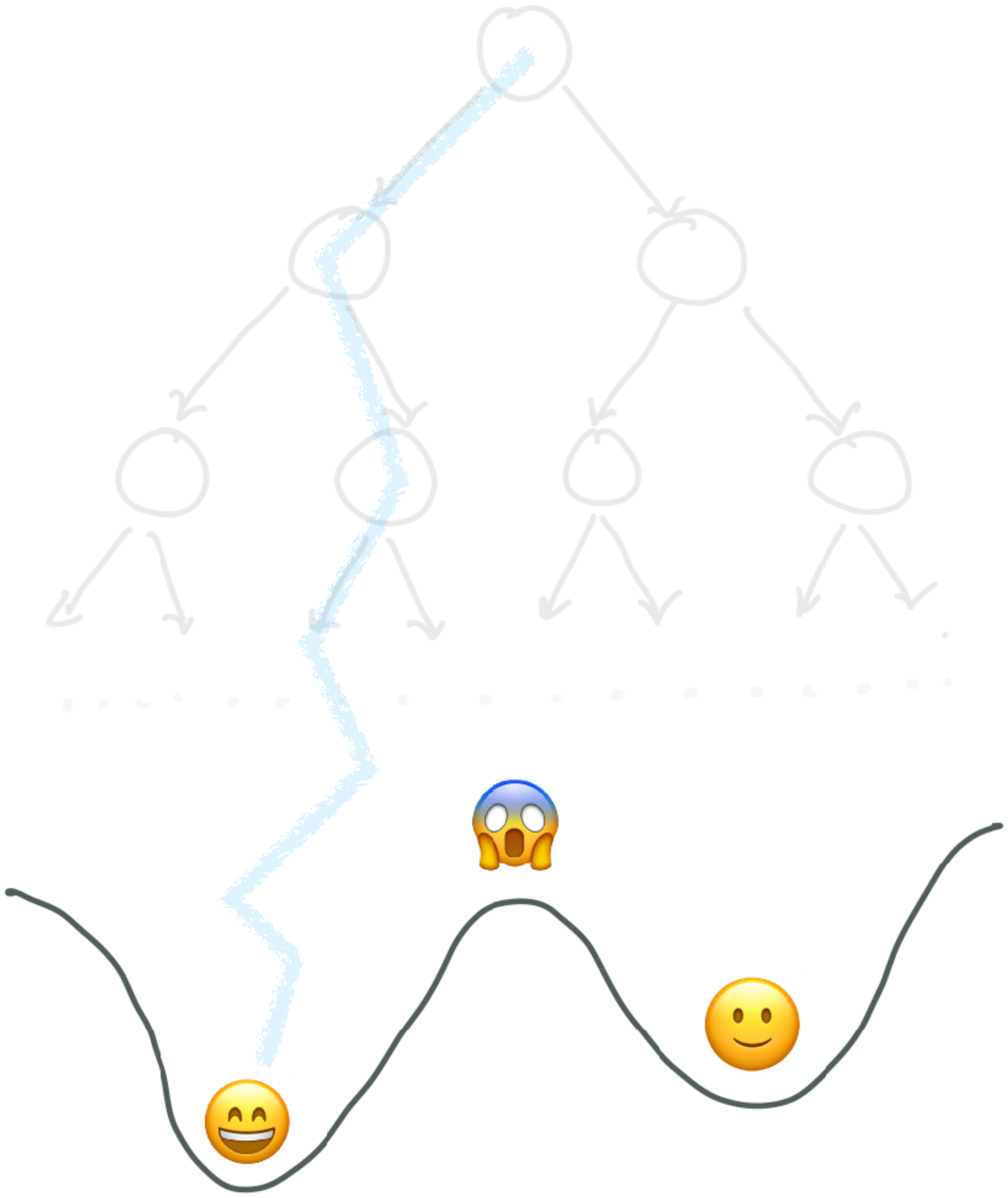
How do we efficiently find the optimal sequence of decisions?

Models

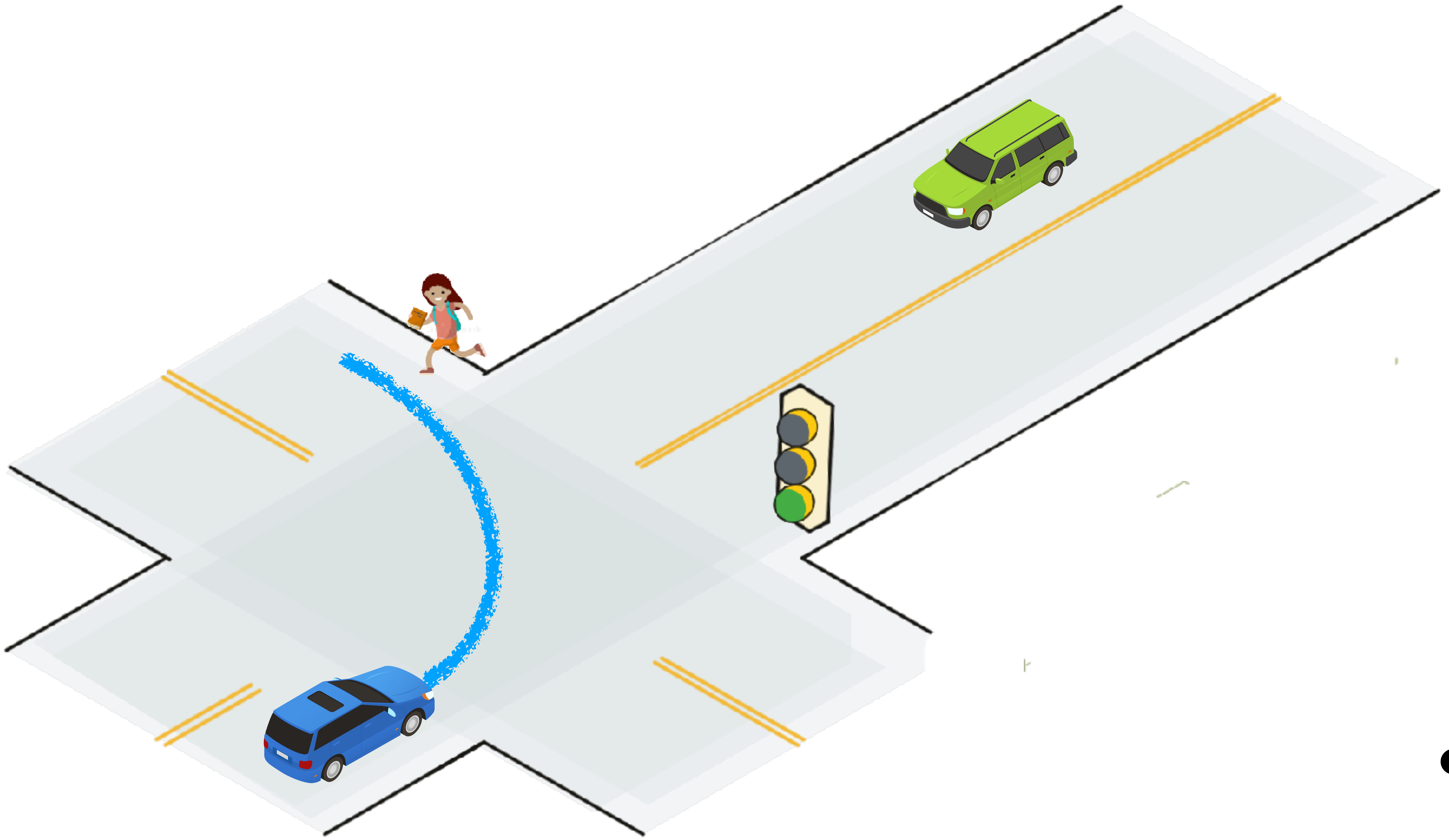
How do decisions affect states?

Values

What are good / bad states?



What are good / bad states?



Bad

- Collision
- Cutting off pedestrians
- Cutting off oncoming car
- Getting stuck in intersection when light turns red
- Excessive braking / braking speed limit

Good

- Completing the turn quickly

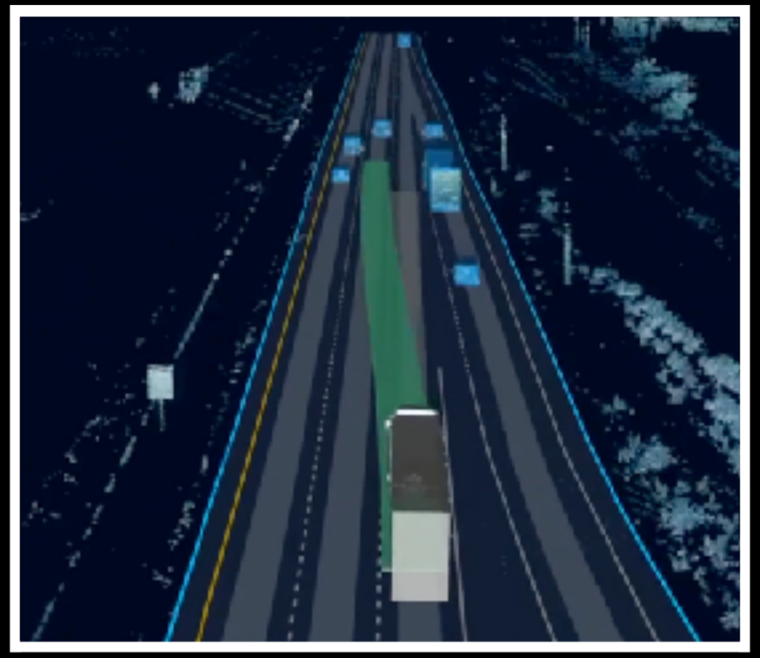
Question:

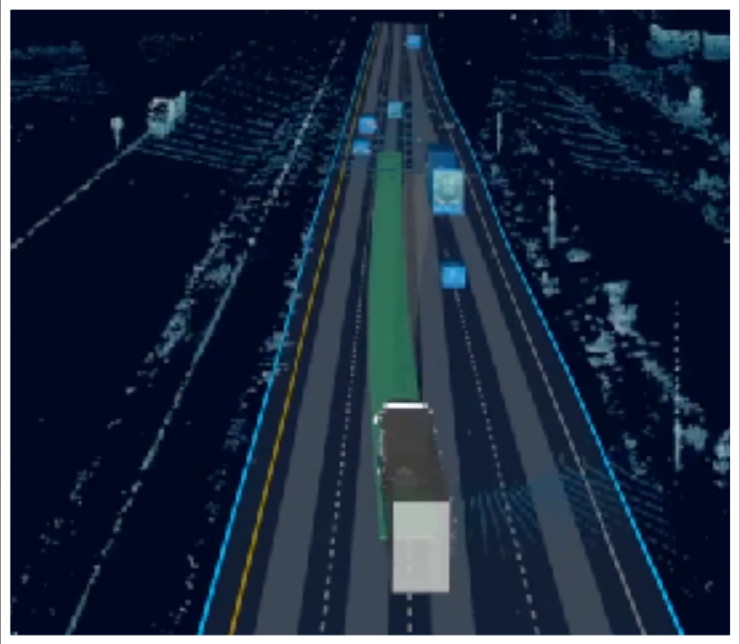
How do we program in these values?

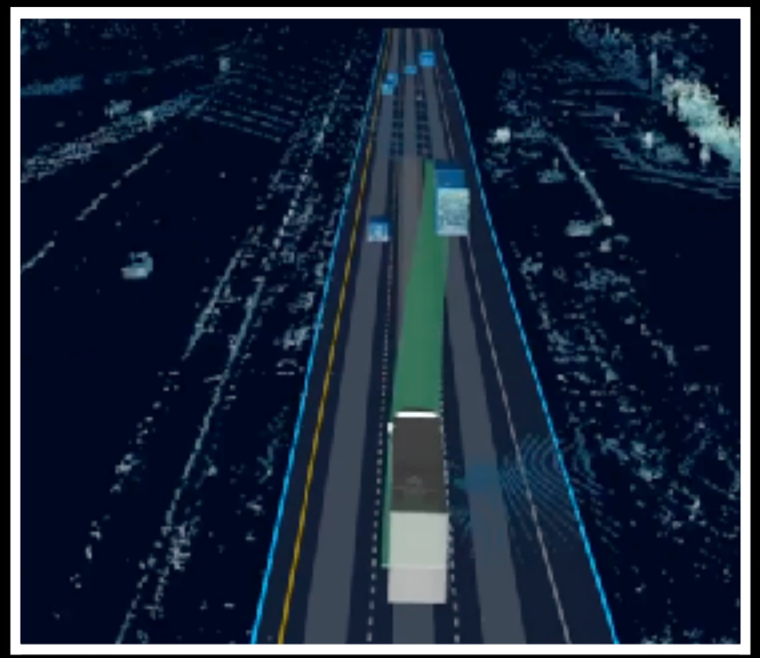
Why don’t we simply *imitate* good human driving?

SUPERVISED LEARNING

#1 Get Expert Data







Input (s) Output (a)

#2
Train
Policy

$$\pi : s \rightarrow a$$

#3 Deploy!


Train ≠ Test

Lesson #1

Feedback drives covariate shift

Models

How do decisions affect states?

Optimization

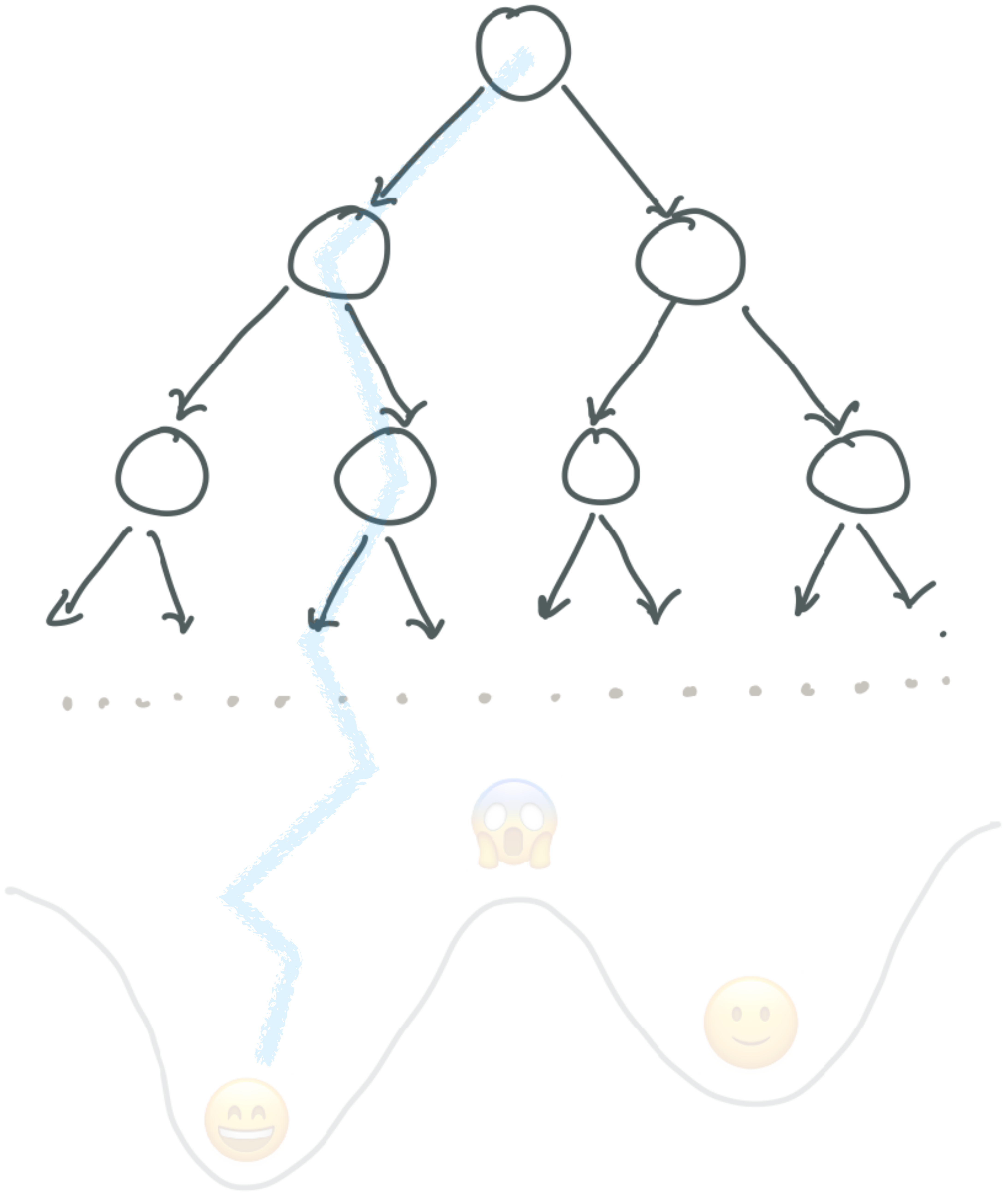
How do we efficiently find the optimal sequence of decisions?

Models

How do decisions affect states?

Values

What are good / bad states?



Activity!

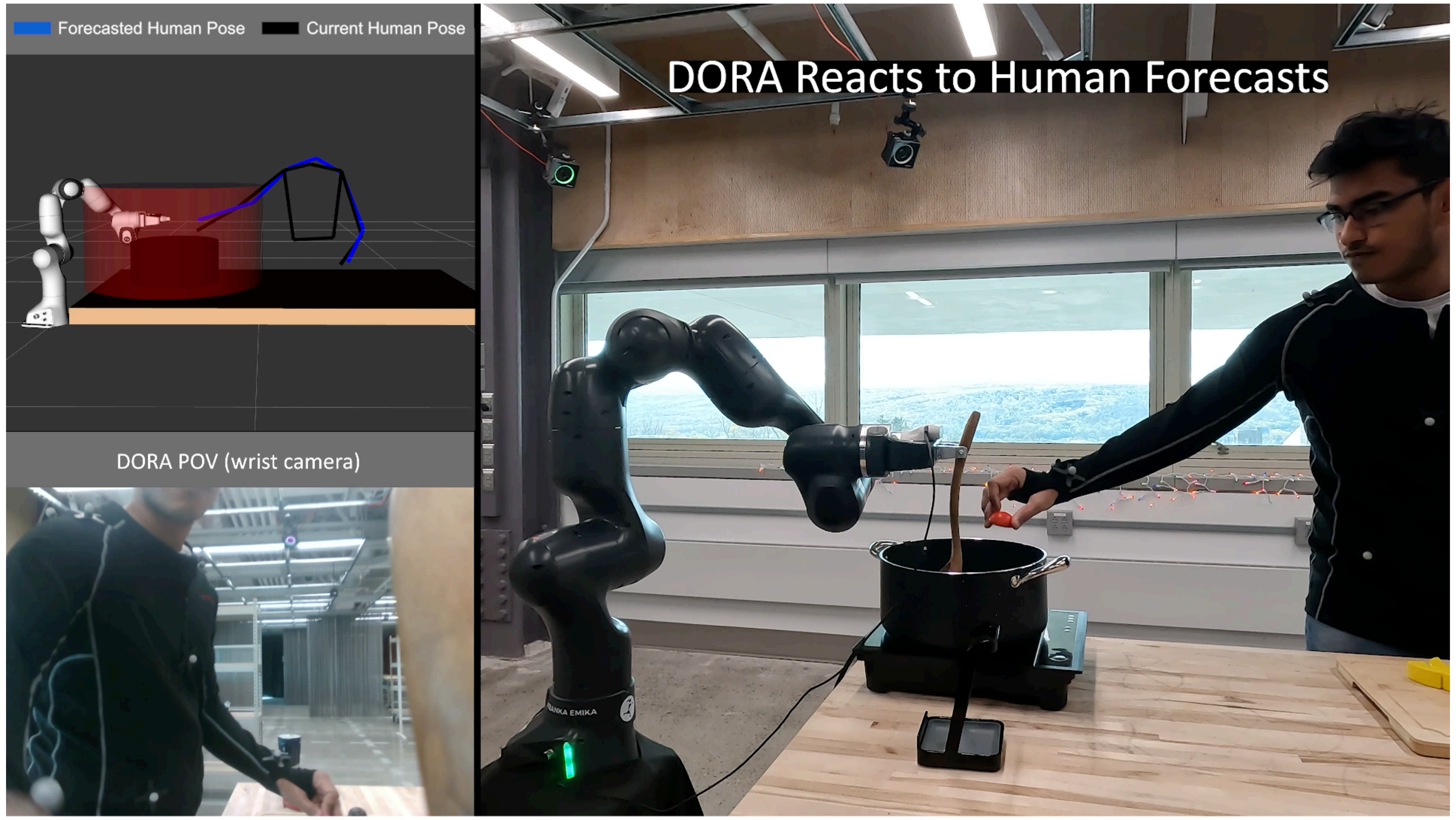

Forecasted Human Pose
Current Human Pose
DORA POV (wrist camera)
DORA Reacts to Human Forecasts
FRANKA EMIKA

Think-Pair-Share

Think (30 sec): How do we train a model of how humans move? Data? Model? Loss?

Pair: Find a partner

Share (45 sec): Partners exchange ideas



Lesson #2

Models are useful fictions

Optimization

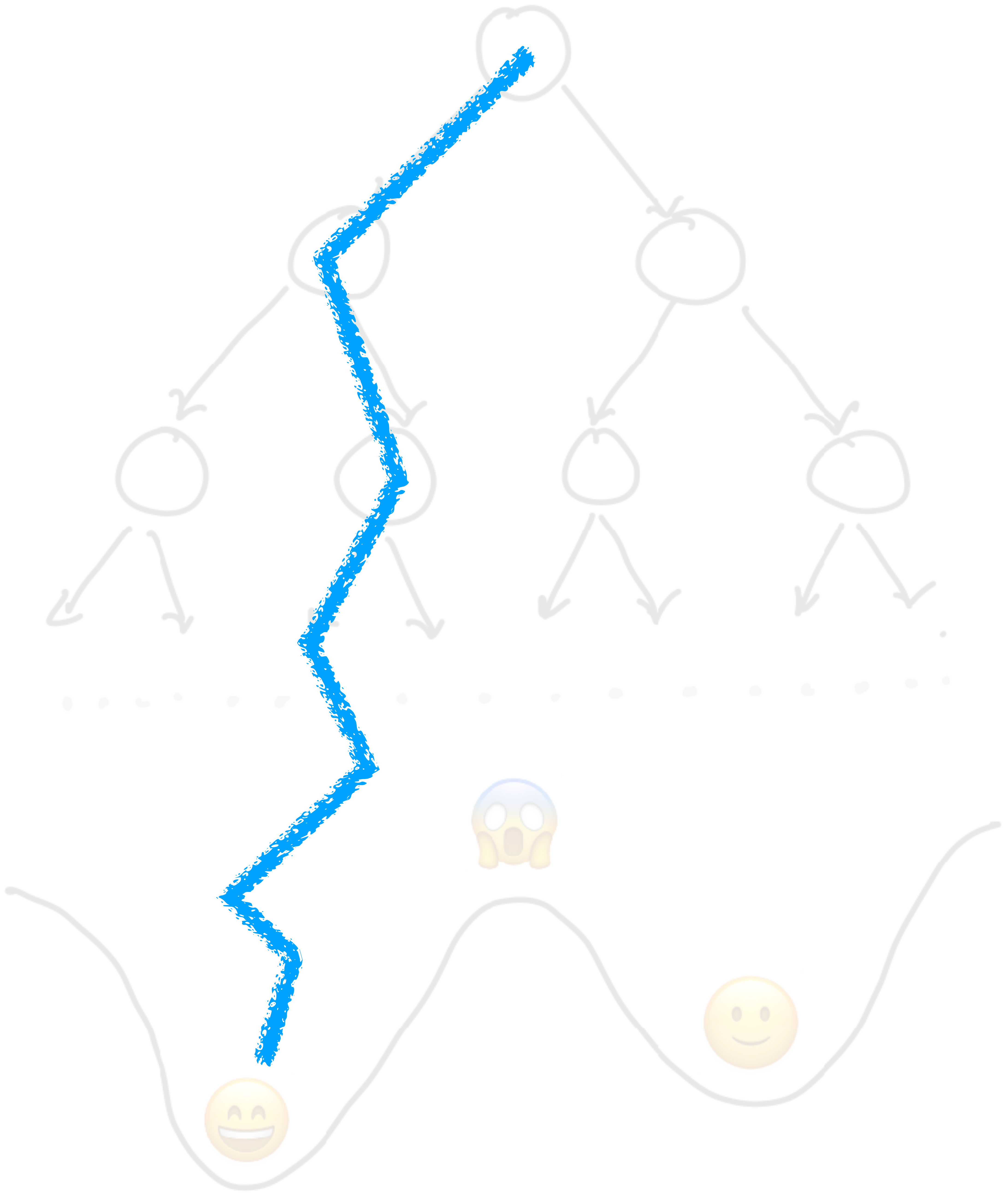
How do we efficiently find the optimal sequence of decisions?

Models

How do decisions affect states?

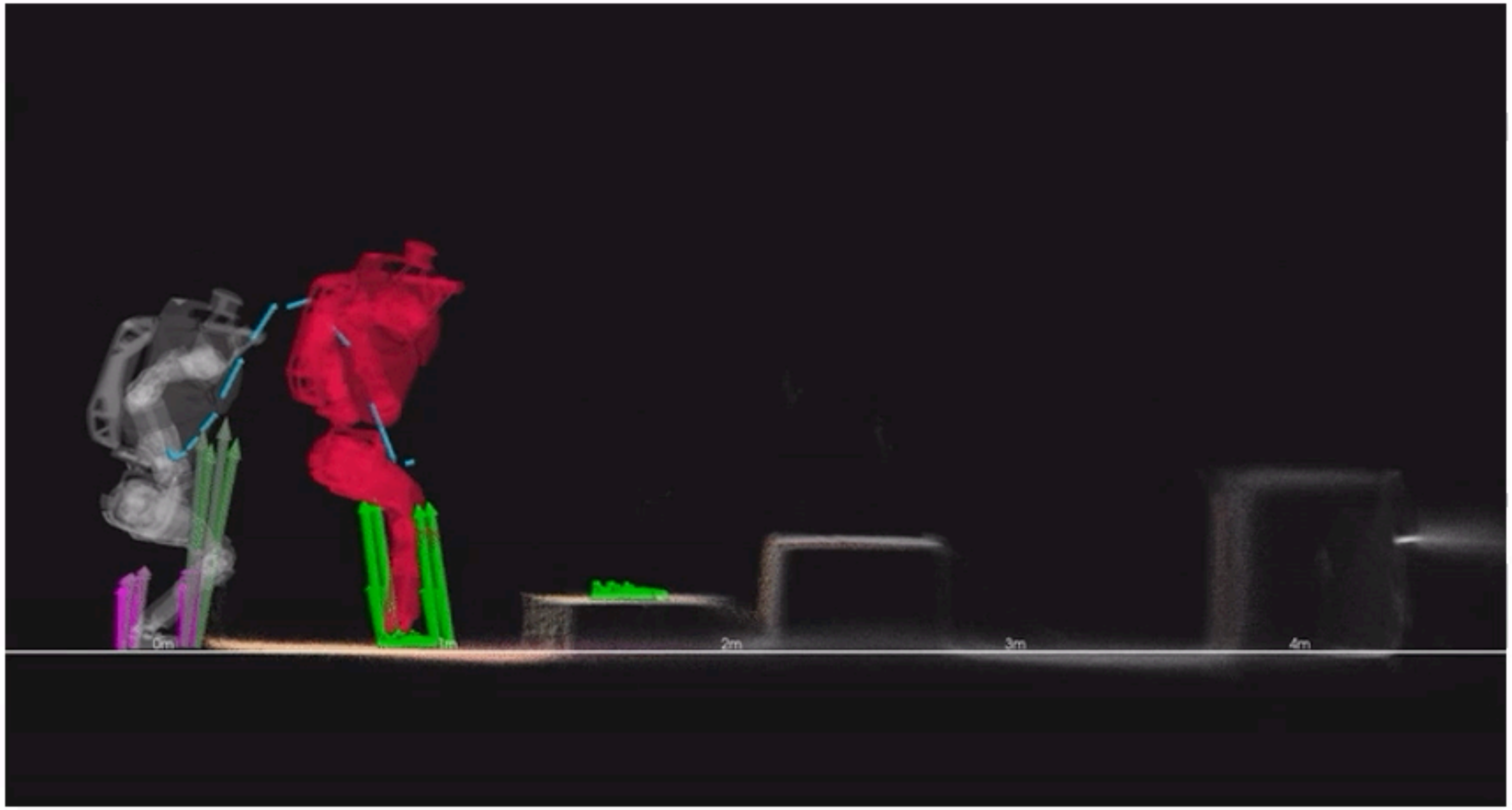
Values

What are good / bad states?



High-dimensional, continuous trajectory optimization

(With hard constraints!)

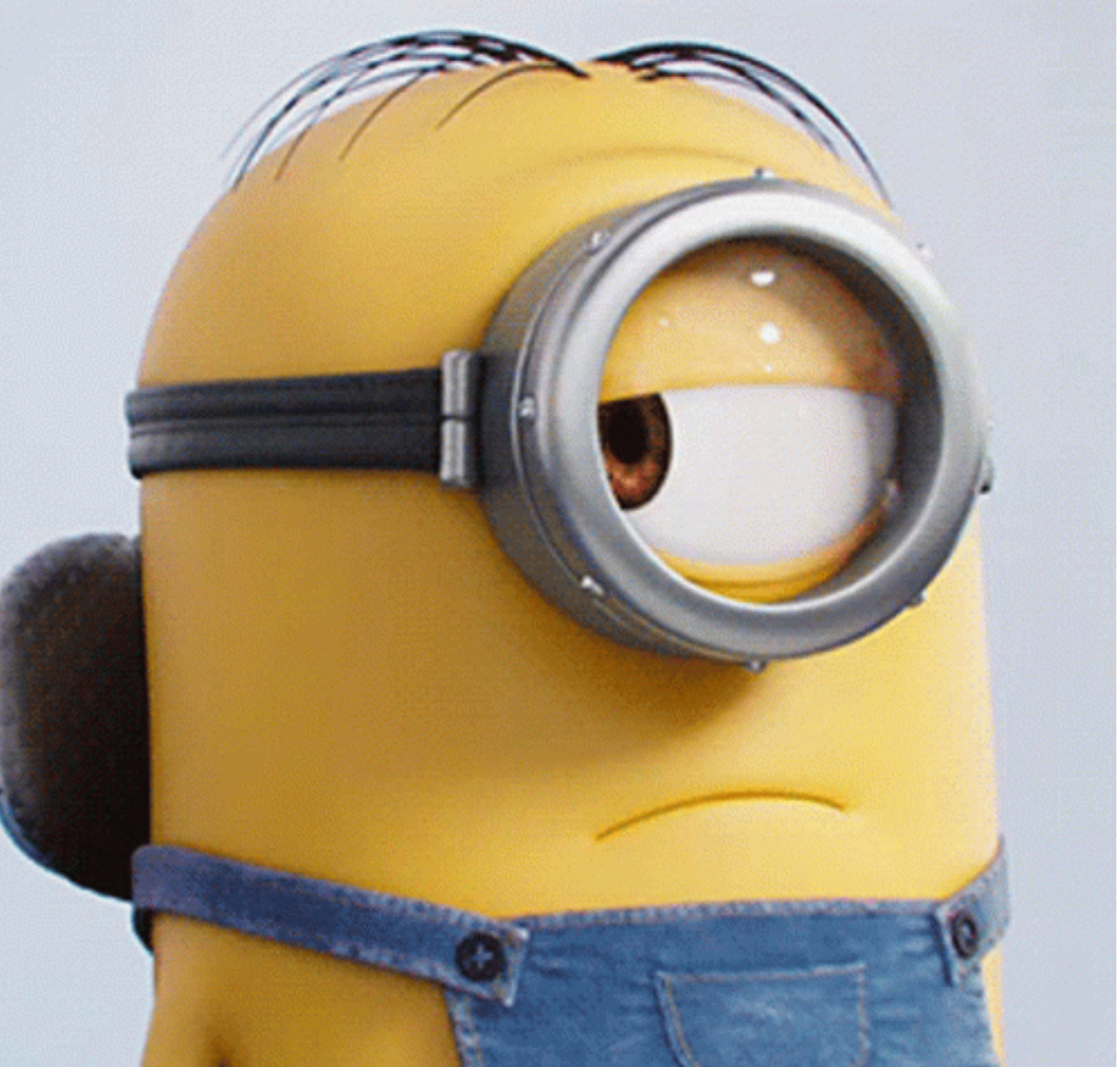


The journey ahead!

Schedule (Tentative)

Date	Lecture	Preread	Resources
08/22/23	Introduction: How should robots learn to make good decisions?		
08/24/23	Interactive Online Learning	Shai Shalev-Shwartz (Pg.108-111)	Arora et al. "Multiplicative Weights", Generalized Weighted Majority video
	Planning		
08/29/23	Markov Decision Process	MACRL Ch. 1	Dan Klein's slides I
08/31/23	Linear Quadratic Regulator: The Analytic MDP	MACRL (Ch 2, Pg. 23-27)	Underactuated robotics, Ch. 8, History of Optimal Control
09/05/23	Iterative Linear Quadratic Regulator	MACRL (Ch 2, Pg. 28-33)	iLQR paper , DDP for helicopter flight
09/07/23	Solving Hard MDPs: Constraints, Long Horizons, and more!	MACRL (Ch 4)	Gordon's notes on Lagrange, ALTRO: AuLa + iLQR,
	Imitation Learning		
09/12/23	Imitation Learning: Feedback and Covariate Shift (Assignment 2 Released)	MACRL (Ch 6, Pg. 53-57)	Three regimes of covariate shift
09/14/23	DAgger: A Reduction to No-Regret Learning	MACRL (Ch 6, full)	DAGGER , Agnostic SysId
09/19/23	Imitation Learning as Inferring Latent Expert Values		EIL, Youtube lec
09/21/23	Inverse Reinforcement Learning: From Maximum Margin to Maximum Entropy	MACRL (Ch 7)	LEARCH , MaxEntIRL , Youtube lec
09/26/23	Scaling Inverse Reinforcement Learning		Guided cost learning , f-divergence IL , Youtube lec
09/28/23	Imitation Learning: The Big Picture		Of Moments and Matching , Youtube lec

	Reinforcement Learning		
10/03/23	Approximate Dynamic Programming : Temporal Difference, Q-learning (Assignment 3 Released)	MACRL (Ch 8, full) , MACRL (Ch 9, full)	Sutton&Barto (Ch. 5, 6) , DQN , Rainbow DQN
10/05/23	Black-box vs White-box Policy Optimization	MACRL (Ch 10, full)	
10/12/23	Nightmares of Policy Optimization	MACRL (Ch 11, full)	
10/17/23	Actor Critic Methods		
10/19/23	Model-based Reinforcement Learning		
10/24/23	Dealing with Uncertainty (Extended Abstracts Due)		
	Frontiers		
10/26/23	Meta Learning		
10/31/23	Offline Reinforcement Learning		
11/02/23	Diffusion Models and Imitation Learning		
11/07/23	Large Language Models and Task Planning		
11/09/23	Multi-agent Forecasting and Imitation Learning		
11/14/23	Learning Visual Representations from Ego Videos		
11/16/23	Visuomotor Skill Learning		
11/21/23	Causal Representation Learning		
11/28/23	Project presentations		
11/30/23	Project presentations		
12/04/23	Recap Lecture		



Logistics

Logistics

Website: https://www.cs.cornell.edu/courses/cs6756/2023fa/

Lectures

Interactive lectures, please read assigned book chapters / papers

Assignments [3 assignments * 15% grade = 45%]

Programming heavy. HW2, HW3 involve PyTorch. Done individually!

Project [45%]

Final project. Pick a research problem, apply techniques from class. Be creative!

Groups of 2. Extended abstract, final presentation, final paper.

Participation [10%]

Activities in lectures, in class polls

Book!

https://macrl-book.github.io/

Modern Adaptive Control and Reinforcement Learning,
James A. Bagnell, Byron Boots, and Sanjiban Choudhury

(Please send me feedback)

Expectations

Assignments are programming / ML heavy!

For final project: Check if you have the resources to train models

Familiar with ML concepts and modern tools (transformers, CNNs).
Also familiarity with linear algebra (SVD etc)

Please pre-read book chapters / pre-watch supplementary video

Which course should you take?

CS 4756 / 5756 (Spring!)

More focus on fundamentals, also covers robot perception.
Good course to take before taking this one.

CS 6756 (this class)

Builds on CS 4756, goes deeper in decision making,
assumes familiarity with ML tools and concepts

Generative AI

The work you do for CS 6756 consists of writing code and natural language descriptions.

To some extent, the new crop of "generative AI" (GAI) tools can do both of these things for you.

However, **we require that the vast majority of the intellectual work must be originated by you,** not by GAI. You may use GAI to look up helper functions, or to proofread your text, but clearly document how you used it.

Generative AI

In this class, for every assignment and final project, you can choose between two options:

Option 1: Avoid all GAI tools. Disable GitHub Copilot in your editor, do not ask chatbots any questions related to the assignment, etc. If you choose this option, you have nothing more to do.

Option 2: Use GAI tools with caution and include a one-paragraph description of everything you used them for along with your writeup. This paragraph must:

1. Link to exactly which tools you used and describe how you used each of them, for which parts of the work.
2. Give at least one concrete example (e.g., generated code or Q&A output) that you think is particularly illustrative of the "help" you got from the tool.
3. Describe any times when the tool was unhelpful, especially if it was wrong in a particularly hilarious way.
4. Conclude with your current opinion about the strengths and weaknesses of the tools you used for real-world compiler implementation.

Remember that you can pick whether to use GAI tools for every assignment, so using them on one set of tasks doesn't mean you have to keep using them forever.

Assignment 0

Simple survey

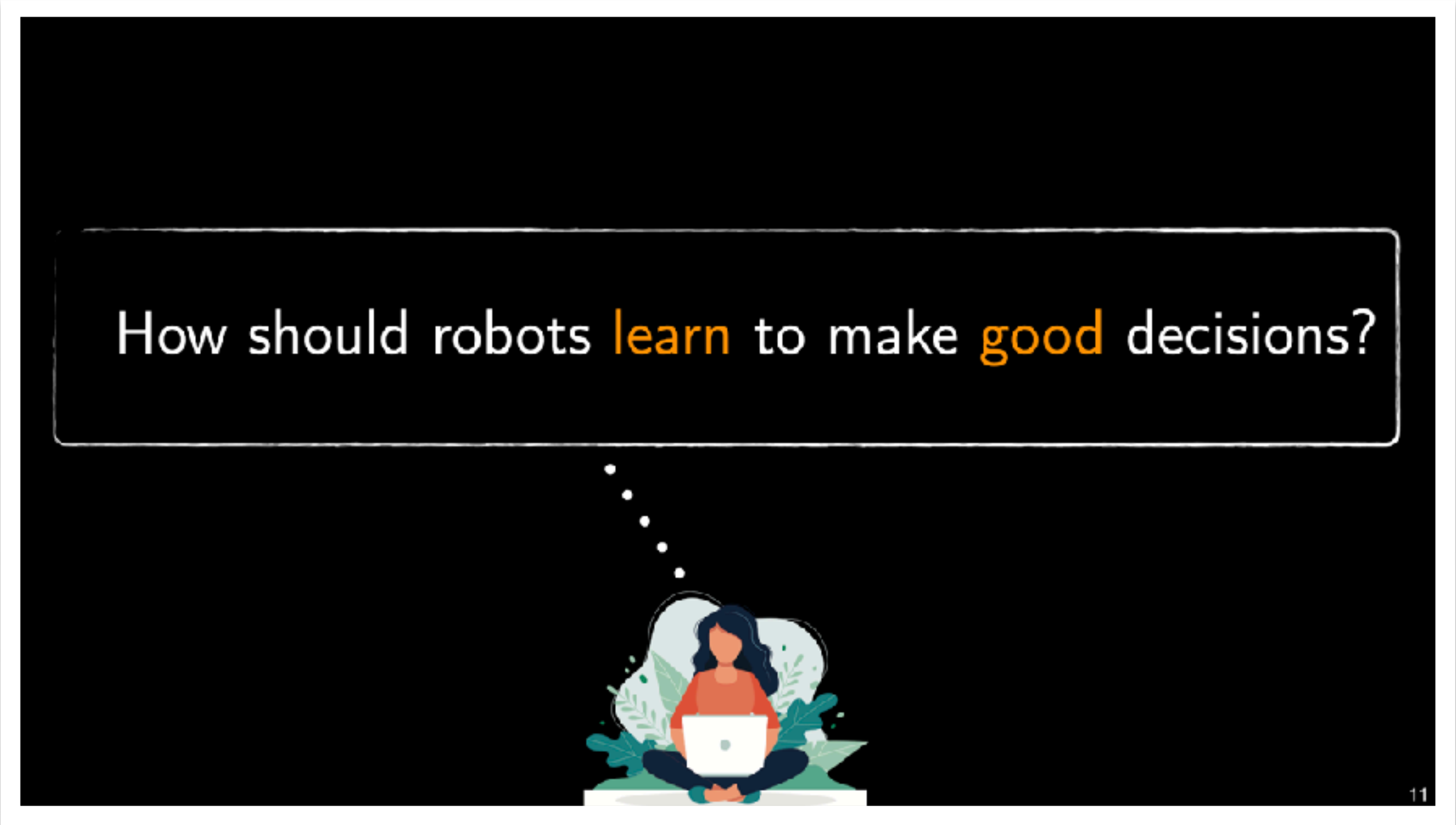
Link in the website:
https://docs.google.com/forms/d/e/1FAIpQLSdMMyzh1o3hA0j4S1sDEjnN91vjcxRBPaeTaoVww82apMCDyg/viewform?usp=sf_link

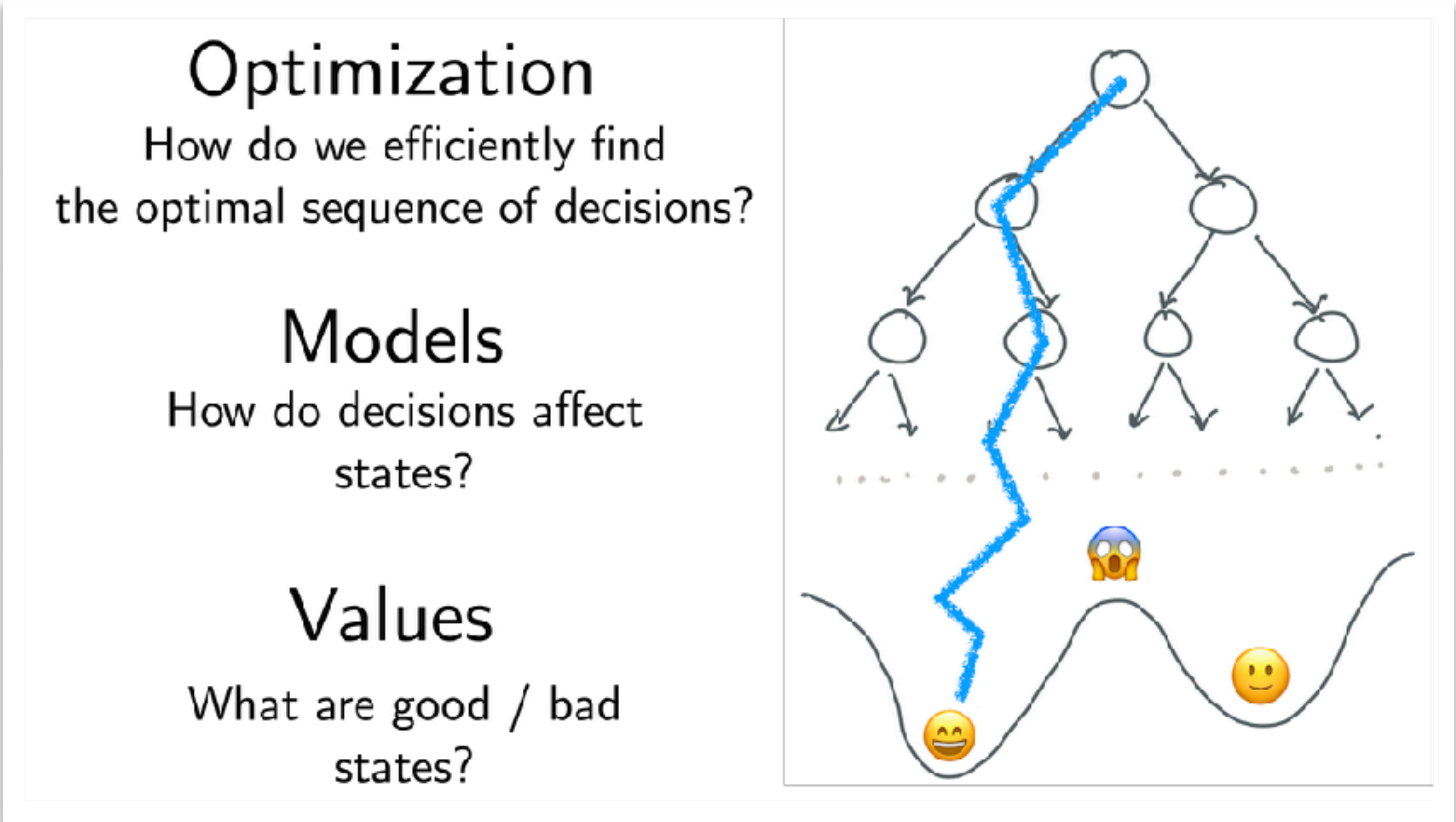
Questions?

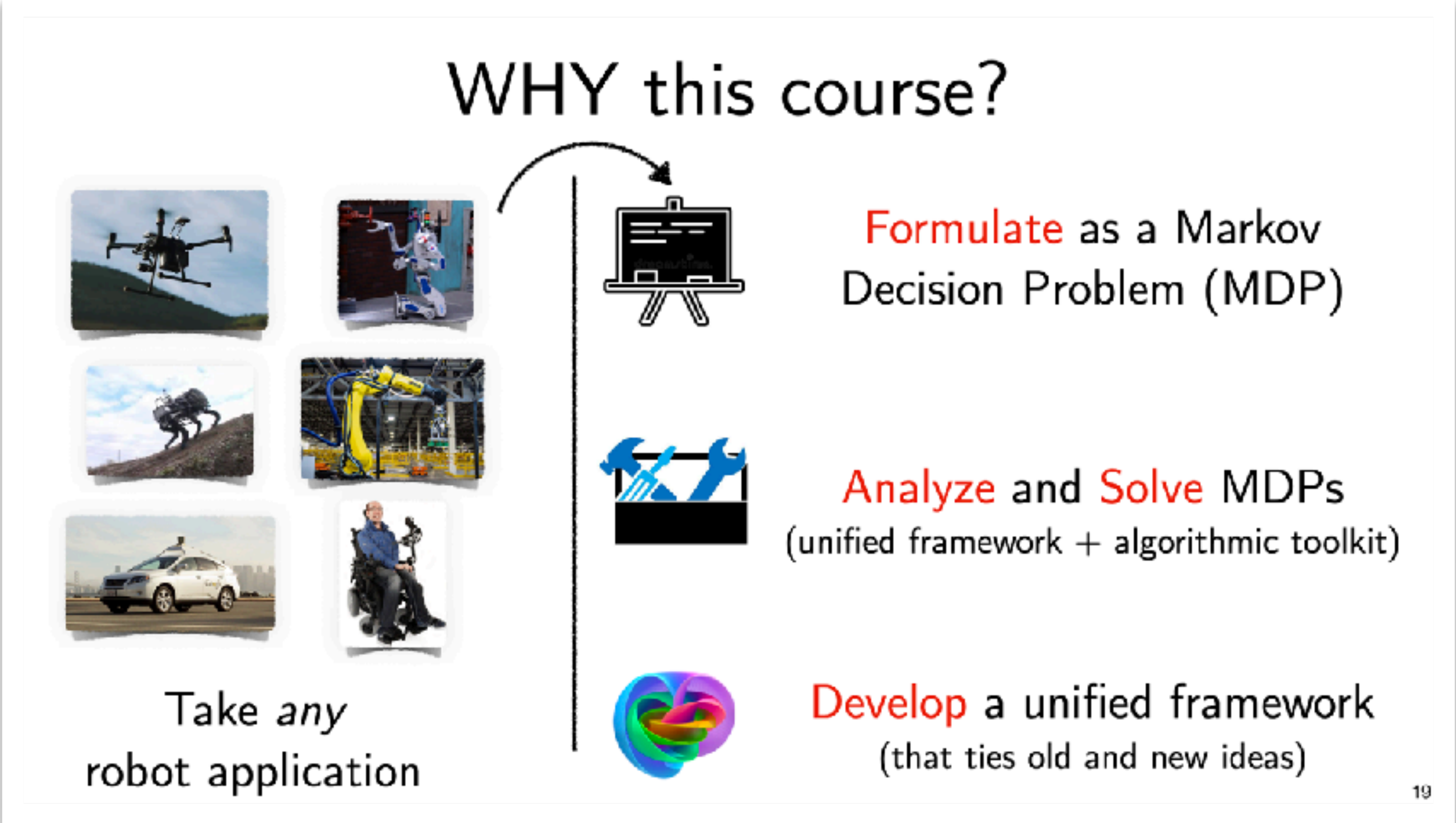
PINs issued every Tues / Thursday

Plenty of space in class so waitlist should get cleared

tl;dr


How should robots learn to make good decisions?
11


Optimization
How do we efficiently find the optimal sequence of decisions?
Models
How do decisions affect states?
Values
What are good / bad states?


WHY this course?
Take *any* robot application
Formulate as a Markov Decision Problem (MDP)
Analyze and Solve MDPs
(unified framework + algorithmic toolkit)
Develop a unified framework
(that ties old and new ideas)
19