CS 6756: Learning for Robot Decision Making

Sanjiban Choudhury

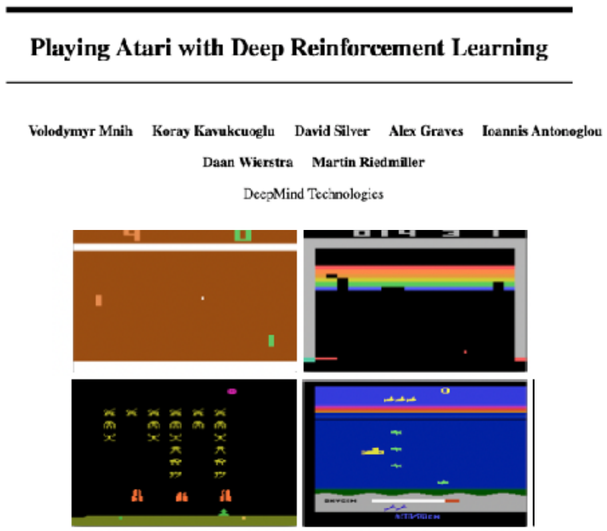
PORTAL

Cornell Bowers CIS Computer Science


WHAT A TIME TO BE ALIVE!

Exciting time for Artificial Intelligence

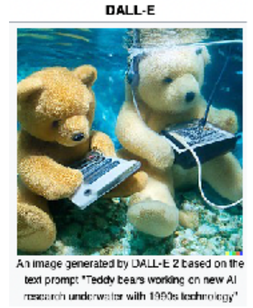
Deep Q Networks

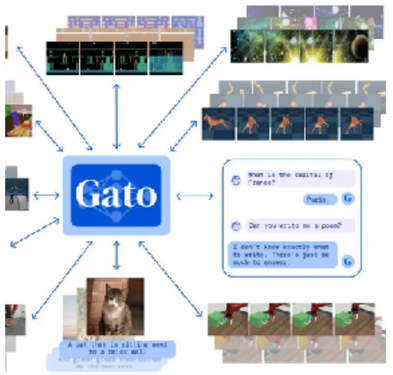


AlphaGo

Transformers

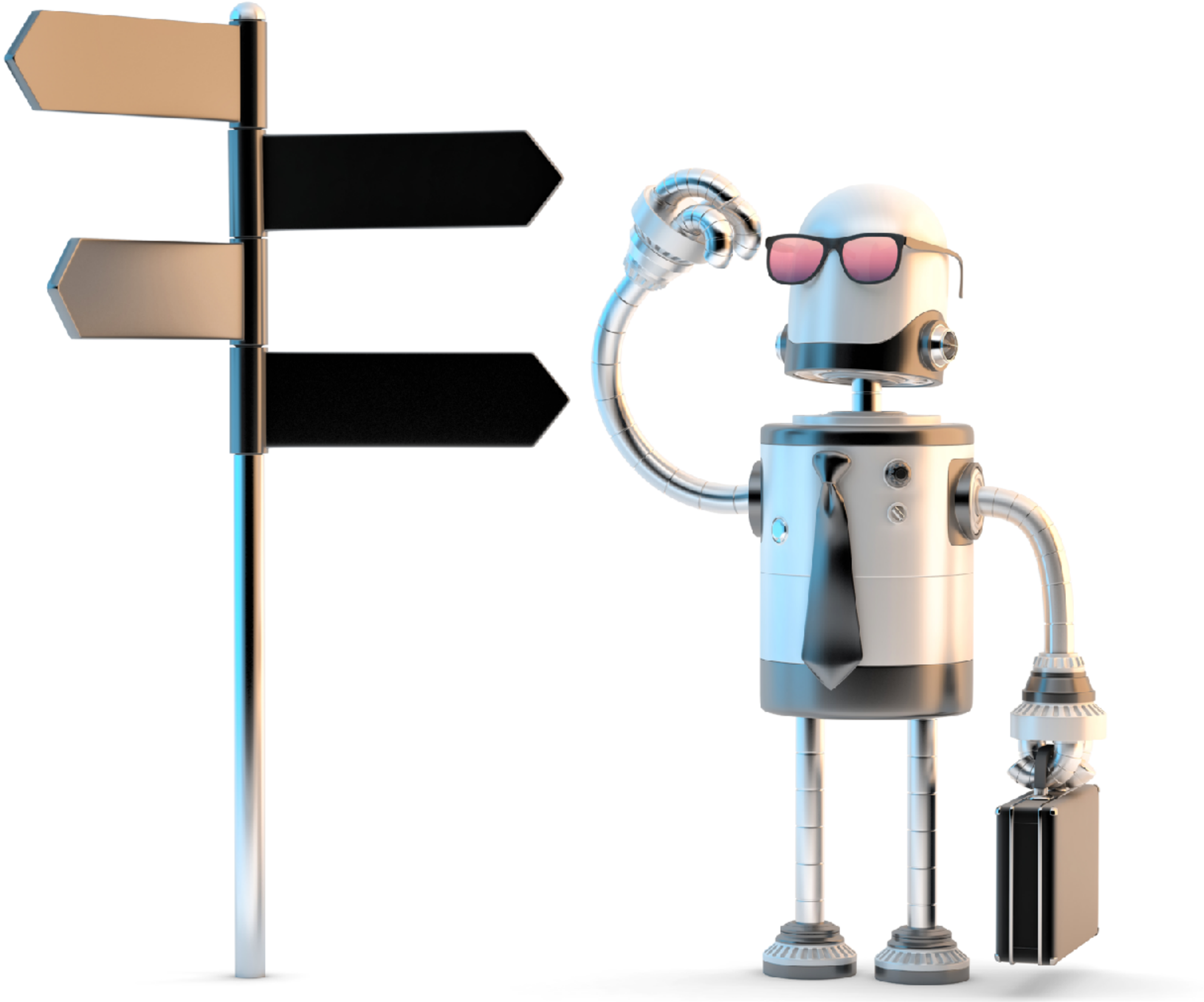
GPT3



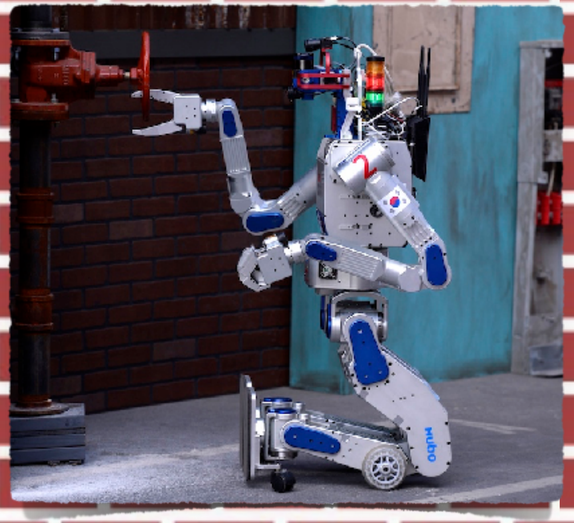


2013 2016 Today

What about robotics?

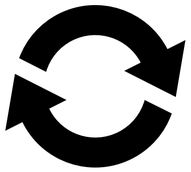


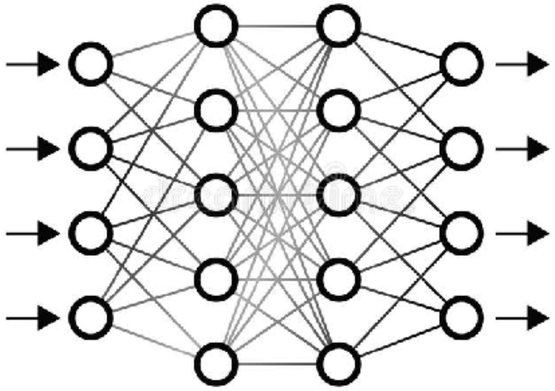
Robotics 1.0: Building things brick by brick



Robotics 2.0: Scale and improve with data

Formulate as a learning problem



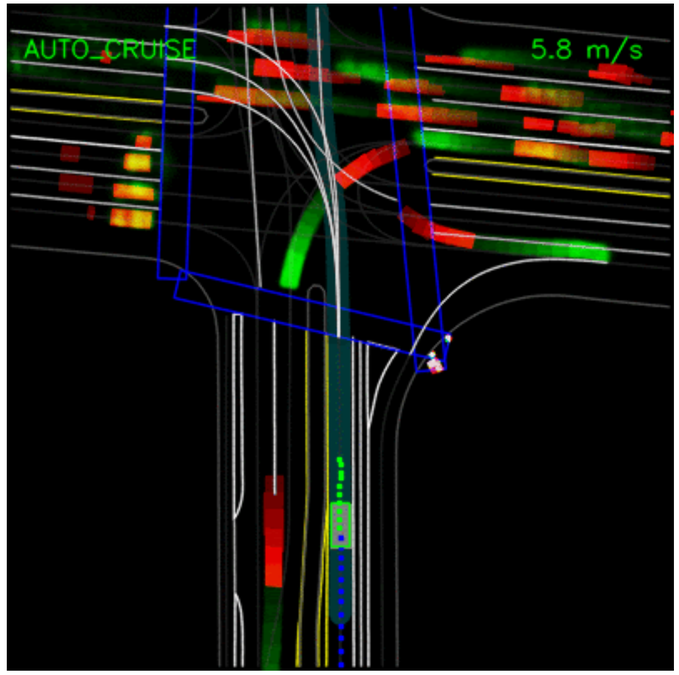


Invest in good ML pipelines

Robotics 2.0: Scale and improve with data

Formulate as a learning problem

Invest in good ML pipelines



Self-driving led the way!

Sanjiban Choudhury

He / Him

Carnegie Mellon
THE ROBOTICS INSTITUTE

PhD

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

Postdoc

Aurora

Startup
(that went public!)

How should robots learn to make good decisions?

Love the ocean!

And traveling with my partner

How should robots learn to make good decisions?

WHY this course?






Take *any*
robot application

Formulate as a Markov Decision Problem (MDP)

Solve MDPs using an all-purpose toolkit

(Imitation/Reinforcement learning, Model based/free)

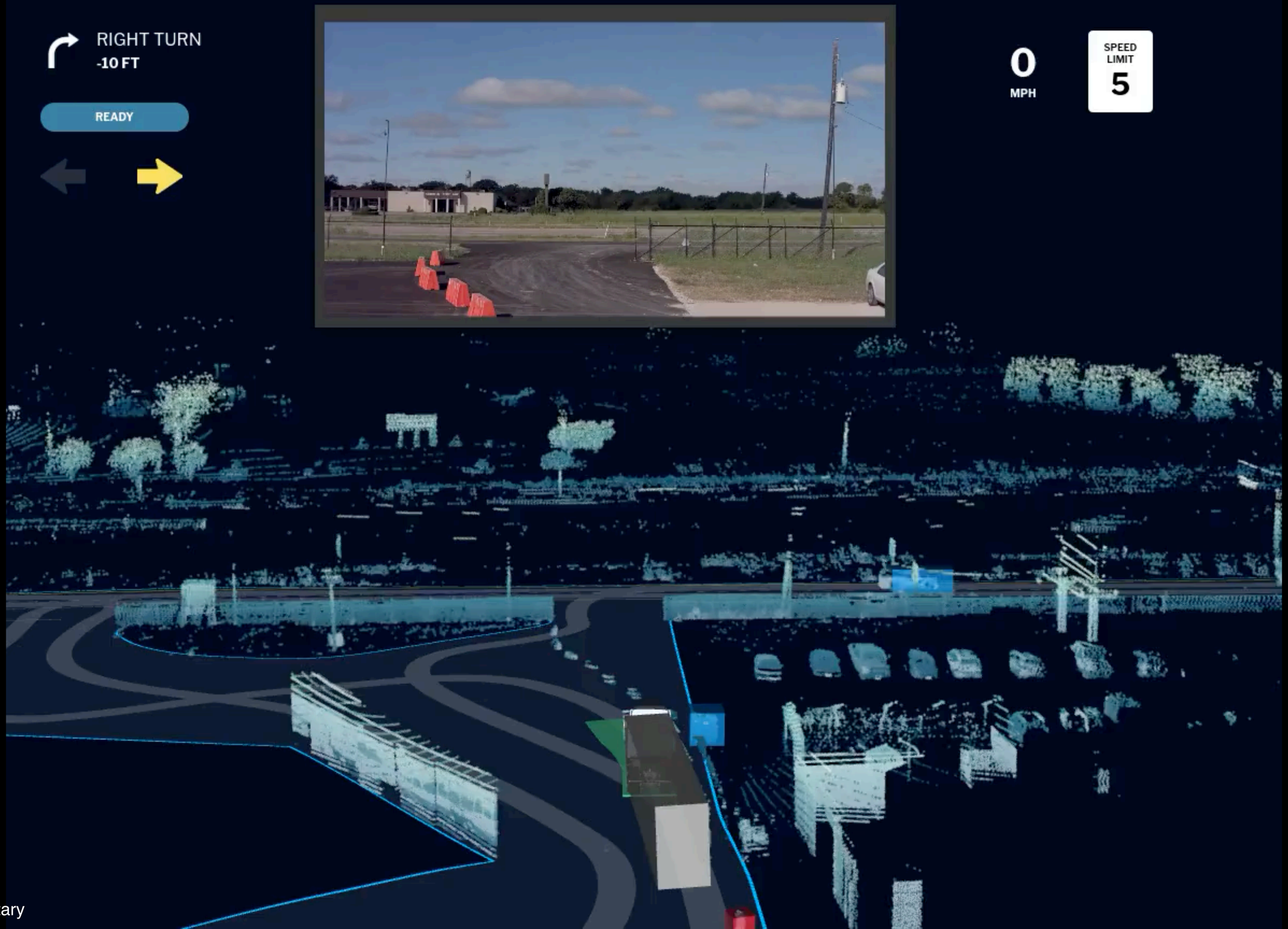
Deploy learners in real-world

(Safety, distribution shift, value alignment)

Belonging

How should robots learn to make good decisions?

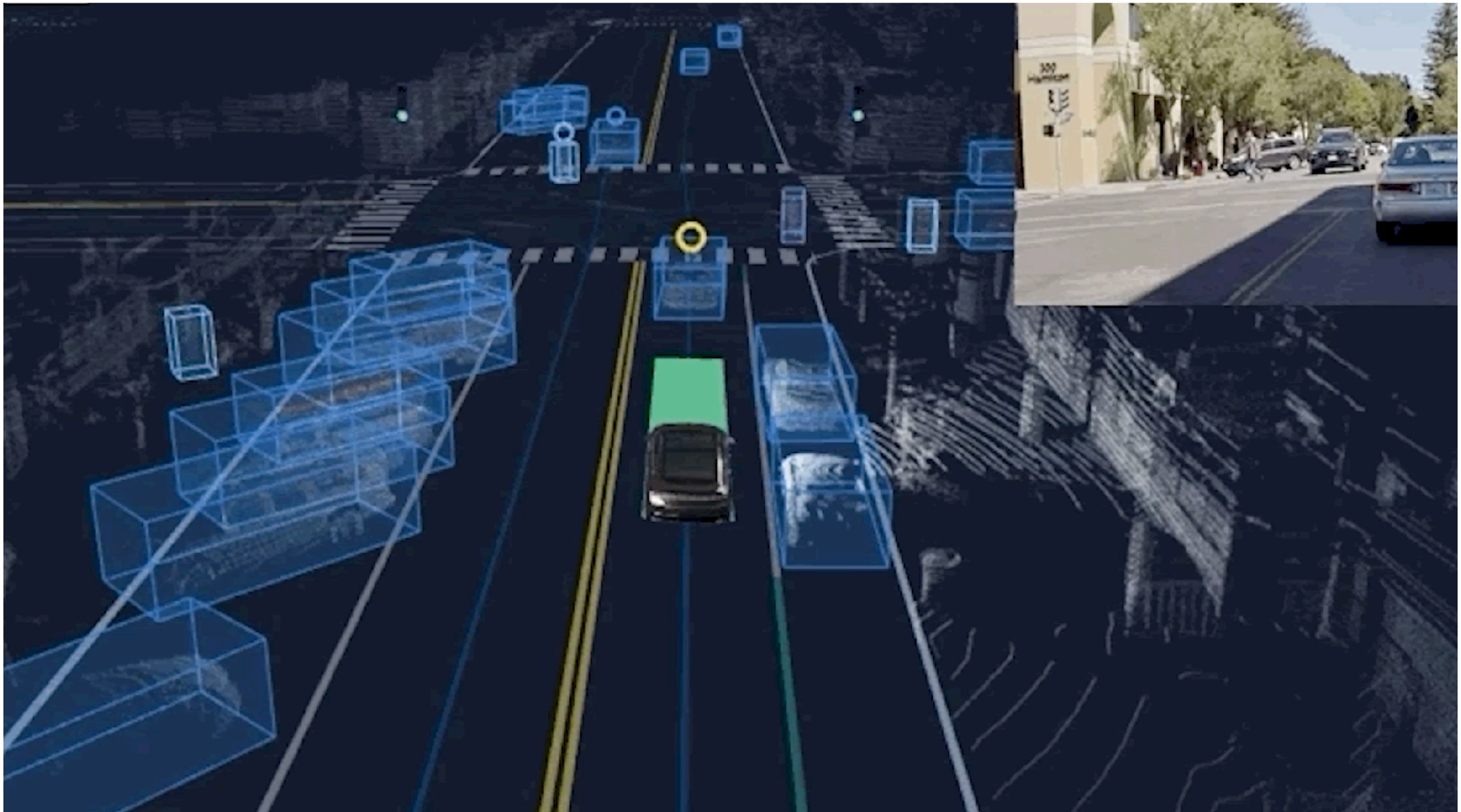
Self-driving



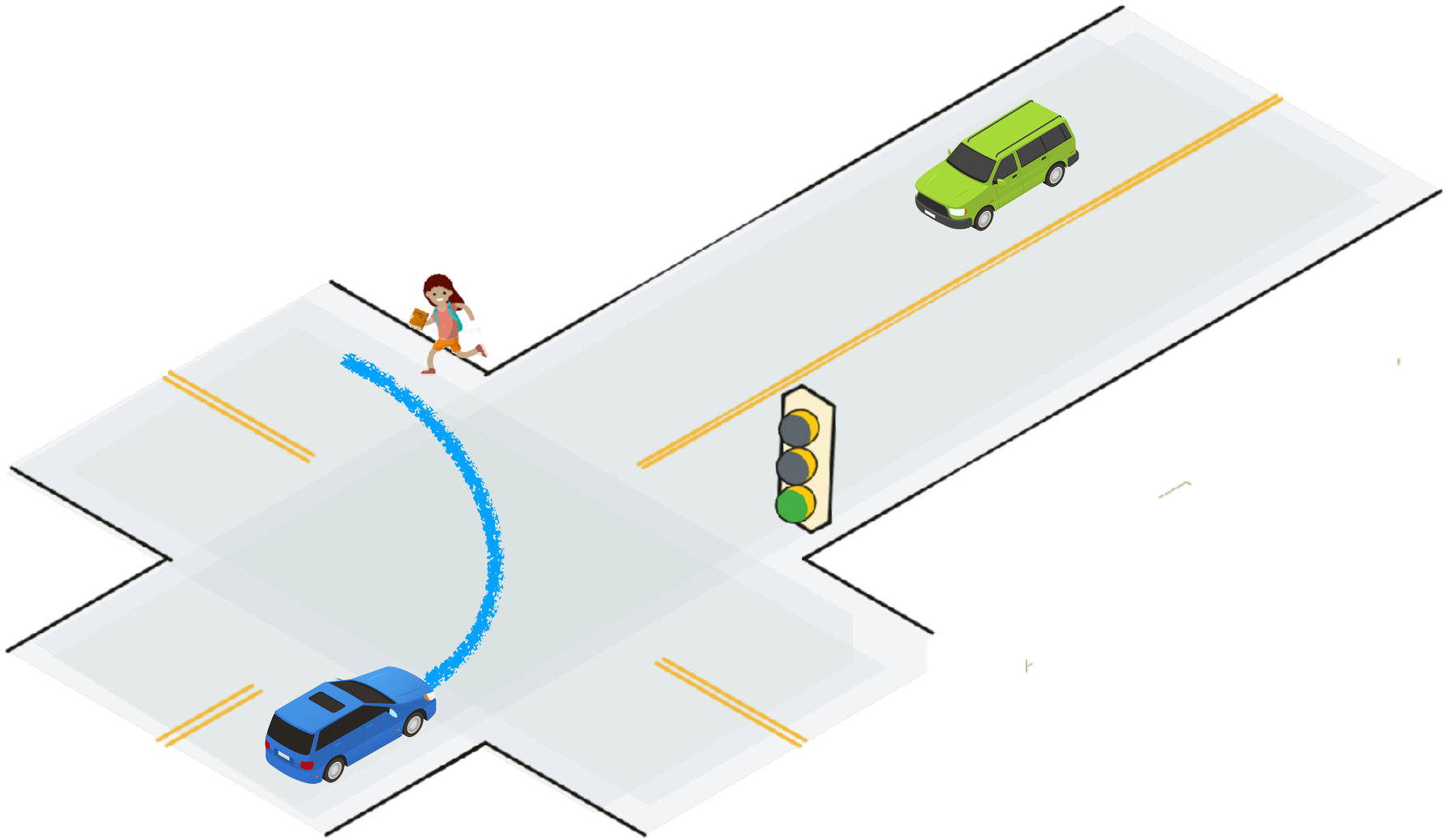
©2021 I Aurora Proprietary

Activity!

Activity: What is “good” behavior in a left turn?



Activity: What is “good” behavior in a left turn?

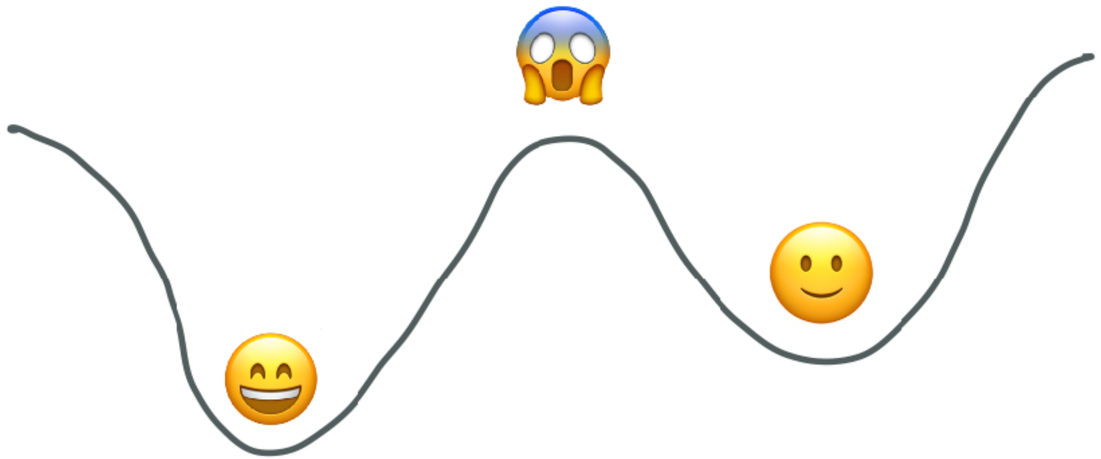


How should robots learn to make good decisions?

Three fundamental questions

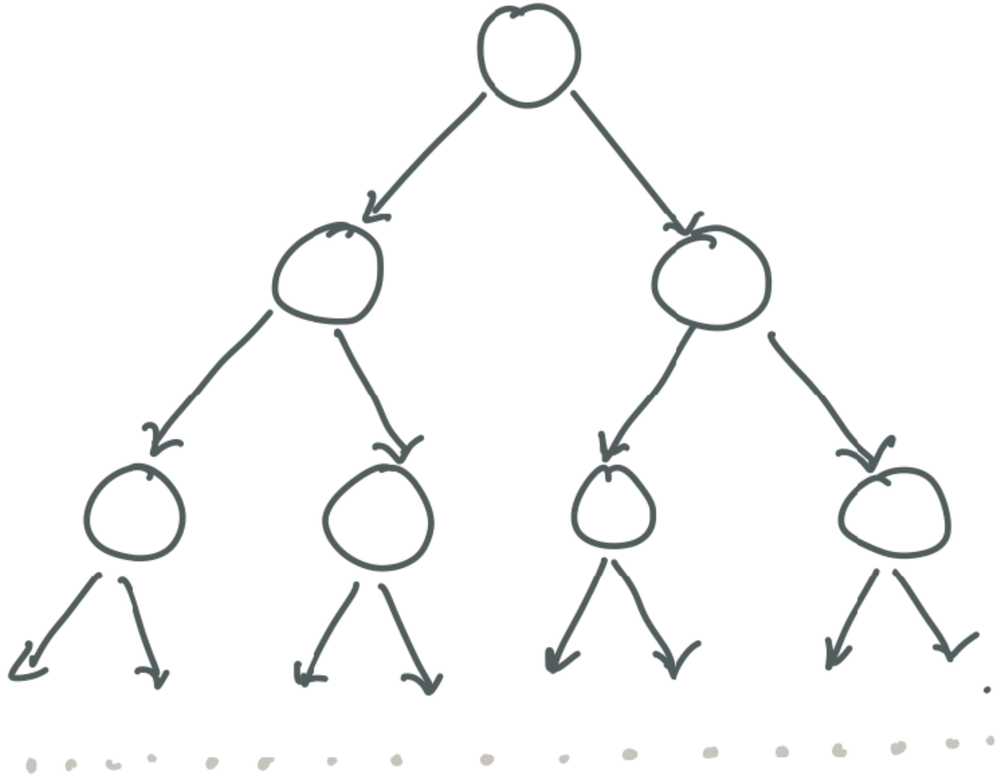
Values

What are good / bad states?



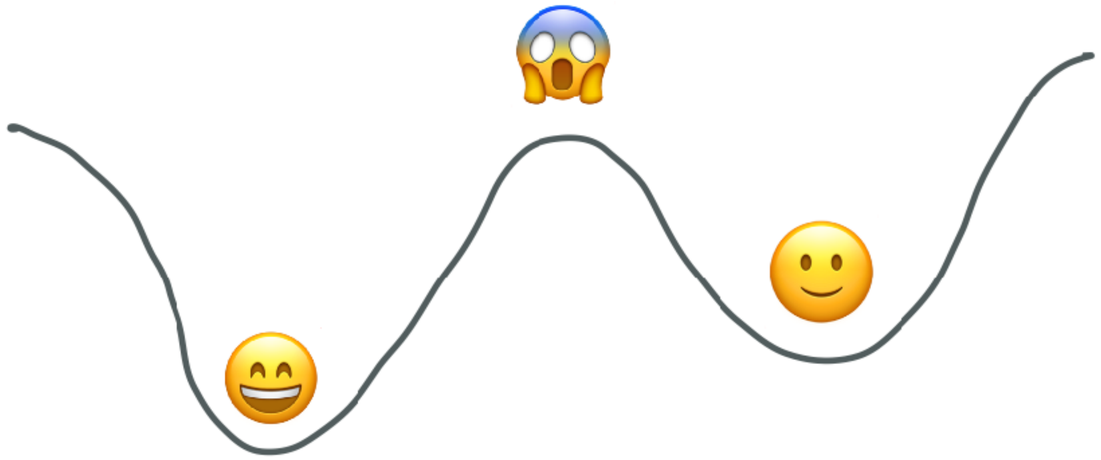
Models

How do decisions affect states?



Values

What are good / bad states?



Optimization

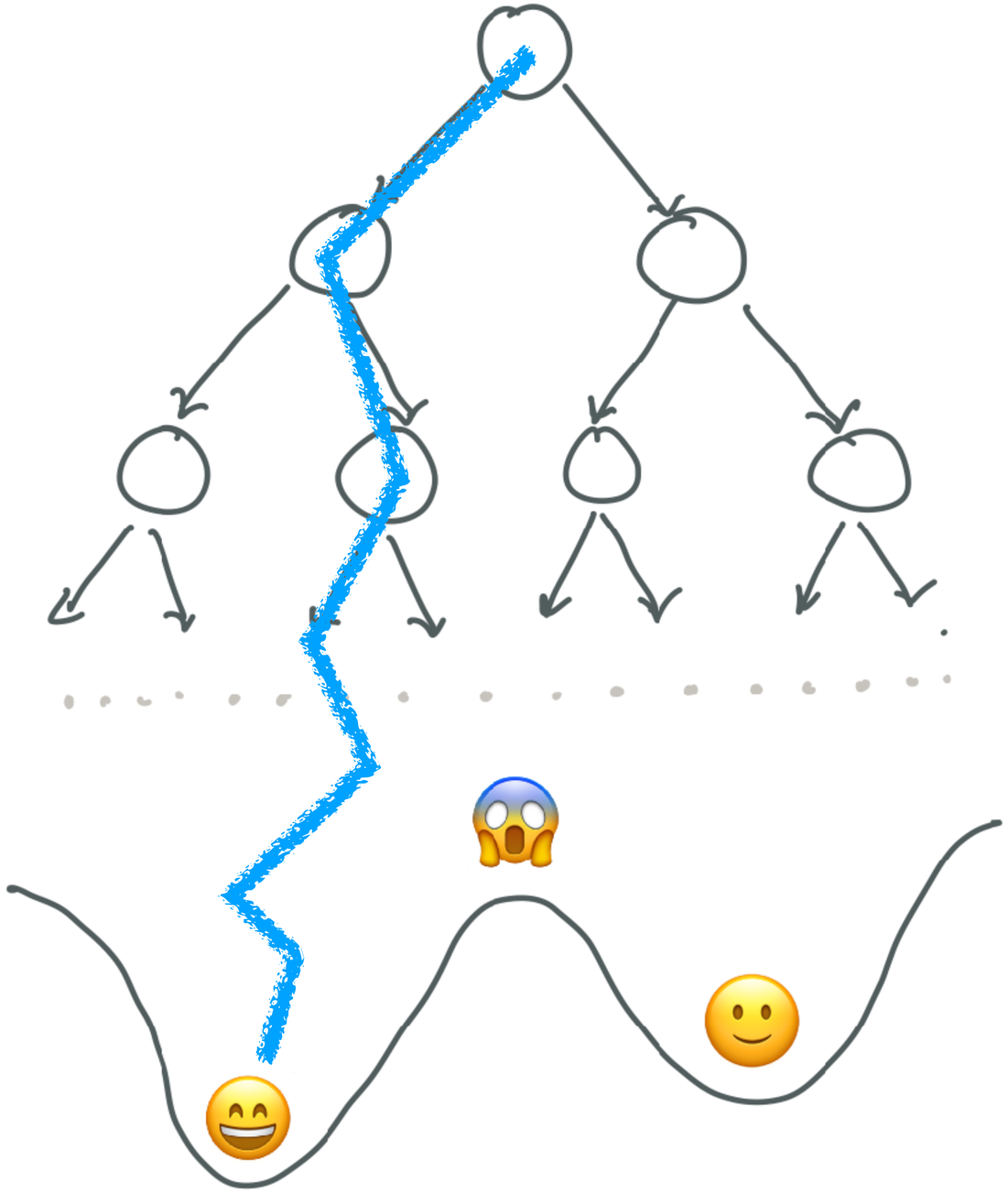
How do we efficiently find
the optimal sequence of decisions?

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

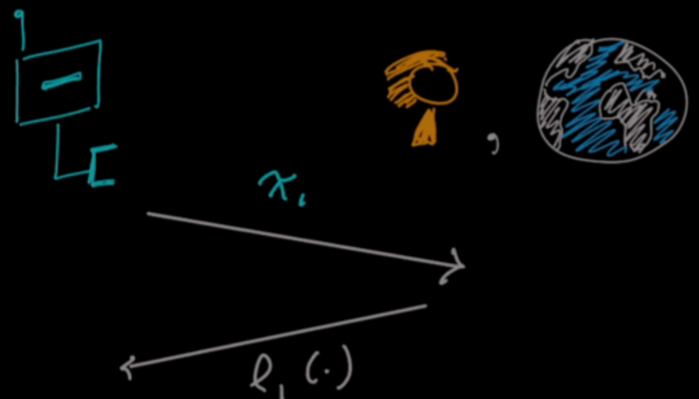
What are good / bad states?

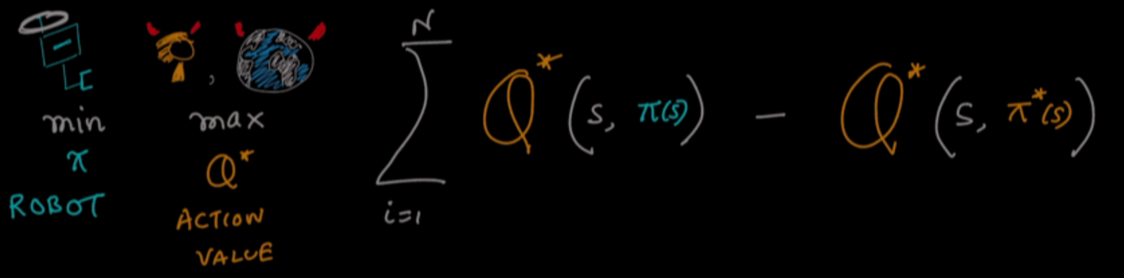
Learning

5 Levels

of

Robot Learning

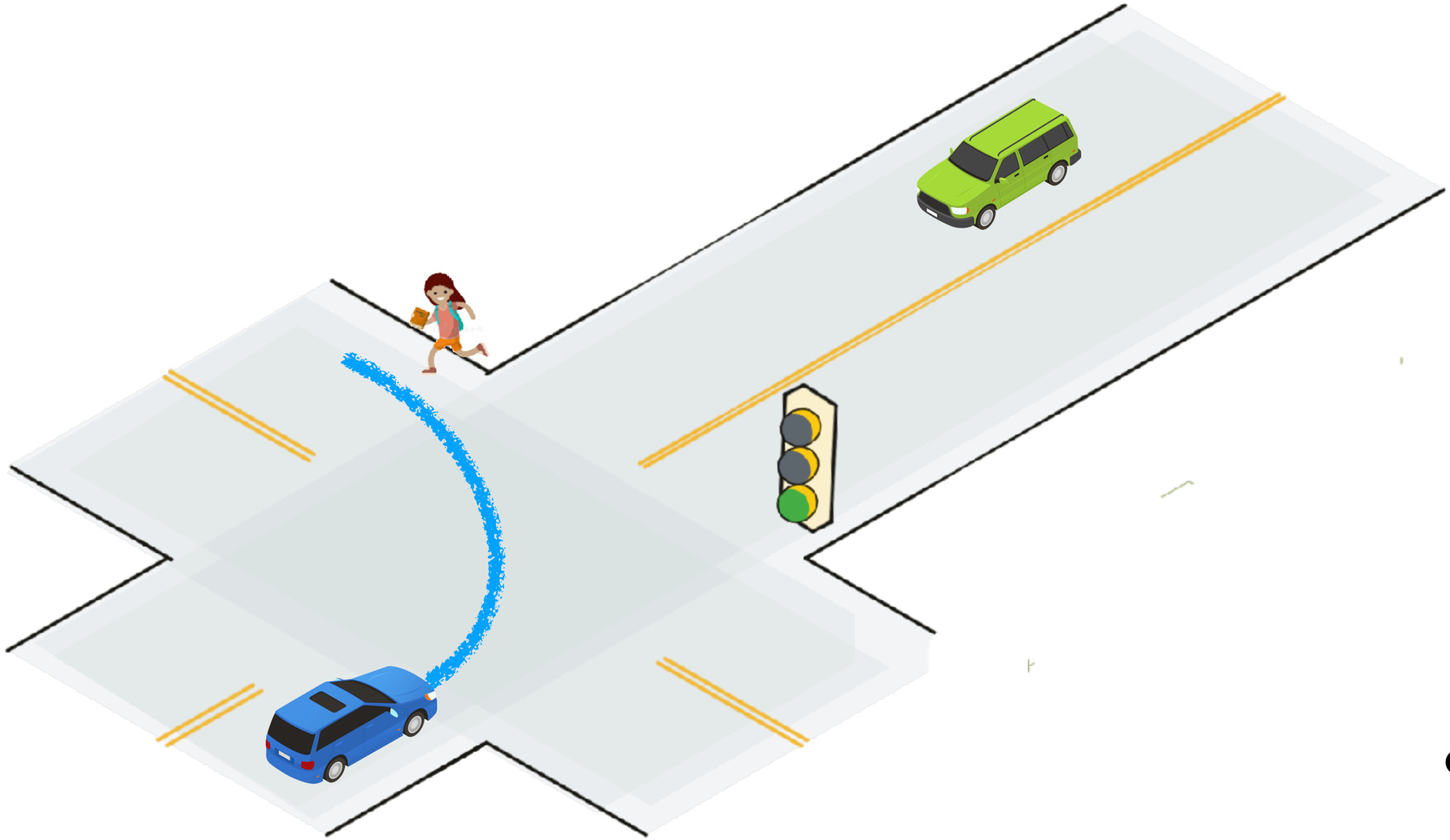




Values

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

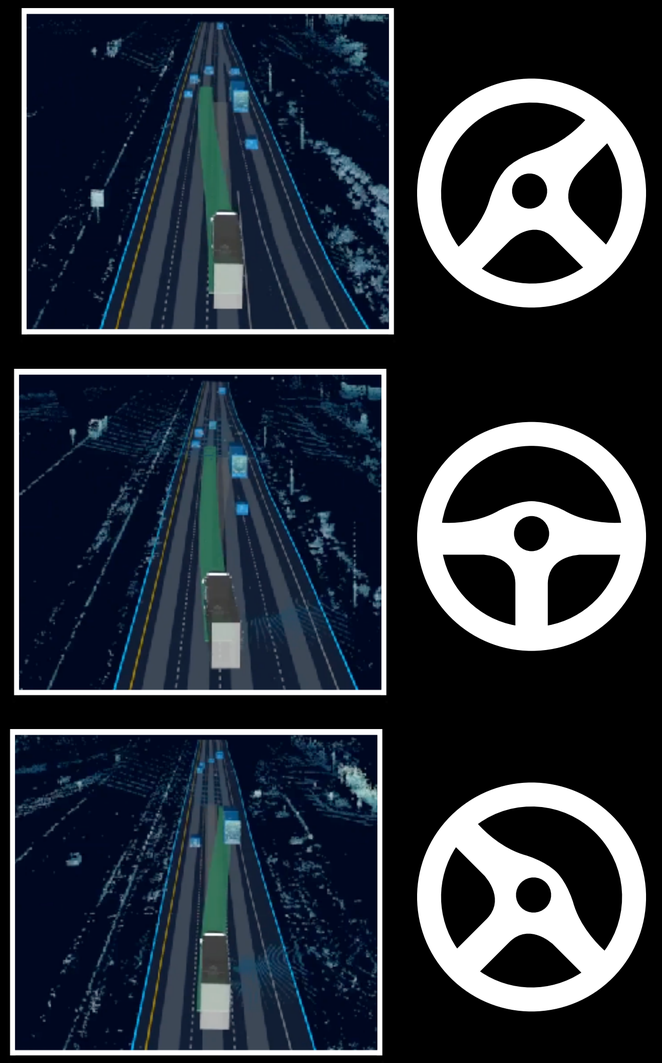
Question:

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

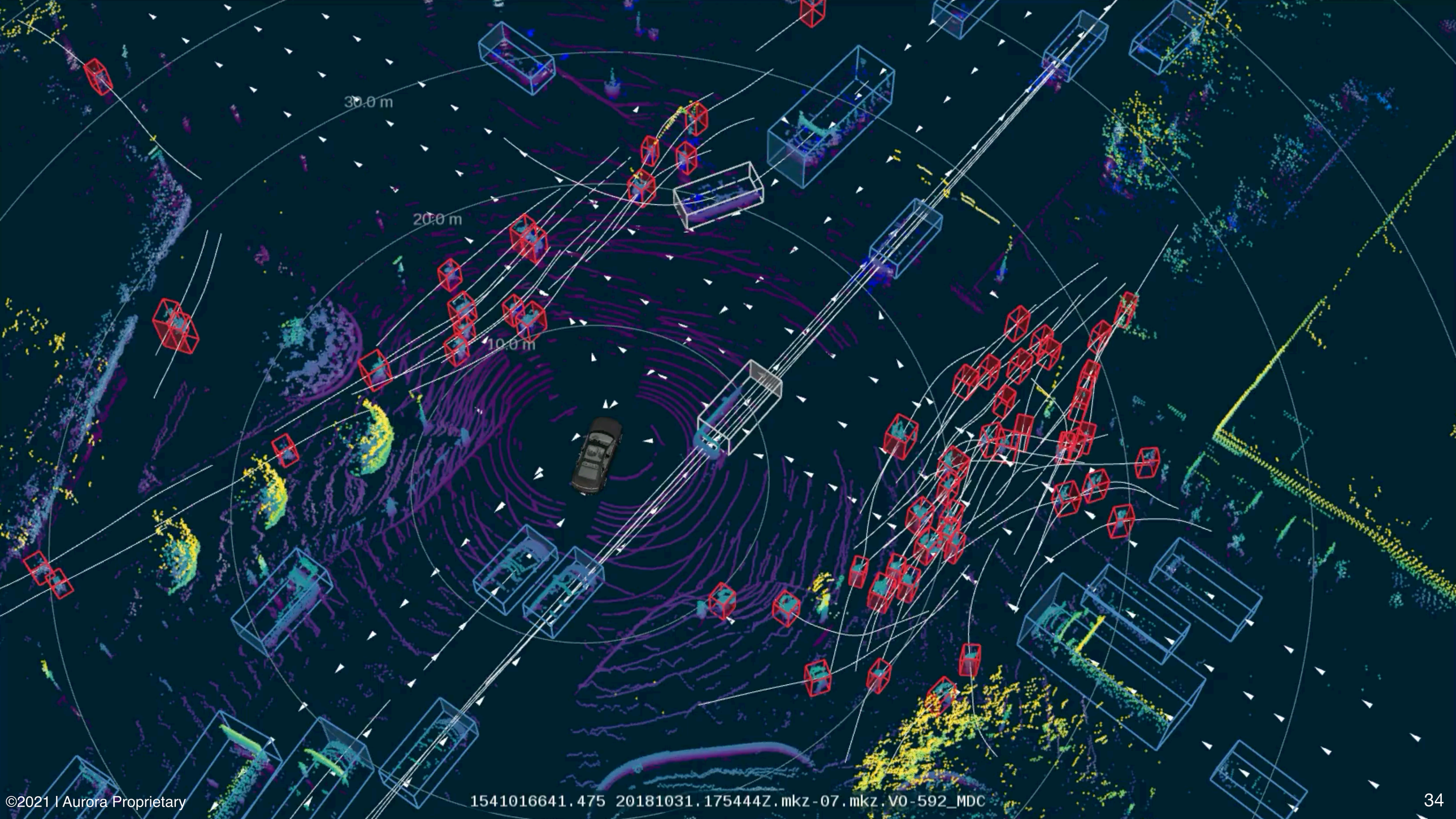
Lesson #1

Feedback drives covariate shift

Models

How do decisions affect states?

Activity!



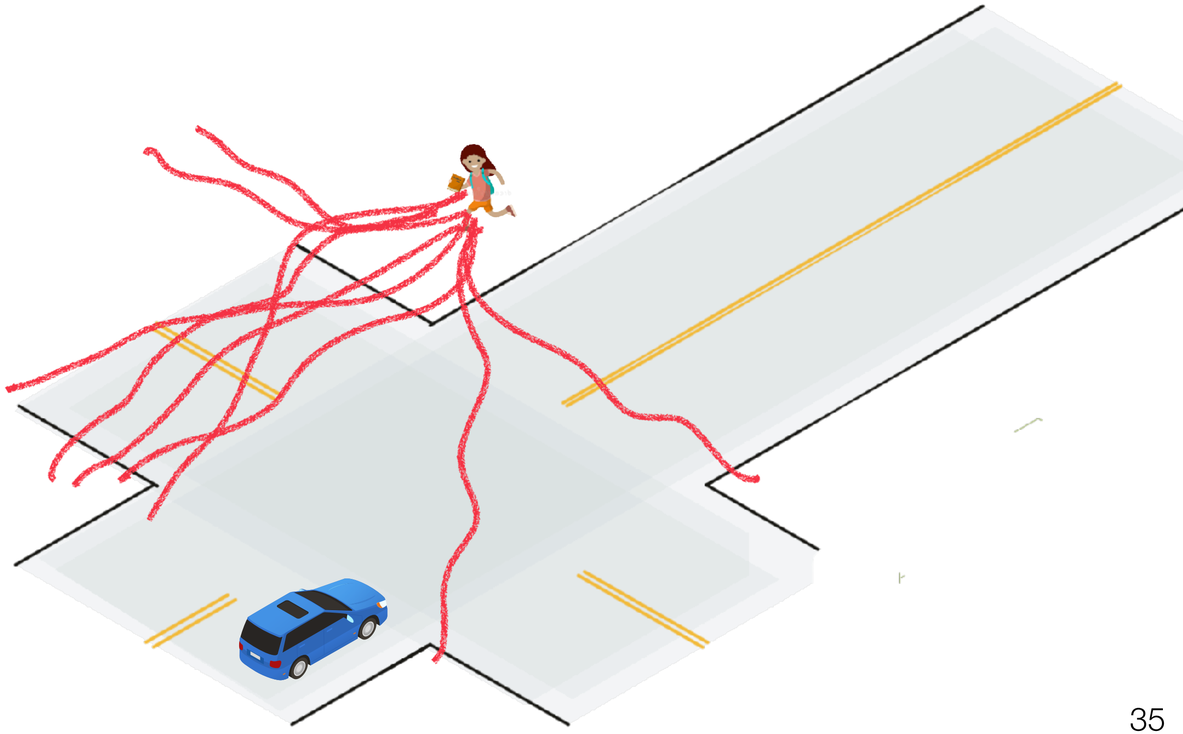
©2021 | Aurora Proprietary

Think-Pair-Share

Think (30 sec): How do we train a model of how pedestrians move?
What are some of the challenges?
What makes for a good model?

Pair: Find a partner

Share (45 sec): Partners exchange ideas



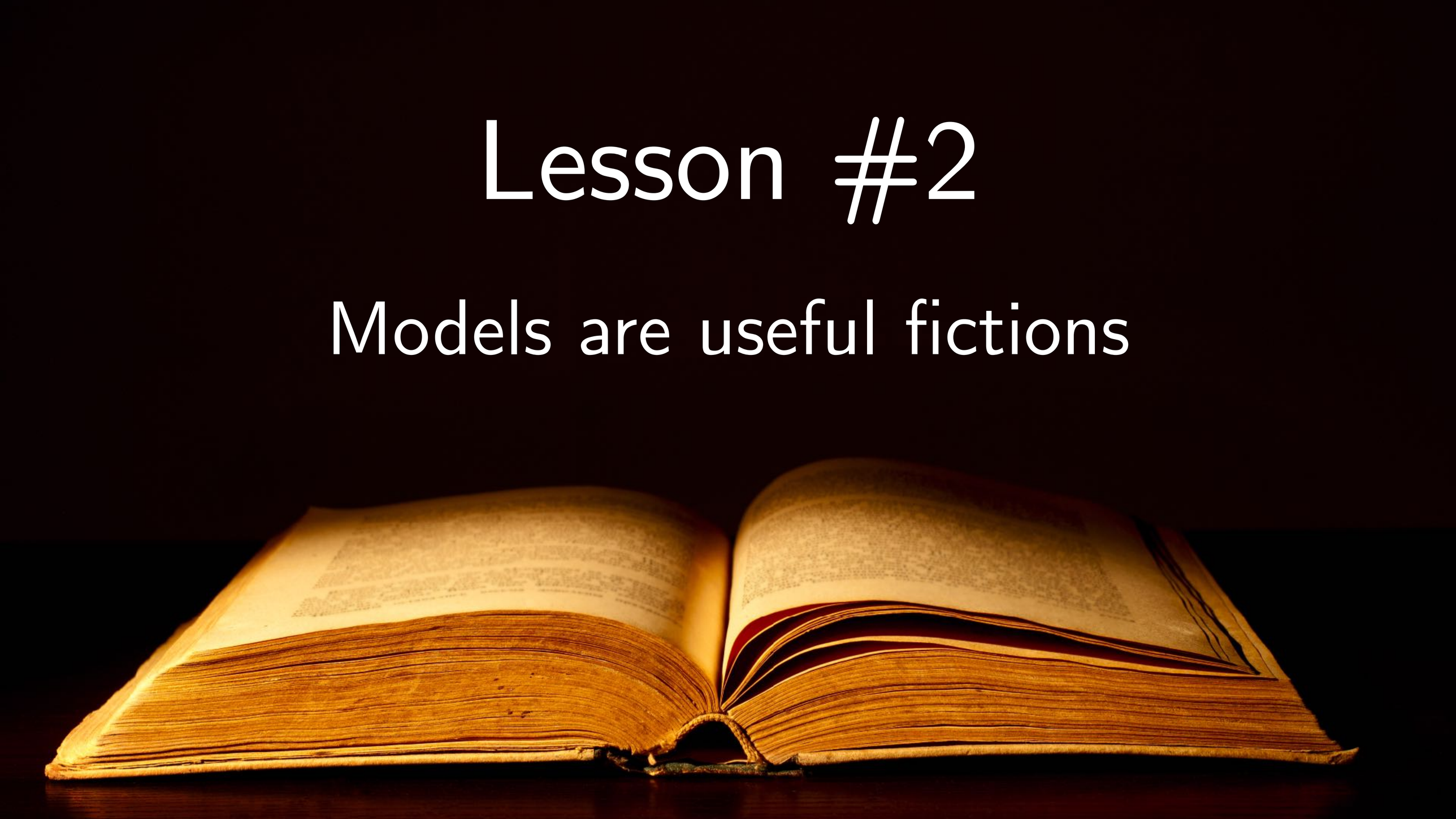
"When solving a problem of interest, do not solve a more general problem as an intermediate step."

Vladmir Vapnik

The Nature of Statistical Learning

Lesson #2

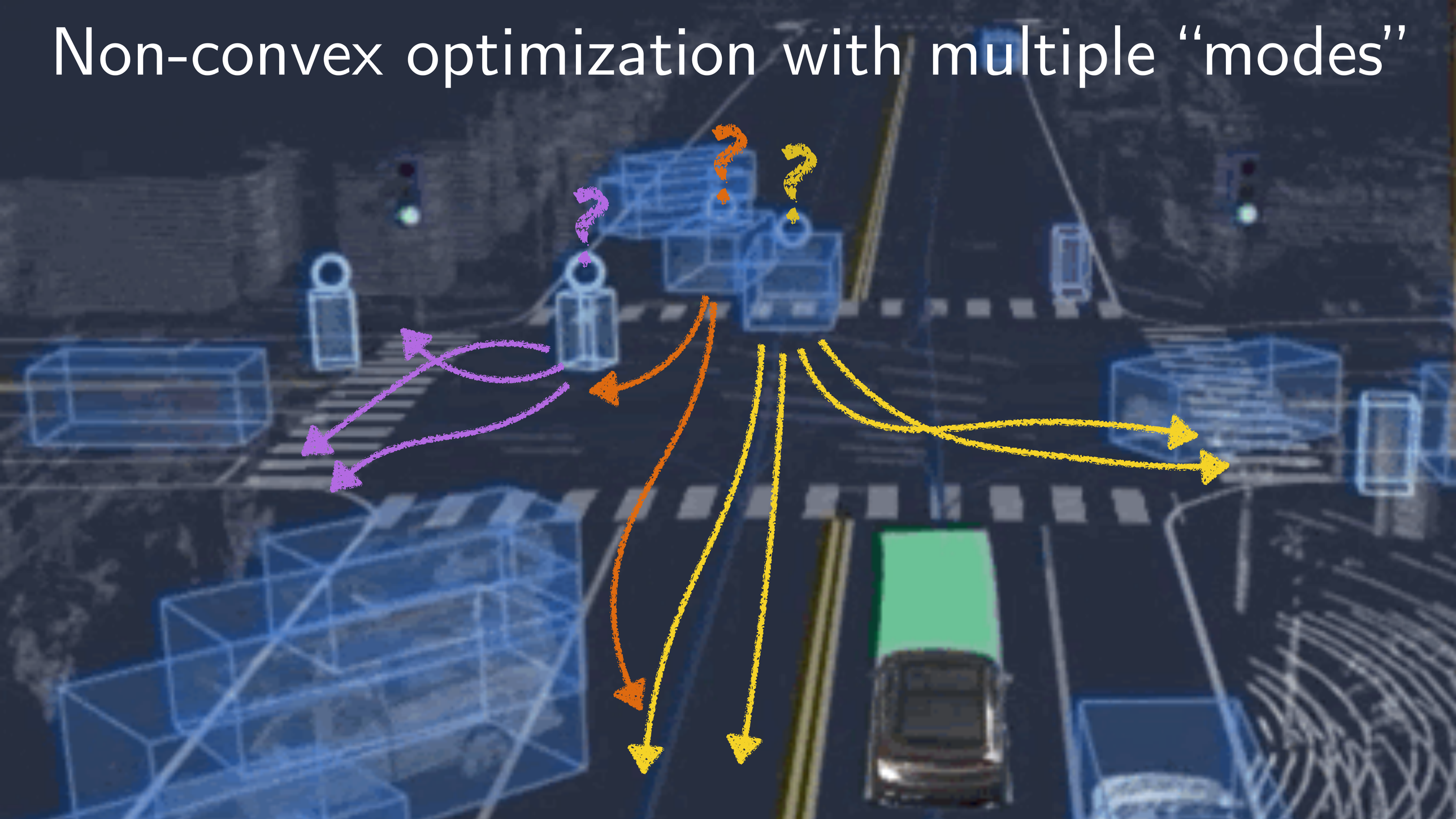
Models are useful fictions



Optimization

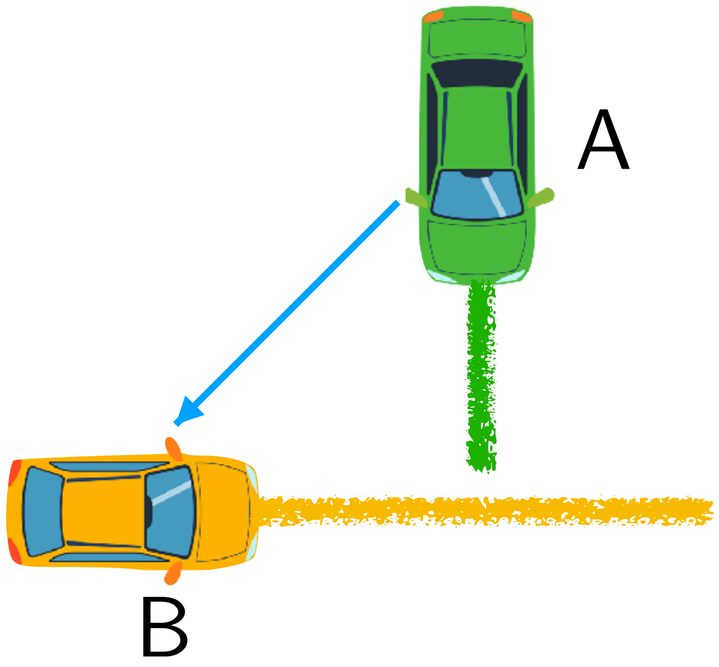
How do we efficiently find the optimal sequence of decisions?

Non-convex optimization with multiple “modes”

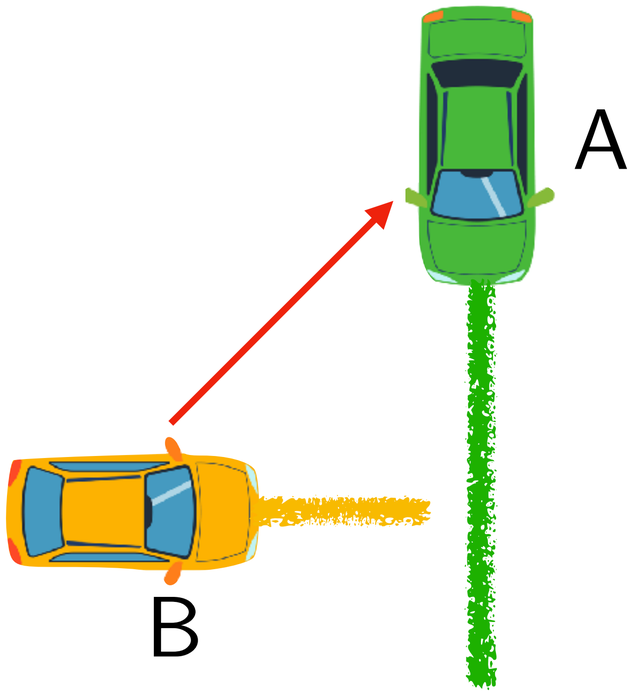


3 discrete modes of space-time paths

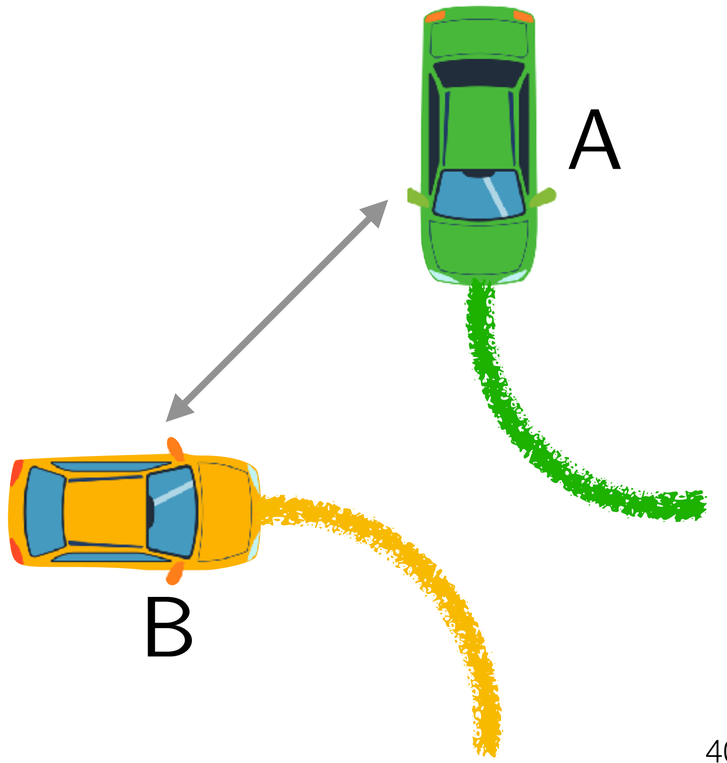
A Yields to B



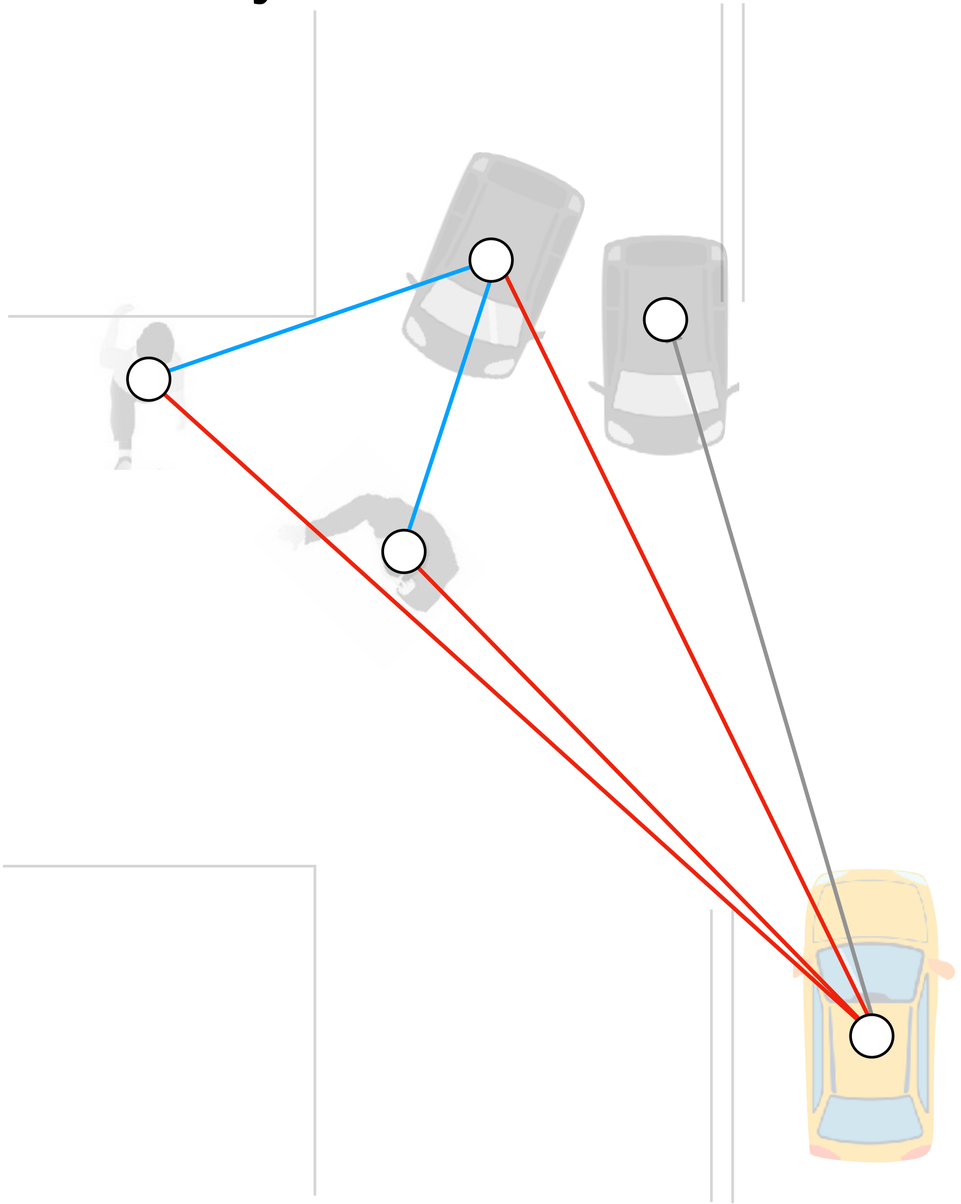
B Yields to A



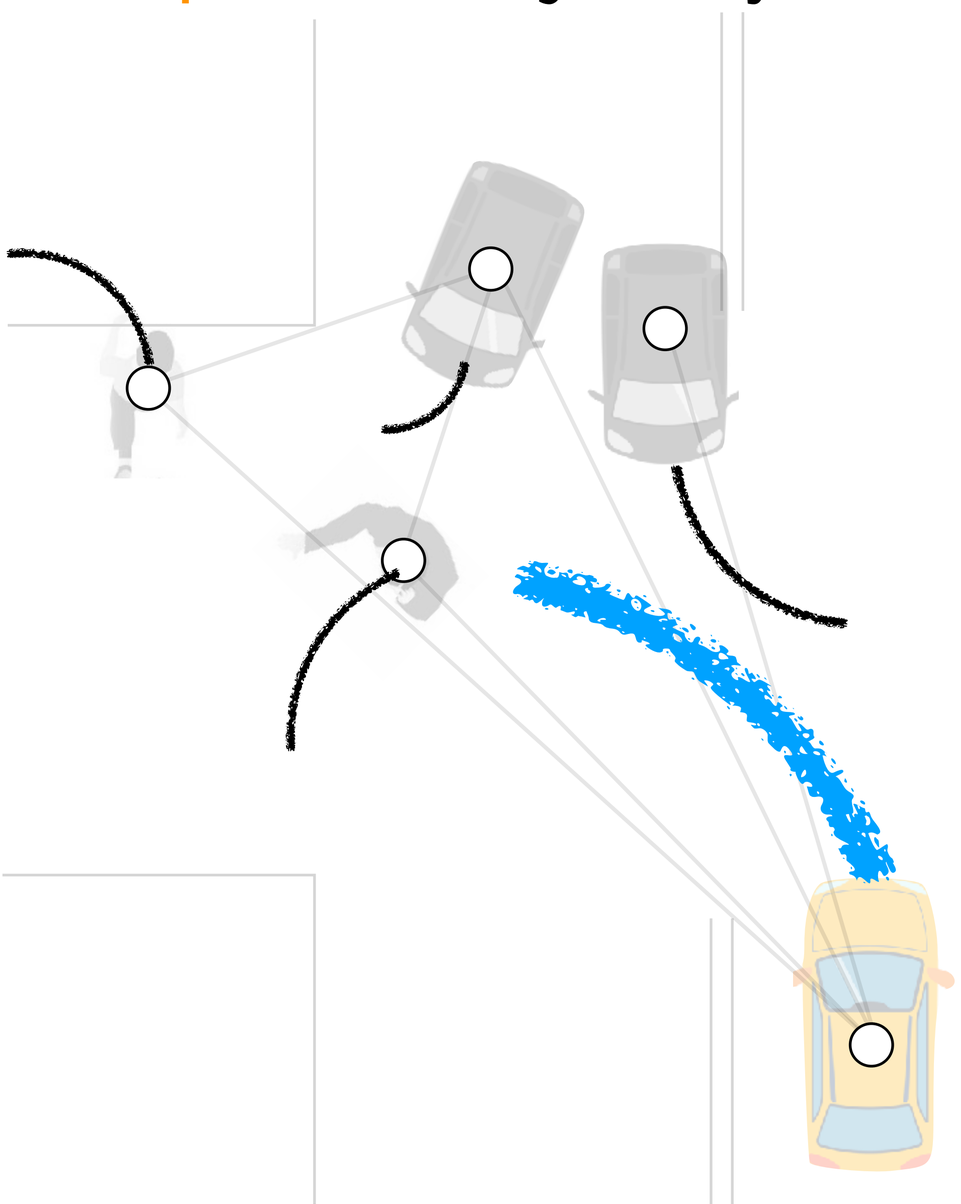
Not Yield



Use learning to recall likely discrete modes

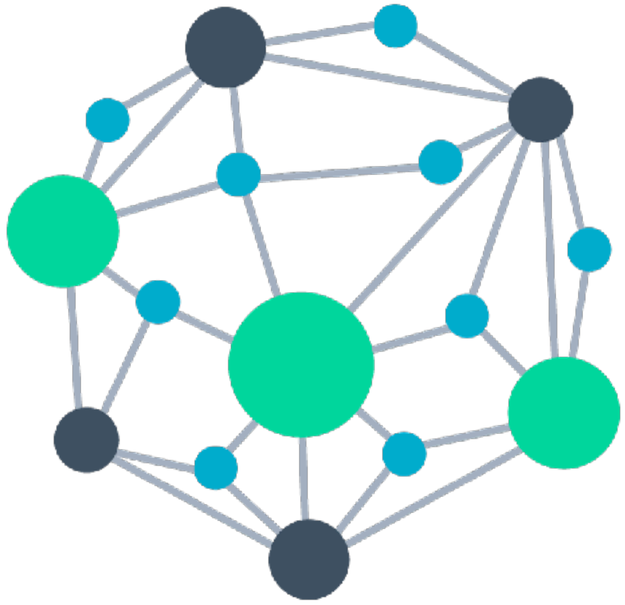


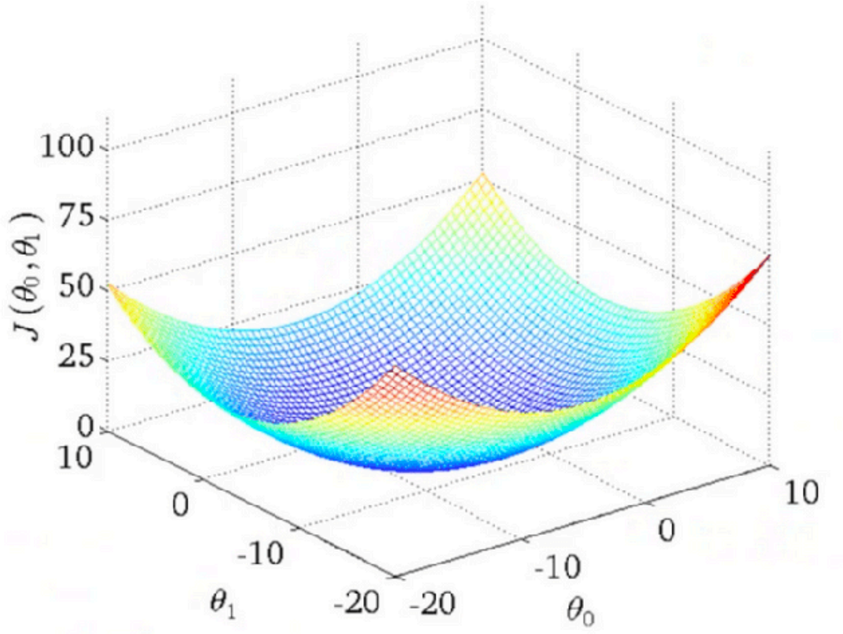
Use optimization to find the precise trajectory



Lesson #3

ML for recall + Optimization for precision





Optimization

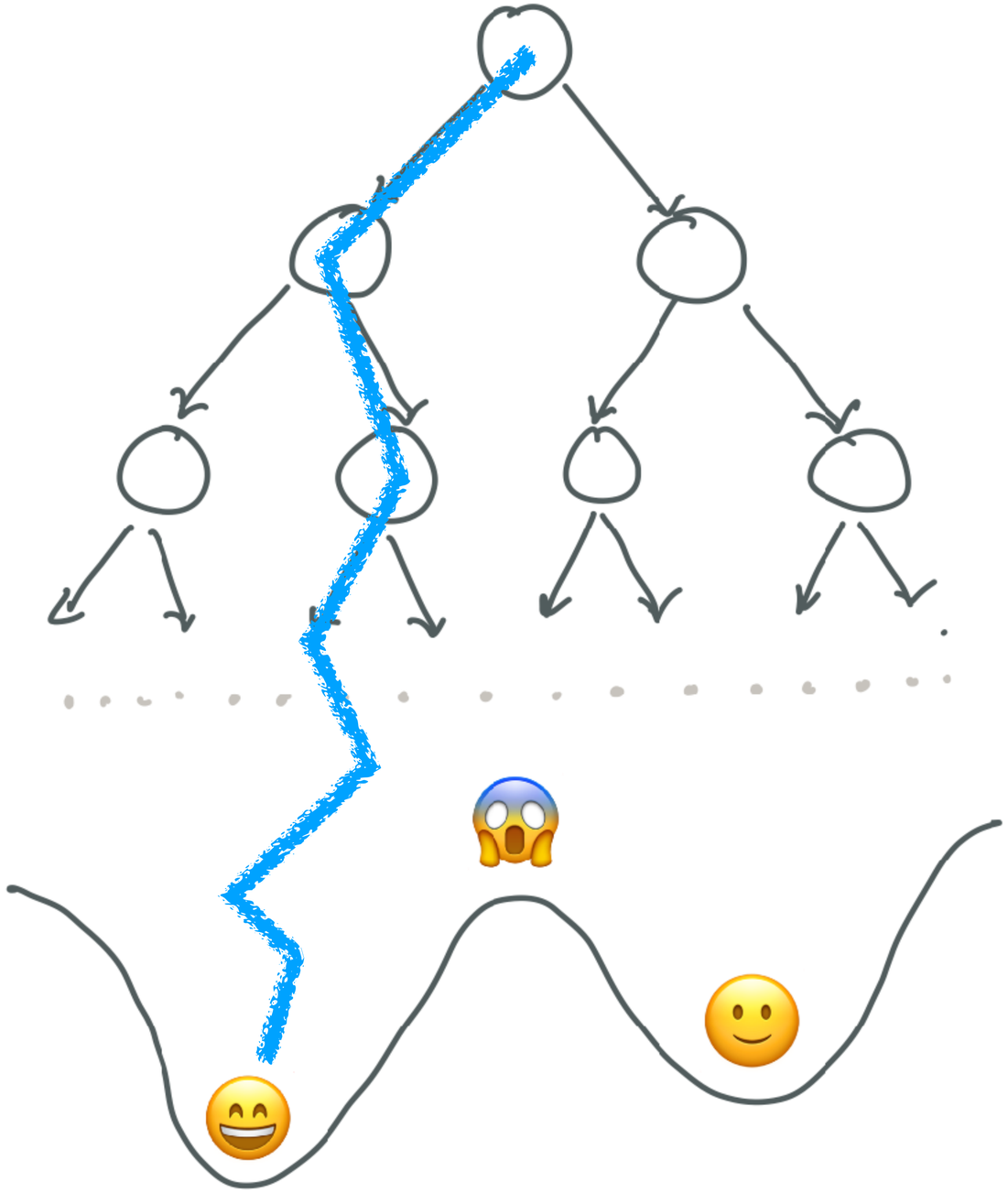
How do we efficiently find
the optimal sequence of decisions?

Models

How do decisions affect
states?

Values

What are good / bad
states?



The journey ahead!

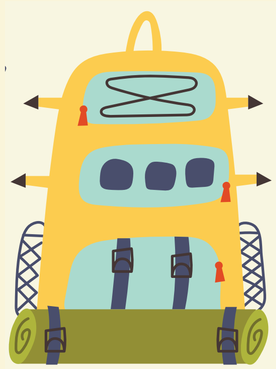
Open Challenges

Reinforcement Learning

(Approximate DP, TD learning, Policy optimization, Trust region)

Imitation Learning

(Behavior cloning, DAGGER, MaxEnt, Distribution matching)

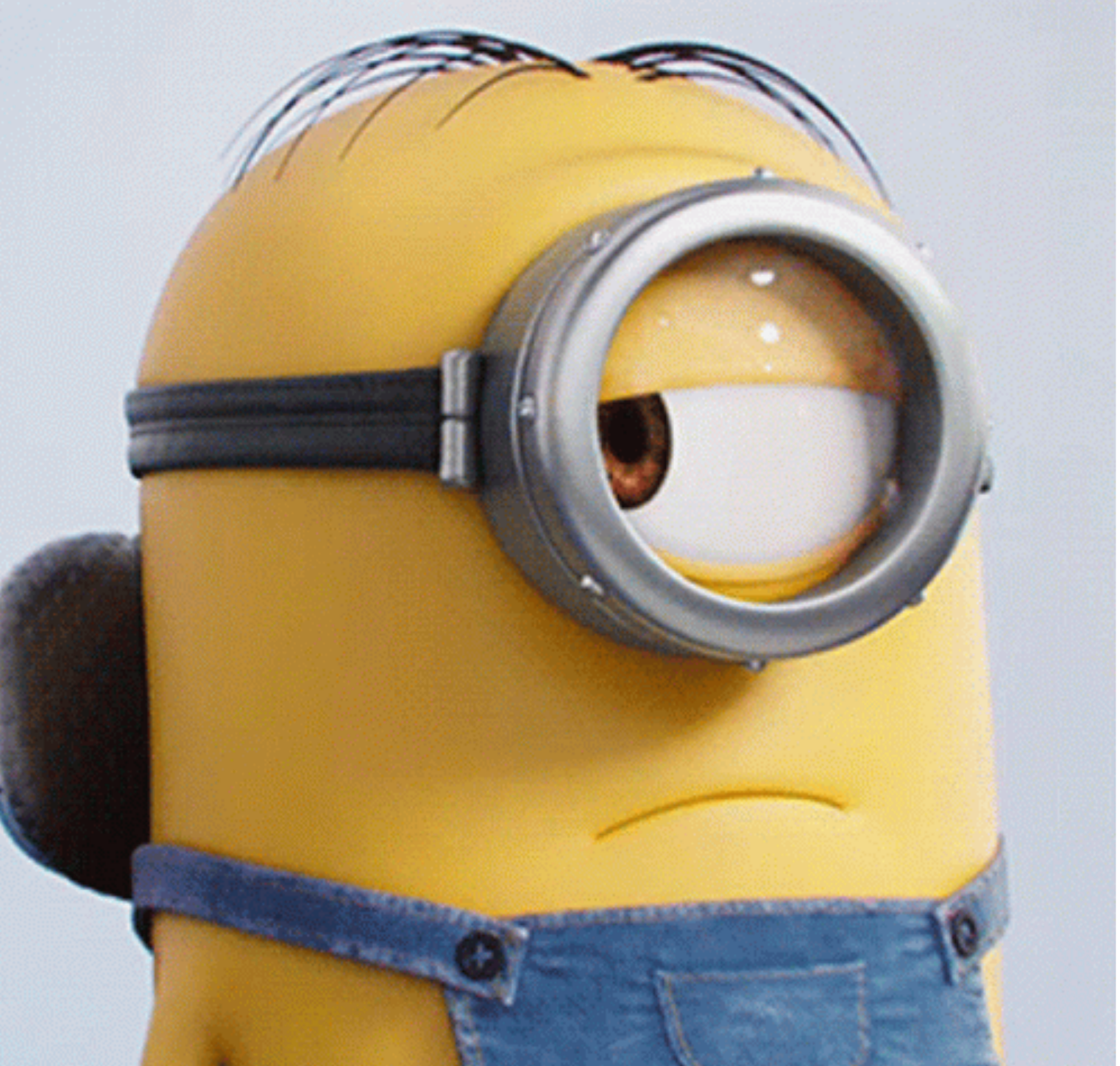


Fundamentals

(MDP, Online learning)

Model Predictive Control

(LQR, iLQR, Augmented Lagrangian)



Logistics

Dhruv Sreenivas

TA Intro!

Second-year CS MS student, advised by Wen Sun.

Generally interested in reinforcement learning, with a focus in

- Imitation learning and offline reinforcement learning, specifically in continuous control
- Deep generative models and connections to model-based RL

Would love to discuss research directions with anyone interested in said topics!

Logistics

Website: https://www.cs.cornell.edu/courses/cs6756/2022fa/

Lectures

Assigned pre-reading (focused), lectures for interaction

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

Python. HW2, HW3 involve PyTorch. Maybe a little theory. 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. Best paper award!

Participation [10%]

Interaction during lectures and/or Online discussions. Help everyone engage!

Course tools

Course Website: The ONE true hub for all information. Please check this frequently and surface any errors or sources of confusion.

Ed: The discussion forum where all announcements are sent, where all student-TA and student-student communications occur.

Gradescope: Where all assignments and projects are submitted.

Canvas: Limited to no use.

Books and other resources

Work-in-progress book

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

(Please feel free to send me feedback)

For other resources, keep checking website

Assignment 0

Simple survey

- What are you hoping to get out of this class?
- Familiarity with concepts
- Etc

Released soon (today..), due end of the week

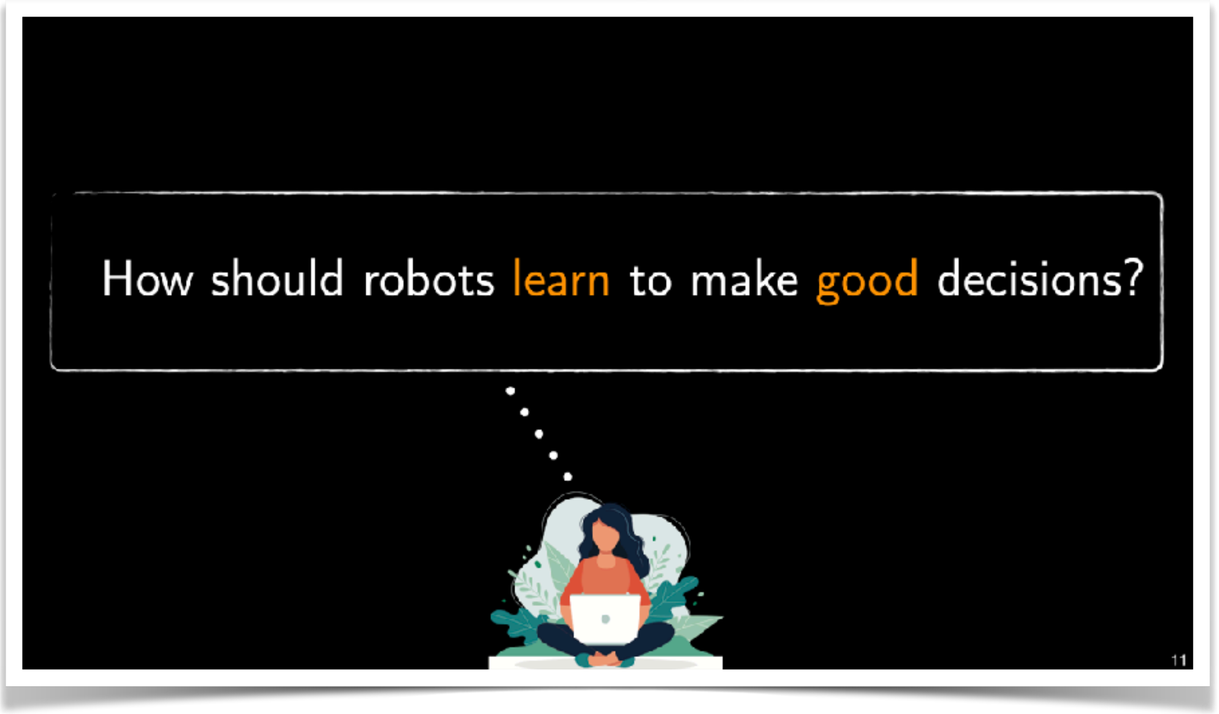
Mandatory!

Questions?

Things that will happen soon:

- Waitlist moved over to enrolled
- Office hours finalized

tl;dr



Optimization

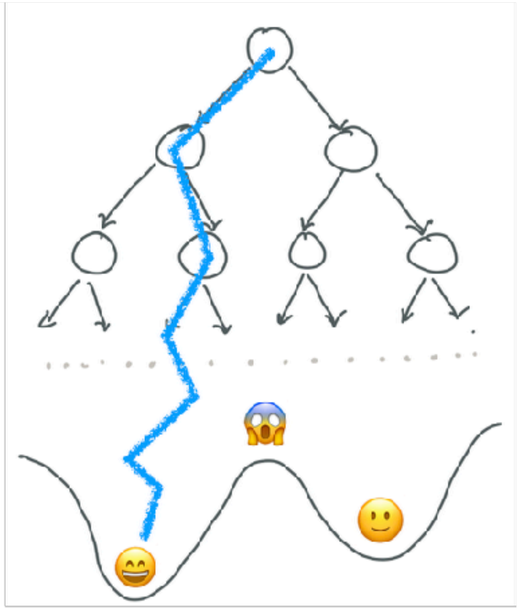
How do we efficiently find the optimal sequence of decisions?

Models

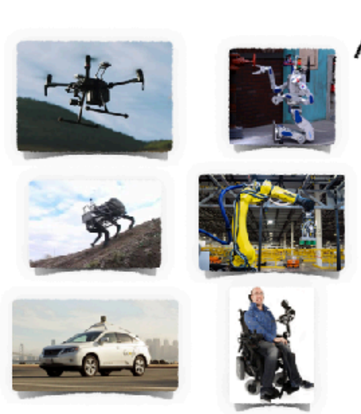
How do decisions affect states?

Values

What are good / bad states?



WHY this course?

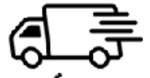


Take *any* robot application

Formulate as a Markov Decision Problem (MDP)

Solve MDPs using all-purpose toolkit

(Imitation/Reinforcement learning, Model based/free)



Deploy learners in real-world

(Safety, distribution shift, value alignment)