Lecture 12

Retrospective & frontiers

Carnegie Mellon University

abajcsy@cmu.edu

intent ROBOTICS LAB

Last Time

[✓] latent-space safety
[✓] *(guest lecture)* OOD in era of large models

This Time

[ ] final project + presentation logistics
[ ] retrospective & frontiers!
[ ] course eval survey

At a glance

Final presentations due ** All presentation slides must be uploaded*	4/21
Presentation talks	4/22 & 4/24
Final report due	5/1

Final *Report* (30% | May 1)

Conference-style paper

~6 pages

IEEE templates in LaTeX and Overleaf *(click image on right to go to Overleaf template)*

Conference Paper Title*

*Note: Sub-titles are not captured in Xplore and should not be used

1st Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

3rd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

4th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

5th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

6th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

***Abstract*—This document is a model and instructions for LaTeX. This and the IEEEtran.cls file define the components of your paper [title, text, heads, etc.]. *CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

This document is a model and instructions for LaTeX. Please observe the conference page limits.

II. Ease of Use

A. Maintaining the Integrity of the Specifications

The IEEEtran class file is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

III. Prepare Your Paper Before Styling

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections III-A–III-E below for more information on proofreading, spelling and grammar.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not number text heads—LaTeX will do that for you.

Identify applicable funding agency here. If none, delete this.

A. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, ac, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

B. Units

- Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as "3.5-inch disk drive".
- Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.
- Do not mix complete spellings and abbreviations of units: "Wb/m^2" or "webers per square meter", not "$webers/m^2$". Spell out units when they appear in text: ". . . a few henries", not ". . . a few H".
- Use a zero before decimal points: "0.25", not ".25". Use "cm^3", not "cc".)

C. Equations

Number equations consecutively. To make your equations more compact, you may use the solidus (/), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in:

$$a + b = \gamma \quad (1)$$

Be sure that the symbols in your equation have been defined before or immediately following the equation. Use "(1)", not

https://www.overleaf.com/latex/templates/ieee-conference-template/grfzhhncsfqn

Final Project Rubric

Criteria	Ratings		Pts
Background and Motivation Does the report offer a clear introduction of the chosen problem or topic of study, and a compelling justification of its importance? Are the problem and the proposed work placed into the broader technical context and connected to prior research efforts?	**25 pts Full Marks**	**0 pts No Marks**	25 pts
Formulation and Analysis Is the problem or topic of study put into a well-defined technical representation? Is there a clear definition of the scope and goals of the project? Are the techniques used to shed light on it applied correctly and appropriately?	**25 pts Full Marks**	**0 pts No Marks**	25 pts
Results and Insights How compelling are the results of your work so far? Is there a valuable technical contribution, in terms of novel research or understanding of existing knowledge? Does the audience walk away from your talk with meaningful new insights?	**25 pts Full Marks**	**0 pts No Marks**	25 pts
Clarity and Rigor Are the ideas systematically explained so that they can be understood by a technically-equipped reader who is not already an expert in this work? Are technical arguments carefully and correctly laid out?	**25 pts Full Marks**	**0 pts No Marks**	25 pts

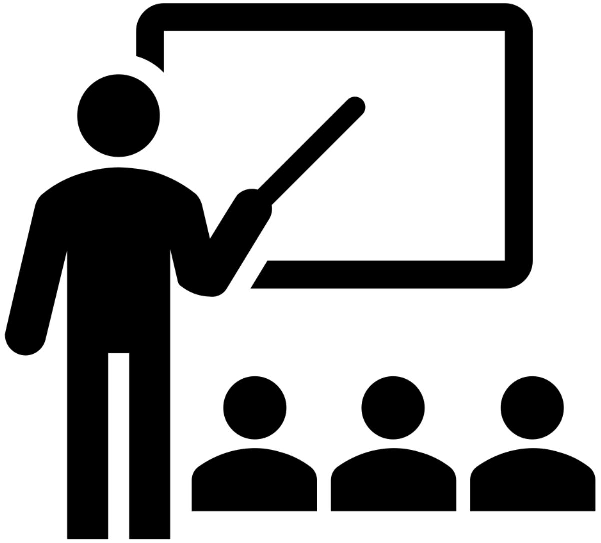
Final *Presentation* (10% | April 22 & April 24)

Conference-style talk

For <u>groups of 1:</u>
    10 minute presentation
    + 5 minute Q&A / transition

For <u>groups of N:</u>
    20 minute presentation
    + 5 minute Q&A / transition

Whole must be class present and in-person!



Day 1 (April 22)

Presenter(s)	Presentation Time
Bowen Jiang, Yilin Wu, Weihao (Zack) Zeng	20 min
Samuel Li	10 min
Sidney Nimako – Boateng	10 min
Xilun Zhang	10 min

Day 2 (April 24)

Presenter(s)	Presentation Time
Jehan Yang, Eliot Xing	20 min
Yumeng Xiu	10 min
Kavya Puthuveetil	10 min

Paper Presentation Rubric

Criteria	Ratings		Pts
Clarity of Exposition How easy is it to understand the ideas you put forward and follow your technical arguments? Could someone give a good 1-minute summary of your work after listening to your talk?	**25 pts Full Marks**	**0 pts No Marks**	25 pts
Narrative Structure How effectively are you walking us through the problem you are investigating? Do you give a compelling motivation for your work? How do you guide us to arrive at the key insight(s)? Does the future work follow naturally and make sense?	**25 pts Full Marks**	**0 pts No Marks**	25 pts
Presentation Design Does the auxilliary content of your presentation (slides, plots, videos, notes, etc.) complement your verbal delivery? Is it possible to take in all visually displayed information while also following what you are saying, or do you have a "wall of text" (or even worse, a "wall of math"!) that forces the audience to choose between reading and listening?	**25 pts Full Marks**	**0 pts No Marks**	25 pts
Results and Insights How compelling are the results of your work so far? Is there a valuable technical contribution, in terms of novel research or understanding of existing knowledge? Does the audience walk away from your talk with meaningful new insights?	**25 pts Full Marks**	**0 pts No Marks**	25 pts
			Total Points: 100

Motivate what you are studying clearly

On technical communication

(presentations, writing,…)

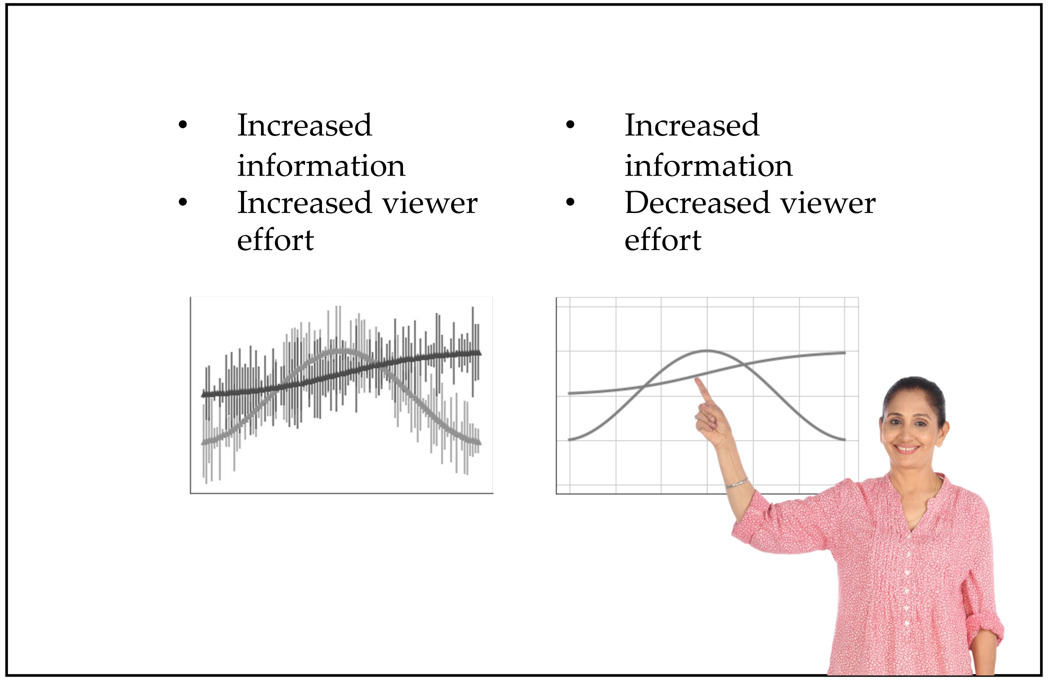
Distill what is the **key idea** and **key takeaways**

On technical communication

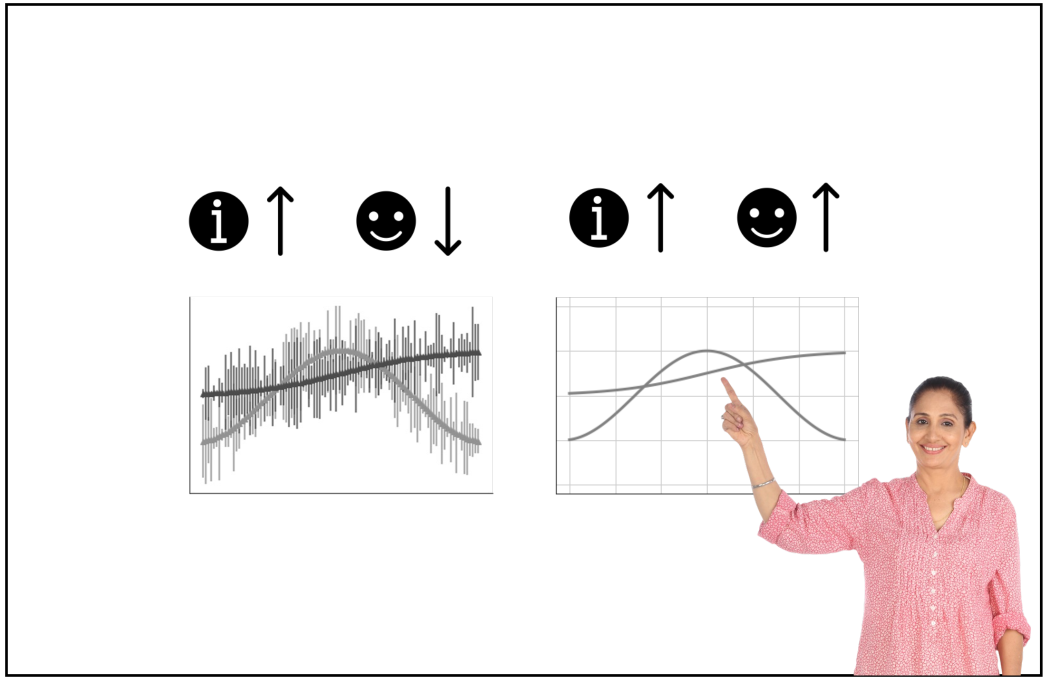
(presentations, writing,…)

Heuristics for good presentations

Sparse; figures over text!



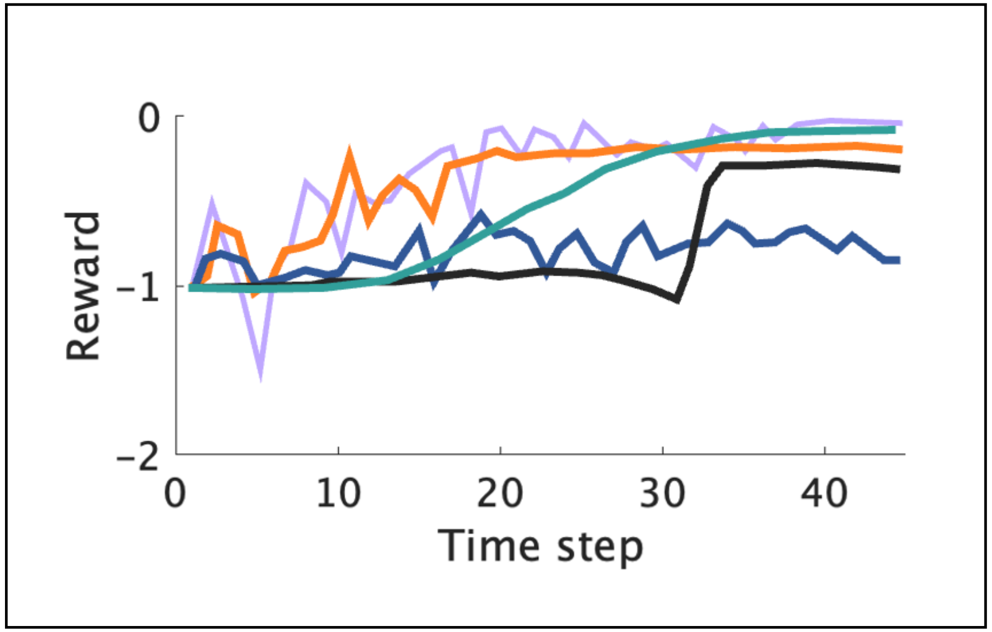
*focus on **slide***

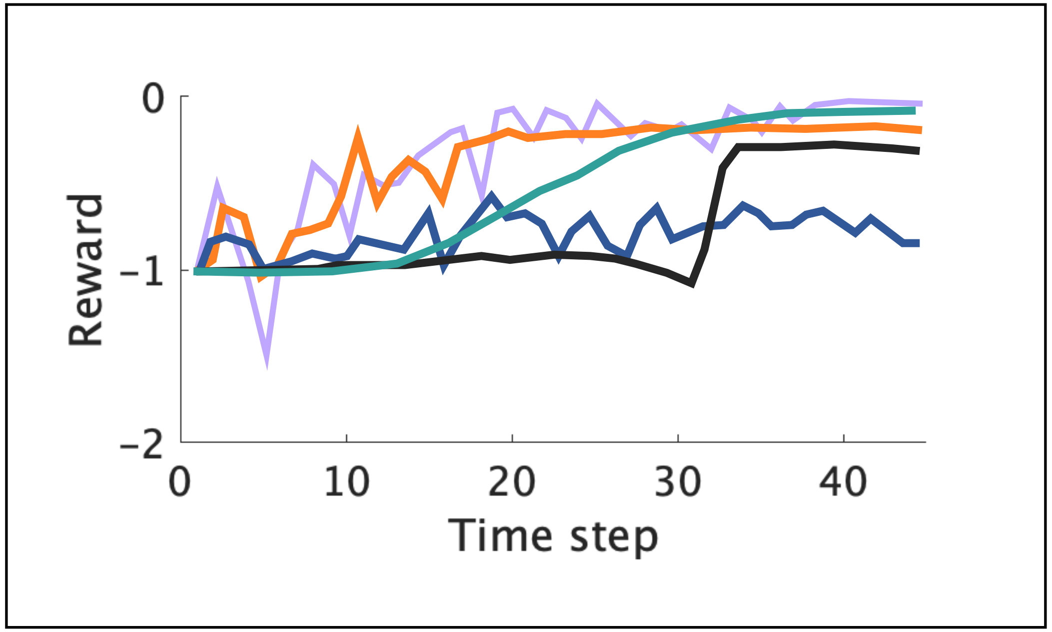


*focus on **speaker***

Heuristics for good presentations

Be visual (e.g., make graphs and break them down)





Heuristics for good presentations

If using equations, explain them and build them up

HJ Reachability

$$\max_{u} \min_{d} \nabla_x V(x,t)^\top f(x,u,d) + \frac{\partial V}{\partial t} = 0$$

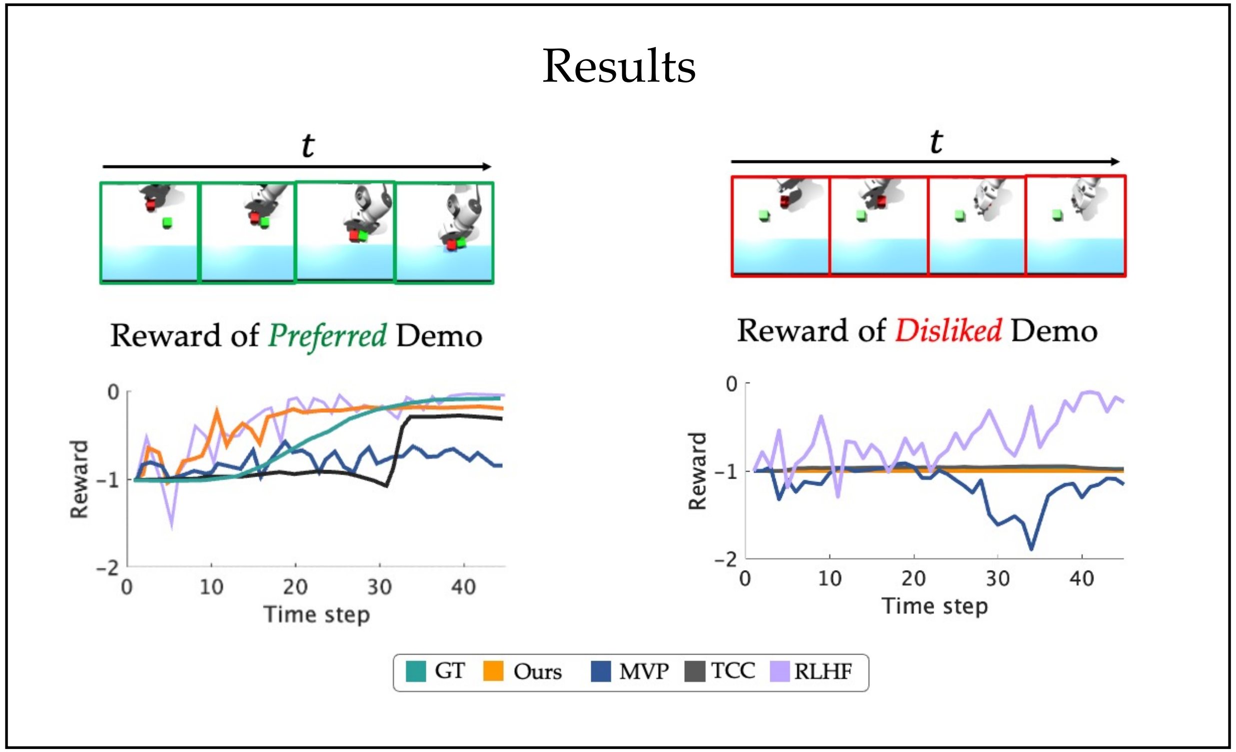
HJ Reachability

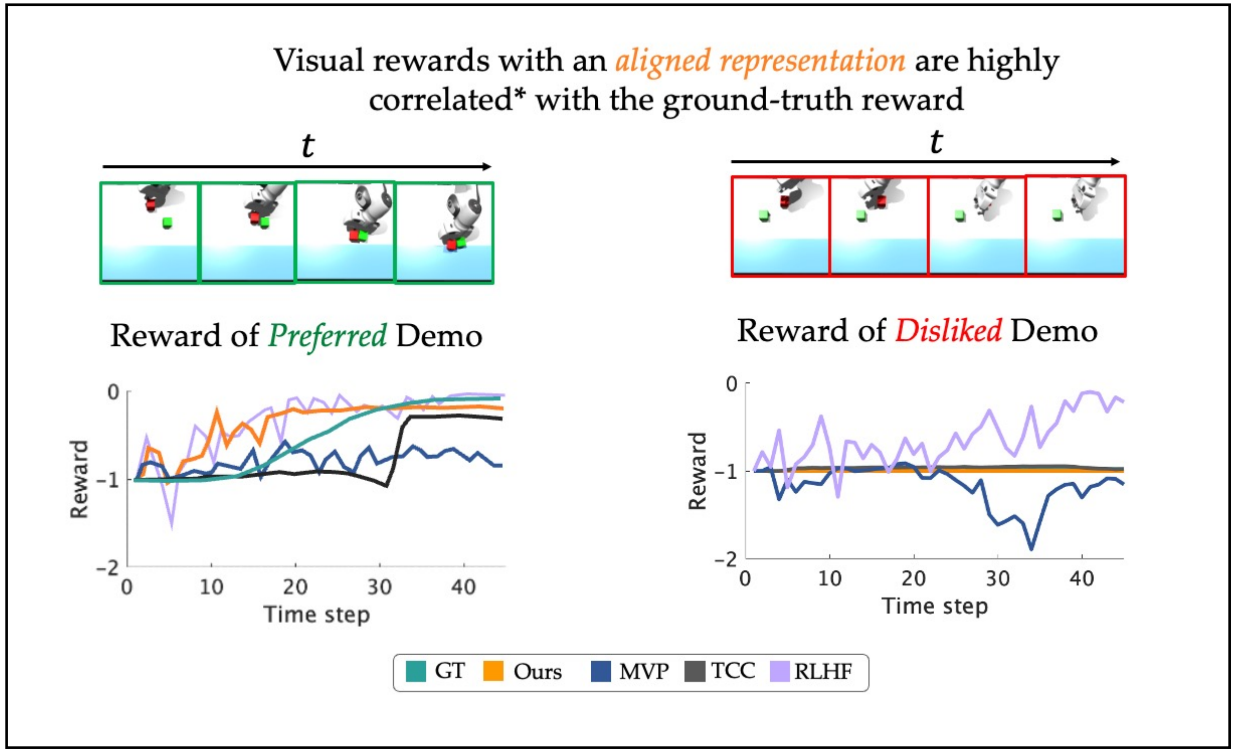
$$f(x,u,d)$$

Player 1 *Player 2*

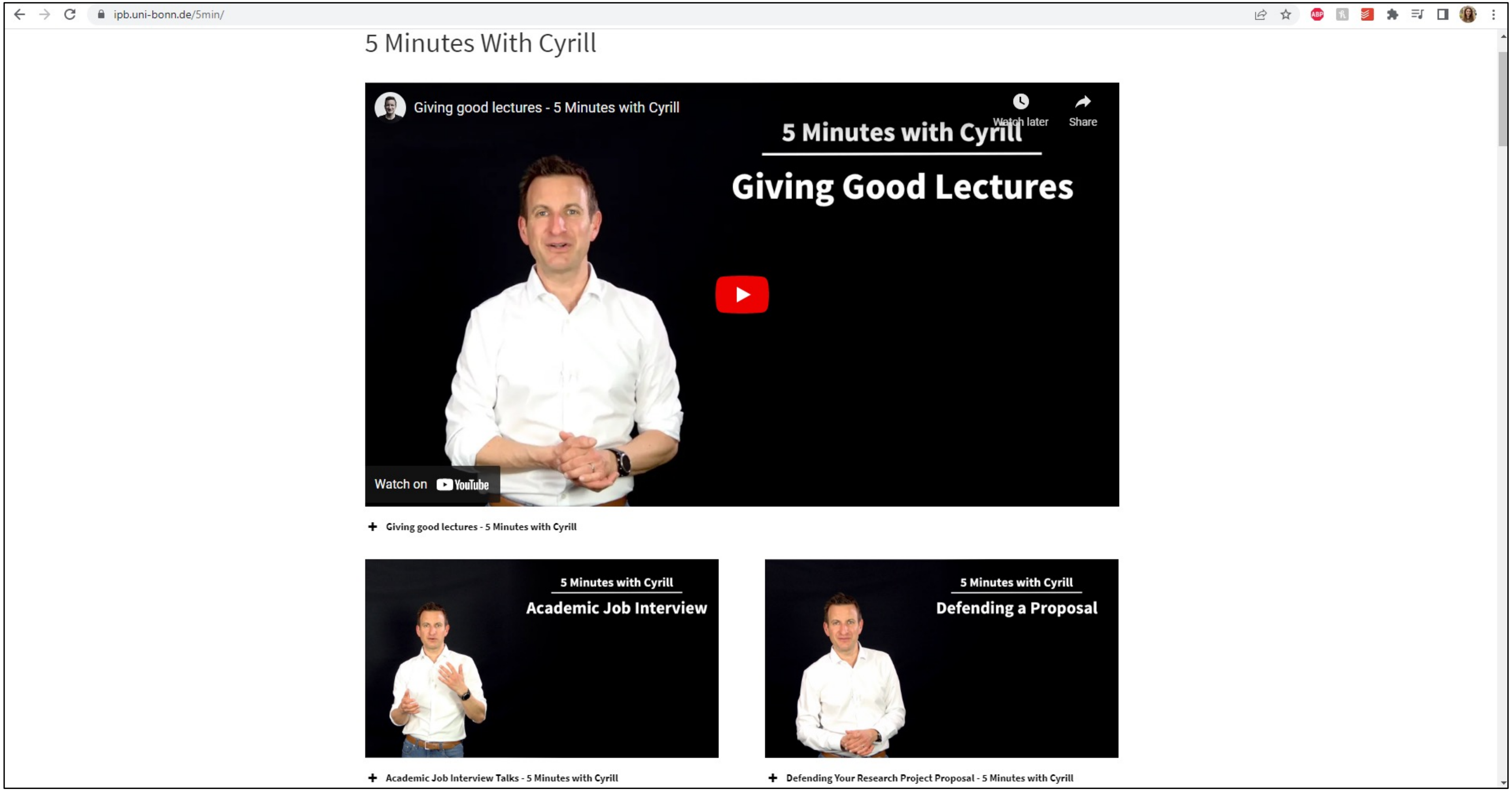
Heuristics for presentations

Use useful titles





Academic-Specific Resource

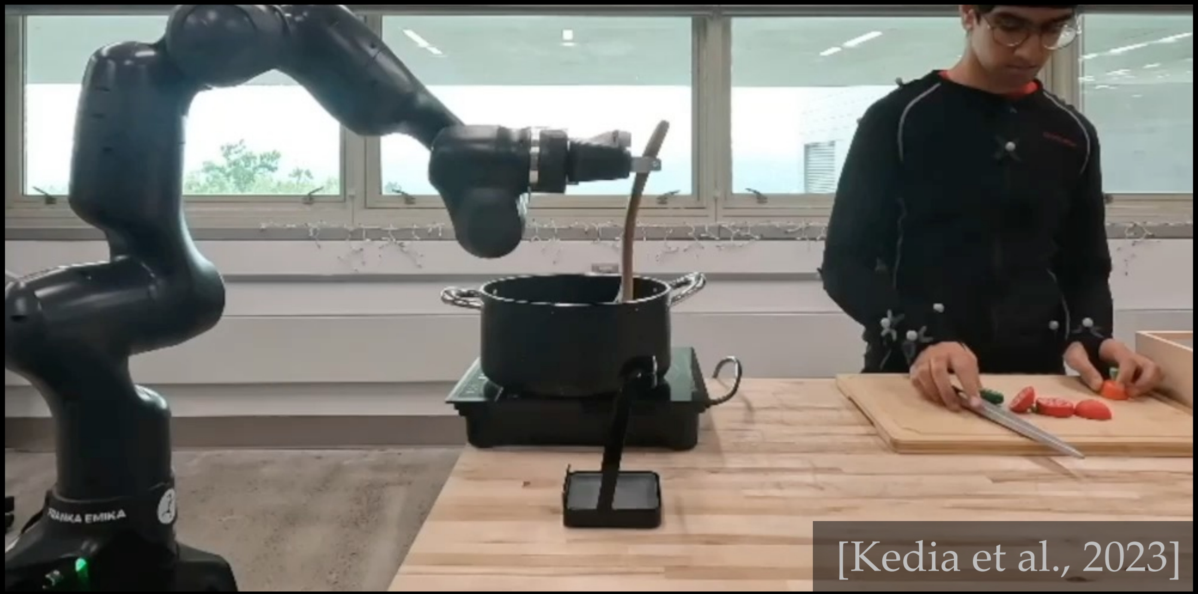


Retrospective & Frontiers


[Waymo, 2023]

[Ren, AZ et al., 2023]

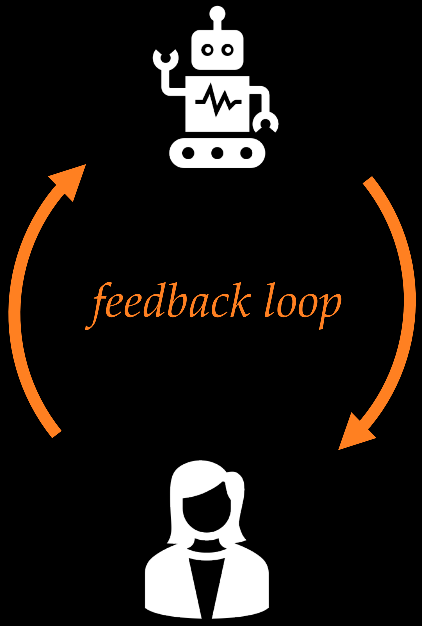
AI is enabling autonomous agents to interact with people at scale


[Kedia et al., 2023]


[DeepMind, 2023]

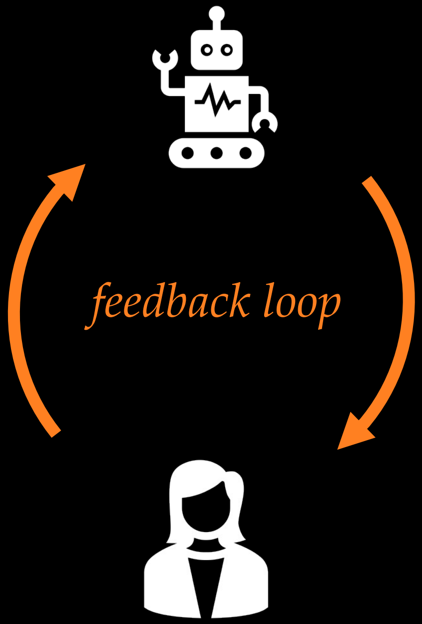
This widespread human—AI interaction has also increased questions about modeling interaction and raised safety & alignment concerns ….

Interaction means there exists a **feedback loop** between human stakeholders and autonomous robots



Influences:

representations
robot decisions
human responses

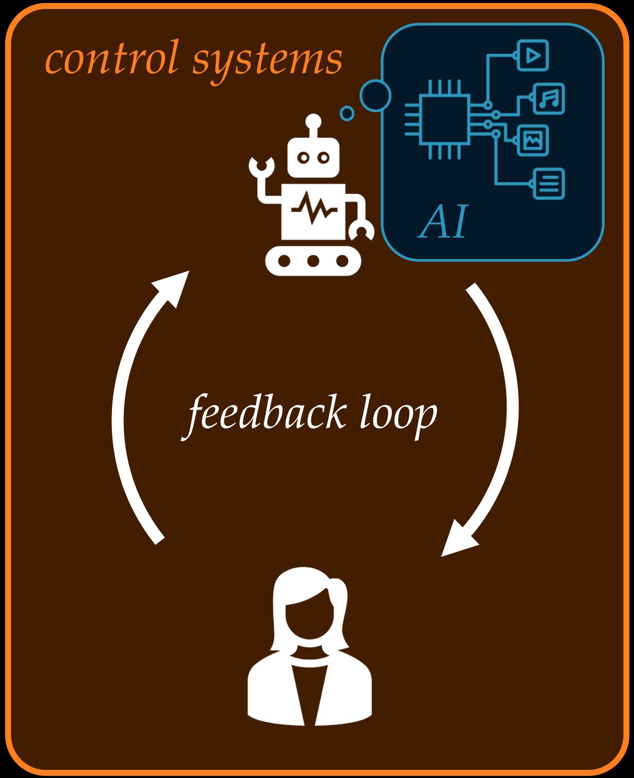


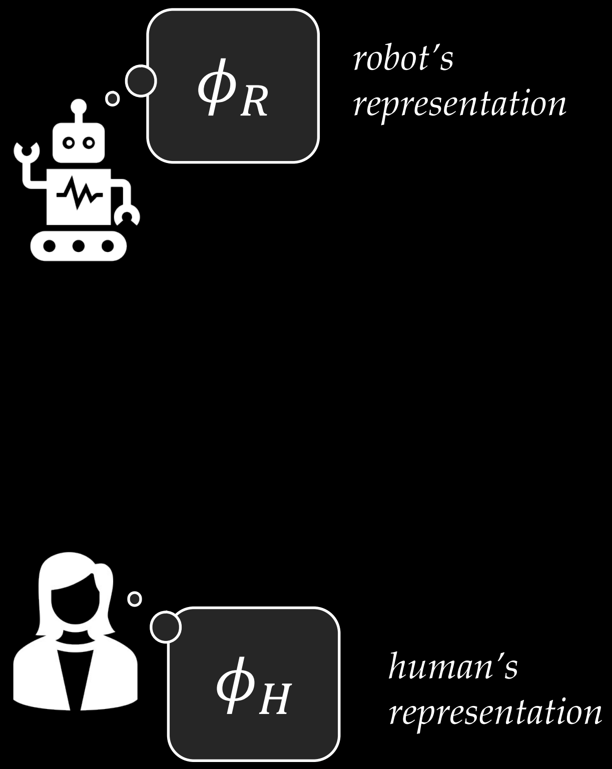
Present at:

training
fine-tuning
deployment

In this class:

formalisms inspired by control & dynamical systems to model **human—robot/AI feedback loops** influenced by robot decisions

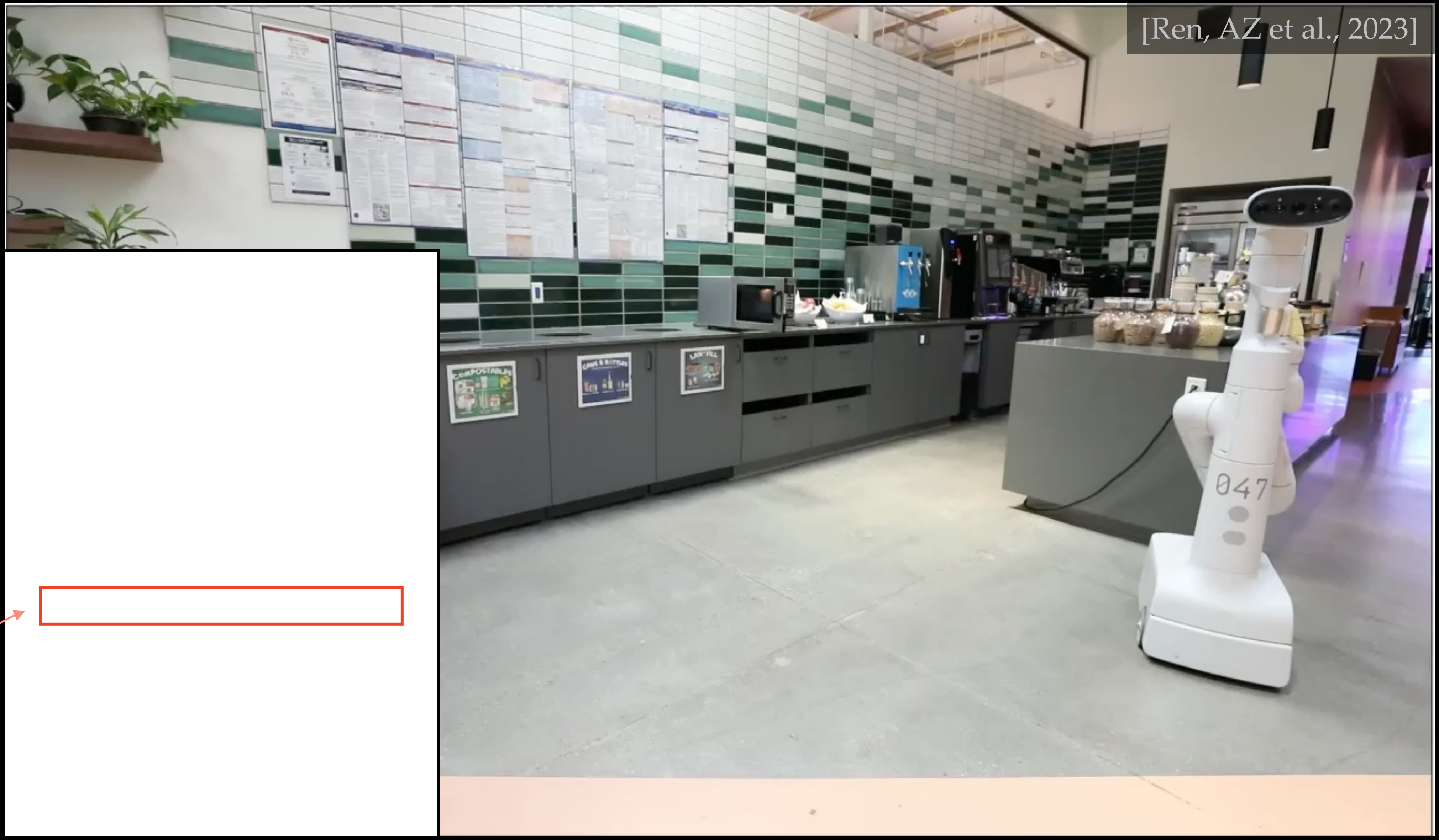




ϕ_R
robot's representation
ϕ_H
human's representation

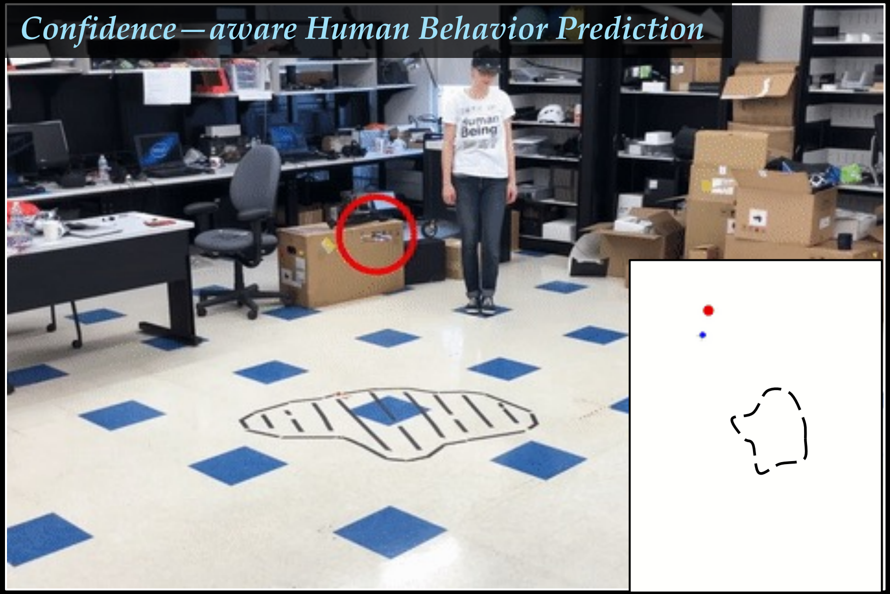
[Autoevolution, "Tesla with FSD Beta Avoids Plastic Bag, Some People Mistake It for a Good Thing"]

Robot's and end-user's *representations* are misaligned!

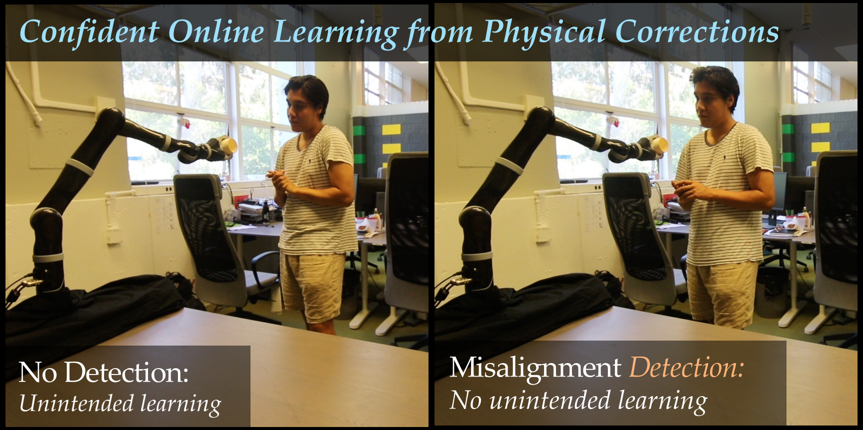


*unsafe to put **metal** in microwave*

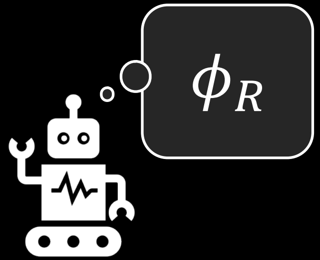
Detecting Representation Misalignment

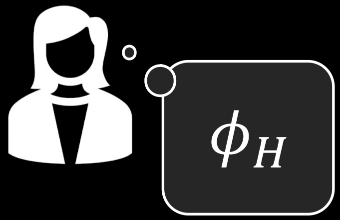


Bajcsy, A. et al. "Confidence-aware motion prediction for real-time collision avoidance." IJRR 2020.*



Bobu, A. et al. "Learning under misspecified objective spaces." CoRL 2018.

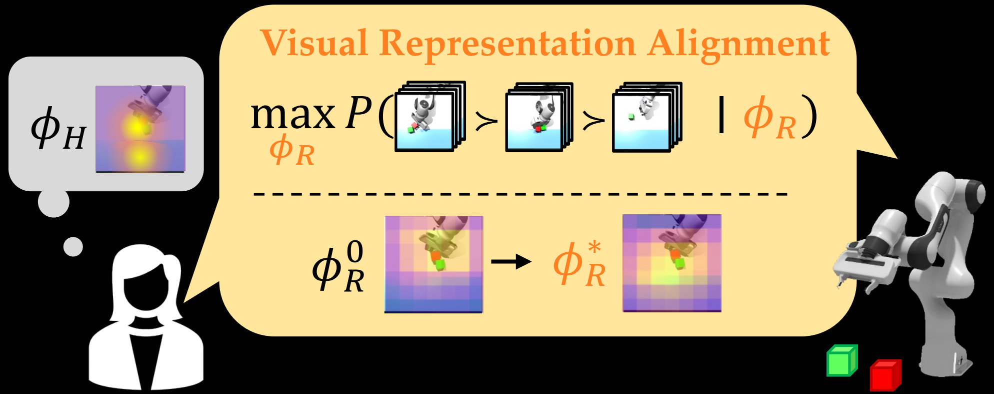




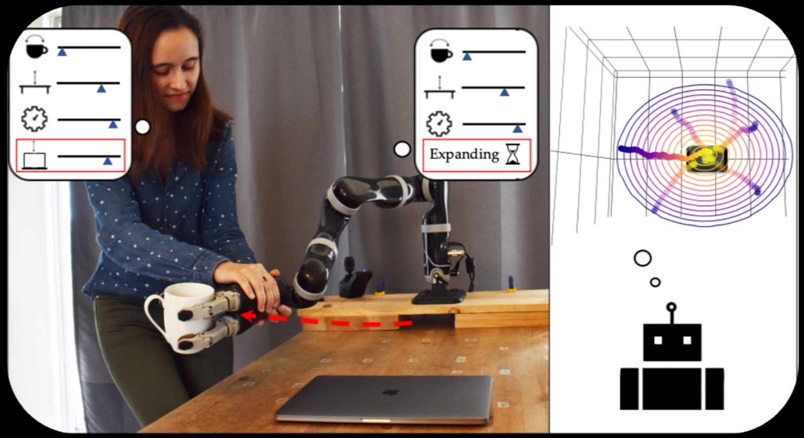
Aligning Robot Representations

Aligning Pre-trained Vision Models w/ Human Feedback

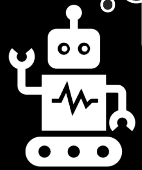
Tian, R. et al. "What Matters to You? Towards Visual Representation Alignment for Robot Learning." ICLR 2024.



New types of human feedback for representation alignment



Bobu, A. et al. "Feature Expansive Reward Learning: Rethinking Human Input." HRI 20201



$\phi_R \mid \pi_R : \phi_R \rightarrow a_R$

physical action, generations for human to rank, …

robot's policy

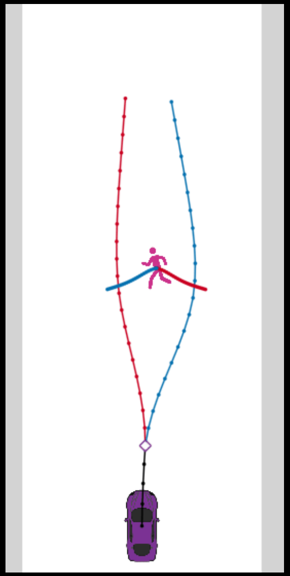
$\phi_H \mid \pi_H : \phi_H \rightarrow a_H$

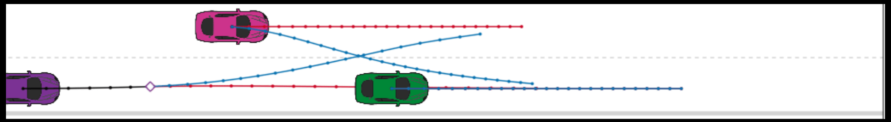
physical action, preference feedback, text prompt…

human's policy

Game-Theoretic & Data-Driven Interaction Models

N-Player Dynamic Games



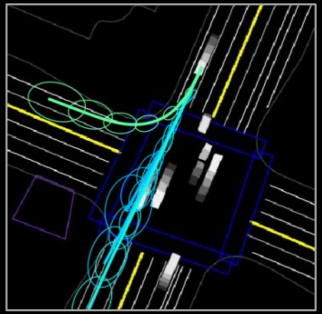


Peters, L. et al. "Contingency Games for Multi-Agent Interaction" RA-L, 2024.

Learning from Corrections

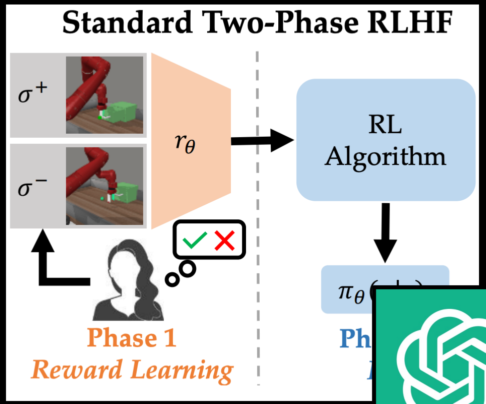
Bajcsy et al., "Learning from pHRI", CoRL 2017

Deep Human Behavior Prediction



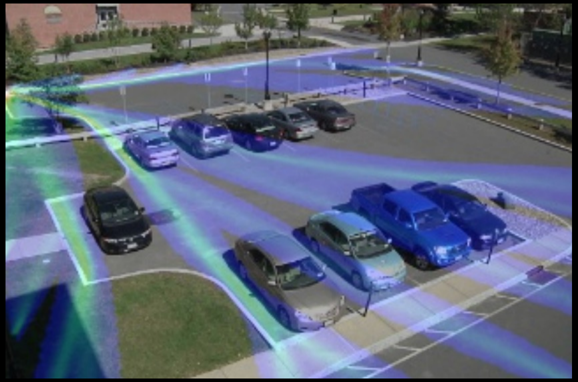
MultiPath, Waymo

RLHF / DPO / CPL

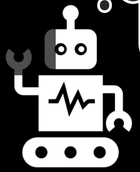


Hejna, J. et al. "Contrastive Preference Learning". ICLR 2024

Inverse Reinforcement Learning



Kitani, et al. Activity Forecasting. ECCV 2012

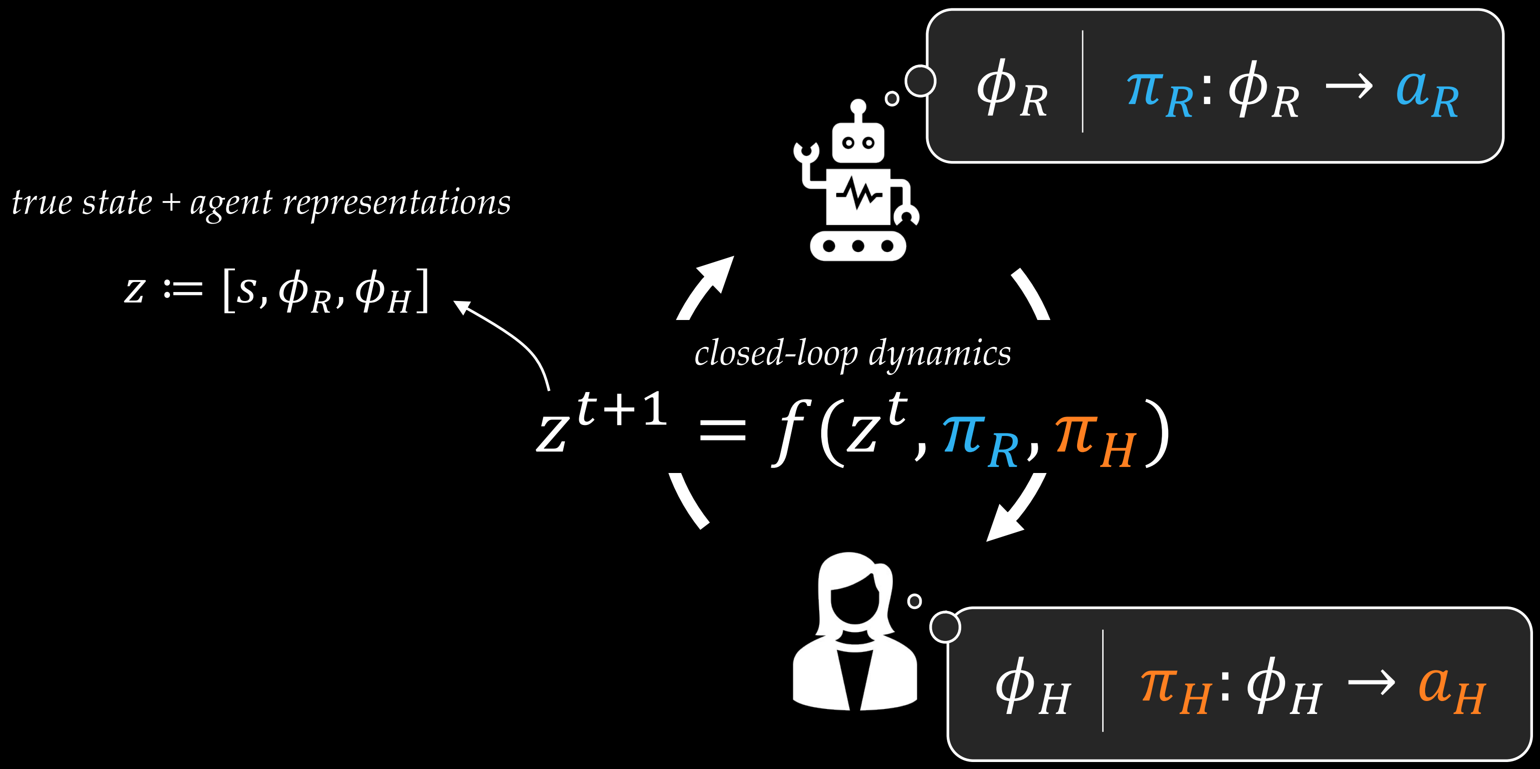


ϕ_R | $\pi_R : \phi_R \rightarrow a_R$

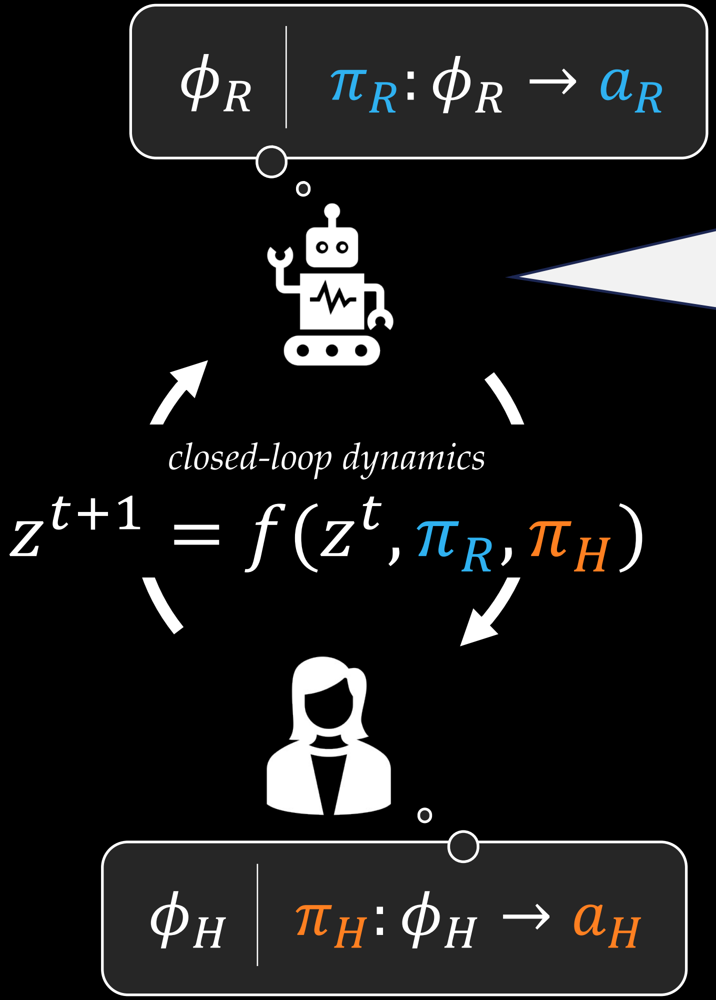
robot's policy

ϕ_H | $\pi_H : \phi_H \rightarrow a_H$

human's policy


true state + agent representations
$z := [s, \phi_R, \phi_H]$
closed-loop dynamics
$z^{t+1} = f(z^t, \pi_R, \pi_H)$
ϕ_R | $\pi_R : \phi_R \to a_R$
ϕ_H | $\pi_H : \phi_H \to a_H$

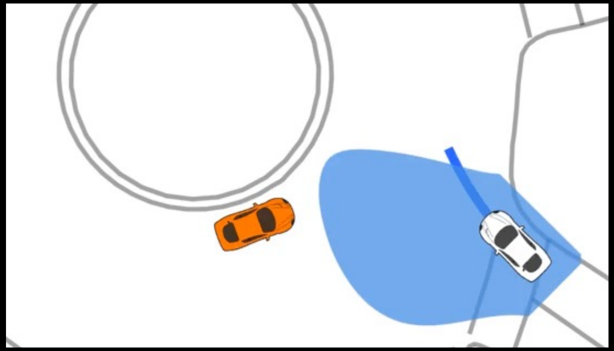
Safety Analysis & Runtime Safety Filtering



fundamental problem:
present actions which do not appear to violate constraints can still steer the system to states of irrecoverable failure in the *future*.

Safety Analysis & Runtime Safety Filtering

Game-theoretic Runtime Safety Filters



Tian, R. et al. "Safety assurances for human-robot interaction via confidence-aware game-theoretic human models." ICRA 2022.

Computationally scalable & data-driven safety

Wabersich, K. P., et al. "Data-driven safety filters" Control Systems Magazine. 2023.

ϕ_R | $\pi_R : \phi_R \rightarrow a_R$

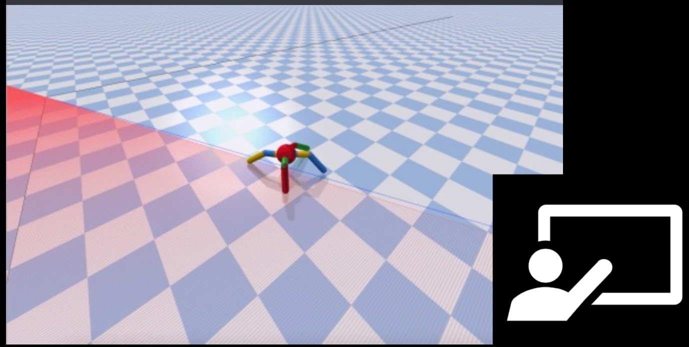
closed-loop dynamics

$$z^{t+1} = f(z^t, \pi_R, \pi_H)$$

ϕ_H | $\pi_H : \phi_H \rightarrow a_H$

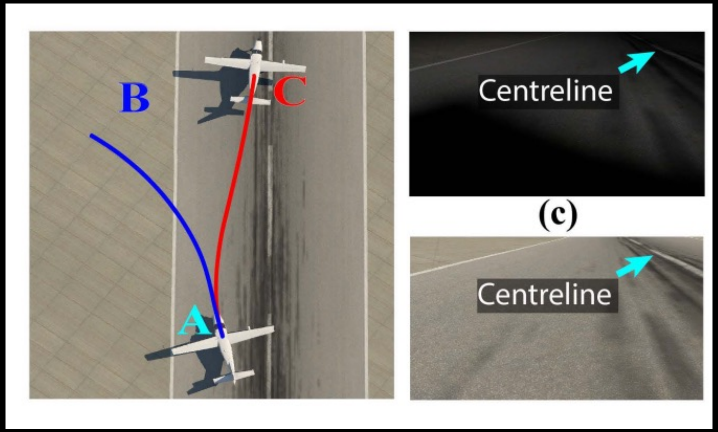
Learning Safety from Demos & Latent-space Safety

Using multi-task data to improve constraint inference.



Kim, K, et al. "Learning shared safety constraints from multi-task demonstrations." NeurIPS 2024.

Closed-loop Failures of Vision-based Controllers

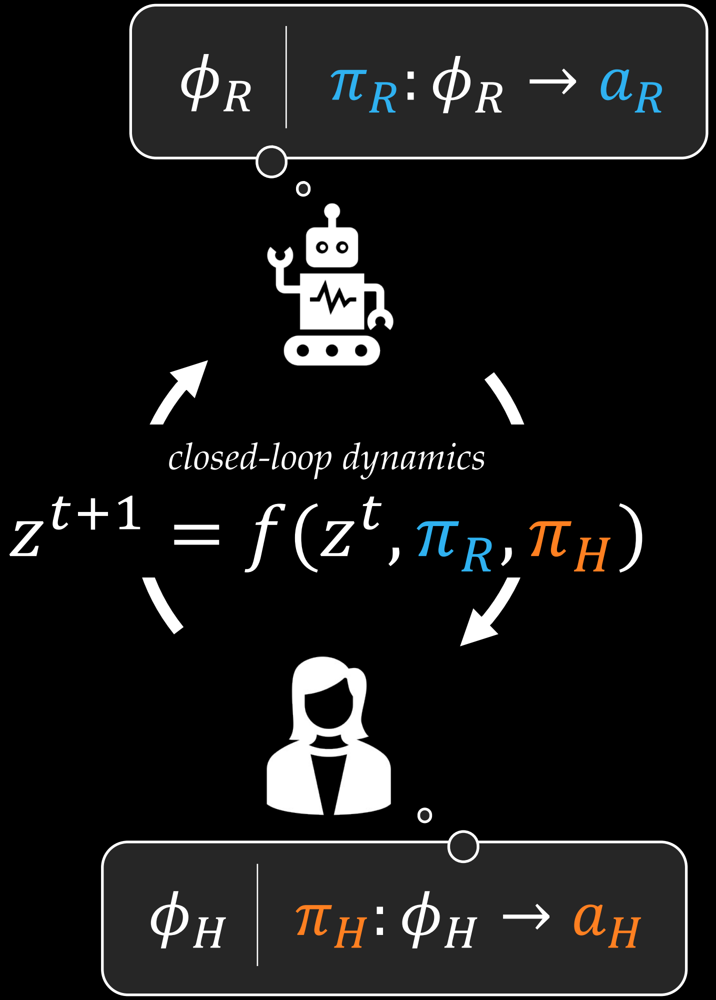


Chakraborty, K. and Bansal, S.. "Discovering Closed-Loop Failures of Vision-Based Controllers via Reachability Analysis." RA-L 2023

Detecting Representation Misalignment

Safety Analysis & Runtime Safety Filtering

Aligning Robot Representations



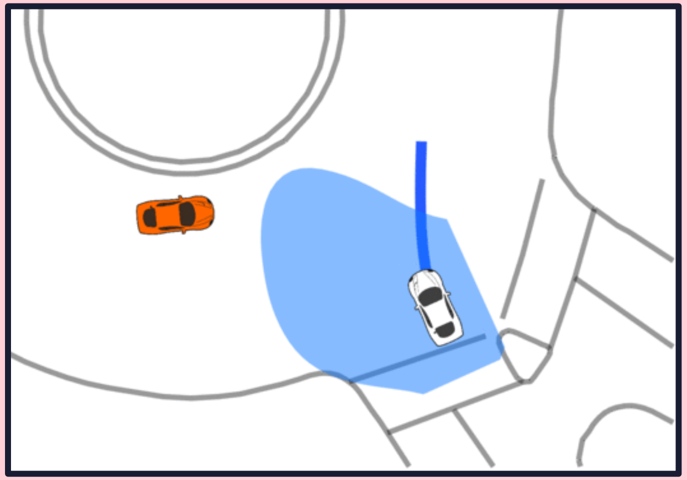
Game-Theoretic & Data-Driven Interaction Models

Learning Safety from Demos & Latent-space Safety

….and more!

Control Systems

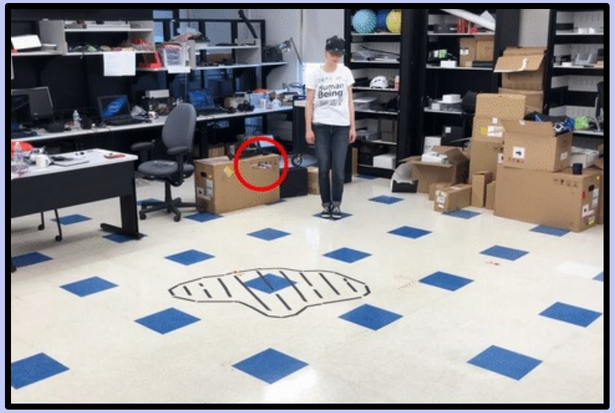
(Safety)



- ☑ *Safety monitoring* coupled with *action (mitigation)*
- ☑ Multi-agent feedback loops
- ☒ Limited to low-D representations & behavior models…

ML / AI

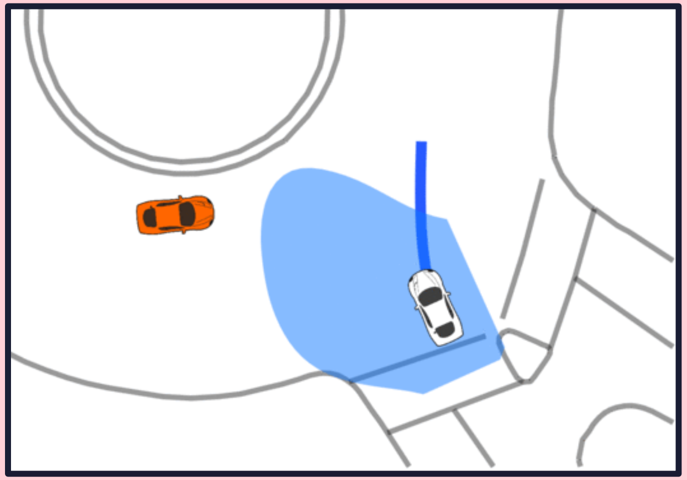
(OOD)



- ☒ Safety at a *component-level* & decoupled from mitigation
- ☑ Expressive behavior generation, multimodal human feedback
- ☑ Rich context & representations

Control Systems

(Safety)

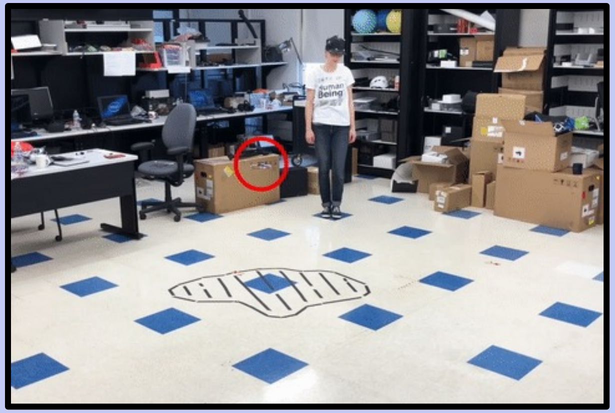


- ☑ *Safety monitoring* coupled with *action (mitigation)*
- ☑ Multi-agent feedback loops
- ☒ Limited to low-D representations & behavior models…

“System-level” approach to human-AI safety

ML / AI

(OOD)



- ☒ Safety at a *component-level* & decoupled from mitigation
- ☑ Expressive behavior generation, multimodal human feedback
- ☑ Rich context & representations

2022

A System-Level View on Out-of-Distribution Data in Robotics

Rohan Sinha, Apoorva Sharma, Somrita Banerjee, Thomas Lew, Rachel Luo,
Spencer M. Richards, Yixiao Sun, Edward Schmerling, Marco Pavone

***Abstract*— When testing conditions differ from those represented in training data, so-called out-of-distribution (OOD) inputs can mar the reliability of learned components in the modern robot autonomy stack. Therefore, coping with OOD data is an important challenge on the path towards trustworthy learning-enabled open-world autonomy. In this paper, we aim to demystify the topic of OOD data and its associated challenges in the context of data-driven robotic systems, drawing connections to emerging paradigms in the ML community that study the effect of OOD data on learned models in isolation. We argue that as roboticists, we should reason about the overall *system-level* competence of a robot as it operates in OOD conditions. We highlight key research questions around this system-level view of OOD problems to guide future research toward safe and reliable learning-enabled autonomy.**

I. Introduction

Machine learning (ML) systems are poised for widespread usage in robot autonomy stacks in the near future, driven by the successes of modern deep learning. For instance, decision-making algorithms in autonomous vehicles rely on ML-based perception and prediction models to estimate and forecast the state of the environment. As we increasingly rely on ML models to contend with the unstructured and unpredictable real world in robotics, it is paramount that we also acknowledge the shortcomings of our models, especially when we hope to deploy robots alongside humans in safety-critical settings.

In particular, ML models may behave unreliably on data that is dissimilar from the training data — inputs commonly termed *out-of-distribution* (OOD). This poses a significant challenge to deploying robots in the open world, e.g., as autonomous vehicles or home assistance robots, as such robots must interact with complex environments in conditions we cannot control or foresee. Coping with OOD inputs remains a key and largely unsolved challenge on the critical path to reliable and safe open-world autonomy. However, there is no generally-agreed-upon precise definition of what makes data OOD; instead, its definition is often left implicit and varies between problem formalisms and application contexts.

In this paper, we concretize the often nebulous notion of the OOD problem in robotics, drawing connections to existing approaches in the ML community. Critically, we advocate for a *system-level* perspective of OOD data in robotics, which considers the impacts of OOD data on downstream decision making and leverages components throughout the full autonomy stack to mitigate negative consequences. To this end, we present robotics research challenges at three timescales crucial to deploying reliable open-world autonomy: (i) real-time decision-making, (ii) episodic interaction with an environment, and (iii) the data lifecycle as learning-enabled robots are deployed, evaluated, and retrained.

We note that this paper represents neither an algorithmic contribution nor a comprehensive survey of existing paradigms and literature on OOD topics in machine learning or robotics; in fact, many of the OOD topics that we discuss, like runtime-monitoring of perception systems [1] or heuristic uncertainty quantification of deep neural networks [2], constitute well-surveyed subfields in their own right. Rather than survey specific styles of analysis or approaches tailored towards particular submodules of the autonomy stack, our goal in this work is to provide an overview of the core considerations and system-wide challenges that we see as essential areas of robotics research activity for the coming years. Our contribution thus is to establish perspective and context to galvanize more research interest in a topic that we view as critical to improving the reliability of autonomous robots.

II. Running Examples

To better describe the challenges that OOD data creates in learning-enabled robotic systems, we use the two future autonomy systems shown in Figure 1 as running examples in this paper. These conceptual examples highlight the plurality of applications and design paradigms used to leverage ML in the design of robotic systems.

Autonomous Drone Delivery Service: Firstly, we consider an autonomous drone delivering packages in a city. As illustrated in Figure 1, this robot uses several learning-enabled components in its autonomy stack. The delivery drone has to make explainable decisions and meet stringent safety requirements by regulatory agencies to be deployed among humans. Crucially, to maintain these reliability requirements in rare and unforeseen circumstances, the drone needs mechanisms to detect and manage OOD inputs.

Robotic Manipulators Assisting in the Home: Secondly, we consider the deployment of robotic manipulators to assist with various tasks in and around the home, as shown in Figure 1. The manipulators' tasks are so diverse and unstructured that we consider a general manipulation policy trained in an end-to-end fashion in a controlled environment, as commonly considered in the reinforcement learning (RL) community. When we deploy these manipulators in people's homes, the environments and contexts that these robots encounter invariably differ from the lab or simulated environments

The authors are with the Autonomous Systems Lab at Stanford University, Stanford, CA. {rhnsinha, apoorva, somrita,

RQ 1 (Averting OOD failures through Runtime Monitoring). Can we leverage *full-stack* sensory information at runtime to detect if a decision system relying on a learned model f will perform poorly, before a failure occurs?

RQ 2 (OOD Aware Decision Making). Can we design decision-making systems compatible with runtime monitors robust to high functional uncertainty?

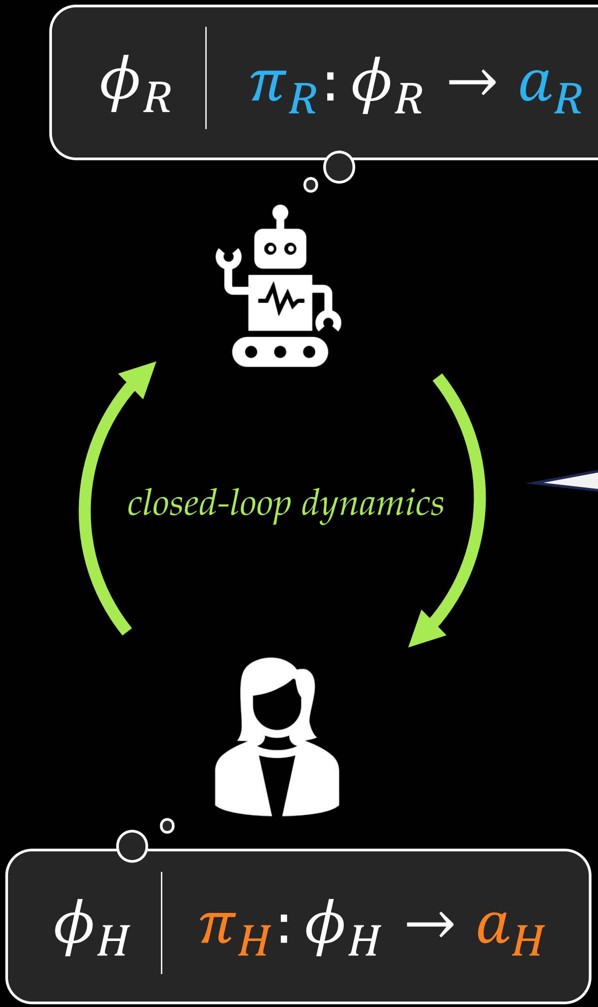
RQ 3 (Temporally Correlated OOD events). Can we develop methods that account for the temporal correlations between inputs when we repeatedly evaluate a learned model f under shifted conditions over the course of an episode?

RQ 4 (Mitigating Distributional Shifts). Can we construct decision-making algorithms that mitigate distributional shifts between the training and deployment conditions to ensure the overall reliability of the deployed system?

RQ 5 (Leveraging Operational Data). How can we use data collected during operation in diverse tasks and contexts to improve the robustness and quality of learned models?

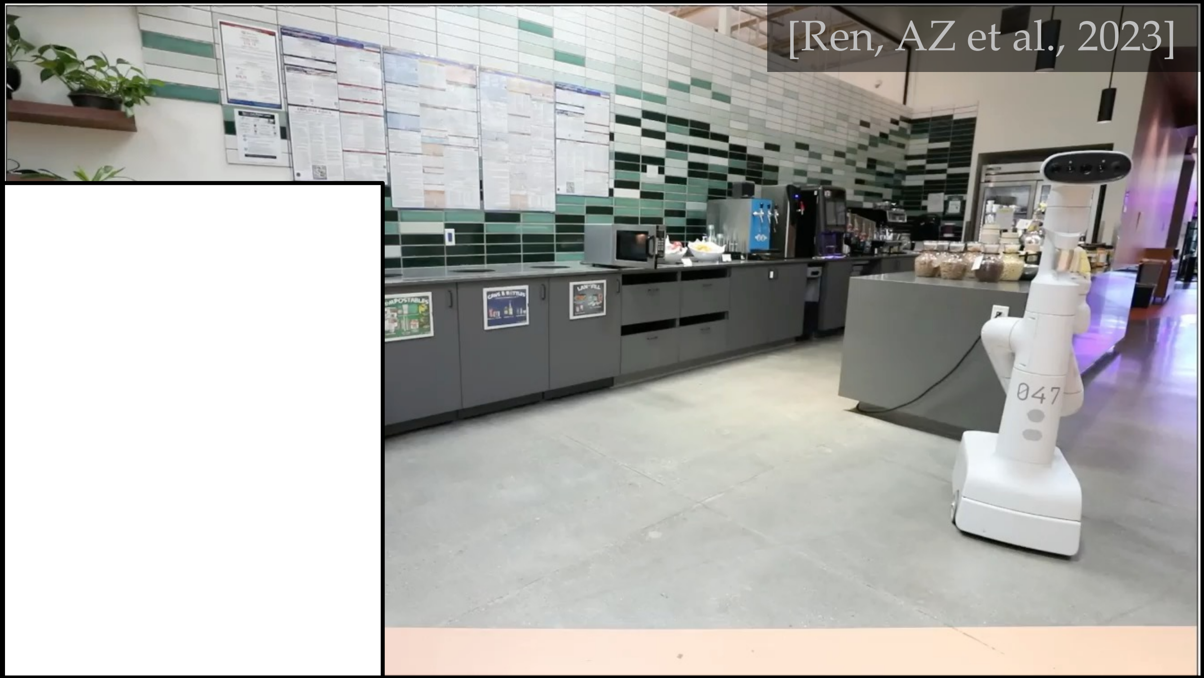
What does this frontier look like in interactive robotics?

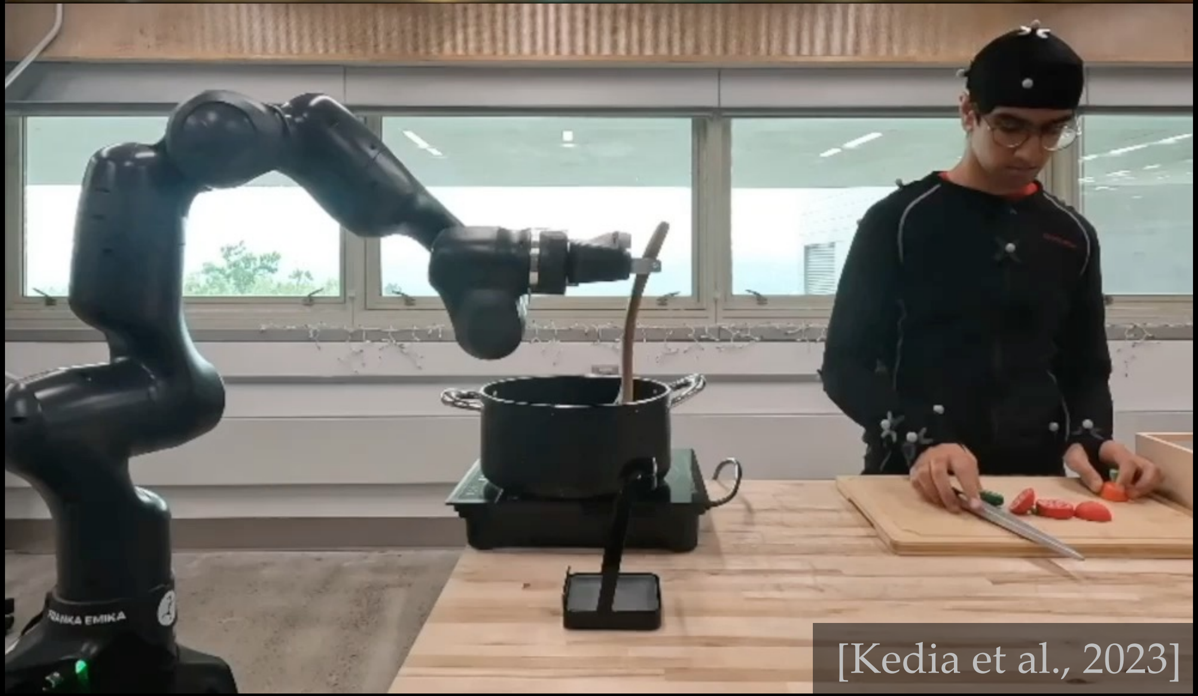
RQ 6 (Efficient Data Collection). How do we select what operational data to use to efficiently improve our models?



Operational life cycle!


[Waymo, 2023]


[Ren, AZ et al., 2023]


[Kedia et al., 2023]


[DeepMind, 2023]

Robot Planner

$$\boldsymbol{a}^{\mathrm{R}} = \pi_{\theta}(s, P_{\phi}, C)$$

cost-based *planner: θ is learned cost fn*
generative *planner: θ is NN weights*

Human Behavior Predictor

$$P_{\phi} : (\boldsymbol{s}, C) \mapsto \tilde{\boldsymbol{a}}^{\mathrm{H}}$$

state history & context

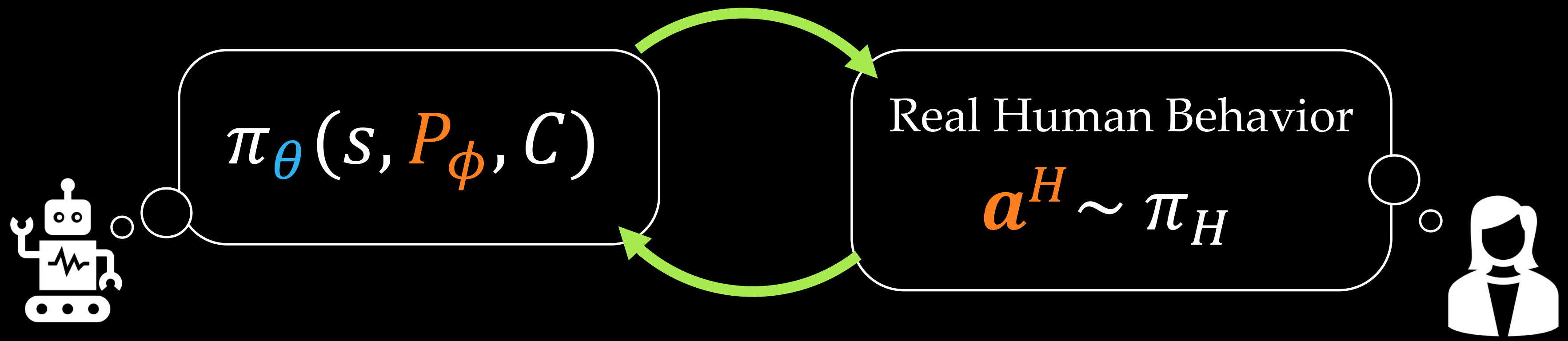
predicted future human actions

Real Human Behavior

$$\boldsymbol{a}^{H} \sim \pi_{H}$$

$$P_\phi : (s, C) \mapsto \tilde{a}^{\mathrm{H}}$$

Which prediction failures *"mattered"* during deployment interactions?
Can we *improve* our interaction models over repeated interactions?



All deployment data

Natural growth over time…

You

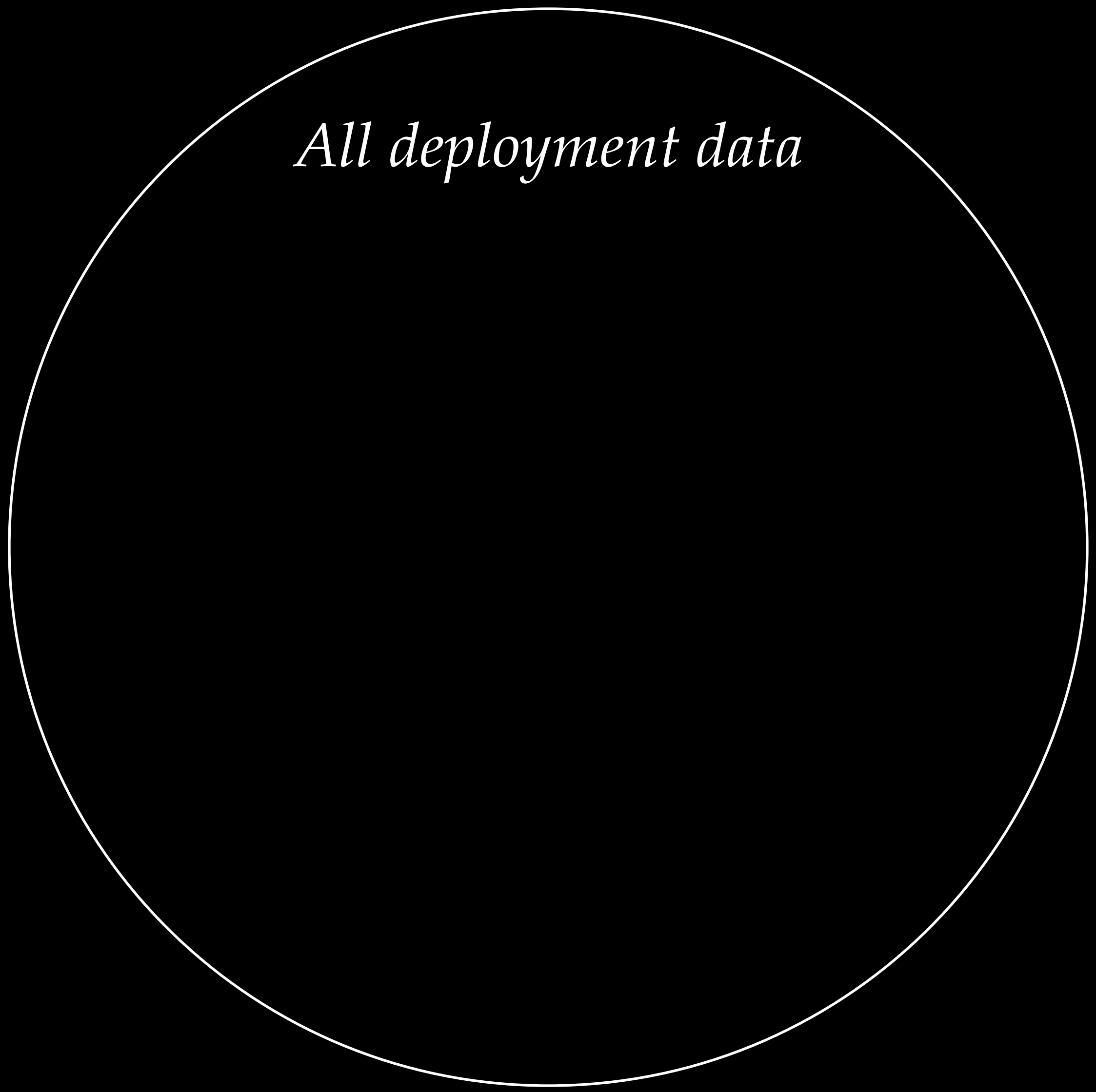
Tell me what happened on Tuesday February 27 2024 in Pittsburgh Pennsylvania, USA

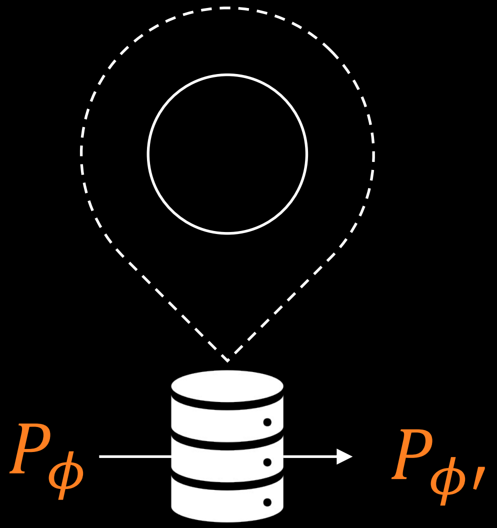
ChatGPT

I can't provide real-time or the most recent updates, including events from specific dates such as February 27, 2024, in Pittsburgh, Pennsylvania, or anywhere else, due to my last training data being in April 2023. To find out what happened on that specific date, I recommend checking the latest news from reliable sources or local news websites for Pittsburgh, such as the Pittsburgh Post-Gazette or Tribune-Review. You can also look for any updates on social media platforms or news aggregators for the most current information.

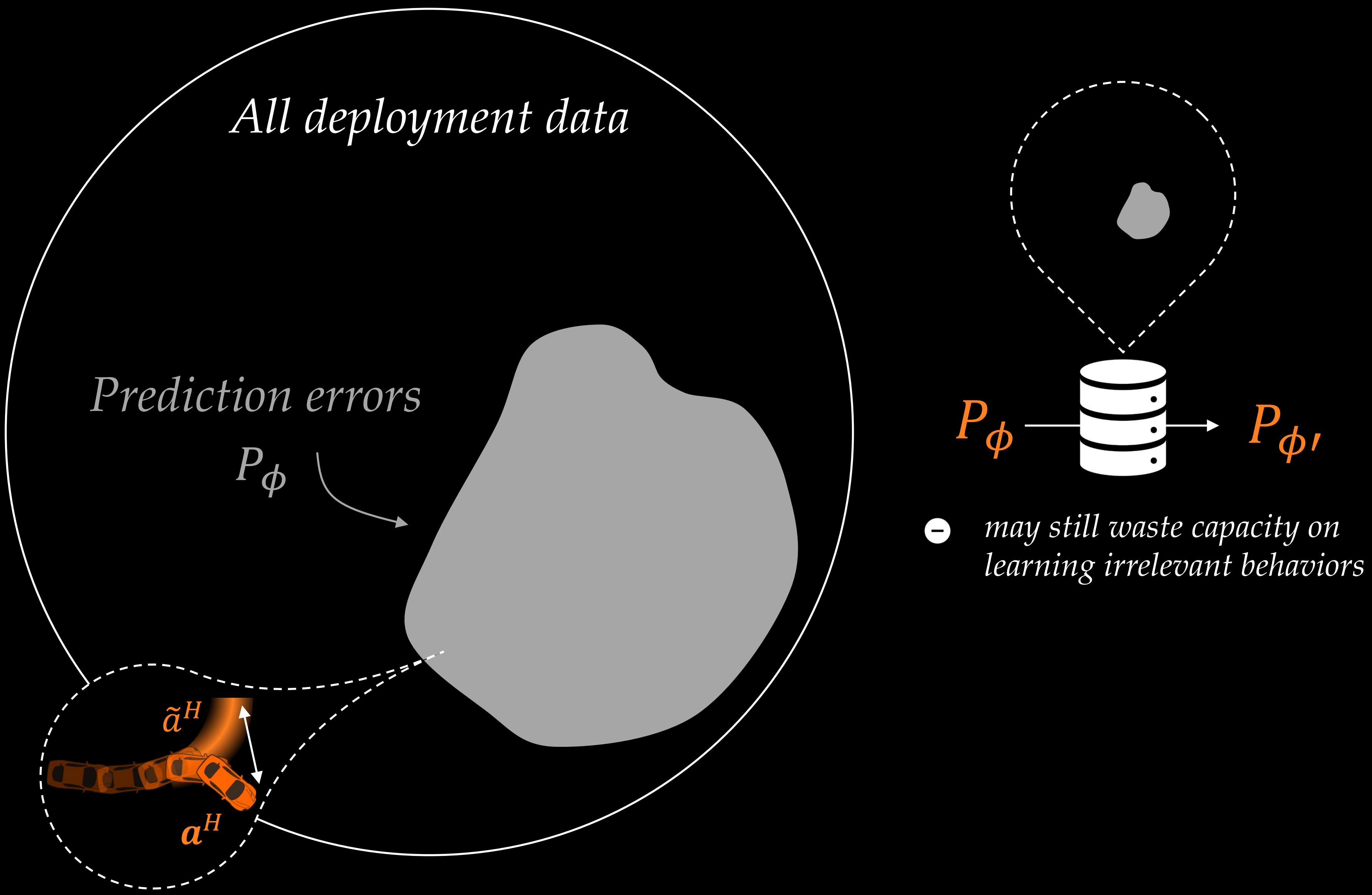
…& human interaction distribution shifts

Source: https://twitter.com/djbaskin

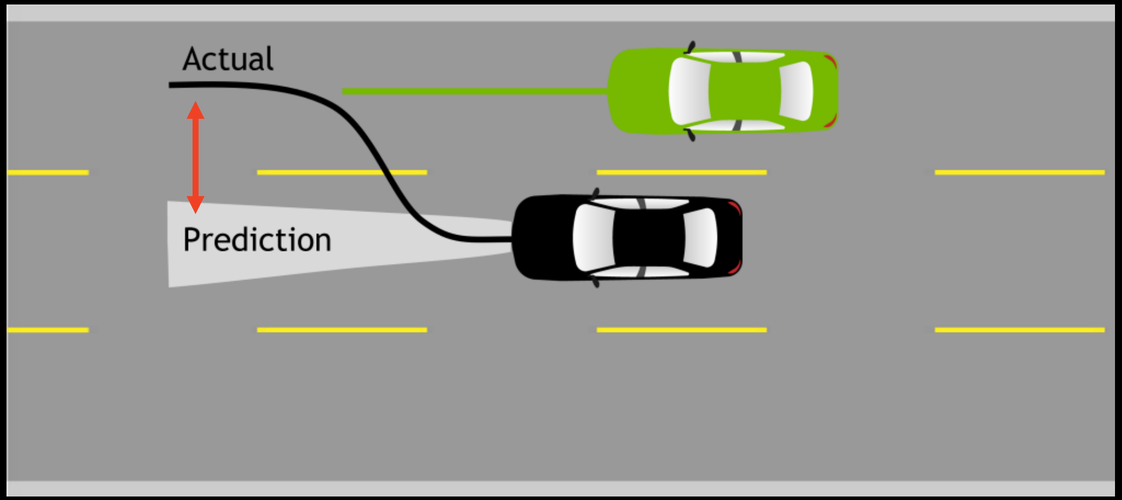




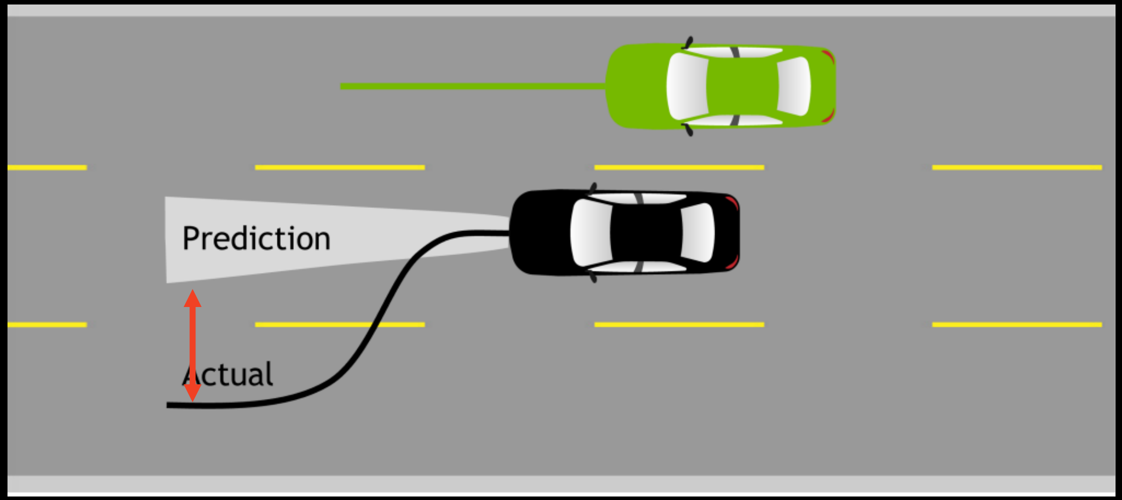
- *limited model capacity* *(esp. if you want fast inference)*


All deployment data
Prediction errors
P_ϕ
$\tilde{a}^H$
a^H
P_ϕ
$P_{\phi'}$
may still waste capacity on learning irrelevant behaviors

Mis-prediction DOES yields system-level failure



Mis-prediction does NOT yield system-level failure

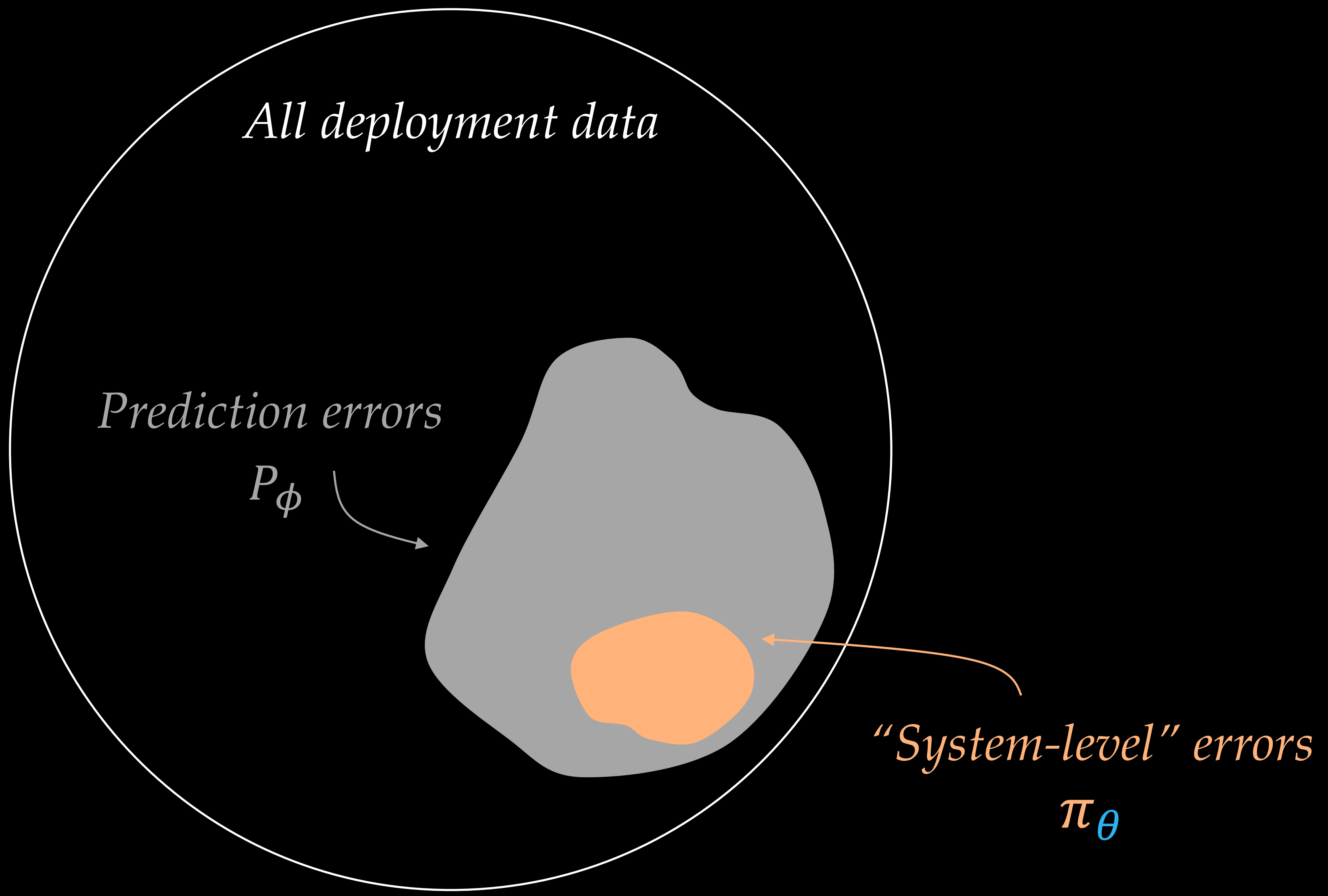


Both mis-predictions have the same L2 prediction error!

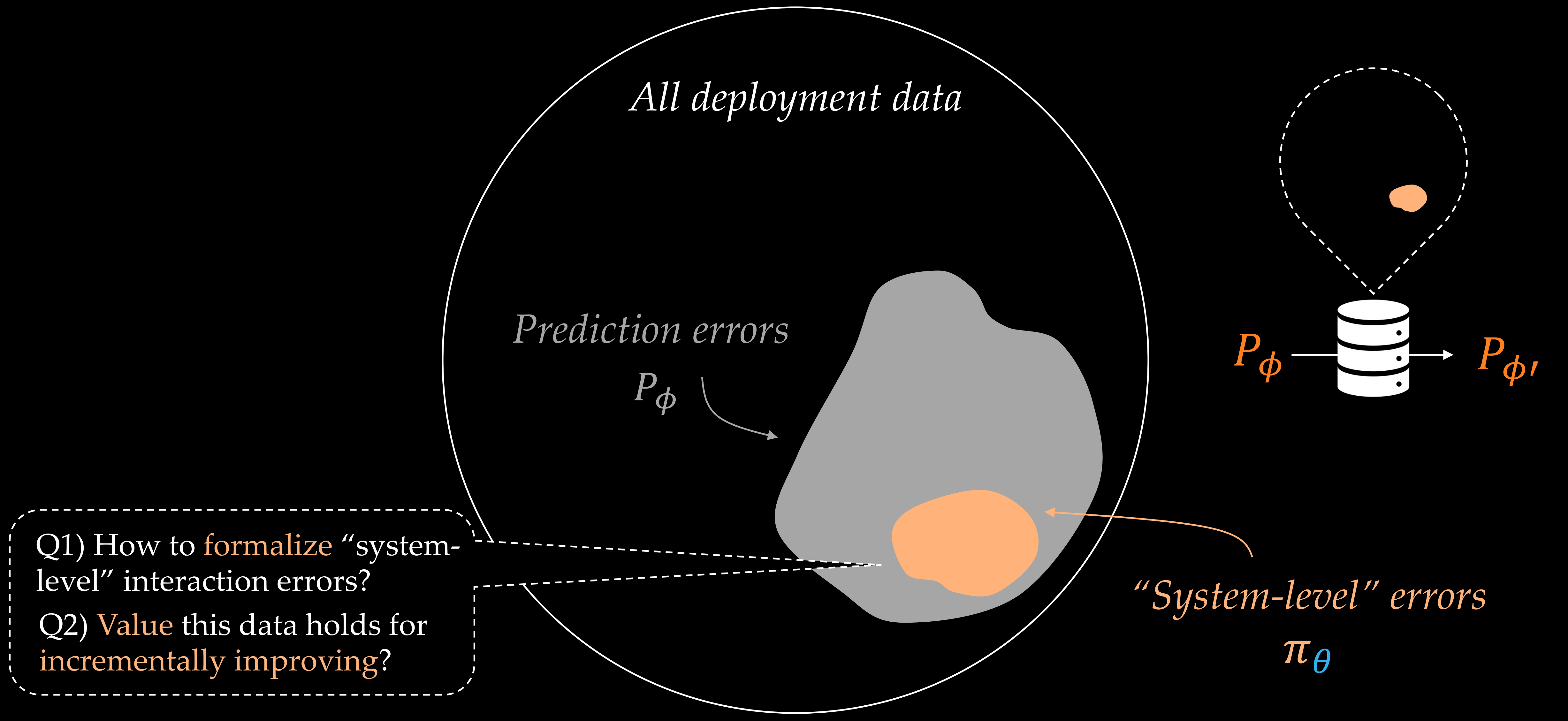
Farid, Alec, et al. "Task-relevant failure detection for trajectory predictors in autonomous vehicles." *CORL* 2023.

Source: https://www.youtube.com/watch?v=7s4-DVIs6Pk c


Google
10X



Sinha, Rohan, et al. "A System-Level View on Out-of Distribution Data in Robotics." *arXiv preprint* (2022).


All deployment data
Prediction errors
P_ϕ
"System-level" errors
π_θ
P_ϕ
$P_{\phi'}$
Q1) How to formalize "system-level" interaction errors?
Q2) Value this data holds for incrementally improving?

Idea 1

Regret precisely characterizes the degree to which prediction errors P_ϕ *degraded* robot performance π_θ

Nakamura, K. et al. "A General Calibrated Regret Metric for Detecting and Mitigating Human-Robot Interaction Failures". (under review) 2024.

Regret

robot's reward w/ parameters θ

$$\max_{\boldsymbol{a}^{\mathrm{R}}}\left[R_{\theta}^{\mathrm{R}}\left(\boldsymbol{a}^{\mathrm{R}}, \hat{\boldsymbol{a}}^{\mathrm{H}_1:\mathrm{H}_\mathrm{M}}, \hat{\boldsymbol{s}}, C\right)\right] - R_{\theta}^{\mathrm{R}}\left(\hat{\boldsymbol{a}}^{\mathrm{R}}, \hat{\boldsymbol{a}}^{\mathrm{H}_1:\mathrm{H}_\mathrm{M}}, \hat{\boldsymbol{s}}, C\right)$$

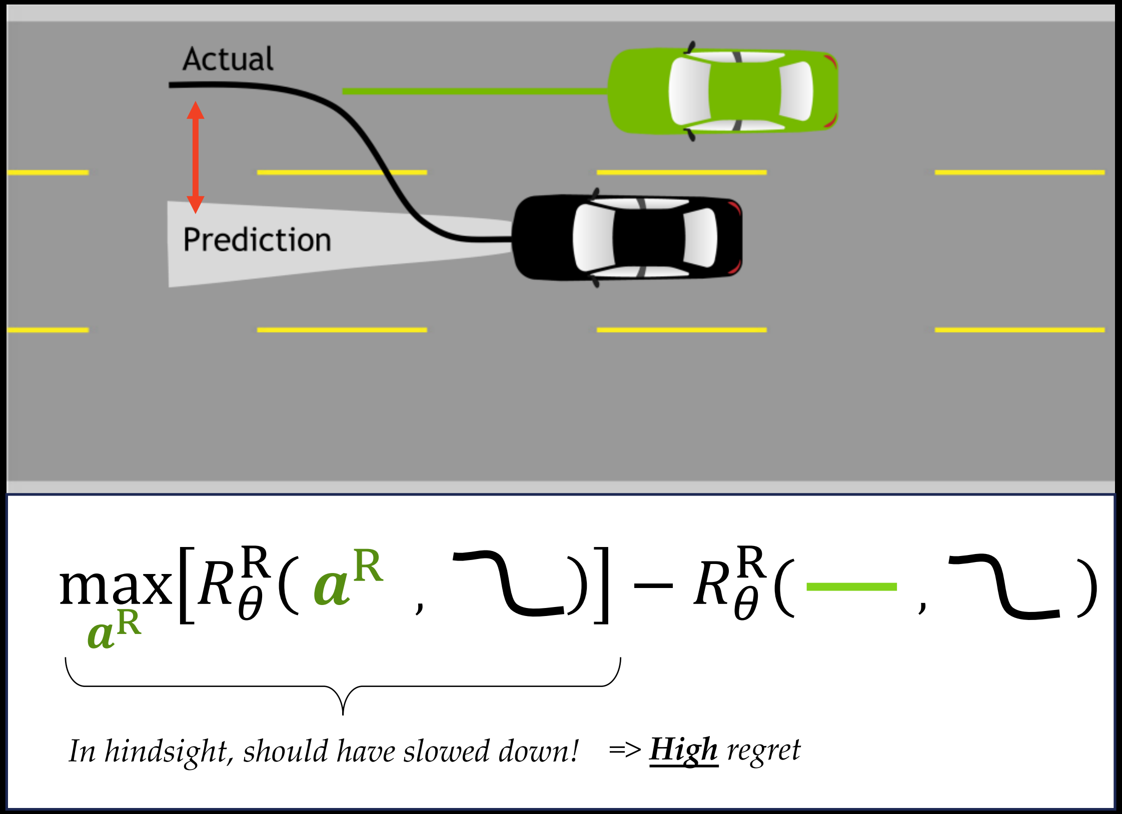
best robot decision <u>in hindsight</u>

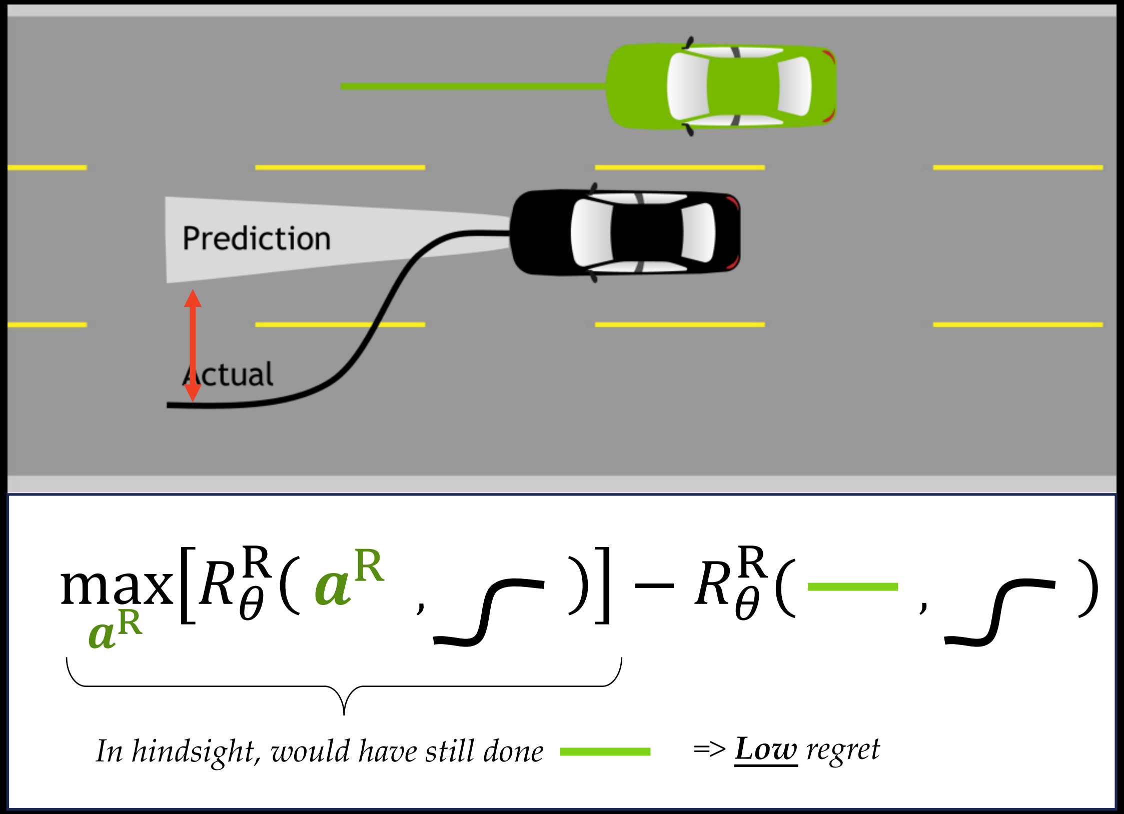
executed robot decision

observed human behavior, joint state, scene context

This mis-prediction would cause high regret!

This would not – robot would *still* do the same thing.





Standard regret can be mis-calibrated between disparate deployment contexts

Deployment Context A

Deployment Context B

Idea 2

evaluating the quality of a robot decision by its likelihood rather than its absolute reward calibrates regret across disparate contexts

Regret

$$\max_{\boldsymbol{a}^{\mathrm{R}}}\left[R_\theta^{\mathrm{R}}\left(\boldsymbol{a}^{\mathrm{R}}, \widehat{\boldsymbol{a}}^{\mathrm{H_1:H_M}}, \hat{\boldsymbol{s}}, C\right)\right] - R_\theta^{\mathrm{R}}\left(\widehat{\boldsymbol{a}}^{\mathrm{R}}, \widehat{\boldsymbol{a}}^{\mathrm{H_1:H_M}}, \hat{\boldsymbol{s}}, C\right)$$

Calibrated Generalized *Regret*

$$\max_{\boldsymbol{a}^{\mathrm{R}}}\left[P_\theta\left(\boldsymbol{a}^{\mathrm{R}} \mid \hat{\boldsymbol{a}}^{\mathrm{H}_1:\mathrm{H}_M}, \hat{\boldsymbol{s}}, C\right)\right] - P_\theta\left(\hat{\boldsymbol{a}}^{\mathrm{R}} \mid \hat{\boldsymbol{a}}^{\mathrm{H}_1:\mathrm{H}_M}, \hat{\boldsymbol{s}}, C\right)$$

params shared with planner

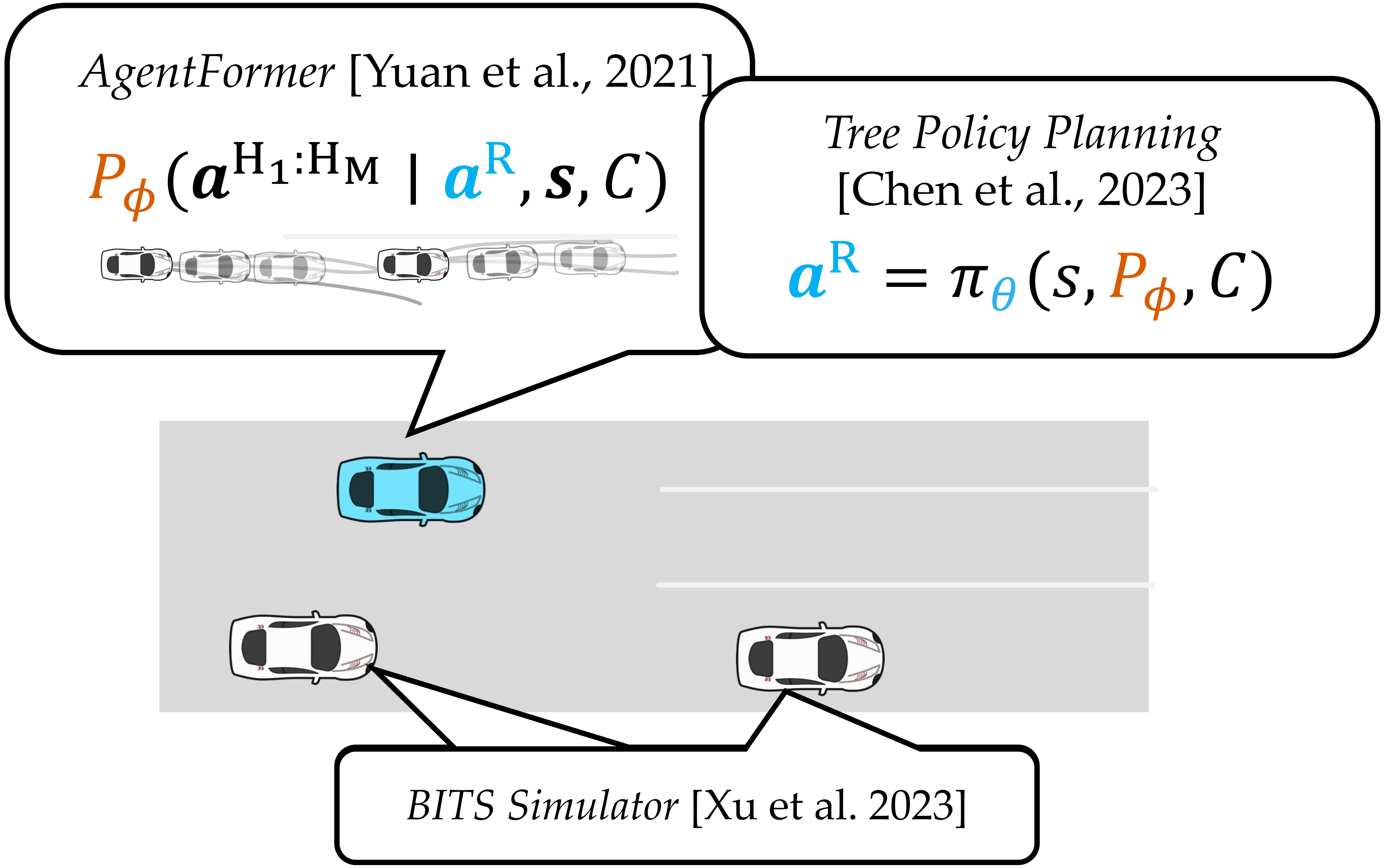
likelihood model for *counterfactual probability* of robot decisions *probability scales anomalies w.r.t. deployment context*

no need for explicit reward model!

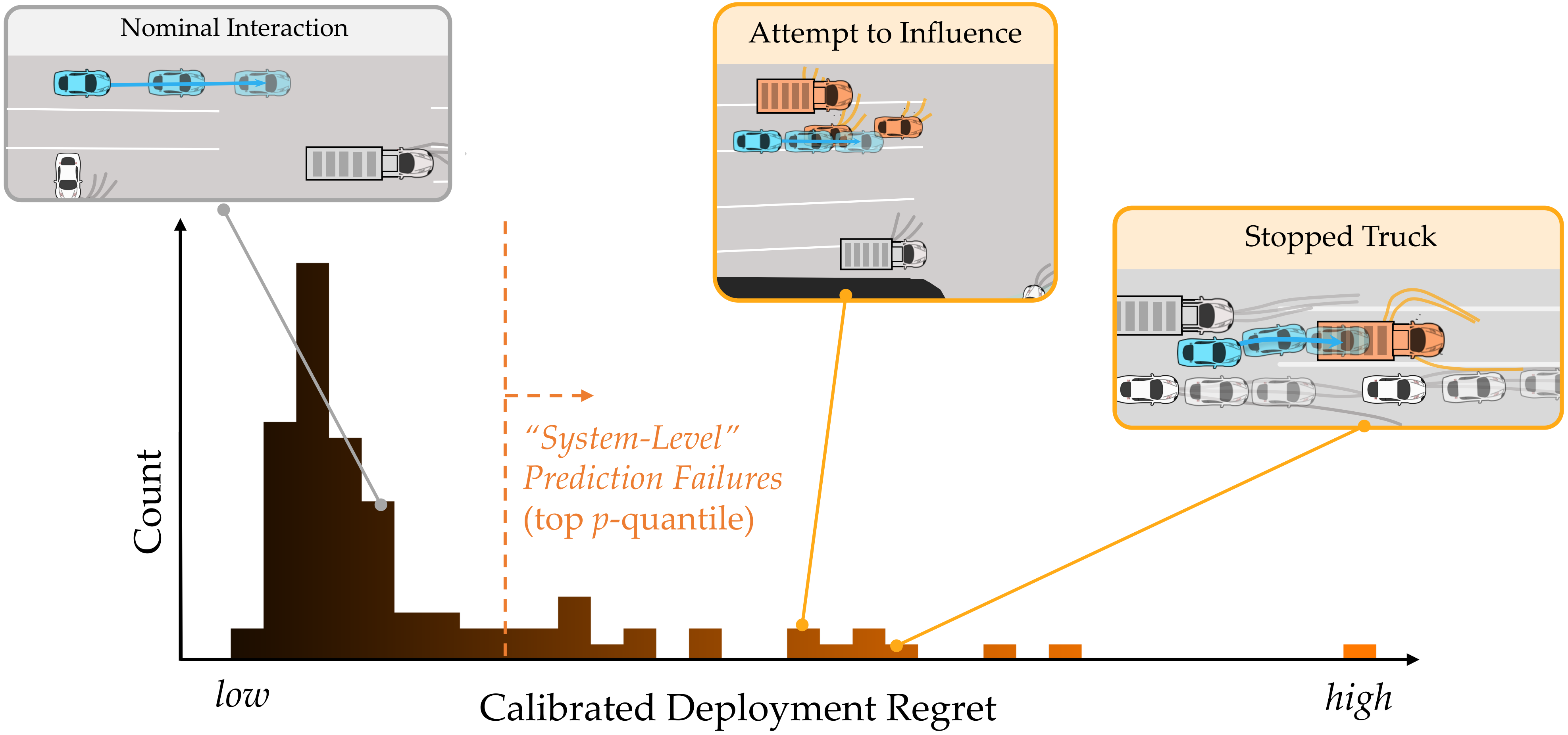
Calibrated Generalized Regret

$$\max_{\boldsymbol{a}^{\mathrm{R}}}\left[P_\theta\left(\boldsymbol{a}^{\mathrm{R}} \mid \hat{\boldsymbol{a}}^{\mathrm{H_1:H_M}}, \hat{\boldsymbol{s}}, C\right)\right] - P_\theta\left(\hat{\boldsymbol{a}}^{\mathrm{R}} \mid \hat{\boldsymbol{a}}^{\mathrm{H_1:H_M}}, \hat{\boldsymbol{s}}, C\right)$$

Closed-Loop Simulation on 100 Held-out Scenes from Nuscenes



Closed-Loop Simulation on 100 Held-out Scenes from Nuscenes



$$\max_{\boldsymbol{a}^{\mathrm{R}}}\left[P_\theta\left(\boldsymbol{a}^{\mathrm{R}} \mid \hat{\boldsymbol{a}}^{\mathrm{H}_1:\mathrm{H}_\mathrm{M}}, \hat{\boldsymbol{s}}, C\right)\right] - P_\theta\left(\hat{\boldsymbol{a}}^{\mathrm{R}} \mid \hat{\boldsymbol{a}}^{\mathrm{H}_1:\mathrm{H}_\mathrm{M}}, \hat{\boldsymbol{s}}, C\right)$$

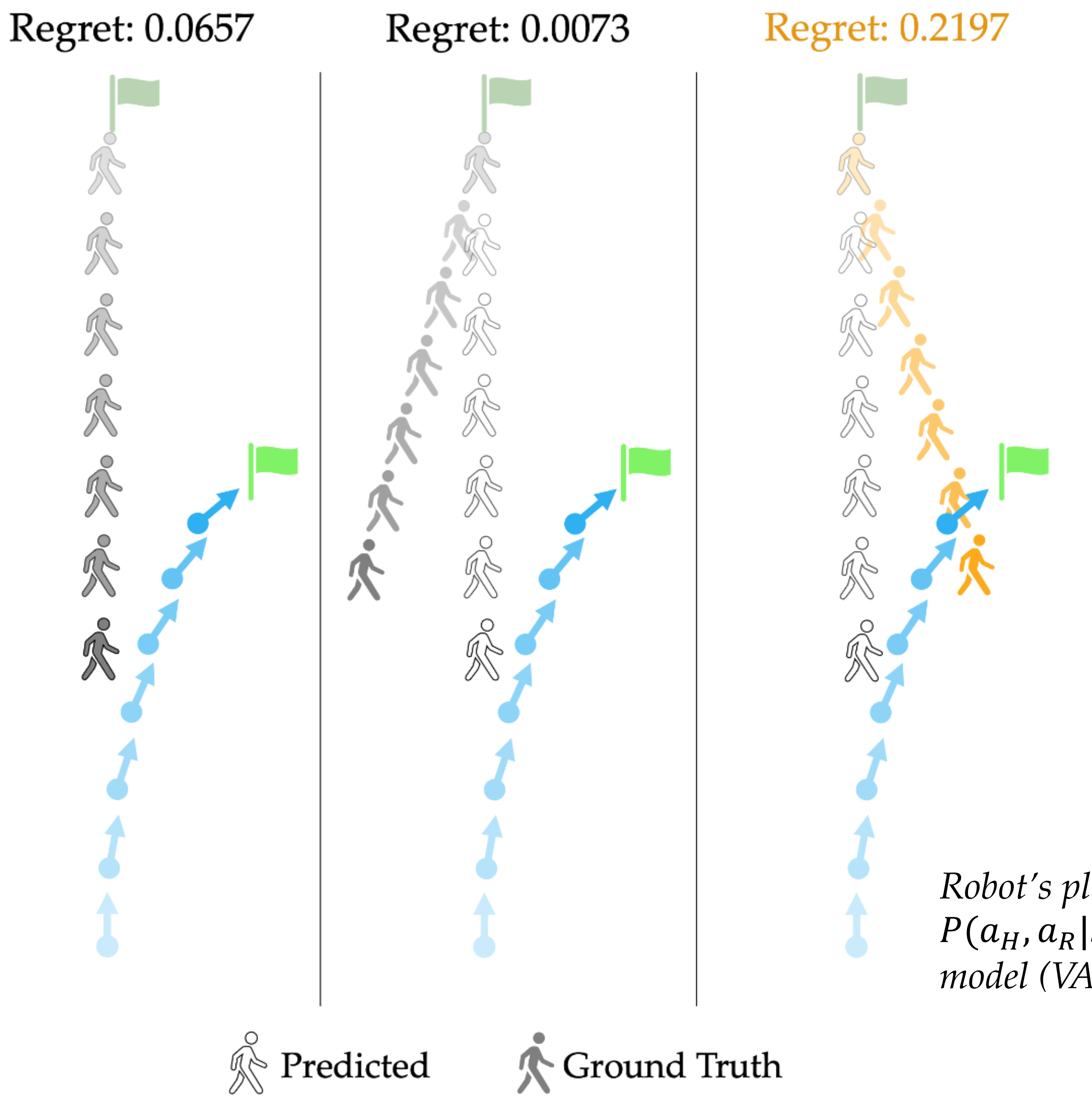
probability scales anomalies w.r.t. deployment context

no need for explicit reward model!

Calibrated Generalized Regret

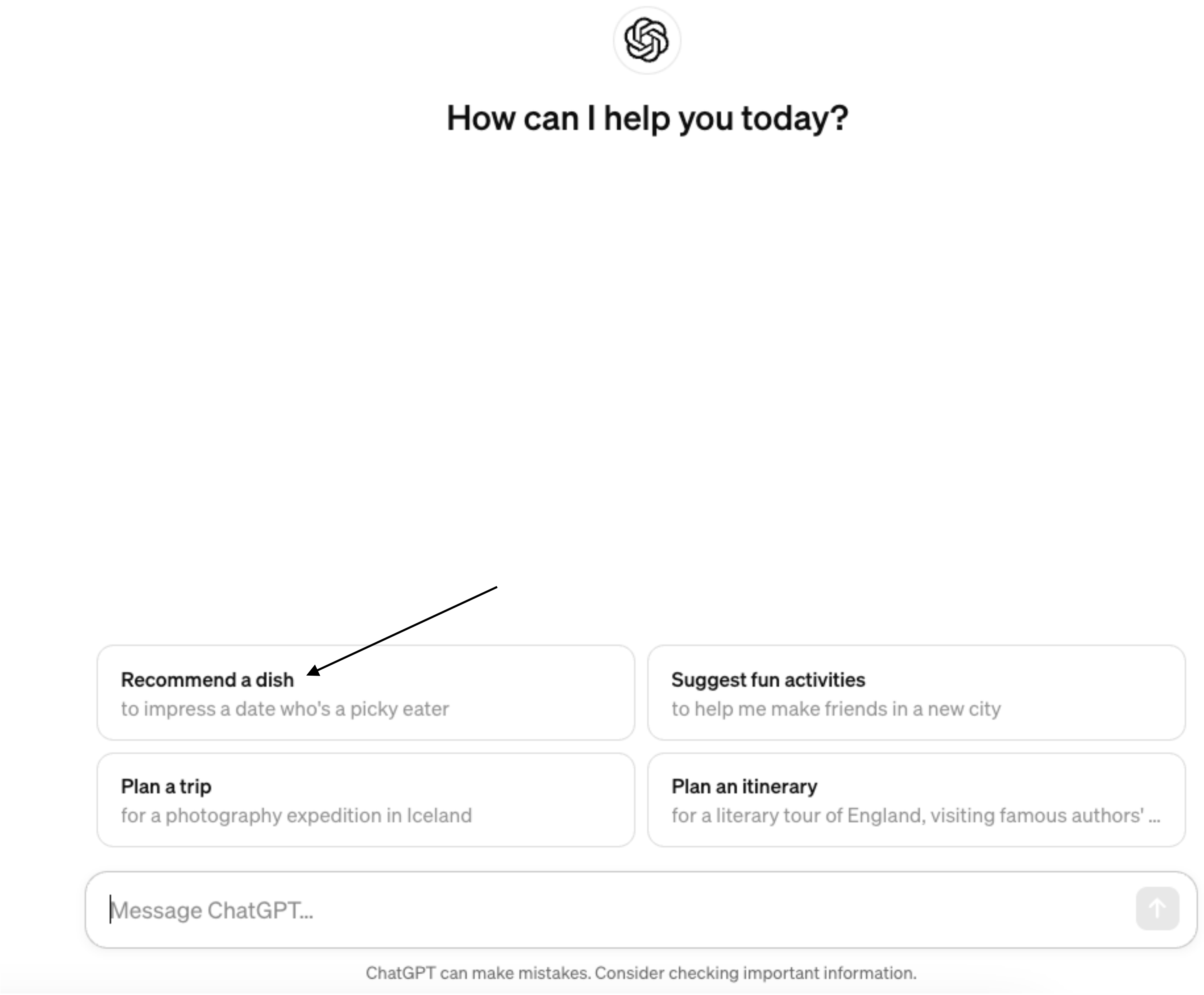
$$\max_{\boldsymbol{a}^{\mathrm{R}}}\left[P_{\theta}\left(\boldsymbol{a}^{\mathrm{R}} \mid \hat{\boldsymbol{a}}^{\mathrm{H}_1:\mathrm{H}_\mathrm{M}}, \hat{\boldsymbol{s}}, C\right)\right] - P_{\theta}\left(\hat{\boldsymbol{a}}^{\mathrm{R}} \mid \hat{\boldsymbol{a}}^{\mathrm{H}_1:\mathrm{H}_\mathrm{M}}, \hat{\boldsymbol{s}}, C\right)$$

! Opens up opportunities for anomaly detection for generative planners which *do not* maintain rewards



Robot's planner approximates $P(a_H, a_R|s)$ via a generative model (VAE)

Illustrative Example



Original Conversation: $(\hat{\boldsymbol{a}}^{R}, \hat{\boldsymbol{a}}^{H}, \hat{\boldsymbol{s}}, C) \in \mathcal{D}$

You

$\hat{a}_0^H$ I'm going to cook for my date who claims to be a picky eater. Can you recommend me a dish that's easy to cook?

Original Conversation: $(\hat{\boldsymbol{a}}^R, \hat{\boldsymbol{a}}^H, \hat{\boldsymbol{s}}, C) \in \mathcal{D}$

$\hat{a}_0^H$

You

I'm going to cook for my date who claims to be a picky eater. Can you recommend me a dish that's easy to cook?

$\hat{a}_0^R$

ChatGPT

Certainly! When cooking for a picky eater, it's usually best to stick with something classic and not overly complex in terms of flavors. A dish that often pleases many palates while being relatively simple to prepare is Chicken Alfredo Pasta. It's creamy, comforting, and has a balance of flavors that are widely appreciated. Here's a straightforward recipe you can try:

Original Conversation: $(\hat{\boldsymbol{a}}^R, \hat{\boldsymbol{a}}^H, \hat{\boldsymbol{s}}, C) \in \mathcal{D}$

$\hat{a}_0^H$ **You**

I'm going to cook for my date who claims to be a picky eater. Can you recommend me a dish that's easy to cook?

$\hat{a}_0^R$ **ChatGPT**

Certainly! When cooking for a picky eater, it's usually best to stick with something classic and not overly complex in terms of flavors. A dish that often pleases many palates while being relatively simple to prepare is Chicken Alfredo Pasta. It's creamy, comforting, and has a balance of flavors that are widely appreciated. Here's a straightforward recipe you can try:

$\hat{a}_0^H$ **You**

My date has an allergy to dairy

Original Conversation: $(\hat{\boldsymbol{a}}^R, \hat{\boldsymbol{a}}^H, \hat{\boldsymbol{s}}, C) \in \mathcal{D}$

$\hat{a}_0^H$ **You**
I'm going to cook for my date who claims to be a picky eater. Can you recommend me a dish that's easy to cook?

$\hat{a}_0^R$ **ChatGPT**
Certainly! When cooking for a picky eater, it's usually best to stick with something classic and not overly complex in terms of flavors. A dish that often pleases many palates while being relatively simple to prepare is Chicken Alfredo Pasta. It's creamy, comforting, and has a balance of flavors that are widely appreciated. Here's a straightforward recipe you can try:

$\hat{a}_0^H$ **You**
My date has an allergy to dairy

$\hat{a}_0^R$ **ChatGPT**
Certainly! If your date is a picky eater and has a dairy allergy, it's best to go for a dish that's simple, versatile, and flavorful without relying on dairy products. One great option is a Lemon Herb Chicken with a side of garlic roasted vegetables. This dish is straightforward, delicious, and avoids dairy. Here's how you can make it:

(Approximate) Counterfactual Optimization: $\max_{\boldsymbol{a}^{\mathrm{R}}}\left[P_{\theta}\left(\boldsymbol{a}^{\mathrm{R}} \mid \hat{\boldsymbol{a}}^{H}, \hat{\boldsymbol{s}}, C\right)\right]$

AB **You**

I had the following conversation with you:

Me: I'm going to cook for my date who claims to be a picky eater. Can you recommend me a dish that's easy to cook?

You: Certainly! When cooking for a picky eater, it's usually best to stick with something classic and not overly complex in terms of flavors. A dish that often pleases many palates while being relatively simple to prepare is Chicken Alfredo Pasta. It's creamy, comforting, and has a balance of flavors that are widely appreciated.

Me: My date has an allergy to dairy

You: Certainly! If your date is a picky eater and has a dairy allergy, it's best to go for a dish that's simple, versatile, and flavorful without relying on dairy products. One great option is a Lemon Herb Chicken with a side of garlic roasted vegetables. This dish is straightforward, delicious, and avoids dairy. Here's how you can make it:

Was there a better response you could have given me to my first question?

(Approximate) Counterfactual Optimization: $\max_{\boldsymbol{a}^{\mathrm{R}}}\left[P_{\theta}\left(\boldsymbol{a}^{\mathrm{R}} \mid \hat{\boldsymbol{a}}^{H}, \hat{\boldsymbol{s}}, C\right)\right]$

Was there a better response you could have given me to my first question?

ChatGPT

Your initial question sought a dish recommendation for a picky eater that's also easy to cook, without mentioning the dairy allergy. My response aimed to provide a universally appealing and simple option with Chicken Alfredo Pasta. However, after learning about the dairy allergy, the suggestion was pivoted to Lemon Herb Chicken with garlic roasted vegetables, which is indeed more suitable given the dietary restriction.

A more thorough initial response could have included a direct inquiry about any dietary restrictions or preferences, which would have allowed for a tailored suggestion from the start. Here's how a better response could have been formulated:

(Approximate) Counterfactual Optimization: $\max_{\boldsymbol{a}^{R}}[P_{\theta}(\boldsymbol{a}^{R} \mid \hat{\boldsymbol{a}}^{H}, \hat{\boldsymbol{s}}, C)]$

Was there a better response you could have given me to my first question?

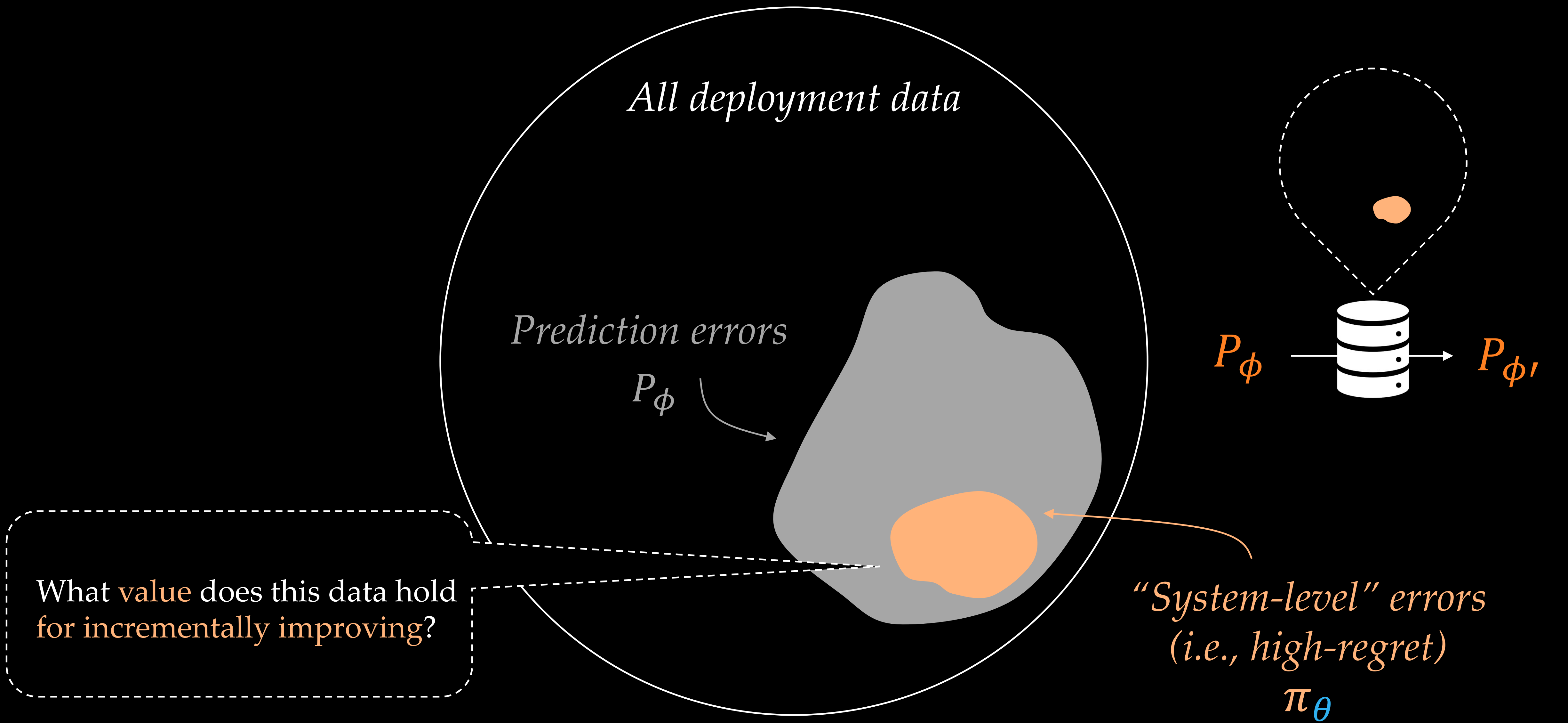
ChatGPT

Your initial question sought a dish recommendation for a picky eater that's also easy to cook, without mentioning the dairy allergy. My response aimed to provide a universally appealing and simple option with Chicken Alfredo Pasta. However, after learning about the dairy allergy, the suggestion was pivoted to Lemon Herb Chicken with garlic roasted vegetables, which is indeed more suitable given the dietary restriction.

A more thorough initial response could have included a direct inquiry about any dietary restrictions or preferences, which would have allowed for a tailored suggestion from the start. Here's how a better response could have been formulated:

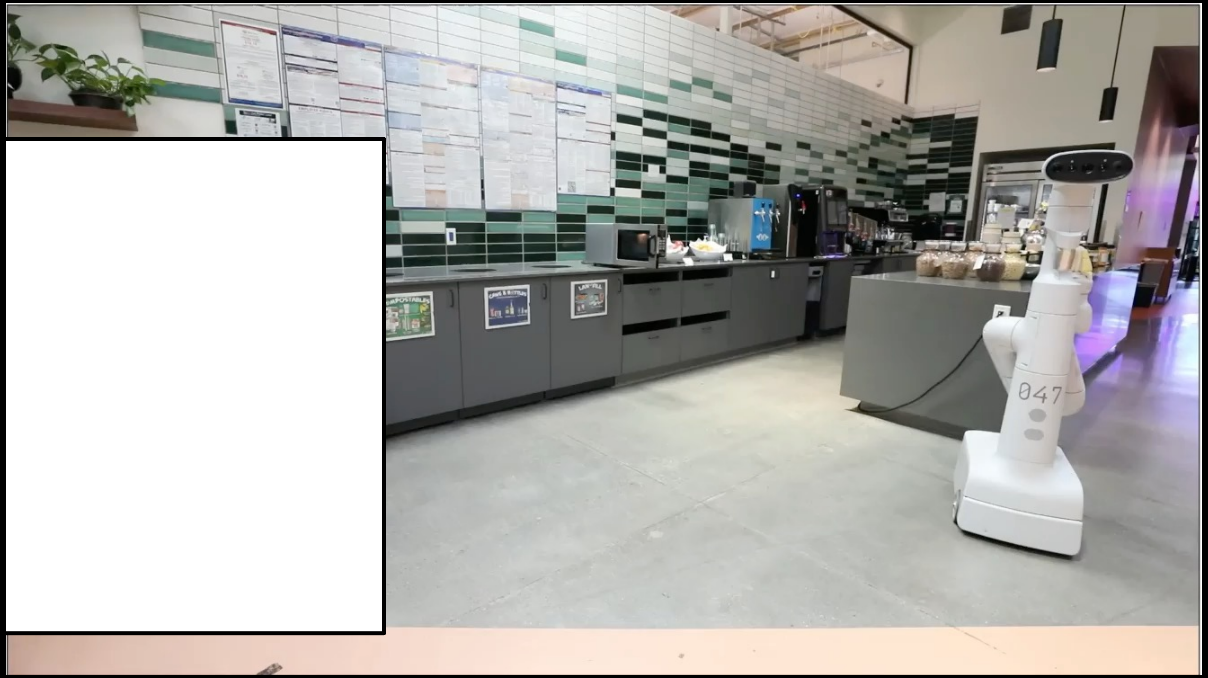
You: I'm going to cook for my date who claims to be a picky eater. Can you recommend me a dish that's easy to cook?

Improved Response: Of course, I'd be happy to help with that! To ensure the recommendation suits your needs, could you tell me if there are any dietary restrictions or particular dislikes I should be aware of? For instance, allergies, vegetarian preferences, or ingredients


All deployment data
Prediction errors
P_ϕ
"System-level" errors
(i.e., high-regret)
π_θ
P_ϕ
$P_{\phi'}$
What value does this data hold for incrementally improving?

Future opportunities beyond autonomous driving domain….

anomaly detection, informed data augmentation, model robustification…

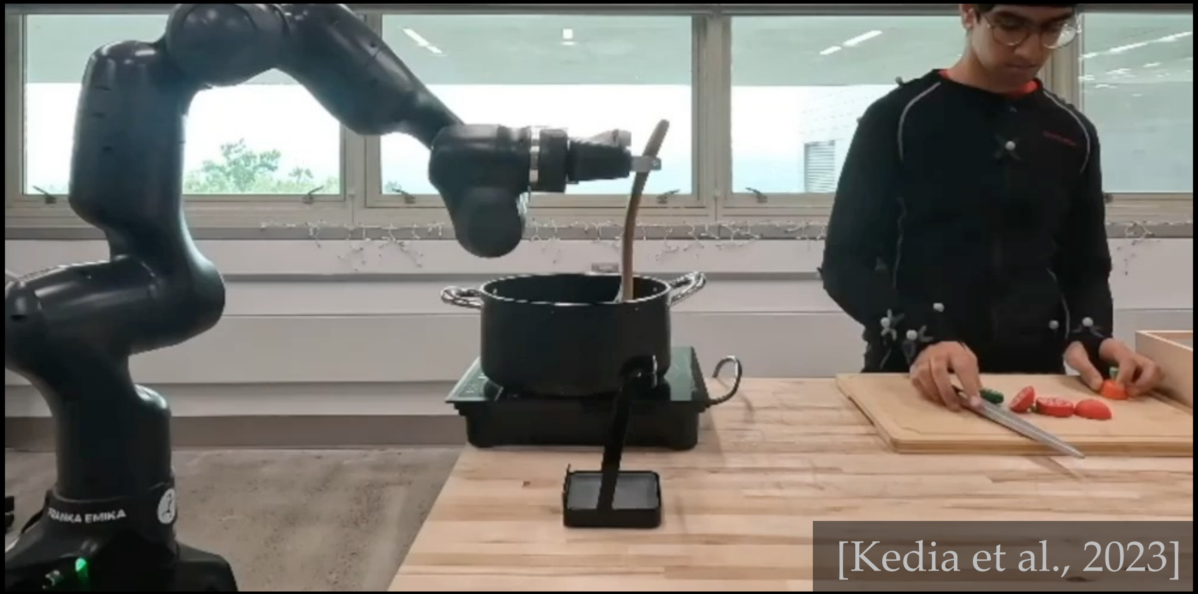


Gemini


[Waymo, 2023]


[Ren, AZ et al., 2023]

More work to be done so autonomous robots can interact reliably at scale


[Kedia et al., 2023]


[DeepMind, 2023]

(5 min) Course Eval Survey! ☺

Carnegie Mellon University

abajcsy@cmu.edu

intent ROBOTICS LAB