Safety Beyond Physical States

Kensuke Nakamura

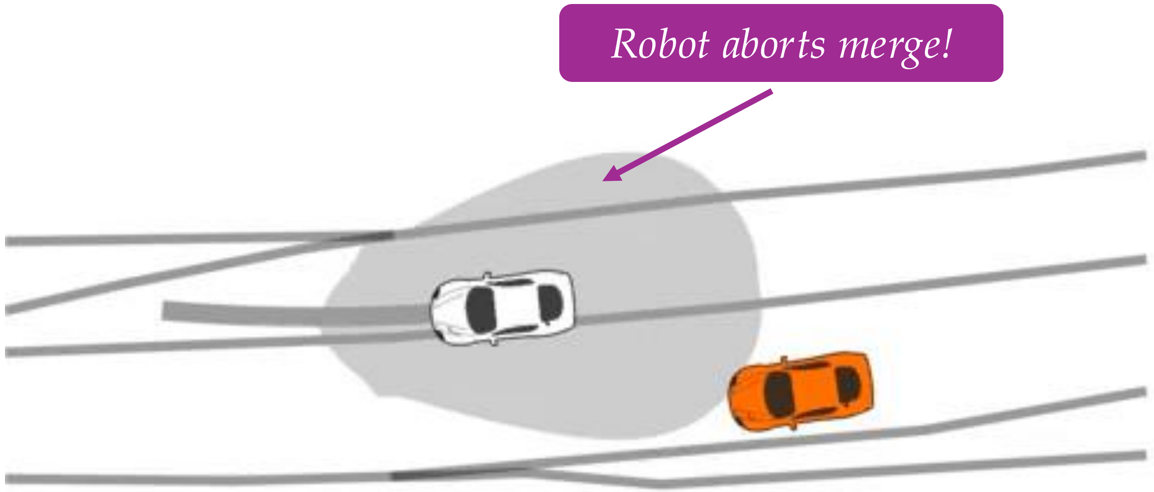
EAIS 16-886 Guest Lecture

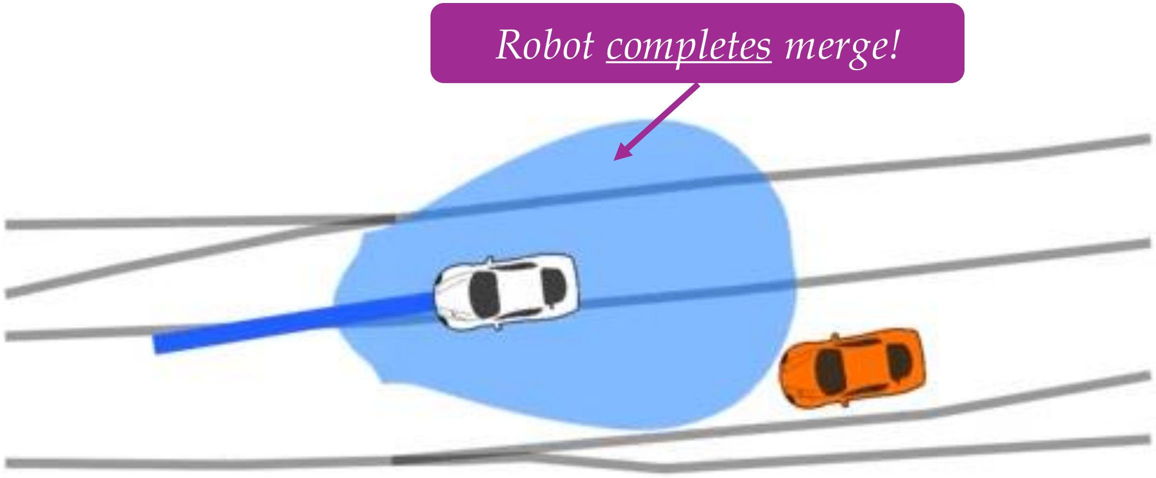
Carnegie Mellon University
Robotics Institute

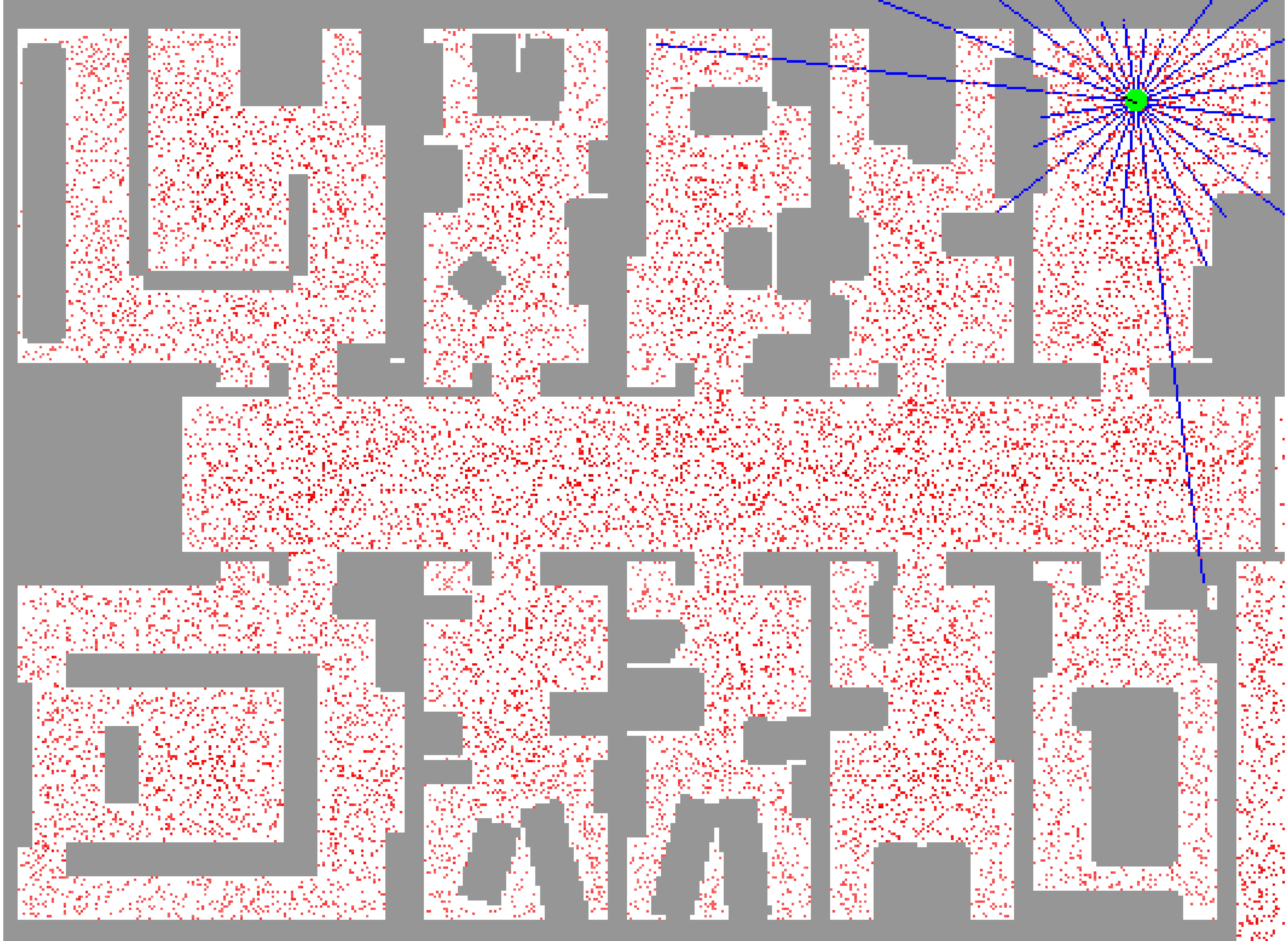
intent
ROBOTICS LAB

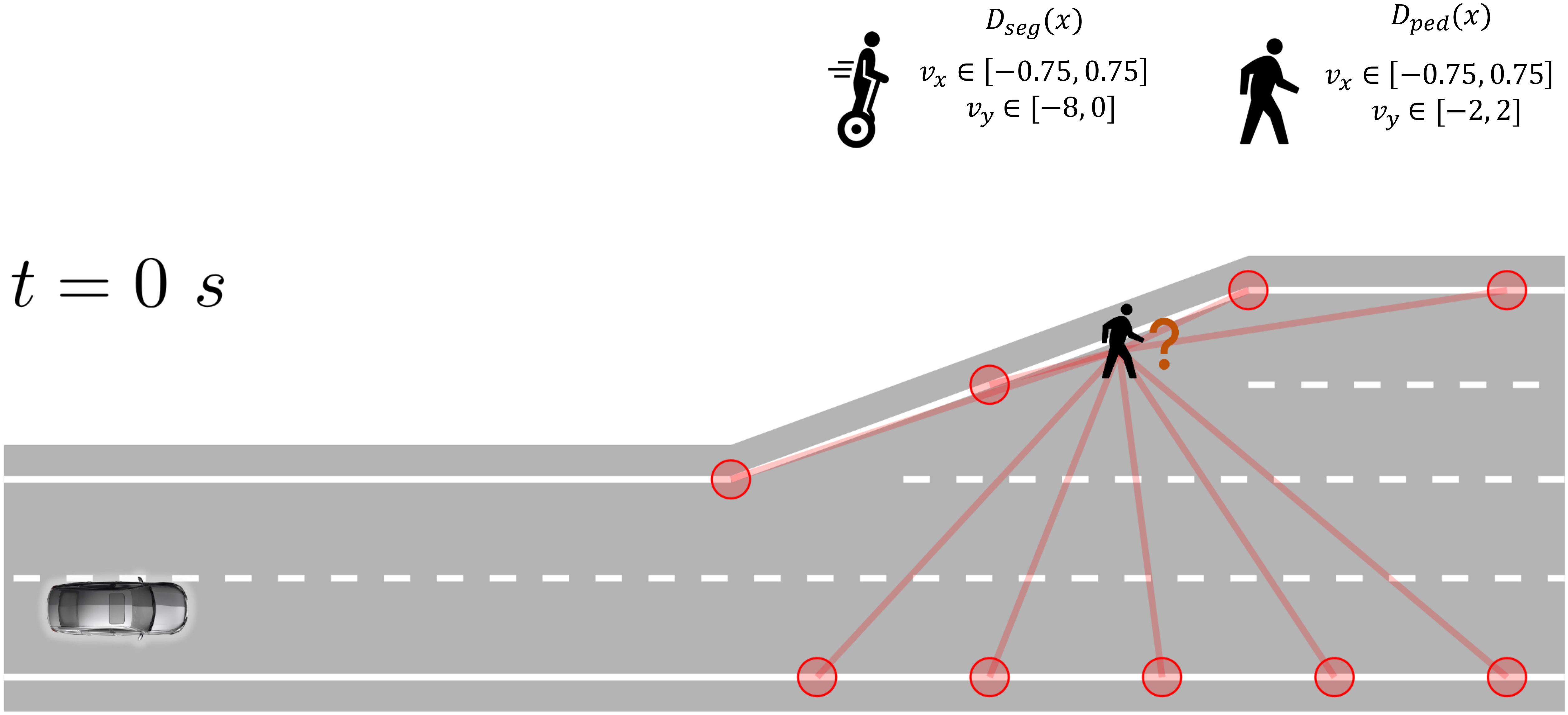

andon
rencester
417


16/30 mph
AUTO


Robot aborts merge!


Robot completes merge!

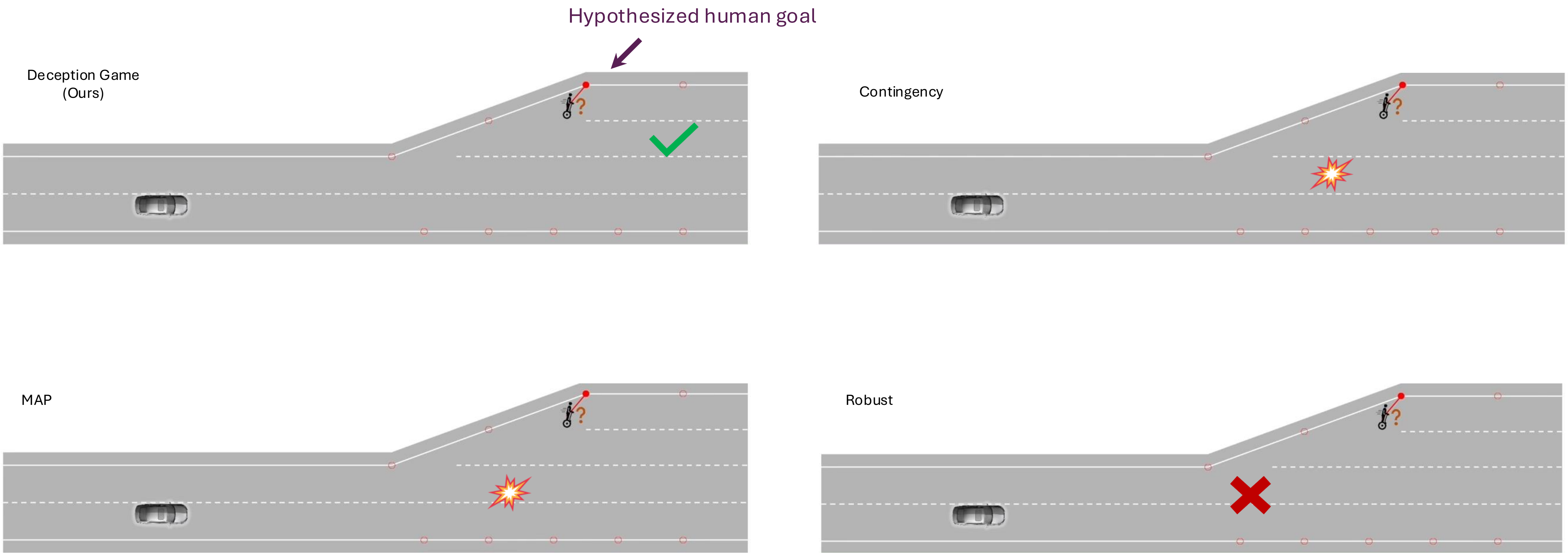




$t = 0\ s$
$D_{seg}(x)$
$v_x \in [-0.75, 0.75]$
$v_y \in [-8, 0]$
$D_{ped}(x)$
$v_x \in [-0.75, 0.75]$
$v_y \in [-2, 2]$
?

Case Study: Biased Prior and Hypothesis Recovery

Even when the robot had a strongly biased and incorrect prior on the human goal, the Deception Game policy was able to safely navigate around an adversarial and deceptive human, unlike the baseline methods



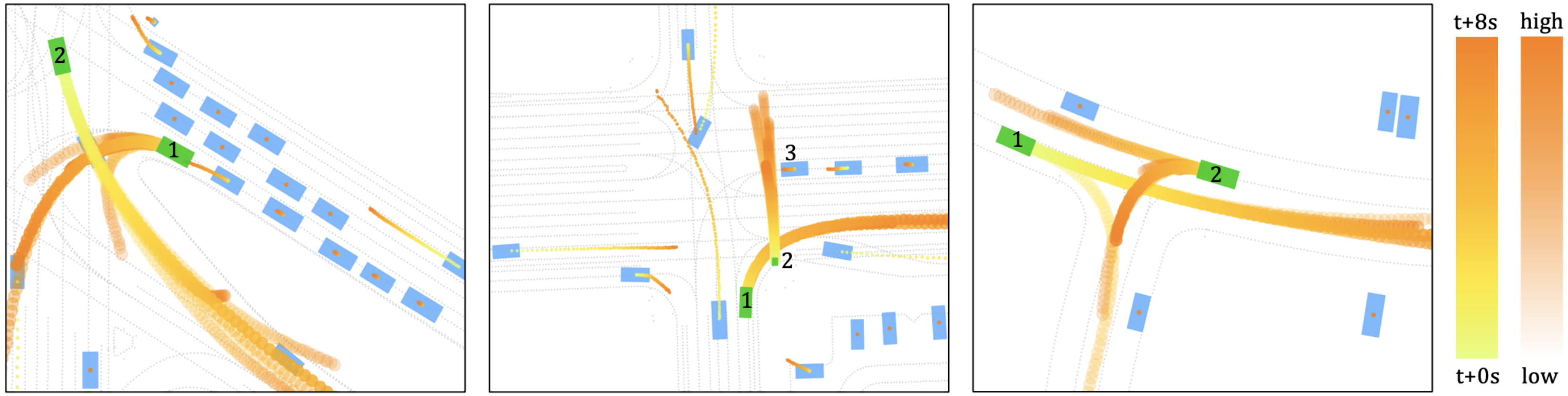
Human uses a pedestrian action

Human uses a Segway action

? Robot is uncertain about the human's type

Implicit Learning Dynamics: Motion Transformer

Input: History of states, map
Output: 64 trajectory predictions + associated weights (GMM)



(a) **V2** is passing the intersection to turn left with high speed. Our model predicts multimodal behaviors for **V1:** turn left or make a U-turn. In any case, **V1** is predicted to yield for **V2**.

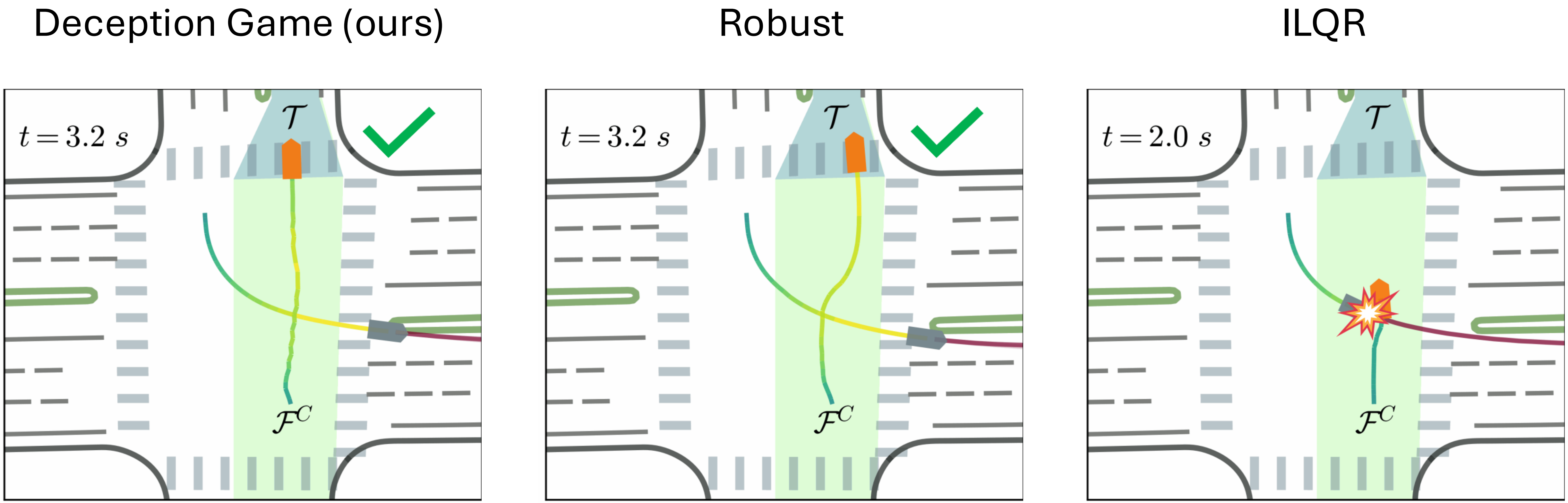
(b) **P2** is passing the road through the crosswalk while **V1** is on the right-turn lane to turn right. Both **V1** and **V3** are predicted to yield for **P2**.

(c) Our model predicts multimodal behaviors for **V1**: go straight and turn right, since it still has a distance to the intersection. **V2** is predicted to yield for **V1** when turning left, since **V1** is moving fast towards the intersection.

Figure 5: Qualitative results of MTR framework on WOMD. There are two interested agents in each scene (green rectangle), where our model predicts 6 multimodal future trajectories for each of them. For other agents (blue rectangle), a single trajectory is predicted by dense future prediction module. We use gradient color to visualize the trajectory waypoints at different future time step, and trajectory confidence is visualized by setting different transparent. Abbreviation: Vehicle (**V**), Pedestrian (**P**).

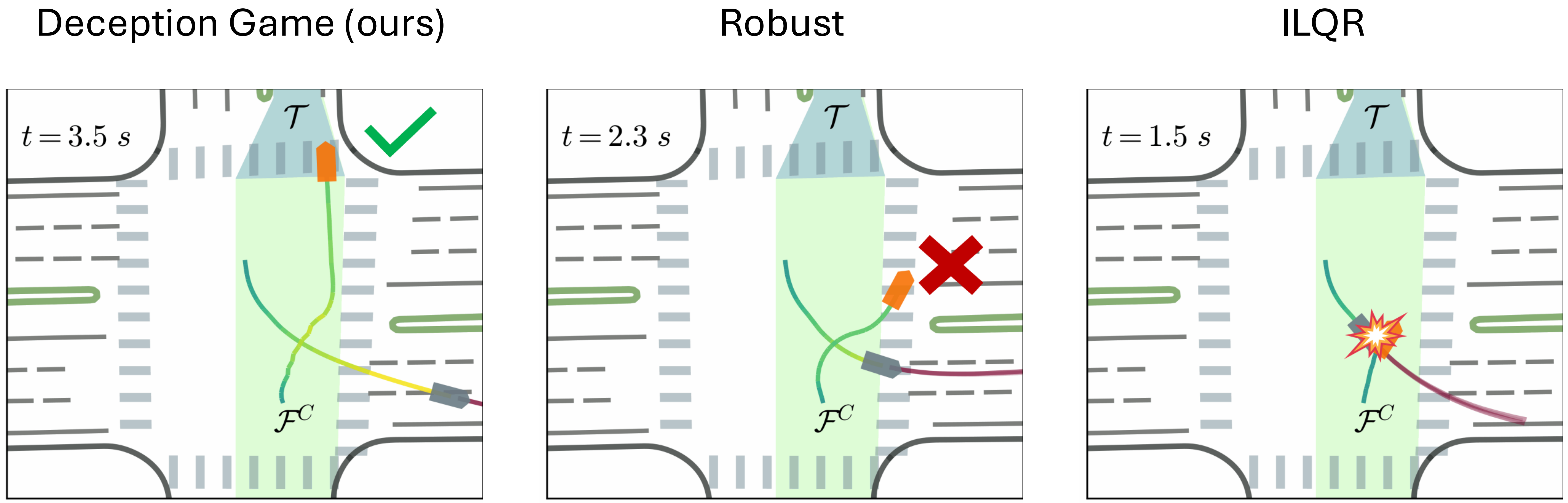
Case Study: Neural Trajectory Predictor

Scenario 1



Case Study: Neural Trajectory Predictor

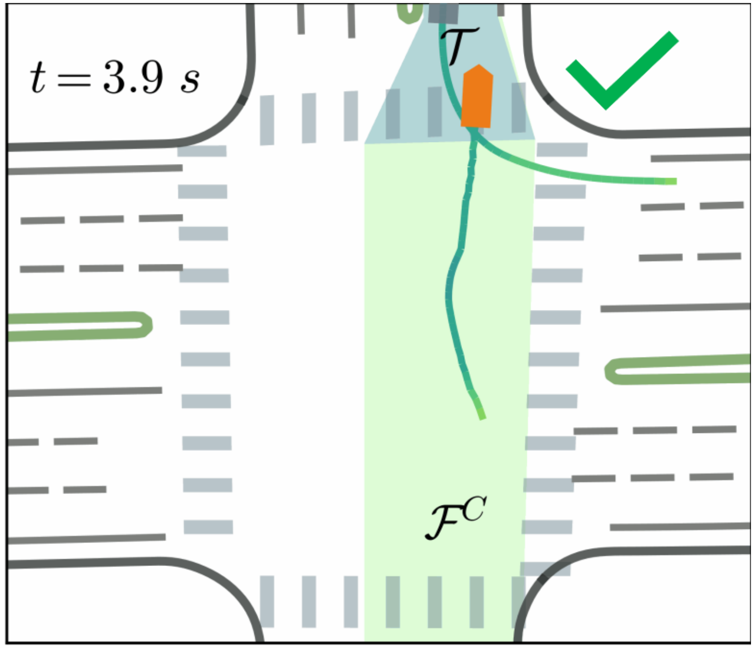
Scenario 2



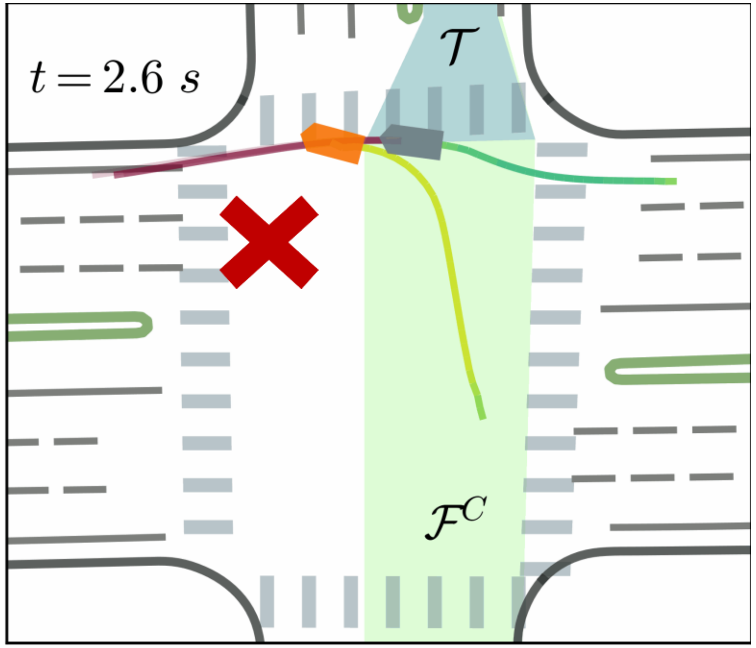
Case Study: Neural Trajectory Predictor

Scenario 3

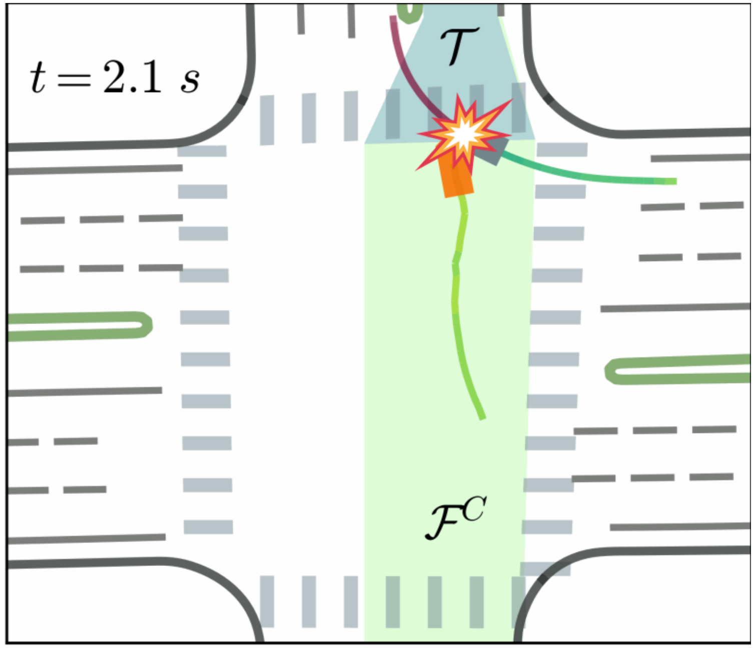
Deception Game (ours)



Robust

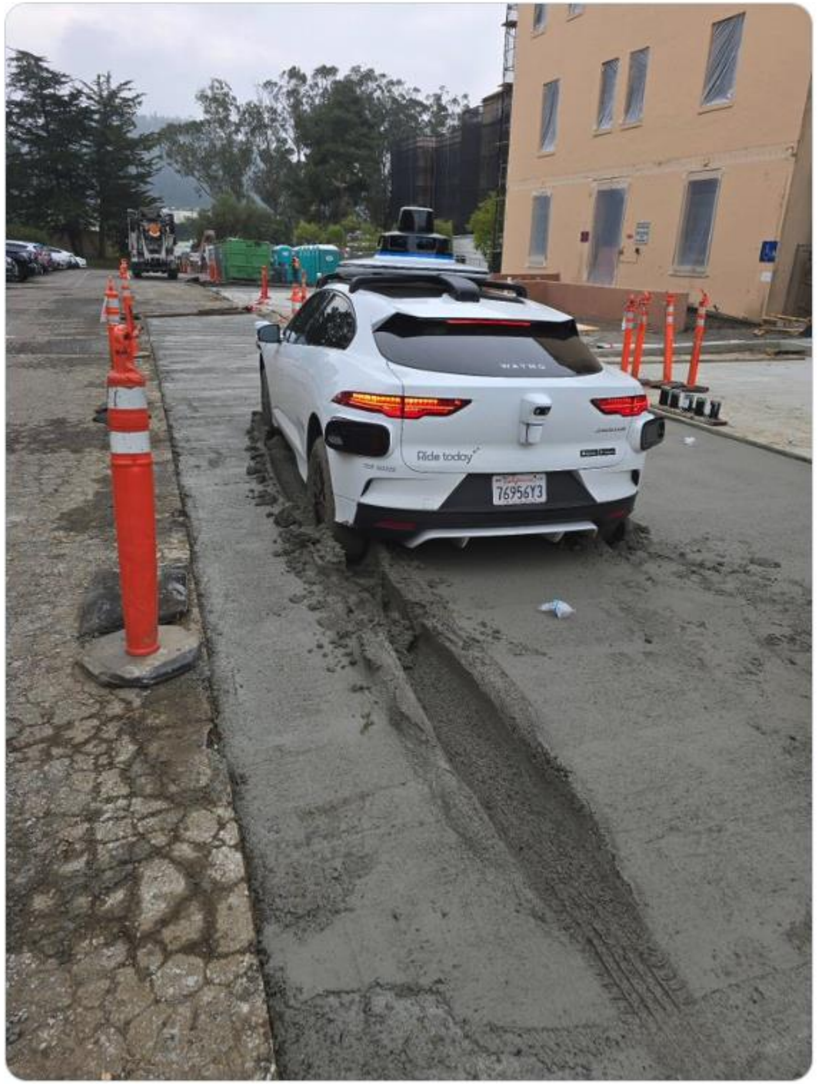


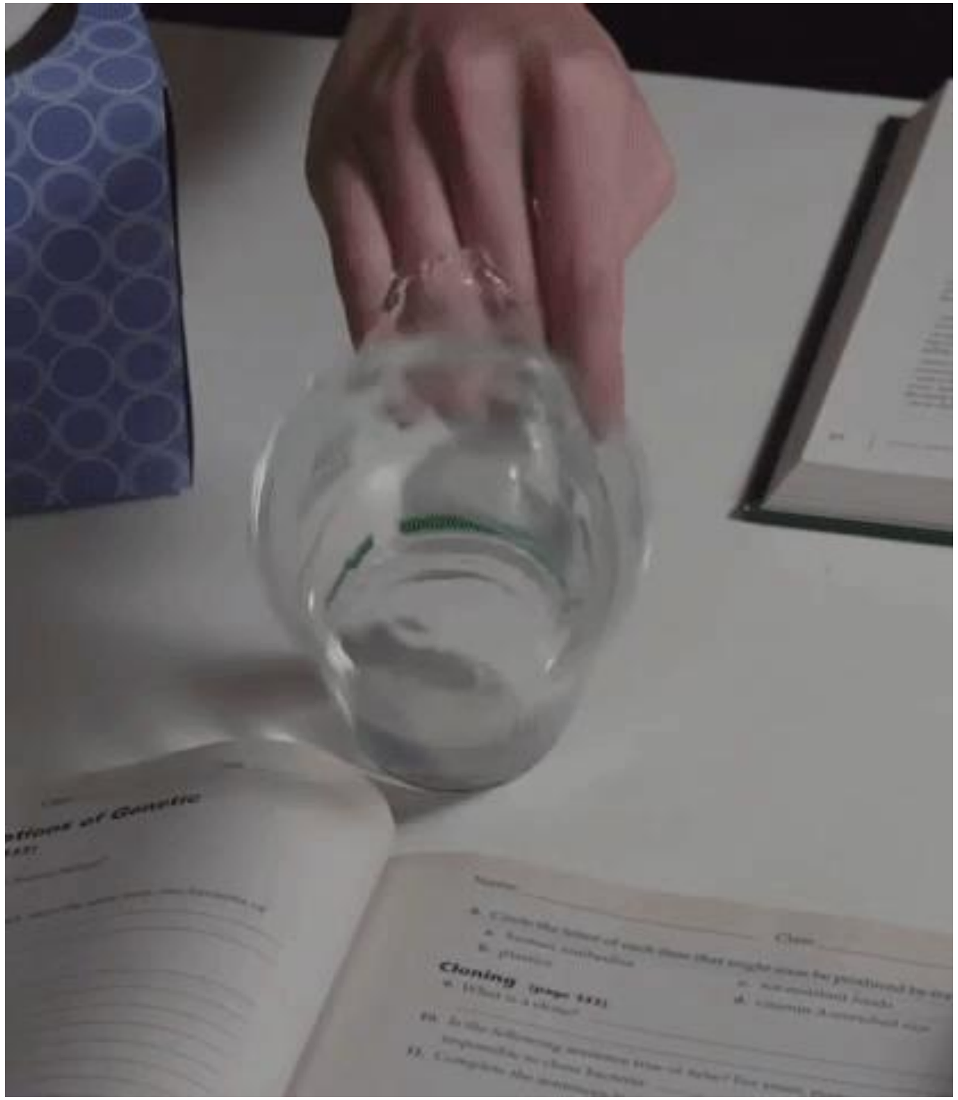
ILQR



So, what's missing for safety in the 'open world'?

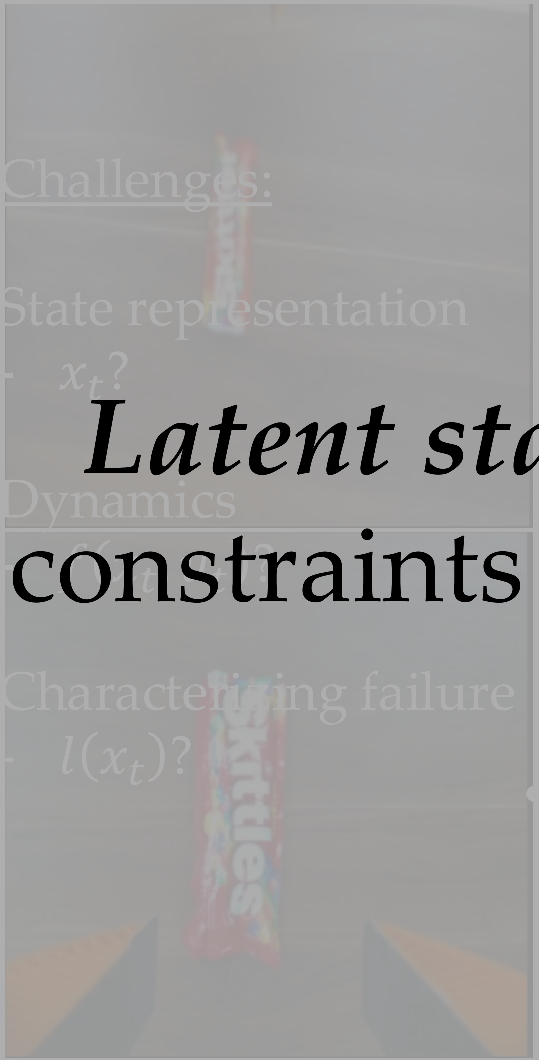
Oops! @Waymo

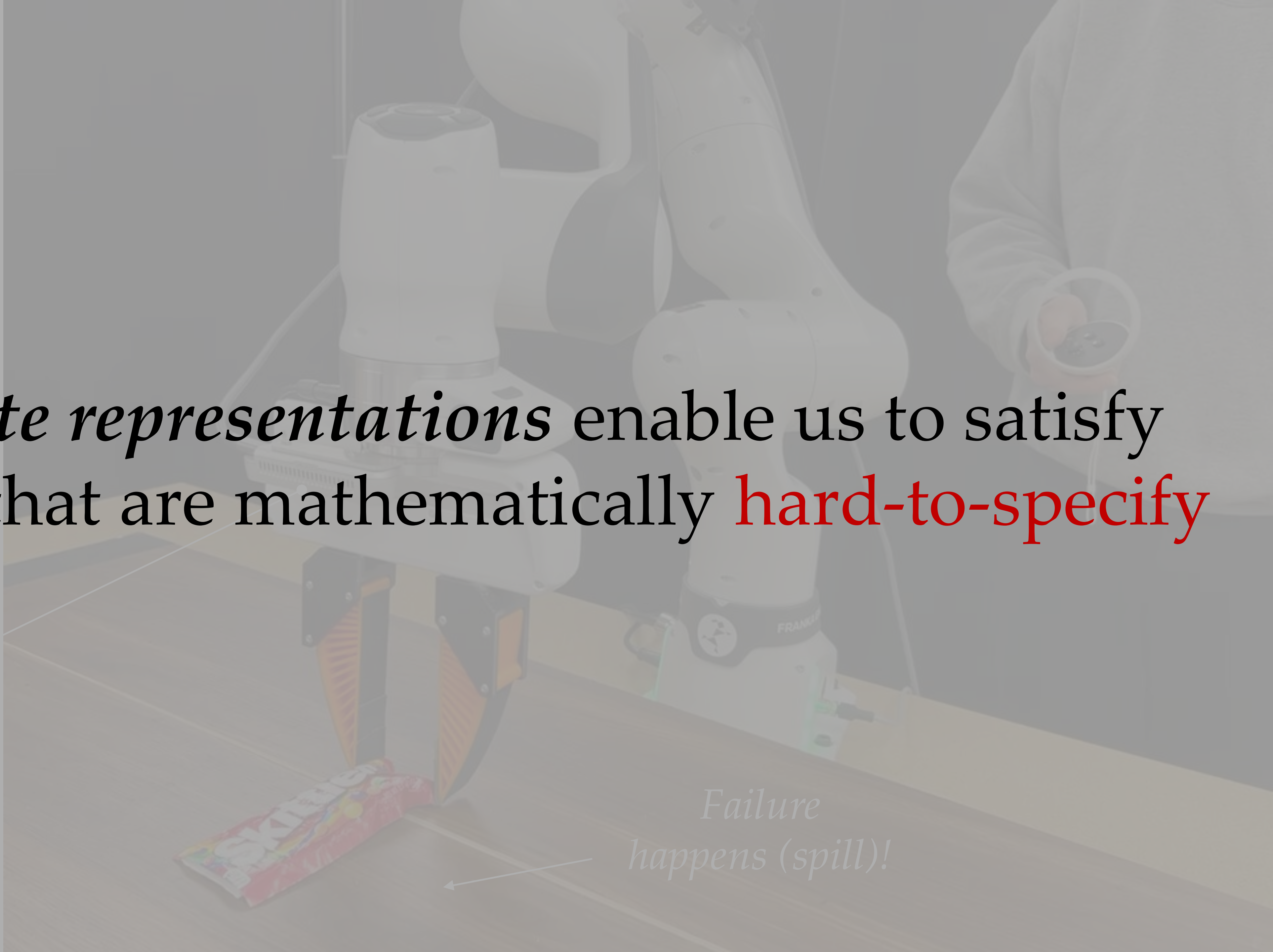




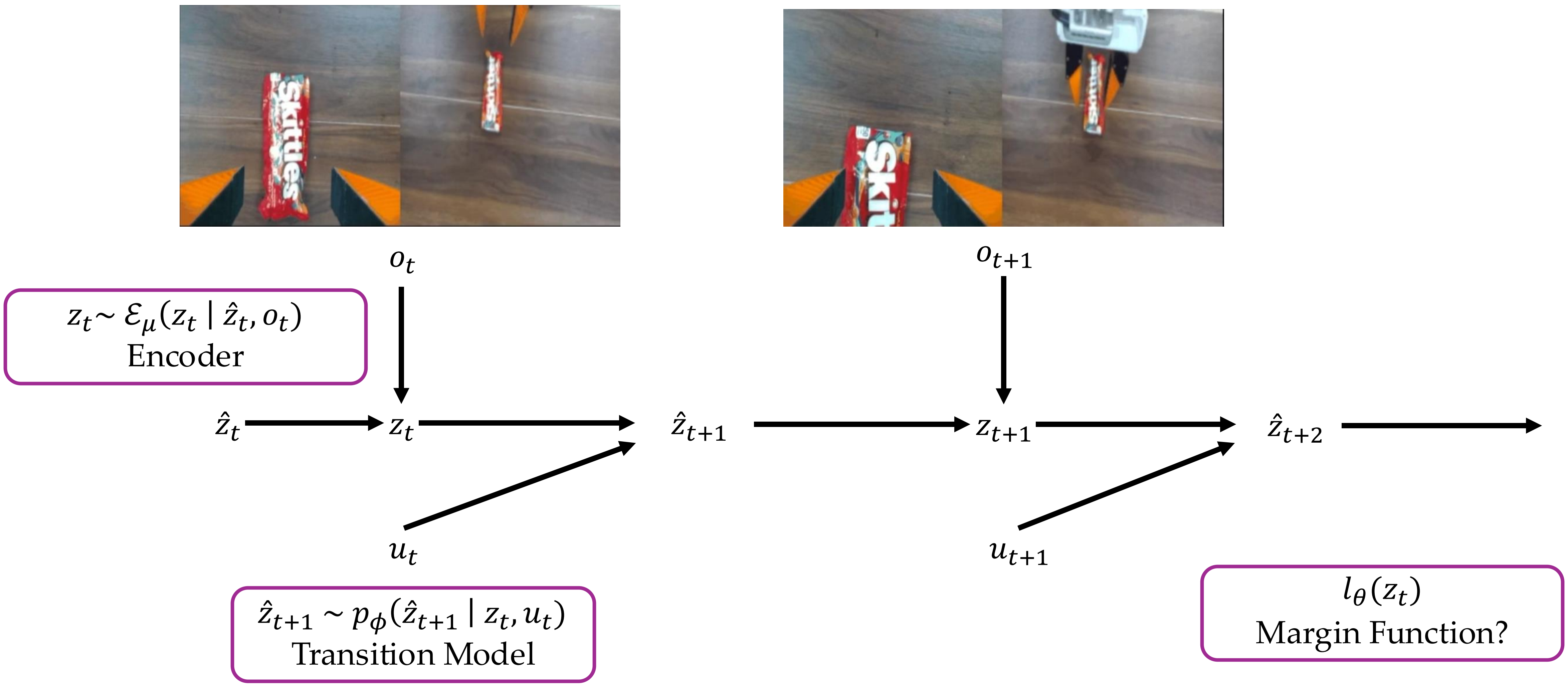
[Guan, et al. "Task Success" is Not Enough. COLM 2024]

Our representations of safety should be *more* than just collisions





Latent state representations enable us to satisfy constraints that are mathematically hard-to-specify



Training objective: minimize difference between $\hat{z}_t$ and z_t (+ auxiliary losses)

Examples: Recurrent state-space models (RSSMs), DINO-WM

Option 3: Pretrained Vision Foundation Model (e.g., DINOv2)

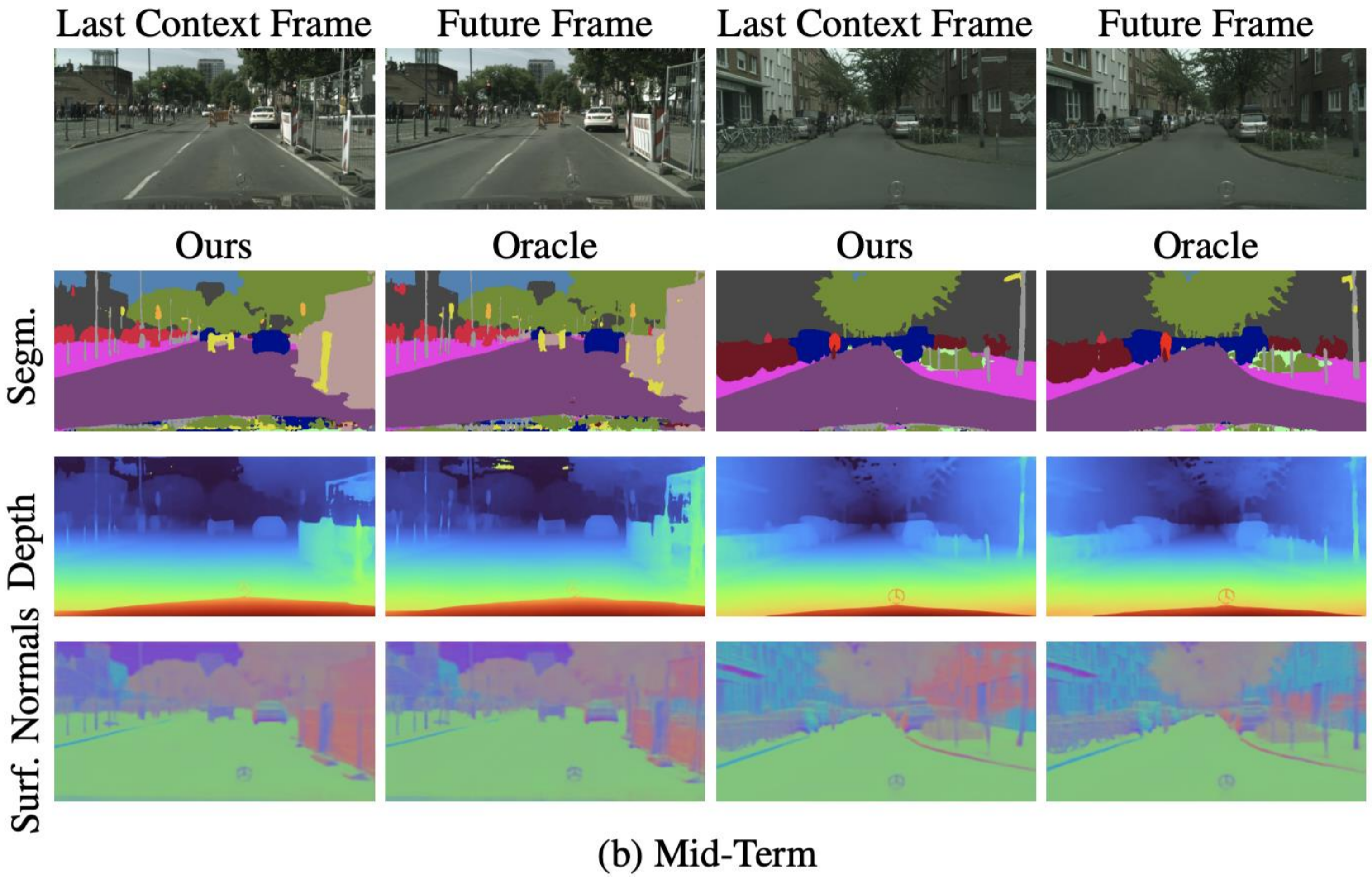
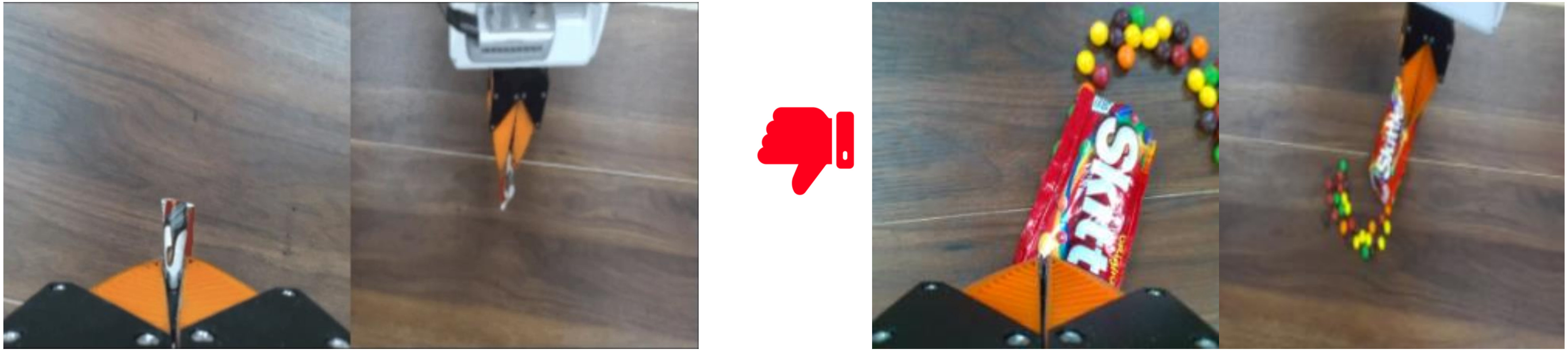


Figure 4. **Visualization of future predictions for semantic segmentation, depth, and surface normals.** Noisy segmentation

DINO-Foresight
Karypidis 2024



z_t

$$\mathcal{L}(\theta) = ReLU(\delta - l_\theta(z_t))$$

"penalize $l(z) < \delta$"

z_t

$$\mathcal{L}(\theta) = ReLU(\delta + l_\theta(z_t))$$

"penalize $l(z) > -\delta$"

Latent Hamilton-Jacobi Safety Bellman Equation

$$V(z) = \min\{l_\theta(z), \max_{u \in \mathcal{U}} \mathbb{E}_{\hat{z}' \sim p_\phi(\cdot|z,u)}[V(\hat{z}')]\}$$

"State" representation:
$z_t \sim \mathcal{E}_\mu(z_t \mid \hat{z}_t, o_t)$

Dynamics:
$\hat{z}' \sim p_\phi(\cdot|z, u)$

Characterizing failure:
$l_\theta(z_t)$

Approximating Safety with Reinforcement Learning

$$V(z) = (1-\gamma)l_\theta(z) + \gamma \min\{l_\theta(z), \max_{u\in\mathcal{U}} \mathbb{E}_{\hat{z}' \sim p_\phi(\cdot|z,u)}[V(\hat{z}')]\}$$

Resets in the world model

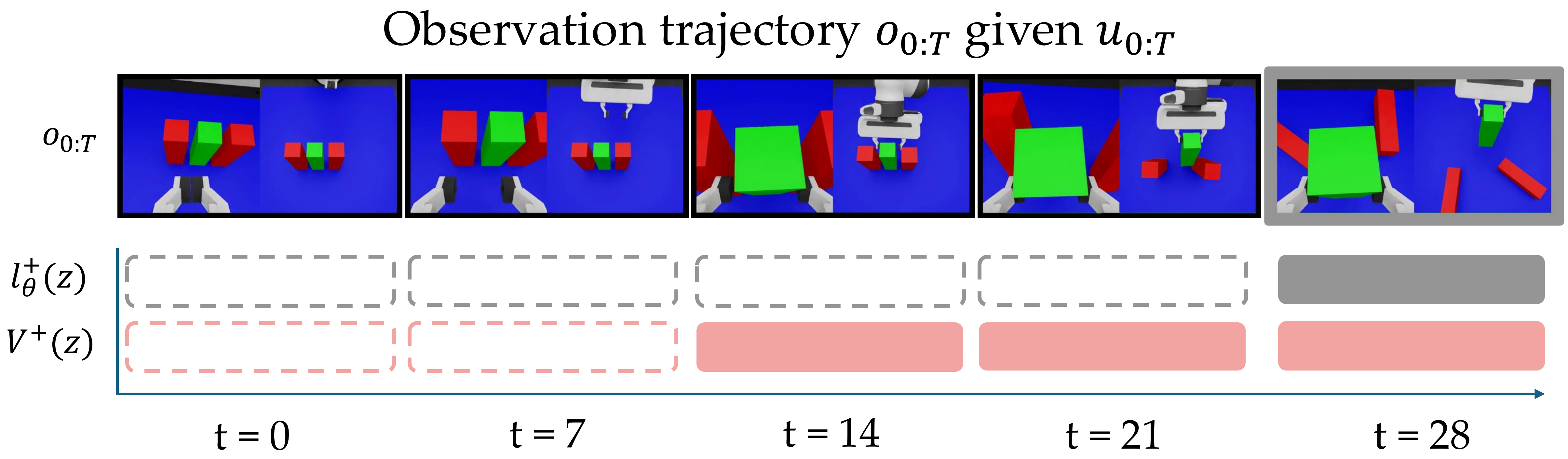
$o_0 \sim ReplayBuffer$

$z_0 \sim \mathcal{E}_\mu(z_0 \mid \hat{z}_0, o_0)$

Short world model rollouts

$$\begin{tikzcd}
u_0 \arrow[dr] & u_1 \arrow[dr] & u_2 \arrow[dr] & u_3 \arrow[dr] & \\
z_0 \arrow[r] \arrow[d] & \hat{z}_1 \arrow[r] \arrow[d] & \hat{z}_2 \arrow[r] \arrow[d] & \hat{z}_3 \arrow[r] \arrow[d] & {} \\
l_\theta(z_0) & l_\theta(z_1) & l_\theta(z_2) & l_\theta(z_3) &
\end{tikzcd}$$

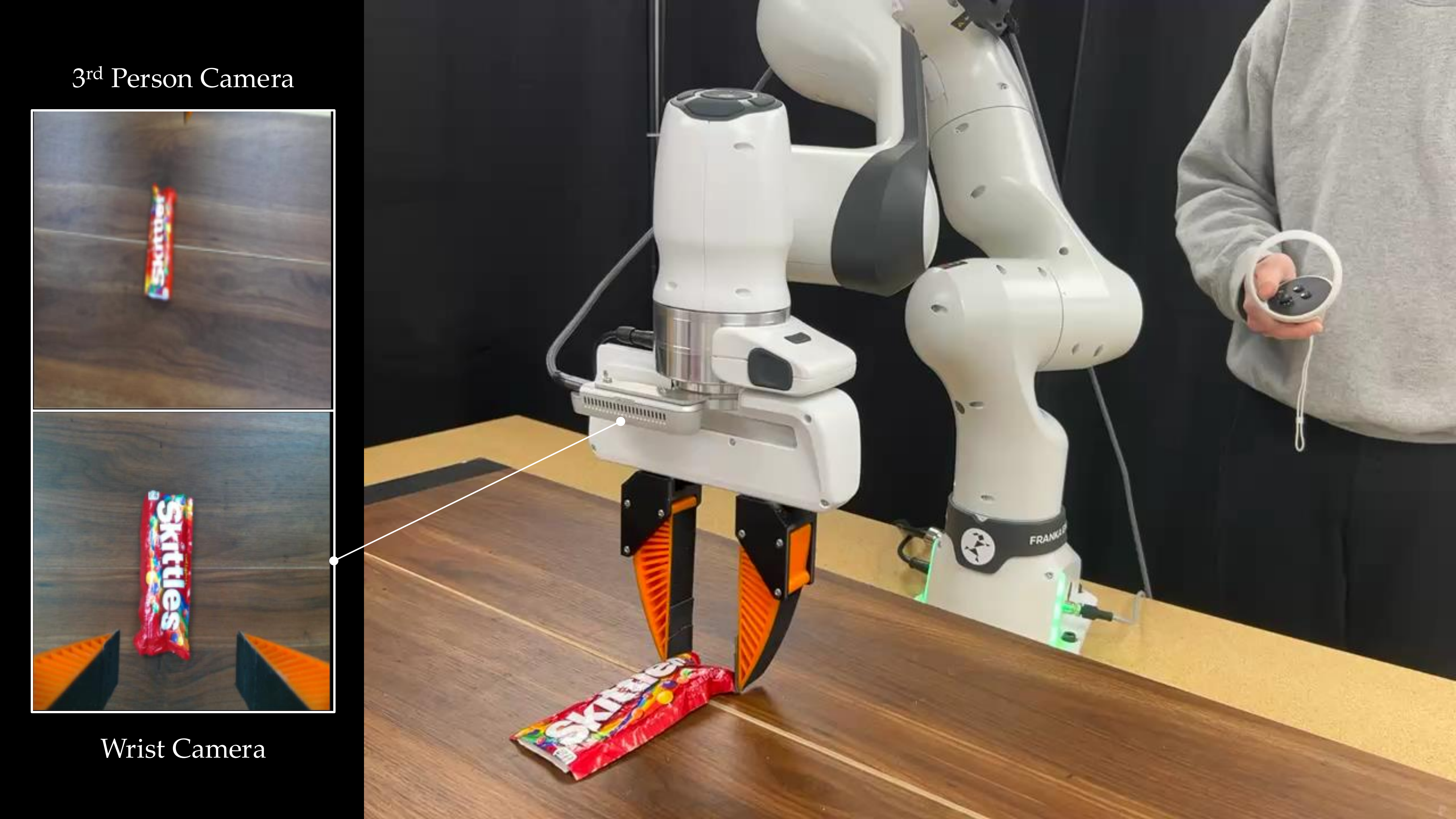
Simulation Experiments



Simulation Experiments

Nominal Policy: Dreamer
Baseline: Safety Q-functions for RL (SQRL)
Ours: LatentSafe

Method	**Safe Success % (↑)**	**Constraint Violation % (↓)**	**Incompletion % (↓)**
Dreamer	64	36	**0**
SQRL ($\epsilon_{\text{risk}} = 0.1$)	68	28	4
SQRL ($\epsilon_{\text{risk}} = 0.05$)	8	22	70
LatentSafe	**80**	**20**	**0**


3rd Person Camera
Wrist Camera
Skittles

World Model: DINO-WM
1300 trajectories
- 1000 random
- 150 safe demos
- 150 unsafe demos
- Manually labeled

Our Latent Safety Filter $(\pi^{\shield}, V^{\shield})$

Freely allows safe grasp…

Sliding motion is ***filtered to slow*** *…*

Unsafe pickup is ***filtered to stop!***

Our Latent Safety Filter $(\pi^{\shield}, V^{\shield})$

Robot POV

$V(z)$