FREA: Feasibility-Guided Generation of Safety-Critical Scenarios with Reasonable Adversariality

Keyu Chen[1], Yuheng Lei[2], Hao Cheng[1], Haoran Wu[1], Wenchao Sun[1], Sifa Zheng[1]

[1]Tsinghua University, [2]The University of Hong Kong.

11/08/2024

What are safety-critical scenarios

Vehicle avoids jaywalker

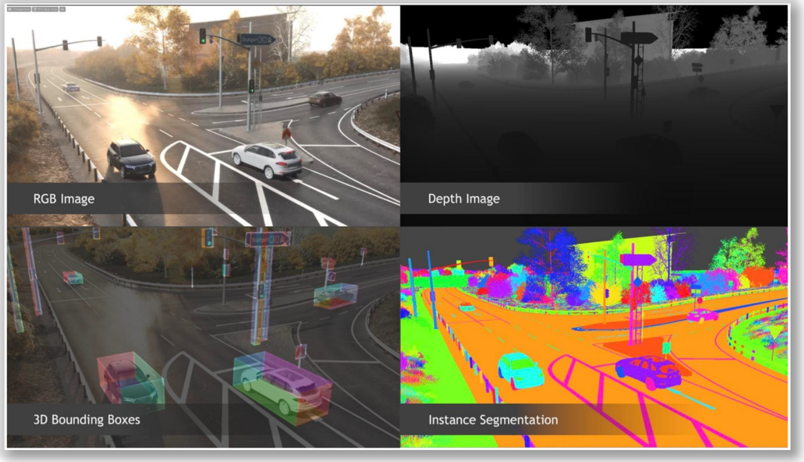
Vehicle avoids unusual objects

Vehicle avoids dumped truck

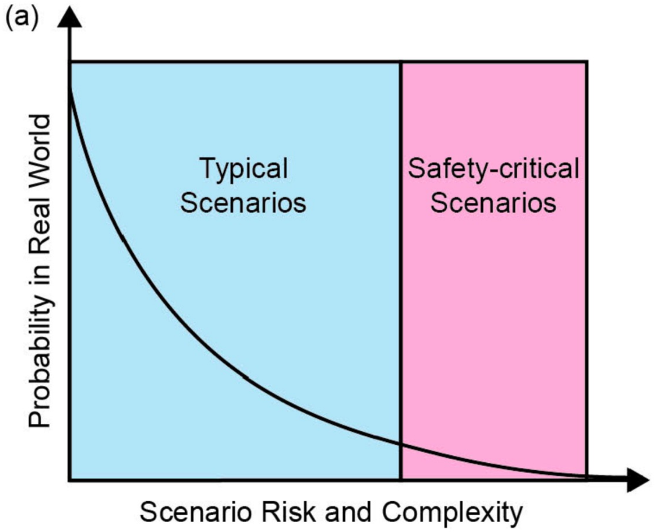
Vehicles encountering glare

Why generate safety-critical scenarios

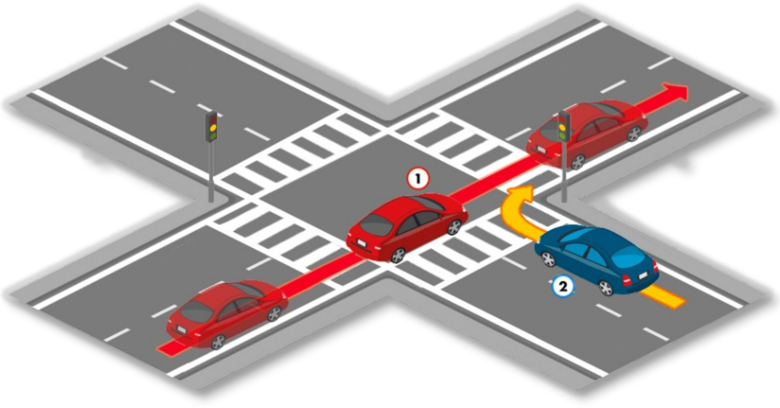
Driving data matters



Long-tailed data



Hand-crafted scenarios

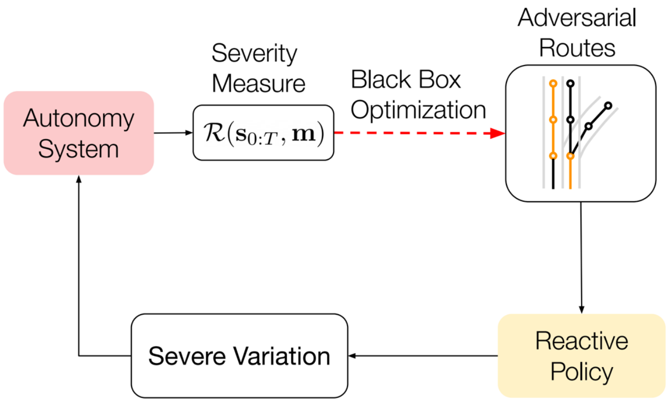


- Learned driving models are **brittle to o.o.d. input** and **require diverse training data**
- Critical scenarios are **rarely observed** in real world
- Hand-crafted scenarios in current simulators needs to be **manually tuned**

How to generate safety-critical scenarios

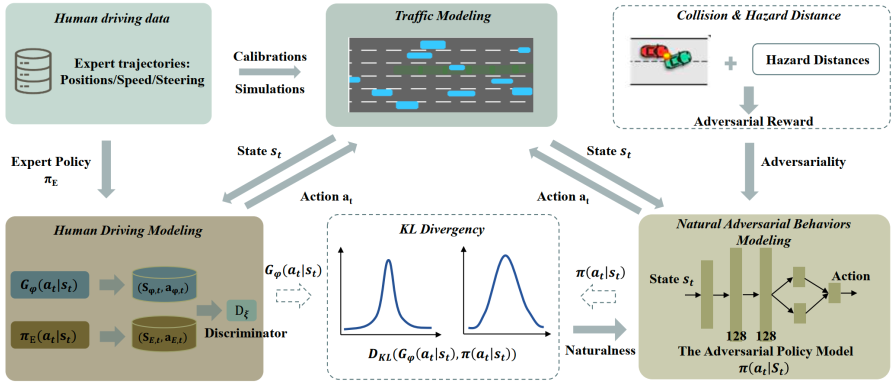
Safety-Critical Scenarios as attacks

Optimization-based[1,2,3]



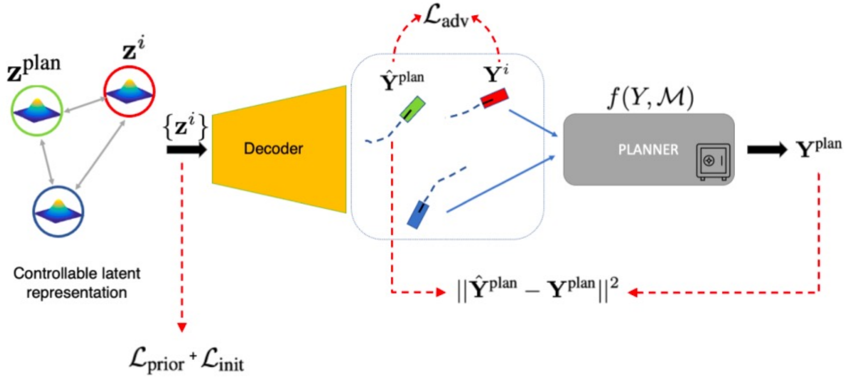
Collision Objective

RL-based[4,5]



Collision Reward

DGM-based[6,7]



Collision Loss

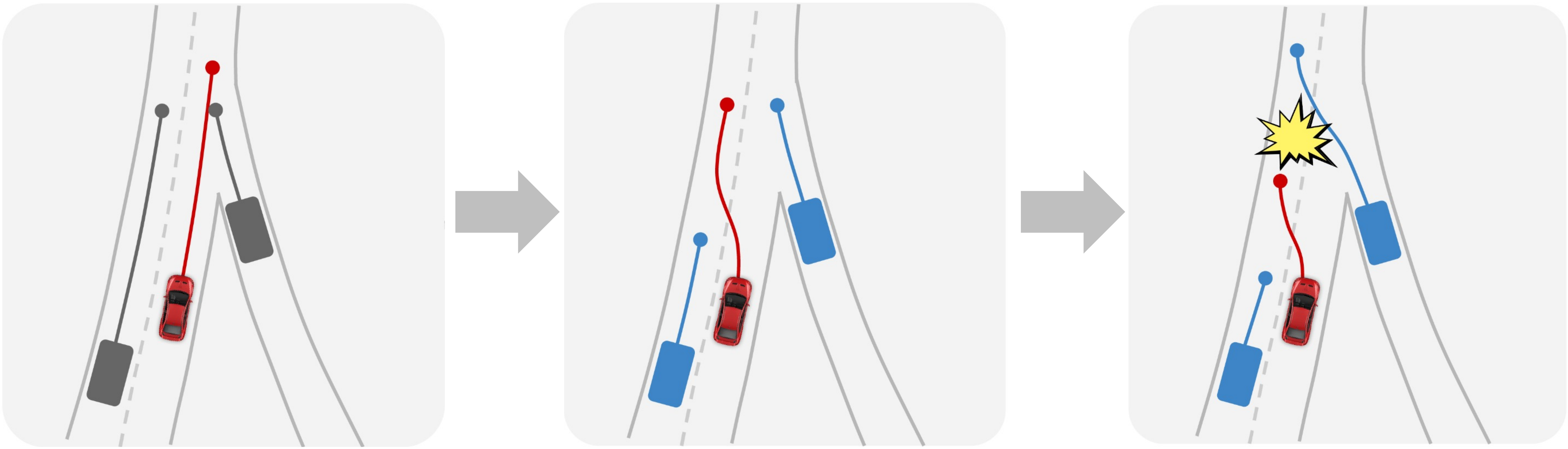
Is collision-driven adversarial attacks the right way?

[1] KING: Generating Safety-Critical Driving Scenarios for Robust Imitation via Kinematics Gradients [Hanselmann et al. ECCV 2022]
[2] MIXSIM: A Hierarchical Framework for Mixed Reality Traffic Simulation [Suo et al. CVPR 2023]
[3] CAT: Closed-loop Adversarial Training for Safe End-to-End Driving [Zhang et al. CoRL 2023]
[4] Dense reinforcement learning for safety validation of autonomous vehicles [Feng et al. Nature 2023]
[5] Adversarial Safety-Critical Scenario Generation using Naturalistic Human Driving Priors [Hao et al. TIV 2023]
[6] CausalAF: Causal Autoregressive Flow for Safety-Critical Driving Scenario Generation [Ding et al. CoRL 2022]
[7] Generating Useful Accident-Prone Driving Scenarios via a Learned Traffic Prior [Rempe et al. CVPR 2022]

Is collision-driven adversarial attacks the right way?

Tsinghua University

Real-world driving behavior is often goal-driven or route-driven

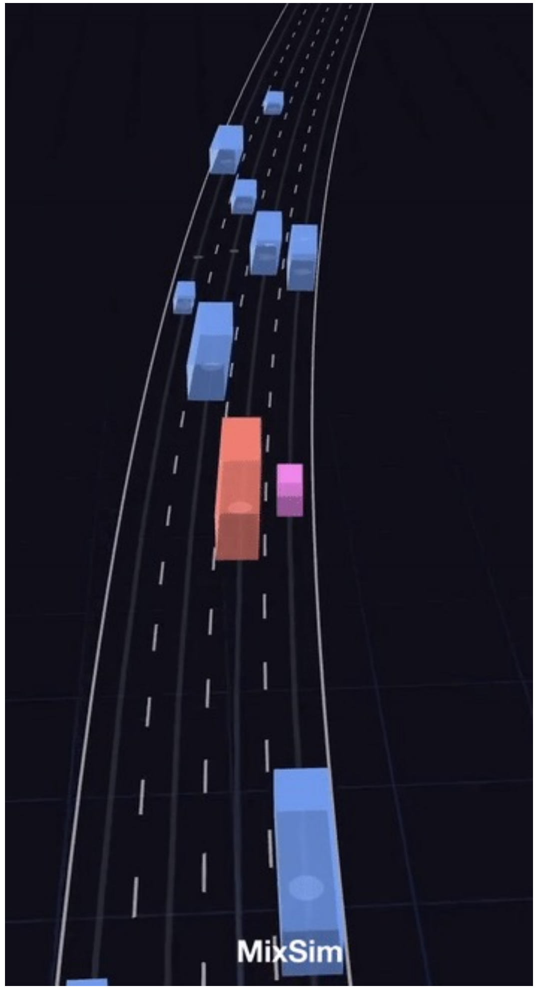


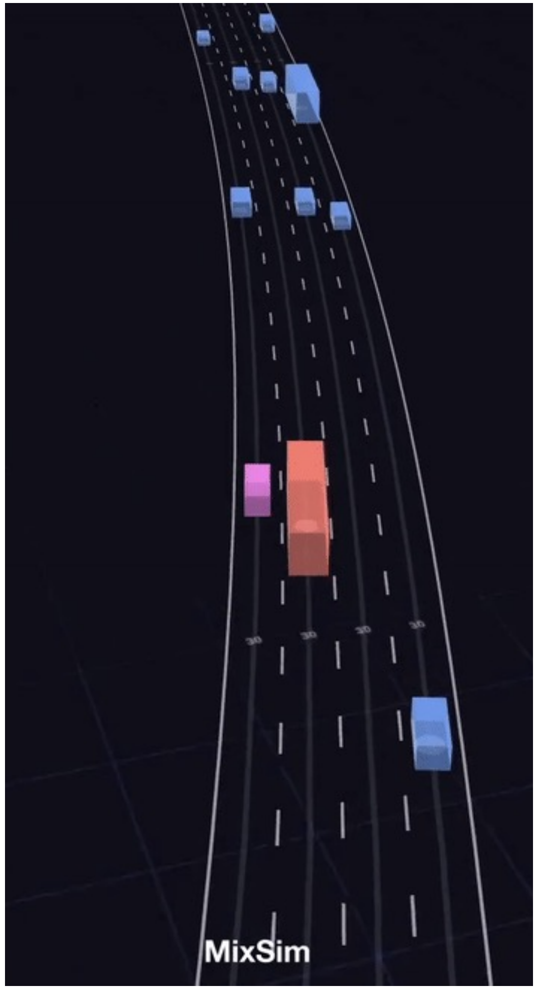
Same driving destinations and aggressive behaviors lead to safety-critical scenarios.

Collision is the consequence, not the objective

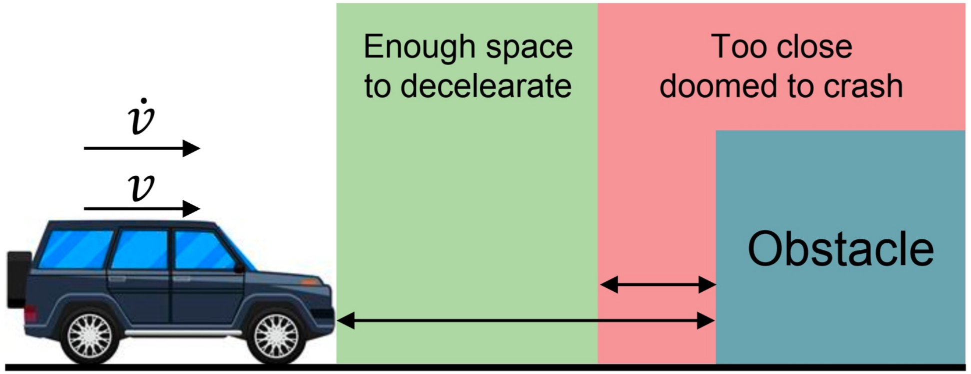
Image Source: MIXSIM: A Hierarchical Framework for Mixed Reality Traffic Simulation [Suo et al. CVPR 2023]

Is collision-driven adversarial attacks the right way?





Collision-driven attacks often create **unreasonable** and **AV infeasible** scenarios

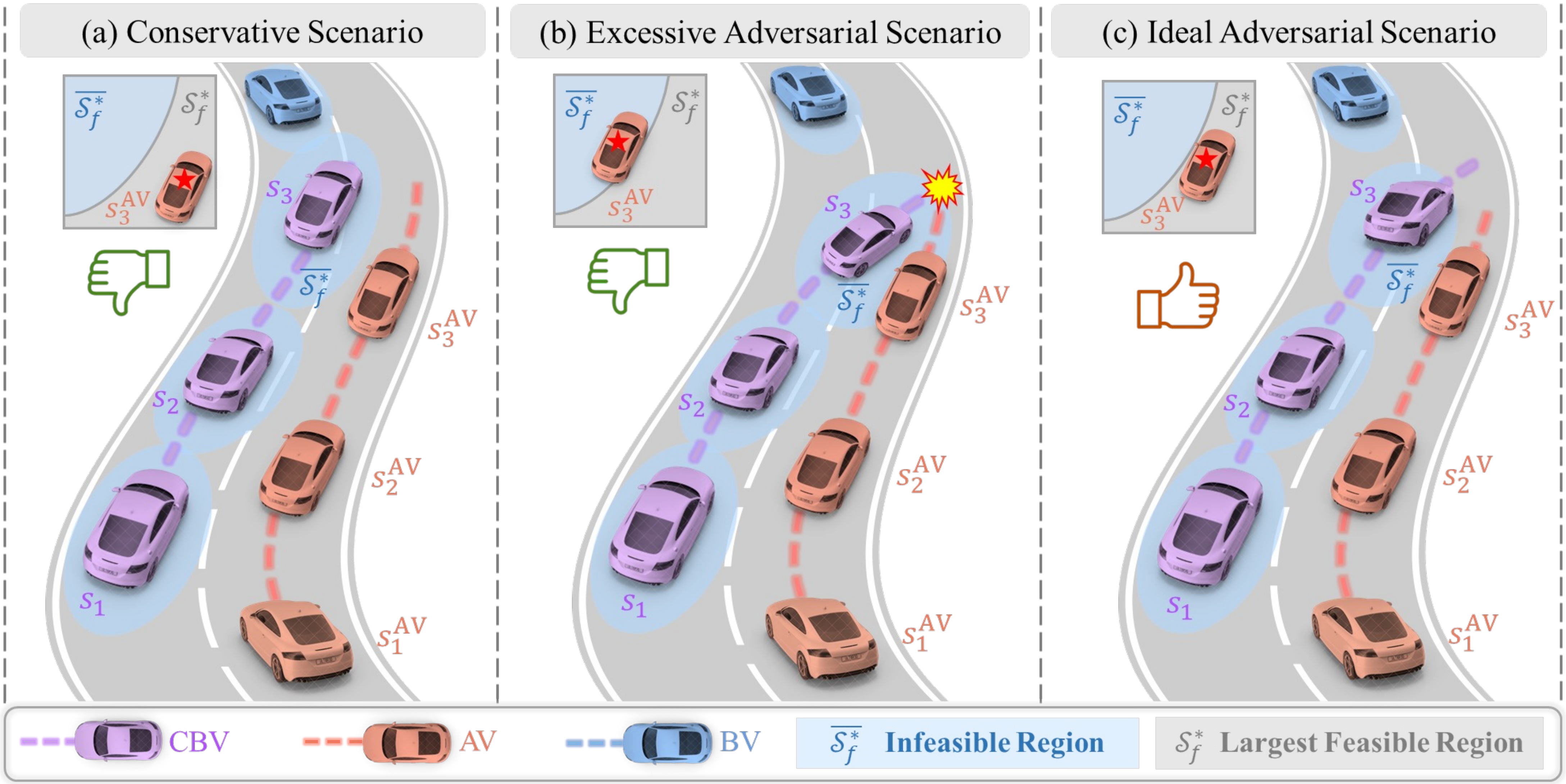


AV infeasible: no policy ensures AV's persistent safety

(AV doomed to crash)

Video Source: MIXSIM: A Hierarchical Framework for Mixed Reality Traffic Simulation [Suo et al. CVPR 2023]
Image Source: Reachability Constrained Reinforcement Learning [Yu et al. ICML 2022]

Create scenarios with reasonable adversariality



Create **adversarial** while **AV-feasible** safety-critical scenarios

Create Reasonable Adversarial Scenarios

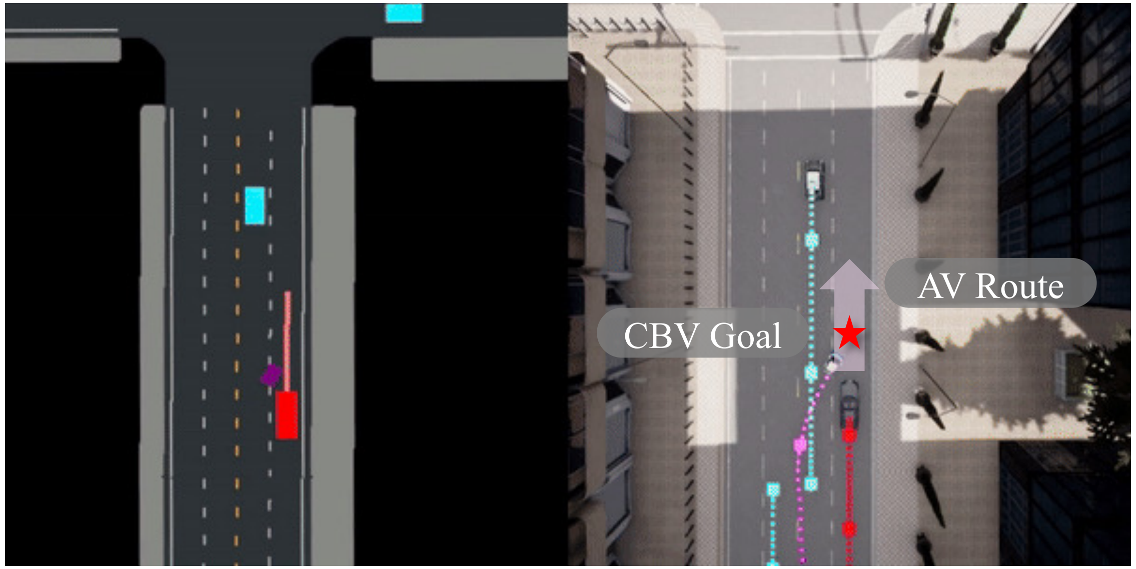
Select Critical Background Vehicle (CBV)

Candidate Selection Rules

- Within 25 meters of the AV
- In the same lane as the AV
- Not the CBV that has reached the goal
- …

The closest candidate is the CBV

Goal-Based RL Reward Setting



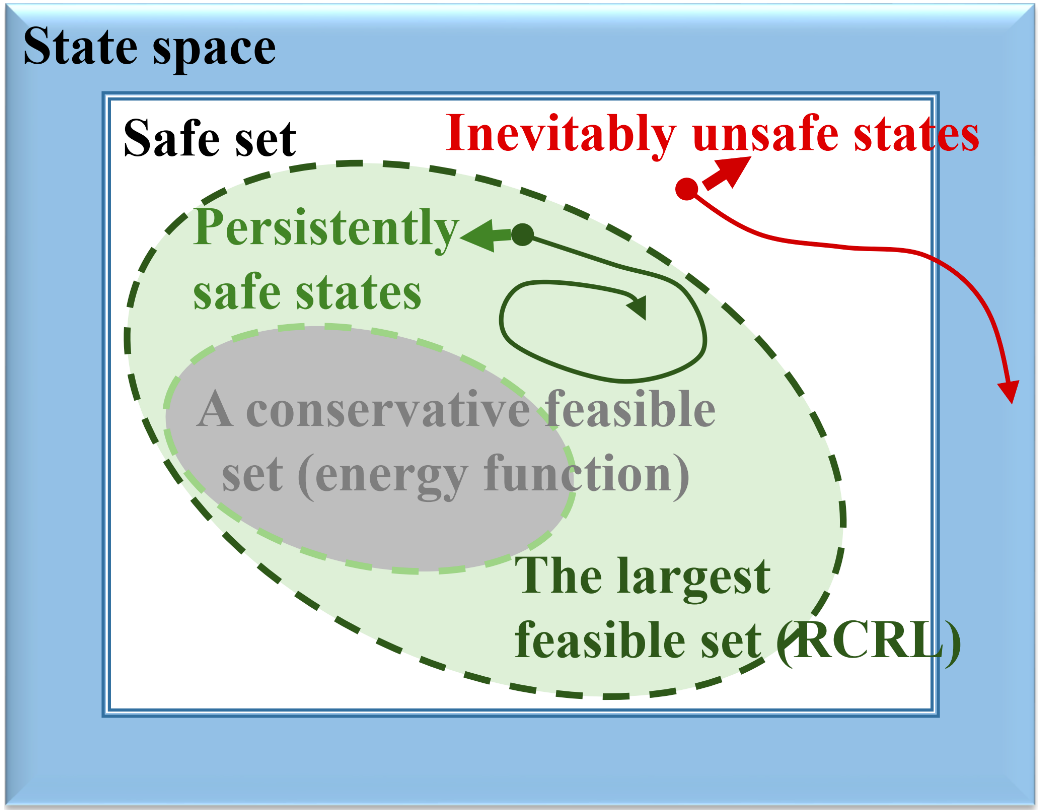
Goal-Based RL Reward

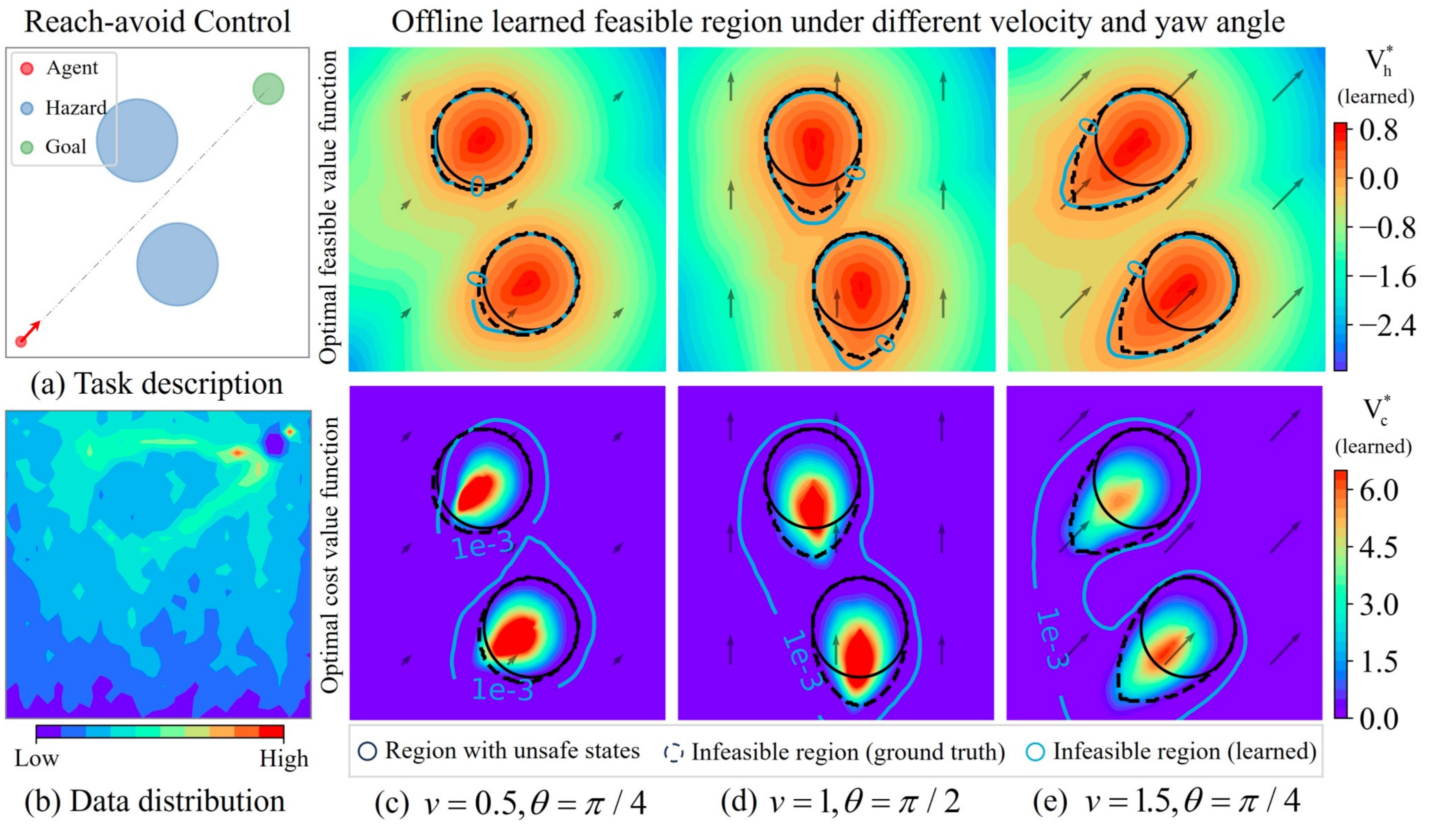
- Goal is the reference point in the AV route

$$R_t = d(\mathrm{CBV}_{t-1}, \mathrm{Goal}_{t-1}) - d(\mathrm{CBV}_t, \mathrm{Goal}_t) + 15 * r_t^{collision} + 15 * r_t^{finish}$$

Encourage the CBV to reach the goal

Establish Largest Feasible Region (LFR) of AV





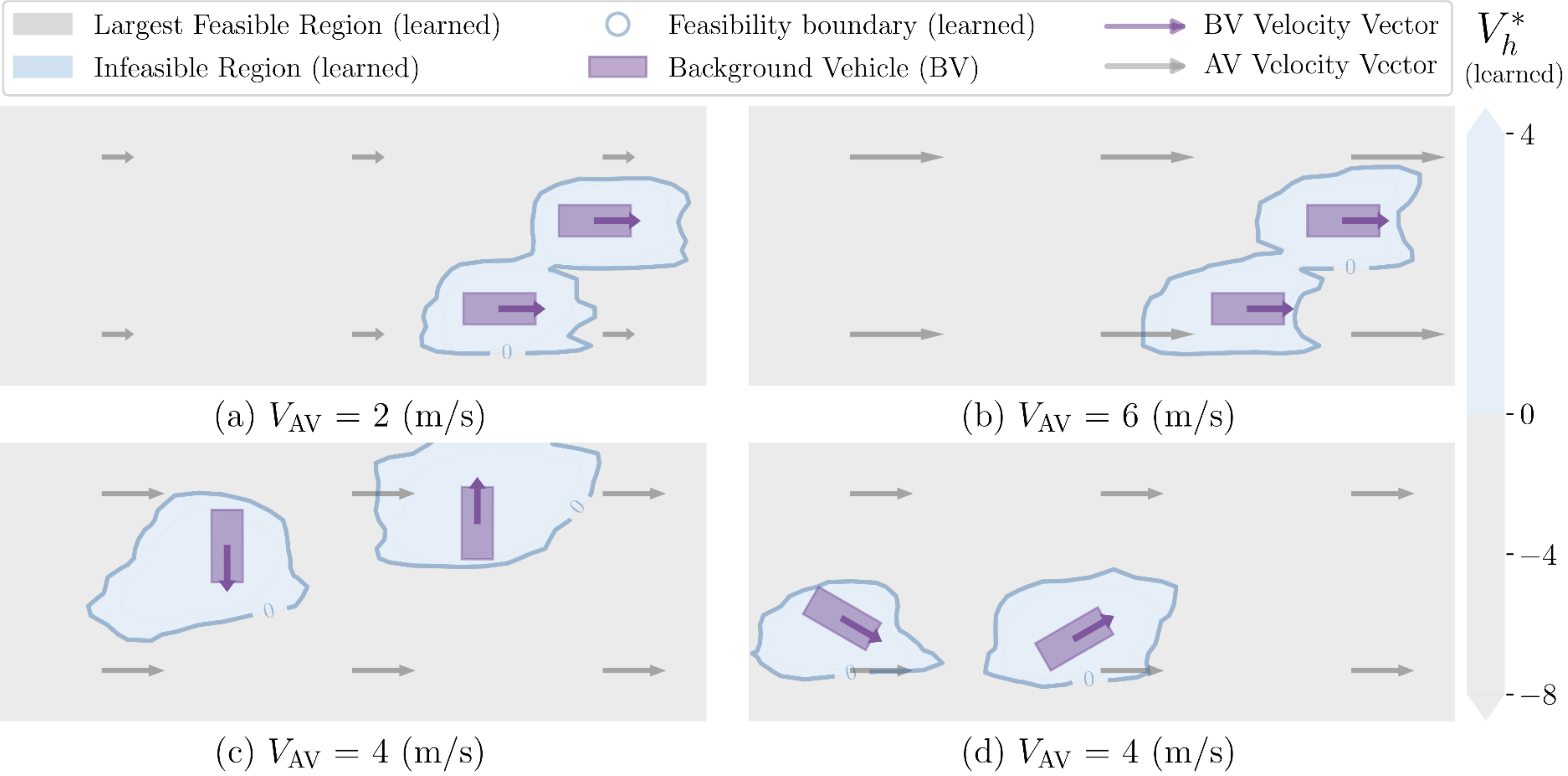
LFR focus on persistent safety instead of the current safety

Image Source:
Reachability Constrained Reinforcement Learning [Yu et al. ICML 2022]
Safe Offline Reinforcement Learning with Feasibility-Guided Diffusion Model [Zheng et al. ICLR 2024]

Create AV-feasible Scenarios

Establish largest feasible region (LFR) of AV

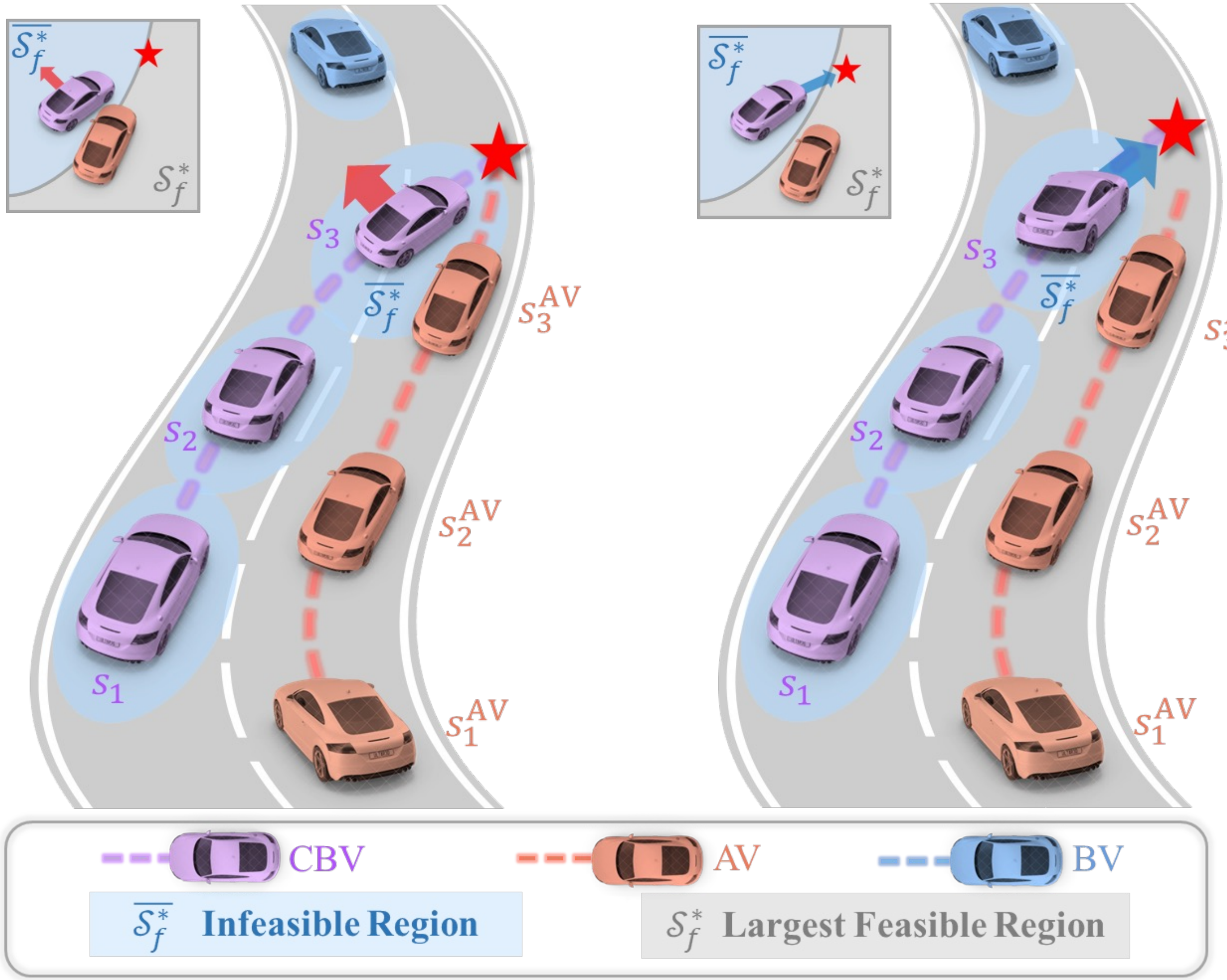
Learning Feasible Value Function from Offline Data[1]



[1] Safe Offline Reinforcement Learning with Feasibility-Guided Diffusion Model [Zheng et al. ICLR 2024]

Create AV-feasible Scenarios

Train CBV with feasibility-dependent objective



Feasibility-depend advantage

$$A(s,a) = \begin{cases} A_r(s,a) & \text{AV-feasible} \\ -A_h^*(s^{AV}, a^{AV}) & \text{AV-infeasible} \end{cases}$$

AV-feasible

Goal-based adversarial advantage function

$$A_r(s,a) = Q_r(s,a) - V_r(s)$$

AV-infeasible

AV's optimal feasibility advantage function

$$A_h^*(s^{AV}, a^{AV}) = Q_h^*(s^{AV}, a^{AV}) - V_h^*(s^{AV})$$

Create AV-feasible Scenarios

清華大學 Tsinghua University

Train CBV with feasibility-dependent objective

Algorithm 1 Feasibility-guided reasonable adversarial policy (*FREA*)

```
Offline Part (Section 3.1)
Initialize feasibility value networks V_h, Q_h.
for each gradient step do
    Update V_h using Eq. (3)    # Optimal feasible state-value function learning
    Update Q_h using Eq. (4)    # Optimal feasible action-value function learning
end for
Online Part (Section 3.2)
Initialize policy parameters θ_0, reward value function parameters ψ_0
for k = 0, 1, 2, ... do
    Collect set of trajectories B_k = {τ_i} with policy π_{θ_k}, where τ_i is a T-step episode.
    Compute reward advantage A_r^{π_{θ_k}}(s, a), using generalized advantage estimator (GAE [23]).
    Compute feasibility advantage using Eq. (9).
    Derive overall advantage using Eq. (6)    # Advantage calculating
    Fit reward value function, by Smooth L1 Loss.    # Value function learning
    Update the policy parameters θ by maximizing Eq. (5).    # Policy learning
end for
```

Experiment

Tsinghua University

Questions need to answer

- Can FREA generate **adversarial** scenarios?
- Can FREA generate **AV-feasible** scenarios?
- How can FREA help **testing AV method**?
- How can FREA help **training AV method**?
- Any **scenario visualization**?

Baselines

- Standard (Rule-based traffic flow)
- PPO (Goal-based Reward)
- FPPO-RS (feasibility penalty)
- KING[1] (Optimization-based method)
- **FREA (Our method)**

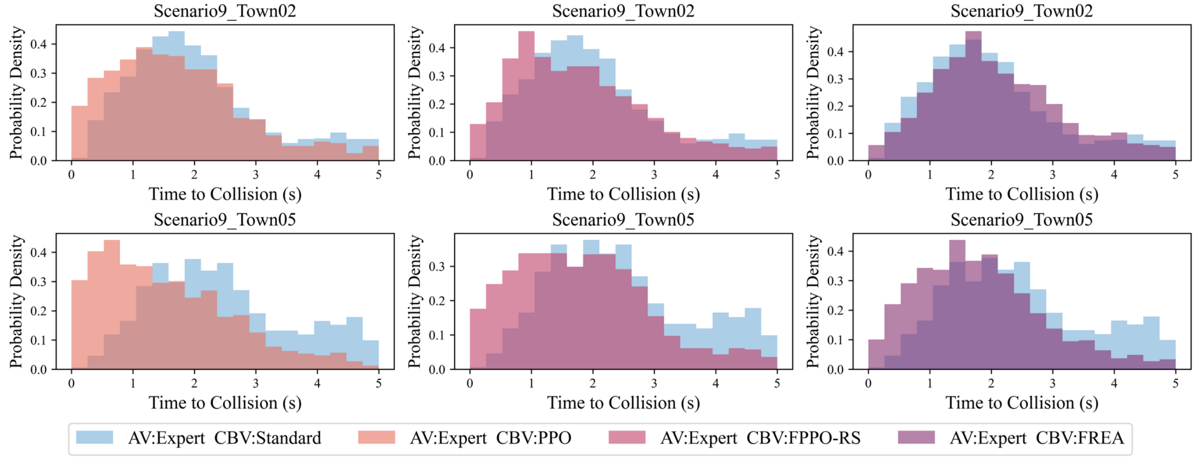
[1] KING: Generating Safety-Critical Driving Scenarios for Robust Imitation via Kinematics Gradients [Hanselmann et al. ECCV 2022]

Experimental Results

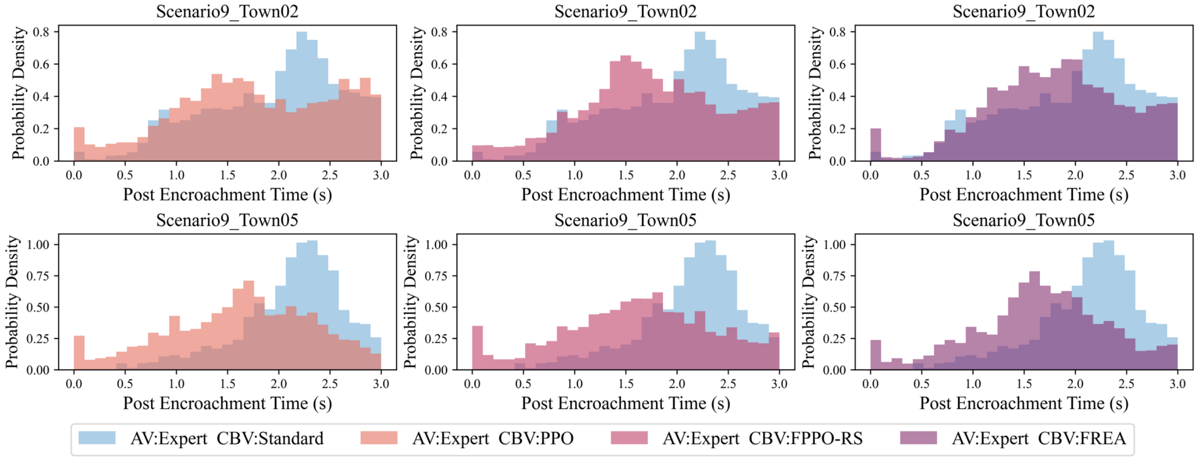
Tsinghua University

Can FREA generate adversarial scenarios?

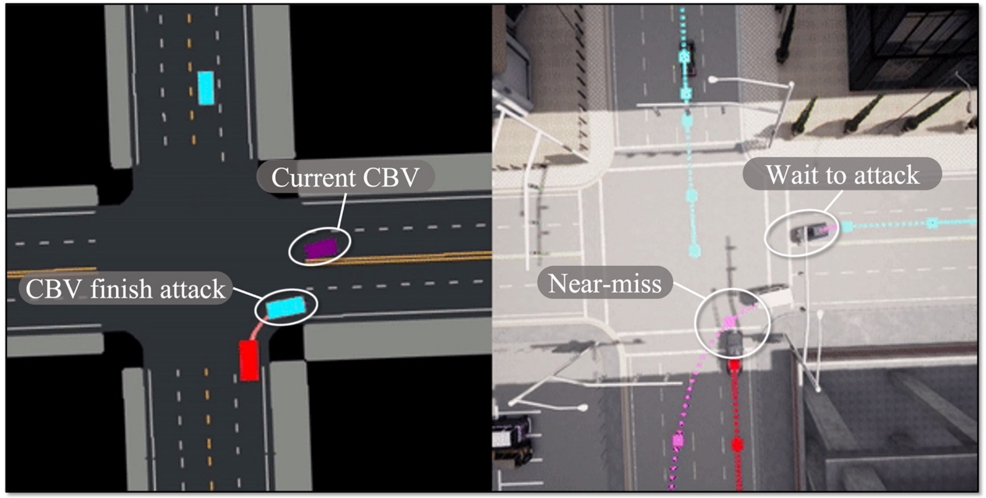
Time To Collision (TTC)

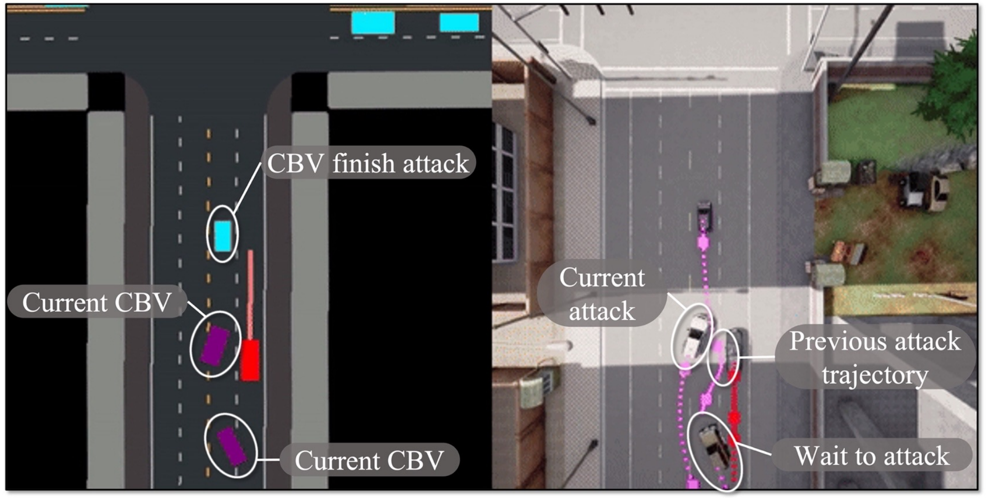


Post Encroachment Time (PET)



Near-Miss Event



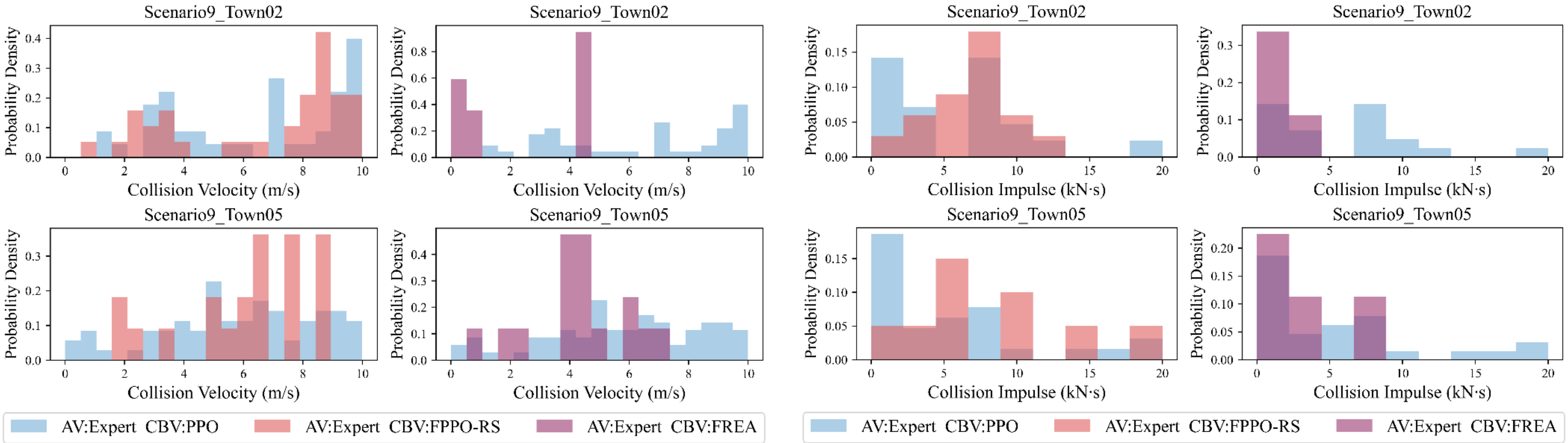


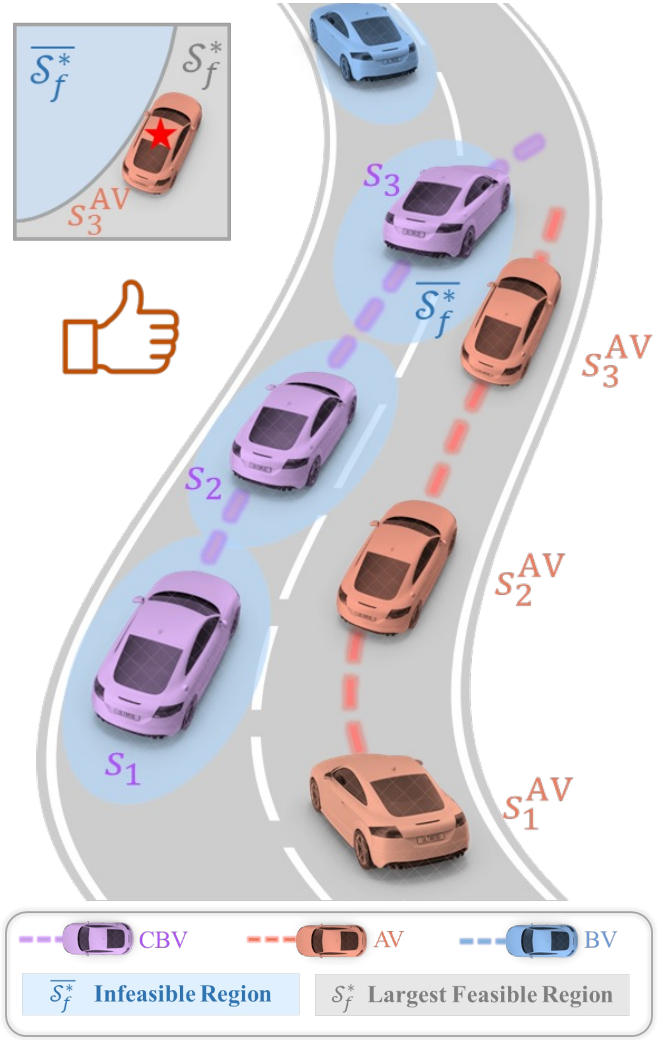
Experimental Results

Can FREA generate AV-feasible scenarios?

Table 1: Feasibility evaluation using Expert [26] as AV under different CBV methods. Results are the average of 10 runs in "Scenario9" with varied seeds.

CBV	Feasibility	Town05 intersections			Town02 intersections		
		CR (↓)	IR (↓)	ID (↓)	CR (↓)	IR (↓)	ID (↓)
KING[5]	✗	76.67%	65.97%	7.54m	N/A	N/A	N/A
PPO	✗	37.5%	35.56%	10.57m	30.0%	51.18%	12.40m
FPPO-RS	✓	11.25%	34.92%	9.13m	24.29%	45.36%	9.16m
FREA	✓	**5.0%**	**31.10%**	**6.25m**	**5.71%**	**27.18%**	**4.94m**





Experimental Results

FREA generalizes well in AV testing

Table 2: Comparative performance of AVs across different maps, using CBV methods pre-trained with various surrogate AVs. Results are the average of 10 runs in "Scenario9" with varied seeds.

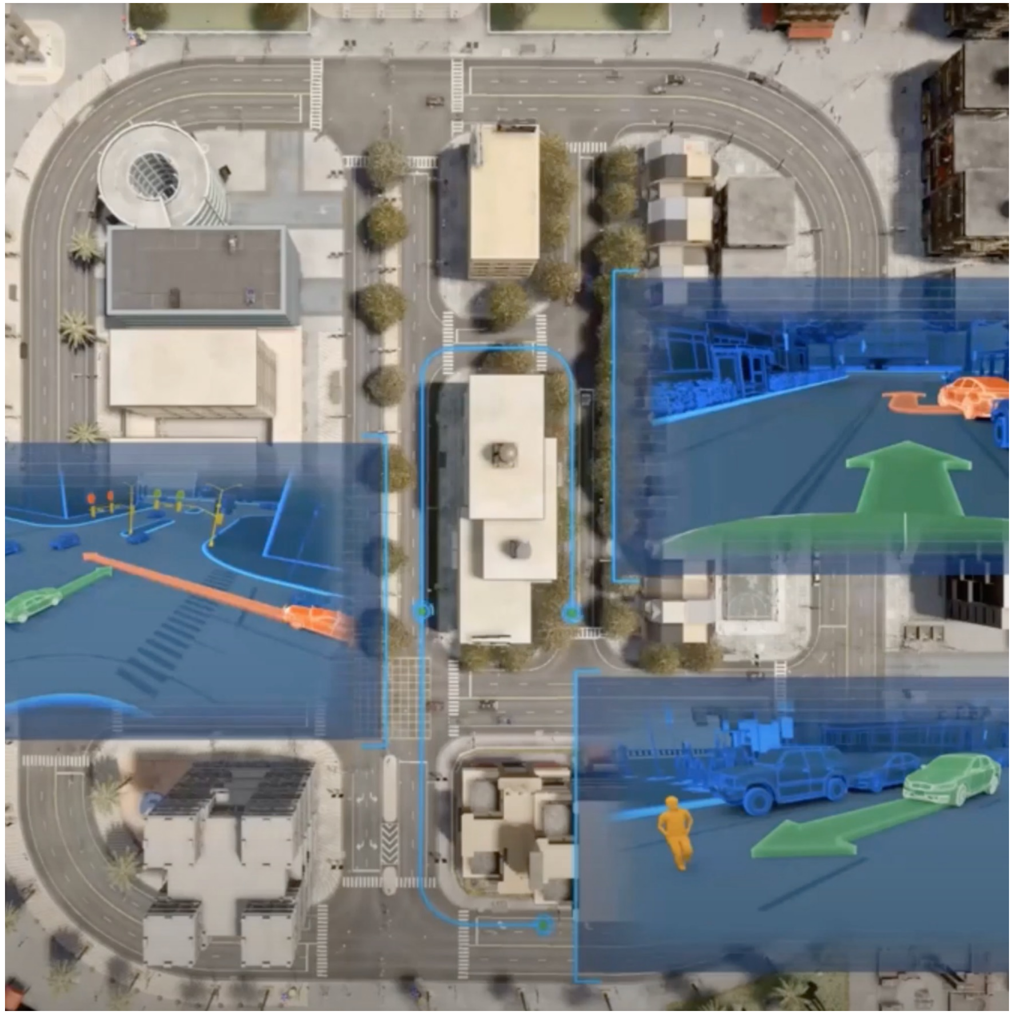
CBV	Surr. AV	AV	Town05 intersections CR (↓)	OR (↓)	RF (↓)	UC (↓)	TS (↓)	OS (↑)	Town02 intersections CR (↓)	OR (↓)	RF (↓)	UC (↓)	TS (↓)	OS (↑)
Standard	✗	Expert	0.0%	0.0m	7.0m	1%	55s	94.0	0.0%	0.0m	6.0m	2%	63s	93.0
		PlanT	1.0%	0.0m	7.0m	6%	70s	90.0	1.0%	0.0m	6.0m	6%	76s	90.0
PPO	Expert	Expert	36.0%	0.0m	6.0m	8%	66s	76.0	40.0%	0.0m	6.0m	15%	66s	72.0
		PlanT	61.0%	1.0m	7.0m	11%	70s	65.0	70.0%	0.0m	6.0m	27%	64s	57.0
PPO	PlanT	Expert	26.0%	0.0m	6.0m	8%	64s	80.0	21.0%	0.0m	6.0m	12%	74s	80.0
		PlanT	45.0%	0.0m	7.0m	7%	69s	72.0	51.0%	0.0m	6.0m	18%	70s	67.0
FREA	Expert	Expert	4.0%	0.0m	7.0m	7%	67s	**89.0**	9.0%	0.0m	6.0m	16%	75s	**83.0**
		PlanT	10.0%	0.0m	7.0m	5%	73s	**86.0**	10.0%	0.0m	7.0m	24%	86s	**79.0**
FREA	PlanT	Expert	5.0%	0.0m	7.0m	5%	62s	**90.0**	14.0%	0.0m	6.0m	15%	75s	**82.0**
		PlanT	9.0%	0.0m	7.0m	6%	73s	**87.0**	17.0%	0.0m	7.0m	18%	83s	**79.0**

FREA-trained AV shows better performance

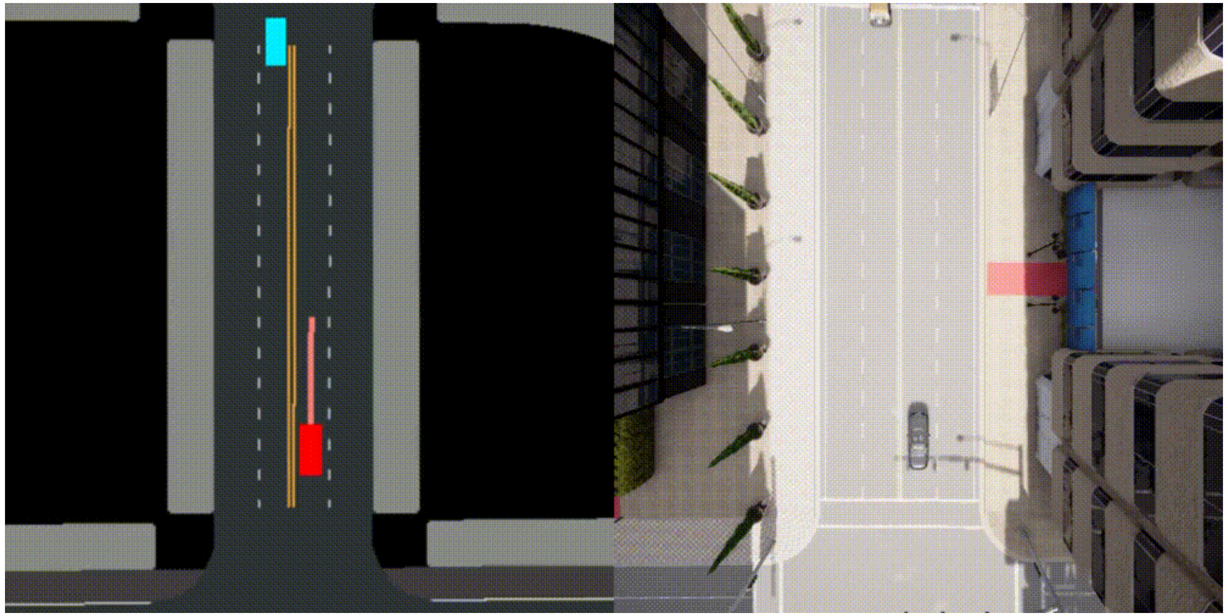
Table 5: Comparative performance of AVs pretrained with various CBV methods across different maps. Results are the average of 10 runs in "Scenario9" with varied seeds.

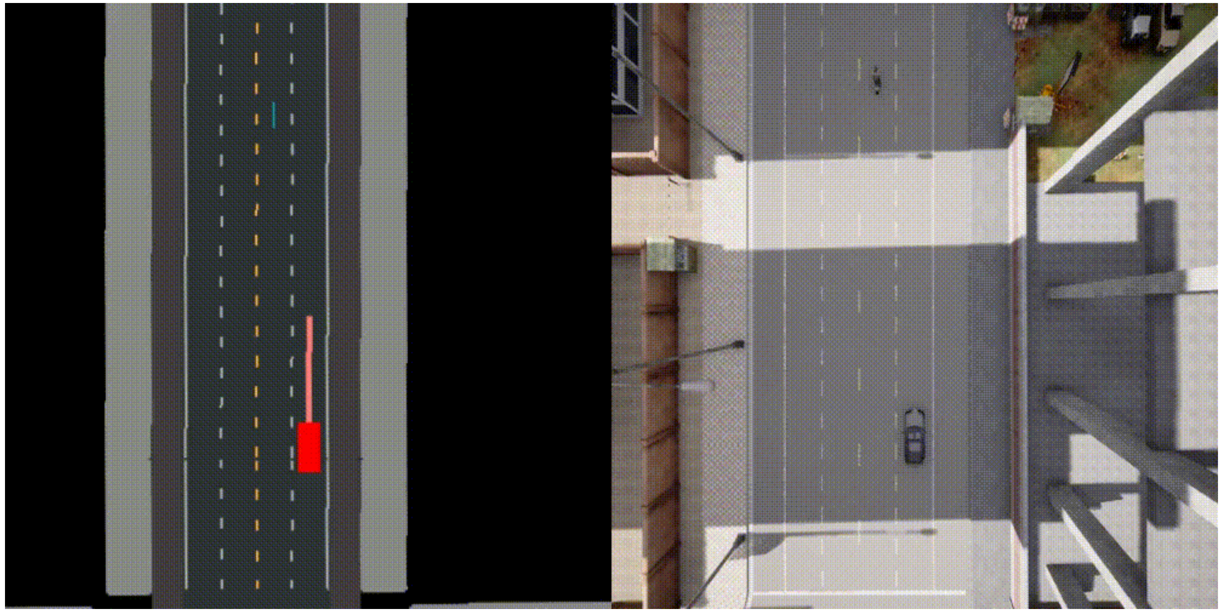
Surr. CBV	CBV	Town05 intersections CR (↓)	OR (↓)	RF (↓)	UC (↓)	TS (↓)	OS (↑)	Town02 intersections CR (↓)	OR (↓)	RF (↓)	UC (↓)	TS (↓)	OS (↑)
Standard	Standard	11%	4m	19m	4%	68s	85	**39%**	8m	20m	18%	57s	70
PPO		17%	11m	17m	6%	66s	82	40%	11m	19m	15%	63s	70
FREA		**3%**	6m	18m	1%	75s	**89**	40%	3m	18m	19%	56s	**71**
Standard	**FREA**	39%	6m	17m	7%	66s	73	79%	3m	17m	35%	45s	51
PPO		36%	15m	21m	6%	66s	74	79%	4m	14m	37%	44s	51
FREA		**31%**	12m	17m	5%	71s	**76**	**73%**	3m	20m	28%	51s	**55**

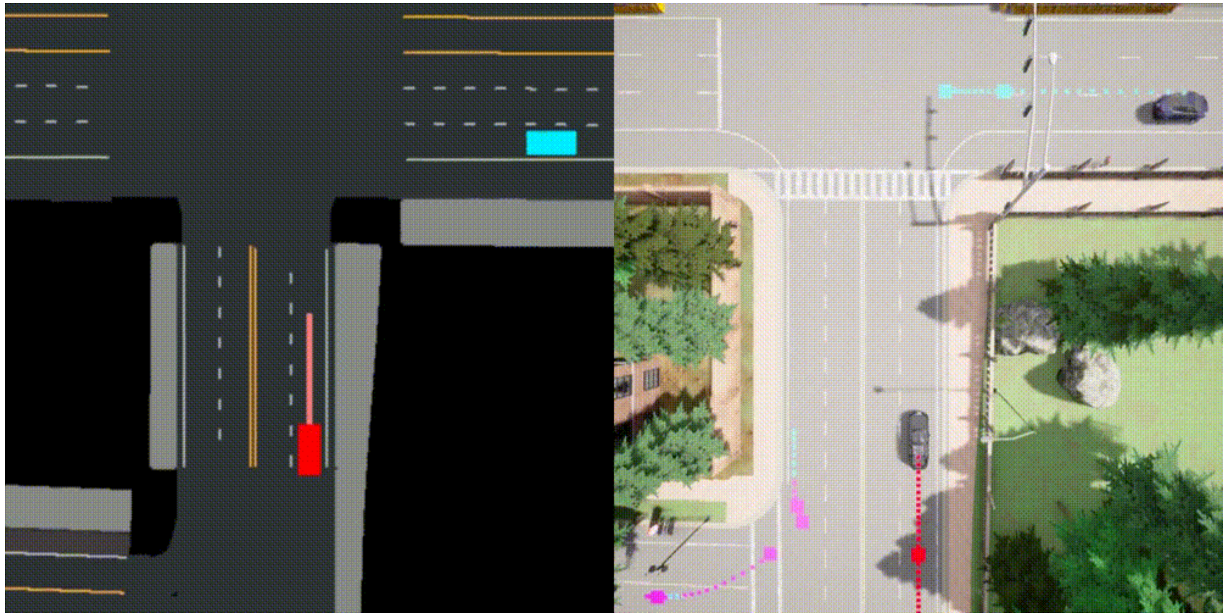
SafeBench

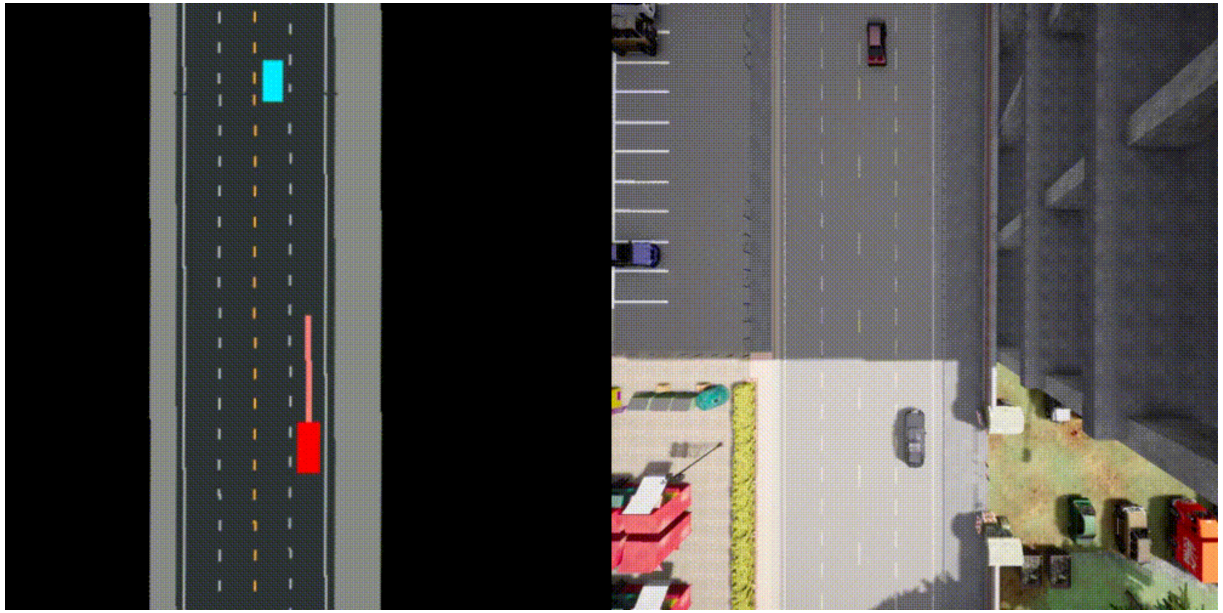


Representative Scenarios









Limitations / Future Work

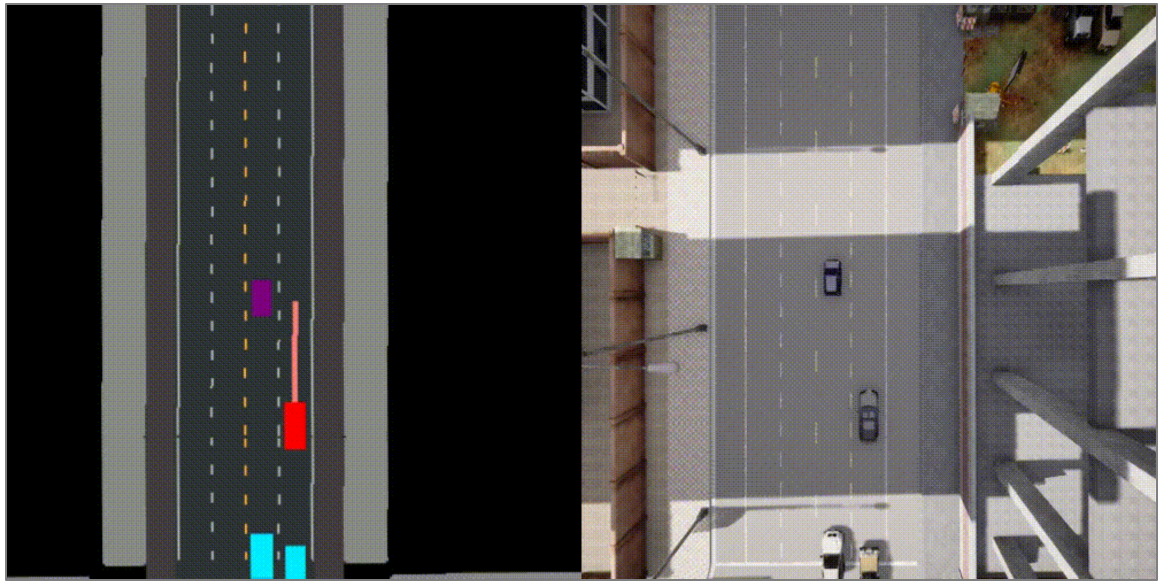
Tsinghua University

Naturalness

- Route-based instead of goal-based
- Agent behavior modeling

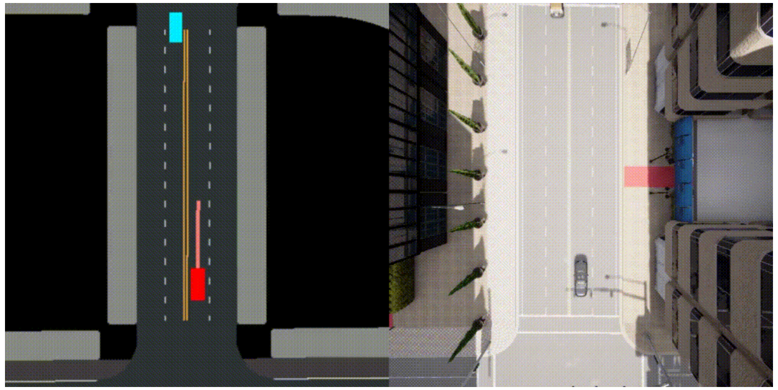
Traffic regulation

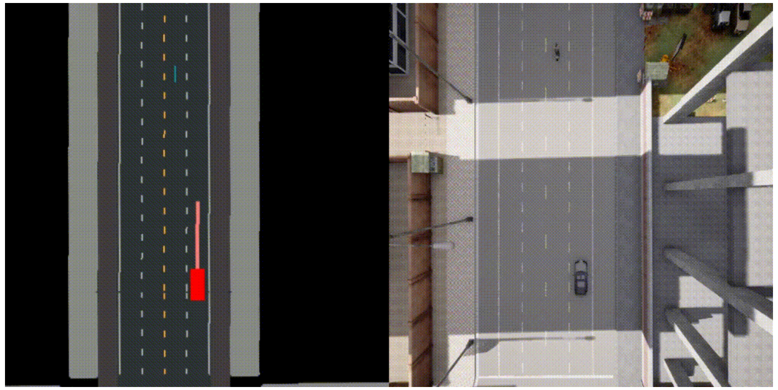
- Integrating LLM for common sense
- Integrating traffic rules

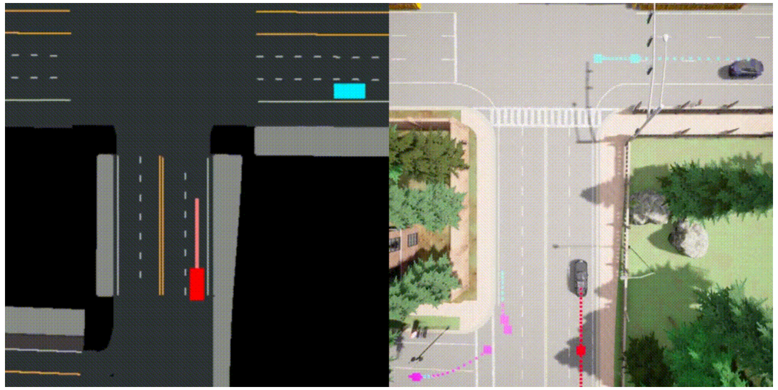


Unnatural Behavior

Lacking Common Sense







FREA: Feasibility-Guided Generation of Safety-Critical Scenarios with Reasonable Adversariality

Q&A