ISAACS: Iterative Soft Adversarial Actor-Critic for Safety

Kai-Chieh Hsu*, Duy Phuong Nguyen*, Jaime F. Fisac

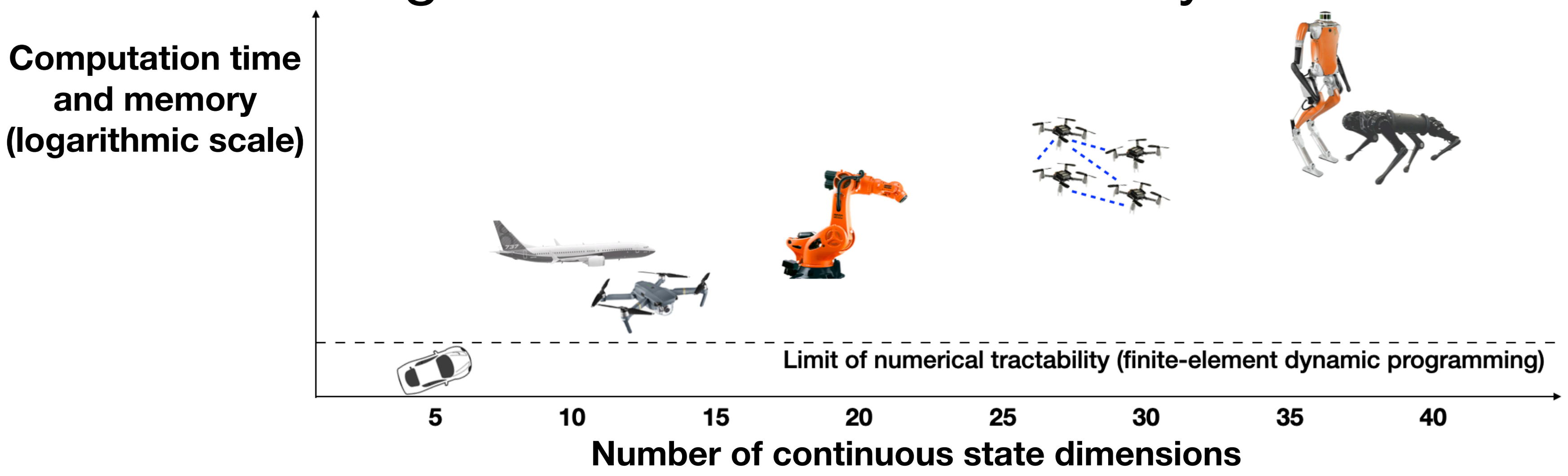
5th L4DC

PRINCETON UNIVERSITY

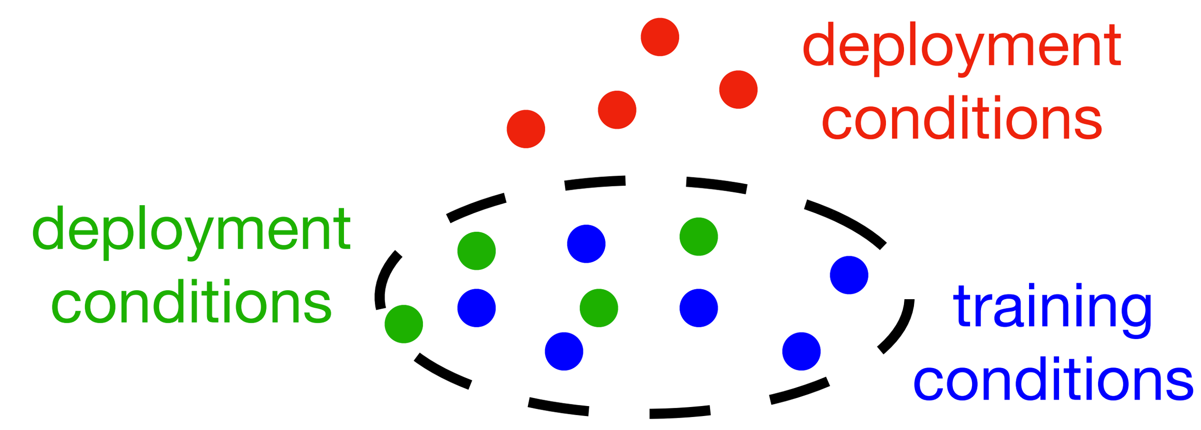
SAFE ROBOTICS LABORATORY

Scalable Safety Analysis in Robotics

Rigorous **robust optimal control** tools scale poorly to high-dimensional nonlinear dynamics.



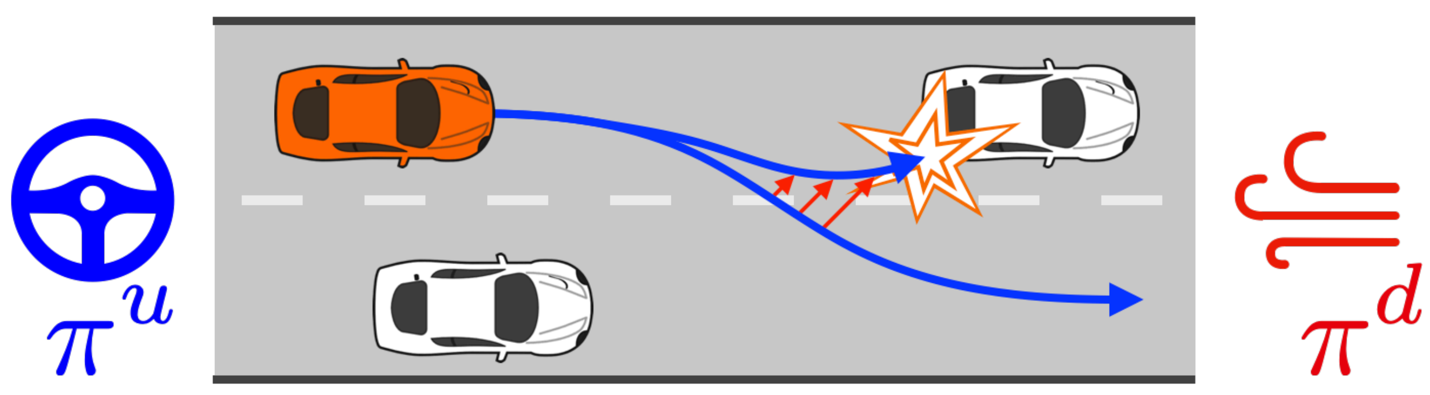
Scalable **deep learning** methods lack guarantees and are brittle to "out-of-distribution" conditions.



Can we devise computational safety tools for uncertain, high-dimensional dynamical systems without renouncing strict safety guarantees?

Offline Safety Synthesis (ISAACS)

1. State-action sequences collected by a series of ***simulated adversarial safety games***



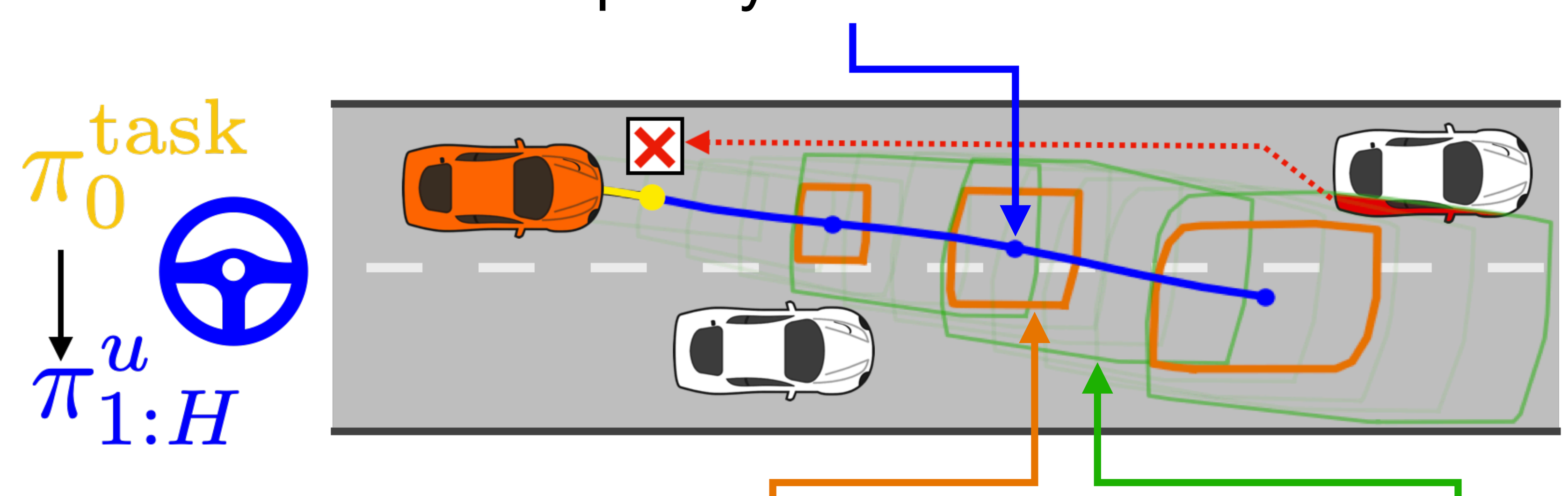
2. Controller/disturbance policies (*actors*) & safety value (*critic*) co-trained via ***safety Isaacs equation***:

$$V(x) = (1-\gamma)g(x) + \gamma \max_{\pi_u} \min_{\pi_d} \mathop{\mathbb{E}}_{u,d} \min\left\{g(x), Q(x,u,d)\right\}$$

3. Leaderboard update via ***cross-play***: rank policy versions by win rate vs. all opponents, keep top *k*

Online Safety Certification

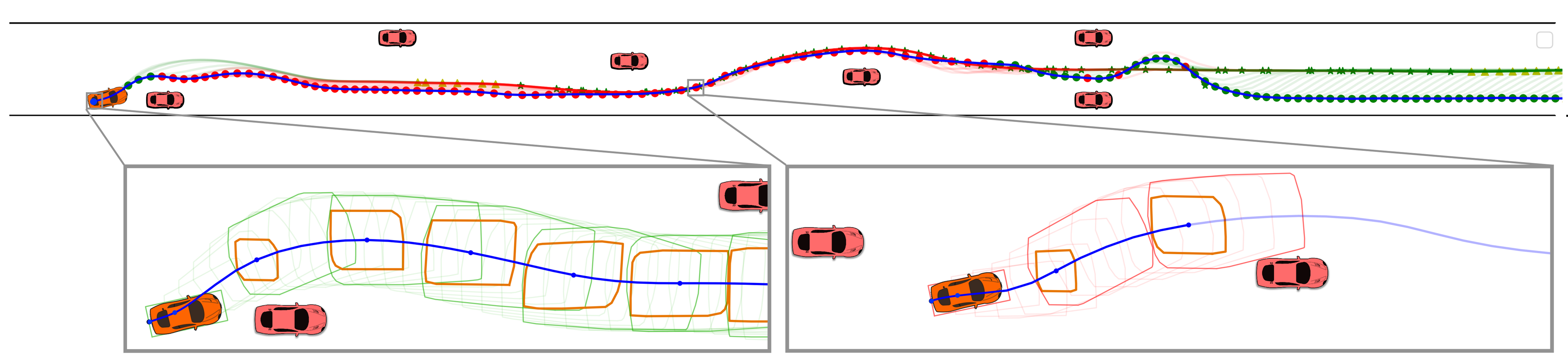
1. ***Reference fallback trajectory*** by rolling out the *untrusted* ISAACS policy after *1 task-driven action*



2. Robust fallback rollout through ***forward-reachable sets (tracking bounds)***

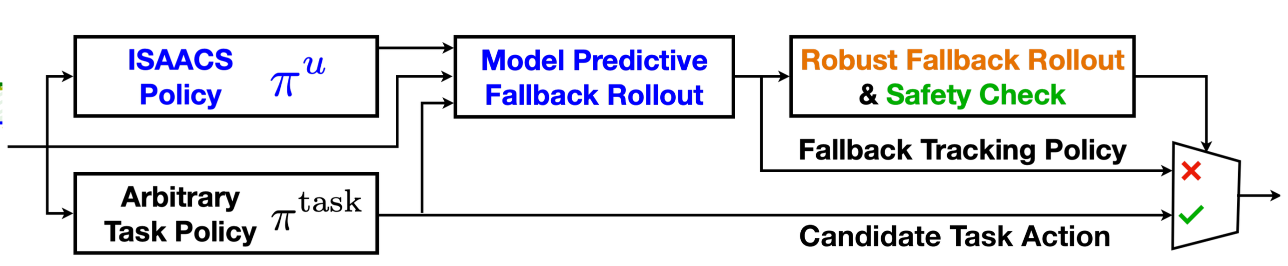
3. Certification by checking ***footprint-augmented*** forward-reachable sets for collisions/violations

Safety Filter



Safe → allow task-driven action

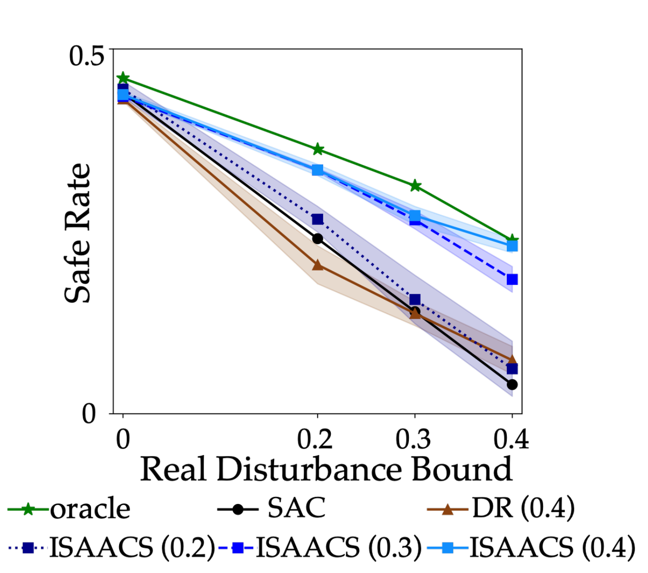
Unsafe → override with fallback action



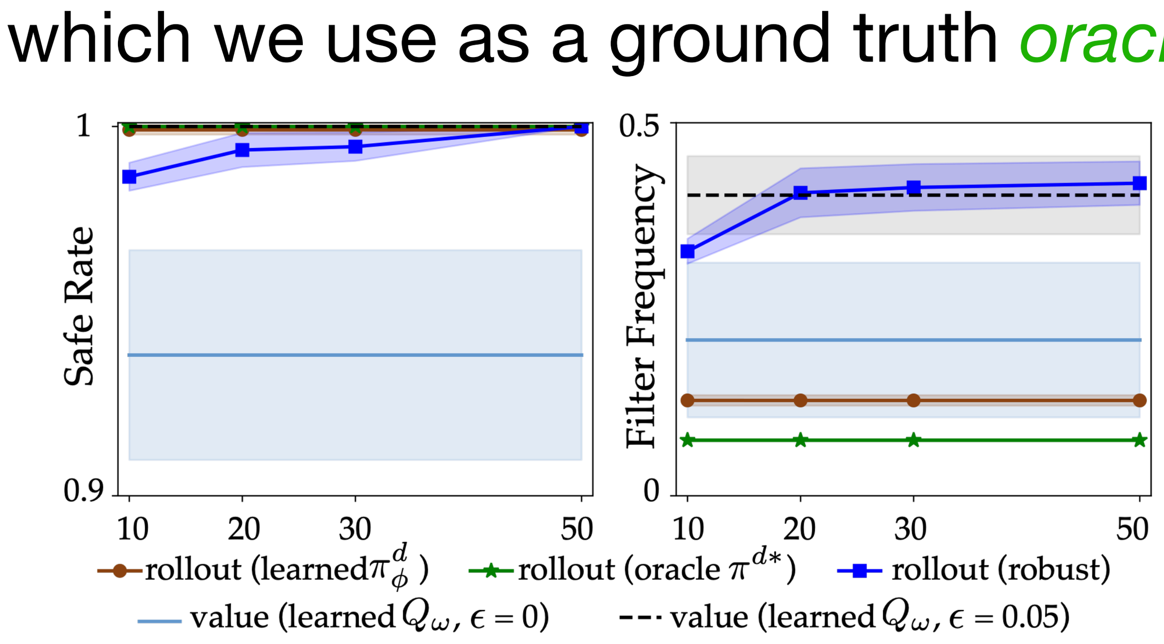
At each time, the task policy's control is allowed *if* it leaves open the option to track the fallback policy *later* and still maintain safety under all uncertainty realizations.

Evaluation: autonomous car

We validate ISAACS on a 5D car system, at the computability limit of "exact" numerical solutions, which we use as a ground truth *oracle*.



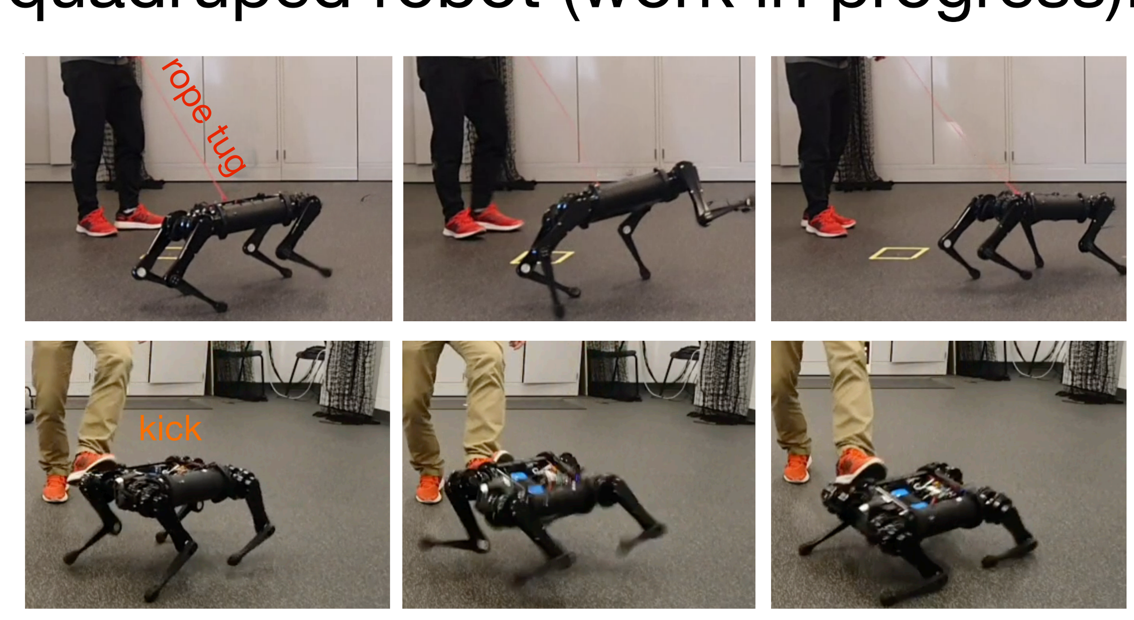
Co-training safety policies and worst-case realizations **boosts robustness** to model error.



The ISAACS robust rollout safety filter achieves a perfect 100% safe rate (0 violations). The direct gameplay rollout/value filters *can* fail, but rarely do!

Sneak peek: quadruped

We are testing ISAACS on a 36D quadruped robot (work in progress).



The ISAACS safety policy, learned purely in simulation and deployed with a value filter, responds to various attacks to prevent falls.