Reinforcement Learning based Control of Imitative Policies for Near-Accident Driving

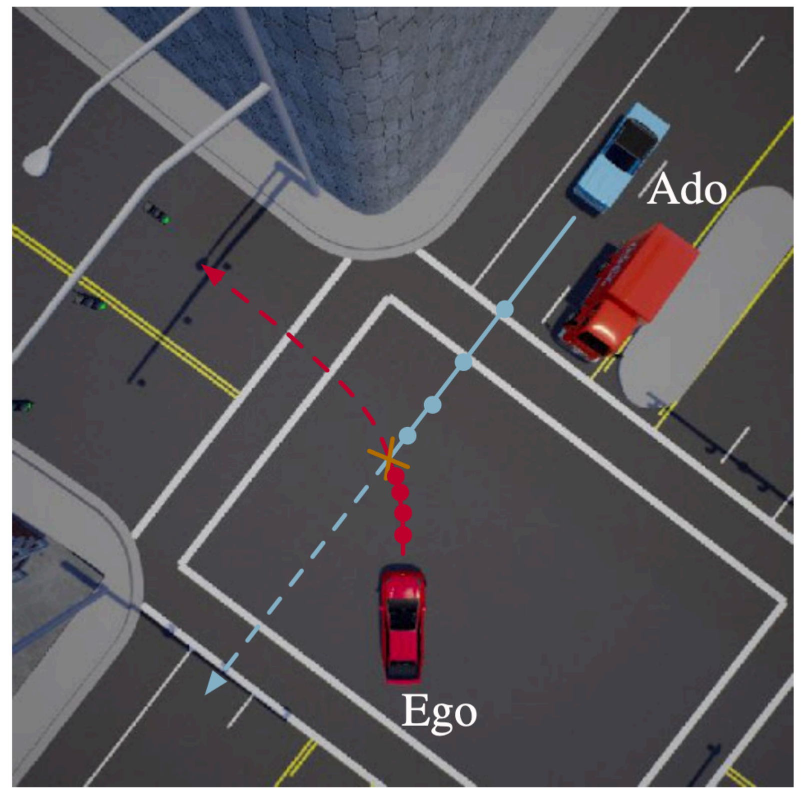
Zhangjie Cao*, Erdem Bıyık*, Woodrow Z. Wang,
Allan Raventos, Adrien Gaidon, Guy Rosman, Dorsa Sadigh

Introduction

Autonomous vehicles often fail in high-risk situations where a potential accident is likely

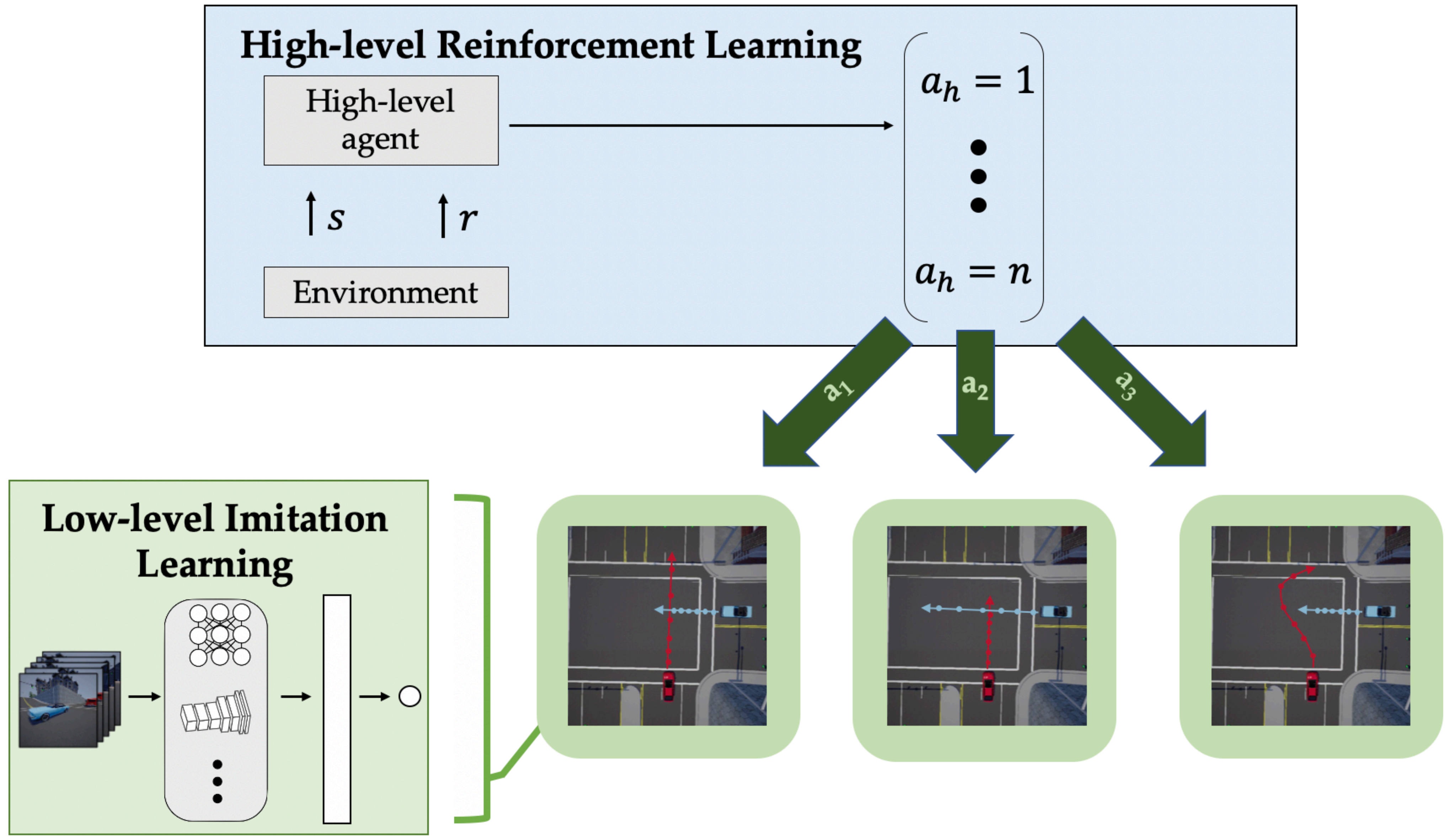
Our goal is to design safe and efficient policies for autonomous vehicles in near-accident scenarios

Problem



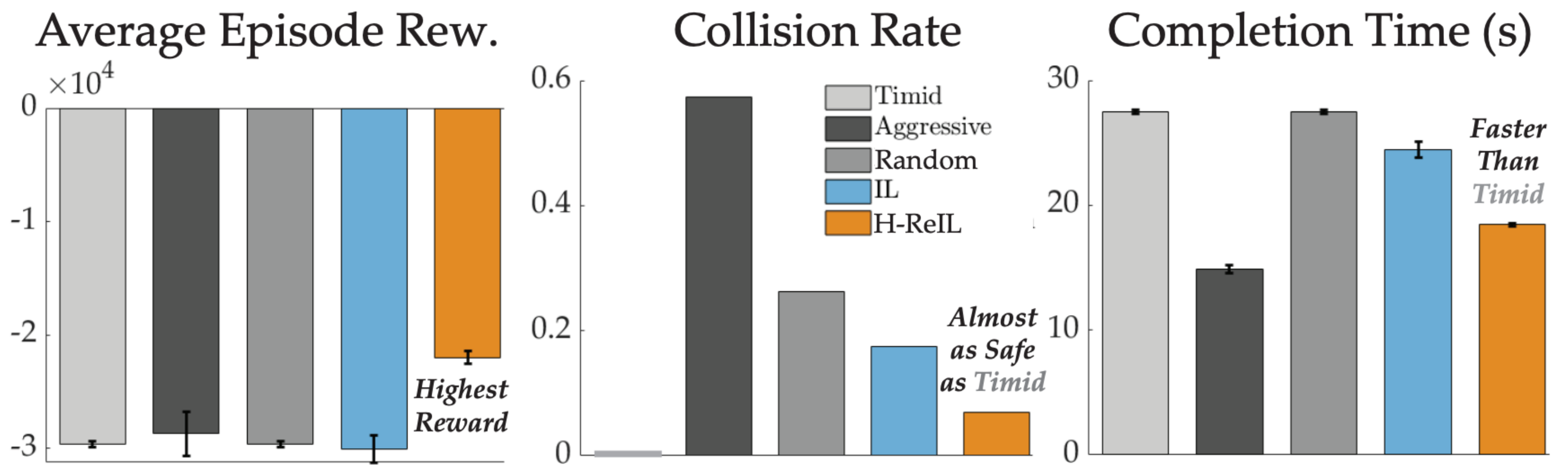
- How do we model rapid phase transitions?
- How do we drive safely and efficiently?

Model



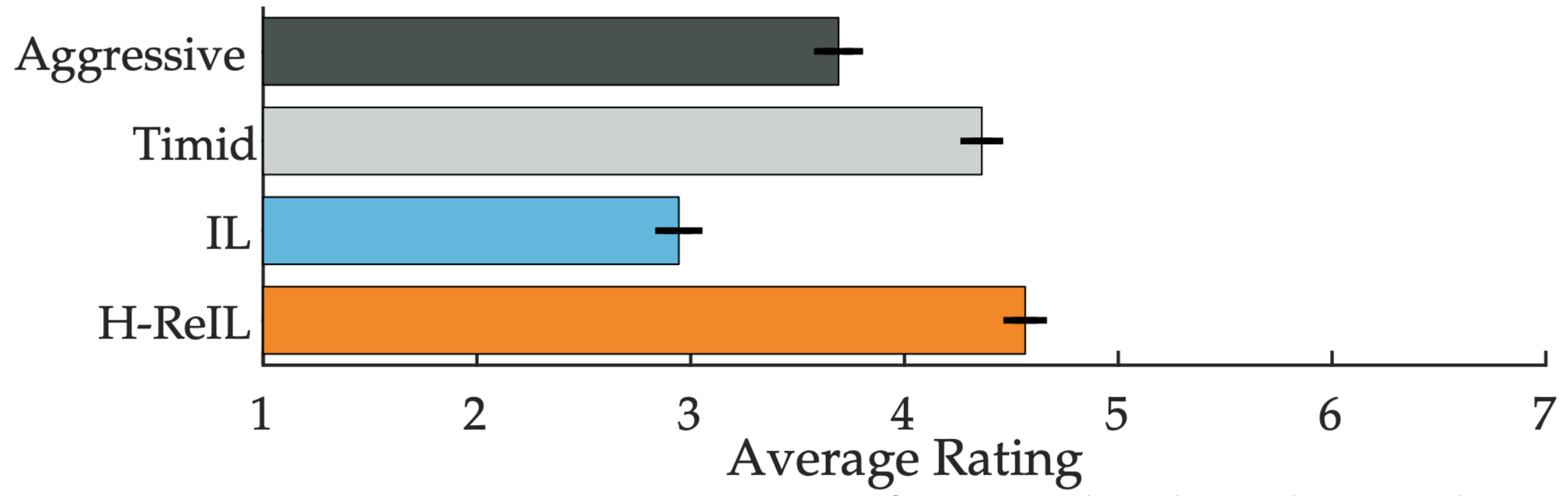
- **Key insight:** Phase transitions can be modeled as optimal switches, learned by reinforcement learning, between different modes of driving styles, each learned through imitation learning
- We propose H-ReIL, which is composed of a high-level policy learned with reinforcement learning that switches between different modes and low-level policies learned with imitation learning
- H-ReIL leverages the benefits of both RL and IL methods

Experimental Results



- Collision rate represents safety and completion time represents efficiency
- H-ReIL achieves a better balance between safety and efficiency compared to baseline methods

User Study Results



- Users rate H-ReIL significantly higher than baseline methods

References

1- Z. Cao, E. Bıyık, W. Wang, A. Raventos, A. Gaidon, G. Rosman, D. Sadigh. "Reinforcement Learning based Control of Imitative Policies for Near-Accident Driving,"in Robotics: Science and Systems, 2020

2-- F. Codevilla, M. Miiller, A. Lopez, V. Koltun, and A. Dosovitskiy, "End-to-end driving via conditional imitation learning," in 2018 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2018

3- T. D. Kulkarni, K. Narasimhan, A. Saeedi, and J. Tenenbaum, "Hierarchical deep reinforcement learning: Integrating temporal abstraction and intrinsic motivation," in Advances in neural information processing systems, 2016