Provably Safe Reinforcement Learning for Motion Planning

Hanna Krasowski

hanna.krasowski@tum.de

Supervisors: Matthias Althoff & Majid Zamani

Collaborators: Prithvi Akella (Caltech) & Aaron Ames (Caltech) & Niklas Kochdumper (Stony Brook) & Jakob Thumm (TUM) & Xiao Wang (TUM)

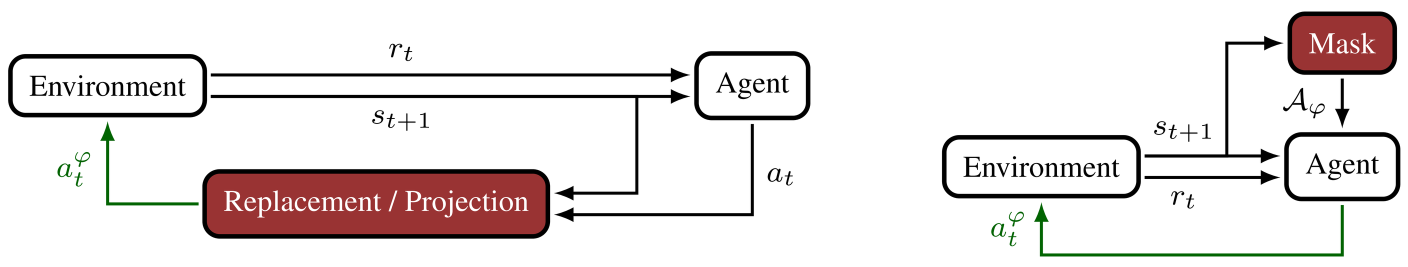
Research Questions

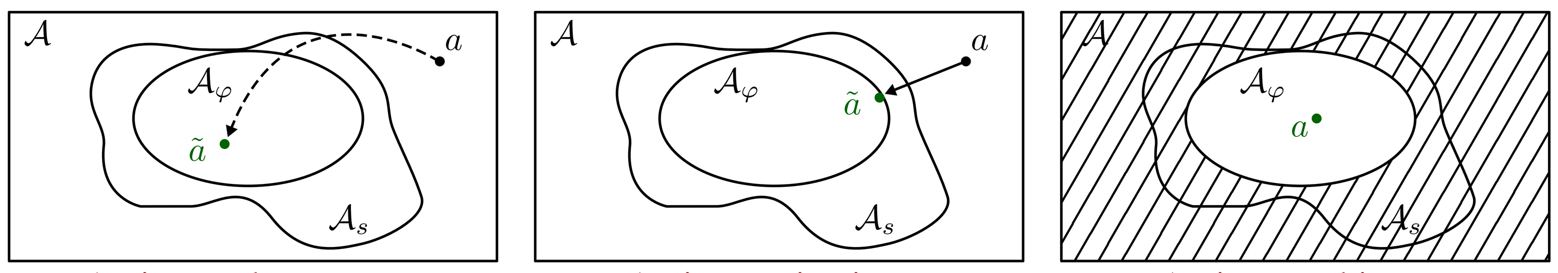
How can we extend reinforcement learning to achieve that safety specifications are always fulfilled?
How can we integrate complex safety specifications of cyber-physical systems in reinforcement learning?

Provably Safe Reinforcement Learning

Def. Provably safe RL provides guarantees for the safety specifications during learning and deployment. The specifications are task-specific, and there are various ways to define them [1].

Three approaches of provably safe RL:

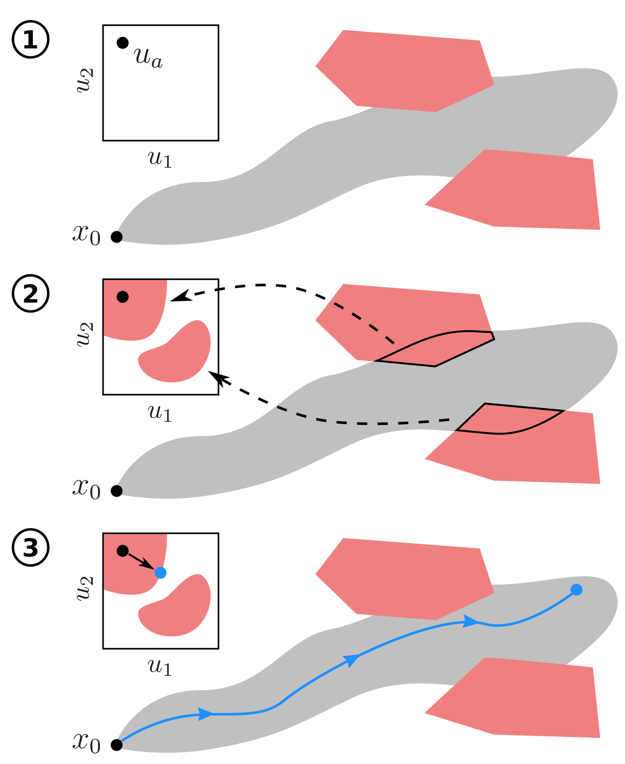




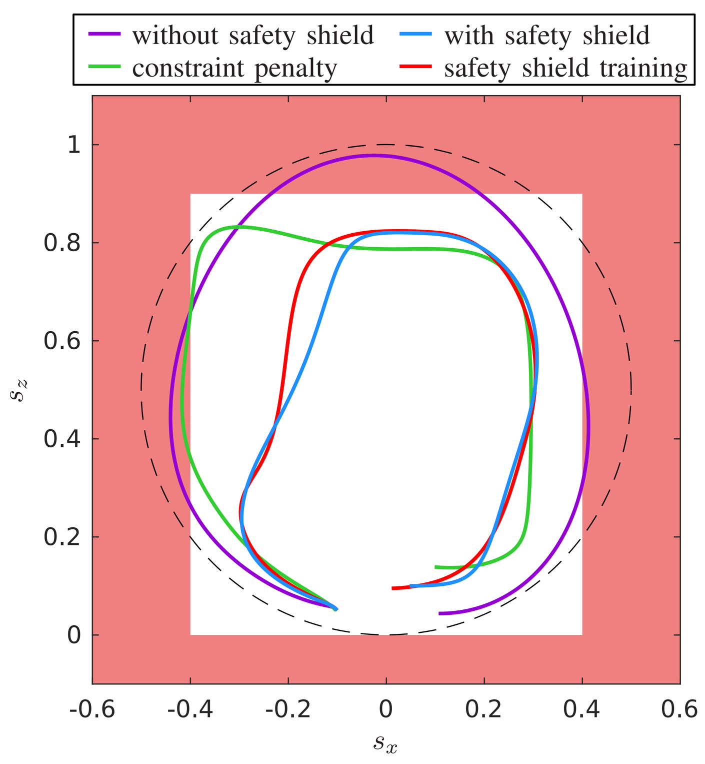
Action replacement — Action projection — Action masking

with the next state s_{t+1}, reward r_t, action a_t, safe action a_t^{φ}, action set $\mathcal{A}$, safe action set $\mathcal{A}_s$, provably safe action set $\mathcal{A}_{\varphi}$, and correction action $\tilde{a}$.

Action Projection with Reachability Analysis



Safety shield concept



Results 2D quadrotor

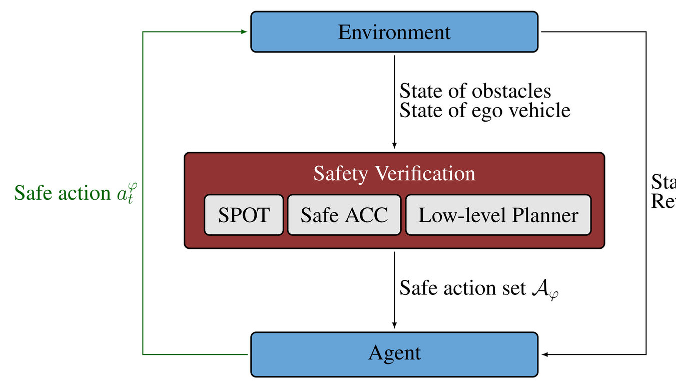
We can formulate the projection to the closest safe control input as an optimization problem [2]:

$$\min_{\alpha\in[-\mathbf{1},\mathbf{1}]} \|\alpha-\alpha_a\|_2^2 \quad \text{subject to} \quad \alpha \in \bigcap_{i=1}^{c}\bigcup_{l=1}^{h_i}\langle d_{il}, b_{il}, E_{il}\rangle_{LS}$$

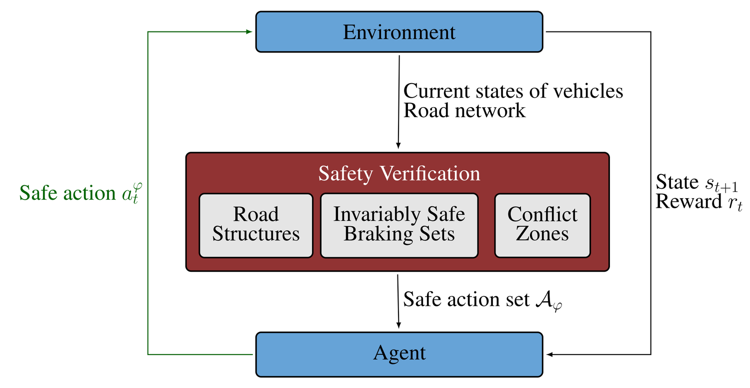
where

- $\alpha_a \in \mathbb{R}^p$ is a parametrization of the input $u_a = c_u + G_u\alpha_a$ and proposed by the reinforcement learning agent,
- $\langle d_{il}, b_{il}, E_{il}\rangle_{LS}$ is the polynomial level set describing the collisions i for the halfspace constraint l of the unsafe set.

Action Masking for Autonomous Driving



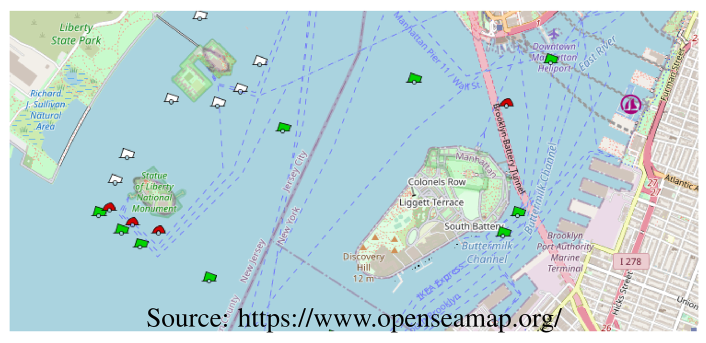
Highway driving



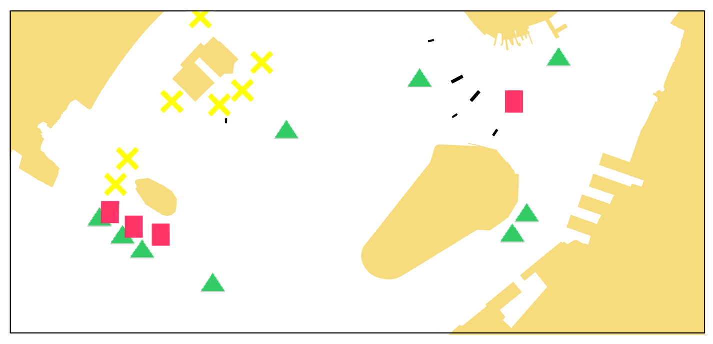
Intersection driving

- Commonroad-RL is an open-source reinforcement learning environment for autonomous driving [3].
- For highway driving, goal reaching decreases by 10 % to 87.5% on highD dataset compared to unsafe baseline [4].
- For urban driving, adding more traffic rules is necessary to increase goal reaching (about 30 % on inD dataset) [5].

Safe Motion Planning for Autonomous Vessels



OpenSeaMap map of the scenario

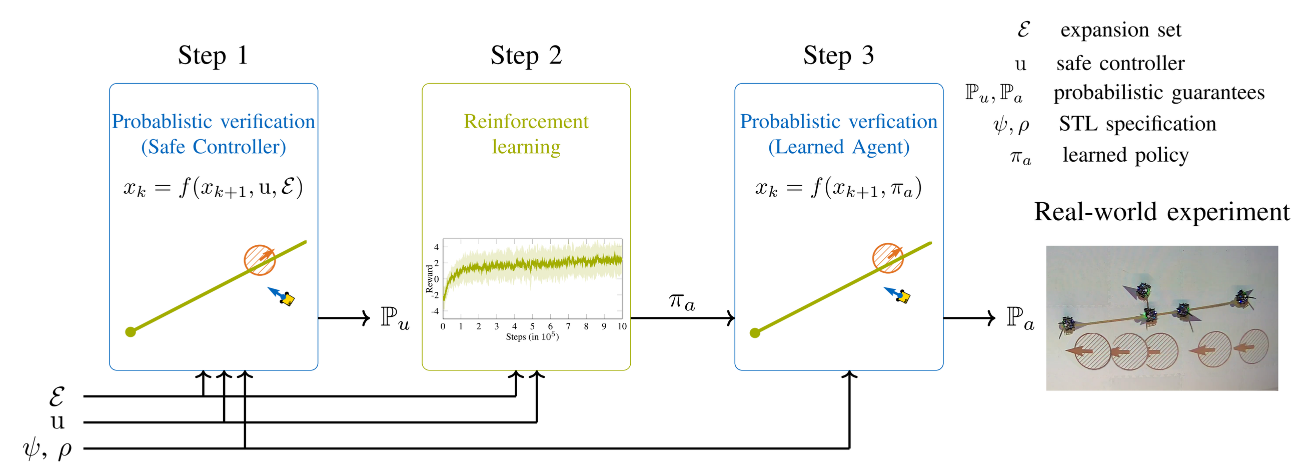


Corresponding CommonOcean scenario

- CommonOcean is a benchmarking suite for motion planning on the water [6].
- Formalization of traffic rules in temporal logic [7], e.g.,

$$\mathrm{G}\Big(\mathrm{keep}(x_{ego}, x_o, *) \implies \big(\mathrm{no\_turning}(x_{ego}, *)\,\mathrm{U}\,\neg\mathrm{keep}(x_{ego}, x_o, *)\big)\Big).$$

Probabilistic Guarantees via Temporal Logic



Modularity and transferability is achieved by separating safety specifications and performance objectives [8].

References

[1] H. Krasowski et al. *Provably Safe Reinforcement Learning: A Theoretical and Experimental Comparison*. 2023. arXiv: 2205.06750.

[2] N. Kochdumper et al. "Provably Safe Reinforcement Learning via Action Projection using Reachability Analysis and Polynomial Zonotopes". In: *Open Journal of Control Systems* (2023).

[3] X. Wang et al. "CommonRoad-RL: A Configurable Reinforcement Learning Environment for Motion Planning of Autonomous Vehicles". In: *ITSC*. 2021.

[4] H. Krasowski et al. "Safe Reinforcement Learning for Autonomous Lane Changing Using Set-Based Prediction". In: *ITSC*. 2020.

[5] H. Krasowski et al. "Safe Reinforcement Learning for Urban Driving using Invariably Safe Braking Sets". In: *ITSC*. 2022.

[6] H. Krasowski et al. "CommonOcean: Composable Benchmarks for Motion Planning on Oceans". In: *ITSC*. 2022.

[7] H. Krasowski et al. "Temporal Logic Formalization of Marine Traffic Rules". In: *IV*. 2021.

[8] H. Krasowski et al. *Probabilistic Guarantees for Safe Reinforcement Learning in Continuous Action Spaces via Temporal Logic*. 2023. arXiv: 2212.06129.