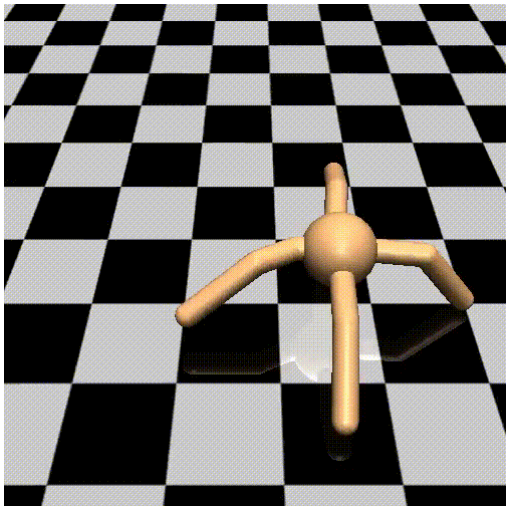
HiSaRL: A Hierarchical Framework for Safe Reinforcement Learning

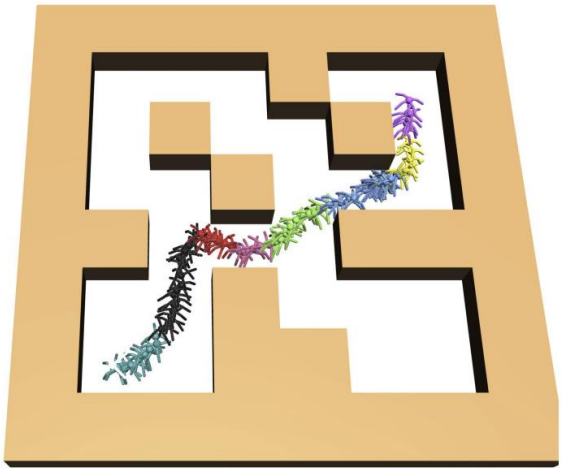
AAAI SAFEAI WORKSHOP 2022

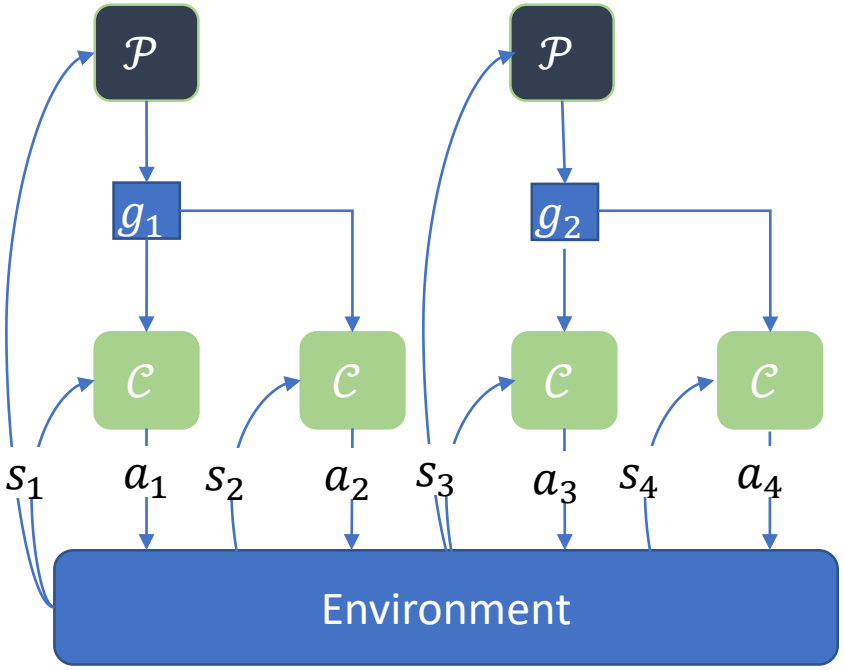
Zikang Xiong, Ishika Agarwal, Suresh Jagannathan

Computer Science Department, Purdue University

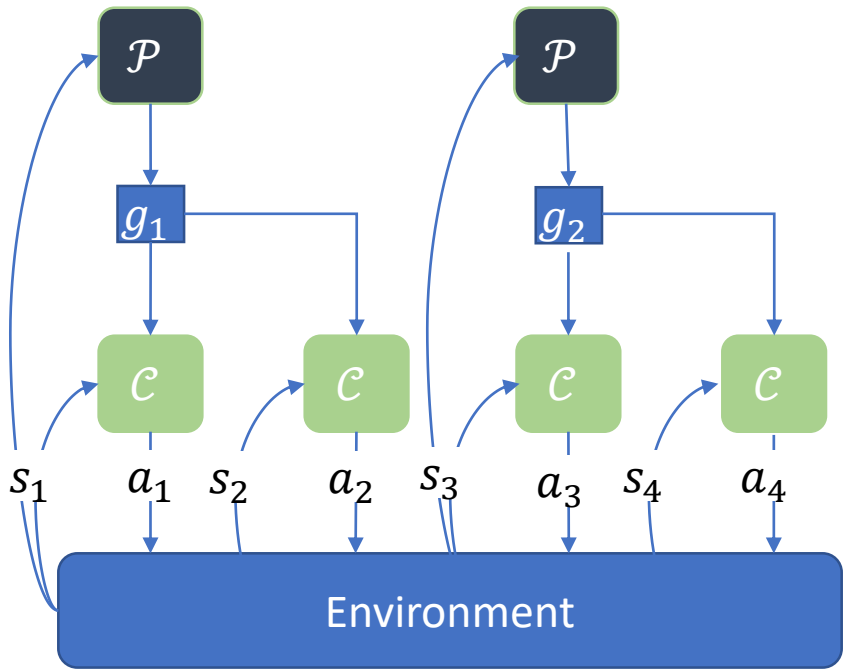
Hierarchical Framework

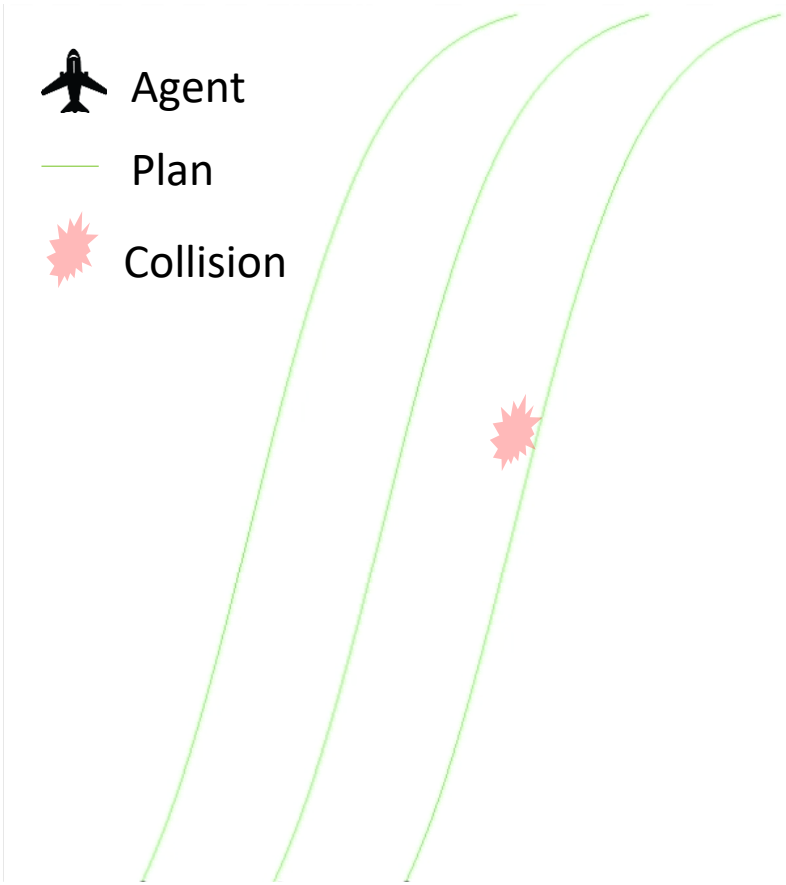




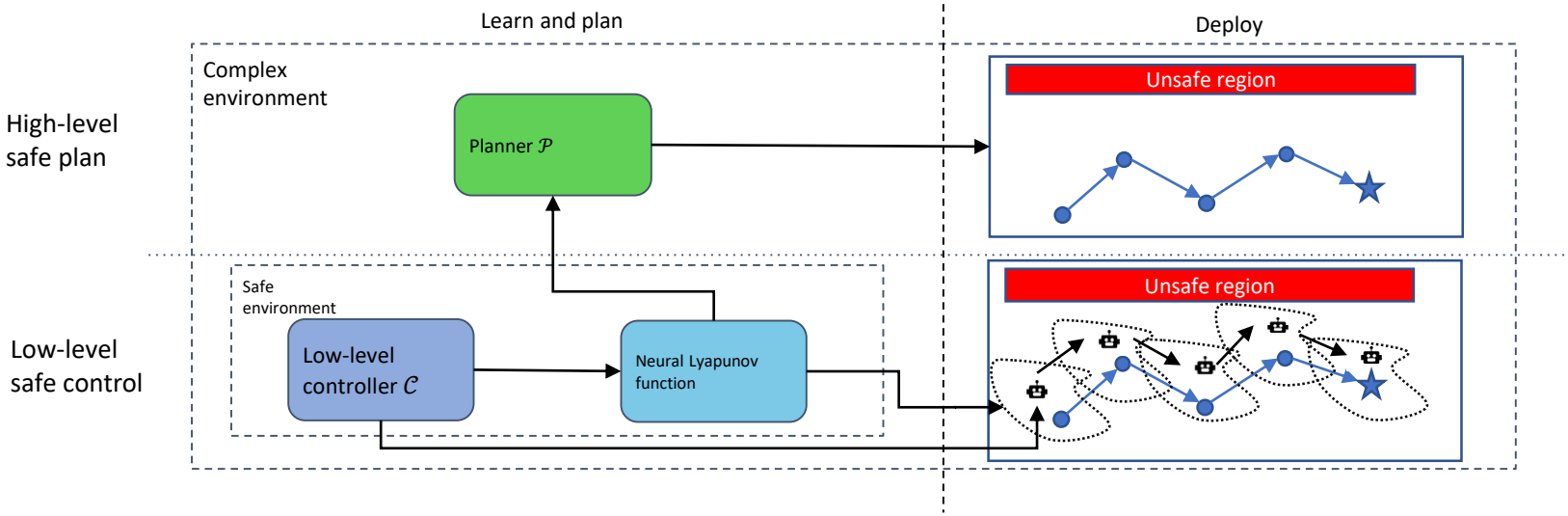


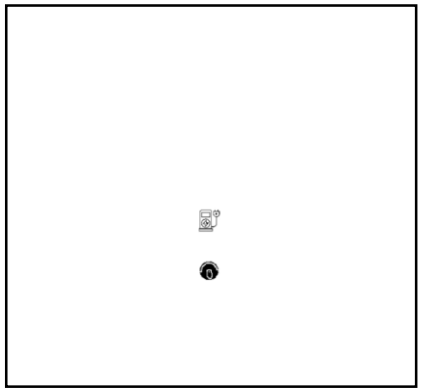
Safe Plan ≠ Safe Framework

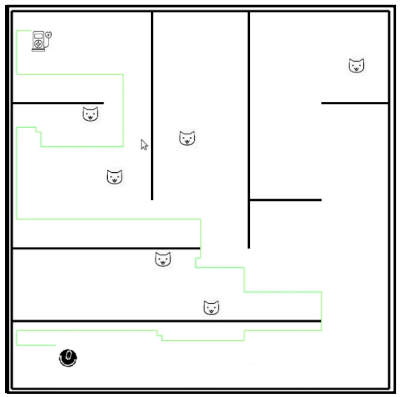




Architecture







Model-free Region of Attraction

Model-free RL

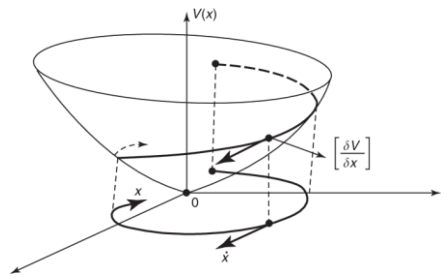
Suppose goal is g, robot position is x_t.

$$R(x_t) = e^{-||g-x_t||}$$

Minimizing $\sum_{t=0}^{T} R(x_t)$ with RL.

+

Neural Lyapunov Function (NLF)

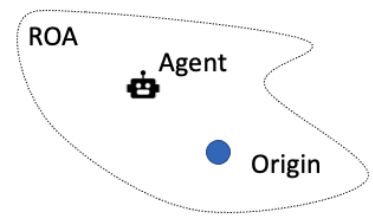


$$\begin{aligned} V(x_o) &= 0 \\ \forall x \neq x_o, V(x) &> 0 \\ V(x_{t+1}) - V(x_t) &< 0 \end{aligned}$$

$$\begin{aligned} L_{lf}(\theta_V) = \mathbb{E}_{s_t \sim (E,\pi)}(&V_{\theta_V}^2(s_o) \\ &+ \max(0, -V_{\theta_V}(s_t)) \\ &+ \max(0, \nabla_\pi V_{\theta_V}(s_t))) \end{aligned}$$

=

Region of Attraction(ROA)

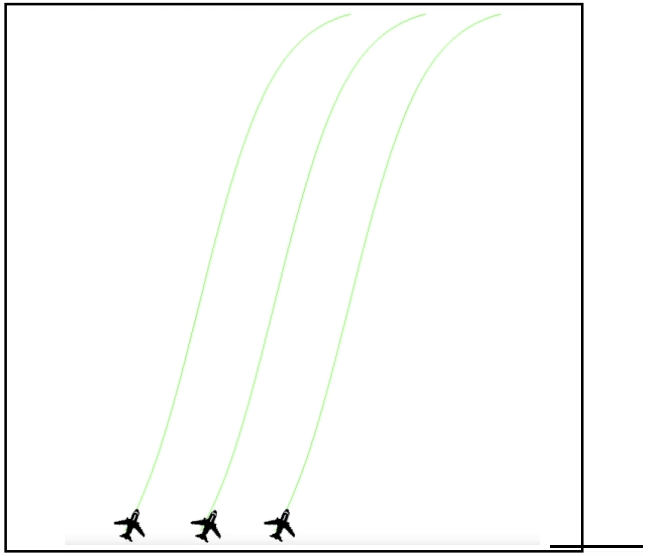


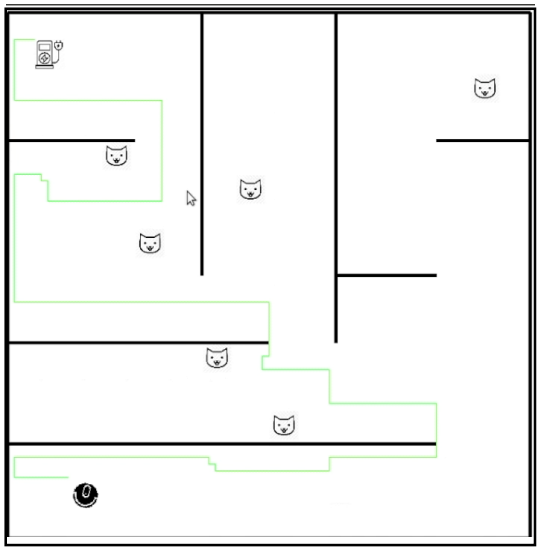
Set goal as origin

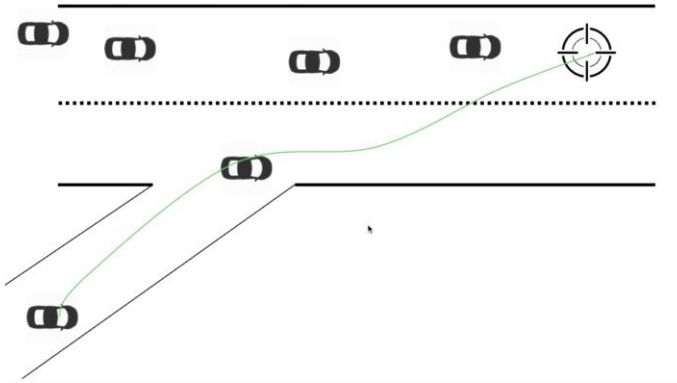
$$\text{ROA} = \{g + x \mid V(x) < C_{ROA}\}$$

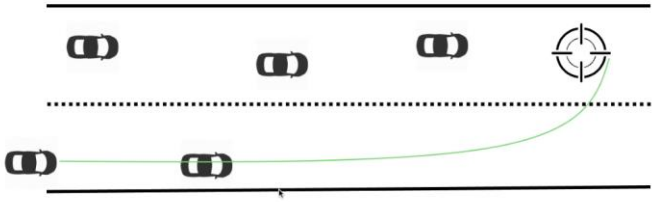
Model-free!

Benchmarks

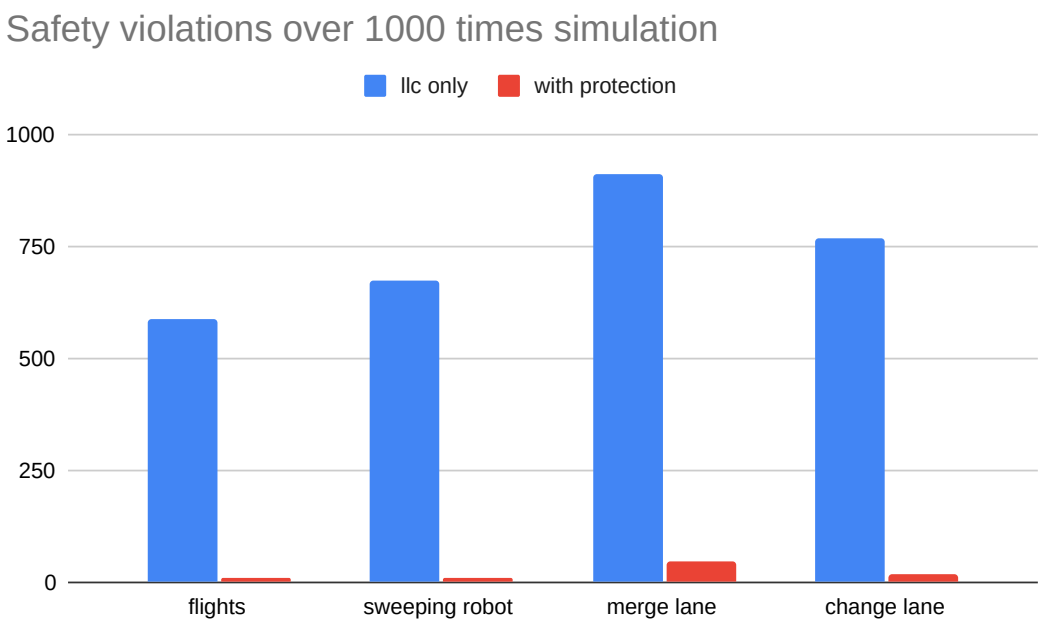






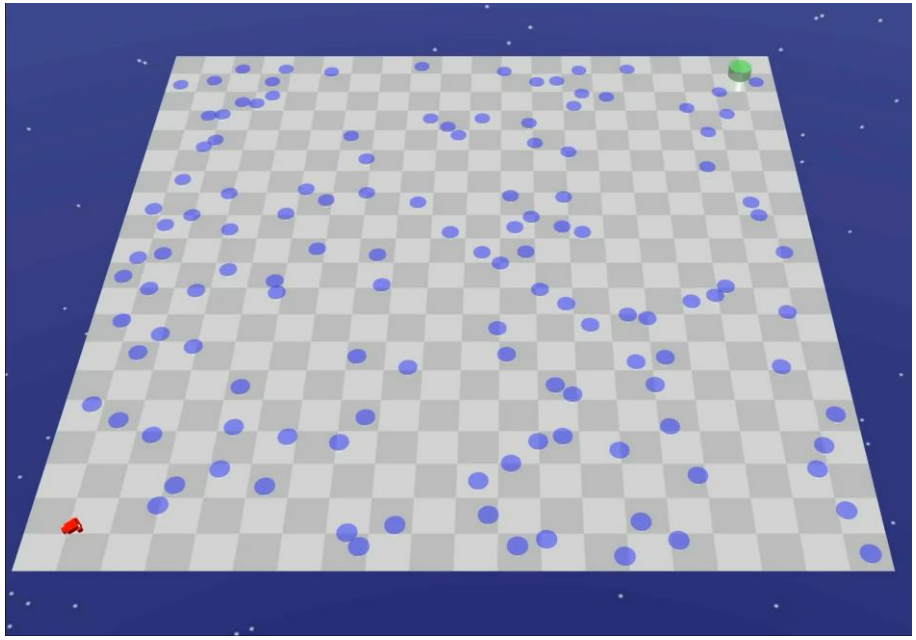


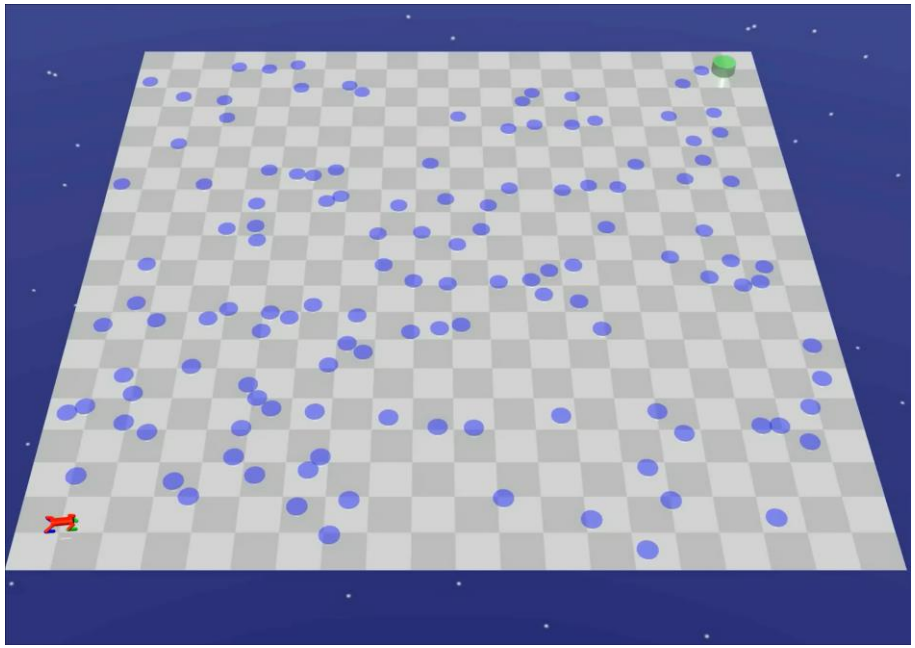
Safety Violation



- Significant safety improvement when comparing with the simple hierarchical structure

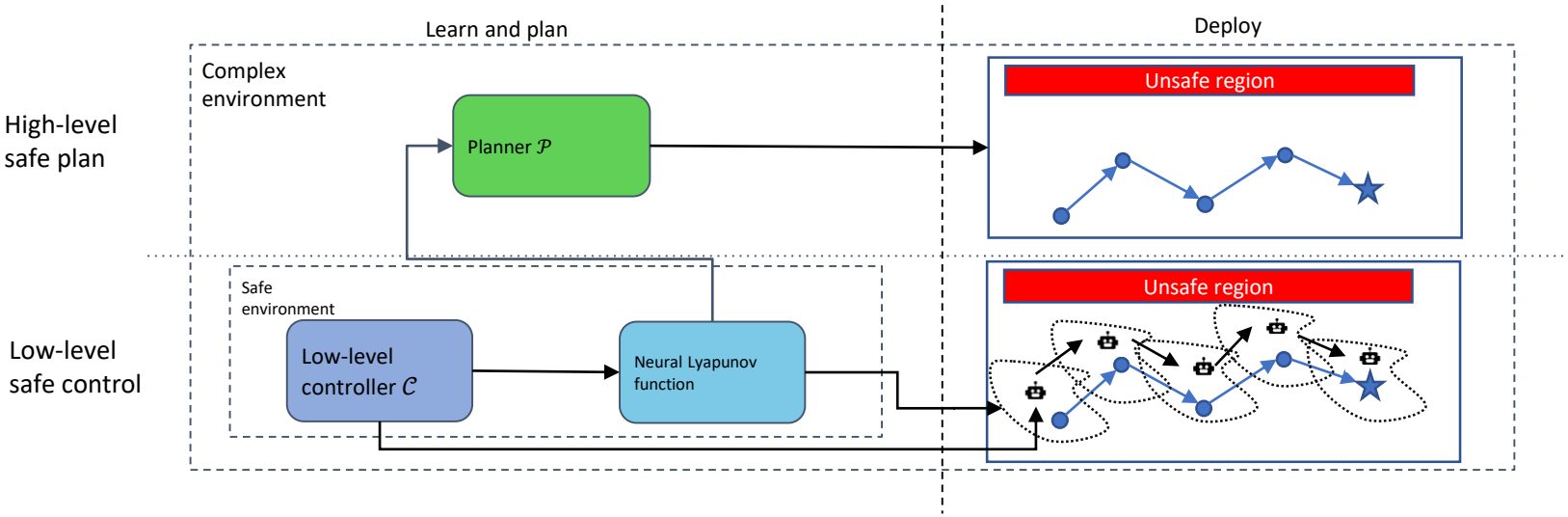
More Recently





- We have extended our approach to challenging benchmarks
  - Complex robot which are hard to model and control
    - large DOF
    - High observation space dimension
    - Implicit observation like raw lidar data
  - Sensitive safety constraints that are easy to violate

Summary



- Model-free hierarchical framework for safe reinforcement learning
- Safely combine planning and control
- Learning-based Lyapunov function and Neural RoA
- Sequentially shielding
- Deployed the trained agent in complex environment with significant safety improvement.
- Some other experiments are provided in paper (Robustness to adversarial attack, runtime planning path repairer)

Thank you!