USC Viterbi
School of Engineering

Reinforcement Learning for the Toyota HSR

Learning from High-level Task Objectives

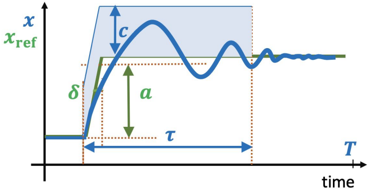
Andrea Dunn Beltran

August 2023

Summer Research Focuses

Exploring Key Areas of Emphasis

Temporal Logic



Overshoot:
$\text{alw}_{[0,T]}(\text{step}(x_{\text{ref}}, t) \Rightarrow \text{alw}_{[0,\tau]}(x(t) - x_{\text{ref}}(t) < c))$

GAZEBO

Simulations

Reinforcement Learning

Robot Operation

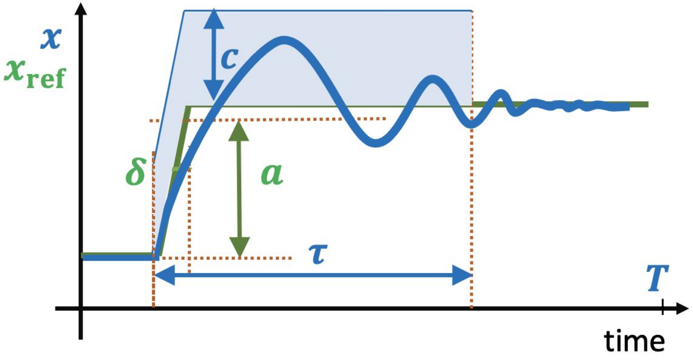
Signal Temporal Logic (STL)

Guiding High-Level Task Objectives

- Formal language for expressing temporal logic-based specifications.

- Uses operators like 'G' (globally) and 'F' (eventually) to define complex temporal patterns.

G[0, T] (A ∧ B)



Overshoot:

$\text{alw}_{[0,T]}\big(\text{step}(x_{\text{ref}}, t) \Rightarrow \text{alw}_{[0,\tau]}(x(t) - x_{\text{ref}}(t) < c)\ \big)$

Example Scenario

Putting Theory into Practice

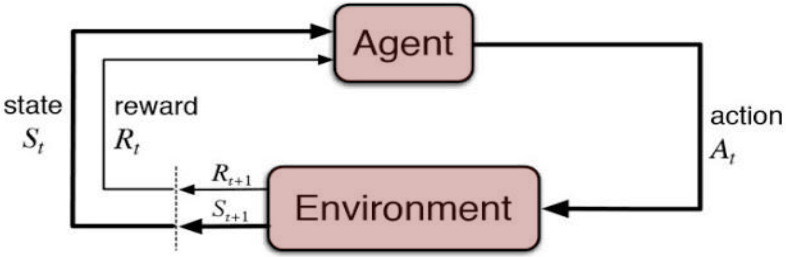
- **Goal:** The robot should navigate through a room to the charging station safely
  - ReachChargingStation ⇛ F (RobotReachesCharger)
  - AvoidObstacles ⇛ G (~RobotCollidesWithObstacle)

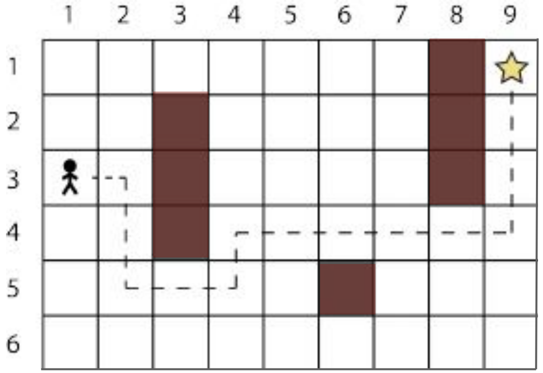
Reinforcement Learning (RL)

Training for Autonomy and Decision-Making

- Agent takes actions to maximize cumulative rewards over time.

- Enables the robot to learn and improve its performance in achieving high-level objectives.





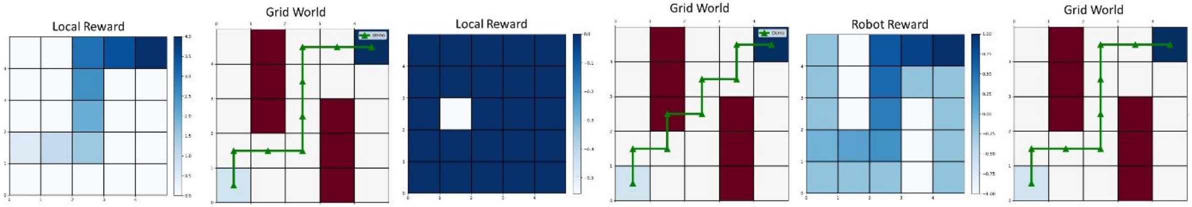
Learning from Demonstrations

Leveraging Real-world Examples for Enhanced Autonomy

- Extracting knowledge from human-provided demonstrations.

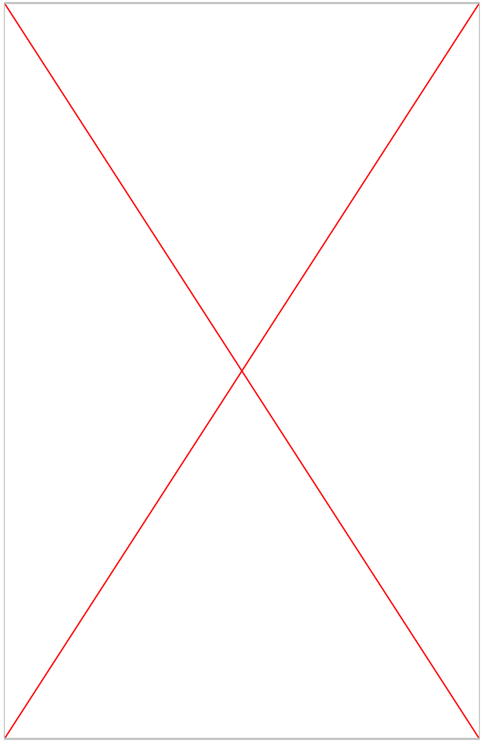
- Guiding the robot's learning process with expert behaviors.

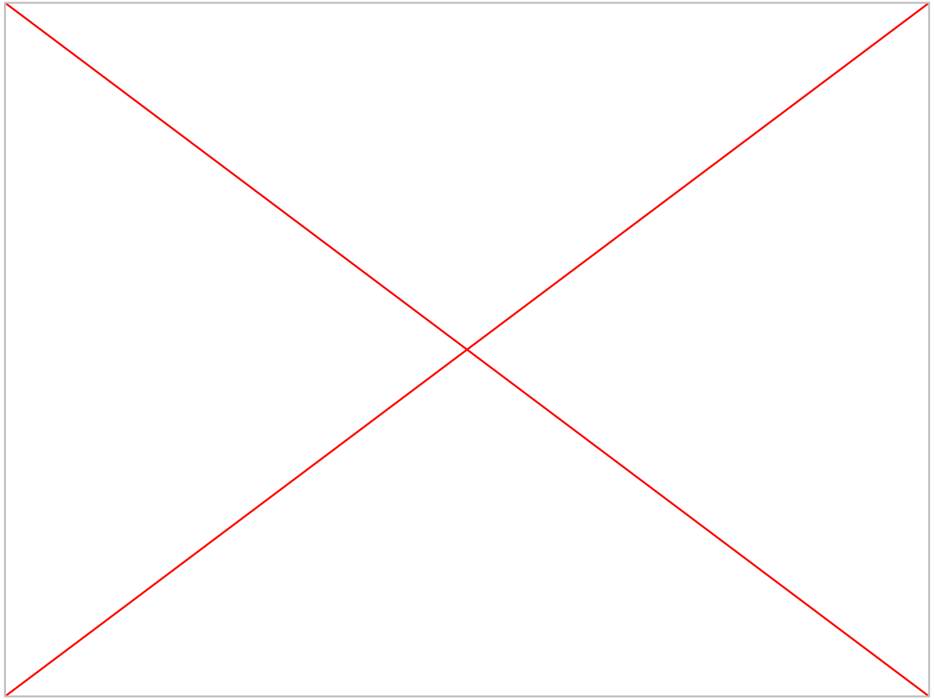
- Complementing reinforcement learning with imitation learning.



(a) Demo 1 (b) Demo 2 (c) Robot Policy

Figure 4: Results: Left figures represent learned rewards. Right figures show the grid-world with start state (light blue), goal (dark blue), obstacles (red) and demonstration/policy (green).





Learning New Tricks

Leveraging Examples for Training and Autonomy

Challenges Faced

The Reality of Working with Robots

Documentation

Setup

Debugging

Future Work

Guiding High-Level Task Objectives

- Speed up RL training for HSR Gazebo simulations
- Integrate demonstration learning to RL model
- Allow connected edward functions and add STL to simulations

Special thanks to