Custom Reinforcement Learning Assignment
Learning-Based Control and Navigation for a Mobile Robot

Overview

The goal of this assignment is to study the application of **reinforcement learning (RL)** to mobile robotics by progressively replacing classical components of a navigation stack with learned policies. The robot is modeled as a **low-degree-of-freedom mobile robot** (unicycle), and the focus is on **state-based, low-dimensional RL**, avoiding image-based observations.

The assignment is divided into two parts of increasing difficulty, enabling a clear comparison between classical control methods and learning-based approaches.
The general context is a ful planning-controller pipeline

- Task Planner: define a set of goal point and the order to reach them
- Road Planner: define the general path between the start position and the next goal divided by a set of step
- Dubins planner: define the dubins curve of one step
- Lyapunov Controller: given the dubins curve control the robot to follow it

Part 1: Learning a Feedback Controller for Trajectory Tracking

In the first part, the dubins planner provides a reference trajectory as a sequence of waypoints. Traditionally, the robot follows this reference using a hand-crafted analytical controller: a Lyapunov-based tracking controller.

The task consists of training an RL policy that **replaces the analytical control law** and directly outputs continuous velocity commands (v, ω). The RL policy observes a low-dimensional state including tracking errors and reference information, and is trained to minimize tracking error while producing smooth and stable control commands.

Part 2: Learning a Control Policy with Planning Awareness

The second part addresses a more challenging problem; the RL policy is extended to handle **local navigation decisions**, such as obstacle avoidance and path optimization. This policy should replace the dubinsa planner and the controller in a unique plocy.

The policy receives as input the robot state, the target pose, and compact local obstacle information (e.g., distances and angles to nearby obstacles). It directly outputs velocity commands (v, ω) and must balance multiple objectives, including: reaching the target efficiently, avoiding collisions, ...

While a global planner may still provide a goal or coarse path, the RL policy implicitly integrates aspects of local planning and control.