Robot Path Planning via Deep Reinforcement Learning

Cem Alpturk[1], Venkatraman Renganathan[2], and Anders Rantzer[3]

[1]Student: `ce5368al-s@student.lu.se`
[2]Supervisor: `venkat@control.lth.se`
[2]Examiner: `anders.rantzer@control.lth.se`

January 2022 - June 2022

1 Introduction

The purpose of this project is to compute a path in real time for a robot operating in a known environment with static and dynamic obstacles, given the initial and target positions. The calculated path should minimize the risk of collision with any obstacles. The robot is modeled as a discrete-time LTI system with additive process noise belonging to a moment based ambiguity set and is controlled with a finite set of inputs. A wide variety of path planning algorithms exist that work on systems that are deterministic, however for this project, the additive process noise requires a different solution. Optimal solutions such as RRT^* [**RRT**] exist but can be too slow to execute in real time.

2 Problem Statement

Given an initial state $x_0 \in \mathcal{X} \in \mathbb{R}^n$ and a set of final goal locations $\mathcal{X}_{goal} \subset \mathcal{X}$, find a measurable control policy $\pi = [\pi_0, ..., \pi_{T-1}]$ with $u_t = \pi_t(x_t)$ that maximizes the finite-horizon expected and discounted reward function subject to constraints:

$$
\begin{aligned}
\underset{\pi}{\text{maximize}} \quad & \mathbb{E}\left[\sum_{n=t}^{N-1} \gamma^{n-t} r(x_n, u_n, x_{n+1})\right] \\
\text{subject to} \quad & x_{t+1} = Ax_t + Bu_t + w_t, \\
& x_o \sim \mathbb{P}_{x_0} \in \mathcal{P}^x := \{\mathbb{P}_{x_0} | \mathbb{E}[x_0] = \bar{x}_0, \mathbb{E}[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = \Sigma_{x_0}\}, \\
& w_t \sim \mathbb{P}_w \in \mathcal{P}^w := \{\mathbb{P}_w | \mathbb{E}[w_t] = 0, \mathbb{E}[w_t w_t^T] = \Sigma_w\}, \\
& u_t \in \mathcal{U}, \\
& \mathcal{X}_t^{free} = \mathcal{X} \setminus \bigcup_{i \in \mathcal{B}} \{x_t | A_i x_t \le b_{it}\}, \\
& \sup_{\mathbb{P}_{x_t} \in \mathcal{P}^x} \mathbb{P}_{x_t}\left(x_t \notin \mathcal{X}_t^{free}\right) \le \alpha_t, \quad \forall t \ge 0
\end{aligned}
$$

(1)

where $\mathcal{P}^x$ is an ambiguity set of marginal state distributions and $\alpha_t \in (0, 0.5]$ is a stage risk budget parameter such that $\sum_{t=0}^{T} \alpha_t \le \alpha$ with α being the user-prescribed total risk budget. The task for the robot is to learn a control policy such that risk of obstacle collision is minimum.

3 Approach

A formulation as in [**pathplanning**] will be adopted in order to control the robot and a solution based on Distributionally Robust Reinforcement Learning framework as in [**distributionally**] will be developed in order to solve the path planning problem in real time. Numerical experiments will be performed for the proposed method using Deep Reinforcement learning. This is supposed to return a sub-optimal solution at the expense of less computation time compared to optimal motion planning algorithms as in [**RRT**]. The main goal will be to complete this task with static obstacles and if time permits, dynamic obstacles will also be taken in to consideration.

4 Contributions

The project could help automated processes in known enclosed spaces to perform efficient path planning and collision risk analysis in the presence of process noise that is not necessarily Gaussian.

5 Resources

The project will be implemented with Python and-or MATLAB. In the case that the training process is too demanding a remote machine, Google Colab or AWS could be used to speed up the project.