Continuous & Mapless Navigation for robots using Policy Gradient Algorithm

Avnish Gupta, Anamika Kumari, Ashwij Kumbla, Shreedhar Kodate

Proposal Summary

Previously, we have proposed to train a Deep Deterministic Policy Gradiet(DDPG) based network for the purpose of continuous & mapless motion planner for the navigation for robots. The proposed mapless motion planner aims to navigate the nonholonomic robot to desired targets without colliding with any obstacles in the environment without any predefined features and demonstrations. The training procedure of this project will be on virtual environments created using the Gazebo simulation platform. Multiple environments with both static and dynamic obstacles will be created and the deployed robot will learn the models of these environments from scratch. The target position for the robot will be randomly initialized after every episode.

Progress

We started by setting up our collaboration space for working on the project. We created a GitHub Repository which will store all Gazebo & ROS files, as well as our DQN files and utilities.

For this project we chose turtlebot3 from ROS as our agent since it provides package for control and real-time state feedback and supports wide range of sensors. We have also set up LIDAR within the gazebo environment.

For the purpose of training the agent we created three different environment with increasing complexity of the map as shown in the following figures. The orange box is the goal box the agent tries to navigate. The blue region is the area which the agent is able to sense using the LIDAR from its current location.

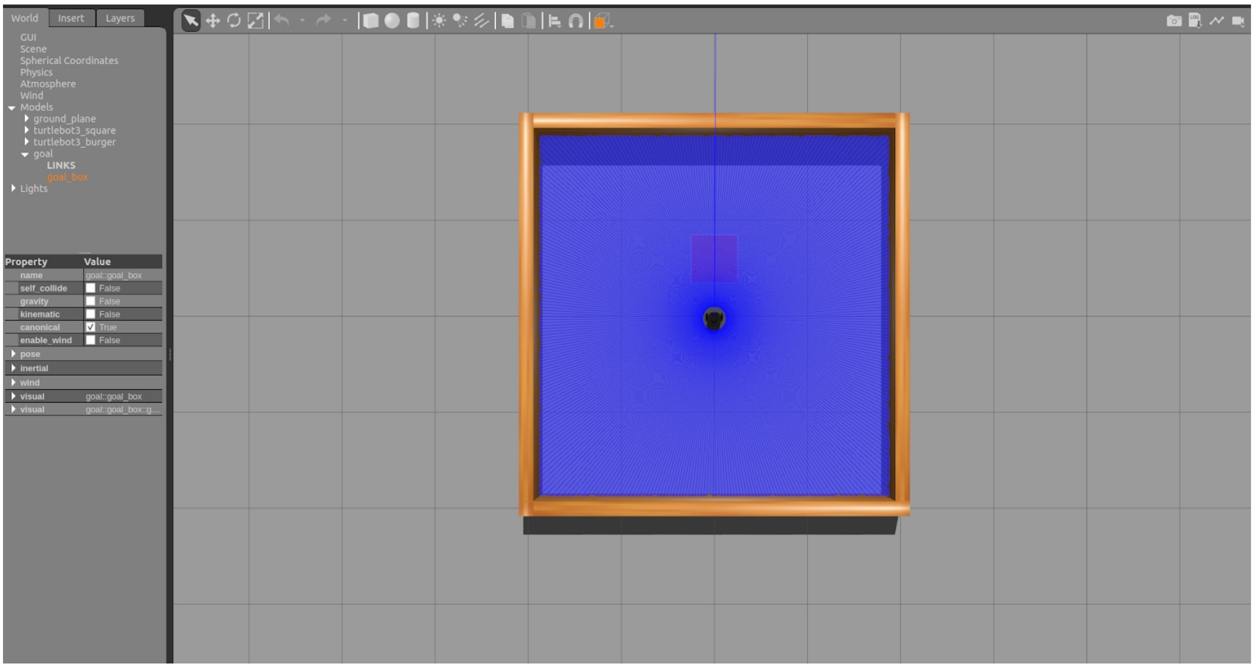

Figure 1: Stage_1 with no obstacles

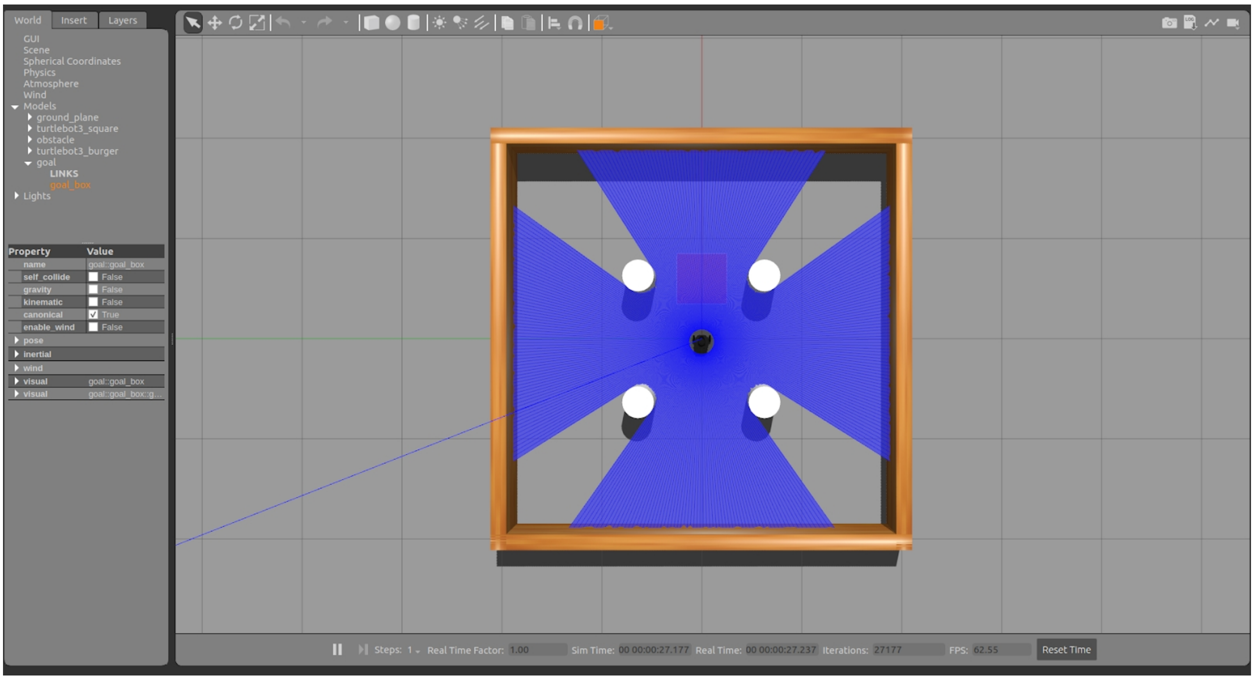

Figure 2: Stage_2 with simple obstacles

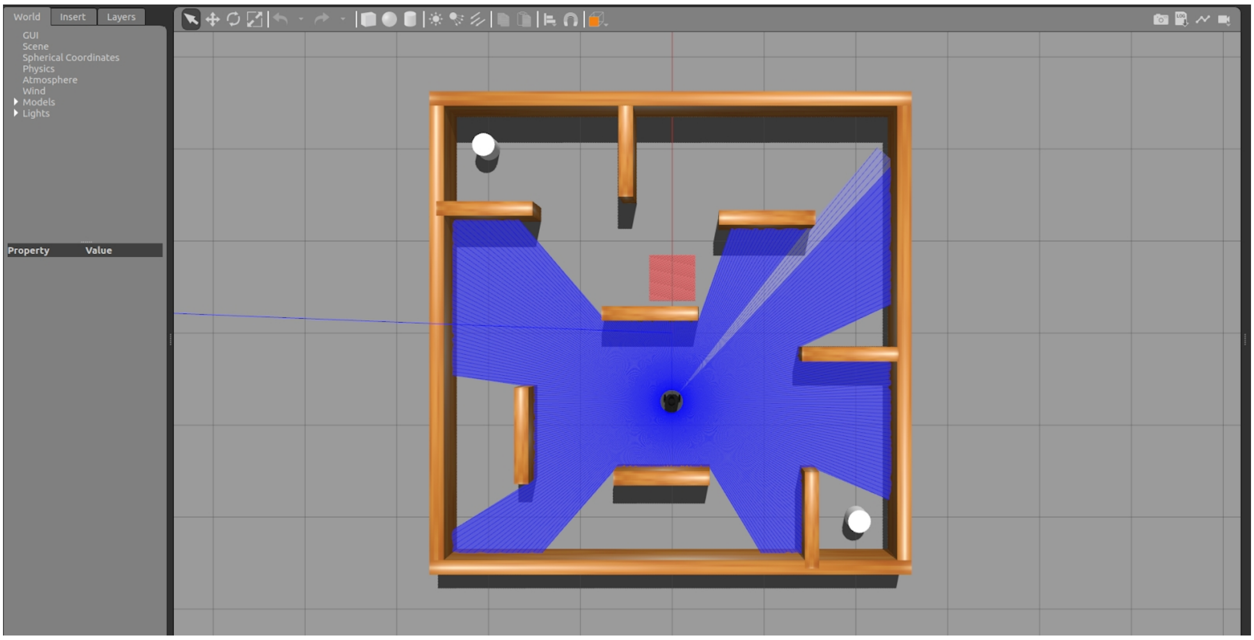

Figure 3: Stage_3 with complex obstacles

We have also determined Deep Networks architecture for the both Actor and Critic network. (FILL THIS)

Problem Formulation

Setup:
The model will be trained on simulated Turtlebot3 on gazebo with different environments.\

Action space:
The 2-dimensional action of everytime step includes the angular and the linear velocities of the differential mobile robot. To constrain the range of angular velocity in (1, 1), a hyperbolic tangent function (tanh) is used as the activation function. Moreover,the range of the linear velocity is constrained in (0, 1) through a sigmoid function. Backward moving is not expected because laser findings cannot cover the back area of the mobile robot. Considering the real dynamics
of the robot we clip the angular velocity at 1 radian per second and linear velocity at 0.5 meter per second.

Reward:

1. Relative distance (d_t) of the robot and the target location is calculated after every step of the robot movement simulation. If d_t is less than the threshold (c_d) then the agent will receive a positive reward of r_arrive and the episode will terminate.
2. If the distance from obstacles (can be inferred from the state) is less than a threshold (c_o) from the obstacles, then a negative reward of r_collision would be given to the RL agent and the episode will terminate.
3. If the state is a non-terminal state, then the reward is calculated as a product of floating point hyperparameter (c_r) and the difference between the relative distances (d_t-1 - d_t).

Architecture:
Our architecture will take a similar approach as work in [1] that is shown below.

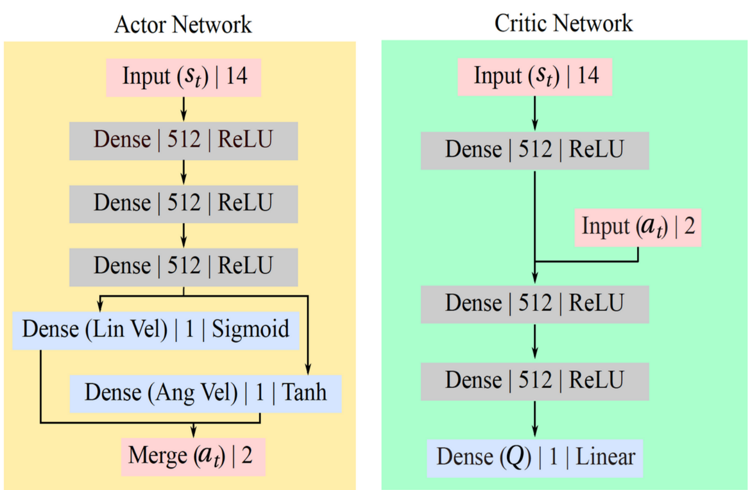

Figure 4: Actor-Critic DDPG network

Inputs:

1. 10-dimensional laser range findings sampled from the raw laser findings between 0 and 360 degrees in a trivial and fixed angle distribution
2. The 2-dimensional target position represented in polar coordinates with respect to the mobile robot coordinate frame.
3. The 2-dimensional previous action

Expected Project Timeline

Week 1:

- Create the simulation environment .
- Implement the network architecture and general code structure.
- Complete!

Week 2:

- Training and tuning the network in the simulation
- Creating the simulation environment for testing
- In Progress!

Week 3:

- Testing in the simulation
- Yet to start

Source

1. Tai, Lei, Giuseppe Paolo, and Ming Liu. "Virtual-to-real deep reinforcement learning: Continuous control of mobile robots for mapless navigation." *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017.

1. Zhang, Daniel, and Colleen P. Bailey. "Obstacle avoidance and navigation utilizing reinforcement learning with reward shaping." *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II*. Vol. 11413. International Society for Optics and Photonics, 2020.

2. http://gazebosim.org/

1. https://www.ros.org/