qRRT-Quality biased incremental RRT

Dhruv Krishna, Siddhant Saoji
Supervisor: Suril V Shah

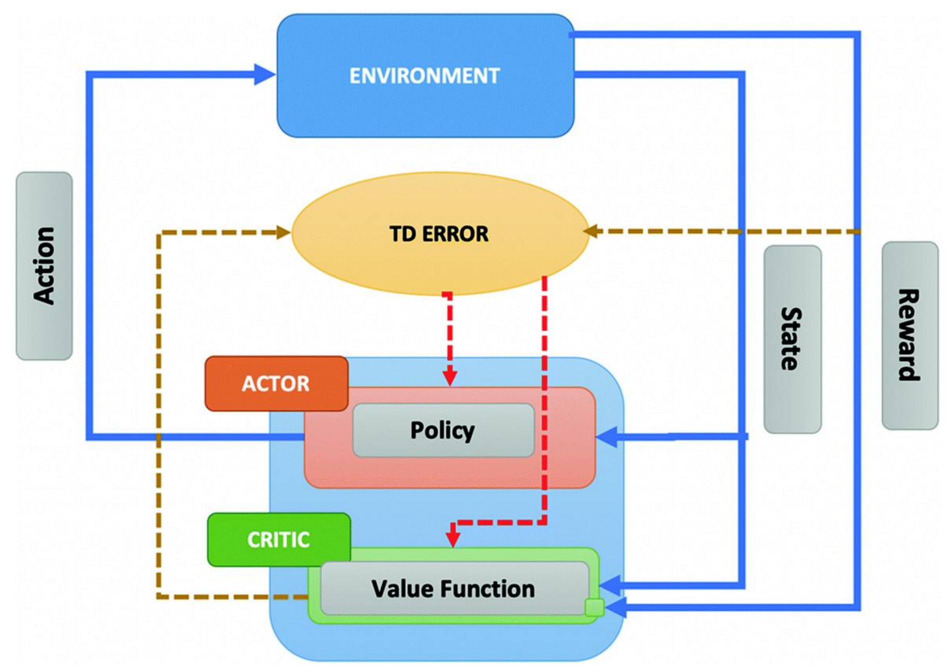
Abstract

In our project, we present a sampling based method for optimal motion planning in the absence of known cost functions. It uses the principle of ‘learning through experience’ to deduce cost to go of regions within the workspace which is used to bias an incremental graph based search algorithm. Iterative improvement of cost information and search bias produces solutions that are proven to be asymptotically optimal.

Quality biased RRT (qRRT) which combines the rapid exploration of RRT and experience based method of Reinforcement Learning to learn region costs uses the sample trajectories generated using RRT to learn region costs using Deep Reinforcement Learning and this information is used to bias search towards low-cost regions of the configuration space.

Introduction

Optimality is a desired feature in all robot motion planning algorithms. Robot complexities and workspace complexities make optimal path planning difficult. Random search methods are good candidates for complex high dimensional systems. Tree-based systems like RRT were observed to have an advantage of rapidly generating feasible solution path however, finding an optimal solution with it is highly unlikely. In systems where costs of actions in the configuration space is either known or calculable, methods like A*, BIT* are guaranteed to converge, but the prior knowledge of the cost functions makes it unsuitable for some systems.



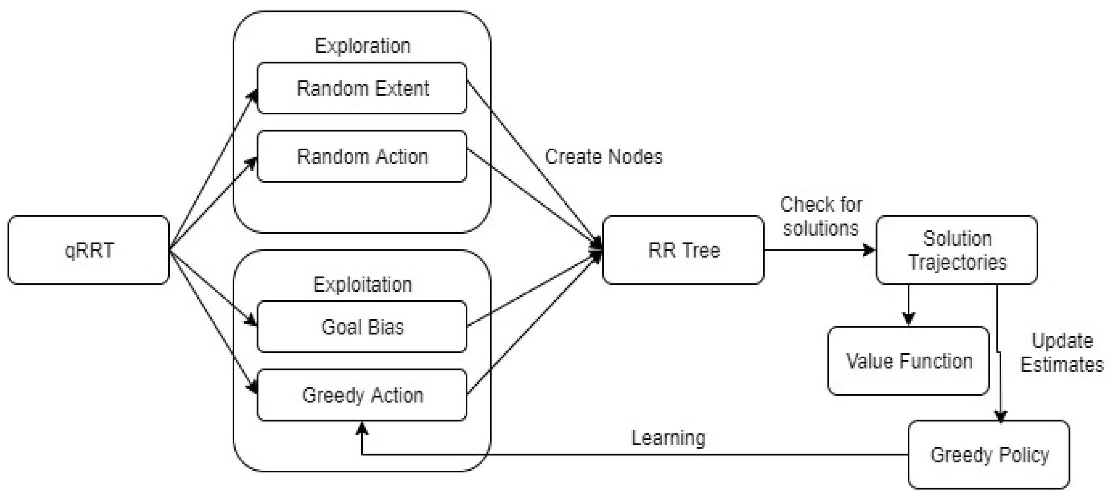
Algorithm

The qRRT algorithm executes a continuously growing RRT tree until the termination condition has been met. The algorithm can broadly be divided into 3 categories: Exploration, Learning and Exploitation.

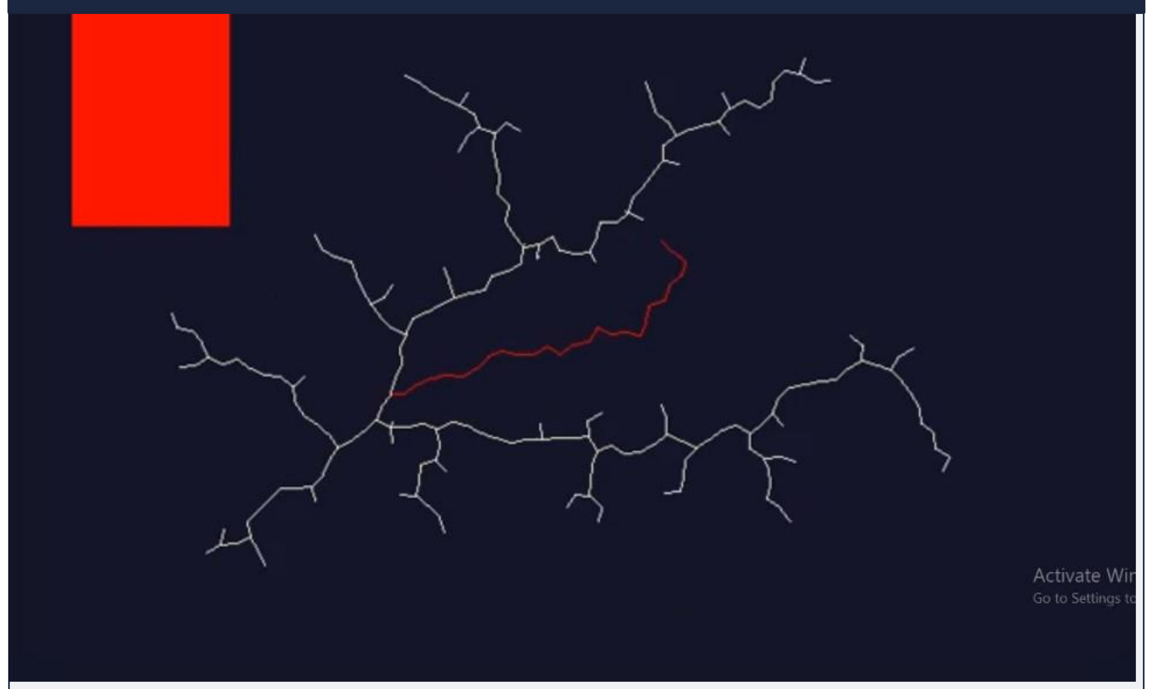
Tree-Based Search:
The tree growth is performed through explorative and exploitative actions that result in the creation of new nodes. The Adaptive Artificial Neural Networks (AANNs) for value function and policy are created and are trained after every episode. If a new node is sufficiently close to the goal then the complete path from the initial point rewarding to the goal is traced back and the AANNs are updated. The solution trajectory generated in each iteration is chosen from trajectories within the RRT-Tree. All the trajectories are evaluated and trajectory with the highest quality is returned.

Exploration:
A random state is selected and the closest node is found on the existing RR Tree. A new node is then created in the direction of the randomly selected point at a distance equal to the node length of the RRT-Tree. After that, the transition cost is calculated using the appropriate reward function which is then used to calculate the cumulative reward of the solution trajectory.

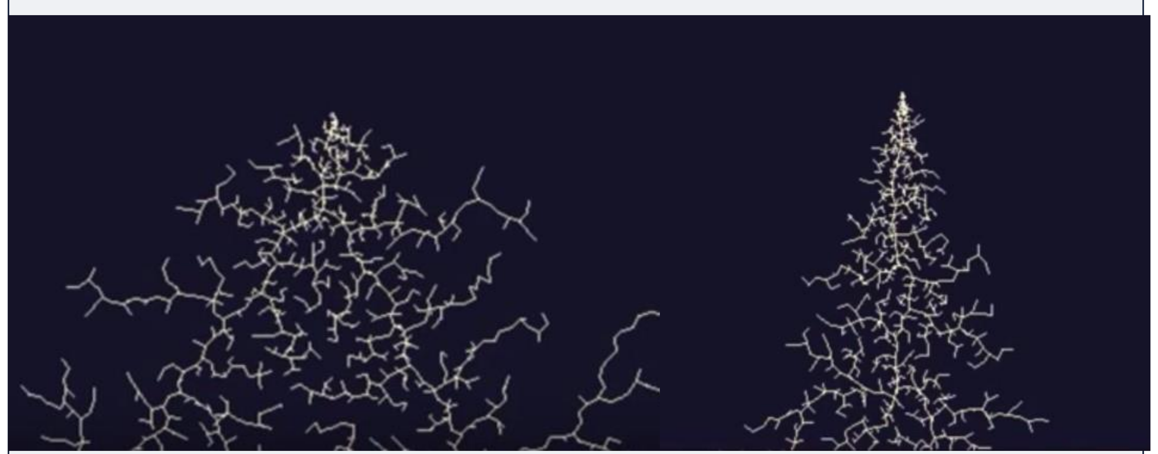
Exploitation:
Explorative actions are performed using greedy actions predicted by the AANNs. The AANN selects a random state on the tree and performs the action given by the greedy policy. This implements the quality biased extend operation.



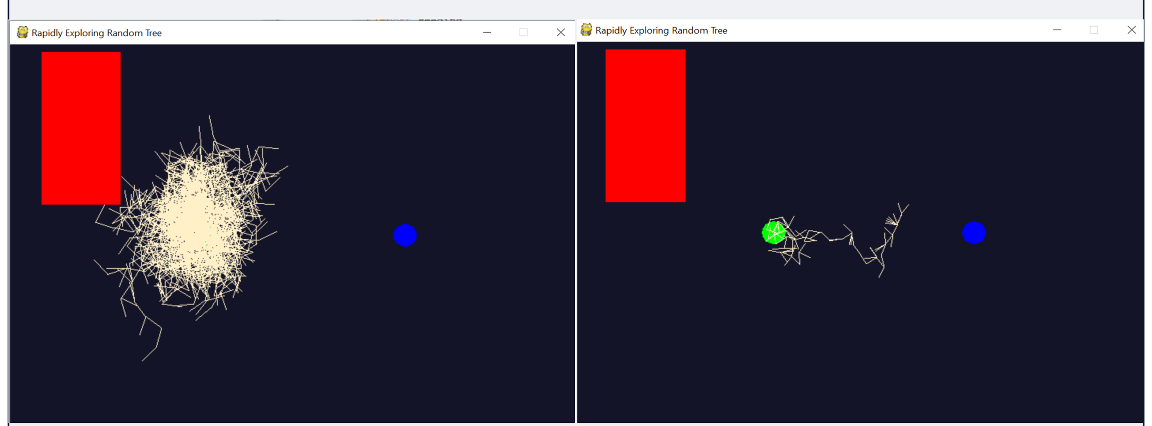
Result



Generating training data:
Multiple instances of RRT were run and the solution trajectories were generated after reaching the termination condition. The AANN was then trained using the generated data to predict the optimal path.



Tree Biasing:
A biased RR Tree was generated using the gaussian sampling method for tree expansion in order to select the random state for expansion. Different variances were sampled and the bias level was selected accordingly.



Before Training **After Training**

Biased RR Tree

Reference

1. LaValle, Steven M. "Rapidly-exploring random trees: A new tool for path planning." (1998).
2. Sutton, Richard S., and Andrew G. Barto. Reinforcement learning: An introduction. MIT press, 1998.
3. Abdul Hafez A. H., Mithun P. , Anurag V. V., Shah S. V., Krishna K. M., "Reactionless visual servoing of a multi-arm space robot combined with other manipulation tasks," Robotics and Autonomous Systems, 91, pp. 1-10, May 2017 (IF: 2.638)

Prometeo 2020

Industry Day

Indian Institute of Technology Jodhpur

भारतीय प्रौद्योगिकी संस्थान जोधपुर

॥ त्वं ज्ञानमयो विज्ञानमयोऽसि ॥