Safe Motion Planning with Learning in the Loop

Emanuel Herberthson
emanuel.herberthson@liu.se
Division of Automatic Control, Linköping University

Daniel Axehill
daniel.axehill@liu.se
Division of Automatic Control, Linköping University

Background

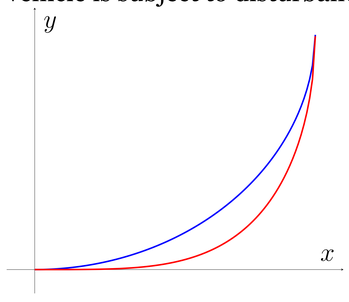
Optimal motion planning remains of significant interest to society, as it provides methods for not only finding feasible paths between start and goal points, essential for applications such as autonomous vehicles, but also for minimizing, for instance, fuel consumption. Extensive research has been done in the area, primarily using sampling-based approaches or deterministic optimal control approaches. With the rise of machine learning and artificial intelligence (AI), new opportunities have emerged to address more complex scenarios and enhance the capabilities of optimal motion planning methods.

Research Project Formulation

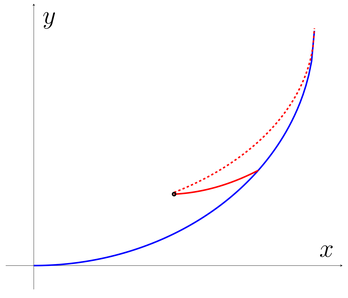
This project advances the group's previous work in optimal motion planning by integrating methods from AI and optimal control into more efficient and powerful approaches than those achieved individually. The primary tools to be integrated with the current state-of-the-art optimal control and motion planning algorithms are reinforcement learning and generative AI/foundation models. Although these AI tools often yield impressive results, they typically do not align well with the need to ensure safety in safety-critical applications. Achieving such impressive results while simultaneously ensuring some level of guaranteed performance and safety is the main research challenge of this project. Several candidate approaches exist to address this challenge, such as explicitly accounting for the uncertainty in the learned components and/or incorporating various forms of supervisory functionality.

Learning Motion Primitives

One optimal control based approach is discretizing the environment into a state lattice, computing motion primitives between pairs of states, and forming a complete path by connecting the states through sequences of motion primitives. The initial setup considered involves truck-trailer vehicles. When certain parameters, such as trailer length, are unknown, it becomes impossible to compute the motion primitives accurately. A possibly salvaging approach is the use of machine learning methods such as reinforcement learning. Starting from an initial policy, which is computed based upon an incorrect trailer length estimate, the policy can be iteratively refined and an optimal policy is learned. Such a policy would also have the advantage of completing a motion primitive in an optimal way if the vehicle is subject to disturbances.



Blue - precomputed primitive which is suboptimal.
Red - optimal, and desired, primitive.

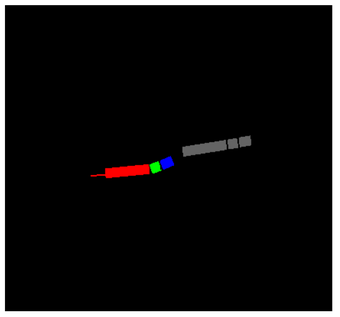


Red line - returning to motion primitive given some controller.
Red dashed - learned policy finishing the primitive optimally.

A particularly interesting aspect of the problem is that the reinforcement learning agent must learn to reverse a system consisting of a truck, a dolly, and a trailer. The agent is thus required to control an unstable system where jack-knifing, joint angles between vehicle-parts that prohibit reversing, must constantly be avoided.

Simulation Environment

A simulation environment has been developed using Python and the library Gymnasium; the latter is specifically designed for creating reinforcement learning environments. In the environment, an agent controlling a truck with a dolly and a trailer is tasked with reaching the predefined goal state while maximizing the rewards received. Initial experiments using a basic RL algorithm have been conducted, indicating that this problem is non-trivial.



Connection to SEDDIT focus areas

Optimal motion planning, and this research project as an immediate consequence, is directly connected to the SEDDIT focus area of achieving zero carbon emissions and resilient transportation systems. Contributions towards energy-optimal and secure motion planning will in turn lead to reduced emissions and more reliable transportation systems.

Acknowledgements

This work was performed within the Competence Center SEDDIT (Sensor Informatics and Decision making for the DIgital Transformation), supported by VINNOVA (Sweden's Innovation Agency) within the research and innovation program Advanced digitalization.

LINKÖPING UNIVERSITY

SEDDIT
Sensor informatics and Decision-making for the Digital Transformation

UPPSALA UNIVERSITET