Paper ID: 85857
oral
student

IAF ASTRODYNAMICS SYMPOSIUM (C1)
Guidance, Navigation & Control (3) (5)

Author: Mr. Yashdeep Chaudhary
The University of Auckland, New Zealand

Dr. Harry Holt
The University of Auckland, New Zealand
Mrs. Cristina Parigini
Deimos Space UK Ltd, United Kingdom
Dr. Roberto Armellin
The University of Auckland, New Zealand

INTEGRATING REINFORCEMENT LEARNING AND REACHABILITY ANALYSIS FOR CONSTRAINT HANDLING IN AUTONOMOUS MULTI-IMPULSE TRANSFERS

Abstract

Reinforcement Learning (RL) holds promise for autonomous spacecraft guidance within the machine learning domain. Under the RL paradigm, the control problem is formulated as a Markov Decision Process (MDP), where optimal control policies are obtained through iterative interactions with the environment. This enables the execution of complex optimal decisions in real-time, while judiciously conserving fuel, time and other critical resources. The inherent capability of RL to efficiently map states to control actions provides a unique insight into state-dependent control for robust trajectories, particularly for environments with constantly evolving or even unknown dynamics, where traditional control methods can falter. Despite its potential to significantly enhance onboard guidance systems, the efficacy of RL in adhering to hard constraints remains a formidable challenge. In practice, RL policies must contend with a multitude of competing objectives like terminal constraints, fuel limitations, and desired time-of-flight thresholds. These objectives imbue the learning process with added complexity, necessitating RL policies to navigate through a narrow corridor of acceptable solutions within an expansive search space. One prevalent strategy to address this issue involves incorporating these constraints into the reward function as penalties. Even though this approach helps in incorporating objectives that would otherwise be very difficult to formulate, it often introduces classical problems related to the weightings and constraint accuracy, ultimately resulting in sub-optimal trajectories.

In this study, we propose a novel integration of RL with reachability analysis to tackle multi-impulse transfer problem. Our approach inherently guarantees the satisfaction of final arrival constraints by segmenting the transfer trajectory and formulating each segment as a two-point boundary value problem (TPBVP), subsequently solved using Lambert solvers. Within this framework, RL is tasked with selecting position vectors (waypoints), under the assumption that departure and arrival points are predetermined. This ensures that every generated trajectory adheres to predetermined departure and arrival constraints, while RL optimises intermediate states to minimise total Δv expenditure. Nonetheless, challenges persist, particularly if waypoint locations lack adequate constraints. To mitigate this challenge, we employ rapid assessment of impulsive reachability analysis (RA). This analysis estimates the set of reachable states within a specified time frame, given an initial state and propulsion specifications. We parameterise this set to restrict the position vectors available for RL, thereby delimiting the solution search space for effective trajectory optimisation. Our experimental validation employs simple planet-to-planet transfers as a test case, demonstrating the efficacy and enhanced performance of RL within these integrated frameworks.