

IAF ASTRODYNAMICS SYMPOSIUM (C1)
Mission Design, Operations & Optimization (1) (6)

Author: Mr. Sungmoon Choi
Iowa State University, United States

Dr. Rodney Anderson
Caltech/JPL, United States

Dr. Ashwati Das-Stuart
NASA Jet Propulsion Laboratory, United States

Dr. Ossama Abdelkhalik
Iowa State University, United States

REINFORCEMENT LEARNING FOR MULTIPLE GRAVITY ASSIST TRAJECTORY
OPTIMIZATION**Abstract**

Traditional preliminary trajectory optimization for multiple gravity assists is typically formulated using the patched-conic approximation, where each transfer arc is solved assuming two-body dynamics. This problem has been widely addressed using global optimization algorithms such as genetic algorithms, particle swarm optimization, and differential evolution. However, these methods optimize the entire set of design variables simultaneously. As a result, the optimizer may deviate away from a solution in which one variable is not optimal while the rest of the variables are close to optimal. Then, the optimizer cannot assess whether the previously discarded solutions contain optimal values for some variables.

Reinforcement learning (RL), a machine learning technique based on the Markov Decision Process (MDP), differs fundamentally from traditional optimization approaches. In RL, the next state depends solely on the previous state and the action taken at that state, rather than the entire history of states and actions. Consequently, the cost of previous trajectory arcs, whether high or low, does not influence the current state. This characteristic makes conventional formulations for multiple gravity-assist trajectory optimization incompatible with RL. To address this limitation, we introduce a novel formulation that treats each point-to-point transfer arc independently while adhering to the MDP framework. We define the state, reward, and action based on principles of astrodynamics, constructing an environment that reflects the problem's physical background.

In this study, we implement Proximal Policy Optimization (PPO), one of the most widely used policy-based RL algorithms. Unlike traditional methods that minimize the total cost of an entire trajectory, PPO optimizes the expected return by evaluating and refining individual trajectory segments along the spacecraft's transfer path. This approach allows the algorithm to seek optimal or near-optimal solutions, even when earlier segments incur high costs. As a result, PPO explores a broader solution space while maintaining a balance between exploration and optimization throughout the trajectory design process. Furthermore, by assessing trajectory segments independently, regardless of whether they are high-cost or low-cost, the RL framework leverages stored knowledge within a neural network to improve decision-making at each step.

In this paper, we implement the proposed formulation and demonstrate its effectiveness; we successfully compute the Cassini trajectory. By leveraging RL's ability to utilize GPU acceleration without requiring modifications to the underlying implementation, the method has the potential to significantly improve computational efficiency during the solution search. This paper also discusses the challenges of implementation, potential applications, and directions for future work.