Paper ID: 102774
student

IAF ASTRODYNAMICS SYMPOSIUM (C1)
Interactive Presentations - IAF ASTRODYNAMICS SYMPOSIUM (IP)

Author: Mr. Simone Rotondi
Sapienza University of Rome, Italy

Ms. Valentina Becchetti
University of Rome "La Sapienza", Italy
Dr. Stefano Carletta
Sapienza University of Rome, Italy
Prof. Alessandro Giuseppi
Sapienza University of Rome, Italy

SAFE REINFORCEMENT LEARNING FOR AUTONOMOUS RENDEZVOUS AND PROXIMITY OPERATIONS: BALANCING SAFETY AND PERFORMANCE IN CUBESAT DOCKING MISSIONS

Abstract

Autonomous Rendezvous and Proximity Operations (RPOD) for CubeSats and small satellites are becoming increasingly important for on-orbit servicing, formation flying, and debris mitigation. These operations present unique challenges: the spacecraft must navigate complex dynamics, operate under strict safety constraints, and adapt to uncertain environmental conditions. Traditional control techniques, such as Linear Quadratic Regulators (LQR) and Model Predictive Control (MPC), offer stability and theoretical guarantees but struggle when exposed to unmodeled dynamics, external disturbances, or actuator faults. By contrast, reinforcement Learning (RL) offers flexibility and adaptability to nonlinear and partially observable environments. Still, its exploration-driven learning often leads to unsafe behaviors, especially in the critical phases of approach and docking, where collisions, misalignments or loss of control could jeopardise the mission.

In this work, we propose the use of Safe Reinforcement Learning (Safe RL) to address the safety-performance trade-off in autonomous RPOD operations. We extend the Soft Actor-Critic (SAC) framework with an auxiliary Safety Critic, which estimates the probability that a given state-action pair will lead to a violation of safety constraints, such as collision risk, excessive relative velocity, or orientation misalignment that could lead to a failure of the mission. By incorporating this Safety Critic, the agent can actively avoid unsafe actions during both training and execution, ensuring safe proximity operations even when adapting to new conditions or system dynamics.

The Safe RL approach is applied to a simulated SE(2) planar RPOD scenario, where a 12U CubeSat performs autonomous approach and docking maneuvers with a target. We compare the Safe RL agent's performance to both classical controllers and standard deep reinforcement learning agents.

This work highlights the potential of Safe Reinforcement Learning as a critical enabler for next-generation CubeSat autonomy, combining the flexibility of learning-based approaches with the robustness and operational safety demanded by space missions. Future work will explore extending the Safety Critic concept to incorporate real-time onboard learning for adaptive anomaly handling.