Paper ID: 57609
oral
student

IAF ASTRODYNAMICS SYMPOSIUM (C1)
Attitude Dynamics (2) (9)

Author: Mr. Jacob Elkins
The University of Alabama, United States, jgelkins@crimson.ua.edu

Dr. Rohan Sood
The University of Alabama, United States, rsood@eng.ua.edu
Dr. Clemens Rumpf
NASA Ames Research Center, United States, clemens.rumpf@nasa.gov

AUTONOMOUS SPACECRAFT ATTITUDE CONTROL USING DEEP REINFORCEMENT LEARNING

Abstract

This work presents a novel method for deriving a neural spacecraft attitude controller using reinforcement learning (RL), a paradigm of machine learning. The developed RL attitude controller can tune itself over time to the local dynamics of the system, allowing the controller to perform effectively in complex space environments. Additionally, the study explores the RL implementation specifics that produced the best results and possible avenues of improvement and presents experimental results for performing large-angle slews in the developed simulation environment.

RL is generally defined as optimizing a *policy* $\pi(a_t|s_t)$, a function mapping from states, s_t, to actions, a_t, by iteratively interacting with the environment to maximize reward, r_t. In this study, the policy and value functions are represented as neural networks with varying architectures and activation functions. An actor-critic variant of Proximal Policy Optimization (PPO) is used to train the policy and value networks. The actor-critic algorithms use a state-action value function $V(s_t)$ to map the current state to the *value*, or discounted sum of future rewards, of the current state. RL simulation environments are commonly broken into *episodes*, discrete intervals of training used in batches to train the agent. The attitude control problem is formulated into this episode format by integrating Euler's equations of rigid body rotation and propagating the spacecraft orientation error quaternion using fourth-order Runge-Kutta integration. The Lockheed Martin LM50 SmallSat is modeled for this study. The agent controls discrete torques about the body-fixed principal axes to maximize a reward function based on the attitude error of the current timestep.

In initial results, feedforward neural networks are used for the policy and value functions, with reward structures based on attitude error quaternion values. Training results, thus far, have produced controllers (agents) that can successfully perform various large-angle (30° to 150°) slews to pointing accuracies as high as 0.25°. Current work is being carried out to further refine the reward structure, policy, and value networks to achieve industry-standard performance.

Autonomous, self-improving attitude control systems could revolutionize spacecraft control and motivate the application of RL to other domains of autonomous control. An agent is trained to successfully learn the generalized mapping from a quaternion and angular rate state vector to control torques for large-angle slew maneuvers. This work proves the feasibility of RL for autonomous spacecraft attitude control and provides a useful entry framework to applying RL to the spacecraft attitude control problem.