Project Proposal: Reinforcement Learning for Spacecraft Mode Control

March 1, 2018

Abstract

1 Background

The high cost of space mission operations has motivated several space agencies to prioritize the development of "autonomous" spacecraft control techniques [**?**]. Of particular note are on-board autonomy techniques, which serve to expand spacecraft capabilities in environments where ground contact is not possible. While these techniques offer considerable promise, concerns remain regarding a variety of factors surrounding their implementation, such as the cost of developing such systems and their tolerance for un-modeled behaviors[**?**]. Many of these issues are addressed in other fields through the application of machine learning (ML) techniques, which can provide autonomous systems with the ability to learn from and adapt to their environments without human intervention. These techniques, coupled with high-fidelity simulation tools, are a potential avenue for reshaping the manner in which flight systems are developed and fielded.

Currently, Bayesian estimation is the de facto method used for autonomy. By obtaining the maximum likelihood estimate of a spacecrafts states given a set of measurements, Kalman Filters provide an optimal solution for guidance. Although this is a robust solution, these filters often fall short when it comes to fault detection. They also demand expertise when being developed, modified, or maintained. Reinforced learning methods solve the optimization problem in a new and scarcely used way. The reason for their lack of popularity in the space sector is that data is not plentiful, whereas ML techniques rely on the abundance of data. Yet, with improved simulation methods, runs can now be generated swiftly and with high fidelity, hence yielding the data needed for reinforced learning.

2 Machine Learning Focus

This project topic is primarily motivated by our group's interest in understanding reinforcement learning techniques and their application to problems in aerospace. However, the broad scope of the topic can include other aspects of machine learning, such as classification and inference, if time allows.

Reinforcement learning is broadly defined as a class of machine learning techniques that maximize the earned reward of a software agent that interacts with a given environment.

3 Problem Statement

At present, we propose parameterizing the spacecraft mode-control problem in the following manner.

3.1 State Description

Environment States

1. Spacecraft Position: $\mathbf{r} \in \mathbb{R}^3$
2. Spacecraft Velocity: $\mathbf{v} \in \mathbb{R}^3$
3. Reference Position: $\mathbf{r_0} \in \mathbb{R}^3$
4. Reference Velocity: $\mathbf{v_0} \in \mathbb{R}^3$
5. Spacecraft Error Indicator: $E \in \{0, 1\}$

On-Board States

1. State Vector: $\mathbf{X} = \begin{bmatrix}\mathbf{r} & \mathbf{v}\end{bmatrix}^T \in \mathbb{R}^6$
2. State Estimate: $\hat{\mathbf{X}} \in \mathbb{R}^6$
3. State Error: $\mathbf{x} \in \mathbb{R}^6$

3.2 Action Model

This project considers a spacecraft with the following operational modes:

1. **Orbit Determination:** In this mode, the spacecraft uses its sensors to reduce the error in its estimate of its orbit. However, perturbing forces such as solar radiation pressure will drive it away from its desired trajectory.
2. **Orbit Control:** In this mode, the spacecraft uses thrusters to command its current state estimate towards its desired trajectory. We model this as bringing the current state-desired error towards the state-estimate error.

3. **Mission Maneuver**: In this mode, the spacecraft conducts a large burn to enter a transfer trajectory or insert itself in an orbit around a planet.

4. **Science Operations:** In this mode, the spacecraft collects a reward proportional to its state error. However, both its orbit determination and orbit control errors increase with time.

5. **Safe Mode:** if the spacecraft enters an error state (i.e., if the error state variable is 1), entering this mode will reset the error state to 0. No other behavior occurs.

4 Proposed Work Plan

The work for this project is envisioned to proceed as follows:

1. Parametereize the spacecraft mode control problem.

2. Develop an environment model for this problem within the Partially-Observable Markov Decision Process framework

3. Implement a "naive" reinforcement learning approach.

4. Implement more complex reinforcement learning approaches.