Paper ID: 57168
interactive
student

IAF ASTRODYNAMICS SYMPOSIUM (C1)
Guidance, Navigation and Control (1) (1)

Author: Mr. Linfeng Li
Beijing Institute of Control Engineering, China Academy of Space Technology (CAST), China,
llfzju@hotmail.com

Prof. Yong Chun Xie
Beijing Institute of Control Engineering, China Academy of Space Technology (CAST), China,
xieyongchun@vip.sina.com
Dr. Yong Wang
Beijing Institute of Control Engineering, China Academy of Space Technology (CAST), China,
wyongzju@163.com
Mrs. Ao Chen
Beijing Institute of Control Engineering, China Academy of Space Technology (CAST), China,
chenao526@foxmail.com

SAFETY ANALYSIS FOR SPACE ROBOTIC MANIPULATION: A REINFORCEMENT LEARNING PERSPECTIVE

Abstract

To leverage the exploration and utilization into deeper space, upgrading spacecraft's autonomous level is the trend for developing space technology. One feasible solution is robotic on-orbit service (OOS). For decades, space robot has shown the potential to assist, cooperate with, or even substitute human's role in space manipulation tasks. The increasing performance needs a more refined safety analysis. However, the traditional stability analysis, which is a major concern for control theory and engineering, cannot thoroughly depict the safety under uncertainty of space environment. With the help of recent advances in deep reinforcement learning, a robotic system can learn a control policy based on large data of interaction behavior collected in complex and uncertain environments. Thus, it is valuable to investigate in the learning process that how decision making and optimization can impact behavior policy's safety property. In this paper, we use the approach of constrained Markov decision processes (CMDPs), which is shown in some literature suitable to merge with the Lyapunov method and have a theoretical guarantee, to dealing with safe reinforcement learning for space robotic manipulation. Specifically, we consider two working conditions: (1) Binding: the end of robot arm is connected with astronaut's feet; (2) Autonomy: robot arm performs task without human interference. We propose a modified version of Safe Policy Iteration (SPI) for the above conditions, based on different Lyapunov function formulations. We further validate the algorithm in a simulation environment for space robotic refueling task. The result shows that the learned policy can constrain the system states within the safety region, thereby increasing the overall safety level.