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

Author: Dr. Simeng Huang
China

Mr. Jie Li
Beijing Institute of Technology, School of Aerospace Engineering, China
Mr. Yuxing Wang
Tsinghua University, China
Prof. Bin Liang
Tsinghua University, China
Prof. Mingming Shi
China
Dr. Yabo Hu
Tsinghua University, Beijing, China
Prof. Yiyong Sun
Beijing Institute of Technology, School of Aerospace Engineering, China

AN ORBITAL GAME CONTROL ALGORITHM FOR ON-BOARD APPLICATION BY BEHAVIOUR CLONING

Abstract

The orbital game between a pursuing spacecraft and an evading spacecraft has been widely studied. However, limited by the on-board computing cability, complex orbital dynamics and dynamic imaging conditions, applications of the current methods are rarely tested in real missions. This paper introduces a novel orbital game control algorithm designed for on-board application via behaviour cloning, a deep supervised learning method to imitate a pre-designed optimal behaviour. The study is carried out in two phases. The first phase is to generate the imaging-constrained optimal game control training dataset by solving various relevant orbital game control problems using an up-to-date indirect method. The second phase is to obtain an orbital game control policy modelled using a deep neural network to imitate the optimal behaviour derived in the first phase, by means of a behaviour cloning method. The first phase is organised as follows. First, the dot-product inequalities of the position and velocity vectors as well as the dot-product inequalities of the position and illumination vectors are used to describe the detection range constraint and the spatial imaging constraint of the pursuer, respectively. Then, the orbital game control problem is transformed into a two-point boundary value problem by using the minimum principle, and the two-point boundary value problem is discretised by the pseudo-general method. Following that, a high-dimensional optimisation problem with inequality constraints can be formulated. Finally, the control sequence of the pursuer is obtained by combining the divide-and-conquer search of the game terminal time and the sequential quadratic programming. The second phase is organised as follows. First, a state space is established by considering the constraints and initial conditions of an extensive array of pursuers and evaders across diverse scenarios. Then, the initial state sample set is formulated and delineated by the boundaries of the state space, while the desired control sequence sample set for the initial sample set is derived with the method delevoped in the first phase. Consequently, the training dataset is constructed by combining the two sample sets. Finally, the control policy model is trained by behaviour cloning in a supervised manner. In particular, to handle the distributional drift during the training, we integrate expert

demonstrations into the training dataset for data augmentation. Upon the on-board implenmention, the pursuer can utilise the current state as the inputs for the control policy, enabling swift generation of the optimal control outputs.