IAF SPACE OPERATIONS SYMPOSIUM (B6)
Interactive Presentations - IAF SPACE OPERATIONS SYMPOSIUM (IPB)

Author: Dr. Siyao Lu
Beijing Institute of technology, China, lsylusiyao@bit.edu.cn

Prof. Ai Gao
Beijing Institute of technology, China, gaoai@bit.edu.cn
Prof. Rui Xu
Beijing Institute of technology, China, xurui@bit.edu.cn
Ms. Zhaoyu Li
Beijing Institute of technology, China, lzyzhaoyu@163.com
Mr. Pan Huang
China, 15011502134@163.com
Mr. Chen Zhao
China, 105010161@qq.com

A MULTI-AGENT PLANNING METHOD ON DEEP REINFORCEMENT LEARNING FOR LUNAR ROVERS COLLABORATED OPERATION WITH UNCERTAINTY

Abstract

International Lunar Research Station will be established around 2035 by China and Russia. At the beginning period, advanced lunar rovers equipped with robotic arms act as constructors of the station. Carrying rovers collaborate with each other to collect lunar soil and lunar water, then transport them to the mixing blender where construction materials are created. However, rovers start from different places, go towards different mining sites, drive through different paths, collect different kinds of resource, arrive to the blender at different time. These situations above make activity planning and path planning for operating rovers full of hard. What's more, obstacles are unknown for several reasons, such as lunar maps' resolution is not high enough for rovers to plan and see in advance and lunar solid leads rovers' speed hard to control as excepted, which leads to uncertainty for rovers. Therefore, we propose a new way of collaborated planning through multi-agent deep reinforcement learning, where simulation environments with randomly generated obstacles are established to train the rovers complete tasks with barriers avoided and decisions made according to circumstances. Actions for environments and tasks are previously trained into neural networks to change the mode of planning and implementing to instant decision making to avoid planning repair. First, to simulate the moon surface, we design a way of establishing the training environment where there're craters of different shape and size, obstacles of different size and location. Both the two kinds of barriers obstruct the path of rovers, and they need to bypass. Second, we propose an architecture on deep reinforcement learning to let rovers decide the next step instantaneously according to the surroundings. Because rovers are supported by battery, training is targeting to minimize the cost of power consumption or some other custom metrics. In our experiment, we intend that our way leads each rover finish the carrying work with no collision while consuming less energy and finishing the work quicklier than the traditional way of planning in advance and planning repairing during implementation.