Abstract of Master Thesis 2020

Collective Transport Behavior in a Robotic Swarm with Hierarchical Reinforcement Learning

M195690 韓 子堯

【Introduction】

Swarm robotics is a field of research in which multiple agents coordinate their behaviors to solve problems that exceed an individual's ability. There are two design methods for developing the robotic swarm controller, the behavior-based design method and the automatic design method. Reinforcement learning (RL) is a typical approach in the automatic design method. RL traditionally refers to a class of learning problems: a robot learns behavior through trial-and-error interactions with an environment and receives positive and negative rewards for its actions. However, when reinforcement learning is utilized as the design method to achieve some multi-objective tasks, the developed controller gets poor performances due to the sparse rewards problem. The sparse rewards problem refers to the problem that it is difficult to obtain positive rewards in exploration, resulting in slow or impossible in learning. Solving this problem is an important issue in RL, and this is also the motivation of this study. In this research, the main objective is to design a training method to develop a robotic swarm controller to achieve a collective transport task, a multi-objective task that is too tricky to be achieved by the conventional RL method with sparse reward settings.

【Method】

In this research, the collective transport task is designed as a multi-objective task, where robots need to learn to push two kinds of food resources, the resource pack and the resource. There are two goals in this task. The first one is pushing the resource packs to the split area, and the second one is pushing the resources to the nest area. A hierarchical RL method is applied to train a robotic swarm controller to generate collective transport behaviors. In this training method, the original task decomposes into two sub-tasks: the decomposition and the transport task. In the decomposition task, robots need to achieve the first goal, and in the transport task, robots need to accomplish the second goal. The sub-controllers are developed to achieve those sub-tasks. A robotic swarm controller loaded with the trained sub-controllers is then developed with sparse reward settings to complete the original collective transport task. In this study, Proximal Policy Optimization, an RL algorithm, is utilized for training all controllers.

【Simulated Results】

The computer simulations are implemented in Unity3D. Fig.1 Shows the simulation environment. Three sets of experiments are conducted to compare the performances between the conventional RL method and the hierarchical training method. Fig.2 shows the episode length in each episode during the training process, representing the time it takes for the robots to complete the task. The controller developed by the conventional RL method with sparse reward settings is failed to accomplish the given task because the episode length is always 6000, the maximum time step in simulations. The conventional RL method with dense reward settings and hierarchical RL method succeed in achieving the collective transport task because the episode lengths converge to about 2500 and 1500, respectively. Moreover, the controller developed with the hierarchical RL method starts to learn to complete the task at about the 1000000th time step, much earlier than the conventional RL method.

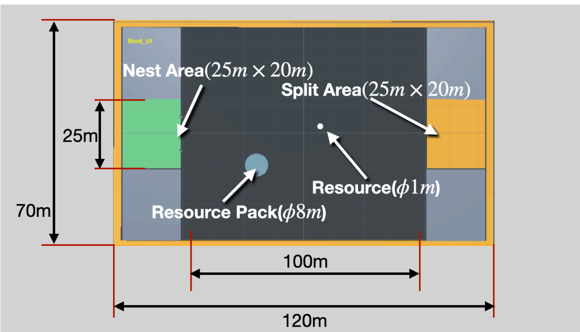

Fig. 1. Top View of Environment

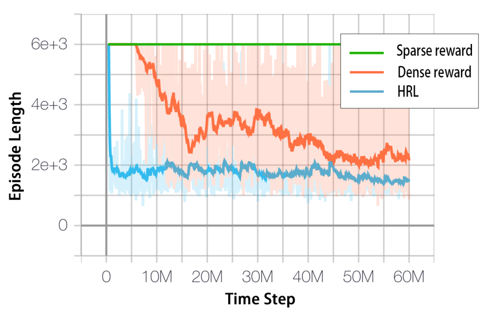

Fig. 2. The Episode Length of Different Training Methods(the lower one is better)