Maze Search Using Reinforcement Learning by a Mobile Robot

Makoto Katoh*, Keiichi Tanaka and Shunsuke Shikichi

Department of Control System Laboratory, Osaka Institute of Technology, Japan

Received: 03 October 2018; Published: 10 October 2018

*Corresponding author: Makoto Katoh, Department of Control System Laboratory, Osaka Institute of Technology, Japan

Abstract

This review presents on research of application of reinforcement learning and new approaches on a course search in mazes with some kinds of multi-point passing as machines. It is based on a selective learning from multi-directive behavior patterns using PS (Profit Sharing) by an agent. The behavior is selected stochastically from 4 kinds of ones using PS with Boltzmann Distribution with a plan to inhibit invalid rules by a reinforcement function of a geometric sequence. Moreover, a variable temperature scheme is adopted in this distribution, where the environmental identification is valued in the first stage of the search and the convergence of learning is shifted to be valuing as time passing. A SUB learning system and a multistage layer system were proposed in this review, and these functions were inspected by some simulations and experiments using a mobile robot.

Keywords: Autonomous; Mobile Robot; Learning; Agent; System Simulation

Introduction

In robots which has begun to spread to not only industrial world but also general home, e.g. cleaning robots etc., recently achievement of complex tasks and adaptation of complex environment has been required and can be done by agents which were concept of distributed artificial intelligent and caught abstractly various robots. Conventionally, as behavior of agents has been controlled by rules designed as if then rules, a lot of rules were required for adaptation to complex environment and achievement of complex tasks. Then, in fact, it is impossible that human designers design an individual rule of each environment.

Then, a lot of reinforcement learning researches, e.g., Q-Learning (QL), Profit Sharing (PS), Instance-Based (IB), which is an unsupervised learning to attain optimal task by learning the environment based on the agent behavior without foresight knowledge on the objects and environments, are paid to attention. The various application areas such as maze search [1], optimal route search [2], a design of dynamic route navigation system using electrical maps [3] have been considered. Especially, a new method of integration with reinforcement learning and A* algorithm which is one of the shortest route search algorithms which do not use learning etc. is groped for in the application to the route search. The advantage of integrating reinforcement learning to such algorithm without learning is that trial and errors of the agent achieves the target even if only the target point is given, and the environment is unknown (Even if the unknown dynamic changes exist).

The reinforcement learning is more effective than the shortest route search algorithms in the case of unknown route as a maze or unknown dynamic change by the way. Then, it is necessary to choose suitable field for them when the field of application of reinforced learning is set. Basic Profit Sharing (PS) has been theoretically considered by Muraoka and Miyazaki [4]. Recently, Kawada proposed the efficient maze search method which improved the action selection machine and the study machine of Profit Sharing (PS). It is an action selection switch type with a premeditated action selection machine, and the method of not strengthening the rule again more than the necessity at learning.

Moreover, it is pointed out that PS is more advantageous than QL in the maze search because the number of steps in PS is convergence which was known from the results of the comparison of numerical value experiment of PS and Q-Learning (QL). Besides, there is a research which is not batch payment but makes the reward installments of two stages in the goal, too.