The Sixteenth International Symposium on Artificial Life and Robotics 2011 (AROB 16th '11),
B-Con Plaza, Beppu,Oita, Japan, January 27-29, 2011

# A Study on Reinforcement Learning Scheme in Cooperative Network-Systems of Sensor Nodes and Mobile Robots

Koichi Hirayama†, Akihide Utani‡, and Hisao Yamamoto‡

†*Graduate School of Engineering, Tokyo City University*
‡*Faculty of Knowledge Engineering, Tokyo City University*
*1-28-1, Tamazutsumi, Setagaya, Tokyo 158-8557, Japan*
*{autani; yamahisa}@tcu.ac.jp*

***Abstract*:** Recently, there is growing expectation for a new network service by a wireless sensor network consisting of many sensor nodes placed in an object area and not a few mobile robots as a result of the strong desire for the development of systems that function flexibly in dramatically changing environments. This study proposes a reinforcement learning scheme in a cooperative network-system of many static sensor nodes and not a few mobile robots, named the Network-Robot System (N-R System) for adaptive cooperation of many static sensor nodes and not a few mobile robots. We evaluate the proposed reinforcement learning scheme by computer simulations on a benchmark problem. In the experiments performed, the performance of the proposed scheme is investigated to verify its effectiveness.

***Keywords*: Reinforcement Learning, Cooperative Systems, Static Sensor Nodes, Mobile Robots.**

## I. INTRODUCTION

Various network services have been provided. They include inter-vehicle communication, which is a network service in intelligent transport systems, natural environmental monitoring, and emergent communication between mobile nodes in such the case of emergency as disaster. As a means of realizing the above network services, autonomous decentralized networks, such as a mobile ad-hoc network and a wireless sensor network, have been intensively studied with great interests. Especially, a wireless sensor network, which is a key network to facilitate ubiquitous information environments, has great potential as a means of realizing a wide range of applications, such as natural environmental monitoring, environmental control in office buildings and factories, object tracking, and precision agriculture [1]. Recently, there is growing expectation for a new network service by a wireless sensor network consisting of many sensor nodes placed in an object area and not a few mobile robots as a result of the strong desire for the development of systems that function flexibly in dramatically changing environments.

This study proposes a reinforcement learning scheme in a cooperative network-system of many static sensor nodes and not a few mobile robots, named the Network-Robot System (N-R System) for adaptive cooperation of many static sensor nodes and not a few mobile robots. We evaluate the proposed reinforcement learning scheme by computer simulations on a benchmark problem. In the experiments performed, the performance of the proposed scheme is investigated to verify its effectiveness. This paper is organized as follows. In Section II, the proposed scheme is outlined and the experimental results on a benchmark problem are reported. Finally, the paper closes with conclusions and ideas for further study in Section III.

## II. PROPOSAL AND RESULTS

In the existing studies, many reinforcement learning algorithms have been proposed as agent-adaptive algorithms. In the existing reinforcement learning algorithms, the agent is reinforced (trained) using only rewards from the environment. Reinforcement learning algorithms can be classified into two approaches. One is "exploitation-oriented" approach, and the other is "exploration-oriented" approach. This study uses the "exploitation-oriented" approach. In the proposed N-R System, many static sensor nodes limited resource, which are compact and inexpensive, are placed in an object area. Each sensor node consists of a sensing function to measure the status of an object, a limited function on information processing, and a simplified wireless communication function, and operates on a resource of a limited power-supply capacity such as a battery. Not a few mobile robots have learning function and learn adaptive actions based on information obtained from static sensor nodes placed in an object area.

The proposed N-R System, which is a cooperative network-system of many static sensor nodes and not a few mobile robots, consists of the following main modules: Perception Module, Learning Module.

**1. Perception Module**

In the perception module, the sensor node perceives the task (request) and disseminates the task (request) information and the route information to the task point to all nodes that include mobile robots.

**2. Learning Module**

In the learning module, each mobile robot learns adaptive actions based on information obtained from static sensor nodes placed in an object area. As the learning scheme, the profit sharing plan is used [2,3].

Through simulation experiments on the delivery problem by mobile robots in the plant shown in Fig.1, the performance of the proposed N-R System is investigated to verify its effectiveness. The N-R System consisting of twenty-five static sensor nodes arranged in the object plant and two mobile robots is assumed. The conditions of simulation, which were used in the experiments performed, are shown in Table1. In the initial stage, each mobile robot has fifty requirements.

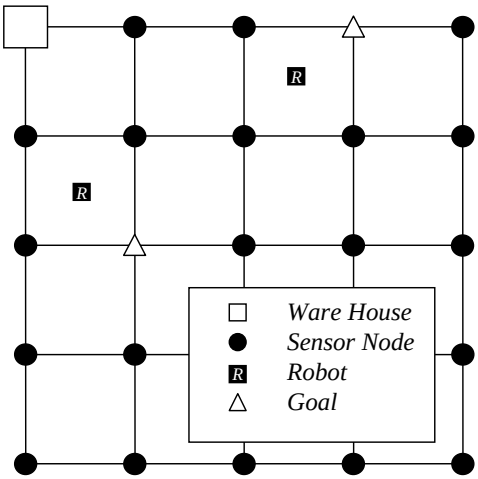


Fig.1. Simulation model

Table1. Conditions of simulation

| | |
|---|---|
| Simulation size | 50m × 50m |
| The number of sensor nodes | 5 × 5 |
| The number of mobile robot | 2 |
| The number of warehouses | 1 |
| The number of requirement points | 2 |
| The number of episodes | 500 |
| Reward | $6.0\times10^5$ |

Experimental results on each pattern shown in Table2 are illustrated in Fig.2 and Fig.3. Fig.3 shows the learning curves of the case that the warehouse functions as a subgoal. In case that the warehouse functions as a subgoal, the profit is shared with each mobile robot by the profit sharing plan [2,3] when it reaches the warehouse. Fig.4 and Fig.5 show the experimental results on Pattern 4 and Pattern 5. From these results, it can be confirmed that the subgoal effectively functions.

Table2. Experimental patterns

| | The total number of requirements | Direct |
|---|---|---|
| Pattern1 | 0～50 | 2 |
| Pattern2 | 51～100 | 2 |
| Pattern3 | 51～100 | 1 |
| Pattern4 | 101～150 | 1 |
| Pattern5 | 101～150 | 0 |
| Pattern6 | 151～ | 0 |

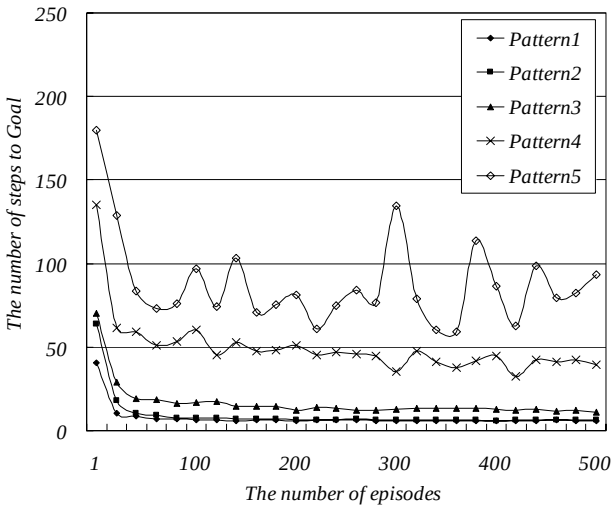


Fig.2. Learning curves

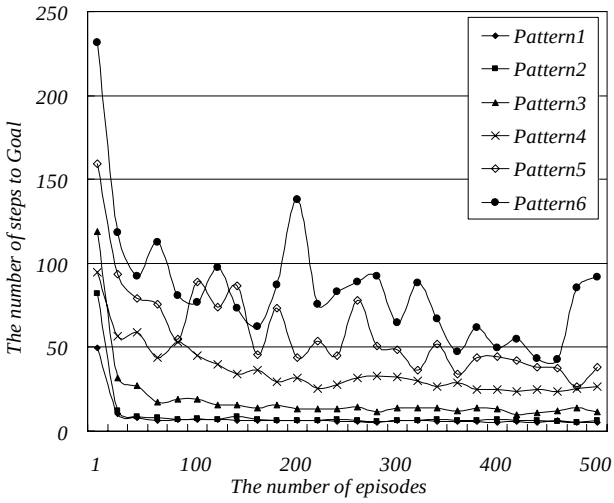


Fig.3. Learning curves (Subgoal is used)

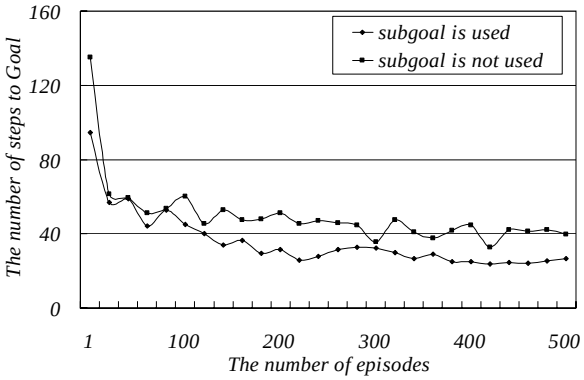


Fig.4. Learning curves on Pattern 4

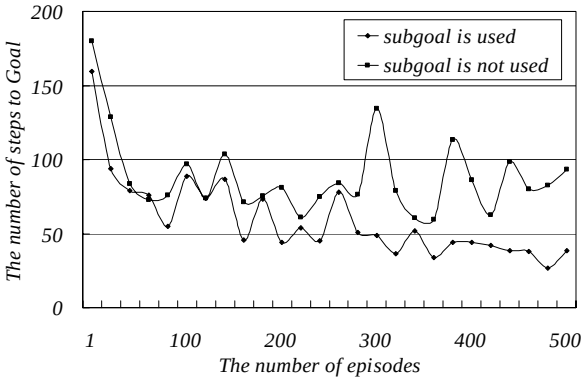


Fig.5. Learning curves on Pattern 5

## III. CONCLUSIONS

In this paper, a reinforcement learning scheme in a cooperative network-system of static sensor nodes and mobile robots, named the Network-Robot System for adaptive cooperation of static sensor nodes and mobile robots, has been proposed. In simulation experiments on the delivery problem by mobile robots in the plant, the performance of the proposed system was investigated to verify its effectiveness. Experimental results indicate that the proposed system has the development potential to facilitate ubiquitous information environments. In future studies, we evaluate the performance of the proposed system in detail.

## REFERENCES

[1]I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," Computer Networks Journal, vol.38, no.4, pp.393-422, 2002.
[2]K. Miyazaki, M. Yamanoto, and S. Kobayashi, "A theory of profit sharing in reinforcement learning," Trans. of the Japanese Society for Artificial Intelligence, vol.9, no.1, pp.104-111, 1994.
[3]K. Miyazaki, H. Kimura, and S. Kobayashi, "Theory and applications of reinforcement learning based on profit sharing," Trans. of the Japanese Society for Artificial Intelligence, vol.14, no.5, pp.800-807, 1999.