Exploring XCS in Multiagent Environments *

Hiroyasu Inoue
ATR NIS
2-2-2 Hikaridai, Seika-cho
Soraku-gun, Kyoto, Japan
hir_inoue@atr.jp

Keiki Takadama
Tokyo Institute of Technology
4259 Nagatsuta-cho
Midori-ku, Kanagawa, Japan
keiki@dis.titech.ac.jp

Katsunori Shimohara
ATR NIS
2-2-2 Hikaridai, Seika-cho
Soraku-gun, Kyoto, Japan
katsu@atr.jp

ABSTRACT

This paper investigates the adaptability of XCS in four different multiagent environments. The environments are realized in a simplified soccer game, and they include (1) single-agent environment, (2) multiagent environment with an opponent, (3) multiagent environment with a teammate, (4) multiagent environment with both an opponent and a teammate. Although XCS generally seems inferior to strength-based XCS in such stochastic environments, experimental results in a specific stochastic environment show that XCS is superior to strength-based XCS. Furthermore, XCS with profit sharing is more effective than one using the bucket brigade in multiagent environments.

Categories and Subject Descriptors

I.2.6 [**Learning**]: Knowledge acquisition

General Terms

Algorithms, experimentation

Keywords

Learning Classifier System, Multiagent, Profit sharing, Bucket brigade

1. INTRODUCTION

Currently, XCS[3] is one of the most popular Learning Classifier Systems (LCS), and its effectivity has been examined. However, XCS has not yet been investigated in multiagent environments. Therefore, this paper aims at investigating the adaptability of XCS in multiagent environments. Particularly, we compare the following four types of XCS: (1) accuracy-based XCS with bucket brigade (original XCS), (2) strength-based XCS with bucket brigade, (3) accuracy-based XCS with profit sharing, and (4) strength-based XCS with profit sharing.

2. CATEGORIES OF MULTIAGENTS

This paper uses a simplified soccer game as a typical example of multiagent environments and classifies soccer environments into four categories: Category 0: single-agent environments, which are ordinary environments that have often been used in investigations of XCS. Category 1: multiagent environments with an opponent, where transitions of states randomly change. Category 2: multiagent environments with a teammate that learns simultaneously. Category 3: multiagent environments with both an opponent and a teammate. These four categorized environments are different in how stochastic they are, which is important in discussing the adaptability of XCS.

3. SIMPLIFIED SOCCER

We employ a simplified soccer game for experiments consisting of discrete squares. The field is a rectangle and has 5 squares in the vertical direction and 9 squares in the horizontal direction. There are a left and a right team and a ball. The vertical line on the right side is the left team's goal, and vice versa. If the ball traverses the left goal, the left team gets a goal, and vice versa. At the beginning of the experiments or after a goal, the agents are randomly set in the field and the ball is set to the center of the field.

The sensors of agents are represented by seven bits whose data are used for the condition parts of XCS. The first four bits are assigned to information of four directions (upper, lower, left and right) of teammates. If any teammate is in a particular direction, the bit corresponding to that direction is set to 1, otherwise to 0. Next, three bits are assigned to the direction of the ball. If the ball is in the upper, lower, left, or right squares, 010, 011, 001, and 000 are set to three bits, respectively. The remaining areas, upper-right, upper-left, lower-right, and lower-left, are set to 100, 101, 110, and 111, respectively.

The simplified soccer game is conducted in discrete time. At every time step, all agents simultaneously move either to the left, right, up, down, or not at all. If an agent goes out of bounds, the agent is returned to the previous square. The agents and the ball can freely move into the same square. If at least one agent moves into the same square as the ball, it is moved (kicked) toward the goal of the agent's team by a distance of 4 squares, and after that it moves vertically and horizontally by less than one square at random. If the

*This research was conducted as part of 'Research on Human Communication' with funding from the National Institute of Information and Communications Technology.

GECCO'05 June 25–29, 2005, Washington, DC, USA.

ball is about to crosse a horizontal line, it goes back before this happens. If more than two agents move into the same square, one agent is selected randomly, and the ball is kicked.

4. EXPERIMENTS

4.1 Agents and environments

We employ the following four types of XCS: (1) accuracy-based XCS with bucket brigade (original XCS), (2) strength-based XCS with bucket brigade[1], (3) accuracy-based XCS with profit sharing, and (4) strength-based XCS with profit sharing.

Table 1 shows parameter settings of XCS. These settings are ordinary ones except θ_{GA} and β. θ_{GA} is 10,000 and larger than usual, but agents take many steps to acquire rewards in multiagent environments, so this is rational. β is 0.02 and smaller than usual. This is because multiagent environemts are stochastic in many ways, and thus deliberate learning is required. In addition, profit sharing is set according to Miyazaki's paper[2], and the common ratio of geometric progression is 0.9. If a team consisting of two agents gets a goal, 1.0 is given to the agent which kicks the ball last and 0.9 is given to the other.

We use four environments, which are classified by the number of agents they include. These environments are one left team agent and no right team agent (1on0), one left team agent and one right team agent (1on1), two left team agents and no right team agent (2on0), and two left team agents and one right team agent (2on1). They correspond to Categories 0, 1, 2, and 3, respectively. Only left team agents are implemented by XCS or modified XCS, and a right team agent is simply a random agent (if any).

From the above description, we have 16 combinations of experiments, which consist of four different types of XCS and four different environments.

5. EXPERIMENTAL RESULTS

Table 2 shows the results of experiments, which include 16 combinations of experiments. We execute each experiment in 10 trials with different random seeds. Each trial is held in 500,000 steps, and the goals made in every 10,000 steps are counted. Consequently, 50 values are acquired in each trial, and these values are averaged over 10 trials. Table 2 shows their maximum values. Acc. and Str. indicate that accuracy-based XCS and strength-based XCS, respectively.

6. DISCUSSIONS AND CONCLUSIONS

Random changing environments: We investigate the adaptability of XCS to multiagent environments with one opponent where transitions of states randomly change. This study is done by comparing results between 1on0 and 1on1, and between 2on0 and 2on1. Note that the existence of an opponent (right team) agent cannot be seen by learning (left team) agents in the experiments, which means that the ball transitions are not predictable. Table 2 shows that the goals of all agents decrease from 1on0 to 1on1 and from 2on0 to 2on2; however, there is no remarkable difference in the tendencies, which implies that there is no difference among the adaptabilities of agents in this multiagent environment.

Simultaneous learning environments: We investigate the adaptability of XCS to multiagent environments with one teammate by comparing the results between 1on0 and 2on0 and between 1on1 and 2on1. Note that the two sets have different number of left team agents, which is the only difference in the sets. Table 2 shows that the goals of all agents increase from 1on0 to 2on0 and from 1on1 to 2on1. These results naturally correspond to the expectation that plural agents can get more goals than a single agent. Furthermore, there are differences among the inclination tendencies. The detail is shown in the rest of this section.

Bucket brigade vs. profit sharing: Accuracy-based XCS with profit sharing is apparently superior to accuracy-based XCS with bucket brigade in environments with a teammate, although profit sharing is inferior to bucket brigade with a single agent. The strength-based XCS also has the same tendency. These results mean that the profit sharing method has high adaptability to simultaneous learning environments. It seems that they have to coordinate their actions for high adaptation, but they cannot predict clearly the actions of other agents. Therefore, the trial-and-error prevents XCS with bucket brigade from acquiring accuracy or strength appropriately because it is a bootstrap system. On the other hand, profit sharing is robust against such situations because it is a non-bootstrap system.

Accuracy-based vs. strength-based: We compare the adaptability of accuracy-based XCS and strength-based XCS. Generally, strength-based XCS seems to be robust against noisy environments. In the above experiments, a multiagent environment, where transitions of states randomly occur, accord to the general noisy environments. However, experimental results do not show remarkable robustness in such environments. Moreover, strength-based XCS is inferior to accuracy-based XCS in 2on0 and 2on1, although there is no remarkable difference in 1on0 and 1on1. This means that strength-based XCS is weak in environments with plural learning agents. This is because strength-based XCS yeilds over-generalizations and creates many strong classifiers that directly contribute to obtaining rewards. Therefore, they tend to behave like greedy agents, which reduces their chances to coordinate cooperative behaviours. However, accuracy-based XCS does not generate such over-generalizations, which enables agents to coordinate cooperative behaviours, thus resulting in a higher frequency of goals.

From the above discussion, we conclude this paper by making the following points. We investigated the adaptability of four different XCSs in four different multiagent environments. Although accuracy-based XCS is generally inferior to strength-based XCS in such stochastic environments, experimental results in a specific environment show that accuracy-based XCS is superior to strength-based XCS. Furthermore, XCS with profit sharing is more effective than

Table 2: Experimental Results: Maximum goals per 10,000 steps

	Left team agents			
	Bucket brigade		Profit sharing	
Environments	Acc.	Str.	Acc.	Str.
1on0	184.3	192.7	72.2	71.2
1on1	153.7	157.4	61.4	52.0
2on0	536.1	422.1	594.7	443.9
2on1	451.1	386.4	496.2	429.2

Table 1: Parameter settings of XCS

Parameter	θ_{sub}	θ_{GA}	θ_{del}	θ_{mna}	δ	N	β	α	ν	ϵ_0	χ	μ	$P_{\#}$
Value	20	10,000	20	5	0.1	200	0.02	0.1	5	0.01	0.5	0.01	0.33

one with bucket brigade in multiagent environments.

7. REFERENCES

[1] T. Kovacs. XCS's strength-based twin. Part I. In *Learning Classifier Systems*, volume LNCS 2661, pages 61–80. Springer, 2003.

[2] K. Miyazaki, M. Yamamura, and S. Kobayashi. On the rationality of profit sharing in reinforcement learning. In *Proceedings of the 3rd International Conference on Fuzzy Logic, Neural Nets and Soft Computing*, pages 285–288, 1994.

[3] S. Wilson. Classifier fitness based on accuracy. *Evolutionary Computation*, 3(2):149–175, 1995.