Submitted to Fuzzy Sets and Systems (Final Version)

Learning fuzzy rules from iterative execution of games

Hisao Ishibuchi, Ryoji Sakamoto, Tomoharu Nakashima

Department of Industrial Engineering, Osaka Prefecture University

1-1 Gakuen-cho, Sakai, Osaka 599-8531, Japan

Corresponding author: Hisao Ishibuchi

Department of Industrial Engineering, Osaka Prefecture University

1-1 Gakuen-cho, Sakai, Osaka 599-8531, Japan

Phone +81-722-54-9350   Fax +81-722-54-9915

E-mail: hisaoi@ie.osakafu-u.ac.jp

Learning fuzzy rules from iterative execution of games

Hisao Ishibuchi, Ryoji Sakamoto, Tomoharu Nakashima

Department of Industrial Engineering, Osaka Prefecture University

1-1 Gakuen-cho, Sakai, Osaka 599-8531, Japan

Abstract

This paper discusses the linguistic knowledge extraction from the iterative execution of a multi-player non-cooperative repeated game. Linguistic knowledge is automatically extracted in the form of fuzzy if-then rules. Our knowledge extraction is mainly based on the on-line incremental learning of fuzzy rule-based systems. In this sense, our linguistic knowledge extraction is the learning of fuzzy rules. We first briefly describe a market selection game, which is formulated as a non-cooperative repeated game with many players and several alternative actions. We also describe some simple strategies for our market selection game. In our market selection game, the payoff of each player depends on the actions of all players. When a particular action is chosen by many players, those players receive low payoff. High payoff is obtained from actions chosen by only a small number of players. This means that minority players with respect to their actions receive high payoff. Next we show how our market selection game can be handled as a pattern classification problem where a single training pattern is successively generated from every round of our game. A fuzzy rule-based classification system is used as a decision-making system by each player for choosing an action in every round. An on-line incremental learning algorithm is proposed for adjusting the fuzzy rule-based classification system. Then we show how our market selection game can be handled as a function approximation problem. A fuzzy rule-based approximation system is used as a value function for approximating the expected payoff from each action. Finally simulation results show that comprehensible linguistic knowledge is extracted by the learning of fuzzy rule-based systems.

Keywords: Knowledge extraction, non-cooperative game, repeated game, learning, linguistic modeling, decision analysis, fuzzy rule-based system.

1. Introduction

Recently various approaches have been proposed for extracting comprehensible linguistic knowledge from numerical data [6,12,13,19,20,22,23,28]. In those studies, emphasis was placed on the comprehensibility or transparency of extracted knowledge as in the field of knowledge discovery and data mining [5,21]. In this paper, we show how linguistic knowledge can be extracted from the iterative execution of a market selection game, which is a multi-player non-cooperative repeated game formulated in our former study [7]. Our linguistic knowledge extraction is mainly based on the on-line incremental learning of fuzzy rule-based systems. In this sense, our linguistic knowledge extraction is the learning of fuzzy rules. We propose two learning schemes. In one scheme, our market selection game is handled as a pattern classification problem where a single training (i.e., labeled) pattern is successively generated from every round of our game. A fuzzy rule-based classification system [3,4,6,8,9,16,17] is used as a decision-making system for choosing an action in the current round based on the result of the previous round. In the other scheme, our market selection game is handled as a function approximation problem where a single input-output pair is successively generated from each round. A fuzzy rule-based approximation system is used as a value function for approximating the expected payoff from each action. In this learning scheme, fuzzy rule-based systems can be viewed as function approximators as in their applications to function approximation problems and control problems [14,26].

Evolution of game strategies has been mainly studied for the Iterated Prisoner's Dilemma (IPD) game [1,2,15,27]. In those studies, game strategies were evolved by genetic operations. Our market selection game is much more complicated than the IPD game in its payoff mechanism. The payoff mechanism in our market selection game cannot be represented in a simple tabular form while it is usually represented by a 2×2 payoff matrix in the IPD game. Moreover our market selection game has several alternative actions while the IPD game usually has only two actions (i.e., "*cooperate*" and "*defect*").

The main characteristic feature of our knowledge extraction task is that the amount of available information gradually increases during the iterative execution of our game. That is, a piece of available information is successively generated from every round of our game. A fuzzy rule-based system, which is used for game-playing by each player, is also successively adjusted by an on-line incremental learning scheme after every round. Such on-line learning of the fuzzy rule-based system affects the game-playing in future rounds. That is, the learning affects the generation of available information in future rounds. In this sense, our knowledge extraction task is dynamic while most of the above-mentioned studies on linguistic knowledge extraction were applied to static tasks where the available information was given in advance.

This paper is organized as follows. In the next section, we briefly describe our market selection game formulated as a non-cooperative repeated game with many players (e.g., 100 players) and several alternative actions (e.g., five markets) in our former work [7]. Some game strategies examined in [7,10,11] are described in Section 3. In Section 4, we propose an on-line incremental learning scheme of fuzzy rule-based classification systems. In this learning scheme, our market selection game is handled as a pattern classification problem where a single training pattern is successively generated from every round of our game. In Section 5, our market selection game is handled as a function approximation problem. A fuzzy rule-based approximation system is used for approximating the value of each action. Simulation results are reported in Section 6 for demonstrating that comprehensible linguistic knowledge can be extracted by our two learning schemes. Finally Section 7 concludes this paper.

2. Formulation of a market selection game

In this section, we briefly describe a market selection game formulated in our former study [7]. Our market selection game involves n players and m markets. All players and markets are located in a two-dimensional world. Fig. 1 shows an example of our market selection game with 100 players (i.e., $n = 100$) and five markets (i.e., $m = 5$). Locations of players and markets in this figure are available from http://www.ie.osakafu-u.ac.jp/~hisaoi/ci_lab_e/index.html. Every player i ($i = 1, 2, ..., n$) is supposed to simultaneously choose one of the m markets in every round of our game as shown in Fig. 2. Let s_i be the action of the i-th player. The action s_i is to choose one of the given m markets: $s_i \in \{1, 2, ..., m\}$.

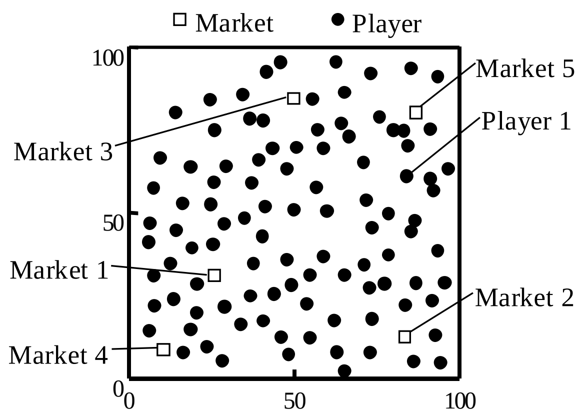


Fig. 1. An example of our market selection game.

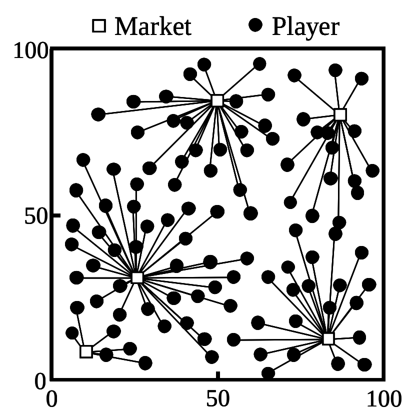


Fig. 2. An example of the market selection by 100 players.

Every player has a single product to be sold in a market in every round of our game. A fixed transportation cost c_{ij} is required for the transportation of the product from the *i*-th player's location to the *j*-th market. In our computer simulations on the market selection game in Fig. 1, we simply defined c_{ij} by the Euclidean distance from the *i*-th player's location to the *j*-th market. The payoff of each player is defined by the market price at the selected market and the transportation cost to that market. It is assumed that the market price at each market is determined by a linear decreasing function of the number of players who choose that market. Let X_j be the number of players who choose the *j*-th market. It should be noted that the equality $X_1 + X_2 + ... + X_m = n$ holds from the definition. The market price p_j at the *j*-th market is defined as

$$p_j = a_j - b_j \cdot X_j \,, \tag{1}$$

where a_j and b_j are positive constants that specify the market price mechanism in the *j*-th market. In our computer simulations on the market selection game in Fig. 1, we used the same linear decreasing function for all the five markets:

$$p_j = 200 - 3X_j \quad \text{for} \quad j = 1, 2, 3, 4, 5. \tag{2}$$

The payoff of the *i*-th player who chooses the market s_i (i.e., the *i*-th player with the action s_i) is defined as

$$r_{i(s_i)} = p_{(s_i)} - c_{i(s_i)} = a_{(s_i)} - b_{(s_i)} \cdot X_{(s_i)} - c_{i(s_i)} \,. \tag{3}$$

In this formulation, $X_{(s_i)}$ is the number of players who choose the market s_i. Thus the payoff of the *i*-th player depends on the actions of all players.

The main characteristic feature of our market selection game is its payoff mechanism in (3). High payoff cannot be obtained from an action that is also chosen by many other players. That is, majority players with respect to their actions receive low payoff. High payoff can be obtained from an action that is chosen by only a small number of players. Such deterioration in payoff due to the concentration of players can be observed in many everyday situations. For example, the chance to pass the entrance examination of a particular department of a university may decrease as the number of applicants to that department increases. Various choices related to plans for the summer vacation also have similar payoff mechanisms. For example, the choice of the same route to a summer resort by many people may decrease their payoff due to heavy traffic jams. The choice of the same resort by many people may also decrease their payoff due to several negative effects such as the hike in travel expenses and the difficulty in booking.

Another important characteristic feature shared by our market selection game and these everyday situations is the dependence of future decision-making on previous results. For example, if the competition to pass the entrance examination of a particular department is unusually tough this year, the number of applicants to that department may decrease next year. If people have a hard time on a highway due to a heavy traffic jam, they will try to avoid that route next time. As shown by these discussions, our market selection game shares some interesting features with many everyday decision-making problems.

3. Some simple strategies

In this section, we briefly explain some simple strategies examined in our previous studies [7,10,11]. We also propose a maximum expected payoff strategy. These strategies will be compared with two on-line incremental learning schemes proposed in later sections.

3.1 Random selection strategy

The point of our market selection game is to avoid the undesired concentration of players to a few markets. The simplest way for avoiding the concentration is to randomly select a market. In this random selection strategy, each market is selected as the action s_i of the i-th player with the probability $1/m$. When all players use this strategy, the undesired concentration is avoided. Fig. 3 is an example of the market selection by the 100 players with the random selection strategy. As shown in Fig. 3, the market selection is totally in disorder. Since the transportation cost is not taken into account, we cannot obtain high payoff from this strategy. We performed computer simulations on our market selection game in Fig. 1 by assigning this strategy to all the 100 players. We performed ten independent trials, each of which consisted of 1000 rounds of our market selection game. The average payoff was 84.8 over the ten trials. This average payoff is used for evaluating the performance of other strategies in this paper.

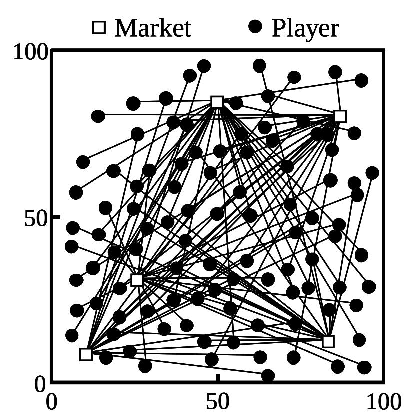


Fig. 3. Example of the market selection using the random selection strategy.

3.2 Minimum transportation cost strategy

While the undesired concentration of players was avoided in Fig. 3, high payoff was not obtained. This is because the transportation cost was not taken into account. The transportation cost can be minimized by choosing the nearest market from each player. The minimum transportation cost strategy always chooses the nearest market in every round of our market selection game. Fig. 2 in the previous section was depicted using this strategy. When all players adopt this strategy, the market selection in Fig. 2 is simply iterated. In this case, the average payoff is 108.0. This average payoff is larger than the result by the random selection strategy (i.e., average payoff 84.8).

3.3 Optimal strategy for the previous actions

Since every player is supposed to simultaneously choose a market in every round of our market selection game, no player knows the best action for the current round. Every player, however, knows the best action in the previous round that has already been completed. The optimal strategy for the previous actions chooses the best action in the previous round as the current action. This strategy first calculates the potential payoff from each market in the previous round. Let r_{ijt} be the actual or potential payoff from the *j*-th market in the *t*-th round for the *i*-th player. When the *j*-th market was actually selected in the *t*-th round by the *i*-th player, r_{ijt} is the actual payoff calculated by (3) in Section 2. For the other markets that were not actually selected, r_{ijt} is calculated by considering how much payoff would have been obtained from the *j*-th market if the *i*-th player had chosen that market in the *t*-th round. That is, the actual or potential payoff is calculated as:

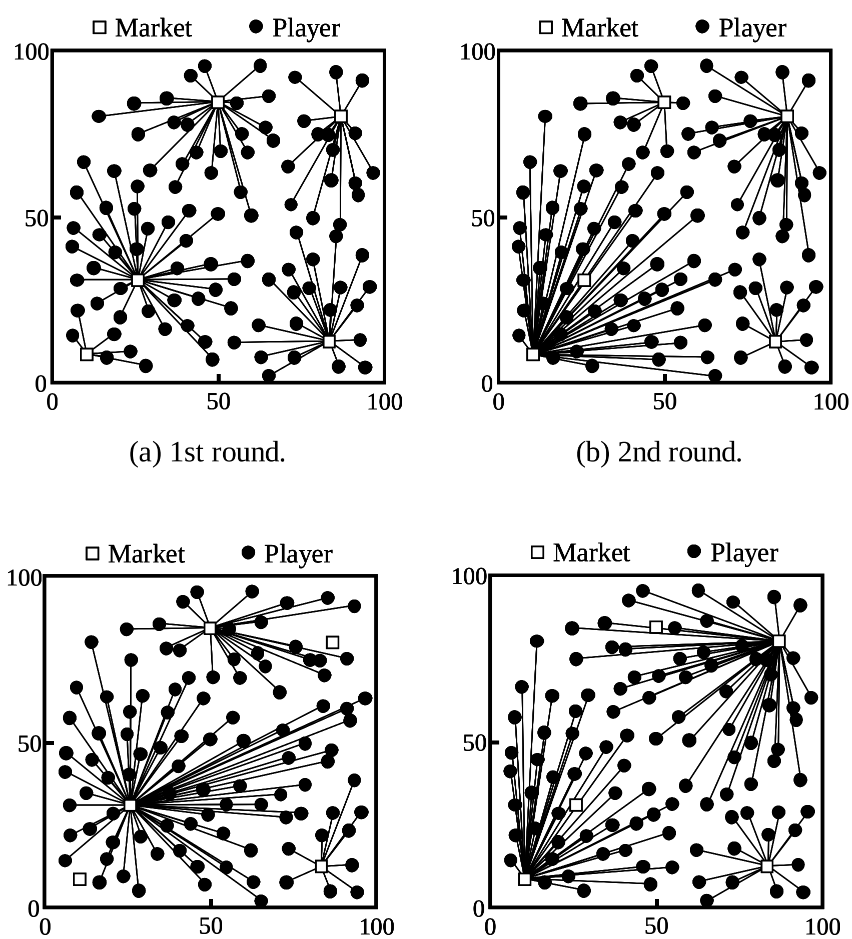
$$r_{ijt} = \begin{cases} a_j - b_j \cdot X_j - c_{ij}, & \text{if } s_i = j \text{ in the } t\text{-th round,} \\ a_j - b_j \cdot (X_j + 1) - c_{ij}, & \text{otherwise,} \end{cases} \tag{4}$$

where X_j is the number of players who actually chose the *j*-th market in the *t*-th round. When the potential payoff is calculated for each market that was not actually chosen by the *i*-th player in the *t*-th round, the *i*-th player is added to X_i as $(X_i + 1)$ in (4).

In the $(t+1)$ th round of our market selection game, the optimal strategy for the previous actions chooses the market with the maximum value of r_{ijt} among the *m* markets. If multiple markets have the same maximum value, this strategy randomly chooses one from those markets. Of course, the selected market is not always the best in the current round. When all the other players do not change their market selection, the optimal strategy for the previous actions is also optimal in the current round. On the other hand, when all players use this strategy, the undesired concentration of players is rapidly self-organized as shown in Fig. 4.

In the first round, no market was selected in the previous round (i.e., $s_i \neq j$ and $X_j = 0$ for all markets in (4)). Thus the optimal strategy for the previous actions chooses the nearest market with the minimum transportation cost (see Fig. 4 (a)) when a_j and b_j are the same for all markets. In the second round, the optimal market of each player is chosen for the actions of the other players in the first round. In Fig. 4 (b), many players choose the left-bottom market with only a few players in the first round. In the third round, the optimal market for the actions in the second round is chosen as shown in Fig. 4 (c) where two markets are selected by no player. In the fourth round, many players choose those two markets (see Fig. 4 (d)). The market selection was iterated until the 1000th round. The average payoff was 47.2 over the 1000 rounds. This average payoff is terribly poor (i.e., much smaller than the result 84.8 by the random selection strategy) due to the undesired concentration of players.

While the average payoff was very small when all players used this strategy, we can obtain the best result when a single player with this strategy plays against all the other players with the minimum transportation cost strategy. We can also obtain good results from the optimal strategy for the previous actions when this strategy is adopted by only a small number of players. Competition among different strategies will be discussed in Section 6.



(a) 1st round. (b) 2nd round.

(c) 3rd round. (d) 4th round.

Fig. 4. Selected markets by the optimal strategy for the previous actions in the first four rounds.

3.4 Maximum expected payoff strategy

In the optimal strategy for the previous actions, only the previous single round was taken into account. Such limited utilization of previous results may cause the synchronized oscillation of the market selection in Fig. 4. In this subsection, the expected payoff from each market is estimated from all the previous results. Let v_{ij} be the value of the *j*-th market (i.e., the expected payoff from the *j*-th market) for the *i*-th player. After the *t*-th round of our market selection game is completed, v_{ij} is updated as

$$v_{ij}^{\text{New}} := (1-\alpha)\cdot v_{ij}^{\text{Old}} + \alpha \cdot r_{ijt} \quad \text{for} \quad j = 1,2,\ldots,m\,, \tag{5}$$

where α is a learning rate ($0 < \alpha \leq 1$), and r_{ijt} is the actual or potential payoff of the *i*-th player from the *j*-th market in the *t*-th round of our market selection game, which is calculated by (4).

In the framework of reinforcement learning [25], only the value of the actually selected action is updated. Such a learning scheme was examined in our previous studies [7,10]. In this paper, we update the values of all markets (i.e., all alternative actions) because the potential payoff from each market can be calculated by (4) even if that market was not actually selected in the previous round.

In every round of our market selection game, the market with the largest value is selected among the *m* markets. When multiple markets have the same largest value, one market is randomly selected from those markets. We use such a simple greedy method because the values of all markets are updated by (5). If the value of only the actually selected market were updated, the greedy method would not work well. In this case (i.e., in the framework of reinforcement learning), some exploration mechanism should be included [25]. Our maximum expected payoff strategy in this subsection always chooses the market with the largest value. Before the first round of our market selection game, the values v_{ij}'s of all markets are specified as the same initial real number. As a result, the market selection is randomly performed in the first round. In our computer simulations, we specified the initial value of each market as $v_{ij}^{\text{Initial}} = 200$, which is the upper bound of the market price (see (2)). The specification of the initial value of each action is very important for facilitating the exploration of actions in reinforcement learning [25]. The effect of the initial value of each action, however, is limited to the first several rounds in our maximum expected payoff strategy when the same initial value is used. When we use random initial values in a wide range (e.g., interval [0, 200]), many rounds are required for adjusting the values v_{ij}'s.

Our maximum expected payoff strategy is illustrated in Fig. 5 where all players adopt this strategy with $\alpha = 0.1$. In the first round, every player randomly chooses a market (see Fig. 5 (a)). When the first round is completed, the value v_{ij} of each market for each player is updated by (5) based on the market selection in the first round. In the second round, every player chooses the market with the largest value for that player (see Fig. 5 (b)). We iterated this computer simulation until the 1000th round. In the tenth round (see Fig. 5 (c)), good coordination of the market selection was realized. Such good coordination of the market selection continued after the tenth round (see Fig. 5 (d)). We performed this computer simulation ten times. The average payoff over ten independent trials was 118.9. This result is better than the average payoff by the other strategies mentioned in the previous subsections.

When $\alpha = 1$, the market selection by the maximum expected payoff strategy is almost the same as the optimal strategy for the previous actions. This is because the value v_{ij} is defined only by the result of the previous single round in the case of $\alpha = 1$ (see the update mechanism in (5)). In this case, we obtained the average payoff 41.2 from ten independent trials. This average payoff is almost the same as the result 47.2 by the optimal strategy for the previous actions. On the other hand, when α is very small, the value v_{ij} cannot be rapidly adjusted to sudden changes of environment.

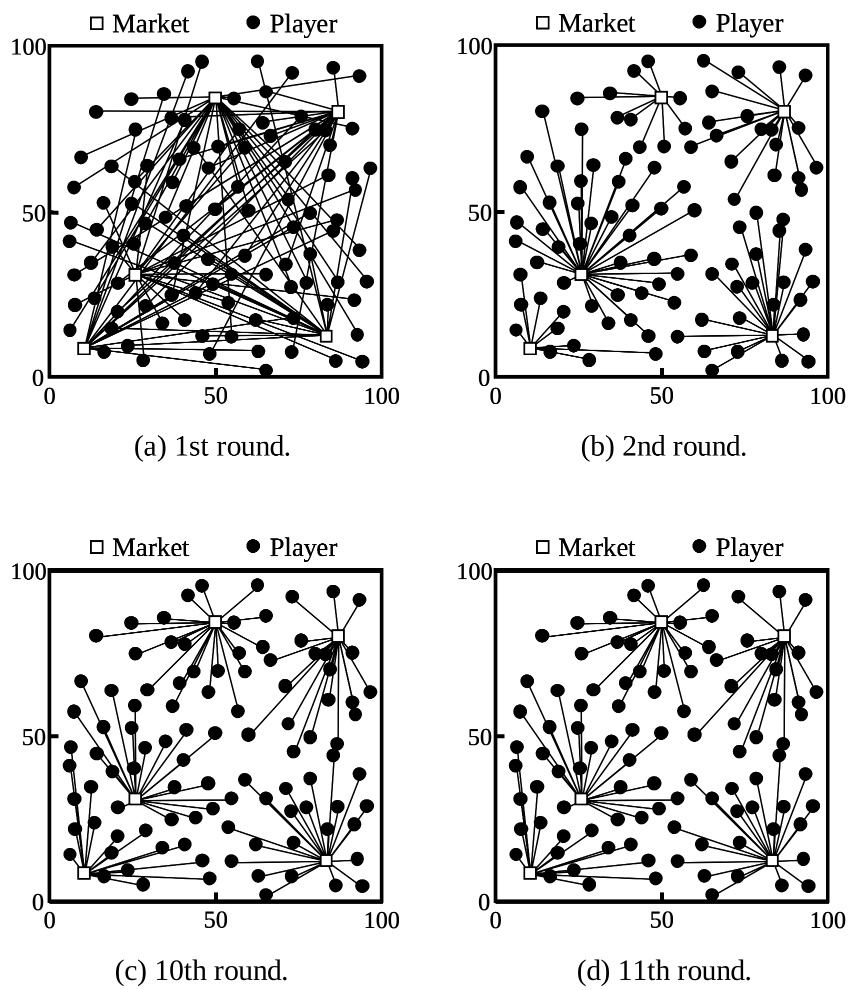


Fig. 5. Selected markets by the maximum expected payoff strategy.

4. Handling as a pattern classification problem

In this section, we show how our market selection game can be handled as a pattern classification problem. An on-line incremental learning scheme is proposed for adjusting fuzzy rule-based classification systems.

4.1 Data acquisition

As we have already explained in the previous section, every player knows the optimal market selection when the current round of our market selection game is completed. It is, however, difficult to effectively utilize the information about the optimal market selection in previous rounds for the market selection in future rounds. As shown in Fig. 4, the optimal strategy for the previous actions leads to the undesired concentration of players to a few markets when all players adopt this strategy. Thus we need a trick for effectively utilizing the information about the optimal market selection in previous rounds.

In this section, we associate the market prices in the *t*-th round to the market selection in the $(t+1)$ th round. That is, we generate a labeled pattern $(\boldsymbol{p}_t, c_{i(t+1)})$ where $c_{i(t+1)}$ is the best market for the ***i***-th player in the $(t+1)$ th round and $\boldsymbol{p}_t$ is the price vector in the *t*-th round:

$$\boldsymbol{p}_t = (p_{t1}, p_{t2}, ..., p_{tm}), \tag{6}$$

where p_{tj} is the market price at the *j*-th market in the *t*-th round of our market selection game. It should be noted that the best market $c_{i(t+1)}$ is identified after the $(t+1)$ th round is completed.

The market selection of each player is performed by a single fuzzy rule-based classification system. The first labeled pattern $(\boldsymbol{p}_1, c_{i2})$ is obtained after the second round is completed. This means that we have no training data until the second round is completed. Thus the market selection in the first two rounds is performed randomly. The first labeled pattern is utilized for the learning of the fuzzy rule-based classification system before the market selection in the third round. When the third round is completed, the second labeled pattern $(\boldsymbol{p}_2, c_{i3})$ is obtained. In this manner, we have $(t-1)$ labeled patterns when the *t*-th round of our market selection game is completed. Those labeled patterns { $(\boldsymbol{p}_1, c_{i2})$, $(\boldsymbol{p}_2, c_{i3})$, ..., $(\boldsymbol{p}_{t-1}, c_{it})$ } can be used in the learning of the fuzzy rule-based classification system for the market selection of the ***i***-th player in the $(t+1)$ th round.

4.2 Fuzzy rule-based classification systems for market selection

For the market selection, we use the following fuzzy if-then rules:

$$\text{Rule } R_k : \text{If } p_1 \text{ is } A_{k1} \text{ and ... and } p_m \text{ is } A_{km} \text{ then } c_k \text{ with } CF_k, \; k = 1,2,...,K, \tag{7}$$

where A_{kj} is an antecedent fuzzy set with a linguistic label, c_k is a consequent market, CF_k is a certainty grade, and K is the number of fuzzy if-then rules. The fuzzy if-then rule R_k in (7) is interpreted as *"If the market prices in the previous round are* $(A_{k1},...,A_{km})$ *then choose the market* c_k *in the current round"*. The certainty grade CF_k is used for representing the weight of the fuzzy if-then rule R_k .

In our computer simulations of this paper, we used two antecedent fuzzy sets "*low*" and "*high*" in Fig. 6 for all the five markets. Thus $2^5 = 32$ fuzzy if-then rules, which were generated by combining these two antecedent fuzzy sets for the five markets, were used for the market selection of each player. That is, the fuzzy rule-based classification system for each player consisted of 32 fuzzy if-then rules in our computer simulations. The membership functions of the two antecedent fuzzy sets were specified in an *ad hoc* manner. When the 100 players are evenly distributed over the five markets, the market price of each market is 140 because 20 players choose each market (see (2) in Section 2). The two membership functions intersect with each other at this market price as shown in Fig. 6. When the number of players choosing a particular market is doubled (i.e., from 20 to 40), the market price decreases from 140 to 80. The membership values of "*high*" and "*low*" are 0 and 1 at this market price, respectively. The upper bound of the market price of each market is 200 (see (2) in Section 2). The membership values of "*high*" and "*low*" are 1 and 0 at this market price, respectively. We also examined a different pair of membership functions with three parameter values (100, 150, 200) instead of (80, 140, 200) in Fig. 6. That is, membership functions with a smaller overlapping area were also examined. Simulation results from these membership functions were almost the same as the case of Fig. 6. We also examined an interval partition where "*low*" and "*high*" were represented by two intervals [0, 140] and [140, 200], respectively. Simulation results by this interval partition were slightly inferior to those by the fuzzy partition in Fig. 6. The automated specification of linguistic labels and their membership functions are beyond the scope of this paper while it is a very important issue. This issue is left for future research on our market selection game. The main difficulty in the handling of this issue for our market selection game is that training data are incrementally obtained from the iterative execution of our game.

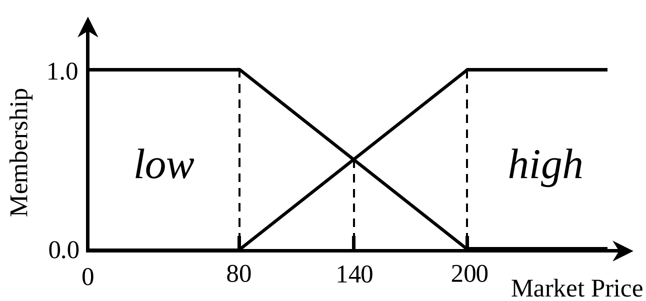


Fig. 6. Two antecedent fuzzy sets "*low*" and "*high*".

For the market selection in the $(t+1)$ th round, the price vector $\boldsymbol{p}_t = (p_{t1}, p_{t2}, ..., p_{tm})$ in the t-th round is used as the input vector to the fuzzy rule-based classification system with the K fuzzy if-then rules in (7). The market selection is performed using a fuzzy reasoning method based on a single winner rule [6,8,9,17]. In this fuzzy reasoning method, first the compatibility of the input vector with each fuzzy if-then rule is calculated as

$$\mu_k(\boldsymbol{p}_t) = \mu_{k1}(p_{t1}) \cdot \mu_{k2}(p_{t2}) \cdot \ ... \ \cdot \mu_{km}(p_{tm}), \tag{8}$$

where $\mu_{kj}(\cdot)$ is the membership function of the antecedent fuzzy set A_{kj}. The winner rule R_{k*} is defined as

$$\mu_{k*}(\boldsymbol{p}_t) \cdot CF_{k*} = \max\{\mu_k(\boldsymbol{p}_t) \cdot CF_k : k = 1,2,...,K\}. \tag{9}$$

The consequent market c_{k*} of the winner rule R_{k*} is chosen for the market selection in the $(t+1)$ th round. When multiple rules have the same maximum value in (9), the winner rule cannot be uniquely specified. In this case, a single rule is randomly chosen from those rules as the winner rule R_{k*} for the market selection in the $(t+1)$ th round.

4.3 Learning algorithm

We have already explained how training patterns can be successively obtained from the iterative execution of our market selection game. We have also explained how the market selection can be performed by the fuzzy rule-based classification system. Our next issue is the learning of the fuzzy rule-based classification system. Our learning algorithm is based on a heuristic procedure [6,8,9,17] for generating fuzzy if-then rules for pattern classification problems.

Let β_{kj} be the discounted sum of the compatibility grades of training patterns labeled as the j-th market (i.e., $c_{it} = j$) with the fuzzy if-then rule R_k. When the t-th round is completed, a training pattern $(\boldsymbol{p}_{t-1}, c_{it})$ is obtained. Using this training pattern, β_{kj} is updated as

$$\beta_{kj}^{\text{New}} = \gamma \cdot \beta_{kj}^{\text{Old}} + \mu_k(\boldsymbol{p}_{t-1}) \cdot \delta_j(c_{it}) \quad \text{for} \quad j = 1,2,...,m, \tag{10}$$

where γ is a kind of a discount rate $(0 \le \gamma \le 1)$ introduced for discounting the effect of the previously obtained training patterns, and $\delta_j(c_{it})$ is the following function for identifying the market c_{it} of the current training pattern $(\boldsymbol{p}_{t-1}, c_{it})$:

$$\delta_j(c_{it}) = \begin{cases} 1, & \text{if } \ c_{it} = j, \\ 0, & \text{if } \ c_{it} \ne j. \end{cases} \tag{11}$$

When $\gamma = 1$, the compatibility grades of the previously obtained training patterns are not discounted in (10). On the other hand, when $\gamma = 0$, β_{kj} is calculated only from the current training pattern $(\boldsymbol{p}_{t-1}, c_{it})$.

The consequent market c_k of the fuzzy if-then rule R_k is determined in the same manner as a heuristic rule generation procedure in our former studies [6,8,9,17]:

$$\beta_{k(c_k)} = \max\{\beta_{kj} : j = 1,2,...,m\}. \tag{12}$$

That is, the consequent market c_k has the maximum discounted sum of the compatibility grades among the m markets. When the consequent market c_k cannot be uniquely determined (i.e., multiple markets have the same maximum value in (12)), we specify c_k as $c_k = \phi$ for indicating that the fuzzy if-then rule R_k is a dummy rule with no effect on the market selection. For the dummy rule R_k, we specify the certainty grade CF_k as $CF_k = 0$. From the definition of the winner rule in (9), we can see that any dummy rule with $CF_k = 0$ is never selected as the winner rule.

When R_k is not a dummy rule, its certainty grade CF_k is defined from β_{kj} and c_k as follows [6,8,9,17]:

$$CF_k = \frac{(\beta_{k(c_k)} - \overline{\beta})}{\sum_{j=1}^{m} \beta_{kj}}, \tag{13}$$

where

$$\overline{\beta} = \sum_{\substack{j=1 \\ j \neq c_k}}^{m} \frac{\beta_{kj}}{(m-1)}. \tag{14}$$

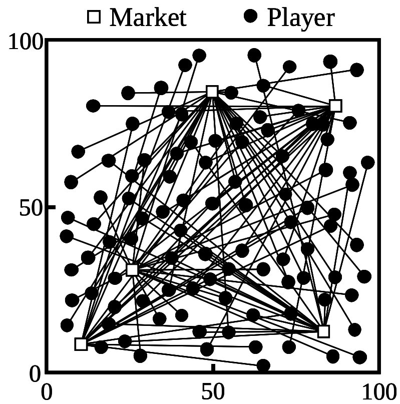
The above heuristic procedure can be easily understood if we consider a two-class pattern classification problem (i.e., market selection with two markets). For example, when $\beta_{k1} > \beta_{k2}$, the consequent market c_k and the certainly grade CF_k are determined as $c_k = 1$ and $CF_k = (\beta_{k1} - \beta_{k2})/(\beta_{k1} + \beta_{k2})$, respectively.

Our fuzzy rule-based classification system is adjusted after every round of our market selection game using the updated β_{kj} in (10). That is, the consequent market and the certainty grade of each fuzzy if-then rule are redefined after every round using the updated β_{kj}. The initial value of β_{kj} is specified as $\beta_{kj} = 0$ because we have no training pattern before the first round. As we have already mentioned, the market selection in the first two rounds is performed randomly. When the second round is completed, each fuzzy if-then rule is adjusted. Then the market selection for the third round is performed using the adjusted fuzzy rule-based classification system.

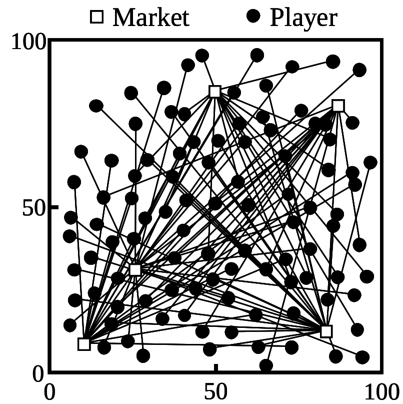
4.4 Computer simulations

We performed computer simulations on the market selection game in Fig. 1 using fuzzy rule-based classification systems with 32 fuzzy if-then rules. Each rule was generated using the two antecedent fuzzy sets "*low*" and "*high*" in Fig. 6. In our computer simulations of this subsection, every player used its own fuzzy rule-based classification system for the market selection. That is, our market selection game was executed by 100 fuzzy rule-based classification systems. The discount rate γ was specified as $\gamma = 0.9$ for all players. The specification of γ is discussed later in this subsection. It will be also discussed in Section 6 in a non-stationary situation with sudden changes of environment.

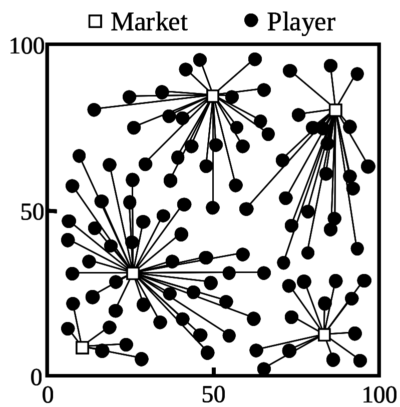
Fig. 7 shows the results in several rounds. As we have already mentioned, the market selection in the first two rounds was randomly performed (see Fig. 7 (a) and Fig. 7 (b)). After the second round, each fuzzy rule-based classification system was adjusted by the first training pattern. As a result, each player used the adjusted fuzzy rule-based classification system in the third round (see Fig. 7 (c)). In some rounds, we observed the undesired concentration of players (see Fig. 7 (d)). Such undesired concentration appeared periodically through 1000 iterations of our market selection game. As in the computer simulations in Section 3, the market selection was iterated until the 1000th round. This computer simulation was performed ten times. The average payoff over ten independent trials was 61.6. This average payoff is not good due to the undesired concentration of players as the optimal strategy for the previous actions (its average payoff was 47.2). These two strategies are based on the information about the optimal market selection in the previous round. While the simulation result was not good when all players adopted the fuzzy rule-based classification strategy, this strategy works very well in competitive situations with other strategies as shown in Section 6. We also performed the same computer simulation using different values of γ. When $\gamma \leq 0.8$, the average payoff was almost the same as the case of the optimal strategy for the previous actions. For example, it was 42.0 when $\gamma = 0.8$. That is, the learning of each fuzzy rule-based system was mainly governed by the currently obtained single training pattern. On the other hand, by increasing the value of γ, we could improve the average payoff. For example, it was 100.6 when $\gamma = 1.0$. In this case, the effect of previously obtained training patterns is not discounted. This leads to poor adaptability of fuzzy rule-based systems to changes of environment. The adaptability of fuzzy rule-based systems will be discussed in Section 6.



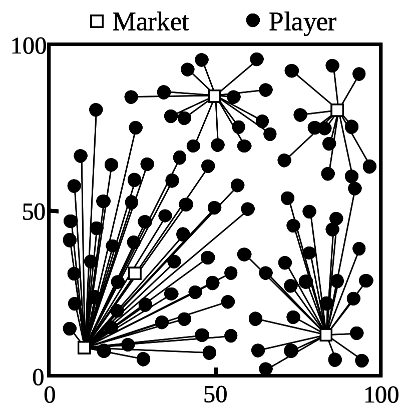
(a) 1st round.



(b) 2nd round.



(c) 3rd round.



(d) 5th round.

Fig. 7. Selected markets by the fuzzy rule-based classification systems.

By monitoring the winner rule in each round during the iterative execution of our market selection game, we can extract a small number of frequently used fuzzy if-then rules for each player. For example, let us consider Player 1 in Fig. 1. We monitored the winner rule for this player in each round of a single trial with 1000 rounds, and counted the number of rounds where each rule was selected as the winner rule. Table 1 shows frequently used fuzzy rules over the 1000 rounds. Since the fuzzy rule-based classification system was successively updated during the execution of our market selection game, we show the final consequent class and the final certainty grade of each rule after the 1000th round. From Table 1, we can see that the market selection of Player 1 was mainly performed by only a few rules. We can also see that all the listed rules have the same final consequent: Market 5 (see Fig. 1). This means that Market 5 was selected independent of the market prices in the previous round. Actually, Player 1 in Fig. 1 almost always chose Market 5 while the market selection of many players was governed by the synchronized oscillation as shown in Fig. 7. Table 1 also suggests that the target market $c_{i(t+1)}$ in the training data set would be almost always Market 5. We monitored the generated training data for Player 1 during the 1000 rounds of our market selection game. The target market was Market 5 in 864 labeled patterns (i.e., 864 rounds) during the execution of our market selection game.

It should be noted that these observations were obtained from a single trial. Simulation results in this subsection strongly depended on the random market selection in the first two rounds. Actually each trial showed totally different behaviors. That is, we obtained different average payoff and different fuzzy rules from each trial. The dependency of simulation results on the random market selection in the first two rounds can be decreased by specifying the initial value of β_{kj} by a random real number instead of the same initial value $\beta_{kj} = 0$. The average payoff was improved from 61.6 to 114.6 when we specified the initial value of β_{kj} by a random real number in the closed interval [0, 1]. In Section 6, we will show different situations where the market selection in each round is determined by the market prices in the previous round.

Table 1. Frequently used fuzzy if-then rules. *"H"* and *"L"* denote the antecedent fuzzy set *"high"* and *"low"*, respectively.

Number of rounds	Antecedent part					Consequent	Certainty
	q_1	q_2	q_3	q_4	q_5	c_k	CF_k
447	*L*	*H*	*H*	*H*	*H*	Market 5	0.95
394	*H*	*H*	*H*	*L*	*H*	Market 5	0.6
77	*H*	*H*	*L*	*L*	*H*	Market 5	0.54
24	*L*	*H*	*H*	*H*	*L*	Market 5	0.97
13	*H*	*L*	*H*	*L*	*H*	Market 5	0.68
11	*H*	*L*	*L*	*L*	*H*	Market 5	0.58
7	*L*	*L*	*H*	*H*	*H*	Market 5	0.95

5. Handling as a function approximation problem

In this section, we show how our market selection game can be handled as a function approximation problem. A fuzzy rule-based approximation system is used for approximating the value of each market for each player.

5.1 Data acquisition

As in the previous section, the market prices are used in the antecedent part of each fuzzy if-then rule. The consequent part is related to the expected payoff from each market (i.e., the value of each market). When the *t*-th round of our market selection game is completed, the actual or potential payoff r_{ijt} of the *i*-th player from the *j*-th market in the *t*-th round is calculated by (4) in Section 3. An *m*-input and single-output fuzzy rule-based approximation system is used for approximating the mapping from the price vector $\boldsymbol{p}_t = (p_{t1}, ..., p_{tm})$ in the *t*-th round to the value v_{ij} of the *j*-th market in the $(t+1)$ th round. Each player uses *m* fuzzy rule-based approximation systems for the market selection.

Each approximation system is used as an approximator of the value of each market.

The first input-output pair $(\boldsymbol{p}_1, r_{ij2})$ is obtained for the *j*-th market when the second round of our market selection game is completed. This input-output pair is used for the learning of the *j*-th fuzzy rule-based approximation system. Such an input-output pair is obtained for each market. That is, a set of *m* input-output pairs is obtained for the learning of *m* fuzzy rule-based approximation systems. Each approximation system is adjusted by the corresponding input-output pair. The market selection in the third round is performed using the adjusted *m* approximation systems. When the third round is completed, the next input-output pair $(\boldsymbol{p}_2, r_{ij3})$ is obtained from the *j*-th market for the learning of the *j*-th fuzzy rule-based approximation system. In this manner, $(t-1)$ input-output pairs $\{(\boldsymbol{p}_1, r_{ij2})$, $(\boldsymbol{p}_2, r_{ij3})$, ..., $(\boldsymbol{p}_{t-1}, r_{ijt})\}$ are obtained for the learning of the *j*-th fuzzy rule-based approximation system before the market selection in the $(t+1)$ th round. As in the previous section, the market selection in the first two rounds is performed randomly.

5.2 Fuzzy rule-based systems for value approximation

For approximating the value v_{ij}, we use the following fuzzy if-then rules:

$$\text{Rule } R_k : \text{If } p_1 \text{ is } A_{k1} \text{ and ... and } p_m \text{ is } A_{km} \text{ then } v_{ij} \text{ is } v_{ijk}, \; k = 1,2,...,K, \tag{15}$$

where v_{ijk} is a consequent real number. These fuzzy if-then rules are the simplest form of the Takagi-Sugeno model [24]. While conventional fuzzy if-then rules in the Takagi-Sugeno model have a consequent linear function, our rules in (15) have a consequent real number. The fuzzy if-then rule R_k in (15) is interpreted as *"If the market prices in the previous round are* $(A_{k1},...,A_{km})$ *then the value of the j-th market in the current round is* v_{ijk} *"*.

We use the same weighted averaging scheme as the Takagi-Sugeno model for calculating the output $\hat{v}_{ij}$ from the fuzzy rule-based approximation system with the *K* fuzzy if-then rules in (15). When the input vector $\boldsymbol{p}_t$ (i.e., the market price vector in the *t*-th round) is presented to the approximation system, the output $\hat{v}_{ij}$ is calculated as follows:

$$\hat{v}_{ij} = \frac{\sum_{k=1}^{K} \mu_k(\boldsymbol{p}_t) \cdot v_{ijk}}{\sum_{k=1}^{K} \mu_k(\boldsymbol{p}_t)} = \sum_{k=1}^{K} \mu_k^*(\boldsymbol{p}_t) \cdot v_{ijk}, \tag{16}$$

where $\mu_k(\boldsymbol{p}_t)$ is the compatibility grade of the input vector $\boldsymbol{p}_t$ with the fuzzy if-then rule R_k defined by (8), and $\mu_k^*(\boldsymbol{p}_t)$ is the normalized compatibility grade:

$$\mu_k^*(\boldsymbol{p}_t) = \frac{\mu_k(\boldsymbol{p}_t)}{\sum_{k=1}^{K} \mu_k(\boldsymbol{p}_t)} . \tag{17}$$

For the market selection in the $(t+1)$ th round, we first calculate the output $\hat{v}_{ij}$ from each fuzzy rule-based approximation system by (16). Then we choose the market with the largest output among the *m* markets. When multiple markets have the same largest output, one market is randomly selected from those markets. The same greedy method was used in the maximum expected payoff strategy in Subsection 3.4.

5.3 Learning algorithm

The learning of each fuzzy rule-based system is performed by updating the consequent real number v_{ijk} of each fuzzy if-then rule. When the *t*-th round is completed, the input-output pair $(\boldsymbol{p}_{t-1}, r_{ijt})$ is obtained from the *j*-th market for the *i*-th player. This input-output pair is used for the learning of the *K* fuzzy if-then rules in the *j*-th fuzzy rule-based approximation system of the *i*-th player.

The learning of each fuzzy if-then rule is performed by updating the consequent real number v_{ijk} using the actual or potential payoff r_{ijt} as

$$\begin{aligned} v_{ijk}^{\text{New}} &= (1 - \alpha \cdot \mu_k^*(\boldsymbol{p}_{t-1})) \cdot v_{ijk}^{\text{Old}} + \alpha \cdot \mu_k^*(\boldsymbol{p}_{t-1}) \cdot r_{ijt} \\ &= (1 - \alpha^*) \cdot v_{ijk}^{\text{Old}} + \alpha^* \cdot r_{ijt} , \end{aligned} \tag{18}$$

where

$$\alpha^* = \alpha \cdot \mu_k^*(\boldsymbol{p}_{t-1}) . \tag{19}$$

This update rule is almost the same as (5) in the maximum expected payoff strategy except that the learning rate α is multiplied by the normalized compatibility grade $\mu_k^*(\boldsymbol{p}_{t-1})$. The amount of the adjustment is proportional to the normalized compatibility grade in (18). Fuzzy if-then rules with small compatibility grades are slightly adjusted while those with large compatibility grades are significantly adjusted. The same update rule was used in fuzzy *Q*-learning for the market selection game in our former studies [7, 10]. While only the fuzzy rule-based system for the actually selected market was adjusted in those studies based on reinforcement learning, all the *m* fuzzy rule-based systems are adjusted in this paper.

The adjustment of each fuzzy if-then rule by (18) tries to decrease the difference between its consequent real number v_{ijk} and the target r_{ijt} (i.e., actual or potential payoff). Thus (18) can be viewed as a kind of local learning [29]. A heuristic method for determining the consequent real number was proposed based on a similar idea to local learning in [18].

It is also possible to adjust the consequent real number for decreasing the difference between the output $\hat{v}_{ij}$ and the target r_{ijt}. In such global learning [29], the following squared error is usually used as an error function to be minimized:

$$E = (r_{ijt} - \hat{v}_{ij})^2 / 2 . \tag{20}$$

An update rule for the consequent real number v_{ijk} is written as

$$v_{ijk}^{\text{New}} = v_{ijk}^{\text{Old}} - \alpha \cdot \frac{\partial E}{\partial v_{ijk}} . \tag{21}$$

From (16), this update rule is rewritten as

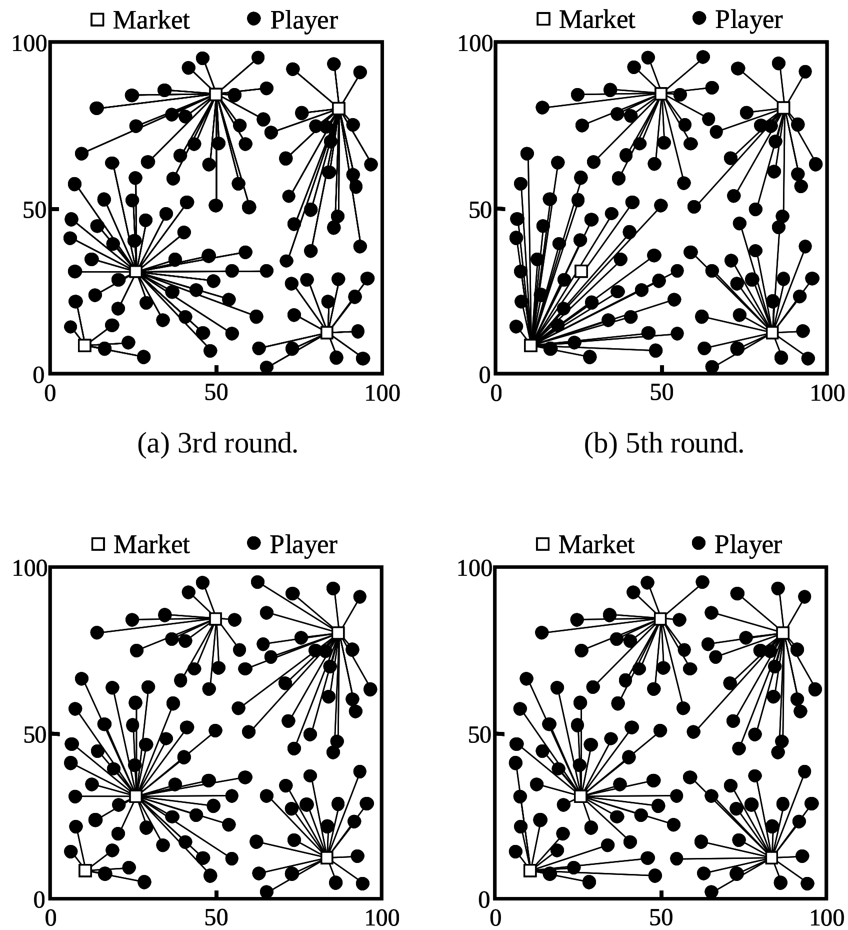
$$\begin{aligned} v_{ijk}^{\text{New}} &= v_{ijk}^{\text{Old}} + \alpha \cdot \mu_k^*(\mathbf{p}_{t-1}) \cdot (r_{ijt} - \hat{v}_{ij}) \\ &= v_{ijk}^{\text{Old}} + \alpha^* \cdot (r_{ijt} - \hat{v}_{ij}) . \end{aligned} \tag{22}$$

In our computer simulations reported in this paper, we used the update rule in (18) based on the concept of local learning. This is because our main aim is to extract comprehensible fuzzy if-then rules. In general, local learning improves the interpretability of fuzzy if-then rules while global learning improve the accuracy of fuzzy rule-based systems (see Yen et al.[29]). We also examined (20)-(22) in some computer simulations. Almost the same results in terms of the average payoff were obtained from these two learning schemes. Of course, trained fuzzy if-then rules were not the same because the consequent real number of each rule was updated in different manners.

5.4 Computer simulation

We performed computer simulations on the market selection game in Fig. 1 using fuzzy rule-based approximation systems with 32 fuzzy if-then rules generated from the two antecedent fuzzy sets "*low*" and "*high*" in Fig. 6. Our market selection game was executed by 100 players with five fuzzy rule-based approximation systems, each of which was used for approximating the value of each market. The learning rate α was specified as $\alpha = 0.1$ for all players. The specification of α will be discussed in Section 6 in non-stationary situations with sudden changes of environment. The consequent real numbers of all fuzzy if-then rules were specified as $v_{ijk} = 200$ before the learning.

Fig. 8 shows the results in several rounds. Since the first input-output pair was obtained after the second round, the market selection in the first two rounds was randomly performed as in the case of the fuzzy rule-based classification strategy (see Fig. 7 (a)-(b) in Section 4). As in the market selection by fuzzy rule-based classification systems in Section 4, each fuzzy rule-based approximation system was adjusted after the second round. In the third round, the market selection was performed by the adjusted fuzzy rule-based system (see Fig. 8 (a)). Good coordination of the market selection was gradually realized as shown in Fig. 8 (b)-(d). The average payoff over ten independent trials with 1000 rounds was 118.8. This average payoff is almost the same as the results by the maximum expected payoff strategy (i.e., average payoff 118.9).



(a) 3rd round. (b) 5th round.

(c) 10th round. (d) 15th round.

Fig. 8. Selected markets by the fuzzy rule-based approximation systems.

As shown in the fuzzy reasoning method in (16), the effect of each fuzzy if-then rule R_k on the calculation of the output $\hat{v}_{ij}$ from the fuzzy rule-based approximation system is proportional to the normalized compatibility grade $\mu_k^*(\boldsymbol{p}_t)$. Thus we can find some influential fuzzy if-then rules for each player by monitoring the normalized compatibility grade of each rule. As in the previous section, let us consider Player 1 in Fig. 1. We monitored the normalized compatibility grade $\mu_k^*(\boldsymbol{p}_t)$ of each fuzzy if-then rule for this player and calculated its sum over 1000 rounds in a single trial. Table 2 shows some influential fuzzy if-then rules. Five fuzzy if-then rules with the same antecedent conditions for the five markets are shown in a single row in this table. The consequent real number of each fuzzy if-then rule in this table is its final value after the 1000th round. From this table, we can see that all fuzzy if-then rules in this table have almost the same consequent real numbers for each market. This means that the value of each market did not depend on the previous market prices in our computer simulations in this subsection. We monitored the marker price of each market during the 1000 rounds. Except for the first 14 rounds, the market price of each market was the same during the 1000 rounds because no player changed its action after the 14th round. Thus the target output for each market was the same after the 14th round. As a result, the consequent real numbers for each market were almost the same in the seven fuzzy if-then rules in Table 2. While Table 2 was obtained from a single trial, almost the same results were obtained from other trials with different random market selection in the first two rounds. Contrary to the case of fuzzy rule-based classification systems, simulation results by fuzzy rule-based approximation systems did not strongly depend on the market selection in the first two rounds. In the next section, we will show different situations where the value of each market is strongly affected by the previous market prices.

Table 2. Influential fuzzy if-then rules for estimating the value of each market.

Sum of compatibility	Antecedent part					Consequent part				
	q_1	q_2	q_3	q_4	q_5	v_1	v_2	v_3	v_4	v_5
41.4	*L*	*L*	*L*	*H*	*L*	70.6	87.3	94.3	58.1	121.8
41.3	*L*	*L*	*L*	*H*	*H*	70.6	87.3	94.3	58.1	121.8
37.7	*L*	*L*	*H*	*H*	*L*	71.5	88.1	95.0	59.1	122.4
37.7	*L*	*L*	*H*	*H*	*H*	71.5	88.1	95.1	59.1	122.4
37.6	*L*	*H*	*L*	*H*	*L*	71.6	88.1	95.1	59.1	122.4
37.6	*H*	*L*	*L*	*H*	*H*	71.6	88.1	95.1	59.1	122.4
37.6	*L*	*H*	*L*	*H*	*H*	71.6	88.1	95.1	59.1	122.4

6. Illustration of linguistic knowledge extraction

In this section, simulation results on various situations are reported for examining the performance of each strategy. First, competition between two strategies is examined. Then, competition among several strategies is examined. Finally, a non-stationary situation with a sudden change of environment is examined. Through computer simulations, we demonstrate that comprehensible fuzzy if-then rules are extracted from our two learning schemes.

6.1 Competition between two strategies

We have already examined the performance of each strategy by computer simulations where a single strategy was adopted by all the 100 players. Table 3 summarizes the average payoff obtained from each strategy in such computer simulations. In this table, good results were obtained from the maximum expected payoff strategy and the fuzzy rule-based approximation strategy. The common feature shared by these strategies is the accumulation of available information in the form of the value (i.e., expected payoff) of each market using on-line incremental learning schemes. For comparison, we used the optimal strategy for the previous actions for finding an equilibrium state of our market selection game. This strategy was applied to a randomly selected single player after each round by changing its market selection. That is, only a randomly selected single player could change its market selection after each round. In this manner, we iterated our market selection game until an equilibrium state was found. We performed this computer simulation 100 times from random initial market selection. Each of all the 100 trials found an equilibrium state where no player changed its action any more. In such an equilibrium state, the market selection of every player was optimal for the actions of all the other players. Among these computer simulations, 27 different equilibrium states were found. The average payoff over the 100 players in those equilibrium states was between 119.0 and 119.1. We also examined our market selection game in the framework of cooperative games. The best average payoff was 119.2 when the average payoff was maximized by genetic algorithms (see [11]). From the comparison of Table 3 with these results, we can see that the results by the maximum expected payoff strategy and the fuzzy rule-based approximation strategy are very good.

Table 3. Average payoff from each strategy when all the players used the same strategy.

Strategy	Average payoff
Random selection	84.8
Minimum transportation cost	108.0
Optimal for previous actions	47.2
Maximum expected payoff	118.9
Fuzzy classification	61.6
Fuzzy approximation	118.8

As we have already mentioned, the performance of each strategy strongly depends on strategies adopted by other players. We examined the performance of each strategy against other strategies by computer simulations where a single player used one strategy and the other 99 players used another strategy. That is, a player with one strategy played against the other 99 players with another strategy.

This performance examination was executed for all combinations of the six strategies. For each combination, our computer simulation was performed 100 times so that all players were selected as a minority player. That is, when the performance of Strategy A was examined against Strategy B, first Player 1 with Strategy A played against the other 99 players with Strategy B for 1000 rounds. Next Player 2 with Strategy A played against the other 99 players with strategy B. In this manner, the performance of Strategy A against Strategy B was evaluated by calculating the average payoff obtained by Strategy A over 100 trials with 1000 rounds. Such evaluation was performed for all combinations of the six strategies. Simulation results are summarized in Table 4. From this table, we can see that the performance of the optimal strategy for the previous actions strongly depends on the strategy of the other 99 players. When the other 99 players also used this strategy (see Table 3), the average payoff was very small (i.e., 47.2). This strategy, however, can play very well against the other strategies. Actually, high average payoff was obtained from this strategy in Table 4 when the other 99 players adopted another strategy. We can observe similar characteristic features in the simulation results by the fuzzy rule-based classification strategy in Table 4

Table 4. Average payoff of a single player with a minority strategy when it played against the other 99 players with a majority strategy.

Strategy of a single player	Strategy of the other 99 players					
	Random	Cost	Optimal	Payoff	Fuzzy C.	Fuzzy A.
Random selection	(84.8)	84.4	84.9	84.7	85.1	84.7
Minimum cost	118.1	(108.0)	121.9	118.2	117.5	118.3
Optimal for previous	115.3	131.8	(47.2)	118.9	104.8	119.1
Maximum payoff	117.9	131.8	114.9	(118.9)	121.8	119.0
Fuzzy classification	117.9	131.7	164.2	118.7	(61.6)	118.7
Fuzzy approximation	118.0	131.7	164.2	119.3	120.4	(118.8)

In Table 4, we can also see that exceptionally high average payoff was obtained from the two fuzzy rule-based strategies when they played against the other 99 players with the optimal strategy for the previous actions. As we have already shown in Fig. 4 of Subsection 3.3, the synchronized oscillation of selected markets was self-organized by the optimal strategy for the previous actions. If such oscillation is recognized during the iterative execution of our market selection game, a player can enjoy high payoff by choosing a market that will not be selected by the other 99 players with the optimal strategy for previous actions. Very high payoff was obtained from the two fuzzy rule-based strategies because they learned such market selection knowledge from the game-playing against the other 99 players with the optimal strategy for the previous actions.

We monitored the winner rule in the fuzzy rule-based classification system for Player 1 in the same manner as in Section 4. Table 5 shows frequently used fuzzy if-then rules when Player 1 with the fuzzy rule-based classification strategy played against the other 99 players with the optimal strategy for the previous actions. From the fuzzy if-then rules in Table 5, we can see that the fuzzy rule-based classification system learned the market selection knowledge that can make use of the synchronized oscillation of selected markets. The trained fuzzy rule-based classification system chooses Market 3 when the previous market prices at Market 3 and Market 5 were *low* and *high* respectively. On the other hand, when they were *high* and *low* respectively, Market 5 is chosen. That is, the fuzzy rule-based classification system learned to choose a market with a lower previous market price between Market 3 and Market 5. We can also find two patterns of the market price vectors (*low, high, low, high, high*) and (*high, low, high, low, low*) from the antecedent parts of the most frequently used two fuzzy if-then rules in Table 5.

Table 5. Frequently used fuzzy if-then rules by Player 1 during the game-playing against the other 99 players with the optimal strategy for the previous actions.

Number of rounds	Antecedent part					Consequent	Certainty
	q_1	q_2	q_3	q_4	q_5	c_k	CF_k
499	*L*	*H*	*L*	*H*	*H*	Market 3	1
494	*H*	*L*	*H*	*L*	*L*	Market 5	1
4	*H*	*H*	*H*	*L*	*L*	Market 5	1

We also monitored the fuzzy reasoning process in each fuzzy rule-based approximation system for Player 1 when Player 1 with the fuzzy rule-based approximation strategy played against the other 99 players with the optimal strategy for the previous actions. In the same manner as in Section 5, we show influential fuzzy if-then rules in Table 6. As in Table 5, the two patterns of the market price vector can be observed from the most influential fuzzy if-then rules in Table 6. We can also see that the fuzzy rule-based approximation systems learned the market selection knowledge that can make use of the synchronized oscillation of selected markets. That is, each fuzzy if-then rule in Table 6 says that high payoff will be obtained in the current round from markets where the previous market prices were *low*. Player 1 with the trained fuzzy rule-based approximation systems in Table 6 chooses a market with a lower previous market price between Market 3 and Market 5. As shown in Table 5 and Table 6, the two fuzzy rule-based strategies learned the same market selection knowledge through different learning schemes.

Table 6. Influential fuzzy if-then rules of Player 1 during the game-playing against the other 99 players with the optimal strategy for the previous actions.

Sum of compatibility	Antecedent part					Consequent part				
	q_1	q_2	q_3	q_4	q_5	v_1	v_2	v_3	v_4	v_5
485.340	*L*	*H*	*L*	*H*	*H*	131.5	64.5	152.6	-1.2	75.6
338.720	*H*	*L*	*H*	*L*	*L*	-45.5	148.5	38.6	106.8	174.6
147.334	*H*	*H*	*H*	*L*	*L*	-45.5	148.5	38.6	106.8	174.6
12.434	*L*	*H*	*L*	*H*	*L*	151.1	103.5	166.1	56.2	111.2
8.716	*H*	*L*	*L*	*L*	*L*	57.1	169.9	106.1	145.7	185.2

6.2 Competition among several strategies

We also examined the competition among all the six strategies. The performance of each strategy was examined by 200 independent trials with 1000 rounds. In each trial, one of the six strategies was randomly assigned to each player. Thus each strategy was used by 16 or 17 players on the average. During each trial, each player continued to use the randomly assigned strategy through 1000 rounds. Average payoff obtained by each strategy over 200 trials is summarized in Table 7.

Table 7. Average payoff from each strategy when the six strategies were randomly assigned to the 100 players.

Strategy	Average payoff
Random selection	84.5
Minimum transportation cost	118.1
Optimal for previous actions	115.9
Maximum expected payoff	118.7
Fuzzy classification	118.8
Fuzzy approximation	118.6

Good results were obtained from the five strategies except for the random selection strategy. This means that the undesired concentration of players to a few markets was avoided. We examined the two fuzzy rule-based strategies by focusing our attention on Player 1. That is, we assigned one of the two fuzzy rule-based strategies to Player 1. The six strategies were randomly assigned to the other 99 players. In the same manner as in the previous subsection, we monitored the winner rule in the fuzzy rule-based classification strategy and the fuzzy reasoning process in the fuzzy rule-based approximation strategy. Final fuzzy if-then rules after the 1000th round are summarized in Table 8 and Table 9. In these tables, all the trained fuzzy if-then rules insisted that Market 5 should be chosen by Player 1 independent of the previous market prices. These learning results suggest that the price at each market would not change so much during the 1000 rounds of our market selection game. We monitored the price at each market during the 1000 rounds. The market price did not change so much. Thus the best market for each player was almost the same during the execution of our market selection game.

Table 8. Frequently used fuzzy if-then rules by Player 1 in the competition among the six strategies.

Number of rounds	Antecedent part					Consequent	Certainty
	q_1	q_2	q_3	q_4	q_5	c_k	CF_k
216	*L*	*L*	*L*	*H*	*H*	Market 5	1
164	*L*	*H*	*L*	*H*	*H*	Market 5	1
94	*L*	*H*	*L*	*H*	*L*	Market 5	1
64	*L*	*L*	*H*	*H*	*H*	Market 5	1
62	*H*	*L*	*L*	*H*	*H*	Market 5	1
61	*L*	*L*	*L*	*H*	*L*	Market 5	1
42	*L*	*L*	*H*	*H*	*L*	Market 5	1

Table 9. Influential fuzzy if-then rules of Player 1 in the competition among the six strategies.

Sum of compatibility	Antecedent part					Consequent part				
	q_1	q_2	q_3	q_4	q_5	v_1	v_2	v_3	v_4	v_5
48.1	*L*	*L*	*L*	*H*	*H*	68.4	87.0	90.7	58.8	122.7
45.1	*L*	*L*	*L*	*H*	*L*	68.7	87.4	91.0	59.3	122.9
44.3	*L*	*H*	*L*	*H*	*H*	68.9	87.4	91.0	59.4	123.0
41.9	*L*	*H*	*L*	*H*	*L*	69.2	87.9	91.3	59.9	123.3
40.1	*H*	*L*	*L*	*H*	*H*	68.4	88.1	92.0	61.3	123.5
39.5	*L*	*L*	*H*	*H*	*H*	70.3	88.3	91.7	60.3	123.5
38.0	*H*	*L*	*L*	*H*	*L*	68.8	88.6	92.5	61.9	123.9

6.3 Adaptation to sudden changes of environment

In the previous computer simulations, the market selection game involved no drastic changes of environment. In this subsection, we examine the adaptability of each strategy through computer simulations on a non-stationary situation with a sudden change of the strategy of the other players. In our computer simulations of this subsection, one of the six strategies was assigned to a single player and another strategy was assigned to the other 99 players. We suddenly changed the strategy of the other 99 players after the 500th round from the minimum transportation cost strategy to the optimal strategy for the previous actions. That is, a single minority player played against the other 99 players with the minimum transportation cost strategy in the first 500 rounds, and it played against those with the optimal strategy for the previous actions in the last 500 rounds. As in Subsection 6.1, the performance of each minority strategy adopted by a single player was examined by computer simulations of 100 trials. In each trial, the examined strategy was assigned to a different player. Over those 100 trials for examining a particular strategy, we calculated the average payoff obtained by the examined minority strategy in each round. Simulation results are summarized in Fig. 9.

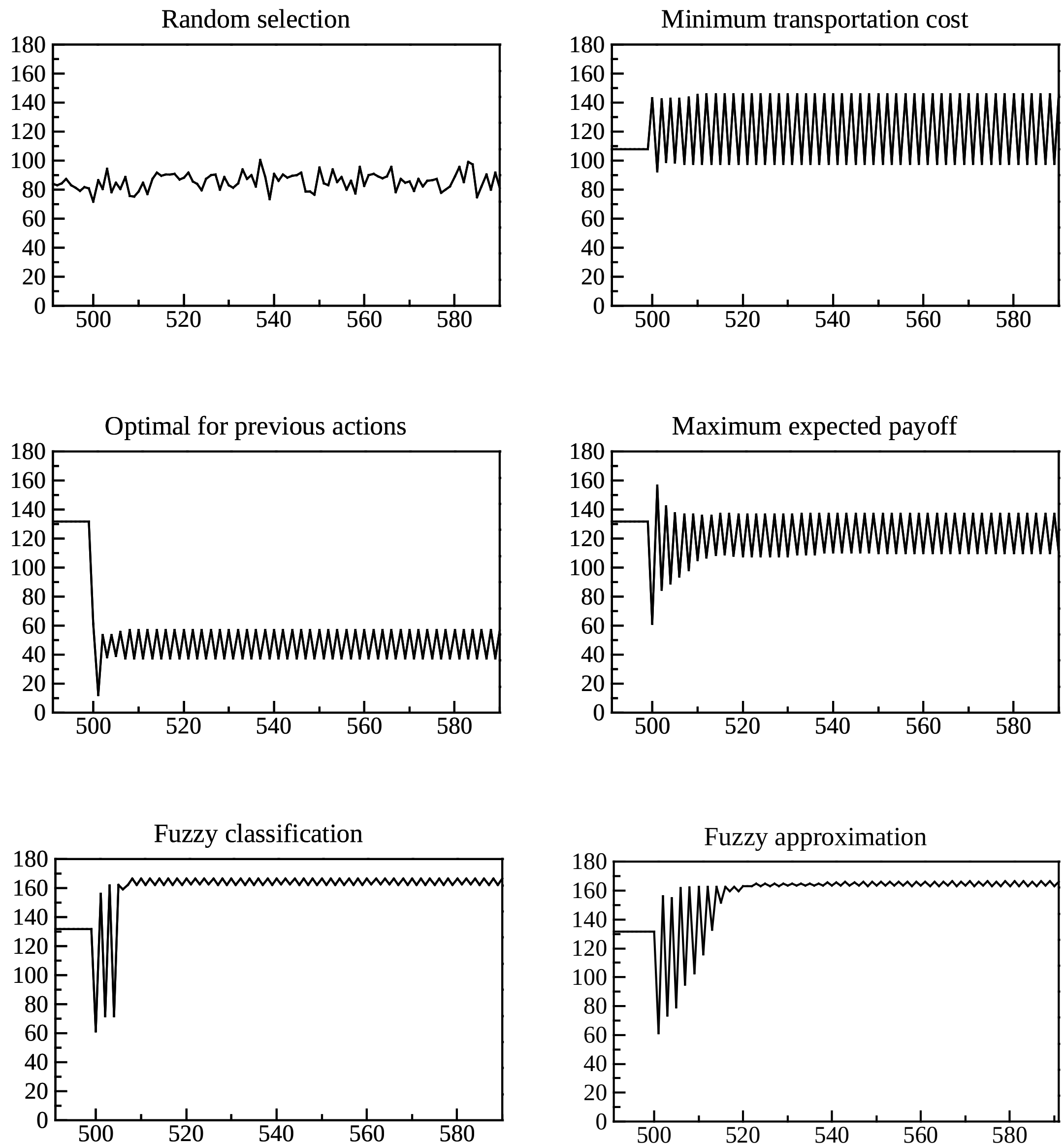


Fig. 9. Average payoff from each strategy adopted by a single player in the non-stationary situation where the strategy of the other 99 players was suddenly changed.

From this figure, we can see that the two fuzzy rule-based strategies could rapidly adapt to the new situation. As in Subsection 6.1, we examined the fuzzy rule-based classification system used for the market selection of Player 1 when the fuzzy rule-based classification strategy was assigned to this player. The final fuzzy rule-based classification system after the 1000th round was almost the same as Table 5 in Subsection 6.1. We also examined the five fuzzy rule-based approximation systems for Player 1 when the fuzzy rule-based approximation strategy was assigned to this player. The final five fuzzy rule-based approximation systems were almost the same as Table 6 in Subsection 6.1. These results suggest that the two fuzzy rule-based strategies could find correct market selection knowledge for the new situation.

We also examined the effect of parameter specifications on the adaptability of the two fuzzy rule-based strategies by computer simulations on the same non-stationary situation. In the fuzzy rule-based classification strategy, we used the discount rate γ for discounting the effect of previous rounds. The value of γ was 0.9 in Fig. 9. When $\gamma = 1$, all the previous rounds have the same effect on the determination of fuzzy if-then rules. On the contrary, $\gamma = 0$ means that the effect of the previous rounds is not accumulated. In this case, fuzzy if-then rules are determined only by the result of the previous single round that has just been completed. That is, the fuzzy rule-based classification strategy with $\gamma = 0$ is almost the same as the optimal strategy for the previous actions. We examined four specifications of γ: $\gamma = 0.4$, 0.8, 0.99, 1.0. Simulation results are summarized in Fig. 10. This figure shows the average payoff at each round over 100 trials. In each trial, a different player adopted the fuzzy rule-based classification strategy. In Fig. 10, this strategy could not rapidly adapt to the new situation when the discount rate was too large (i.e., 0.99 and 1.0).

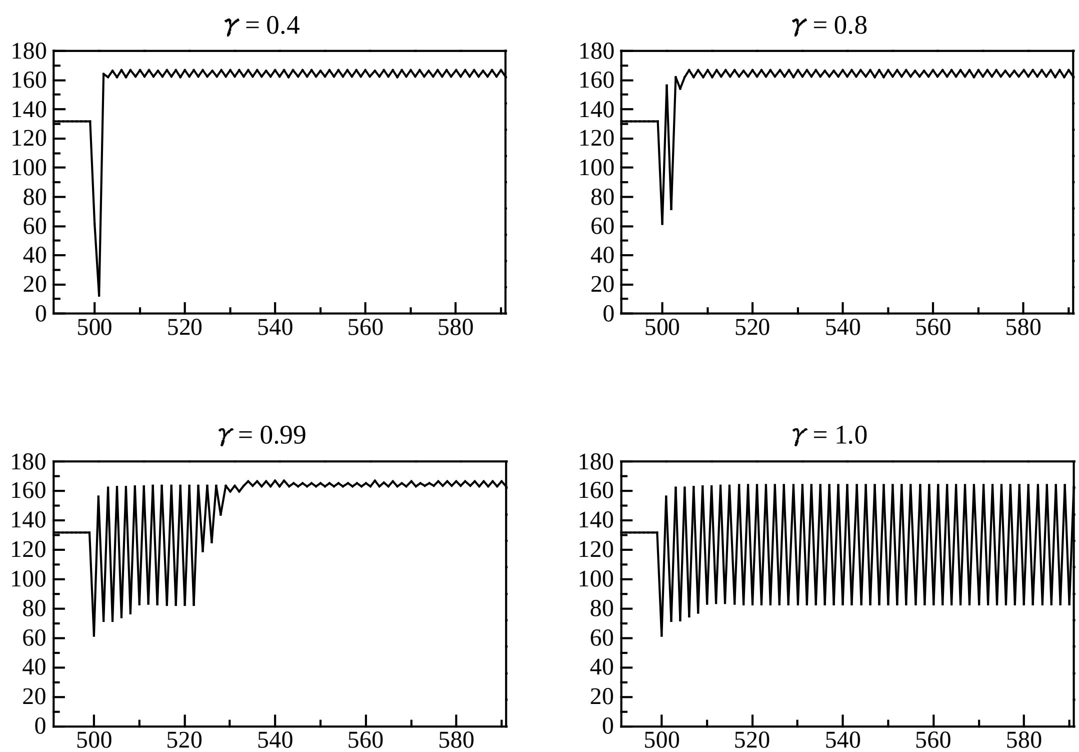


Fig. 10. Effect of the discount rate γ on the adaptability of the fuzzy rule-based classification strategy.

In the fuzzy rule-based approximation strategy, we used the learning rate α for updating the consequent real number of each fuzzy if-then rule. In Fig. 9, the value of α was 0.1. When the learning rate is large, the consequent real number is mainly determined by the result of the previous single round. On the other hand, when the learning rate is very small, all the previous results have a significant effect on the determination of the consequent real number. We examined four specifications of the learning rate α: $\alpha = 0.001$, 0.01, 0.1, 1.0. Simulation results are summarized in Fig. 11 in the same manner as in Fig. 10. When the learning rate was too small (i.e., 0.001 and 0.01), the fuzzy rule-based approximation strategy could not rapidly adapt to the new situation.

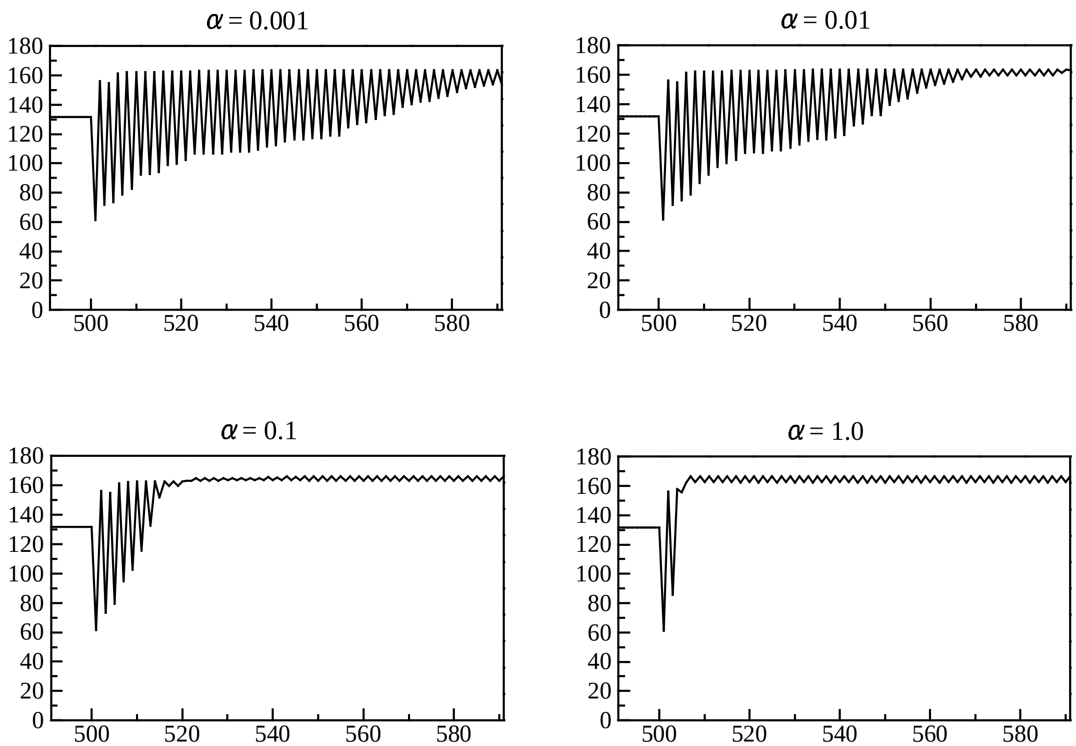


Fig. 11. Effect of the learning rate α on the adaptability of the fuzzy rule-based approximation strategy.

7. Conclusion

In this paper, we illustrated how linguistic knowledge can be extracted from the iterative execution of our market selection game. Linguistic knowledge for game-playing was extracted in the form of fuzzy if-then rules. Our linguistic knowledge extraction was based on the learning of fuzzy rule-based systems. We proposed two on-line incremental learning schemes of fuzzy rule-based systems for the linguistic knowledge extraction. In one scheme, fuzzy rule-based classification systems were used as decision-making systems. In the other scheme, fuzzy rule-based systems were used for approximating the value of each action (i.e., expected payoff from each market). Through computer simulations, we demonstrated that high payoff can be obtained from these two learning schemes. We also showed that our two learning schemes can predict the synchronized oscillation of selected markets by other players. We showed that linguistic knowledge with respect to such prediction can be obtained by extracting frequently used rules and influential rules. We also demonstrated that our two learning schemes can rapidly adapt to sudden changes of environment.

As we mentioned in Section 2, our market selection game has a special payoff mechanism: High payoff can be obtained from actions that are not selected by many players. This payoff mechanism is a general characteristic feature of many everyday decision-making problems. Thus our learning schemes may be employed for analyzing or simulating various decision-making problems. High comprehensibility of extracted knowledge is one advantage of our learning schemes. In this paper, we used very primitive tricks for extracting a small number of important rules. That is, we picked up frequently used rules in fuzzy rule-based classification systems or influential rules in fuzzy rule-based approximation systems. Combination of more sophisticated rule extraction methods involving rule selection and input selection with our learning schemes is left for future research. Utilization of more efficient learning methods including the automated generation of membership functions is also left for future research. Competition against more complicated strategies will be examined using our market selection game with other environmental changes in future research.

References

[1] R. Axelrod, *The Evolution of Cooperation*, Basic Books, New York, 1984.

[2] R. Axelrod, The evolution of strategies in the Iterated Prisoner's Dilemma, in: L. Davis (ed.), *Genetic Algorithms and Simulated Annealing*, Morgan Kaufmann, Los Altos, 1987, pp. 32-41.

[3] O. Cordon, M. J. del Jesus, and F. Herrera, Genetic learning of fuzzy rule-based classification systems cooperating with fuzzy reasoning methods, *International Journal of Intelligent Systems* 13 (1998) 1025-1053.

[4] O. Cordon, M. J. del Jesus, and F. Herrera, A proposal on reasoning methods in fuzzy rule-based classification systems, *International Journal of Approximate Reasoning* 20 (1999) 21-45.

[5] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, Knowledge discovery and data mining: Toward a unifying framework, *Proc. of 2nd International Conference on Knowledge Discovery and Data Mining* (August 2-4, 1996, Portland) 82-88.

[6] H. Ishibuchi, T. Murata, and I. B. Turksen, Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems, *Fuzzy Sets and Systems* 89 (1997) 135-149.

[7] H. Ishibuchi, T. Nakashima, H. Miyamoto, and C. H. Oh, Fuzzy *Q*-learning for a multi-player

non-cooperative repeated game, *Proc. of 6th IEEE International Conference on Fuzzy Systems* (July 1-5, 1997, Barcelona) 1573-1579.

[8] H. Ishibuchi, T. Nakashima, and T. Murata, Performance evaluation of fuzzy classifier systems for multi-dimensional pattern classification problems, *IEEE Trans. on SMC - Part B: Cybernetics* 29 (1999) 601-618.

[9] H. Ishibuchi, K. Nozaki, and H. Tanaka, Distributed representation of fuzzy rules and its application to pattern classification, *Fuzzy Sets and Systems* 52 (1992) 21-32.

[10] H. Ishibuchi, C. H. Oh, and T. Nakashima, Competition between strategies for a market selection game, *COMPLEXITY INTERNATIONAL* (On-line Journal: http://www.csu.edu.au/ci/) 6 (1999).

[11] H. Ishibuchi, R. Sakamoto, and T. Nakashima, Evolution of unplanned coordination in a market selection game, *IEEE Trans. on Evolutionary Computation* 5 (2001) 000-000.

[12] Y. Jin, Fuzzy modeling of high-dimensional systems: Complexity reduction and interpretability improvement, *IEEE Transactions on Fuzzy Systems* 8 (2000) 212-221.

[13] Y. Jin, W. von Seelen, and B. Sendhoff, On generating FC^3 fuzzy rule systems from data using evolution strategies, *IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernetics* 29 (1999) 829-845.

[14] B. Kosko, Fuzzy systems as universal approximators, *Proc. of 1st IEEE International Conference on Fuzzy Systems* (March 8-12, 1992, San Diego) 1153-1162.

[15] K. Lindgren, Evolution phenomena in simple dynamics, in: C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen (eds.), *Artificial Life II*, Addison-Wesley, Reading, 1991, pp. 295-312.

[16] D. Nauck and R. Kruse, A neuro-fuzzy method to learn fuzzy classification rules from data, *Fuzzy Sets and Systems* 89 (1997) 277-288.

[17] K. Nozaki, H. Ishibuchi, and H. Tanaka, Adaptive fuzzy rule-based classification systems, *IEEE Trans. on Fuzzy Systems* 4 (1996) 238-250.

[18] K. Nozaki, H. Ishibuchi, and H. Tanaka, A simple but powerful heuristic method for generating fuzzy rules from numerical data, *Fuzzy Sets and Systems* 86 (1997) 251-270.

[19] V. de Oliveira, Semantic constraints for membership function optimization, *IEEE Trans. on Systems, Man, and Cybernetics - Part A: Systems and Humans* 29 (1999) 128-138.

[20] W. Pedrycz and V. de Oliveira, Optimization of fuzzy models, *IEEE Trans. on Systems, Man, and Cybernetics - Part B: Cybernetics* 26 (1996) 627-637.

[21] G. Piatetsky-Shapiro, R. Brachman, T. Khabaza, W. Kloesgen, and E. Simoudis, An overview of issues in developing industrial data mining and knowledge discovery applications, *Proc. of 2nd International Conference on Knowledge Discovery and Data Mining* (August 2-4, 1996, Portland) 89-95.

[22] M. Setnes, R. Babuska, U. Kaymak, and H. R. van Nauta Lemke, Similarity measures in fuzzy rule base simplification, *IEEE Trans. on Systems, Man, and Cybernetics - Part B: Cybernetics* 28 (1998) 376-366.

[23] M. Setnes, R. Babuska, and B. Verbruggen, Rule-based modeling: Precision and transparency, *IEEE Trans. on Systems, Man, and Cybernetics - Part C: Applications and Reviews* 28 (1998) 165-169.

[24] T. Takagi and M. Sugeno, Fuzzy identification of systems and its applications to modeling and control, *IEEE Trans. on Systems, Man, and Cybernetics* 15 (1985) 116-132.

[25] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, 1998.

[26] L. Wang, Fuzzy systems are universal approximators, *Proc. of 1st IEEE International Conference on Fuzzy Systems* (March 8-12, 1992, San Diego, CA) 1163-1170.

[27] X. Yao and P. Darwen, An experimental study of *N*-person Iterated Prisoner's Dilemma games, *Informatica* 18 (1994) 435-450.

[28] J. Yen and L. Wang, Simplifying fuzzy rule-based models using orthogonal transformation methods, *IEEE Trans. on Systems, Man, and Cybernetics - Part B: Cybernetics* 29 (1999) 13-24.

[29] J. Yen, L. Wang, and C. W. Gillespie, Improving the interpretability of TSK fuzzy models by combining global learning and local learning, *IEEE Trans. on Fuzzy Systems* 6 (1998) 530-537.