Paper

Artificial adaptive agent model characterized by learning and fairness in the ultimatum games

Tomohiro Hayashida, Ichiro Nishizaki, and Hideki Katagiri

Abstract—This paper examines the result of the experimental research on the ultimatum games through simulation analysis. To do so, we develop agent-based simulation system imitating the behavior of human subjects in the laboratory experiment by implementing a learning mechanism involving a concept of fairness. In our agent-based simulation system, mechanisms of decision making and learning are constructed on the basis of neural networks and genetic algorithms.

Keywords—artificial adaptive agents, simulation, games, behavior of players.

1. Introduction

In this paper, we develop a multi-agent simulation system for analyzing behavior of players in the ultimatum games. In the subgame perfect equilibrium of the ultimatum game, player *1* who is a proposer obtains almost all the payoff which is divided between players *1* and *2*, and player *2* accepts the offer of player *1*. It is known from the results of the experimental investigations of the past that the subgame perfect equilibrium does not accurately forecast the ultimatum play, and the payoff is divided almost equally. A diversified range of experiments have been accumulated in order to examine why outcomes of the games deviate from the subgame perfect equilibrium [4, 8, 11, 13, 14, 15, 17, 20, 21], where the following issues are focused on: fairness of players, the number of rounds of the game, difference in nations or races, the right to be player *1*, the structural power of player *1*, anonymity of play, punishment for unfair proposals, magnitude of payoff, and so forth.

Bolton [2] tries to explain the experimental results by using a utility function of a player which is influenced not only by a payoff of the player but also by a payoff of the opponent; the utility function is defined by the payoff of the player and the ratio of the payoff of the player to that of the opponent. Moreover, Bolton and Ockenfels [3] extend this model to games with incomplete information. Rabin [18] define a fairness equilibrium by using a utility function of the payoff of self and the kindness to the opponent, and consider some economic examples. Fehr and Schmidt [7] consider fairness, competition and cooperation in the economic environment by using a utility function defined by the payoff of self and a difference between the payoff of self and that of the opponent. Costa-Gomes and Zauner [5] attempt to explain the experimental data of Roth *et al.* [20] by a utility function with the payoffs of two players and a random disturbance term.

Concerning approaches without any utility function, Roth and Erev [19] propose a simple learning model based on reinforcement learning. Gale *et al.* [9] show that replicator dynamics leads not to the subgame perfect equilibrium, but to the Nash equilibria; they suggest that researchers should give attention to not always the subgame perfect equilibrium but also the Nash equilibria in evaluating the experimental data. Incorporating the quantal response equilibria (QRE) model [16], Yi [22] attempts to explain the experimental result of the ultimatum games.

Abbink *et al.* [1] compare an approach based on the utility function with an approach based on adaptive learning; they argue the abilities and limitations of both approaches. From these research results, it seems to be desirable to incorporate both concepts of fairness and learning for modeling the behavior of players in the ultimatum game. In this paper, we develop a simulation system with artificial adaptive agents which have a decision making and learning mechanism based on neural networks (e.g., [12]) and a genetic algorithm (e.g., [10]). By employing the utility function proposed by Fehr and Schmidt [7] as a fitness function of the genetic algorithm, fairness is incorporated in the learning mechanism of the artificial agents. In our system for simulation of the ultimatum games, an action of an agent is determined by a vector of outputs from a nonlinear function with several input data that agents can know after playing a stage game; this decision mechanism is implemented by a neural network. The artificial agents with chromosomes consisting of the synaptic weights and thresholds characterizing the neural network are evolved so as to obtain larger payoffs through a genetic algorithm, and then this learning mechanism develops agents with better performance.

To imitate the behavior of human subjects in a laboratory experiment and examine the result of the experiment by using the agent-based simulation system, we use the data from the experiment by Roth *et al.* [20], and identify the standard set of the parameters in the utility function incorporating fairness by Fehr and Schmidt [7]. Moreover, by varying the values of the parameters, we evaluate the effect of each individual parameter on the behavior of the artificial agents.

The organization of this paper is as follows. In Section 2, we describe the ultimatum game and briefly review the experimental result of the ultimatum game by Roth *et al.* [20]. Section 3 is devoted to describing the agent-based simulation system with the learning mechanism and the utility function incorporating fairness. In Section 4, we examine the results of the simulations; finally in Section 5, we give a summary of the simulations and some concluding remarks.

2. The ultimatum game

We deal with an ultimatum game in which two players divide $10. In this game, player *1* who is the first mover makes an offer $(x_1, x_2 = 10 - x_1)$, $x_2 \in \{1, \ldots, 10\}$, and player *2* who is the second mover accepts or rejects the offer (x_1, x_2); if player *2* accepts, players *1* and *2* obtain $\pi_1 = x_1$ and $\pi_2 = x_2$, respectively; otherwise they obtain nothing, $\pi_1 = \pi_2 = 0$. It is noted that the offer of 0 by player *1* is removed from a list of possible offers; this setting is also used in the ultimatum game in [6]. In Fig. 1, a game tree of the ultimatum game is depicted. Because player *2*'s payoff by acceptance is larger than that by rejection in any of player *2*'s nodes, the subgame perfect equilibrium is player *1*'s offer of $(\$9, \$1)$ and acceptance of player *2*; the pair of the equilibrium payoffs is $(\$9, \$1)$.

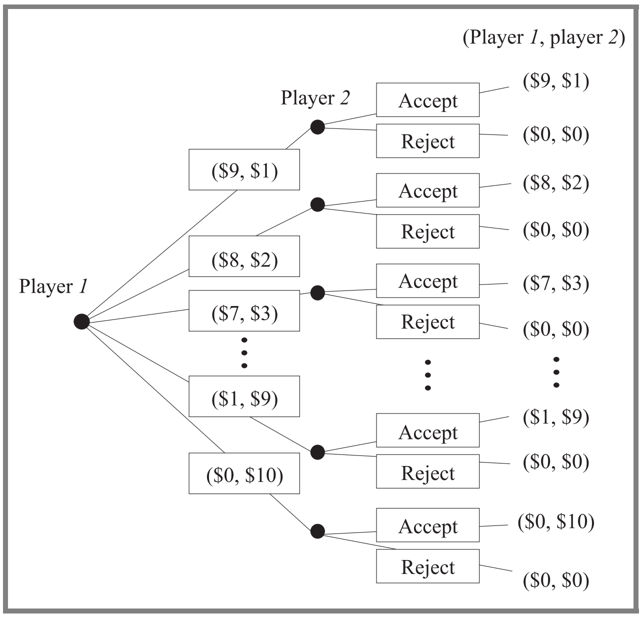


Fig. 1. Game tree of the ultimatum game.

We review and summarize the result of the experiment by Roth *et al*. [20], where the experiment about bargaining and market behavior is conducted in four countries: Israel, Japan, the United State, and Yugoslavia. As a practical matter, in the experiment $10 is represented as 1000 tokens, and all offers are made in multiples of 5 tokens. There are three sessions of the ultimatum game; in each session, about 20 subjects are recruited and the game is played 10 rounds; a pair of players are randomly matched at each round. Because any offer is available only in increments of $1 in our simulation, we use the discretized data of the experiment shown in [5] and we take the average after pooling all the data of the four countries. The result of the experiment compatible with the setting of our simulation is shown in Fig. 2, where the data of $1 correspond to offers from 0 token to 150 tokens, the data of $2 correspond to offers from 155 tokens to 250 tokens, and so forth, because the minimum offer is set to $1 in our setting.

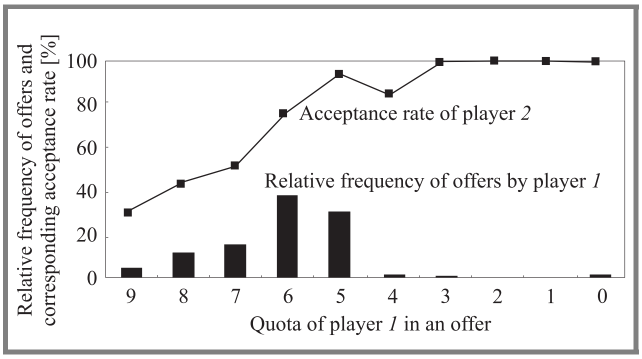


Fig. 2. Summary of the result of the experiment with human subjects.

As can be seen in Fig. 2, only about 4% of all the offers correspond to the subgame perfect equilibrium and the rate of acceptance for the corresponding offers is 31%. This fact is not consistent with the subgame perfect equilibrium prediction. The offers of $4 or $5 account for 68% of all the offers, and therefore player *1* seems to be making relatively fair offers. The rate of acceptance by player *2* for offers larger than or equal to $4 is over 75%.
Although any offers by player *1* a quota of player *2* of which is smaller than or equal to $1 brings a positive payoff to player *2*, player *2* rejects it at the rate of 70%. This behavior can be interpreted as punishment for unfair proposals by player *1*. By using utility functions defined by not only the payoff of self but also the payoff of the opponent, explanations of such behavior have been attempted [2, 3, 5, 7, 18].

3. Agent-based simulation model

In this paper, because it is supposed that human behavior is adaptive, we employ a simulation model which is a natural framework to implement the adaptive behavior of individuals. In our simulation model, each agent has a decision making mechanism built by a neural network (e.g., [12]) and a learning mechanism based on a genetic algorithm (e.g., [10]).

3.1. Decision making by a neural network

An agent corresponds to a neural network which is characterized by synaptic weights between two nodes in the neural network and thresholds which are parameters for the output function of nodes. Because a structure of neural networks is determined by the number of layers and the number of nodes in each layer, an agent is prescribed by the fixed number of parameters if these numbers are fixed. Forming a string consisting of these parameters which is identified with an artificial agent, we think of the string as a chromosome of the agent in an artificial genetic system of our simulation model. In our simulation model for analyzing behavior of players in the ultimatum games, two types of agents are required. The first one which is called agent *1* corresponds to player *1* who makes an offer to player *2*; the second one which is called agent *2* corresponds to player *2* who accepts or rejects the offer by player *1*.

3.1.1. Decision making of agent *1*

The structure of a neural network of agent *1* is depicted in the diagram of Fig. 3a. Agent *1* makes an offer corresponding to the largest output among all the ten outputs of the neural network, where the outputs out_s, $s = 1, \ldots, 10$ correspond to from the offer $(9,1)$ to the offer $(0,10)$; the offer $(10-s^*, s^*)$ with the largest output out_{s^*} is chosen as the next offer of agent *1*.
Inputs of the neural network for agent *1* is summarized as follows.

[Input 1] an offer by agent *1* in the last game: $x_1 \in \{9,8,\ldots,0\}$[1] .

[Input 2] a payoff obtained by agent *1* in the last game: $\pi_1 \in \{9,8,\ldots,0\}$.

[Input 3] agent *2*'s choice between acceptance and rejection in the last game: $y_2 \in \{0,1\}$; 0 means rejection, and 1 means acceptance.

[Input 4] a payoff obtained by agent *2* in the last game: $\pi_2 \in \{1,2,\ldots,10\}$.

[Input 5] the average payoff obtained by agent *1* in the past all games: $\bar{\pi}_1 \in [0,9]$.

[Input 6] the average payoff obtained by agent *2* in the past all games: $\bar{\pi}_2 \in [0,10]$.

[Input 7] the average offer of agent *1* in the past all games: $\bar{x}_1 \in [0,9]$.

[Input 8] the average rate of acceptance by agent *2* for the offers in the past all games: $\bar{y}_2 \in [0,1]$.

It should be noted that the average payoff $\bar{\pi}_1$ and the average offer $\bar{x}_1$ are memorized and updated by agent *1*, and similarly the average payoff $\bar{\pi}_2$ and the average rate of acceptance $\bar{y}_2$ are memorized and updated by agent *2*.

3.1.2. Decision making of agent *2*

Agent *2* also makes a decision in a way similar to agent *1*; 8 inputs of all the 9 inputs of the neural network for agent *2* are the same as those of agent *1*, and the other one is an offer by agent *1* in the current game. The output layer of the neural network of agent *2* consists of two nodes which correspond to the choices of acceptance and rejection for the offer. The structure of the neural network of agent *2* is depicted in the diagram of Fig. 3b. Inputs of the neural network for agent *2* are summarized as follows.

[Inputs 1–8] the same as the inputs of agent *1*.

[Input 9] an offer by agent *1* in the current game: $\hat{x}_1 \in \{9,8,\ldots,0\}$.

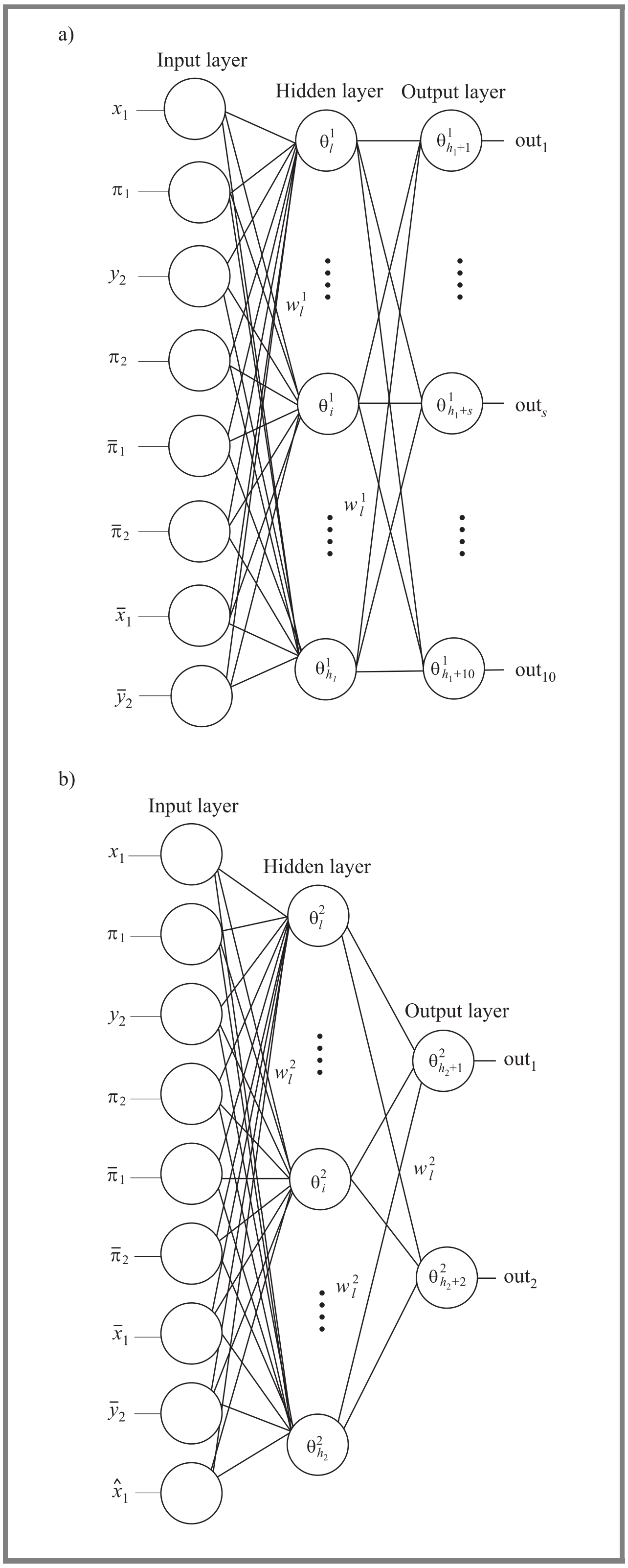


Fig. 3. Neural networks for (a) agent *1* and (b) agent *2*.

[1]Although an offer is represented by a pair (x_1, x_2) in the previous section, the offer (x_1, x_2) is identified only with x_1 because $x_2 = 1 - x_1$.

3.2. Utility function incorporating fairness

For modeling the behavior of players in the ultimatum game, we consider that it is appropriate to incorporate both learning and fairness in the artificial adaptive agent model simultaneously. To implement the concept of fairness, we use utility function as a fitness function in the genetic algorithm which is the basis of the learning mechanism of our agent-based simulation system. As a utility function appropriate for this purpose, we employ the utility function proposed by Fehr and Schmidt [7]; the parameters of the function seems to be easy to interpret because the function is linear, and the excess of the payoff of the opponent over that of self and the reciprocal excess are separated.

When players i and j obtain payoffs π_i and π_j, respectively, the utility u_i of player i is represented as

$$u_i(\pi_i,\pi_j) = \pi_i - \alpha_i \max\{\pi_j - \pi_i, 0\} - \beta_i \max\{\pi_i - \pi_j, 0\},\quad i,j = 1,2,\ i \neq j, \tag{1}$$

where α_i and β_i are coefficients; the utility $u_i(\pi_i,\pi_j)$ of player i consists of the payoff of self, the penalty for the excess of the payoff of the opponent over that of self, and the penalty for the reciprocal excess. When the payoff of self exceeds that of the opponent, i.e., $\pi_i > \pi_j$, $u_i(\pi_i,\pi_j) = \pi_i - \beta_i(\pi_i - \pi_j)$; when the payoff of the opponent exceeds that of self, i.e., $\pi_j > \pi_i$, $u_i(\pi_i,\pi_j) = \pi_i - \alpha_i(\pi_j - \pi_i)$.

3.3. Evolutionary learning through the genetic algorithm

An agent is prescribed by the fixed number of parameters in our agent-based simulation system. Forming a string consisting of these parameters, we use the string as a chromosome in an artificial genetic system. As we mentioned above, for agent *1*, there are the 8 units in the input layer and the 10 units in the output layer. Let h_1 be the number of units in the hidden layer. Then because the number of links between nodes is $18h_1$ and the number of units in the hidden and the output layers is $h_1 + 10$, the neural network corresponding to an agent can be governed by the synaptic weights w_l^1, $l = 1,\dots,18h_1$ and the thresholds θ_l^1, $l = 1,\dots,h_1+10$. Similarly, for agent *2*, the neural network is also governed by the synaptic weights w_l^2, $l = 1,\dots,11h_2$ and the thresholds θ_l^2, $l = 1,\dots,h_2+2$, where h_1 and h_2 are the numbers of nodes in the hidden layers. These parameters and the input values determine an action of the agent, and the synaptic weights and the thresholds are adjusted through the genetic algorithm so that the initial population evolves into the population of agents obtaining larger payoffs.

We separately arrange two subpopulations of agents *1* and *2*; there are N agents in each subpopulation. One agent is selected from each subpopulation, and two agents make a pair for playing the game. Agents repeatedly paly the ultimatum game, and accumulate the payoffs obtained in each stage game. Because the value of the utility Eq. (1) is directly used as a fitness in the artificial genetic system, agents obtaining larger utilities are likely to survive.

We start by describing how the parameters prescribing an agent are initialized. In the experiment, because experimenters explain a procedure of the ultimatum game, subjects should understand the payoff structure of the game. Thus, it is not appropriate that artificial agents start to play the game without any prior knowledge of the game; we give the artificial agents some knowledge of the game before playing it. We implement this by adjusting the parameters of the neural network which are the synaptic weights and the thresholds through the error back propagation algorithm (e.g., [12]) with the teacher signals.

A chromosome of agent *1* consists of the synaptic weights w_l^1, $l = 1,\dots,18h_1$ and the thresholds θ_l^1, $l = 1,\dots,h_1+10$, and that of agent *2* consists of the synaptic weights w_l^2, $l = 1,\dots,11h_2$ and the thresholds θ_l^2, $l = 1,\dots,h_2+2$. Initial values of the parameters w_l^i and θ_l^i are set to random values in $[-1,1]$ before the adjustment by the error back propagation algorithm.

We give teacher signals to the neural network for agent *1* so as to make offers yielding larger payoffs of self. Because the outputs $\text{out}_1^1, \text{out}_2^1, \dots, \text{out}_{10}^1$ of the neural network for agent *1* correspond to the offers $(9,1),(8,2),\dots,(0,10)$, the teacher signals of $1, 8/9, \dots, 0$ are given to the outputs $\text{out}_1^1, \text{out}_2^1, \dots, \text{out}_{10}^1$ for any set of the inputs given at random.

For agent *2*, based on the experimental results, the parameters of the neural network is adjusted by using the error back propagation algorithm such that the possibility of ac-

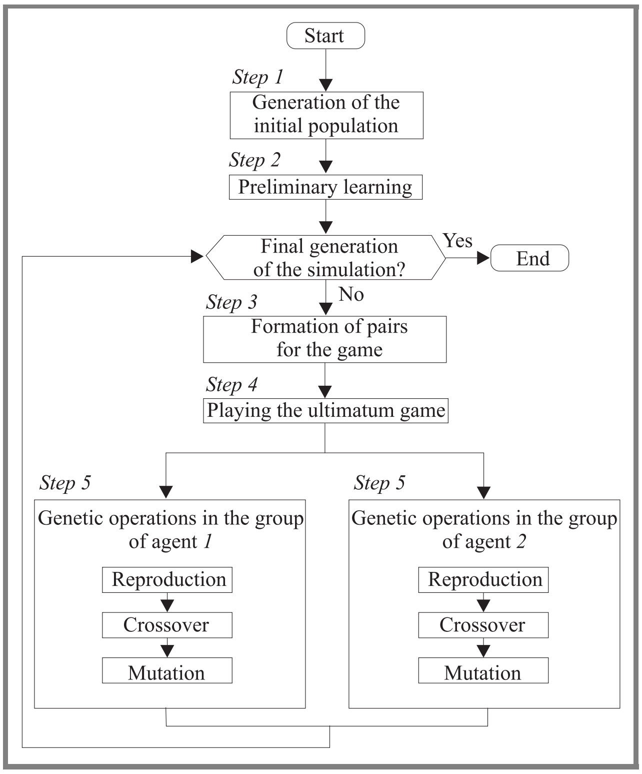


Fig. 4. Flowchart of the agent-based simulation model.

ceptance is equal to that of rejection for the most unfair offer $(10, 0)$, the possibility of acceptance increases as the payoff of agent *2* in an offer becomes larger, and agent *2* perfectly accepts the profitable offers with agent *2*'s quotas larger than 6. To be more precise, when input 9 corresponding to agent *1*'s quota in an offer is $\hat{x}_1$ and inputs 1 to 8 are randomly given, the teacher signal, $\frac{1}{12}(10-\hat{x}_1)+0.5$, is given to the output out_1^2 corresponding to acceptance of an offer, and the teacher signal to the output out_2^2 is a complement of $\frac{1}{12}(10-\hat{x}_1)+0.5$ on 1.

We arrange 30 sets of the teacher signals for each of agents *1* and *2*, and the parameters of the neural networks are adjusted by the error back propagation algorithm. In our agent-based simulation system, there are two subpopulations of N agents for agents *1* and *2*, and one agent from the subpopulation of agent *1* and one from the subpopulation of agent *2* are randomly chosen, and then one pair for playing the game is formed. The game is played by N pairs of agents *1* and *2* according to the above mentioned decision making mechanism. The utilities of agents *1* and *2* are determined by Eq. (1) depending on the outcome of the game; these utilities are directly used as the fitness in the genetic algorithm. Because it is known that the algorithm works effectively by enlargement and reduction of the values of the fitness [10], in our system the fitness is linearly scaled. A procedure of the simulation model is summarized in the following and is diagrammatically shown in Fig. 4.

Step 1: **Generation of the initial population**. For each of agents *1* and *2*, N individuals are generated.

Step 2: **Preliminary learning**. By using the error back propagation algorithm with given teacher signals, the parameters of the neural network for each individual are adjusted.

Step 3: **Formation of pairs for the game**. A pair for playing the game is formed by selecting one agent from each of the subpopulations of agents *1* and *2*; by repeating this operation, N pairs are formed.

Step 4: **Playing the ultimatum game**. In each of the N pairs, the ultimatum game is played; the decision of each agent is determined by the outputs of the neural network; and artificial agents obtain their utilities depending on an outcome (π_1, π_2) of the game.

Step 5: **Genetic operation**. The two subpopulations for agents *1* and *2* are formed again by gathering the same type of artificial agents from the N pairs for playing the game; the genetic operations are executed to each subpopulation consisting of N individuals. The utility Eq. (1) of agent *1* or *2* is directly used as the fitness of an artificial agent in the genetic algorithm, and the fitness is linearly scaled.
If the number of periods reaches a given final generation of the simulation, the procedure stops.

Step 5-1: **Reproduction**. As a reproduction operator, the roulette wheel selection is adopted. By a roulette wheel with slots sized by the probability

$$p_{ij}^s = \frac{f_{ij}}{\sum_{j=1}^N f_{ij}}, i = 1, 2, \qquad (2)$$

each chromosome is selected into the next generation, where f_{ij} is the fitness of the jth individual of agent i.

Step 5-2: **Crossover**. A single-point crossover operator is applied to any pair of chromosomes with the probability of crossover p^c. Namely, a point of crossover on the chromosomes is randomly selected and then two new chromosomes are created by swapping subchromosomes which are the right side parts of the selected point of crossover on the original chromosomes.
For offsprings of agent *1*, the average payoff and the average offer are given by averaging those of the parents with the probabilities corresponding to the sizes of the swapped subchromosomes; similarly, for agent *2*, the average payoff and the average rate of acceptance are calculated.

Step 5-3: **Mutation**. With a given small probability of mutation p^m, each gene which represents a synaptic weight or a threshold in a chromosome is randomly changed. The selected gene is replaced by a random number in $[-1, 1]$.

4. Results of the simulations

We develop the artificial agents in this agent-based simulation system so as to imitate the behavior of human subjects in a laboratory experiment by Roth *et al.* [20], and examine the result of the experimental research through the simulation analysis. First, we identify the standard set of the four parameters α_1, β_1, α_2, β_2 in the utility function (1) by minimizing the error of mean square between the result of the simulation and that of the experiment. After identified the the standard set of the parameters, we examine effect of the parameters characterizing the behavior of the human subjects.

Artificial adaptive agents have mechanisms of decision making and learning based on a neural network and a genetic algorithm, and the parameters of the neural network and the genetic algorithm are set to the following values:

– the number of nodes in the neural network for agent *1*:
8 in the input layer, 10 in the hidden layer, 10 in the output layer;

– the number of nodes in the neural network for agent *2*:
9 in the input layer, 11 in the hidden layer, 2 in the output layer;

– the size of subpopulations for agents *1* and *2*: $N = 100$;
– the maximal generation of the genetic algorithms: $MaxGen = 3000$;
– the parameters of genetic operations: crossover $p^c = 0.5$, mutation $p^m = 0.001$, generation gap $g = 0.8$.

In this paper, the simulation system is executed 100 runs for each setting of the parameters. Because all preparatory runs converge at certain level until 2500 periods, we set the maximal generation of the simulation to 3000 periods. Numerical data of the simulation are given by averaging each observed value in the last 150 generations of the 100 runs.

4.1. Identification of the standard set of the parameters

By varying values of the parameters, we find the standard set of the parameters approximating the behavior of human subjects. There are four parameters α_1, β_1, α_2 and β_2 in the utility function, and especially, the parameter β_2 is the penalty coefficient for the excess of the payoff of player *2*'s self over the payoff of the opponent in the utility function of player *2*. When player *1* makes an offer such that the payoff of player *2* is larger than the payoff of player *1*'s self, i.e., (x_1, x_2), $x_1 < x_2$, this penalty is valid. Although it is true that such an offer is unfair, it is not natural that player *2* is penalized for accepting the offer. From this reason, fixing the value of β_2 at $\beta_2 = 0$, the values of α_1 and β_1 are varied from 0 to 1 at intervals of 0.1, and the value of α_2 is varied from 0 to 2.

In order to find the standard set of the parameters imitating the behavior of human subjects and successfully approximating the result of the experiment, we use the error of mean square which is represented by

$$E(\alpha_1,\beta_1,\alpha_2,\beta_2) = \sum_{x_1=9}^{0} (p_{x_1}^{\text{sim}} - p_{x_1}^{\text{sub}})^2 + \sum_{x_1=9}^{0} (q_{x_1}^{\text{sim}} - q_{x_1}^{\text{sub}})^2, \tag{3}$$

where $p_{x_1}^{\text{sub}}$ and $p_{x_1}^{\text{sim}}$ are the fraction of the human subjects and the artificial agents making an offer $x_1 \in \{9, 8, \ldots, 0\}$ in the experiment and in the simulation, respectively; $q_{x_1}^{\text{sub}}$ and $q_{x_1}^{\text{sim}}$ are the fraction of the human subjects and the artificial agents accepting the offer x_1 in the experiment and in the simulation, respectively.

By executing 100 runs for all the 2541 cases of parameter variations, it is found that the standard set of the parameters is $(\alpha_1, \beta_1, \alpha_2, \beta_2) = (0.4, 0, 1.1, 0)$ minimizing the error of mean square Eq. (3) and the minimum is $E(\alpha_1, \beta_1, \alpha_2, \beta_2) = 0.1263$; at the standard set of the parameters $(\alpha_1, \beta_1, \alpha_2, \beta_2) = (0.4, 0, 1.1, 0)$, the distribution of offers by agent *1* and the rate of acceptance for any offer are given in Fig. 5 with the behavior of human subjects in the experiment. The values of the parameters in the utility function of agent *1* are $\alpha_1 = 0.4$ and $\beta_1 = 0$, and the penalty is not larger than 40% of the excess of the payoff of an agent over that of the other. In contrast, the value of the coefficient $\beta_2 = 1.1$ in the utility function of agent *2* are considerably large, and therefore it appears that agent *2* strongly ask the opponent for a fair offer compared with agent *1*.

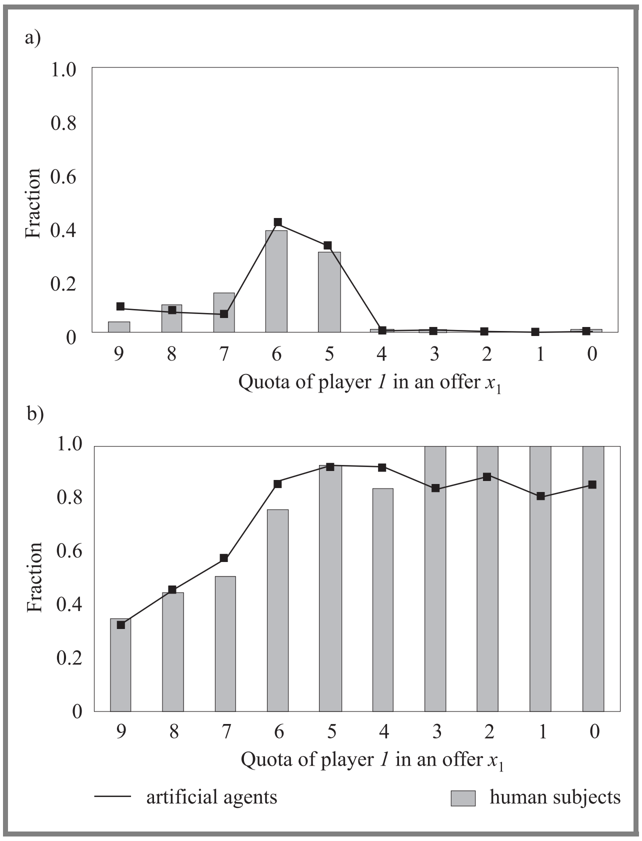


Fig. 5. Behavior of artificial agents and human subjects at the standard set of the parameters: (a) distribution of offers; (b) rate of acceptance.

The behavior of agents *1* and *2* can be characterized by the distribution of offers and the rate of acceptance, respectively. As can be seen in Fig. 5, all in all, the behavior of artificial agents in the simulation successfully approximates that of human subjects in the experiment. We will begin by examining the offers by agent *1*. The frequencies of the offers in which the quota of agent *1* is larger than 4, $x_1 > 4$, by the artificial agents in the simulation are similar to those by the human subjects in the experiments. For the offers such that the quota of agent *1* is smaller than or equal to 4, $x_1 \leq 4$, the behavior of the artificial agents in the simulation is almost the same as that of the human subjects in the experiment. Next, we look into the rate of acceptance. For the offers in which the quota of agent *1* is larger than or equal to 4, $x_1 \geq 4$, both of the rates of the simulation and the experiment denote a similar tendency; for the other offers, $x_1 \leq 3$, however, the rate of the simulation is slightly smaller than that of the experiment. This is attributed to the fact that as seen in the graph of Fig. 5a, the offers in which the quota of agent *1* is smaller than or equal to 4 are hardly proposed in the simulation and therefore the artificial agents cannot sufficiently learn how they respond such offers.

In Fig. 6, we show transitions of the average offer by agent *1* and the average rate of acceptance by agent *2* in the early

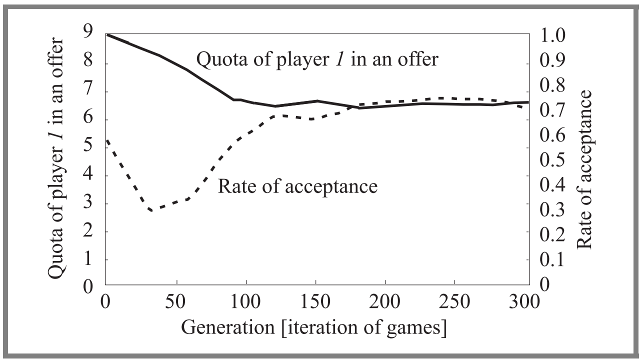


Fig. 6. Transitions of the average offer and acceptance rate.

generations of the simulation. As seen in the figure, at the beginning of the simulation, agent *1* makes the offer $(9,1)$ and agent *2* accepts it with the probability of about 0.55; it is conceivable that the couple of these actions is due to the preliminary learning of the neural network. Just after the start of the learning by the genetic algorithm, agent *2* begins to reject extremely unfair offers such as the offer $(9,1)$. However, as a quota x_1 of agent *1* in an offer (x_1,x_2) decreases by a high incidence of rejection of unfair offers by agent *2*, the rate of acceptance of the offer increases; after 150 generations, the average offer by agent *1* converges to an appreciably fair offer $(6.15,3.85)$ and the average rate of acceptance by agent *2* also converges to about 0.68.

Table 1
Utilities of agents *1* and *2* at the standard set of the parameters

x_1	9	8	7	6	5	4	3	2	1	0
u_1	9	8	7	6	5	3.2	1.4	−0.4	−2.2	−4.0
u_2	−7.8	−4.6	−1.4	1.8	5	6	7	8	9	10

In Table 1, the utilities of agents *1* and *2* are shown at the standard set of the parameters $(\alpha_1,\beta_1,\alpha_2,\beta_2) = (0.4,0,1.1,0)$. As seen in the table, the utility of agent *1* is an increasing function with the quota x_1 of agent *1*; the utility of agent *2* is a decreasing function with x_1. Especially, because the value of α_2, which is the penalty coefficient for the excess of the payoff of agent *1* over the payoff of agent *2*'s self in the utility of agent *2*, is relatively large, the utilities of agent *2* become negative when the quota x_1 of agent *1* is larger than or equal to 7, $x_1 \geq 7$. From the fact that the utility of agent *2* is zero when agent *2* rejects an offer, it is preferable for agent *2* to reject such offers. Such behavior of agent *2* can be interpreted as the punishment for unfair proposals by agent *1*. As can be seen in Fig. 6, through the repeated rejections by agent *2*, agent *1* gradually lowers a quota of agent *1*'s self in offers. This process can be explained by the learning of agent *1*. For the offer $(6,4)$, conversely it is advantageous for agent *2* to accept it. It seems to be for this reason that the frequency of the offer $(6,4)$ is the largest.

From the result of the simulation, it is conceivable that the developed agent-based simulation system successfully approximates the behavior of human subjects in the experiment by incorporating the fairness in the learning mechanism of the artificial agents. Moreover, while Abbink *et al.* [1] conclude that a fairness motive is a better explanation for why player *2* rejects unfair offers compared with learning, our result is consistent with their argument.

While we have found the standard set of the parameters imitating the behavior of human subjects by varying values of the parameters, we should examine effects of individual parameters on the behavior of the artificial agents. To verify that the behavior of agent *1* is mainly revised through the learning, by varying values of the parameters α_1 and β_1 in the utility function of agent *1*, we examine change of the behavior of the artificial agents. Moreover, while we suppose that fairness and the corresponding punishment largely explain the behavior of agent *2*, to confirm this argument, we also investigate change of the behavior of the artificial agents by varying value of the parameter α_2 in the utility function of agent *2*.

4.2. Effect of learning on the behavior of agent 1

Fixing the values of α_1, α_2 and β_2 at the standard setting $\alpha_1 = 0.4$, $\alpha_2 = 1.1$ and $\beta_2 = 0$, we vary the value of β_1 from 0 to 0.5 at intervals of 0.1. The result of this treatment is given in Fig. 7 showing the average offer and the average rate of acceptance.

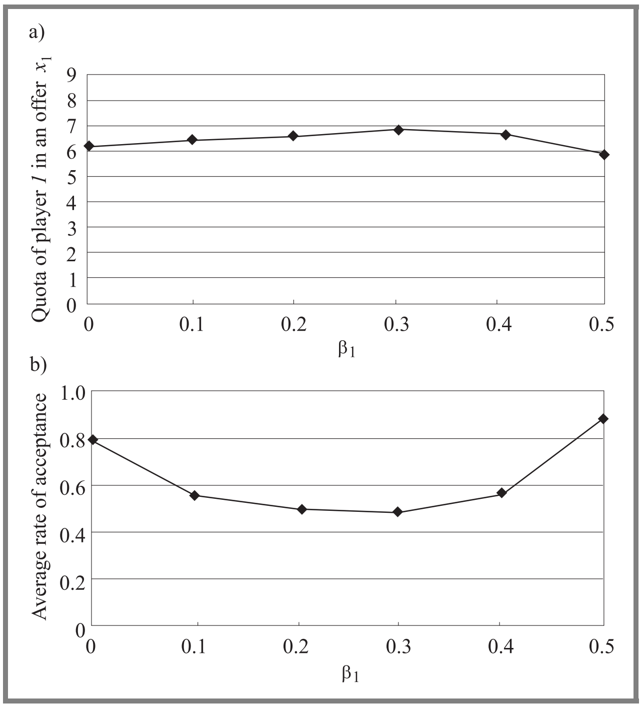


Fig. 7. Change of the behavior with respect to the parameter β_1: (a) average offer; (b) rate of acceptance.

As can be seen in Fig. 7, the average quota x_1 of agent *1* in offers remains almost the same in the range $0.0 \le \beta_1 \le 0.4$; when $\beta_1 = 0.5$, because the number of agent *1* making fairer offers increases, the average quota x_1 decreases below 6. Although the average rate of acceptance is relatively high when $\beta_1 = 0.0$ compared with the cases of $\beta_1 = 0.1, 0.2, 0.3, 0.4$, the average rates of acceptance are almost the same when $\beta_1 = 0.1, 0.2, 0.3, 0.4$. On the other hand, when $\beta_1 = 0.5$, because the average quota x_1 decreases and offers become fair, the average rate of acceptance obviously rises.

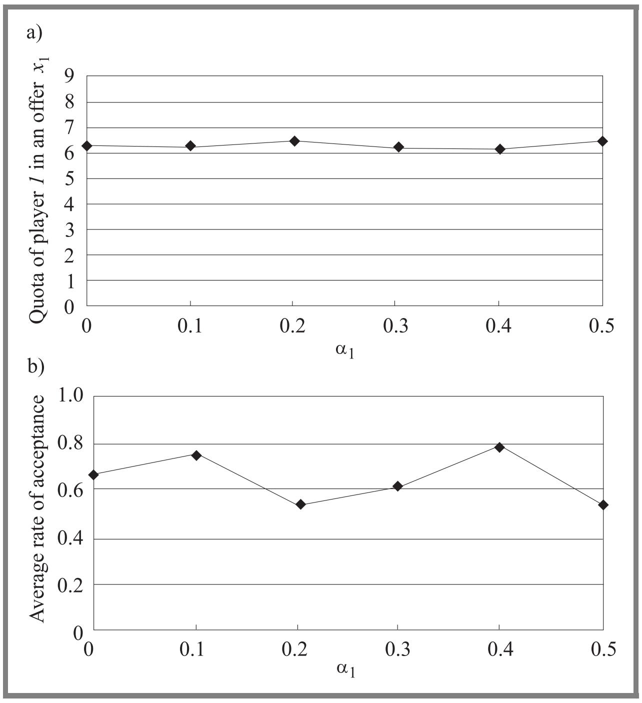


Fig. 8. Change of the behavior with respect to the parameter α_1: (a) average offer; (b) rate of acceptance.

For sensitivity with respect to the parameter α_1, the average offer and the average rate of acceptance are similarly given in Fig. 8. As can be seen in Fig. 8, the average quota x_1 of agent *1* in offers remains almost same in the range $0.0 \le \alpha_1 \le 0.5$. Moreover, from the graph of Fig. 8a, it is found that there is little linkage between the average rate of acceptance of agent *2* and the change of the value of α_1.

From the above observation, the sensitivity of the behavior of the artificial agents to the change of parameter β_1 or α_1 from zero is not so high, and it would be said that introduction of the parameter β_1 or α_1 does not have a major function in explanation of the behavior of the artificial agents. Thus, the effect of the parameter of fairness is relatively small, and it appears that the behavior of agent *1* is mainly revised through the learning.

When $\pi_1 > \pi_2$, the utility function of agent *1* is represented as $u_1(\pi_1, \pi_2) = (1 - 2\beta_1)\pi_1 + 10\beta_1$. If $\beta_1 < 0.5$, because the coefficient of π_1 is smaller than one, the influence of agent *2*'s decision of acceptance or rejection on the utility of agent *1* is evidently larger than that of the offer by agent *1*'s self. For the parameter α_1, because agent *1* rarely makes offers such that the quota of agent *1* is smaller than or equal to 4, $x_1 \le 4$, and α_1 is valid when $\pi_1 < \pi_2$, the parameter α_1 has little influence on the behavior of agent *1*. From this viewpoint, it is also found that the behavior of agent *1* be strongly affected by learning through a series of actions of agent *2*.

4.3. Effect of punishment on the behavior of agent 2

To observe effect of the punishment on the behavior of agent *2*, we conduct an additional treatment by varying the value of α_2 from 0 to 2 at intervals of 0.1, fixing the values of α_1, β_1 and β_2 at $\alpha_1 = 0$, $\beta_1 = 0$ and $\beta_2 = 0$. The result of this treatment is given in Fig. 9.

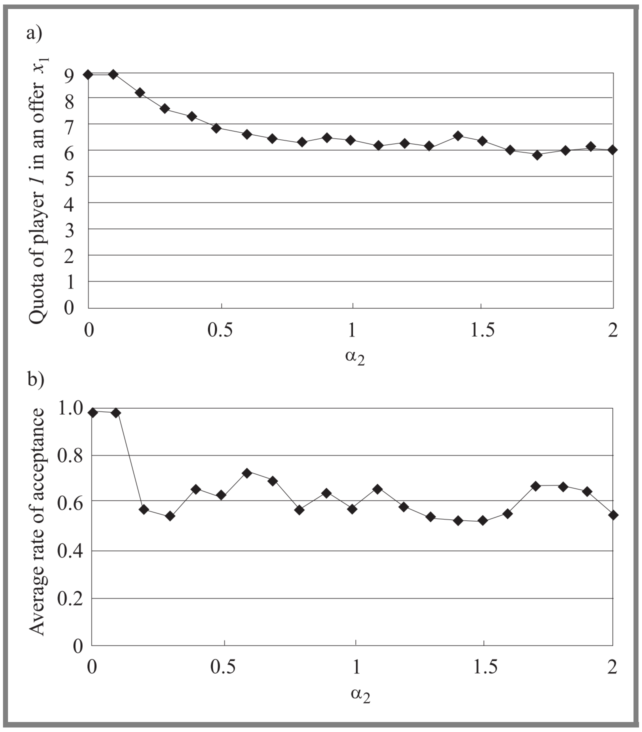


Fig. 9. Change of the behavior with respect to the parameter α_2: (a) average offer; (b) rate of acceptance.

As can be seen in Fig. 9, the average quota x_1 of agent *1* in offers specifically decreases in the range $0.1 \le \alpha_2 \le 0.8$, and the average rate of acceptance steeply drops from $\alpha_2 = 0.1$ to 0.2. It is just conceivable that the behavior of the artificial agents is very sensitive to the change of parameter α_2 from zero, and the behavior of agent *1* is mainly explained by introduction of the parameter α_2 of the fairness and punishment.

5. Conclusions

We have developed agent-based simulation system for analyzing the behavior of human subjects in the experiment. The learning mechanism incorporating the concept of fairness in the system efficiently works, and it is shown that our artificial adaptive agents successfully approximates the behavior of human subjects in the laboratory experiment by Roth *et al*. [20]. Through the simulation analysis, we have verified that the behavior of agent *1* is mainly revised through the learning, and fairness and corresponding punishment largely explain the behavior of agent *2*.

References

[1] K. Abbink, G. E. Bolton, A. Sadrieh, and F.-F. Tang, "Adaptive learning versus punishment in ultimatum bargaining", *Games Econom. Behav.*, vol. 37, pp. 1–25, 2001.

[2] G. E. Bolton, "A comparative model of bargaining: theory and evidence", *Amer. Econom. Rev.*, vol. 81, pp. 1096–1136, 1991.

[3] G. E. Bolton and A. Ockenfels, "ERC: a theory of equity, reciprocity, and competition", *Amer. Econom. Rev.*, vol. 90, pp. 166–193, 2000.

[4] G. E. Bolton and R. Zwick, "Anonymity versus punishment in ultimatum bargaining", *Game Econom. Behav.*, vol. 10, pp. 95–121, 1995.

[5] M. Costa-Gomes and K. G. Zauner, "Ultimatum bargaining behavior in Israel, Japan, Slovenia, and the United States: a social utility analysis", *Game Econom. Behav.*, vol. 34, pp. 238–269, 2001.

[6] J. Duffy and N. Feltovich, "Does observation of others affect learning in strategic envirnments? An experimental study", *Int. J. Game Theory*, vol. 28, pp. 131–140, 1999.

[7] E. Fehr and K. M. Schmidt, "A theory of fairness, competition and cooperation", *Q. J. Econom.*, vol. 114, pp. 817–868, 1999.

[8] R. Forsythe, J. L. Horowitz, N. E. Savin, and M. Sefton, "Fairness in simple bargaining experiments", *Games Econom. Behav.*, vol. 6, pp. 347–369, 1994.

[9] J. Gale, K. G. Binmore, and L. Samuelson, "Learning to be imperfect: the ultimatum game", *Games Econom. Behav.*, vol. 8, pp. 56–90, 1995.

[10] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning.* Reading: Addison Wesley, 1989.

[11] W. Güth, R. Schmittberger, and B. Schwarze, "An experimental analysis of ultimatum bargaining", *J. Econom. Behav. Organ.*, vol. 3, pp. 367–388, 1982.

[12] M. H. Hassoun, *Fundamentals of Artificial Neural Networks*. Cambridge: The MIT Press, 1995.

[13] E. Hoffman, K. A. McCabe, K. Shachat, and V. L. Smith, "Preferences, property rights, and anonymity in bargaining games", *Games Econom. Behav.*, vol. 7, pp. 346–380, 1994.

[14] E. Hoffman, K. A. McCabe, and V. L. Smith, "On expectations and the monetary stakes in ultimatum games", *Int. J. Game Theory*, vol. 25, pp. 289–301, 1996.

[15] D. Kahneman, J. L. Knetsch, and R. H. Thaler, "Fairness and the assumptions of economics", *J. Bus.*, vol. 59, pp. S285–S300, 1986.

[16] R. D. McKelvey and T. R. Palfrey, "Quantal response equilibria for normal form games", *Games Econom. Behav.*, vol. 10, pp. 6–38, 1995.

[17] J. Neelin, H. Sonnenschein, and M. Spiegel, "A further test of noncooperative bargaining theory: comment", *Amer. Econom. Rev.*, vol. 78, pp. 824–836, 1988.

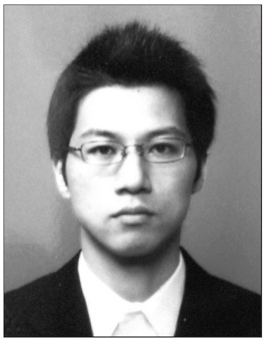
[18] M. Rabin, "Incorporating fairness into game theory and economics", *Amer. Econom. Rev.*, vol. 83, pp. 1281–1302, 1993.

[19] A. E. Roth and I. Erev, "Learning in extensive form games: experimental data and simple dynamic models in the intermediate term", *Games Econom. Behav.*, vol. 8, pp. 163–212, 1995.

[20] A. Roth, V. Prasnikar, M. Okuno-Fujiwara, and S. Zamir, "Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: an experimental study", *Amer. Econom. Rev.*, vol. 81, pp. 1068–1095, 1991.

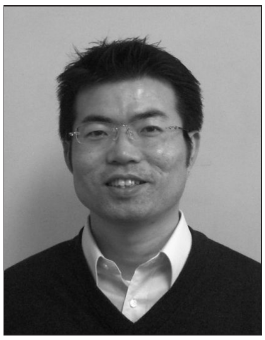
[21] E. Weg and V. Smith, "On the failure to induce meager offers in ultimatum game", *J. Econom. Psychol.*, vol. 14, pp. 17–32, 1993.

[22] K.-O. Yi, "Quantal-response equilibrium models of the ultimatum bargaining game", *Games Econom. Behav.*, vol. 51, pp. 324–348, 2005.



Tomohiro Hayashida is currently Assistant Professor at Department of Artificial Complex Systems Engineering, Graduate School of Engineering of the Hiroshima University, Japan. His current research interests are agent-based simulation analysis, game theory, and decision making theory.

e-mail: hayashida@hiroshima-u.ac.jp
Department of Artificial Complex Systems Engineering
Graduate School of Engineering, Hiroshima University
1-4-1 Kagamayama, Higashi-Hiroshima, 739-8527, Japan



Hideki Katagiri is currently Associate Professor at Department of Artificial Complex Systems Engineering, Graduate School of Engineering of the Hiroshima University, Japan. His current research interests are fuzzy stochastic programming, soft computing and marketing science.

e-mail: katagiri-h@hiroshima-u.ac.jp
Department of Artificial Complex Systems Engineering
Graduate School of Engineering, Hiroshima University
1-4-1 Kagamayama, Higashi-Hiroshima, 739-8527, Japan

Ichiro Nishizaki – for biography, see this issue, p. 35.