Evolution of Strategies with Different Representation Schemes

in a Spatial Iterated Prisoner's Dilemma Game

Hisao Ishibuchi, *Senior Member*, *IEEE*, Hiroyuki Ohyanagi, and Yusuke Nojima, *Member*, *IEEE*

Graduate School of Engineering, Osaka Prefecture University

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

Corresponding author: Prof. Hisao Ishibuchi

Graduate School of Engineering, Osaka Prefecture University,

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

Phone +81-72-254-9350, FAX +81-72-254-9915

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

Evolution of Strategies with Different Representation Schemes

in a Spatial Iterated Prisoner's Dilemma Game

Hisao Ishibuchi, *Senior Member*, *IEEE*, Hiroyuki Ohyanagi, and Yusuke Nojima, *Member*, *IEEE*

Abstract—The iterated prisoner's dilemma (IPD) game has been frequently used to examine the evolution of cooperative behavior among agents in the field of evolutionary computation. It has been demonstrated that various factors are related to the evolution of cooperative behavior. One well-known factor is spatial relations among agents. The IPD game is often played in a two-dimensional grid-world. Such a spatial IPD game has a neighborhood structure, which is used to choose opponents for the IPD game and parents for genetic operations. Another important factor is the choice of a representation scheme to encode the strategy of each agent. Different representation schemes often lead to different results. Whereas the choice of a representation scheme is known to be important, a mixture of different representation schemes has not been examined for the spatial IPD game in the literature. That is, a population of homogeneous agents with the same representation scheme has been usually assumed in the literature. In this paper, we introduce the use of different representation schemes in a single population to the spatial IPD game in order to examine the evolution of cooperative behavior under more general assumptions. With the use of different representation schemes, we can examine the evolution of cooperative behavior in various settings such as partial interaction through the IPD game, partial interaction through crossover, full interaction through the IPD game and crossover, and no interaction between different sub-populations of agents.

I. INTRODUCTION

The prisoner's dilemma is a well-known non-zero-sum game. In this game, two players independently choose one of the two actions: cooperate and defect. If both players cooperate, they can enjoy high payoff. However, one player's cooperation leads to the worst result for that player (and the best result for its opponent) if the opponent defects. Thus each player tends to defect, which leads to mutual defection with low payoff. The dilemma for the players is that they will eventually receive low payoff from mutual defection whereas higher payoff can be obtained from mutual cooperation.

The evolution of cooperative behavior has been frequently studied for the iterative version of the prisoner's dilemma game in the field of evolutionary computation since the late 1980s [1] and the early 1990s [2], [3]. Various strategy representations have been studied for the iterated prisoner's dilemma (IPD) game such as a binary string, a neural network, and a decision tree. IPD game strategies are evolved through selection and variation operators (crossover and mutation). The fitness of an agent in a population is defined by its average payoff obtained from the IPD game against other agents in the same population. A number

of techniques have been introduced to the IPD game for further examining the evolution of cooperative behavior such as the speciation of strategies [4], individual recognition [5], and partner selection [6]. The IPD game has been also extended to various situations such as a multi-player version [7]-[9], a spatial version [10]-[15], stochastic strategies [13], [14], random pairing [15], multiple objectives [16], multiple choices [17], and noisy games [18]. See [19], [20] for a review of studies on the evolution of cooperative behavior among agents in the IPD game.

Recently the IPD game has been used to examine the effect of the choice of a representation scheme on the evolution of cooperative behavior in some studies [21]-[23]. Those studies examined a wide range of strategy representations such as a finite-state machine, a feed-forward neural network, and a lookup table. Experimental results showed that the choice of a representation scheme had a large effect on the evolution of cooperative behavior.

In almost all studies on the evolution of cooperative behavior in the IPD game, a single representation scheme of strategies was assumed so that an arbitrary pair of strategies can be used for the IPD game and genetic operations. That is, a mixture of different representation schemes has not been examined as a population. In this paper, we examine the evolution of cooperative behavior in a population of heterogeneous agents with different representation schemes.

The main novelty of this paper is the use of different representation schemes in a single population. In most cases, strategies with different representation schemes are not recombined to generate new strategies. Whereas we use the word "population" to refer to a set of agents as in many other studies on the IPD game, "ecology" may be a more appropriate word when strategies of agents with different representation schemes are not recombined.

The motivation behind the use of different representation schemes is the fact that different species interact with each other in many real-world situations whereas they are not recombined. The use of different representation schemes for evolutionary optimization in the literature also motivated us to examine such a situation in the framework of the IPD game. For example, it was shown by Skolicki and De Jong [24] that multi-representation island models outperformed standard evolutionary algorithms on difficult multi-modal function optimization problems.

Another important factor that has a large effect on the evolution of cooperative behavior is spatial relations among agents. It is well-known that the use of a two-dimensional grid-world often facilitates the evolution of cooperative behavior. A single neighborhood structure was used for both local opponent selection in the IPD game and local parent selection in genetic operations in the literature [10]-[14]. In [15], we used two neighborhood structures motivated by the idea of structured demes [25]-[28]. One is for the interaction among agents through the IPD game. Each agent in a cell plays against only its neighbors defined by this neighborhood structure. That is, this neighborhood structure is for local opponent selection. The other is for the mating of strategies. A new strategy for an agent is generated from a pair of parents in its neighboring cells defined by the

second neighborhood structure. That is, this neighborhood structure is for local parent selection.

Whereas a single neighborhood structure has been usually used in the spatial IPD game (and in cellular algorithms in the field of evolutionary computation in general), there exist a number of real-world situations with two neighborhood structures. For example, neighboring plants fight with each other for water and sunlight in one neighborhood structure. This neighborhood structure is much smaller than the other one where they can disperse their pollen. Another example is territorial animals. One neighborhood structure can be used to model their territories. The same neighborhood structure, however, cannot be used to model their behavior in the breeding season because territorial animals often move beyond their territories to find their mates.

The use of two neighborhood structures was examined in [29] in a continuous prisoner's dilemma model: One for local opponent selection and the other for local payoff comparison. In [30], [31], two neighborhood structures were used for optimization problems: One for local fitness evaluation and the other for local parent selection.

This paper is an extended version of our previous studies [32], [33]. We examined the evolution of cooperative behavior in a mixture of heterogeneous agents in a short (i.e., 4-page) journal paper [32] where we used only 3-bit and 5-bit strategies. In a conference paper [33], we used real number strings of length 3 and length 5 as stochastic strategies as well as deterministic 3-bit and 5-bit strategies. Due to the page limitation, we reported only a small number of experimental results with no detailed discussions in [32], [33]. In this paper, we report much more experimental results together with detailed discussions. We also examine the effect of two types of interaction between agents (i.e., the IPD game and crossover) on the evolution of cooperative behavior. For this purpose, we use real number strings as deterministic strategies where real numbers in the unit interval [0, 1] are rounded to 0 or 1 based on the threshold value 0.5. Such a deterministic strategy, which behaves like a deterministic binary strategy in the IPD game, can be recombined with a stochastic strategy. Thus we can examine the following four situations of the interaction between agents with different representation schemes: Full interaction through the IPD game and crossover, partial interaction only through the IPD game, partial interaction only through crossover, and no interaction.

The rest of this paper is organized as follows. First we explain the IPD game and some representation schemes in Section II. Next we explain our spatial IPD game in a grid-world with two neighborhood structures in Section III. Then we report experimental results in Section IV. Finally we conclude this paper in Section V.

II. IPD GAME

We use a typical payoff matrix in Table I. When both agents cooperate, each receives three points. When both agents defect, each receives one point. The highest payoff of five is obtained by defecting when the opponent cooperates. In this case, the opponent receives the lowest payoff of zero.

An agent's strategy determines its next action based on a finite history of previous rounds of the game. Binary strings have

often been used to represent strategies where "1" and "0" usually mean "cooperate" and "defect", respectively. In Table II, we show a 3-bit strategy "101" called TFT (Tit-for-Tat), which determines its next action based on the opponent's action in the previous round of the game. An agent with this strategy cooperates at the first round and then cooperates at each round only when the opponent cooperated in the previous round. In Table III, the same TFT strategy is represented by a 5-bit strategy "10011".

TABLE I

PAYOFF MATRIX OF ITERATED PRISONER'S DILEMMA GAME

Agent's Action	Opponent's Action	
	C: Cooperate	D: Defect
C: Cooperate	Agent: 3 Opponent: 3	Agent: 0 Opponent: 5
D: Defect	Agent: 5 Opponent: 0	Agent: 1 Opponent: 1

TABLE II

THREE-BIT STRING "101" REPRESENTING THE TIT-FOR-TAT STRATEGY

Agent's first action: Cooperate		1
Opponent's previous action	Suggested action	
D: Defect	D: Defect	0
C: Cooperate	C: Cooperate	1

TABLE III

FIVE-BIT STRING "10011" REPRESENTING THE TIT-FOR-TAT STRATEGY

Agent's first action: Cooperate			1
Actions on the preceding round		Suggested action	
Player	Opponent		
D: Defect	D: Defect	D: Defect	0
C: Cooperate	D: Defect	D: Defect	0
D: Defect	C: Cooperate	C: Cooperate	1
C: Cooperate	C: Cooperate	C: Cooperate	1

Some 3-bit strategies have special names as

000: ALLD (Always Defect),

001: STFT (Suspicious TFT),

010: ATFT (Anti TFT),

101: TFT (Tit-for-Tat),

111: ALLC (Always Cooperate).

Real number strings can be used to represent stochastic strategies where each real number in the unit interval [0, 1] shows the

probability of "cooperate". Binary strings such as "101" in Table II and "10011" in Table III can be viewed as a special case of real number strings. For example, an agent with a real number string "0.8 0.1 1.0" of length 3 cooperates at the first round with the probability 0.8. When the opponent defected in the previous round, this agent cooperates with a small probability 0.1. This agent always cooperates (i.e., with the probability 1.0) when the opponent cooperated in the previous round. As shown in this example, the probability of each action in the current round is determined by the result of the previous round in the case of stochastic strategies.

In this paper, we examine four types of strings for representing IPD game strategies: Binary strings of length 3 and 5, and real number strings of length 3 and 5. Whereas binary strings are always used for representing deterministic strategies, we examine two cases about the usage of real number strings: Stochastic and deterministic strategies. When real number strings are used as stochastic strategies, each real number is handled as the probability of cooperation as we have already explained. However, when real number strings are used as deterministic strategies, they are interpreted as binary strings by viewing real numbers less than 0.5 and more than or equal to 0.5 as 0 and 1, respectively.

Since real number strings are used for stochastic and deterministic strategies, we have six representation schemes of IPD game strategies as follows:

    1. Deterministic 3-bit strategies.

    2. Deterministic 5-bit strategies.

    3. Stochastic length 3 real number strategies.

    4. Stochastic length 5 real number strategies.

    5. Deterministic length 3 real number strategies.

    6. Deterministic length 5 real number strategies.

Whereas any pair of strategies can play the IPD game, we do not recombine a pair of binary and real number strategies. We do not recombine a pair of strategies of different length, either. It should be noted that a pair of stochastic and deterministic real number strategies can be recombined when they have the same length. This means that we can examine the effect of the interaction through crossover between strategies with different representation schemes on the evolution of cooperative behavior.

III. SPATIAL IPD GAME

We use an 11×11 two-dimensional grid-world in our spatial IPD game (with the torus structure). Since a single agent is located in each cell of the grid-world, the population size is 121. In the grid-world, each agent plays the IPD game against only

its neighbors defined by a neighborhood structure for interaction. Let $N_{\mathrm{IPD}}(i)$ be the set of Agent i and its neighbors. $N_{\mathrm{IPD}}(i)$ is the neighborhood structure for local opponent selection. Agent i plays the IPD game against only agents in $N_{\mathrm{IPD}}(i)$. In computational experiments, we examine seven specifications of the size of $N_{\mathrm{IPD}}(i)$: 5, 9, 13, 25, 41, 49 and 121. When the size is 121, the neighborhood structure is actually the same as the whole grid-world. Fig. 1 shows the other six neighborhood structures. The standard non-spatial IPD game can be viewed as the case where $N_{\mathrm{IPD}}(i)$ is the same as the whole grid-world (i.e., the size of $N_{\mathrm{IPD}}(i)$ is 121).

The IPD game is played between an agent and one of its neighbors including the agent itself for a prespecified number of rounds (e.g., 100 rounds in our computational experiments in this paper). After an agent completes the IPD game against a prespecified number of its neighbors, the fitness value of the agent is calculated as the average payoff obtained from each round of the game. When the size of $N_{\mathrm{IPD}}(i)$ is five, the fitness value of each agent is calculated after the IPD game is completed against all of its five neighbors. However, when the size of $N_{\mathrm{IPD}}(i)$ is larger than five, Agent i randomly selects five opponents from $N_{\mathrm{IPD}}(i)$ to calculate the fitness value. That is, each agent plays the IPD game against the same number of neighbors independent of the neighborhood size for local opponent selection. In other words, the same computation load is always used for the fitness evaluation of each agent. It is possible to evaluate the fitness of each agent from the IPD game against its all neighbors. In our former study on the IPD game with homogeneous agents [15], we examined these two cases (i.e., five neighbors and all neighbors) and obtained similar results. Since the choice of five neighbors significantly decreases the computation time, we perform the IPD game against five neighbors in this paper.

(a) Size 5.   (b) Size 9.   (c) Size 13.

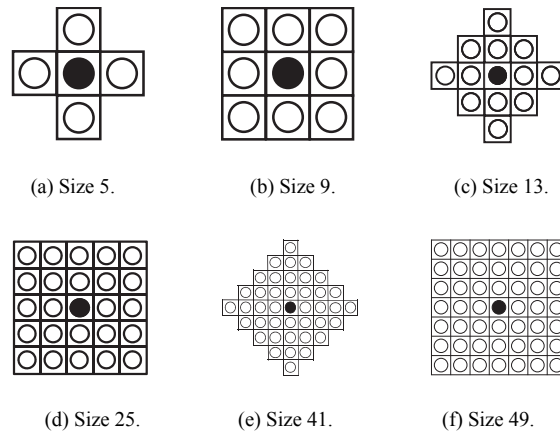(d) Size 25.   (e) Size 41.   (f) Size 49.

Fig. 1. Examples of neighborhood structures.

A new strategy of an agent is generated by genetic operations from two parents selected from its neighbors. Let $N_{\mathrm{GA}}(i)$ be the set of Agent i and its neighbors. $N_{\mathrm{GA}}(i)$ is the neighborhood structure for local parent selection. It should be noted that $N_{\mathrm{GA}}(i)$

for local parent selection is not always the same as $N_{IPD}(i)$ for local opponent selection in the IPD game.

Let $f(s_i)$ be the fitness value of Agent i with Strategy s_i. We use the roulette wheel selection scheme with a linear scaling to select two parents of a new strategy for Agent i:

$$p_i(s_j) = \frac{f(s_j) - f_{min}(N_{GA}(i))}{\sum\limits_{k \in N_{GA}(i)} \{f(s_k) - f_{min}(N_{GA}(i))\}}, \ j \in N_{GA}(i), \tag{1}$$

where $p_i(s_j)$ is the selection probability of Strategy s_j of Agent j as a parent of a new strategy for Agent i, and $f_{min}(N_{GA}(i))$ is the minimum fitness value among the neighbors in $N_{GA}(i)$. A new strategy for Agent i is generated by applying crossover and mutation to the selected two parents. We use the roulette wheel selection scheme with the linear scaling in (1) as in our former study [15] on the IPD game with homogeneous agents. This is not necessarily the best choice with respect to the average payoff over all agents. We could have used other selection schemes in our computational experiments.

After new strategies for all agents are generated, the current population of strategies is replaced with the newly generated strategies (i.e., the current strategy in each cell is replaced with the newly generated strategy for that cell). The fitness evaluation of each strategy through the IPD game and the generation update by genetic operations are iterated for a prespecified number of generations (e.g., 1000 generations in our computational experiments).

It should be noted that $N_{GA}(i)$ in (1) excludes any neighbors that cannot be recombined with Agent i when each agent has a different representation scheme. This means that each agent does not change its representation scheme by genetic operations. In each execution of our computational experiment, first a representation scheme is assigned to each agent. The assigned representation scheme is not changed during 1000 generations.

IV. EXPERIMENTAL RESULTS

A. Conditions of Computational Experiments

As explained in Section II, we examined six representation schemes: Deterministic 3-bit binary strategies in Table II, deterministic 5-bit binary strategies in Table III, deterministic and stochastic length 3 real number strategies, and deterministic and stochastic length 5 real number strategies.

An initial population was randomly generated. A real number in the unit interval [0, 1] was randomly chosen in the case of real number strategies. Each agent played the IPD game for 100 rounds against each of its five neighbors. The evolution of strategies was continued for 1000 generations. As explained in Section III, we examined the seven neighborhood structures for local opponent selection and local parent selection. This means that we examined all the possible 49 combinations of those seven neighborhood structures for local opponent selection and local parent selection. For each of the 49 combinations, we

report average results over 1000 runs for 1000 generations with 121 agents.

For binary strings, we used the one-point crossover and the bit-flip mutation. For real number strings, we used the blend crossover (BLX-α [34]) with $\alpha = 0.25$ and the uniform mutation. The same crossover probability 1.0 and the same mutation probability $1/(121 \times 5)$ were used for both binary and real number strings in our computational experiments.

Let us briefly explain the blend crossover. It generates an offspring $\mathbf{z} = (z_1\, z_2\, ...\, z_n)$ from two parents $\mathbf{x} = (x_1\, x_2\, ...\, x_n)$ and $\mathbf{y} = (y_1\, y_2\, ...\, y_n)$ by randomly choosing a real number for z_i in the interval $[\min\{x_i, y_i\} - \alpha|x_i - y_i|,\, \max\{x_i, y_i\} + \alpha|x_i - y_i|]$ for $i = 1, 2, ..., n$. If the chosen value of z_i is meaningless as a probability, it is repaired as follows: $z_i = 0$ (if $z_i < 0$) and $z_i = 1$ (if $z_i > 1$). When $\alpha = 0$, the value of z_i is chosen as a real number between x_i and y_i, which leads to slow evolution of cooperative behavior. If α is too large (i.e., $\alpha = 10$), the blend crossover often generates a real number outside the unit interval [0, 1]. This parameter was specified as $\alpha = 0.25$ in our computational experiments.

The uniform mutation for real number strings replaces some elements of the newly generated offspring $\mathbf{z}$ with random real numbers in the unit interval [0, 1]. This operation is applied to each element with a small mutation probability (e.g., $1/(121 \times 5)$ in our computational experiments).

### B. Experimental Results Using Homogeneous Agents

We first show experimental results with homogeneous agents. That is, a single representation scheme was used by all the 121 agents in our computational experiments reported in this subsection. The average payoff of those agents over 1000 runs is summarized in Figs. 2-4 for each combination of the two neighborhood structures. Deterministic binary strategies, stochastic and deterministic real number strategies were used in Fig. 2, Fig. 3 and Fig. 4, respectively.

In Fig. 2 with binary strategies, high average payoff was obtained independent of the specifications of the two neighborhood structures. High average payoff, however, was obtained from stochastic real number strategies in Fig. 3 only when the two neighborhood structures were small. The average payoff was severely decreased by the use of a large neighborhood structure $N_{\mathrm{IPD}}(i)$ for local opponent selection in Fig. 3. It is clear in Fig. 3 that the size of $N_{\mathrm{IPD}}(i)$ for local opponent selection had much larger effects on the average payoff than the size of $N_{\mathrm{GA}}(i)$ for local parent selection.

In Figs. 2-4, higher average payoff was obtained from deterministic strategies in Fig. 2 and Fig. 4 than stochastic strategies in Fig. 3. We can also see that deterministic real number strategies in Fig. 4 led to lower average payoff than deterministic binary strategies in Fig. 2.

We performed the Wilcoxon signed-ranks test [35] to examine whether there is a significant difference in the average payoff between Fig. 2 (a) with deterministic 3-bit strategies and Fig. 4 (a) with deterministic length 3 real number strategies. For each

of the 49 combinations of the two neighborhood structures in Fig. 2 (a) and Fig. 4 (a), we calculated a difference score (i.e., the difference in the average payoff between deterministic 3-bit strategies and deterministic length 3 real number strategies). The null hypothesis is that the sum of the ranks of the positive difference scores is equal to the sum of the ranks of the negative difference scores. Let us briefly explain the test procedure [35]. First each difference score is ranked using its absolute value as follows. If the absolute value is zero, the difference score is not ranked. A rank of 1 is assigned to the difference score with the smallest absolute value (except for 0). A rank of 2 is assigned to the difference score with the second smallest absolute value. In this manner, a rank is assigned to each difference score. If there are multiple difference scores with the same absolute value, the average value of possible ranks is assigned to all difference scores with the same absolute value. For example, if we have two difference scores with the second smallest absolute value, their possible ranks are rank 2 and rank 3. Thus their ranks are 2.5. After the ranking, the sign of each difference score is reassigned to its rank. A positive sign means that a higher average payoff was obtained from deterministic 3-bit strategies than deterministic length 3 real number strategies. The sum of the ranks with positive signs was 1188 while that with negative signs was 37. Let T be the smaller value between 1188 and 37 (i.e., $T = 37$). Using the SPSS software, the critical T value for the significance $\alpha = 0.05$ was calculated as $T = 415$ for 49 samples. Thus the null hypothesis was rejected. The p-value was calculated as being less than 0.0001 using the SPSS software. Thus the difference is statistically significant between Fig. 2 (a) and Fig. 4 (a).
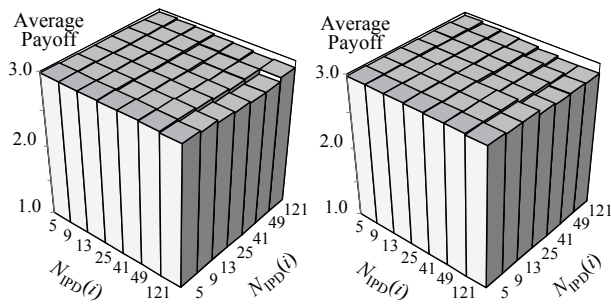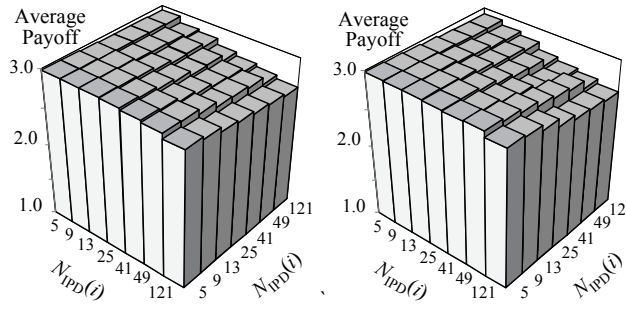

(a) Deterministic 3-bit strategies.　　　(b) Deterministic 5-bit strategies.

Fig. 2. Average payoff by homogeneous agents with binary strategies.
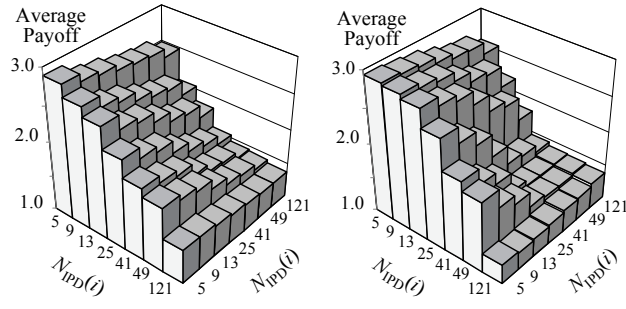
(a) Stochastic length 3 strategies.　　(b) Stochastic length 5 strategies.

Fig. 3. Average payoff by homogeneous stochastic real number strategies.

(a) Deterministic length 3 strategies.　　(b) Deterministic length 5 strategies.

Fig. 4. Average payoff by homogeneous deterministic real number strategies.

In the same manner, we performed the Wilcoxon signed-ranks test to compare the average payoff between Fig. 2 (b) and Fig. 4 (b). The p-value was also calculated as being less than 0.0001. Thus we can say that the difference is statistically significant between Fig. 2 (b) and Fig. 4 (b).

Each round of our IPD game has four possible results: {(Agent's action, Opponent's action)} = {(D, D), (C, D), (D, C), (C, C)}. In each case, the average payoff over the agent and its opponent is calculated from Table I as follows: 1.0 for (D, D), 2.5 for (C, D), 2.5 for (D, C), and 3.0 for (C, C). These calculations show that the maximum average payoff 3.0 is obtained only when all agents always cooperate.

As shown in Fig. 3, the choice of the two neighborhood structures had a large effect on the average payoff when we used stochastic real number strategies. We further examined the case of stochastic length 5 real number strategies using the four extreme combinations of the neighborhood structures: $(|N_{GA}(i)|, |N_{IPD}(i)|) = (5, 5), (5, 121), (121, 5), (121, 121)$. On the left-hand side of Fig. 5, we show the average payoff at each generation for each of the four combinations. Fig. 5 demonstrates that the use of the smallest neighborhood structure $N_{IPD}(i)$ with five neighbors for local opponent selection facilitated the

evolution of cooperative behavior.

The box plot on the right-hand side of Fig. 5 shows the average payoff at the 1000th generation for each of the four neighborhood combinations. The box plot was depicted in the following manner for each neighborhood combination. First we calculated the average payoff over 121 agents at the 1000th generation for each of 1000 runs. Then we found the minimum value, 25 percentile, median, 75 percentile and the maximum value from the calculated 1000 average payoff values. Each box in the box plot shows the range between the 25 and 75 percentiles. The median is shown by the bold line in the box while the maximum and minimum values are shown by the vertical line. For example, we can see from the box plot for (5, 5) in Fig. 5 that the average payoff at the 1000th generation was higher than 2.5 in all the 1000 runs when we used the smallest neighborhood structure with five neighbors for local parent selection and local opponent selection. In the case of (5, 121) with 5 parent selection neighbors and 121 opponent selection neighbors, the average payoff was close to 3.0 (i.e., 100% mutual cooperation) in some runs whereas the median was about 1.25. In Fig. 6, we show the histogram of the 1000 average payoff values for this case (i.e., (5, 121) in Fig. 5). Whereas the average payoff was less than 1.5 in many runs, the average payoff close to 3.0 was also obtained in about 90 runs (i.e., about 9% of 1000 runs) with this setting of the two neighborhood structures.
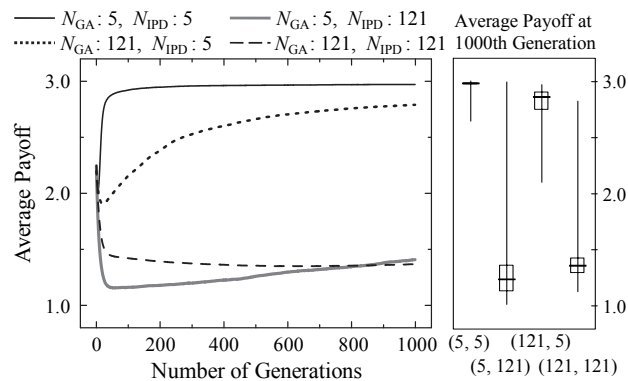

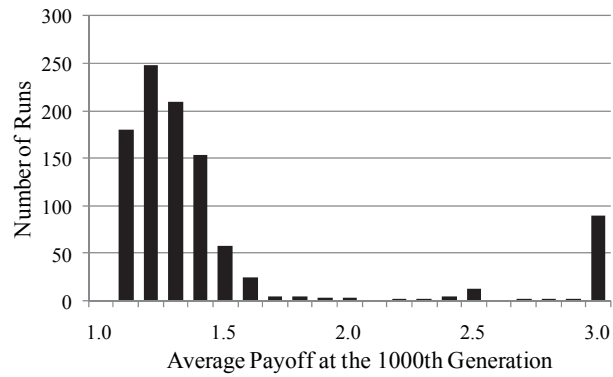Fig. 5. Average payoff by homogeneous stochastic length 5 real number strategies.

Fig. 6. Histogram of the average payoff at the 1000th generation of each of 1000 runs in the case of (5, 121) in Fig. 5.

Let us further discuss the results in Fig. 5 with stochastic length 5 real number strategies. Since initial strategies were randomly generated using real numbers in [0, 1], the expected average probability of cooperation over all agents in the initial generation is 0.5. In this case, the expected average payoff over all agents is calculated as $(1.0 + 2.5 + 2.5 + 3.0)/4 = 2.25$ since all the four possible results (D, D), (D, C), (C, D) and (C, C) with the average payoff 1.0, 2.5, 2.5 and 3.0 have the same probability. We can see in Fig. 5 that the average payoff at the initial generation was about 2.25.

Since the expected average probability of cooperation over all agents is 0.5 in the initial generation in Fig. 5, the expected payoff from "D: defect" and "C: cooperate" can be calculated as $(1 + 5)/2 = 3.0$ and $(0 + 3)/2 = 1.5$, respectively. These calculations suggest that stochastic strategies with less tendency of cooperation are more likely to receive higher average payoff. As a result, the second generation may include more strategies with less tendency of cooperation. That is, "D: defect" is more likely to be chosen in the second generation, which decreases the average payoff over all agents. These discussions explain why the average payoff decreased from 2.25 in the first few generations in Fig. 5.

When the opponent selection neighborhood $N_{IPD}(i)$ is very small, small groups of adjacent agents with high tendency of cooperation are also likely to receive higher average payoff. Thus the number of those agents started to increase just after the rapid decrease in the average payoff in the first few generations when the size of the neighborhood structure for local opponent selection was very small in Fig. 5.

In the same manner as Fig. 5, we show experimental results in Fig. 7 for the case of deterministic length 5 real number strategies. Much higher average payoff was obtained in Fig. 7 by using real number strings as deterministic strategies than Fig. 5 with stochastic length 5 real number strategies. We also show experimental results for the case of deterministic 5-bit binary strategies in Fig. 8. The use of binary strategies in Fig. 8 further increased the average payoff from Fig. 7 where real number strings were used as deterministic strategies. In Fig. 8, the average payoff close to 3.0 (i.e., 100% mutual cooperation) was obtained in many runs in all the four combinations of the neighborhood structures. However, the average payoff close to 1.0

(i.e., 100% mutual defection) was also obtained in some runs in Fig. 8 as shown in the box plot.
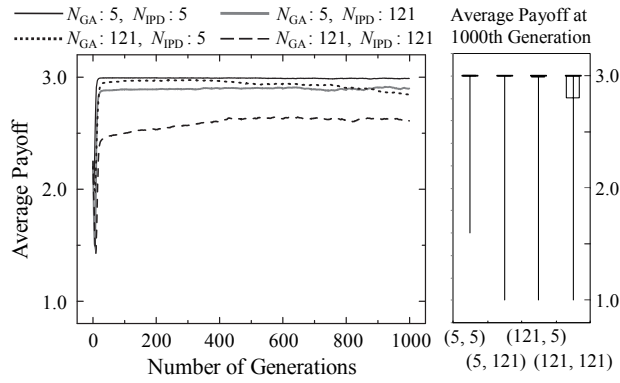

Fig. 7. Average payoff by homogeneous deterministic length 5 real number strategies.
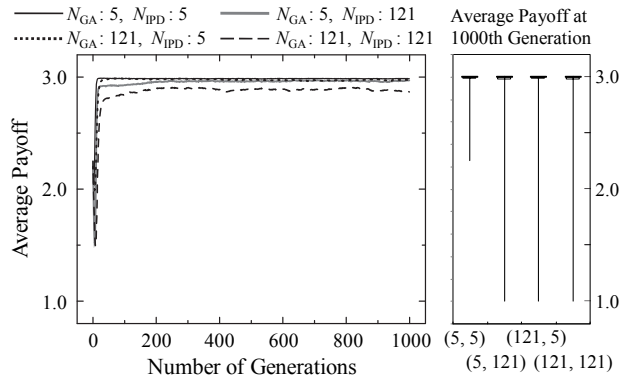

Fig. 8. Average payoff by homogeneous 5-bit binary strategies.

Figs. 5, 7 and 8 suggest that 1000 generations seem to be enough in many cases in our computational experiments. This is because the changes in the average payoff were small in the last 500 generations in those figures if compared with very fast increase/decrease in the first 100 generations.

In our computational experiments, the number of rounds in the IPD game was specified as 100. Let us discuss whether 100 rounds are sufficient to calculate the average payoff. The choice of the next action by each of the six strategies in this paper is based on the result of the current round of the IPD game. As we have already explained, each round of the IPD game has the four possible results: (D, D), (D, C), (C, D) and (C, C). When the IPD game is performed for 100 rounds between deterministic strategies, some or all of these four results appear cyclically over 100 rounds. Thus 100 rounds seem to be enough to evaluate the average payoff of an agent against its opponent when they use deterministic strategies.

In the case of stochastic strategies, no cycles appear over 100 runs. In order to examine whether 100 rounds are enough or not

for stochastic strategies, we randomly generated 100 pairs of stochastic length 3 real number strategies (i.e., 200 different strategies). The IPD game was performed by each of the 100 pairs. The average payoff of each of the 200 strategies was calculated over 100, 1000 and 10000 rounds. In Fig. 9 (a), we show the relation between the 100-round average payoff (the horizontal axis) and the 10000-round average payoff (the vertical axis) of each of the 200 strategies. For comparison, we also show the relation between the 1000-round average payoff and the 10000-round average payoff in Fig. 9 (b). In Fig. 9, the average payoff over a different number of rounds was not the same due to the stochastic nature of strategies. Whereas it is unclear in Fig. 9 whether 100 rounds are enough for stochastic strategies, we specified the number of rounds as 100 since the effect of the stochastic nature of strategies on the 100-round average payoff does not look so large in Fig. 9 (a). We did not specify it as 1000 or 10000 in order to prevent too much increase of computation time in this paper.

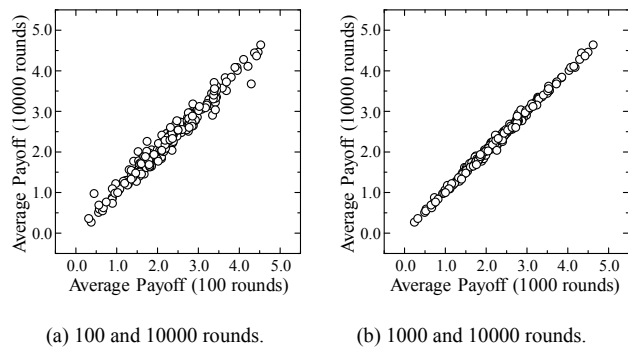| (a) 100 and 10000 rounds. | (b) 1000 and 10000 rounds. |

Fig. 9. Average payoff of stochastic length 3 real number strategies.

We also performed the same computational experiments using all the 64 pairs of the eight ($8=2^3$) deterministic 3-bit strategies. Experimental results are summarized in Fig. 10 where the average payoff over a different number of rounds is almost the same among the three settings.
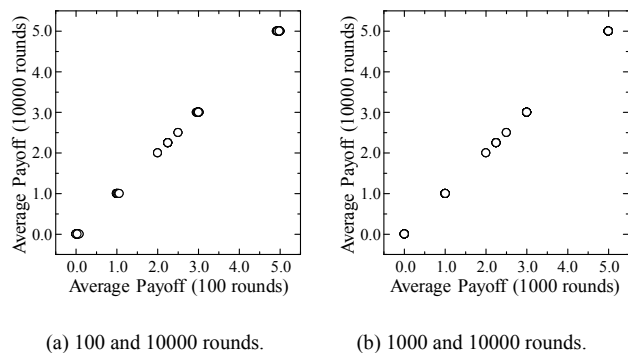

| (a) 100 and 10000 rounds. | (b) 1000 and 10000 rounds. |

Fig. 10. Average payoff of deterministic 3-bit strategies.

C. Experimental Results Using Heterogeneous Agents

Next we report experimental results where a different representation scheme was used by each agent. In this subsection, we randomly divided the 121 cells of the 11×11 grid-world into two sub-populations of almost the same size (i.e., 61 and 60 cells). One representation scheme was assigned to all cells in one sub-population, and another one to the other cells. We also examined the case where the same representation scheme was assigned to both sub-populations. Then the evolution of cooperative behavior was examined in the same manner as in the previous subsection using the following three rules:

Rule 1: Local opponent selection is independent of the population subdivision. That is, Agent i can play the IPD game against any neighbor in $N_{\text{IPD}}(i)$ even when they are in different sub-populations.

Rule 2: Local parent selection is performed within each sub-population. That is, $N_{\text{GA}}(i)$ includes only parent selection neighbors in the same sub-population as Agent i. As a result, strategies in different sub-populations cannot be recombined.

Rule 3: Population subdivision and representation scheme assignment do not change during the evolution of cooperative behavior over 1000 generations. That is, no agent changes its representation scheme over 1000 generations.

We examined all the 4×4 pairs of the deterministic binary string-based and stochastic real number-based representation schemes (i.e., deterministic 3-bit, deterministic 5-bit, stochastic length 3 real number, and stochastic length 5 real number strategies). Experimental results are summarized in Fig. 11 where a plot at each field shows the average payoff of the corresponding 50% row agents in a mixture with 50% column agents. For example, the rightmost top plot in Fig. 11 (i.e., Fig. 11 (d)) shows the average payoff of 50% agents with 3-bit strategies in a mixture with 50% agents with stochastic length 5 real number strategies. In each of the four diagonal plots (i.e., Fig. 11 (a), (f), (k), (p)), the same representation scheme was assigned to the two sub-populations. Even in those plots, strategies of two agents from different sub-populations were not recombined in Fig. 11 due to the above-mentioned second rule (i.e., Rule 2).
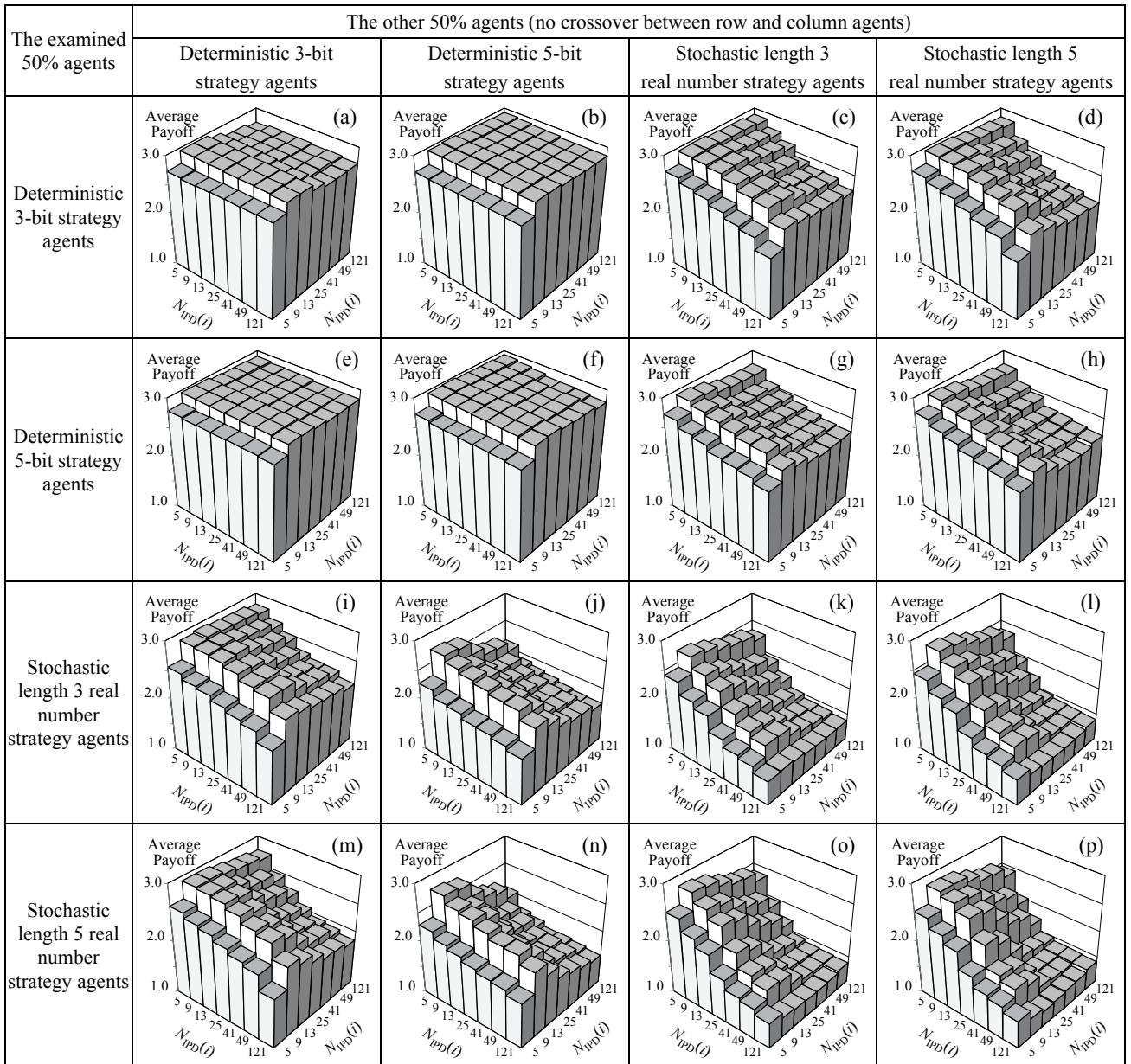
Fig. 11. Average payoff of 50% row agents in a mixture with 50% column agents. The 121 cells were randomly divided into two sub-populations of almost the same size (i.e., 61 and 60 cells). In each of the 16 plots, two sub-populations consist of the corresponding row and column agents. The IPD game was performed between agents independent of the population subdivision whereas crossover was performed between agents in the same sub-population. In each of the four diagonal plots, the same representation scheme was assigned to the two sub-populations. Even in those plots, no crossover was performed between two agents in different sub-populations.

In Fig. 11, high average payoff was not obtained from the smallest neighborhood structure $N_{GA}(i)$ with five neighbors for local parent selection. This observation can be explained as follows. From the second rule (i.e., Rule 2), two parents of a new strategy for Agent i should be neighbors in $N_{GA}(i)$ in the same sub-population as Agent i. When the size of $N_{GA}(i)$ is five, the

number of qualified neighbors is five or less. It is possible that Agent i has only a single qualified neighbor (i.e., Agent i itself). In this case, we cannot use any crossover operation. As a result, a new strategy for Agent i is always generated by the mutation operation from its own strategy. In this case, local parent selection has no selection pressure towards strategies with high fitness values. Moreover, due to the linear scaling in (1), the same neighbor is selected as two parents even when Agent i has two neighbors in $N_{GA}(i)$. Thus no crossover is used to generate a new strategy for Agent i when $N_{GA}(i)$ does not include more than two neighbors.

The four plots in the top row of Fig. 11 (i.e., Fig. 11 (a)-(d)) show that the average payoff of 50% 3-bit strategy agents was decreased by the other 50% agents with stochastic strategies. The same observation is obtained with respect to the average payoff of 50% 5-bit strategy agents from the four plots in the second row (i.e., Fig. 11 (e)-(h)).

In the third row (i.e., Fig. 11 (i)-(l)), the average payoff of 50% stochastic length 3 real number strategy agents was increased by the other 50% agents with deterministic strategies. We have the same observation in the bottom row (i.e., Fig. 11 (m)-(p)) with respect to the average payoff of 50% stochastic length 5 real number strategy agents.

We can obtain other interesting observations from more careful examination of Fig. 11. Let us focus on the case with a mixture of 50% 3-bit strategy agents and 50% 5-bit strategy agents. Experimental results for this case are Fig. 11 (b) for 3-bit strategy agents and Fig. 11 (e) for 5-bit strategy agents. The average payoff in these plots is lower than the homogeneous case in Fig. 2. For comparison, we performed computational experiments after changing the first rule as follows:

Rule 1': Local opponent selection is performed within each sub-population. That is, Agent i can play the IPD game against a neighbor in $N_{IPD}(i)$ only when they are in the same sub-population.

Experimental results are shown in Fig. 12 where a mixture of 50% 3-bit strategy agents and 50% 5-bit strategy agents were used. Since no IPD game was performed between 3-bit and 5-bit strategy agents, there was no interaction between them. That is, the evolution of cooperative behavior was performed independently in each sub-population. From Fig. 11 (b), Fig. 11 (e) and Fig. 12, we can see that the IPD game between 3-bit and 5-bit strategy agents increased the average payoff of agents in each sub-population. We performed the Wilcoxon signed-ranks test to compare the average payoff between Fig. 11 (b) and Fig. 12 (a). The p-value was calculated as being less than 0.0001. We also performed the same test to compare the average payoff between Fig. 11 (e) and Fig. 12 (b). The p-value was also calculated as being less than 0.0001. Thus we can say that the increase in the average payoff by the interaction through the IPD game from Fig. 12 to Fig. 11 (b) and Fig. 11 (e) is statistically significant.

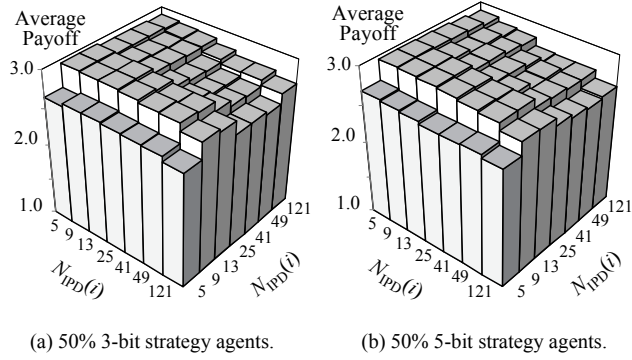(a) 50% 3-bit strategy agents.   (b) 50% 5-bit strategy agents.

Fig. 12. Average payoff of 50% 3-bit and 50% 5-bit strategy agents in the 11×11 grid-world with no IPD game between 3-bit and 5-bit strategy agents.

In Fig. 12, the evolution of cooperative behavior was performed independently within each sub-population. Thus the population size can be viewed as the half of the 11×11 grid-world. We examined the effect of population size by performing computational experiments with homogeneous agents in the 8×8 grid-world, which is about a half of the 11×11 grid-world. Experimental results are summarized in Fig. 13. It should be noted that Fig. 13 was obtained from the same computation experiments as Fig. 2 except for the size of the grid-world. The use of the smaller grid-world in Fig. 13 decreased the average payoff in Fig. 2 with the larger grid-world (compare Fig. 2 with Fig. 13). We can also see that similar results were obtained in Fig. 12 and Fig. 13 except for the leftmost row with five neighbors in $N_{GA}(i)$. These observations suggest that the population size has a large effect on the evolution of cooperative behavior. That is, the decrease in the average payoff in Fig. 12 from Fig. 2 is partially explained by the decrease in the population size. Another reason is the decrease in the number of neighbors in $N_{GA}(i)$ especially when the size of $N_{GA}(i)$ was five.

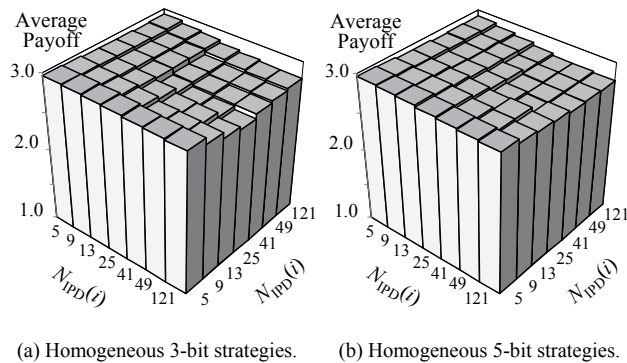(a) Homogeneous 3-bit strategies.   (b) Homogeneous 5-bit strategies.

Fig. 13. Average payoff of 3-bit and 5-bit strategies in the homogeneous situation in the 8×8 grid-world. Except for the size of the grid-world, all conditions are the same as in Fig. 2.

We performed the Wilcoxon signed-ranks test for the pair-wise comparison among Fig. 2, Fig. 12 and Fig. 13. We used the average payoff values for the following 5×6 combinations of the two neighborhood structures in the Wilcoxon signed-ranks test: $N_{GA}(i)$ with 9, 13, 25, 41, 49 neighbors, and $N_{IPD}(i)$ with 5, 9, 13, 25, 41, 49 neighbors. Test results are summarized in Table IV where the *p*-value for each pair-wise comparison is shown. From Table IV, we can say that Fig. 2 is significantly different from Fig. 12 and Fig. 13. However, the difference is not statistically significant between Fig. 12 and Fig. 13.

TABLE IV
THE *P*-VALUES BY THE WILCOXON SIGNED-RANKS TEST

	3-bit strategies in (a)	5-bit strategies in (b)
Fig. 2 and Fig. 12	less than 0.0001	less than 0.0001
Fig. 2 and Fig. 13	less than 0.0001	less than 0.0001
Fig. 12 and Fig. 13	0.393	0.517

Let us focus on another combination of two representation schemes. In Fig. 11 (c) and Fig. 11 (i), we can see that similar results were obtained from 50% 3-bit strategy agents and 50% stochastic length 3 real number strategy agents when they were used as two sub-populations. This is an interesting observation since totally different results were obtained from these two representation schemes in the case of homogeneous agents (compare Fig. 2 (a) with Fig. 3 (a), and also compare Fig. 11 (a) with Fig. 11 (k)).

The similarity in the average payoff between Fig. 11 (c) and Fig. 11 (i) is more clearly demonstrated in Fig. 14 and Fig. 15 where we show the average payoff for each of the four extreme neighborhood combinations. For comparison, we also show experimental results in Fig. 16 and Fig. 17 for the case of no interaction (i.e., no execution of the IPD game between 50% 3-bit strategy agents and 50% stochastic length 3 real number strategy agents based on Rule 1' as in Fig. 12). Figs. 14-17 suggest that the similarity between Fig. 14 and Fig. 15 in the evolution of cooperative behavior was realized by the interaction through the IPD game between heterogeneous agents with different representation schemes.

From the comparison between Fig. 15 and Fig. 17, we can see that the average payoff of stochastic length 3 real number strategy agents was increased by the interaction through the IPD game against 3-bit strategy agents in Fig. 15. The interaction through the IPD game also increased the average payoff of 3-bit strategy agents in Fig. 14 in the later generations when the size of $N_{IPD}(i)$ for local opponent selection was 5 (compare the average payoff around the 1000th generation between Fig. 14 and Fig. 16).

We also examined the percentage of each 3-bit strategy in Fig. 14 and Fig. 16. Since the best results were obtained when $|N_{GA}(i)| = 121$ and $|N_{IPD}(i)| = 5$ in Fig. 14 and Fig. 16 (i.e., dotted lines), we show the average percentage of each 3-bit

strategy under this setting in Fig. 18 and Fig. 19. In these figures, the following strategies are examined: 000 (ALLD), 100 (Cooperation only in the first round), 101 (TFT), and 111 (ALLC). It should be noted that the total percentage of 3-bit strategies was 50% (rather than 100%) since the other 50% were stochastic length 3 real number strategies.
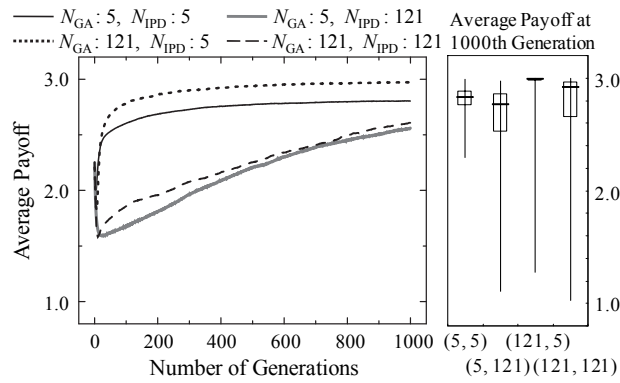

Fig. 14. Results by 50% 3-bit strategy agents with the IPD game against 50% stochastic length 3 real number strategy agents.
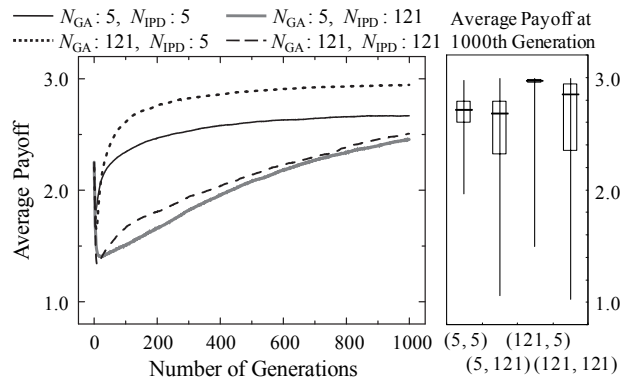

Fig. 15. Results by 50% stochastic length 3 real number strategy agents with the IPD game against 50% 3-bit strategy agents.
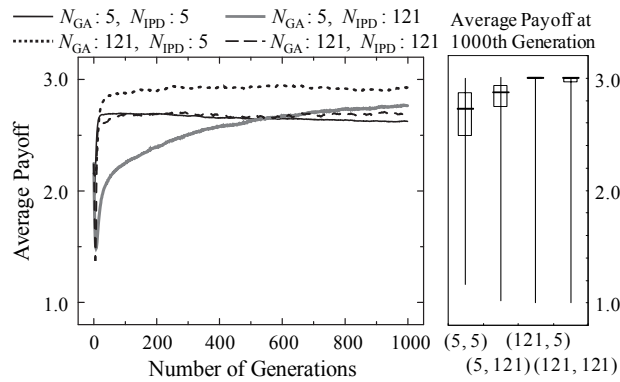

Fig. 16. Results by 50% 3-bit binary strategy agents with no IPD game against 50% stochastic length 3 real number strategy agents.
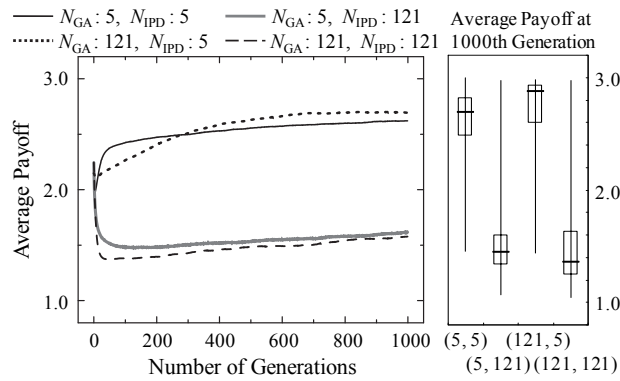
Fig. 17. Results by 50% stochastic length 3 real number strategy agents with no IPD game against 50% 3-bit binary strategy agents.

In Fig. 18, deterministic 3-bit strategy agents played the IPD game against not only deterministic 3-bit but also stochastic length 3 strategy agents. Deterministic 3-bit strategy agents in Fig. 19, however, did not play the IPD game against stochastic length 3 strategy agents. Since all the other conditions are the same between Fig. 18 and Fig. 19, the difference between these two figures is due to the difference in the interaction with other agents through the IPD game.
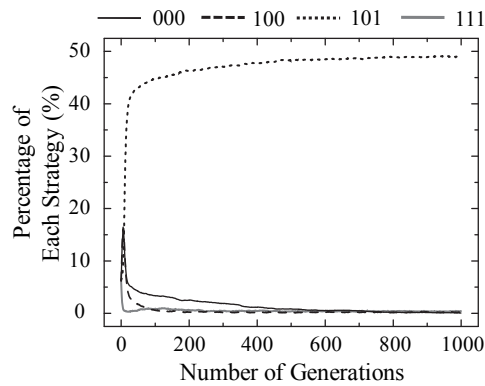

Fig. 18. Percentage of each strategy in a mixture of 50% deterministic 3-bit strategy agents and 50% stochastic length 3 real number strategy agents (with the IPD game between deterministic and stochastic strategy agents).
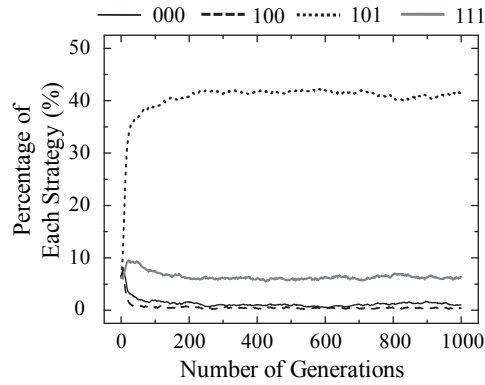
Fig. 19. Experimental results with no IPD game between deterministic and stochastic strategy agents (all the other conditions are the same as Fig. 18).

In order to explain the behavior of each strategy in Fig. 18, let us calculate the expected average payoff of TFT "101", ALLC "111" and ALLD "000" from the IPD game against stochastic strategies in the initial population. Since initial stochastic strategies can be viewed as randomly choosing D or C on average, the action by TFT can be viewed as being random (except for the first round). Thus the expected average payoff of TFT can be calculated as about 2.25 (as explained for Fig. 5). The average expected payoff of ALLC and ALLD can be also calculated as 1.5 and 3.0, respectively. These values explain the sharp increase of ALLD in the first few generations in Fig. 18. At the same time, the percentage of TFT also increased in Fig. 18. The increase of TFT and ALLD decreases the expected average payoff of ALLD. Thus the percentage of ALLD gradually decreased in Fig. 18.

In Fig. 20, we show experimental results of a single run in detail (one out of the 1000 runs in Fig. 18). As we have just explained, ALLD increased in the first few generations. Then all the deterministic 3-bit strategies converged to TFT soon.

In Fig. 19 with no interaction between deterministic and stochastic strategy agents, ALLC remained even after 1000 generations together with TFT. The existence of the ALLC may give a chance of survival to ALLD. As a result, 0.99% agents adopt ALLD in Fig. 19 at the 1000th generation whereas 0.05% agents adopt ALLD in Fig. 18.

Fig. 19 is average results over 1000 runs. In many runs, almost all deterministic 3-bit strategies converged to TFT. However, they converged to ALLC in some runs as shown in Fig. 21 and to ALLD in a few other runs as in Fig. 22.
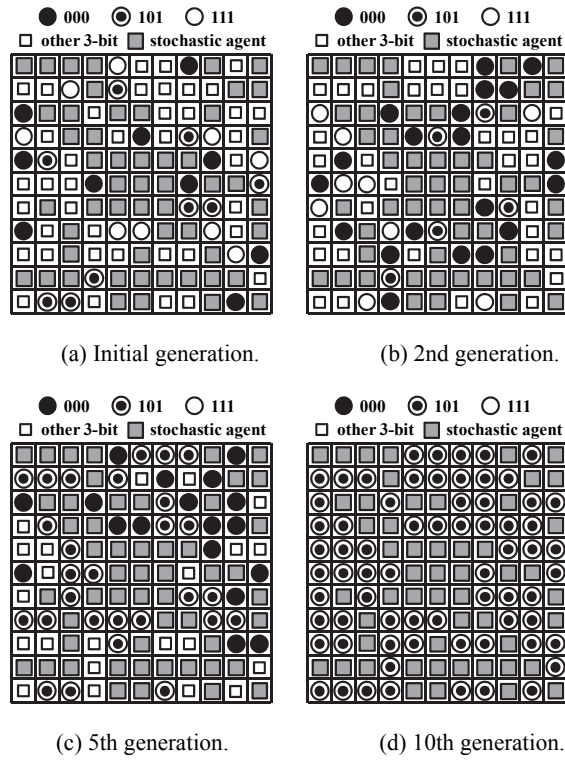
(a) Initial generation.　　(b) 2nd generation.

(c) 5th generation.　　(d) 10th generation.

Fig. 20. Experimental results of a single run in Fig. 18 with the IPD game between deterministic and stochastic strategy agents.

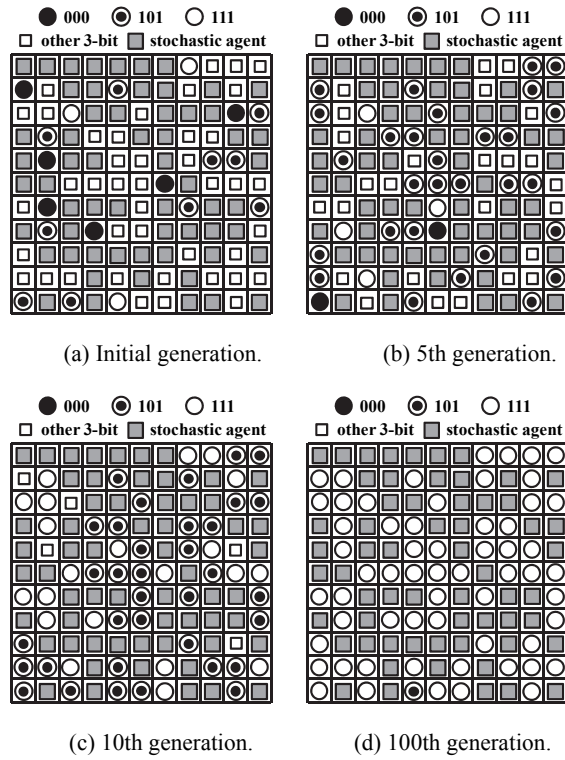(a) Initial generation.　　(b) 5th generation.

(c) 10th generation.　　(d) 100th generation.

Fig. 21. Experimental results of a single run in Fig. 19 without the IPD game between deterministic and stochastic strategy agents.

(a) Initial generation.   (b) 5th generation.
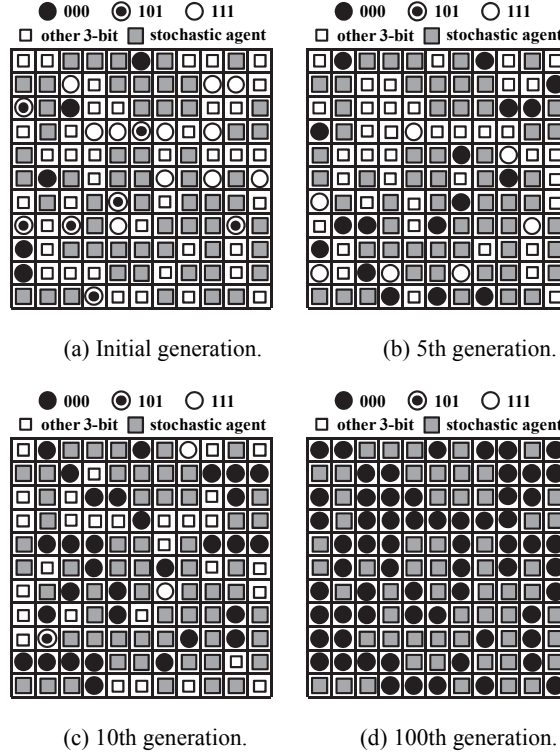
(c) 10th generation.   (d) 100th generation.

Fig. 22. Experimental results of another run in Fig. 19 without the IPD game between deterministic and stochastic strategy agents.

D. Use of Sub-Populations with Different Size

In the previous subsection, the 121 cells were randomly divided into two sub-populations of almost the same size. We also examined other settings as shown in Fig. 23 and Fig. 24. Fig. 23 used a mixture of 75% deterministic 3-bit strategy agents and 25% stochastic length 3 strategy agents whereas Fig. 24 used 25% deterministic 3-bit strategy agents and 75% stochastic length 3 strategy agents. The IPD game was played between agents independent of their representation schemes.

From the comparison between Fig. 23 and Fig. 24, we can see that the increase in the percentage of deterministic 3-bit strategy agents from 25% to 75% increased not only their own average payoff from Fig. 24 (a) to Fig. 23 (a) but also the average payoff of stochastic strategy agents from Fig. 24 (b) to Fig. 23 (b). However, the increase in the percentage of stochastic length 3 strategy agents from 25% to 75% decreased not only their own average payoff from Fig. 23 (b) to Fig. 24 (b) but also the average payoff of deterministic 3-bit strategy agents from Fig. 23 (a) to Fig. 24 (a). It is interesting to observe that the two plots in Fig. 24 (and also in Fig. 23) have some similarity to each other as Fig. 11 (c) is similar to Fig. 11 (i).

(a) 75% deterministic 3-bit agents.   (b) 25% stochastic length 3 agents.
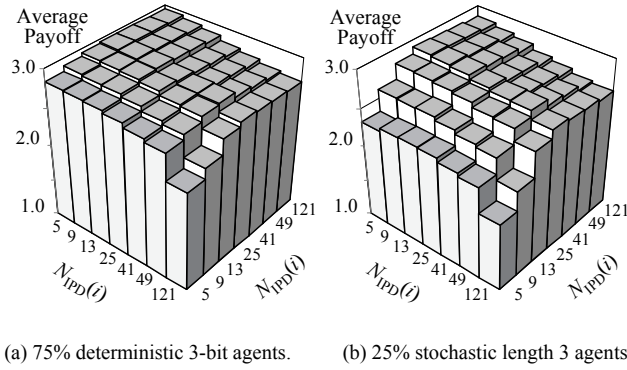
Fig. 23. Average payoff of 75% deterministic 3-bit strategy agents and 25% stochastic length 3 real number strategy agents in the 11×11 grid-world. (With the IPD game between agents with different representation schemes).

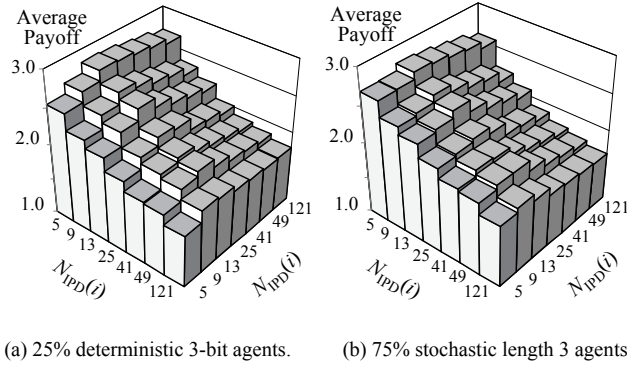(a) 25% deterministic 3-bit agents.   (b) 75% stochastic length 3 agents.

Fig. 24. Average payoff of 25% deterministic 3-bit strategy agents and 75% stochastic length 3 real number strategy agents in the 11×11 grid-world. (With the IPD game between agents with different representation schemes).

E. Effects of Interaction through Crossover and IPD Game

To examine the effects of the two types of interaction between sub-populations (i.e., crossover and the IPD game) on the evolution of cooperative behavior, we performed computational experiments under the following four settings with respect to the interaction between two sub-populations:

(i) No interaction.

(ii) Partial interaction only through the IPD game.

(iii) Partial interaction only through crossover.

(iv) Full interaction through the IPD game and crossover.

We used deterministic and stochastic length 3 real number strategies since they can be recombined. We examined the following three cases of agent assignments: (1) A mixture of 50% deterministic and 50% stochastic strategy agents, (2) 100%

deterministic strategy agents divided into two sub-populations of the same size, and (3) 100% stochastic strategy agents divided into two sub-populations of the same size. These three cases were examined under each of the above-mentioned four settings with respect to the interaction between two sub-populations.

Experimental results are summarized in Figs. 25-28. Each plot in these figures shows the average payoff of 50% row agents in a mixture with 50% column agents. In each of the two diagonal plots (a) and (d) in each figure, the same representation scheme was used in the two sub-populations. Thus those plots show the effects of crossover and the IPD game on the evolution of cooperative behavior among homogeneous agents. The other two off-diagonal plots in each figure show their effects among heterogeneous agents.

In Fig. 25 with no interaction between sub-populations, the two plots in each row are the same because the column agents had no effects on the average payoff of row agents.
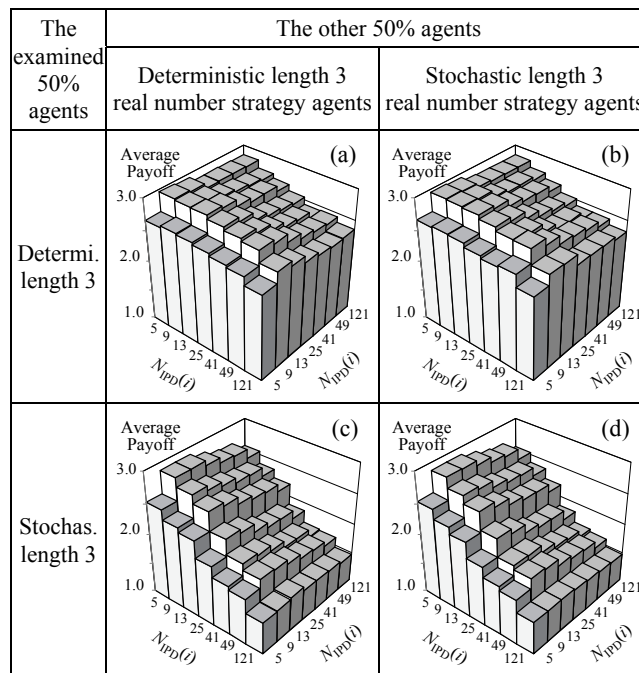

The examined 50% agents	The other 50% agents	
	Deterministic length 3 real number strategy agents	Stochastic length 3 real number strategy agents
Determi. length 3	(a)	(b)
Stochas. length 3	(c)	(d)

Fig. 25. No interaction between two sub-populations.

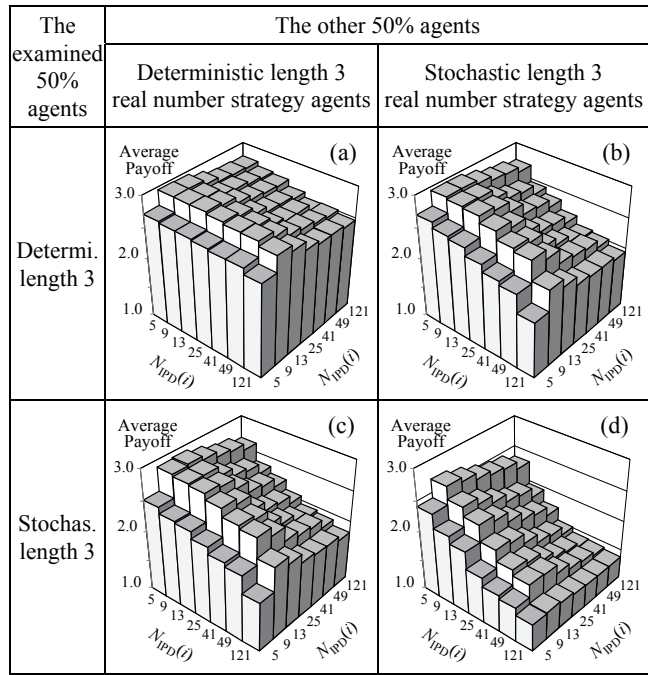| The examined 50% agents	The other 50% agents	
	Deterministic length 3 real number strategy agents	Stochastic length 3 real number strategy agents
Determi. length 3	(a)	(b)
Stochas. length 3	(c)	(d)

Fig. 26. Partial interaction between two sub-populations through the IPD game (no crossover between two sub-populations).
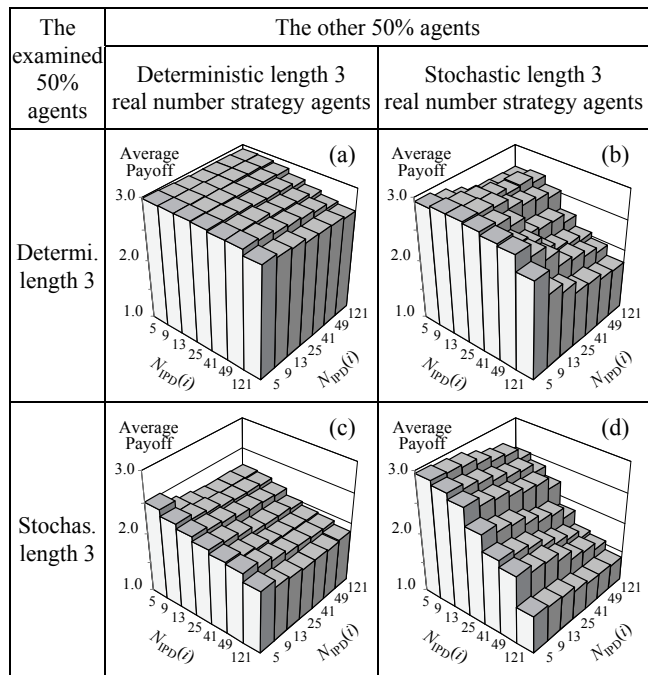
The examined 50% agents	The other 50% agents	
	Deterministic length 3 real number strategy agents	Stochastic length 3 real number strategy agents
Determi. length 3	(a)	(b)
Stochas. length 3	(c)	(d)

Fig. 27. Partial interaction between two sub-populations through crossover (no IPD game between two sub-populations).

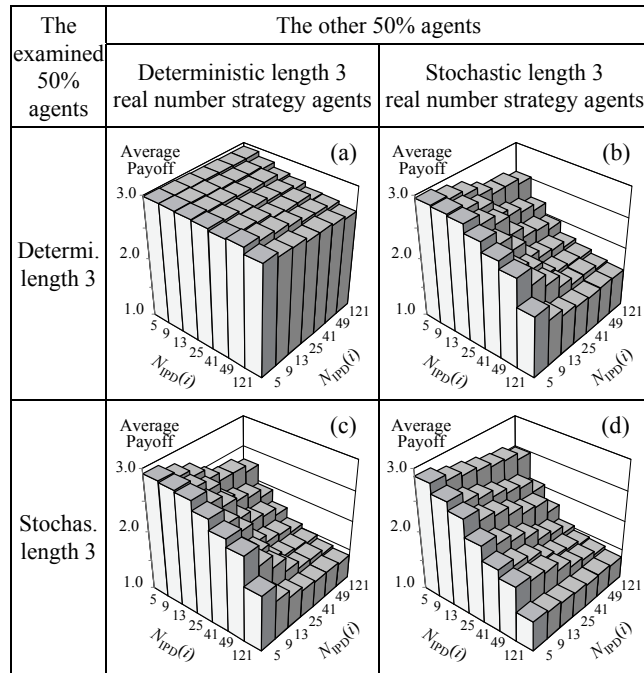| The examined 50% agents	The other 50% agents	
	Deterministic length 3 real number strategy agents	Stochastic length 3 real number strategy agents
Determi. length 3	(a)	(b)
Stochas. length 3	(c)	(d)

Fig. 28. Full interaction between two sub-populations through the IPD game and crossover.

In Fig. 26 with the interaction only through the IPD game (no crossover), similar results were obtained from 50% deterministic and 50% stochastic strategy agents in the two off-diagonal plots (b) and (c). This case is further examined in Fig. 29 and Fig. 30. From the comparison between Fig. 29 and Fig. 30, we can see that similar results were obtained from 50% deterministic and 50% stochastic strategy agents when they interacted with each other through the IPD game.

In Fig. 27, there was the interaction through only crossover. Under this setting, totally different results were obtained from 50% deterministic and 50% stochastic strategy agents in the two off-diagonal plots (b) and (c). In Fig. 28 with full interactions through crossover and the IPD game, similar results were obtained from 50% deterministic and 50% stochastic strategy agents in the two off-diagonal plots (b) and (c). From Figs. 25-28, we can see that similar results were obtained from different representation schemes when they interacted with each other through the IPD game.
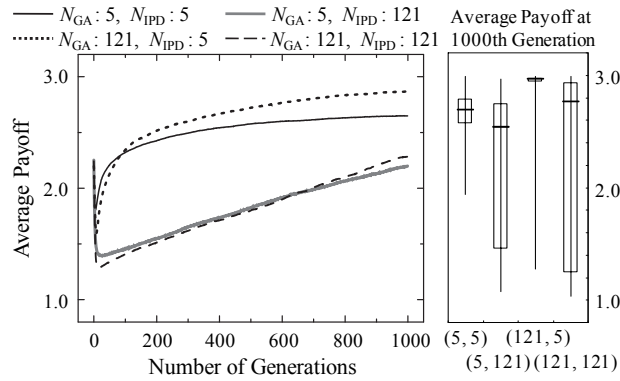
Fig. 29. Results by 50% stochastic length 3 strategy agents with the interaction through the IPD game against 50% deterministic length 3 strategy agents (with no crossover between deterministic and stochastic strategies).
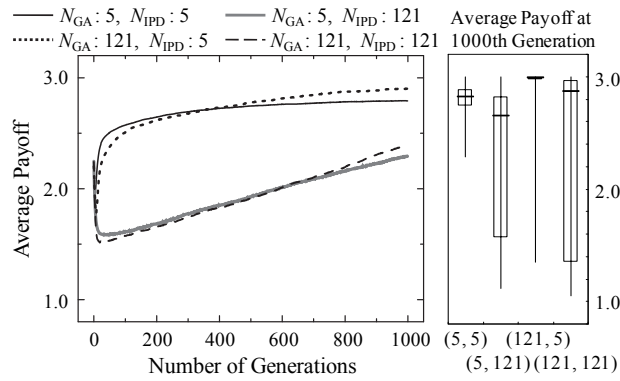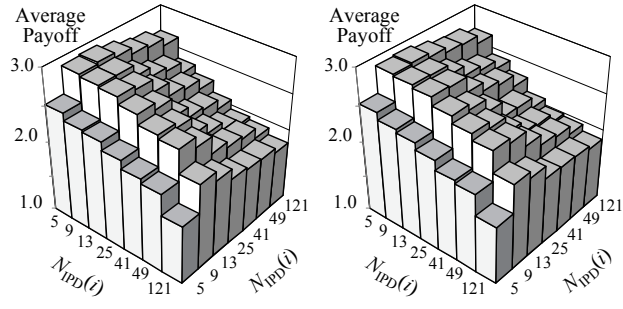

Fig. 30. Results by 50% deterministic length 3 strategy agents with the interaction through the IPD game against 50% stochastic length 3 strategy agents (with no crossover between deterministic and stochastic strategies).

### F. Sensitivity of Results to the Setting of Experiments

Any experimental results of computational experiments usually depend on their setting. In order to examine the sensitivity of our experimental results in this paper to the setting of computational experiments, we show experimental results obtained from different settings. We used a mixture of 50% stochastic and 50% deterministic length 3 strategy agents with the interaction only through the IPD game.

In Figs. 31-34, experimental results with different settings are compared. Fig. 31 shows experimental results with our basic setting. That is, Fig. 31 (a) and Fig. 31 (b) are the same as Fig. 26 (c) and Fig. 26 (b), respectively. In Fig. 32, we used the 5×24 grid-world instead of 11×11. In Fig. 33, we used elitism with a single elite individual instead of no elitism in our basic setting. In Fig. 34, binary tournament selection instead of roulette wheel selection was used.
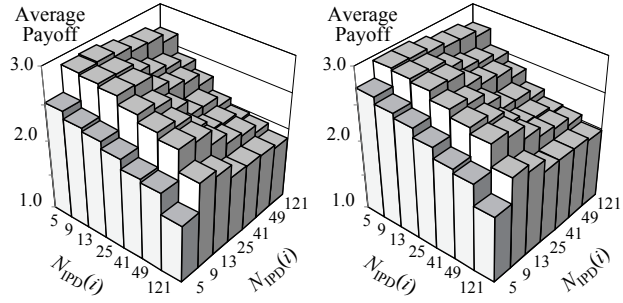
(a) 50% stochastic length 3 agents.   (b) 50% deterministic length 3 agents.

Fig. 31. Experimental results using our basic setting (i.e., the 11×11 grid-world, no elite individual and the roulette wheel selection scheme).
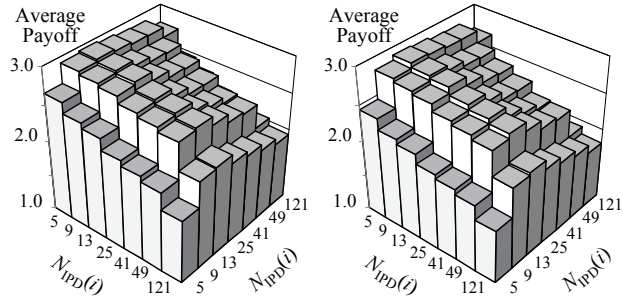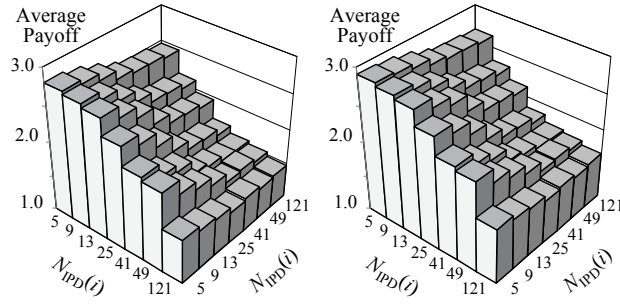

(a) 50% stochastic length 3 agents.   (b) 50% deterministic length 3 agents.

Fig. 32. Experimental results using the 5×24 grid-world.

(a) 50% stochastic length 3 agents.   (b) 50% deterministic length 3 agents.

Fig. 33. Experimental results using the elitism with a single elite individual.

(a) 50% stochastic length 3 agents.   (b) 50% deterministic length 3 agents.

Fig. 34. Experimental results using binary tournament selection.

We can see that experimental results in Fig. 32 and Fig. 33 are similar to Fig. 31. Moreover, the two plots in each figure are similar to each other in Figs. 31-34. The most prominent difference among Figs. 31-34 is high average payoff in Fig. 34 from the case of $|N_{GA}(i)| = 5$. In our computational experiments except for Fig. 34, we always used roulette wheel selection with the linear scaling in (1). Thus the same neighbor is selected as two parents when $N_{GA}(i)$ includes only one or two neighbors from the same sub-population as Agent i. In this case, no crossover operation is used to generate a new strategy for Agent i.

In Fig. 34 with binary tournament selection with replacement, different parents can be selected even when $N_{GA}(i)$ includes only two neighbors from the same sub-population as Agent i. Whereas higher average payoff was obtained in Fig. 34 than Figs. 31-33 in the case of $|N_{GA}(i)| = 5$, the average payoff in Fig. 34 was not higher than Figs. 31-33 for other settings with $|N_{GA}(i)| > 5$.

Finally we examined the effect of mutation in Figs. 35-37. In Fig. 35, the mutation probability was specified as 0. The average payoff in Fig. 35 with no mutation was clearly decreased from Fig. 31 with mutation. In Fig. 36, the mutation probability was specified as $5/(121 \times 5)$. This specification is five times larger than the basic setting $1/(121 \times 5)$. Higher average payoff was obtained in Fig. 36 with the larger mutation probability than Fig. 31 with the basic setting. The use of a too large mutation probability, however, decreased the average payoff as shown in Fig. 37 where the mutation probability was specified as $20/(121 \times 5)$.

As shown in Figs. 35-37, we can see that the specification of the mutation probability has large effects on the average payoff in our computational experiments. However, we can still obtain the same observations from Figs. 35-37 as Figs. 31-34. For example, the two plots in each of Figs. 35-37 are similar to each other. That is, similar results were obtained from different representation schemes when we have the interaction through the IPD game between them. We can also observe from Figs. 35-37 that the size of the neighborhood structure $N_{IPD}(i)$ for local opponent selection has much larger effects on the average

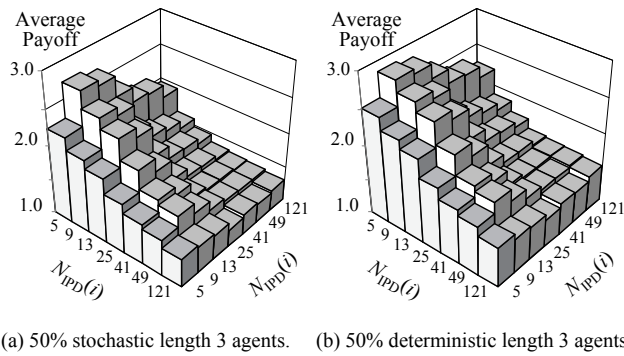payoff than the size of $N_{GA}(i)$ for local parent selection.

(a) 50% stochastic length 3 agents.   (b) 50% deterministic length 3 agents.

Fig. 35. Experimental results with no mutation (i.e., the mutation probability was specified as 0).

(a) 50% stochastic length 3 agents.   (b) 50% deterministic length 3 agents.

Fig. 36. Experimental results with a large mutation probability (i.e., the mutation probability was specified as 5/(121 × 5)).

(a) 50% stochastic length 3 agents.   (b) 50% deterministic length 3 agents.
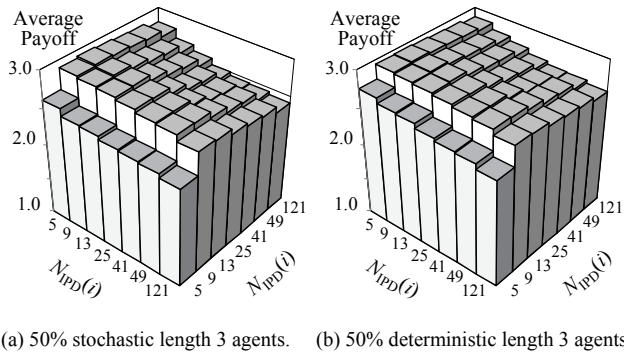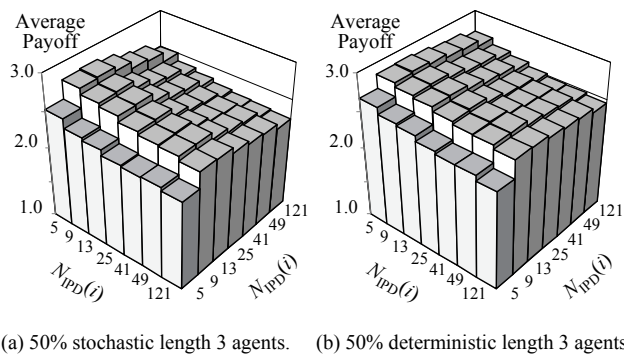
Fig. 37. Experimental results with a too large mutation probability (i.e., the mutation probability was specified as 20/(121 × 5)).

V. CONCLUDING REMARKS

We discussed the evolution of cooperative behavior in a spatial IPD game. Our model for game strategy evolution in the spatial IPD game has two characteristic features: One is the use of different neighborhood structures for local opponent selection and local parent selection. The other is the use of heterogeneous agents with different representation schemes in a single population. These two characteristic features of our model make it possible to examine various new aspects related to the evolution of cooperative behavior in the spatial IPD game. In this paper, we obtained the following observations from our computational experiments:

(1) The use of two representation schemes in a single population often decreased their average payoff from the homogeneous case with only a single representation scheme in a population.

(2) When each of two representation schemes was used by 50% agents in a single population, their experimental results were very similar to each other. This was the case even when their experimental results were totally different from each other in their homogeneous experiments with only a single representation scheme in a population.

(3) The interaction through the IPD game between different representation schemes had a large effect on their experimental results. Without the interaction through the IPD game, similar results were not obtained from different representation schemes.

(4) The effect of the interaction through crossover on the evolution of cooperative behavior was different from that through the IPD game.

(5) Whereas the use of different representation schemes in a single population often degraded the average payoff from the homogeneous case of 100% agents with the same representation scheme, the interaction between different representation schemes through the IPD game helped the evolution of cooperative behavior in some experiments. This observation was obtained from the comparison between the two settings: One is with the IPD game and the other is without IPD game between agents with different representation schemes.

In our computational experiments, we used the six representation schemes. They are different but similar to each other. One future research issue is the examination of the use of a wide variety of representation schemes such as a neural network and a decision tree in a single population. Whereas we examined a mixture of two representation schemes (i.e., two sub-populations) in this paper, it is possible to examine a mixture of more than two representation schemes (i.e., more than two sub-populations). The location of agents with the same representation scheme may have large effects on the evolution of cooperative behavior. Whereas we randomly divided agents into different sub-populations, it may be more realistic to assume that agents with the same representation schemes are closely located in a gird-world. The use of much larger grid-worlds such as 21×21 and

41×41 is another future research issue. In such a large grid-world, we will be able to examine more neighborhood structures. It would be also interesting to change the number of agents with each representation scheme during the evolution of cooperative behavior depending on the average payoff over agents with the same representation scheme. As in [24], the use of different representation schemes in a single population can be examined not only for the evolution of cooperative behavior in IPD games but also for other application areas of evolutionary computation such as optimization and genetics-based machine learning.

REFERENCES

[1] R. Axelrod, "The evolution of strategies in the iterated prisoner's dilemma," in L. Davis (ed.), *Genetic Algorithms and Simulated Annealing*, Morgan Kaufmann, pp. 32-41, 1987.

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

[3] D. B. Fogel, "Evolving behaviors in the iterated prisoner's dilemma," *Evolutionary Computation*, vol. 1, no. 1, pp. 77-97, 1993.

[4] P. Darwen and X. Yao, "Automatic modularisation by speciation," *Proc. of 3rd IEEE International Conference on Evolutionary Computation*, pp. 88-93, Nagoya, Japan, May 1996.

[5] P. H. Crowley, L. Provencher, S. Sloane, L. A. Dugatkin, B. Spohn, L. Rogers, and M. Alfieri, "Evolving cooperation: The role of individual recognition," *BioSystems*, vol. 37, no. 1, pp. 49-66, 1996.

[6] D. Ashlock, M. D. Smucker, E. A. Stanley, and L. Tesfatsion, "Preferential partner selection in an evolutionary study of prisoner's dilemma," *BioSystems*, vol. 37, no. 1, pp. 99-125, 1996.

[7] S. Bankes, "Exploring the foundations of artificial societies: Experiments in evolving solutions to iterated *N*-player prisoner's dilemma," in R. A. Brooks and P. Maes (eds.), *Artificial Life IV*, MIT Press, Cambridge, pp. 337-342, 1994.

[8] X. Yao and P. J. Darwen, "An experimental study of *N*-person iterated prisoner's dilemma games," *Informatica*, vol. 18, no. 4, pp. 435-450, 1994.

[9] Y. G. Seo, S. B. Cho, and X. Yao, "The impact of payoff function and local interaction on the *N*-player iterated prisoner's dilemma," *Knowledge and Information Systems*, vol. 2, no. 4, pp. 461-478, November 2000.

[10] M. A. Nowak, R. M. May, and K. Sigmund, "The arithmetics of mutual help," *Scientific American*, vol. 272, no, 6, pp. 76-81, June 1995.

[11] A. L. Lloyd, "Computing bouts of the prisoner's dilemma," *Scientific American*, pp. 80-83, 1995.

[12] M. Oliphant, "Evolving cooperation in the non-iterated prisoner's dilemma: The importance of spatial organization," in R. A. Brooks and P. Maes (eds.), *Artificial Life IV*, MIT Press, Cambridge, pp. 349-352, 1994.

[13] P. Grim, "Spatialization and greater generosity in the stochastic prisoner's dilemma," *BioSystems*, vol. 37, no. 1, pp. 3-17, 1996.

[14] K. Brauchli, T. Killingback, and M. Doebeli, "Evolution of cooperation in spatially structured populations," *Journal of Theoretical Biology*, vol. 200, no. 4, pp. 405-417, October 1999.

[15] H. Ishibuchi and N. Namikawa, "Evolution of iterated prisoner's dilemma game strategies in structured demes under random pairing in game playing," *IEEE Trans. on Evolutionary Computation*, vol. 9, no. 6, pp. 552-561, December 2005.

[16] S. Mittal and K. Deb, "Optimal strategies of the iterated prisoner's dilemma problem for multiple conflicting objectives," *IEEE Trans. on Evolutionary Computation*, vol. 13, no. 3, pp. 554-565, June 2009.

[17] S. Y. Chong and X. Yao, "Multiple choices and reputation in multi-agent interactions," *IEEE Trans. on Evolutionary Computation*, vol. 11, no. 6, pp. 689-711, December 2007.

[18] S. Y. Chong and X. Yao, "Behavioral diversity, choices and noise in the iterated prisoner's dilemma," *IEEE Trans. on Evolutionary Computation*, vol. 9, no. 6, pp. 540-551, December 2005.

[19] L. A. Dugatkin, *Cooperation among Animals - An Evolutionary Perspective*, Oxford University Press, New York, 1997.

[20] G. Kendall, X. Yao, and S. Y. Chong (eds.), *The Iterated Prisoners' Dilemma: 20 Years*, World Scientific, Singapore, 2007.

[21] D. Ashlock, E. Y. Kim, and N. Leahy, "Understanding representational sensitivity in the iterated prisoner's dilemma with fingerprints," *IEEE Trans. on Systems, Man, and Cybernetics: Part C*, vol. 36, no. 4, pp. 464-475, July 2006.

[22] D. Ashlock and E. Y. Kim, "Fingerprinting: Visualization and automatic analysis of prisoner's dilemma strategies," *IEEE Trans. on Evolutionary Computation*, vol. 12, no. 5, pp. 647-659, October 2008.

[23] D. Ashlock, E. Y. Kim, and W. Ashlock, "Fingerprint analysis of the noisy prisoner's dilemma using a finite-state representation," *IEEE Trans. on Computational Intelligence and AI in Games*, vol. 1, no. 2, pp. 154-167, June 2009.

[24] Z. Skolicki and K. De Jong, "Improving evolutionary algorithms with multi-representation island models," *Lecture Notes in Computer Science 3242: Parallel Problem Solving from Nature - PPSN VIII*, pp. 420-429, Springer, Berlin, September 2004.

[25] D. S. Wilson, "Structured demes and the evolution of group-advantageous traits," *The American Naturalist*, vol. 111, no. 977, pp. 157-185, January-February 1977.

[26] D. S. Wilson, "Structured demes and trait-group variation," *The American Naturalist*, vol. 113, no. 4, pp. 606-610, April 1979.

[27] M. Slatkin and D. S. Wilson, "Coevolution in structured demes," *Proc. of the National Academy of Sciences*, vol. 76, no. 4, pp. 2084-2087, April 1979.

[28] B. Charlesworth, "A note on the evolution of altruism in structured demes," *The American Naturalist*, vol. 113, no. 4, pp. 601-605, April 1979.

[29] M. Ifti, T. Killingback, and M. Doebeli, "Effects of neighbourhood size and connectivity on the spatial continuous prisoner's dilemma," *Journal of Theoretical Biology*, vol. 231, no. 1, pp. 97-106, November 2004.

[30] H. Ishibuchi, T. Doi, and Y. Nojima, "Effects of using two neighborhood structures in cellular genetic algorithms for function optimization," *Lecture Notes in Computer Science 4193: Parallel Problem Solving from Nature - PPSN IX*, pp. 949-958, Springer, Berlin, September 2006.

[31] H. Ishibuchi, N. Tsukamoto, and Y. Nojima, "Examining the effect of elitism in cellular genetic algorithms using two neighborhood structures," *Lecture Notes in Computer Science 5199: Parallel Problem Solving from Nature - PPSN X*, pp. 458-467, Springer, Berlin, September 2008.

[32] H. Ohyanagi, Y. Wakamatsu, Y. Nakashima, Y. Nojima, and H. Ishibuchi, "Evolution of cooperative behavior among heterogeneous agents with different strategy representations in an iterated prisoner's dilemma game," *Artificial Life and Robotics*, vol. 14, no. 3, pp. 414-417, December 2009.

[33] H. Ishibuchi, H. Ohyanagi, and Y. Nojima, "Evolution of cooperative behavior in a spatial iterated prisoner's dilemma game with different representation schemes of game strategies," *Proc. of IEEE International Conference on Fuzzy Systems*, pp. 1568-1573, August 2009.

[34] L. J. Eshelman and J. D. Schaffer, "Real-coded genetic algorithms and interval-schemata," *Foundations of Genetic Algorithms 2*, Morgan Kaufman, San Mateo, pp. 187-202, 1993.

[35] D. J. Sheskin, Handbook of Parametric and Nonparametric Statistical Procedures (4th ed.), Chapman & Hall, Boca Raton, FL, 2

Hisao Ishibuchi (M'93–SM'10) received the B.S. and M.S. degrees in precision mechanics from Kyoto University, Kyoto, Japan, in 1985 and 1987, respectively, and the Ph.D. degree in computer science from Osaka Prefecture University, Sakai, Osaka, Japan, in 1992.

Since 1987, he has been with Osaka Prefecture University, where he was a Research Associate, an Assistant Professor, and an Associate Professor. He is currently a Professor with the Department of Computer Science and Intelligent Systems. His current research interests include evolutionary multiobjective optimization, evolutionary game, and multiobjective genetic fuzzy systems.

Dr. Ishibuchi received the Best Paper Award from the Genetic and Evolutionary Computation Conference in 2004, the IEEE International Conference on Fuzzy Systems in 2009, and the World Automation Congress in 2010. He also received the 2007 Japan Society for the Promotion of Science Prize. He is currently the IEEE Computational Intelligence Society Vice-President for technical activities for 2010–2011. He is also an Associate Editor for a number of international journals such as the IEEE TRANSACTIONS ON FUZZY SYSTEMS, the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, the IEEE TRANSACTIONS ON SYSTEMS, MAY AND CYBERNETICS–PART B, and the IEEE COMPUTATIONAL INTELLIGENCE MAGAZINE.

Hiroyuki Ohyanagi received the B.S. Degrees in computer science and intelligent systems from Osaka Prefecture University,

Osaka, Japan, in 2009. He is currently a master's course student in Department of Computer Science and Intelligent Systems,

Osaka Prefecture University. His research interests are iterated prisoner's dilemma game and evolutionary multiobjective

optimization.

Yusuke Nojima (M'00) received the B.S. and M.S. Degrees in mechanical engineering from Osaka Institute of Technology, Osaka, Japan, in 1999 and 2001, respectively, and the Ph.D. degree in system function science from Kobe University, Hyogo, Japan, in 2004.

Since 2004, he has been with Osaka Prefecture University, Osaka, Japan, where he was a Research Associate and is currently an Assistant Professor in Department of Computer Science and Intelligent Systems. His research interests include multiobjective genetic fuzzy systems, evolutionary multiobjective optimization, parallel distributed data mining, and ensemble classifier design.

Dr. Nojima received the Best Paper Award from the IEEE International Conference on Fuzzy Systems in 2009, and the World Automation Congress in 2010.