The Fourteenth International Symposium on Artificial Life and Robotics 2009 (AROB 14th '09),
B-Con Plaza, Beppu, Oita, Japan, February 5 - 7, 2009

Evolution of Cooperative Behavior among Heterogeneous Agents with Different Strategy Representations in an Iterated Prisoner's Dilemma Game

Hiroyuki Ohyanagi, Yoshihiko Wakamatsu, Yusuke Nakashima, Yusuke Nojima, and Hisao Ishibuchi

Osaka Prefecture University, 1-1 Gakuen-cho, Naka-ku, Sakai, Osaka 599-8531, Japan
(Phone: +81-72-254-9198; Fax: +81-72-254-9915)
({ohyanagi@ci., wakamatsu@ci., ysknksm@ci., nojima@, hisaoi@}cs.osakafu-u.ac.jp)

Abstract: The iterated prisoner's dilemma (IPD) game has been frequently used to examine the evolution of cooperative behavior among agents. When the effect of representation schemes of IPD game strategies was examined, the same representation scheme was usually assigned to all agents. That is, a population of homogeneous agents was usually used in computational experiments in the literature. In this paper, we focus on a slightly different situation where each agent does not necessarily use the same representation scheme. That is, a population can be a mixture of heterogeneous agents with different representation schemes. In computational experiments, we use binary strings of different length (i.e., three-bit and five-bit strings) for representing IPD game strategies. We examine the evolution of cooperative behavior among heterogeneous agents in comparison with the case of homogeneous ones for the standard IPD game with the typical payoff values 0, 1, 3 and 5. Experimental results show that the evolution of cooperative behavior is slowed down by the use of heterogeneous agents. It is also demonstrated that the faster evolution of cooperative behavior is achieved among majority agents than minority ones in a heterogeneous population.

Keywords: Iterated prisoner's dilemma (IPD) game, evolution of cooperative behavior, evolution of game strategies, genetic algorithms, representation, coding schemes.

I. INTRODUCTION

The evolution of cooperative behavior among agents in the iterated prisoner's dilemma (IPD) game has been discussed in many studies since the late 1980s [1] and the early 1990s [2], [3]. A player's strategy, which can be represented in various manners such as a binary string, a real-number string, a finite-state machine and a neural network, is evolved by selection, crossover and mutation in those studies. The fitness of a player in a population is defined by its average payoff obtained in iteratively playing the prisoner's dilemma game against other players in the same population. Various techniques and concepts have been introduced to the IPD game such as the speciation of strategies [4], individual recognition [5], and partner selection [6]. The IPD game has also been extended to various cases such as a multi-player version [7], [8], a spatial version [9], [10], stochastic strategies [11], [12], and random paring [13], [14]. See [15] for various studies on the evolution of cooperative behavior among agents in the IPD game.

Recently the IPD game has been used for examining the effect of the choice of a representation scheme on the evolution of game strategies [16], [17]. Those studies compared various representation schemes such as finite-state machines, cellularly encoded finite-state machines, feed-forward neural networks, if-skip-action lists, parse trees storing two types of Boolean functions, lookup tables, Boolean function stacks, and Markov chains. Experimental results showed that the choice of a representation scheme had a dominant effect on the evolution of game strategies.

When the effect of the choice of a representation scheme was examined, the same representation scheme was usually assigned to all agents in a population. That is, a population of homogeneous agents was usually used in computational experiments. In this paper, we focus on a slightly different situation where each agent does not necessarily use the same representation scheme. For example, some agents can use binary strings as their game strategies even when all the others use feed-forward neural networks. That is, a population can be a mixture of heterogeneous agents. Our aim is to show the effect of mixing different representation schemes.

II. IPD GAME AND GAME STRATEGIES

In this paper, we examine the evolution of cooperative behavior among heterogeneous agents in comparison with the case of homogeneous ones through computational experiments on the standard IPD game with the typical payoff values 0, 1, 3 and 5 (see Table 1).

The prisoner's dilemma game with the payoff matrix in Table 1 is played for a prespecified number of rounds (100 rounds in our computational experiments) between a pair of randomly selected agents from the current population. The random choice of two agents and the game playing between them are repeated in each generation until every agent plays the IPD game against a prespecified number of opponents (five opponents in our computational experiments). The fitness of each agent is defined by the average payoff per round obtained in the current population. New strategies for the next generation are generated by genetic operations.

We use three-bit and five-bit binary strings for representing IPD game strategies. Examples of those binary strings are shown in Table 2 and Table 3. Table 2 shows a three-bit binary string "101" which represents the so-call TFT (Tit-for-Tat) strategy. The same strategy is represented by a five-bit binary string in Table 3. One of these two representation schemes is assigned to each agent in our computational experiments.

Table 1. Payoff matrix of our IPD game.

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

Table 2. A three-bit binary string (TFT)

Player's First Action: Cooperate		1
Opponent's Action on the preceding Round	Suggested Action	
D: Defect	D: Defect	0
C: Cooperate	C: Cooperate	1

Table 3. A five-bit binary string (TFT)

Player'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

When we use both representation schemes, the current population is a mixture of three-bit and five-bit binary strings. The game playing between binary strings with different length involves no additional difficulties. Thus we assume no restriction on the choice of two agents for the game playing. That is, a pair of strings (i.e., agents) is randomly selected from the current population for the game playing with no restriction.

On the other hand, we always choose a pair of strings of the same length for crossover. That is, we use a mating restriction where binary strings of different length are never recombined. The current population can be viewed as two sub-populations: One with three-bit binary strings and the other with five-bit binary strings. Genetic operations are separately performed in each sub-population to generate the next sub-population. That is, the current population can be viewed as two separate sub-populations in the genetic operation phase whereas it is handled as a single population in the IPD game playing phase in our computational experiments.

In each sub-population, a pair of binary strings is selected based on the following selection probability:

$$P(s_i) = \frac{fitness(s_i) - f_{\min}(\Psi)}{\sum_{j \in \Psi} (fitness(s_j) - f_{\min}(\Psi))}, \qquad (1)$$

where s_i is the i-th string, $fitness(s_i)$ is the average payoff of s_i obtained by the IPD game in the current population, Ψ is a sub-population including s_i, and $f_{\min}(\Psi)$ is the minimum average payoff among strings in the sub-population Ψ. Eq.(1) is a standard roulette wheel selection with the linear scaling based on the minimum fitness value. It should be noted that the selection is separately performed in each sub-population.

We apply the standard one-point crossover operation to the selected pair of strings (with the probability 1.0 in our computational experiments). One of the generated two strings by the crossover operation is randomly chosen as an offspring. The standard bit-flip mutation operation is applied to the selected offspring (with the probability 0.002 per bit). By iterating the selection, crossover and mutation, we generate the same number of offspring as the sub-population size. The current sub-population is entirely replaced with the newly generated offspring. Thus the sub-population size is constant throughout the evolution of IPD game strategies.

III. COMPUTATIONAL EXPERIMENTS

We examined the following five situations in our computational experiments:

(1) Homogeneous case with 100% three-bit strings,
(2) Homogeneous case with 100% five-bit strings,
(3) 25% three-bit strings and 75% five-bit strings,
(4) 50% three-bit strings and 50% five-bit strings,
(5) 75% three-bit strings and 25% five-bit strings.

We used the following conditions in computational experiments in this paper:

[Overall computational experiment setting]
Number of runs: 1000 for each case.
[Genetic algorithm setting]
Population size: 100,
Initial strings: Randomly generated binary strings,
Selection: Roulette wheel selection in Eq.(1),
Crossover probability: 1.0 (One-point),
Mutation probability: 0.002 per bit (Bit-flip),
Generation gap: 100% (i.e., no elite individuals),
Termination condition: 1000 generations.
[IPD game setting]
Number of opponents: 5 (Randomly chosen),
Number of rounds: 100 (between the same agents).

First we compare the two homogeneous cases with each other in Fig. 1 where the average payoff at each generation is shown. This figure shows the effect of the choice of a representation scheme on the evolution of cooperative behavior. In Fig. 1, slightly faster evolution of cooperative behavior was achieved by shorter strings (i.e., by three-bit than five-bit). Since the representation schemes are very similar to each other, we obtained similar results in Fig. 1. The choice of a representative scheme, however, often has a dominant effect [16], [17]. We used these very similar representation schemes in order to highlight the effect of mixing them.

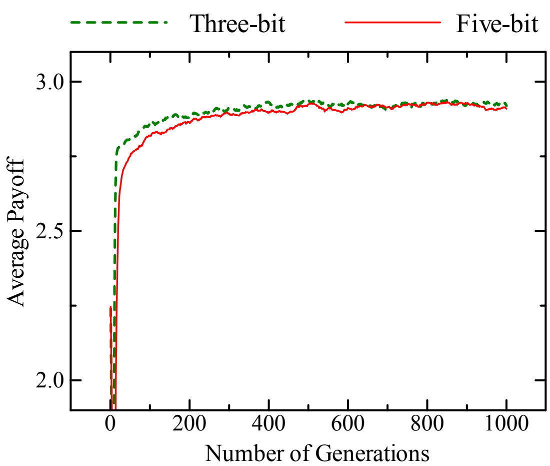

Fig. 1. Average payoff from homogeneous populations.

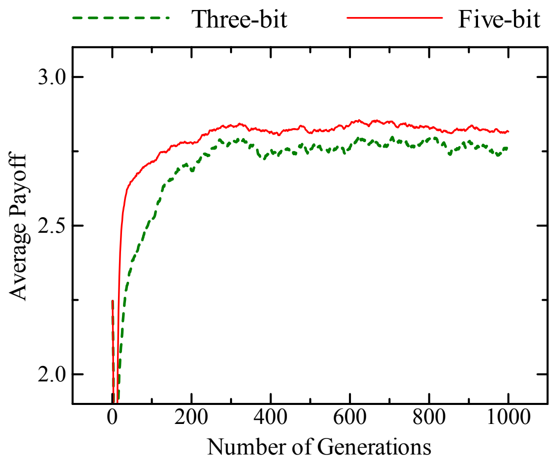

Fig. 2. Average payoff from inhomogeneous populations with 25% three-bit strings and 75% five-bit strings.

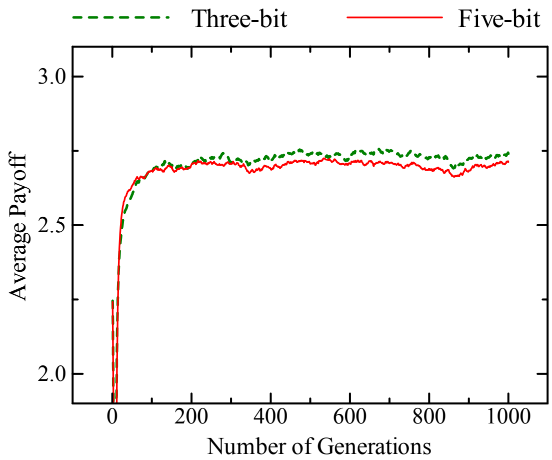

Fig. 3. Average payoff from inhomogeneous populations with 50% three-bit strings and 50% five-bit strings.

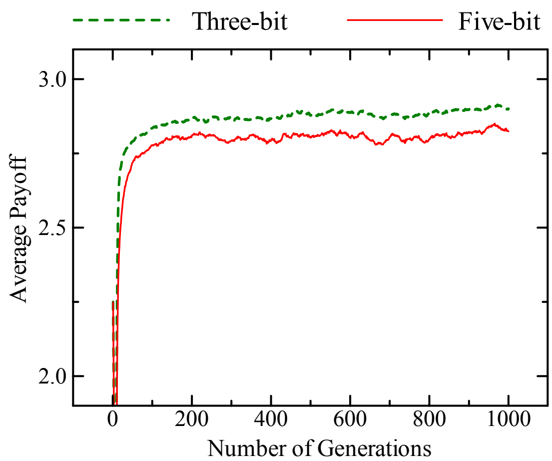

Fig. 4. Average payoff from inhomogeneous populations with 75% three-bit strings and 25% five-bit strings.

Experimental results on the three inhomogeneous cases are shown in Figs. 2-4. The average payoff was calculated in each sub-population in these figures. From the comparison of Fig. 1 with Figs. 2-4, we can see that cooperative behavior was more easily evolved among homogeneous agents in Fig. 1 than heterogeneous ones in Figs. 2-4. That is, mixing different representation schemes in a population slowed down the evolution of cooperative behavior. This negative effect was the most severe in Fig. 3 where the number of agents with each representation scheme was the same. An interesting observation is that better results were always obtained by majority agents (i.e., five-bit strings in Fig. 2 and three-bit strings in Fig. 4) when the number of agents with each representation scheme was different. Another interesting observation is that much better results were obtained in Fig. 4 than Fig. 2 whereas there was no large difference in Fig. 1 between the two schemes.

VI. CONCLUSION

In this paper, we examined the effect of mixing different representation schemes on the evolution of cooperative behavior in the IPD game. We used very similar representation schemes: three-bit and five-bit binary strings. We obtained similar results from these two representation schemes when they were separately used in homogeneous populations. Their simultaneous use in a single population, however, clearly slowed down the evolution of cooperative behavior. This negative effect of mixing different representation schemes affected the minority agents more severely. The worst results (i.e., the most severe negative effect) were obtained when the number of agents with each representation scheme was the same. As future research, we are planning to further examine the effect of mixing different representation schemes on the evolution of cooperative behavior in various situations such as using more than two types of agents and/or totally different representation schemes. We will also discuss potential positive effects of mixing different representation schemes on the evolution of IPD game strategies.

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, Los Altos, pp. 32-41, 1987.
[2] K. Lindgren, "Evolution phenomena in simple dynamics," in C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen (eds.), *Artificial Life II*, Addison-Wesley, Reading, 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. 237-242, 1994.
[8] X. Yao and P. Darwen, "An experimental study of *N*-person Iterated Prisoner's Dilemma games," *Informatica*, vol. 18, no. 4, pp. 435-450, 1994.
[9] M. A. Nowak, R. M. May, and K. Sigmund, "The arithmetics of mutual help," *Scientific American*, pp. 50-53, 1995.
[10] A. L. Lloyd, "Computing bouts of the Prisoner's Dilemma," *Scientific American*, pp. 80-83, 1995.
[11] P. Grim, "Spatialization and greater generosity in the stochastic Prisoner's Dilemma," *BioSystems*, vol. 37, no. 1, pp. 3-17, 1996.
[12] 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, 1999.
[13] H. Ishibuchi and N. Namikawa, "Evolution of cooperative behavior in the iterated prisoner's dilemma under random pairing in game playing," *Proc. of 2005 Congress on Evolutionary Computation*, pp. 2637-2644, Edinburgh, UK, 2005.
[14] 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, 2005.
[15] L. A. Dugatkin, *Cooperation among Animals - An Evolutionary Perspective*, Oxford University Press, New York, 1997.
[16] D. Ashlock, E. Y. Kim, and N. Leahy, "Understanding representational sensitivity in the iterated prisoner's dilemma with fingerprints," *IEEE Trams. on Systems, Man, and Cybernetics: Part C,* vol. 36, no. 4, pp. 464-475, 2006.
[17] 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, 2008.

