Effect of voluntary participation on an alternating and a simultaneous prisoner's dilemma

H. Yamamoto [*]

Faculty of Business Administration, Rissho University, Osaki 4-2-16, Shinagawa-ku, Tokyo 141-8602, Japan

I. Okada

Faculty of Business Administration, Soka University, Tangi 1-236, Hachioji City, Tokyo 192-8577, Japan

T. Taguchi and M. Muto

College of Systems Engineering and Science, Shibaura Institute of Technology, Fukasaku 307, Minuma-ku, Saitama City, Saitama, 337-8570, Japan

We studied the evolution of cooperation in the framework of evolutionary game theory, implementing voluntary participation in the prisoner's dilemma. Although previous studies have tried to overcome the dilemma by introducing voluntary participation called a "loner," the question of which strategies among various strategies including voluntary participation are adaptive under competitive circumstances is still an unsolved puzzle. Here we have developed a model that consists of all possible strategies using a one-period memory of past actions. This model enables us to analyze a "melting pot" of strategies, wherein several strategies interact and compete with each other. Our results revealed that one strategy, in which one escapes if a partner defects or cooperates if a partner becomes a loner, dominates and maintains cooperation in an alternating prisoner's dilemma game. However, the so-called "win-stay, lose-shift" strategy dominates in a simultaneous prisoner's dilemma game. Our simulations clearly show that voluntary participation in the prisoner's dilemma game works in the alternating situation rather than the simultaneous one.

I. INTRODUCTION

Cooperation in mutual competition is a basic mechanism for the prosperity of human society. However, the simplest model of cooperation in game theory predicts that cooperation will not emerge among rational people because cooperative behaviors incur costs to cooperators, and free riding is a better option. A previous study [1] showed that some mechanisms can promote cooperation; above all, direct reciprocity, e.g., "I will help you because you helped me in the past," is one of the simplest mechanisms to promote cooperation. Many studies have employed an iterated prisoner's dilemma game (PDG) to analyze direct reciprocity [2–9]. Among them, some papers have considered two forms of PDG, a simultaneous game and an alternating one. This is because, on the one hand, reciprocal cooperation in a real situation has a behavioral time lag [10], while, on the other hand, mutual help exists when the actions of two players are simultaneous [11]. A theoretical study [12] has shown the differences in the dominant strategies between alternating games and simultaneous ones.

In the simple setting of a PDG, players perform one of two actions: cooperation or defection. Hauert *et al.* [13,14] have proposed a third option in which a player does not participate in the game and instead falls back on a small income that does not depend on others. Those unwilling to join the game are termed "loners." Loners can foil defectors and overcome a social dilemma. The effects of loners have been widely analyzed using public goods games [13–19] and PDGs [20–23].

Although these previous studies focusing on loners considered three strategic types (a pure cooperator, a pure defector, and a pure loner), these papers assume that players use a one-period memory of past actions to decide their actions, as represented by the tit-for-tat strategy [7,12,24,25]. However, no study has tried to analyze a strategy space that contains the loner and the memory of past actions to capture these actions' effects on the evolution of cooperation in detail yet.

Many studies on reciprocal cooperation have analyzed strategies that are robust against an invasion of perfect defectors or perfect cooperators [26]. Perfect defectors can be defined as first-order free riders that do not contribute to public goods. Perfect cooperators are second-order free riders that avoid distinguishing between cooperation and defection. In our real society, however, various strategies interact with each other. Therefore, an imperative next step in studies on reciprocal cooperation with voluntary participation would be to develop an analytical tool that can handle a melting pot of multiple strategies that coexist, interact, and compete.

A pioneering work performed by Zagorosky *et al.* [27] exhaustively analyzed an alternating prisoner's dilemma by employing a Markov process that is described by the past actions of the two players. Their results showed that a strategy called "forgiver" triumphed. Forgivers cooperate whenever their opponents have cooperated. They defect once their opponents have defected, but subsequently, they attempt to reestablish cooperation after mutual defection.

---

[*]hitoshi@ris.ac.jp

In a study on indirect reciprocity, Yamamoto *et al.* [28] and Uchida *et al.* [29] developed a "norm ecosystem" approach that describes the coevolution of strategies and cooperation among various strategies.

In this paper, we exhaustively analyze the coevolution of strategies and cooperation in a PDG with voluntary participation. In this approach, we must consider a combinatorial explosion problem because the total number of strategies becomes $3^9 = 19\,683$ if the strategy space that deals with the past actions of the two players is extended to voluntary participation. Analyzing the dynamics of all strategies mathematically is difficult. To consider this point, we adopted an agent-based model with a genetic algorithm (GA) [30] to analyze adaptive strategies. Although some approaches have been developed to analyze adaptive strategies in evolutionary computation [31], most approaches do not have a process to produce a new candidate of solutions beyond the initial candidates, except for a mutation. Our model must explore an adaptive strategy among about 20 000 strategies. A GA can be harmonistically compatible with both the diversity of the strategies and the evolution of adaptive strategies.

II. MODEL

We model a PDG with voluntary participation that consists of finite individuals with no spatial structure. A set of the players' actions is defined as $A = \{C, D, L\}$, and the payoff matrix of the game is represented as M, where C means cooperation, D means defection, and L means loner. Players have their own strategies to decide their actions. The strategy of each player adopts a deterministic finite automaton [32] determined by the player's and that player's partner's actions; thus, the strategy space of the model consists of the player's previous action (C, D, L) and the partner's previous action (C, D, L). The focal player chooses from three options of actions in each combination of actions in the previous interaction. Therefore, $3^9 = 19\,683$ types of strategies are in the voluntary PDG.

The payoff of the simultaneous game is described in the following scenario. In each round, two players choose C, D, or L. Both players get a payoff, $b - c$, when cooperation occurs, where $b > c > 0$. If a player defects and the partner cooperates, the former gets b and the latter gets $-c$. Nothing happens in mutual defection. If either player chooses a loner option, both of the players get a payoff, δ where $b - c > \delta > 0$.

In the alternating PDG, the two players have different roles from each other. One of the players is the "leader" and the other is the "follower." First, the leader chooses an action in accordance with the actions of the previous round. Then the follower chooses an action in accordance with the action of the leader in the present round and the follower's own action in the previous round. The calculation of their payoffs is the same as that of the simultaneous game. Thus, the payoff matrices M in simultaneous and alternating PDGs are given by

$$M = \begin{pmatrix} b-c & -c & \delta \\ b & 0 & \delta \\ \delta & \delta & \delta \end{pmatrix}. \qquad (1)$$

The first row of the matrix M represents the cooperator's payoffs, the second row represents the defector's payoffs, and the third row represents the loner's payoff. The first column of M is the player's payoff in the case that the player's partner is a cooperator, the second column is said payoff in the case that said partner is a defector, and the third column is said payoff in the case that said partner is a loner.

Our simulation runs through G generations. Each player plays a number of P games in each generation. In each game, a player makes a pair with another player randomly chosen from the whole population, and the selected pair remains fixed along R rounds. Note that the number of players, N, must be an even number. In each round, each pair plays a game where the payoff is described as M. Therefore, player i receives an accumulative payoff U_i calculating RP times of M in each generation.

We set players of a pair as i and j to define the actions of the players. Player i decides its action using the following procedure. Let $I_i(t) \in A$ be the intention of i's action at round t and $B_i(t) \in A$ be i's executed action at round t. The player's strategy, S_i, is a mapping that satisfies $S_i: A \times A \to A$. $I_i(t)$ is defined as $I_i(t) = S_i[B_i(t-1), B_j(t-1)]$ in the simultaneous PDG. In the alternating PDG, $I_{\text{leader}}(t)$ is defined as $I_{\text{leader}}(t) = S_{\text{leader}}[B_{\text{leader}}(t-1), B_{\text{follower}}(t-1)]$, and $I_{\text{follower}}(t)$ is defined as $I_{\text{follower}}(t) = S_i[B_{\text{follower}}(t-1), B_{\text{leader}}(t)]$. If an implementation error occurs in an action at round t, $B_i(t)$ should be chosen at random from two other actions except for the original $I_i(t)$. Otherwise, $B_i(t) = I_i(t)$. The probability of the implementation error is set to a static constant e.

Players each develop their own strategy S_i using GA at the end of each generation. In our model, the strategy of an agent consists of nine loci corresponding to the combinations of previous actions, and each locus has three possible "alleles" (C, D, L). For example, a locus of the gene representing an action for both i and j chooses cooperation in their previous actions. In the process of GA, each agent randomly selects two agents from N agents (including itself) to become its parents. To choose parents, we adopt a roulette selection method. This roulette selection sets a probability distribution of all agents as $\Pi_i = (U_i - U_{\min})^s / \sum_j (U_j - U_{\min})^s$, where U_i denotes the agent i's accumulated payoff in a generation given by $U_i = bW - cV + \delta Z$, with W being the amount of cooperation i received in the generation, V being the amount of cooperation i gave, and Z being the number of loners i or j chose. $U_{\min}$ means a minimum value of the accumulated payoffs among all U_i. s represents selection pressure. In this paper, s is set to 1.0. Finally, each player updates its strategy using a uniform crossover technique. With a constant mutation rate m, each locus is inverted to two other alleles at random to maintain the diversity of the strategy space.

III. RESULTS

The initial strategies of players and their actions in the first round were given randomly. The static parameters of our agent-based simulation were set as $N = 200$, $P = 200$, $R = 200$, $e = 0.01$, $m = 0.01$, and $G = 200$. By setting $b = 1$, the conditions $c + \delta < 1$, $0 < c < 1$, and $0 < \delta < 1$ expressed the full space of M, satisfying the PDG with voluntary participation.

In observing the results, we developed a visualizing method that maps the ratio of all the players' actions to an
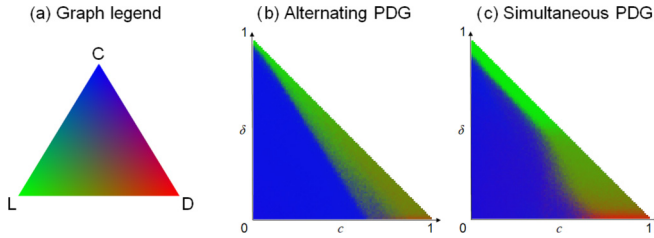
FIG. 1. Phase diagrams of action ratios in population at 200th generation. *x* axis represents *c*, *y* axis represents δ.

RGB color chart [see Fig. 1(a)]. Each vertex of the triangle shows perfect domination of cooperation (blue), loners (green), and defection (red). The center of the triangle shows that all strategies exist equally. For example, if an action ratio is $(C, D, L) = (0.6, 0.2, 0.2)$, the color on the graph is expressed as $(R, G, B) = (51, 51, 153)$ in the RGB color chart. If the cooperation ratio equals one, (R, G, B) becomes $(0, 0, 255)$ and is mapped in blue. Figures 2 and 3 also adopt the graph legend described by Fig. 1(a).

First, Fig. 1 shows the phase diagrams of the average action ratios at the 200th generation of 50 runs with different random seeds. In accordance with the values of *c* and δ, there are two phases that cooperation dominates (shown in blue) and a mixture of three actions. In the case of the alternating PDG [Fig. 1(b)], the boundary of the areas is clearer than that in the simultaneous PDG [Fig. 1(c)]. Although an area is dominated by loners for $\delta > 0.5$ in the simultaneous PDG, the correspondence is narrow and unclear in the alternating PDG.

Next, we analyze distributions of strategies in the population. Observing the ratios of each strategy is difficult because
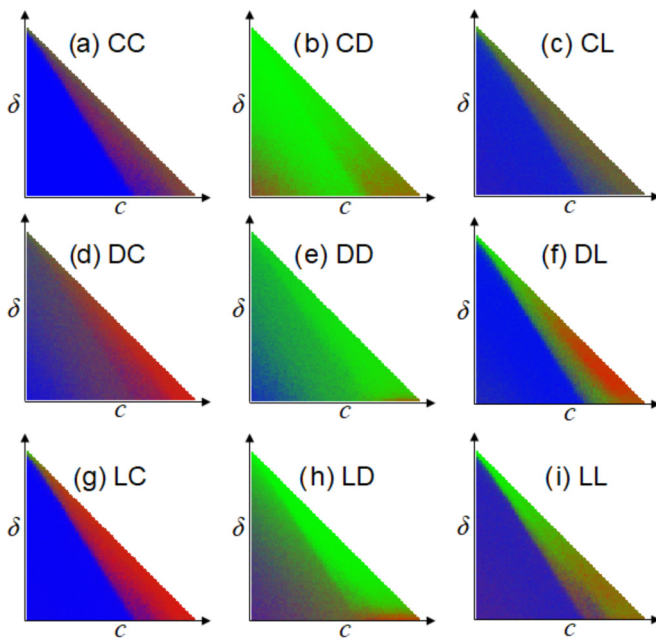

FIG. 2. Phase diagrams of strategy distributions at 200th generation in alternating PDG. *x* axis represents *c*, *y* axis represents δ. For example, panel (a) shows the distribution of actions in a case wherein both players chose cooperation in their previous actions.
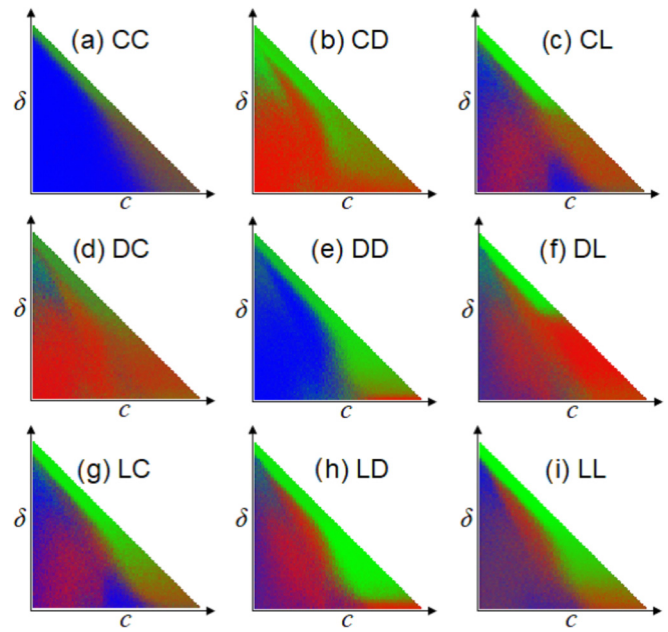

FIG. 3. Results for simultaneous PDG corresponding to those of Fig. 2.

we must consider $3^9 = 19\,683$ strategies; thus, we observe the ratios of alleles of each locus to understand the properties of adaptive strategies.

Figure 2 shows the distributions of strategies in the alternating PDG. Notably, cooperative behaviors after a partner became a loner in the previous round [Figs. 2(c), 2(f), and 2(i)] dominated; in other words, when a partner refuses to participate in the game, the other player behaves conciliatorily and gives the incentive of cooperation to the partner. Further, players who were exploited (when they cooperated while their partners defected) became loners. This allele is expressed in defection in both the win-stay, lose-shift strategy and the forgiver strategy that triumphs in alternating PDG with compulsory participation [27]. However, introducing voluntary participation enables the allele to be a loner. Generally, from Fig. 2, we find that the alternating PDG with voluntary participation favors nonbelligerent strategies that avoid interactions with defectors rather than belligerent strategies that respond by defecting when faced with defection. Finally, a dissimilar result is observed when both *c* and δ are close to zero. This area means that both the cost of cooperation and the benefit of being a loner are quite low. When both players of a pair participate in the game [see Figs. 2(a), 2(b), 2(d), and 2(e)], a strategy represented as CDCC dominates. The strategy corresponds to generous tit-for-tat [12]. This strategy has been shown to be a robust winning strategy in the alternating PDG with compulsory participation [12,27] or spatial interaction in PDGs [33]. Thus, the effect of voluntary participation disappears in the area.

Figure 3 shows the distributions of strategies in the simultaneous PDG. When compared with the alternating PDG, the alleles of defection increase in the area that cooperative behavior dominates. For example, when a partner became a loner in the previous round [3(c), 3(f), and 3(i)], cooperation became the majority in the alternating PDG. In contrast,

defection (shown in red) increases in the simultaneous PDG and vice versa. Additionally, Figs. 3(a) and 3(e) show cooperation (blue) and Figs. 3(b) and 3(d) show defection (red). This corresponds to the win-stay, lose-shift strategy. Hence, the dominant strategy in this situation is a suspicious strategy that responds by defecting to seek an opportunity for exploitation.

IV. CONCLUSION

We described various strategy spaces in PDGs with voluntary participation. In a compulsory PDG that never allows loners, previous studies revealed that a forgiver or generous tit-for-tat strategy triumphs in an alternating PDG, and the win-stay, lose-shift strategy triumphs in a simultaneous PDG. By including loners, who do not participate in the game, we determined that different strategies triumph in both alternating and simultaneous games.

In the alternating PDG, a strategy that can be described as "escape from interaction if a partner defected, or cooperate if a partner escaped from interaction" dominates the population. This strategy can be regarded as pacifism because players rarely defect in the game. In this PDG, introducing voluntary participation can bring tolerant and conciliatory situations. However, in the simultaneous PDG, the win-stay, lose-shift strategy is superior in the wide parameter space. Further, when a player's partner became a loner, the player rarely chose to cooperate but rather to defect. In this PDG, the effect of introducing voluntary participation is not as clear compared with the alternating PDG.

Earlier literature has argued that online social network services (SNSs) resemble a feature of public goods [34,35]. While some works have tried to model free riders on SNSs as defectors [36,37], our model makes it possible to describe nonparticipation explicitly in SNSs by introducing the concept of loners.

The present version of this model has some limitations that should be the subject of future work. Although the two phases of the actions are observed in the alternating PDG, the boundary condition of the phases remains a theoretical calculation. Further, the boundary is not clear in the simultaneous PDG. We need to explore the reason for this difference. While we have focused on a voluntary PDG with no spatial structure to understand the effects of loners in this paper, studies on structured populations have been a major factor in the study of the evolution of cooperation [38–41]. The effect of structures on voluntary participation should be the subject of future investigations. Another factor that is out of the scope of this research is a comparison with an adaptive architecture, i.e., a way for individuals to adapt to their environments. Besides GA, some architectures have been argued for evolutionary computation [29,31,42]. This point also provides an interesting direction for potential future research.

[1] M. A. Nowak, Science **314**, 1560 (2006).

[2] R. Axelrod and W. D. Hamilton, Science **211**, 1390 (1981).

[3] J. M. Smith, *Did Darwin Get It Right?* (Springer, Boston, 1988), pp. 202–215.

[4] K. Lindgren and M. G. Nordahl, Physica D **75**, 292 (1994).

[5] J. W. Weibull, *Evolutionary Game Theory* (MIT Press, Cambridge, MA, 1997).

[6] M. Nakamaru, H. Matsuda, and Y. Iwasa, J. Theor. Biol. **184**, 65 (1997).

[7] M. D. Cohen, R. L. Riolo, and R. Axelrod, Ration. Soc. **13**, 5 (2001).

[8] J. M. Pacheco, A. Traulsen, H. Ohtsuki, and M. A. Nowak, J. Theor. Biol. **250**, 723 (2008).

[9] Y. Liu, X. Chen, L. Zhang, L. Wang, and M. Perc, PLoS ONE **7**, e30689 (2012).

[10] R. L. Trivers, Q. Rev. Biol. **46**, 35 (1971).

[11] M. Milinski, Nature (London) **325**, 433 (1987).

[12] M. A. Nowak and K. Sigmund, J. Theor. Biol. **168**, 219 (1994).

[13] C. Hauert, S. D. Monte, J. Hofbauer, and K. Sigmund, Science **296**, 1129 (2002).

[14] C. Hauert, S. D. Monte, J. Hofbauer, and K. Sigmund, J. Theor. Biol. **218**, 187 (2002).

[15] G. Szabó and C. Hauert, Phys. Rev. Lett. **89**, 118101 (2002).

[16] D. Semmann and H.-J. Krambeck, Nature (London) **425**, 390 (2003).

[17] H. Brandt, C. Hauert, and K. Sigmund, Proc. Natl. Acad. Sci. U. S. A. **103**, 495 (2006).

[18] T. Sasaki, I. Okada, and T. Unemi, Proc. R. Soc. B Biol. Sci. **274**, 2639 (2007).

[19] H. De Silva, C. Hauert, A. Traulsen, and K. Sigmund, J. Evol. Econ. **20**, 203 (2010).

[20] J. M. Orbell and R. M. Dawes, Am. Sociol. Rev. **58**, 787 (1993).

[21] J. Batali and P. Kitcher, J. Theor. Biol. **175**, 161 (1995).

[22] G. Szabó and C. Hauert, Phys. Rev. E **66**, 062903 (2002).

[23] C. Chu, J. Liu, C. Shen, J. Jin, and L. Shi, PLoS One **12**, e0171680 (2017).

[24] M. Frean, Proc. R. Soc. B Biol. Sci. **257**, 75 (1994).

[25] C. Wedekind and M. Milinski, Proc. Natl. Acad. Sci. USA **93**, 2686 (1996).

[26] H. Ohtsuki and Y. Iwasa, J. Theor. Biol. **239**, 435 (2006).

[27] B. M. Zagorsky, J. G. Reiter, K. Chatterjee, and M. A. Nowak, PLoS ONE **8**, e80814 (2013).

[28] H. Yamamoto, I. Okada, S. Uchida, and T. Sasaki, Sci. Rep. **7**, 44146 (2017).

[29] S. Uchida, H. Yamamoto, I. Okada, and T. Sasaki, Front. Phys. **6**, 14 (2018).

[30] J. Holland, *Adaptation in Natural and Artificial Systems* (University of Michigan Press, Ann Arbor, 1975).

[31] M. A. Javarone, Eur. Phys. J. B **90** (2017).

[32] J. E. Hopcroft, R. Motwani, and J. D. Ullman, *Introduction to Automata Theory, Languages and Computation*, 3rd ed. (Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 2006).

[33] H. Fort and E. Sicardi, Physica A **375**, 323 (2007).

[34] J. Fulk, A. J. Flanagin, M. E. Kalman, P. R. Monge, and T. Ryan, Commun. Theory **6**, 60 (1996).

[35] J. Fulk, R. Heino, A. J. Flanagin, P. R. Monge, and F. Bar, Organ. Sci. **15**, 569 (2004).

[36] M. A. Javarone, R. Interdonato, and A. Tagarelli, in *Complex Networks VII*, edited by H. Cherifi, B. Gon calves, R. Menezes, and R. Sinatra (Springer, Cham, 2016), pp. 227–239.

[37] F. Toriumi, H. Yamamoto, and I. Okada, J. Artif. Soc. Soc. Simul. **19**, 6 (2016).

[38] M. A. Nowak and R. M. May, Nature (London) **359**, 826 (1992).

[39] F. C. Santos and J. M. Pacheco, Phys. Rev. Lett. **95**, 098104 (2005).

[40] M. A. Javarone and A. E. Atzeni, Comput. Soc. Netw. **2**, 15 (2015).

[41] M. A. Javarone, Eur. Phys. J. B **89**, 42 (2016).

[42] M. A. Javarone, Physica A **463**, 262 (2016).