

Reactive strategies in indirect reciprocity

Hisashi Ohtsuki*

Department of Biology, Faculty of Sciences, Kyushu University, 6-10-1, Hakozaki, Fukuoka, 812-8581, Japan

Received 4 July 2003; received in revised form 20 October 2003; accepted 5 November 2003

Abstract

Evolution of reactive strategy of indirect reciprocity is discussed, where individuals interact with others through the one-shot Prisoner's Dilemma game, changing their partners in every round. We investigate all of the reactive strategies that are stochastic, including deterministic ones as special cases. First we study adaptive dynamics of reactive strategies by assuming monomorphic population. Results are very similar to the corresponding evolutionary dynamics of direct reciprocity. The discriminating strategy, which prescribes cooperation only with those who cooperated in the previous round, cannot be an outcome of the evolution. Next we examine the case where the population includes a diversity of strategies. We find that only the mean 'discriminatoriness' in the population is the parameter that affects the evolutionary dynamics. The discriminating strategy works as a promoter of cooperation there. However, it is again not the end point of the evolution. This is because retaliatory defection, which was prescribed by the discriminating strategy, is regarded as another defection toward the society. These results caution that we have to reconsider the role of retaliatory defection much more carefully.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Indirect reciprocity; Reputation; Adaptive dynamics; Stochastic strategy

1. Introduction

Explaining the evolution of cooperation or altruism is one of the major problems in evolutionary biology. In theoretical studies on this issue, the Prisoner's Dilemma game is often adopted to model socially dilemmatic situations. Two players involved in this game each has two alternative actions; to cooperate (C) or to defect (D). If both choose to cooperate, each of them gains R , which is larger than the gain when both choose to defect, P . When one of them cooperates and the other defects, however, the defector can enjoy the benefit, T , which is larger even than R , whereas the cooperator results in getting only S , which is the worst payoff. In their seminal paper, Axelrod and Hamilton (1981) showed that the strategy called 'Tit For Tat' (TFT), which prescribes cooperation only with those who cooperated last time, has a remarkable stability against invasions in the Iterated Prisoner's Dilemma (IPD), provided that the interaction with an opponent lasts long enough. The

success of the Tit For Tat suggests the importance of reciprocal altruism (Trivers, 1971).

However, in the social group of primates especially of humans, interactions occur among many individuals, and partners with whom one interacts are not always the same. Then we face the question whether reciprocity works in such a situation or not. Alexander (1987) was the first that referred to reciprocity there as *indirect reciprocity*, which Trivers (1971) himself mentioned as 'generalized reciprocity'. In indirect reciprocity, the player who cooperated gains the return back not from the beneficiary, but he will receive reciprocation from a third person who knows of his good deed through direct observation or reputation. Nowak and Sigmund (1998a, b) was the first that conducted extensive theoretical studies of this issue. Using computer simulations, they found that the 'discriminating strategy' can evolve. The player who uses this strategy cooperates only with the person who has a good social reputation due to his previous cooperation with others. They also studied the evolutionary dynamics analytically and showed the supremacy of the discriminating strategy.

Leimar and Hammerstein (2001) pointed out potential problems concerning the adaptive motivation of

*Tel.: +81-92-642-2641; fax: +81-92-642-2645.

E-mail address: ohtsuki@bio-math10.biology.kyushu-u.ac.jp (H. Ohtsuki).

players using the discriminating strategy. First, in the Nowak and Sigmund's settings (they adopted a *public-goods-game-type* payoff; see Section 3), the gain when a defector changes his action to cooperation when his opponent is to cooperate is exactly the same as the one when his opponent is to defect. This means that the opponent's propensity toward cooperation has nothing to do with his own payoff. Second, the player who cooperated is supposed to acquire good reputation irrespective of his opponent's reputation. In other words, reputation of the opponent has no effects on his own reputation. These two assumptions will lead to the conclusion that players have no motivations at all to classify members in the population into two groups; those who have good reputation and those who do not.

In the present paper, we will study the attainability of strategies. We examine what strategy can emerge among infinitely many reactive strategies in indirect reciprocity. A strategy is called reactive when it prescribes actions based solely on the opponents' behavior. In particular, we will concentrate on stochastic strategies (Nowak and Sigmund, 1989; Nowak, 1990), which prescribe one's action not deterministically but stochastically. This will be investigated by adaptive dynamics (Dieckmann and Law, 1996; Hofbauer and Sigmund, 1998; Page and Nowak, 2002), where monomorphic population is assumed in strategies. We will also study the effect of errors or noise on the dynamics. With such errors we can clarify the difference between the discriminating strategy and such strategies as All C, and can obtain more or less different results. Later in this paper, we will study the general case in which population includes a wide range of strategies.

2. Adaptive dynamics in indirect reciprocity

First, we begin with a population game model. Assume that the population size is infinite. In each 'round' of the game, every player meets an opponent randomly chosen from the population and engages in the one-shot Prisoner's Dilemma game. The payoff matrix is

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{array} \quad (1a)$$

The following two inequalities hold:

$$T > R > P > S, \quad 2R > T + S. \quad (1b)$$

Eq. (1b) states that alternating unilateral cooperation yields less than mutual cooperation. The probability with which the game proceeds into the next round is given by $w \in [0, 1]$.

In indirect reciprocity, players choose their action by assessing reputation or status of the partner. Here we limit our attention to what a player did in the previous round and assume that it will be conveyed to other players as his social information. We assume that a player knows the action of a randomly chosen opponent in the previous round with probability q (for the treatment of incomplete information, see also Nowak and Sigmund, 1998a; Panchanathan and Boyd, 2003). We call q 'information transfer' in the population as it represents the availability of social information. Then a player decides his own action (= to cooperate or to defect) by using that information. In the first round of the game, no information is available on the previous move of the opponent.

We assume that each player has a reactive strategy as a decision rule; behavior of a player is determined solely by his opponent's move in the previous round and it does not depend on what he himself did. For example, a strategy which prescribes cooperation only with those who cooperated in the previous round is called 'discriminating strategy'. There is also an opposite strategy, which chooses to cooperate only with those who defected in the previous round. The strategy which always ordains cooperation irrespective of opponents is called 'All C strategy', and the one which prescribes defection against any opponents is referred to as 'All D strategy'. Such strategies are deterministic. In the real world, however, it is more realistic to expect that a player increases or decreases his propensity toward cooperation, taking social information of his opponent into account. It would be rare that players adopt all-or-nothing rules such as deterministic strategies. Therefore, we assume that a reactive strategy of each player comprises stochasticity.

Let us begin with the mathematical formalism of the model. Each player in the population has his own strategy expressed by a triple $\mathbf{p} = (p_0, p_1, p_2) \in [0, 1]^3$, where p_0 represents the probability with which the player cooperates with the opponent whose information he does not have. This occurs with probability $(1 - q)$; it corresponds to the case in which social information is not available. With probability q , a player knows what his opponent did in the previous round. If the opponent cooperated then, the player cooperates with him with probability p_1 . If the opponent did not cooperate, the player cooperates with him with probability p_2 . In the first round of the game, we assume that a player with \mathbf{p} strategy cooperates with probability p_0 , since no social information is available yet. We mention that Nowak and Sigmund (1990) adopted similar formulation for stochastic strategies in *direct* reciprocity.

We note that an All C player corresponds to $(1, 1, 1)$, an All D player $(0, 0, 0)$, and discriminating strategies $(p_0, 1, 0)$ ($p_0 \in [0, 1]$) in our framework. In particular, we

will refer to $(1, 1, 0)$ as *nice-discriminator* and $(0, 1, 0)$ as *suspicious-discriminator*.

Consider the situation in which the population is monomorphic and all the individuals have the same strategy \mathbf{p} . What we are interested in now is evolutionary change of \mathbf{p} . In order to study this, we use *adaptive dynamics* (Dieckmann and Law, 1996; Hofbauer and Sigmund, 1998; Page and Nowak, 2002). Adaptive dynamics describes the evolution driven by mutation and natural selection that determines which mutants can spread and take over the population. It is assumed that the mutant that is locally best-response to the present population always invades. This leads to calculating the gradient of payoff function with respect to a mutant's strategy. Let $W(\mathbf{p}'|\mathbf{p})$ be the payoff of a rare mutant using \mathbf{p}' in the population of \mathbf{p} . The adaptive dynamics on the trait \mathbf{p} are given by

$$\dot{\mathbf{p}} = \nabla_{\mathbf{p}'} W(\mathbf{p}|\mathbf{p}) \quad (2a)$$

or in components

$$\dot{p}_i = \frac{\partial}{\partial p'_i} W(\mathbf{p}|\mathbf{p}). \quad (2b)$$

(Note that similar formulation can be derived from quantitative genetics, where the variance–covariance matrix of trait \mathbf{p} is multiplied on the right-hand side of Eq. (2a). This matrix plays a role of ‘speed factor’ of the evolution. Eq. (2a) corresponds to the case in which the variance–covariance matrix is replaced with an identity matrix there. See also Iwasa et al. (1991).)

By $W_n(\mathbf{p}'|\mathbf{p})$, we will denote \mathbf{p}' player's payoff in the \mathbf{p} population in round n . For $0 \leq w < 1$, we use

$$W(\mathbf{p}'|\mathbf{p}) = \sum_{n=1}^{\infty} w^{n-1} W_n(\mathbf{p}'|\mathbf{p}). \quad (3a)$$

When $w = 1$, however, the accumulated payoff defined as Eq. (3a) does not converge. Hence we use the average payoff defined as

$$W(\mathbf{p}'|\mathbf{p}) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N W_n(\mathbf{p}'|\mathbf{p}) \quad (3b)$$

for $w = 1$ instead.

Next we calculate the payoff function $W(\mathbf{p}'|\mathbf{p})$. Let c_n, c'_n, c''_n denote the average cooperation rate of a wild-type individual toward another wild-type individual ($\mathbf{p} \rightarrow \mathbf{p}$) in round n , that of a mutant toward a wild-type individual ($\mathbf{p}' \rightarrow \mathbf{p}$) in round n , and that of a wild-type individual toward a mutant ($\mathbf{p} \rightarrow \mathbf{p}'$) in round n , respectively.

The wild-type individual E_1 who plays the game with another wild-type individual E_2 in round $(n+1)$ chooses his action by using the social reputation of E_2 . Since mutants are so rare in the population, we can assume that E_2 played the game with another wild-type individual, say E_3 , in round n . Hence E_1 's action in round $(n+1)$ is determined by how E_2 behaved toward

E_3 in round n . Because the probability with which E_2 cooperated with E_3 in round n is given by c_n , E_1 's propensity to cooperate in round $(n+1)$, that is c_{n+1} , is given by

$$c_{n+1} = (1-q)p_0 + q\{c_n p_1 + (1-c_n)p_2\} \\ = u(\mathbf{p})c_n + v(\mathbf{p}), \quad (4a)$$

where $u(\mathbf{p}) = q(p_1 - p_2)$, and $v(\mathbf{p}) = (1-q)p_0 + qp_2$. One can see that $v(\mathbf{p})$ represents \mathbf{p} player's propensity to cooperate even with defectors and that $u(\mathbf{p})$ represents the difference of cooperativeness toward cooperators from that toward defectors. Therefore, we call each $v(\mathbf{p})$ and $u(\mathbf{p})$, \mathbf{p} player's ‘forgiveness’ and ‘discriminatoriness’ respectively. Similar calculations give us the following recursions on c'_n and c''_n :

$$c'_{n+1} = u(\mathbf{p}')c_n + v(\mathbf{p}'), \quad (4b)$$

$$c''_{n+1} = u(\mathbf{p})c'_n + v(\mathbf{p}). \quad (4c)$$

By solving these recursions, we can derive $W(\mathbf{p}'|\mathbf{p})$ analytically (see Appendix A). Since we are mostly interested in the case where $q < 1$, we assume that $u = u(\mathbf{p}) \in (-1, 1)$ holds from here on.

3. Results

3.1. When $w = 1$

First we will take a look at the case where $w = 1$. Using Eq. (2a) and results in Appendix A, we have the following adaptive dynamics on the trait \mathbf{p} :

$$\dot{\mathbf{p}} = \begin{pmatrix} \dot{p}_0 \\ \dot{p}_1 \\ \dot{p}_2 \end{pmatrix} \\ = \left(\frac{1+u}{1-u} vG_1 + G_2 + uG_3 \right) \begin{pmatrix} 1-q \\ q \frac{v}{1-u} \\ q \left(1 - \frac{v}{1-u} \right) \end{pmatrix}, \quad (5)$$

where $u = u(\mathbf{p})$, $v = v(\mathbf{p})$, $G_1 = R - S - T + P$, $G_2 = S - P$ (< 0), and $G_3 = T - P$ (> 0).

At the limit $q \rightarrow 1$, Eq. (5) above is the same (except for a multiplier) as what Nowak and Sigmund (N & S) (1990) derived (Eq. [16] in N & S), in which the strategic evolution not in *indirect* reciprocity but in *direct* reciprocity was discussed. This is very interesting in that *direct* reciprocity and *indirect* reciprocity are very similar. It does not matter so much whether the return of an altruistic act is brought by the beneficiary himself or by a third individual.

Note that the direction of the evolution described by Eq. (5) is the same as the one in Nowak and Sigmund's (Eq. [24] in N & S), provided that the parameter w in N & S is replaced with q . Thanks to those similarities, we

can make use of the Nowak and Sigmund's results extensively. We note that $1 - q$, $q v / (1 - u)$, and $q(1 - (v / (1 - u)))$ on the right-hand side of Eq. (5) are all nonnegative. Then the sign of $((1 + u) / (1 - u))vG_1 + G_2 + uG_3$ characterizes the whole dynamics. As for the direction of the evolution, two equalities,

$$\frac{q}{1 - q} \dot{p}_0 - \dot{p}_1 - \dot{p}_2 = 0, \quad (6a)$$

$$p_0 \dot{p}_0 + \left(p_1 - \frac{1}{q}\right) \dot{p}_1 + p_2 \dot{p}_2 = 0, \quad (6b)$$

hold. From these we can see that the orbit of the dynamics lies in the intersection of a plane and a sphere; in short, it draws an arc.

Consider first the case where G_1 vanishes. This is true when individuals play the *public goods game*. The *public goods game* is a special version of the Prisoner's Dilemma game in its payoff structure. In this game, cooperation costs the donor of help ($c > 0$) and it brings benefit $b(> c)$ to the beneficiary, whereas defection yields nothing to both. Hence $R = b - c$, $S = -c$, $T = b$, and $P = 0$, so $G_1 = R - S - T + P = 0$.

When G_1 is equal to zero, the sign of $(G_2 + uG_3)$ is of great importance. A two-dimensional plane defined as $u = -G_2/G_3$ separates the strategy space $[0, 1]^3$ into two regions (Fig. 1). Following Nowak and Sigmund (1990), we call the region in which selection favors cooperativeness in all of the components of \mathbf{p} , '*C-region*', and the region in which selection drives the population toward unconditional defection '*D-region*'. Note that the *C-region* in Fig. 1 contains the edge of discriminators. We also note that the *D-region* always includes the vertex of All D strategy. Eq. (5) shows that there exist no regions where p_1 increases but p_2 decreases. If the population was initially in the neighborhood of the edge of discriminating strategies for example, then p_2 is to increase due to advantageous mutation. Hence the discriminatoriness which the population had gradually diminishes. This suggests that the discriminating strategy is not a destination of the evolution.

The *C-region* exists if and only if

$$q > -\frac{G_2}{G_3} \quad (7)$$

holds. When the game is a public-goods-game type, inequality (7) becomes

$$bq > c. \quad (8)$$

This means that the population of discriminating strategies cannot be invaded by uncooperative strategies if the expected return bq exceeds the cost of help, c . This is the stability condition for the (nice-) discriminating strategy derived by Nowak and Sigmund (1998a) (Eq. [59]). However, even when inequality (8) holds, Fig. 1 shows that the population with the discriminating

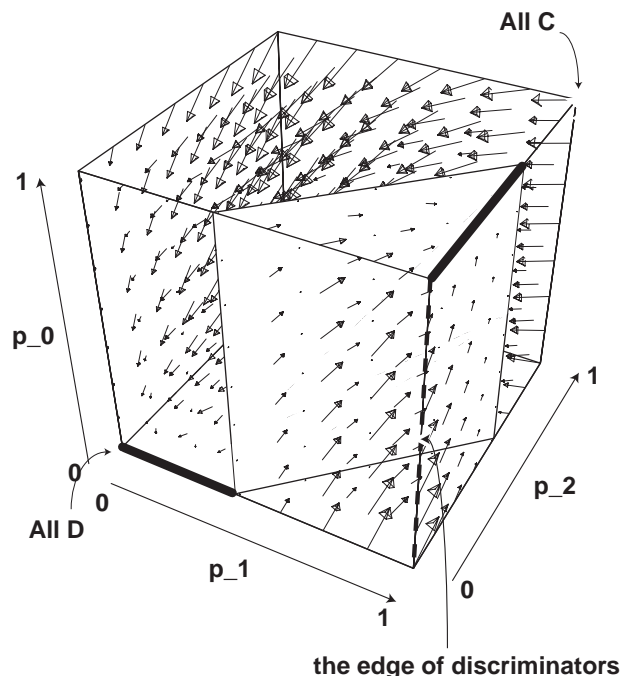


Fig. 1. Adaptive dynamics in indirect reciprocity when $w = 1$, $G_1 = R - S - T + P = 0$. Each axis represents the conditional probability with which a player cooperates with his partner, provided that the partner cooperated in the previous round (p_1), defected in the previous round (p_2), and that he knows nothing about the partner's behavior in the previous round (p_0), respectively. When the flow defined as Eq. (5) is to go outside the strategy space, $[0, 1]^3$, it is properly projected inward so as to ensure that the strategy is always in $[0, 1]^3$. This does not lose the spirit of adaptive dynamics: 'the best to the present population always invades'. The white plane defined as $u = -G_2/G_3$ separates the strategy space into the *C-region* (front-side of the plane) and the *D-region* (back-side). All points on that plane are equilibria of the dynamics. Two solid lines on the edge of the cube also consist of equilibria. All C corresponds to $(1, 1, 1)$ and All D to $(0, 0, 0)$. The dotted line shown in the figure represents discriminating strategies, $(p_0, 1, 0)$. In the *C-region*, all p_i 's increase. The population which started from the *C-region* ultimately reaches either an equilibrium on the plane or the upper neutral edge. In the *D-region*, all p_i 's decrease. The population which started from the *D-region* ultimately arrives at either an equilibrium on the plane or the lower neutral edge. No regions in which p_1 increases but p_2 decreases exist. Other parameters: $q = 0.8$, $T = 4$, $R = 3$, $P = 1$, and $S = 0$, so $G_2 = S - P = -1$, $G_3 = T - P = 3$, and $G_4 = P = 1$.

strategy (except for the deterministic one: $(1, 1, 0)$) is vulnerable to invasions by more cooperative ones.

Now we turn to the case where G_1 is not zero. In this case the strategy space is divided by the surface

$$v = -\frac{G_2 + uG_3}{G_1} \frac{1 - u}{1 + u} \quad (9)$$

into *C-* and *D-regions*. Again, in no regions does selection lead the population to the discriminating strategies, since increase and decrease of p_1 and p_2 synchronize. We note that the sign of G_1 plays a decisive role here.

(i) When $G_1 > 0$. The *C-region* becomes bigger (Fig. 2) compared with when $G_1 = 0$. This means that relatively

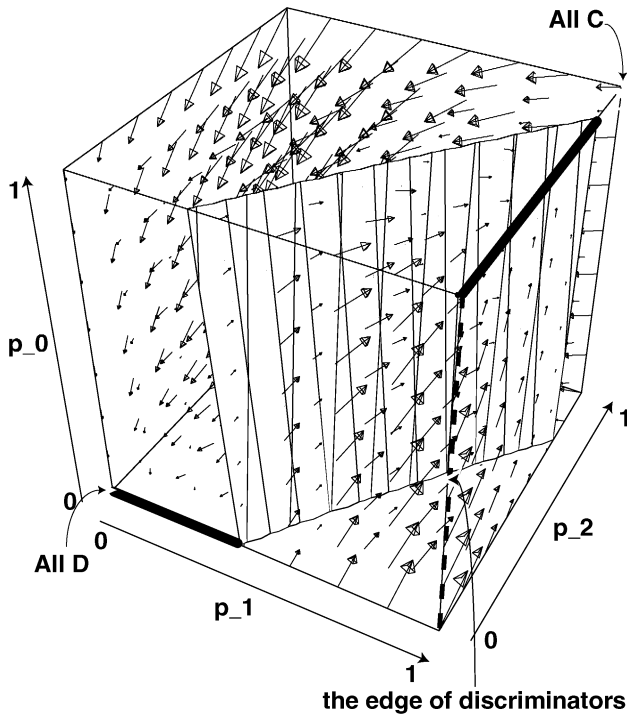


Fig. 2. Adaptive dynamics when $w = 1$, and $G_1 = R - S - T + P = 0.5 > 0$. When G_1 becomes positive, the plane separating the C -region and the D -region moves backward, and the C -region becomes bigger compared with the case where $G_1 = 0$. Parameters: $q = 0.8$, $T = 4$, $R = 3.5$, $P = 1$, and $S = 0$, so $G_2 = S - P = -1$, $G_3 = T - P = 3$, and $G_4 = P = 1$.

larger R (= reward of mutual cooperation) makes cooperation evolve more easily. In order for strategy \mathbf{p} to be in the C -region, v must be larger than the right-hand side of Eq. (9). This says that the greater the forgivingness is, the more likely cooperation evolves. As for the discriminating strategies $(p_0, 1, 0)$, it is in the C -region if and only if

$$p_0 > -\frac{G_2 + qG_3}{G_1(1+q)}. \quad (10)$$

(ii) When $G_1 < 0$. The C -region becomes smaller (Fig. 3) compared with when $G_1 = 0$. This means that the small reward of mutual help makes it difficult for cooperation to evolve. In order for strategy \mathbf{p} to be in the C -region, v must be smaller than the right-hand side of Eq. (9). This tells us that the less the forgivingness, the more likely cooperation evolves. The discriminating strategy $(p_0, 1, 0)$ is in the C -region if and only if

$$p_0 < -\frac{G_2 + qG_3}{G_1(1+q)}. \quad (11)$$

3.2. When $w < 1$

Next we will study the case where $0 \leq w < 1$. The adaptive dynamics is given by

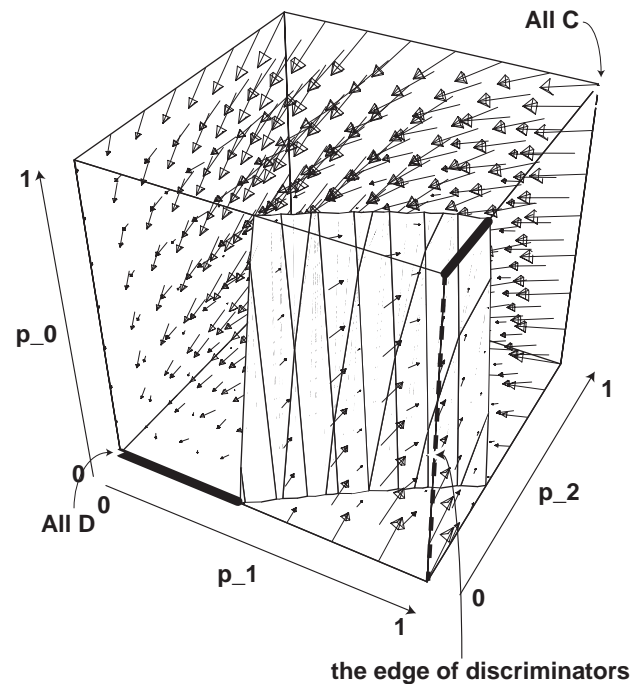


Fig. 3. Adaptive dynamics when $w = 1$, $G_1 = R - S - T + P = -0.5 < 0$. When G_1 becomes negative, the plane separating the C -region and the D -region moves forward, and the C -region becomes smaller compared with the case where $G_1 = 0$. Parameters: $q = 0.8$, $T = 4$, $R = 2.5$, $P = 1$, and $S = 0$, so $G_2 = S - P = -1$, $G_3 = T - P = 3$, and $G_4 = P = 1$.

$$\dot{\mathbf{p}} = \begin{pmatrix} \dot{p}_0 \\ \dot{p}_1 \\ \dot{p}_2 \end{pmatrix} = G_1 \begin{pmatrix} f_0(\mathbf{p}) \\ f_1(\mathbf{p}) \\ f_2(\mathbf{p}) \end{pmatrix} + (G_2 + wuG_3) \frac{1}{1-w} \times \begin{pmatrix} 1-wq \\ wq \frac{(1-w)p_0 + wv}{1-wu} \\ wq \left(1 - \frac{(1-w)p_0 + wv}{1-wu} \right) \end{pmatrix}. \quad (12)$$

For the analytical forms of $f_0(\mathbf{p})$, $f_1(\mathbf{p})$, and $f_2(\mathbf{p})$, see Appendix A for details. We note that $(1-wq)$, $wq\{(1-w)p_0 + wv\}/(1-wu)$, and $wq(1 - \{(1-w)p_0 + wv\}/(1-wu))$ are nonnegative. When G_1 is not equal to zero, the above expression is so complex that we know little from the model. Here we pay attention to when $G_1 = 0$.

Similarly to the previous analysis, we have

$$\frac{wq}{1-wq} \dot{p}_0 - \dot{p}_1 - \dot{p}_2 = 0, \quad (13a)$$

$$p_0 \dot{p}_0 + \left(p_1 - \frac{1}{wq} \right) \dot{p}_1 + p_2 \dot{p}_2 = 0, \quad (13b)$$

so the orbit is an arc again. When $G_1 = 0$, the sign of $(G_2 + wuG_3)$ is of importance in that it determines the landscape of this dynamics. The C -region, in which all

the components of \tilde{p} are positive, is defined as

$$wu > -\frac{G_2}{G_3} \quad (14)$$

and is not empty if and only if

$$wq > -\frac{G_2}{G_3} \quad (15)$$

holds. Again, in no regions does p_1 increase and does p_2 decrease simultaneously. This implies that we cannot expect the discriminating strategy to fix in the population owing to mutation: there are no evolutionary paths to it.

In the *public goods game*, Eq. (15) corresponds to

$$b(wq) > c. \quad (16)$$

Inequality (16) states that there is a possibility for cooperation to evolve only when the cost a player pays to the discriminator is smaller than the expected return from him. When c is so large that even discriminators are not inclined to choose cooperation, cooperation never evolves in that population.

4. Two kinds of errors

In the previous section, we investigated stochastic strategies, including deterministic ones. In an error-prone world, however, it may be very difficult to perform one's own action always correctly. In this section, we consider two kinds of errors (see also [Leimar and Hammerstein, 2001](#)) and investigate their effects on the outcome of the evolution.

4.1. Perception error

When players are to use the reputation of others, it is probable that they sometimes misperceive or misunderstand it. As a result, they might behave incorrectly. We call this kind of error 'perception error'. An observer may misunderstand that those who actually cooperated had defected from their opponent. Similarly, he may misunderstand that those who actually defected had cooperated with their opponent. We assume that those will happen with probability δ_C and δ_D , respectively. Note that δ_C and δ_D are very small. When all players in the population are supposed to commit such errors, the probability that a p strategist chooses to cooperate with those who cooperated in the previous round is modified as

$$(1-q)p_0 + q(1-\delta_C)p_1 + q\delta_C p_2, \quad (17a)$$

and the one that he cooperates with those who defected in the previous round is

$$(1-q)p_0 + q\delta_D p_1 + q(1-\delta_D)p_2. \quad (17b)$$

Consider the strategy $\tilde{p} = (\tilde{p}_0, \tilde{p}_1, \tilde{p}_2)$ satisfying

$$\begin{aligned} \tilde{p}_0 &= p_0, \\ (1-q)\tilde{p}_0 + q\tilde{p}_1 &= (1-q)p_0 + q(1-\delta_C)p_1 + q\delta_C p_2, \\ (1-q)\tilde{p}_0 + q\tilde{p}_2 &= (1-q)p_0 + q\delta_D p_1 + q(1-\delta_D)p_2. \end{aligned} \quad (18)$$

Note that such a strategy always uniquely exists. We can see that a \tilde{p} strategist's behavior under the perception-error-free condition realizes that of a p player with perception errors. Therefore, we can easily understand the effect of perception error. With that error, [Fig. 1](#) turns to [Fig. 4](#). Two lines of neutral equilibria in [Fig. 1](#) vanish. All the equilibria in [Fig. 4](#) are the plane separating the phase space, and the corner of All D = (0, 0, 0). For example, suppose that the initial population consists of the nice-discriminators only. Unlike [Fig. 1](#), with perception errors the population is to evolve to the equilibrium Z in [Fig. 4](#). At Z, any perturbations toward the *C-region* does not affect the subsequent evolutionary outcome very well, for the population is sure to go back near Z. However, any perturbations toward the *D-region* causes invasion by less cooperative strategies, then the population goes away from that equilibrium. Hence Z is sector stable but is locally unstable. This suggests that the population of Z (and of course the

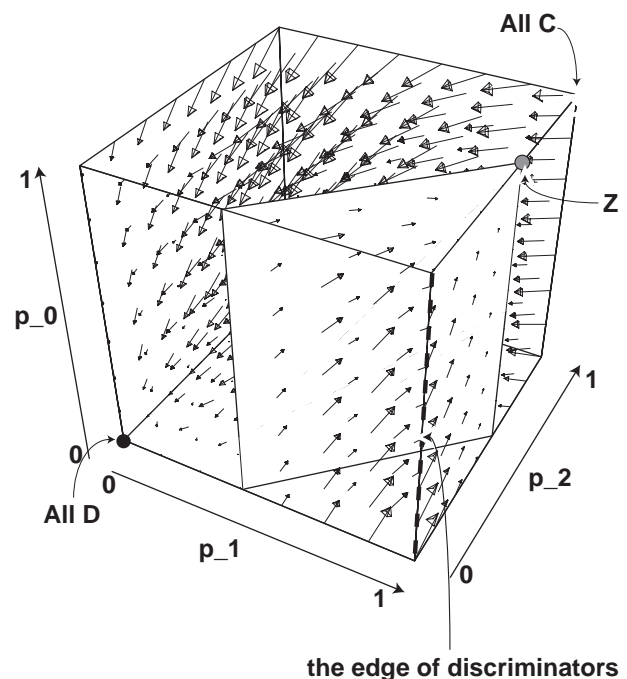


Fig. 4. Adaptive dynamics when perception errors are introduced. Two neutral lines of equilibria in [Fig. 1](#) vanish due to the errors. Only the white plane and All D (0, 0, 0) are the equilibria of the dynamics. The population which started from the *C-region* will go either to an equilibrium on the plane or to Z (in gray). However, these equilibria are unstable against any perturbations toward the *D-region*. Error rates are $\delta_C = \delta_D = 0.025$. Other parameters are the same as [Fig. 1](#).

population of the nice-discriminator) cannot be stably maintained with the perception error.

4.2. Implementation error

Another error is ‘implementation error’. This error occurs in executing one’s intended behavior. Players who choose cooperation may sometimes fail to perform it for some reason. Similarly, players who choose defection might sometimes cooperate with their opponent. We denote those probabilities by ε_C and ε_D , respectively. Note that ε_C and ε_D are very small. We can easily see that a player with \mathbf{p} strategy under the condition in which implementation errors are present behaves as if a $\hat{\mathbf{p}}$ player, where

$$\hat{p}_i = (1 - \varepsilon_C - \varepsilon_D)p_i + \varepsilon_D \quad (i = 0, 1, 2), \quad (19)$$

behaved under the implementation-error-free condition. Fig. 5 illustrates how the dynamics varies after implementation errors are introduced. We set $\varepsilon_C = 0.025$, and $\varepsilon_D = 0$, because usually people who do not want to cooperate never cooperate accidentally. Again, the neutral line of equilibria in the C -region goes away from the figure. As before, the nice-discriminator $(1, 1, 0)$ goes to the unstable equilibrium, Z. Consequently, a cooperative regime becomes susceptible to defectors.

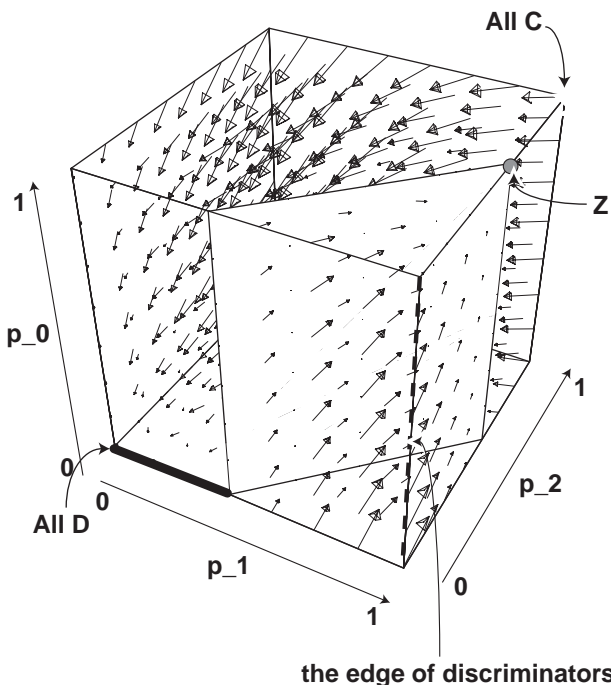


Fig. 5. Adaptive dynamics when implementation errors are introduced. We set $\varepsilon_D = 0$ because those who intended to defect from others seldom cooperate with them accidentally. We set $\varepsilon_C = 0.025$. The neutral line in the C -region in Fig. 1 disappears. If the population was at first at the nice-discriminator, $(1, 1, 0)$, then it goes to the unstable equilibrium, Z (in gray). Similarly to Fig. 4, a cooperative regime cannot be maintained stably. Other parameters are the same as Fig. 1.

5. When different strategies coexist in the population

So far we investigated the evolutionary dynamics of reactive strategies, assuming that the population is monomorphic. But this assumption seems restrictive. Looking back at the real world, a variety of individuals are there, each having different propensity toward cooperation. Some are very generous but others are not. We will consider whether the discriminating strategy has an advantage in such a mixture. In this section, we will study what strategy does well under such a general condition that the population comprises various strategies.

To begin with, let us assume that the density distribution of reactive strategies over $[0, 1]^3$ be ϕ . We assume that ϕ is normalized so that the integral of ϕ over $[0, 1]^3$ is always unity. For later use, we denote the population mean of strategies by \mathbf{m} and the variance–covariance matrix of strategies by \mathbf{H} , respectively. The components of \mathbf{m} and \mathbf{H} are given by

$$m_i = \int p_i \phi(\mathbf{p}) d\mathbf{p} \quad (i = 0, 1, 2), \quad (20a)$$

$$h_{jk} = \int (p_j - m_j)(p_k - m_k) \phi(\mathbf{p}) d\mathbf{p} \quad (j, k = 0, 1, 2). \quad (20b)$$

The integrals are calculated over $[0, 1]^3$.

Let $W(\mathbf{p}'|\phi)$ be the payoff of a \mathbf{p}' strategist when the strategy distribution is ϕ , and let $W_n(\mathbf{p}'|\phi)$ denote that in round n only. We define $W(\mathbf{p}'|\phi)$ by using $W_n(\mathbf{p}'|\phi)$ similarly to Eqs. (3a) and (3b). Complicated algebra (see Appendix B for details) leads to the following remark.

Remark 1. For fixed \mathbf{p}' (and for fixed w and q), $W(\mathbf{p}'|\phi)$ depends only on the mean and the variance–covariance matrix of ϕ . Furthermore, when $G_1 = R - S - T + P = 0$, $W(\mathbf{p}'|\phi)$ depends only on the mean of ϕ .

This remark states that the third or higher moments of ϕ do not affect $W(\mathbf{p}'|\phi)$. Hence we need to know only \mathbf{m} and \mathbf{H} in order to study the effect of heterogeneity of strategies in the population. By comparing $W(\mathbf{p}'|\mathbf{p})$ in Appendix A with $W(\mathbf{p}'|\phi)$ in Appendix B, we have the following result.

Remark 2. Assume $G_1 = 0$. Let \mathbf{m} be the mean of ϕ , which was defined as Eq. (20a). Take the strategy \mathbf{p} satisfying

$$\mathbf{p} = \mathbf{m}.$$

Then for any strategy \mathbf{p}'

$$W(\mathbf{p}'|\phi) = W(\mathbf{p}'|\mathbf{p})$$

holds.

This remark shows that when $G_1 = 0$ the distribution ϕ is thoroughly represented by \mathbf{p} . It suggests that strategies which coexist in the population contribute to the payoff only additively. Nonlinearity appears only when $G_1 \neq 0$.

Let us begin our analysis. We examine the *public goods game*, where G_1 vanishes. First we focus on the case with $w = 1$. Define the average discriminatoriness and the average forgivingness of the population as U and V , respectively:

$$U = \int u(\mathbf{p})\phi(\mathbf{p}) \, d\mathbf{p} = \int q(p_1 - p_2)\phi(\mathbf{p}) \, d\mathbf{p} = q(m_1 - m_2), \quad (21a)$$

$$V = \int v(\mathbf{p})\phi(\mathbf{p}) \, d\mathbf{p} = \int \{(1 - q)p_0 + qp_2\}\phi(\mathbf{p}) \, d\mathbf{p} = (1 - q)m_0 + qm_2. \quad (21b)$$

According to Appendix B, the payoff function is given by

$$\begin{aligned} W(\mathbf{p}'|\phi) &= (bU - c)\left(u' \frac{V}{1 - U} + v'\right) + bV \\ &= (bU - c)\left\{(1 - q)p'_0 + q \frac{V}{1 - U} p'_1 \right. \\ &\quad \left. + q\left(1 - \frac{V}{1 - U}\right)p'_2\right\} + bV. \end{aligned} \quad (22)$$

Because $(1 - q)$, $V/(1 - U)$, and $(1 - (V/(1 - U)))$ are all nonnegative, the sign of $(bU - c)$ is of critical importance. It changes at the threshold given by

$$U_0 = \frac{c}{b}. \quad (23)$$

When $U > U_0$, larger p'_i 's will be favored. It is All C strategy that gains the largest payoff there. When $U < U_0$ holds, on the other hand, smaller p'_i 's are favored. Hence All D strategy is the best in that case. The discriminating strategy is neither the best nor the worst in both cases; it is moderately good. When U is exactly the same as U_0 , the population is at an equilibrium and ϕ does not change any more.

When we assume that the payoff in the game is equal to the fitness (strictly speaking, when we assume that the fitness of an individual is given by his payoff plus basic fitness), we have the following differential equation on ϕ :

$$\dot{\phi}(\mathbf{p}) = \left[W(\mathbf{p}|\phi) - \int W(\mathbf{r}|\phi)\phi(\mathbf{r}) \, d\mathbf{r} \right] \phi(\mathbf{p}). \quad (24)$$

One possible way to investigate it is to use the Price equation (Price, 1970), which tells us the development of the mean of strategies, \mathbf{m} . However, we also need to know the development of \mathbf{H} in order to examine the change of \mathbf{m} , and we have to know the third-order moment of ϕ in order to examine \mathbf{H} , and so on. Hence

we will try to understand the dynamics (24) in a more intuitive way.

Figs. 6(a)–6(c) illustrate the change of the frequency of strategies when they are distributed uniformly at the initial population. Note that we pay attention to the evolution of p_1 and p_2 in those figures: the evolution of p_0 is not sketched. In this simulation, $b = 3$ and $c = 1$, hence $U_0 = 1/3$. We can see that the frequency of All D strategy rapidly grows. This is because U never exceeds U_0 . According to Fig. 6(d), U increases temporarily, but then it begins to decrease.

In order for cooperation to evolve, U must exceed U_0 . Consider the discriminating strategies $\mathbf{p} = (p_0, 1, 0)$. Since $u(\mathbf{p}) = q(1 - 0) = q$, which gives the maximum of u in all possible strategies, the existence of discriminating strategies in the population greatly contributes to the size of U . When the discriminatoriness becomes sufficient thanks to them and U exceeds U_0 , then the selection favors larger p'_i 's. According to Eq. (22), however, All C ($= (1, 1, 1)$) strategy gains larger payoff than the discriminating strategies. As a result, a plenty of discriminators in the group will help All C strategy evolve. Fig. 7 illustrates this process. When the initial distribution has a peak around discriminating strategies, large U owing to that distribution nourishes All C strategy more than discriminators. Consequently, the frequency of All C strategists increases. However, since All C strategy does not have any discriminatoriness in itself (i.e. $u(\text{All C}) = 0$), the average discriminatoriness U begins to decrease and asymptotically approaches the level at which $U = U_0$. Therefore, we state that discriminators are not a destination of the evolution but only an activator of unconditional cooperation.

This scenario becomes much clearer when we adopt the ‘best-response dynamics’ (Matsui, 1992; Hofbauer and Sigmund, 1998) instead of the selection dynamics given by Eq. (24). In the best-response dynamics, a fraction of players randomly sampled from the population change their strategies to the one that is the best-response to the present population. This dynamics includes the process in which players seek the best strategy in their environment, so it is a kind of learning dynamics. Suppose that there are so many discriminators that $U > U_0$ holds initially. In this case the best response is All C $= (1, 1, 1)$, so the frequency of All C gradually increases. Since $u(\text{All C}) = 0$, the upsurge of All C diminishes U , and ultimately U reaches U_0 . The population with large U is sure to go back to the equilibrium. Moreover, when a strategy with small u invades this population, U becomes less than U_0 . Once this occurs, the best response to the population becomes All D $= (0, 0, 0)$ so defectors spread in the population. Since the discriminatoriness of All D strategy, $u(\text{All D})$, is also equal to zero, U will decrease further. Hence the equilibrium represented by $U = U_0$ is unstable. The population can never be back to this point.

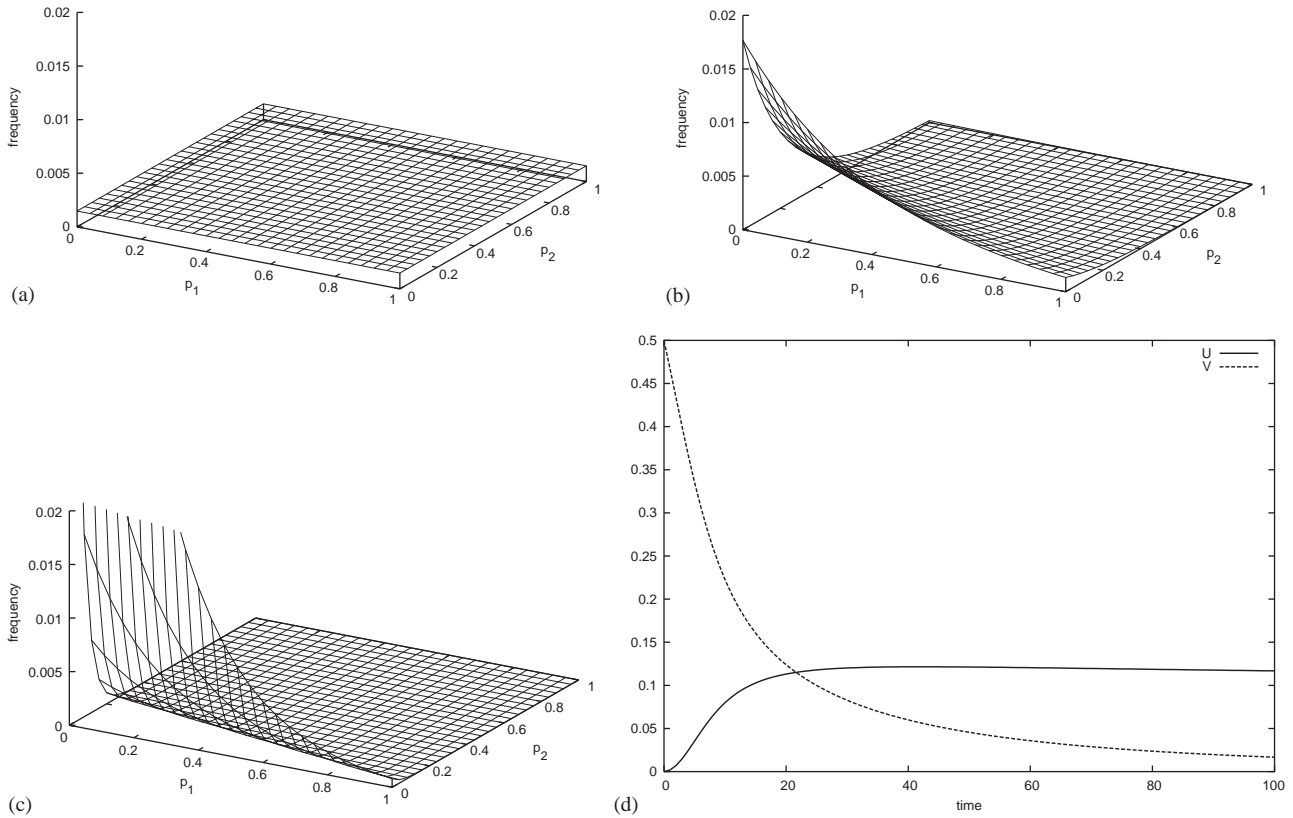


Fig. 6. Simulation results of Eq. (24) when the game is the *public goods game* and $w = 1$. We adopted $26^3 = 17576$ strategies; $(p_0, p_1, p_2) = (n_0/25, n_1/25, n_2/25)$ ($0 \leq n_0, n_1, n_2 \leq 25$). The payoff function is given by Eq. (22). Other parameters: $b = 3$, $c = 1$, and $q = 0.8$, hence $U_0 = c/b = 1/3$ is the threshold at which the landscape of the dynamics changes. (a)–(c) The change of frequency of each strategy when the initial distribution is the uniform distribution. Figures describe the marginal distribution of p_1 and p_2 ; the evolution of p_0 is not sketched. As time passes ((a) $t = 0$, (b) $t = 10$, and (c) $t = 50$), the frequency of All D strategy explodes. (d) The change of the average discriminatoriness (U), and the average forgivingness (V) of the population. The solid line represents U , and the dotted line V . Because the population starts from the uniform distribution, $U = 0$ holds at the initial state. The figure shows that U temporarily increases, but soon it comes to decrease. Due to lack of variance in strategies, however, the decrease of U is very slow. Since U never exceeds the threshold $U_0 = 1/3$, the sign of $(bU - c)$ is always negative. This causes selection to favor All D strategy.

When $w \neq 1$, the payoff function is

$$w(\mathbf{p}'|\phi) = (bwU - c) \times \left[\frac{1-wq}{1-w} p'_0 + \frac{wq}{1-w} \frac{(1-w)m_0 + wV}{1-wU} p'_1 + \frac{wq}{1-w} \left(1 - \frac{(1-w)m_0 + wV}{1-wU} \right) p'_2 \right] + b \left(m_0 + \frac{w}{1-w} V \right). \quad (25)$$

Note that coefficients of p'_i 's in the square brackets in Eq. (25) are all positive. Hence the discussion similar to the one for $w = 1$ holds. In this case, what is important is the sign of $(bwU - c)$. The threshold of U , at which the tendency of the evolutionary dynamics changes, is given by

$$U_0 = \frac{c}{bw}. \quad (26)$$

Again, discriminators are not an evolutionary outcome, but they are only a promoter of unconditional cooperation.

6. Discussion

Since [Trivers \(1971\)](#) proposed reciprocal altruism as a mechanism realizing cooperation between unrelated individuals, many game theorists and biologists have investigated both in theories ([Axelrod, 1984](#); [Axelrod and Hamilton, 1981](#); [Boyd and Richerson, 1992](#); [Hauert et al., 2002](#); [Lotem et al., 1999](#); [Nowak and Sigmund, 1998b](#); [Zahavi, 1995](#)) and in experiments ([Clements and Stephens, 1995](#); [Fehr and Gächter, 2000](#); [Stephens et al., 2002](#); [Wedekind and Milinski, 2000](#); [Wilkinson, 1984](#)) how cooperative behavior has evolved.

In this paper, we have discussed the evolution of reactive strategies in indirect reciprocity. In the context of *direct* reciprocity, [Nowak and Sigmund \(1990\)](#) extensively studied the invasion dynamics involved. Apparently, the strength and stability of Tit For Tat strategy give us the evidence that the evolution of cooperation between a fixed dyad is based on reciprocity. However, the result in the [Nowak and Sigmund](#)

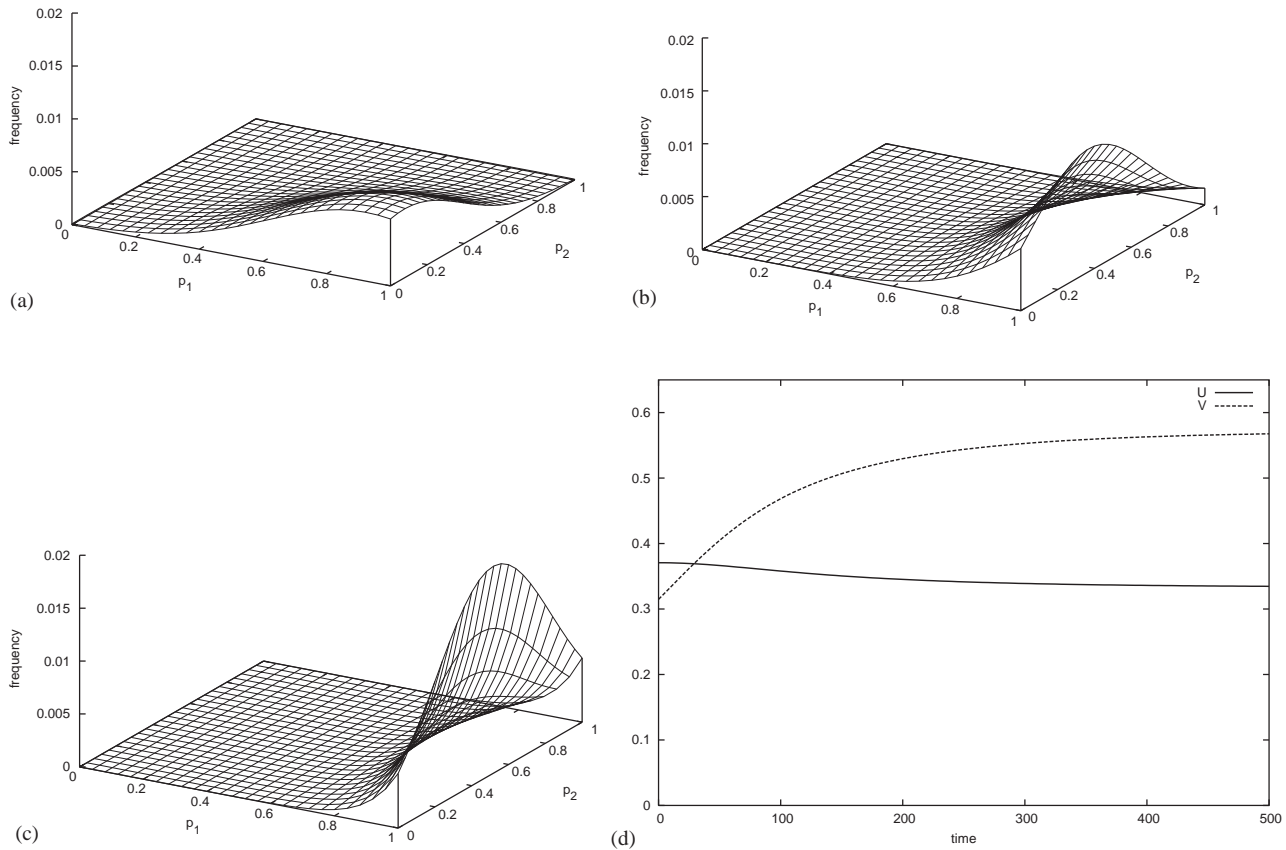


Fig. 7. Simulation results of Eq. (24) when the game is the *public goods game* and $w = 1$. We adopted 26^3 strategies which are the same as those in Fig. 6. The payoff function is given by Eq. (22). Other parameters are the same as Fig. 6. (a)–(c) The change of frequency of each strategy when the initial distribution has a gentle peak near the discriminating strategy. Figures describe the marginal distribution of p_1 and p_2 ; the evolution of p_0 is not sketched. As time passes ((a) $t = 0$, (b) $t = 100$, and (c) $t = 400$), the peak moves toward All C strategy. (d) The change of the average discriminativeness (U), and the average forgivingness (V) of the population. Because U is larger than the threshold $U_0 = 1/3$, selection favors All C strategy the best. However, as the peak shifts toward All C, which has less discriminativeness than the discriminating strategy, U gradually approaches the threshold, then the population will fall into an equilibrium.

(1990) was somewhat counterintuitive; a strategy does not evolve toward TFT. They showed that TFT is only the pivot of the evolution. They also pointed out that the victory of TFT in Axelrod's (1984) tournament was due to the special composition of the population.

Meanwhile, in the context of *indirect* reciprocity, the discriminating strategy has been attracting great interest as the one which gives us a plausible explanation of the evolution of indirect reciprocity. The experiment conducted by Wedekind and Milinski (2000) also gave us the evidence that humans actually use the discriminating strategy. The present paper is motivated by the question whether the discriminating strategy can be attained through the strategic evolution, as TFT cannot. It is an extension of the Nowak and Sigmund's (1990) paper to the *indirect* version of reciprocity.

First we note that our results in adaptive dynamics have great similarity to the one where help is directly returned back from the beneficiary. Although situations are quite different, the results are similar. It is very surprising. In *direct* reciprocity those who play the game

with a mutant of course played the game with the same mutant in the previous round, because the opponent of the game is always fixed. On the other hand, in *indirect* reciprocity those who play the game with a mutant played the game with a wild-type player in the previous round. This is because in indirect reciprocity the partner of the game is never the same. The result we derived in this paper suggests that the difference between those two is negligible; it makes quite little difference. The reason why those two different approaches lead to the similar results lies in the 'information transfer' q . Without social information, a player cannot expel cheaters because the partner of the game is never the same in the two consecutive rounds hence he has no ways to know what his opponent did in the previous interaction. When social reputation of others is available (this is equivalent to that q is large), however, they can assess and monitor each other by using that information, thus it becomes possible to eject only defectors from the population. The fact that Eq. (5) is the same as the one in Nowak and Sigmund (1990) at the limit $q \rightarrow 1$ is eloquent of this argument.

We mainly investigated the case where G_1 vanishes. This condition holds for the *public goods game*. First, we found that the discriminating strategy is not a destination of the evolution. Strategies which are more generous toward defectors are always able to invade the population of discriminating strategies, except for that of the nice discriminators' (1, 1, 0). But this exception is also excluded when we introduce perception errors or implementation errors into the model. When the discriminating strategies are in the *C-region*, they will be invaded by a more cooperative strategy toward defectors. This suggests that players using the discriminating strategy have temptation to change their strategy to a more cooperative one. When one reacts to defectors by defection, this defection works as punishment to them. Thus defectors are able to gain only the payoff P , whereas the others with the discriminating strategy keep enjoying the benefit of mutual cooperation, R . That is how the population of the discriminating strategy is stably maintained against social exploiters. However, those who reacted to defectors by defection will lose their own good reputation. Then they will be defected against in the next round due to that bad reputation. Hence the incentive to punish defectors vanishes and discriminators are inclined to cooperate even with defectors.

Second, the whole strategy space is divided into the *C-region* and the *D-region* and no tendencies exist to discriminate others by using their reputation. This result also elucidates that players have no adaptive motivations to adopt the discriminating strategies. Usually, discriminating others by social information is costly to a player, since he has to keep sensitive to it. We predict that such a cost further prevents the evolution of the discriminating strategies.

We also investigated the effect of two kinds of errors. It is not until some defections are actually performed that we see the difference between indiscriminate altruists and discriminators. We found that those errors urge the population of discriminators toward All C strategy more. In other words, generosity toward defectors, that is p_2 , increases easily due to them. This result seems inconsistent with what Lotem et al. (1999) derived, in which the existence of 'phenotypic defectors', who are unable to help others due to their poor phenotypic conditions, stabilizes the discriminating strategy. The reason for the difference is as follows. In Lotem et al. (1999), they adopted an 'image score' (Nowak and Sigmund, 1998b) as a measure of social reputation. Since it ranges from -5 to $+5$ (i.e. 11 ranks), one defection against a defector causes decrease of his image score only by one unit, which undermines his reputation quite little. On the contrary, the defection is even favorable for him because he can avoid paying the costs of help. Thus defectors in the population further promote the evolution of discriminators. However,

when the focal group adopts such a severe reputation rule as the one in our model, say a binary rule (either good or bad), one defection directly results in his bad reputation, then it causes him to lose his future benefit $+b$ which otherwise he could gain. In other words, he misses b in exchange for saving the cost c . Since b is larger than c , not cooperating with defectors brings a bad result to him.

In Section 5, we studied the effects of heterogeneity in strategies, assuming that a variety of different reactive strategies coexist in the population. We found that the payoff of each player is dependent only on the mean, the variance, and the covariance of the distribution ϕ . This result made it easy for us to study how each strategy evolves under the general condition. In particular, we concentrated on the *public goods game*, where payoff is dependent only on the mean of ϕ . Under such settings, however, the discriminating strategies are not an end point of the evolution, either. Eq. (22) tells us that the discriminating strategies are intermediately good ones in terms of payoff; neither the worst nor the best. Instead, they play an important role in their contribution to the average discriminatoriness in the population, U . The quantity U is of critical importance in the population in that it determines the direction of the evolution. The existence of strategies with high discriminatoriness leads the society toward a cooperative one. Hence we can understand the discriminating strategies as a promoter of cooperative behavior. We are sure that this result remains unchanged even when two kinds of errors above are considered.

Explaining the evolution of cooperation is explaining how individuals in a society expel cheaters, free-riders, or social parasites, who enjoy social welfare with paying nothing. In the Prisoner's Dilemma game, the action D dominates C; a player gains a larger payoff if he chooses D (= defection) irrespective of his opponent's decision. Therefore, the mutual cooperation is always vulnerable to invasions by social exploiters. It is very fragile. The discriminating strategy, like Tit For Tat strategy, attempts detecting those free-riders by using social reputation and retaliates against them by the subsequent defection. Thus the cooperative society would be maintained. However, as mentioned above, it is a problem who will punish the defector. This is the well-known 'second-order free-rider problem'. 'Contribute Tit For Tat' (Boyd, 1989) or 'standing strategy' (Sugden, 1986), and 'reputation discriminator (RDISC)' (Panathan and Boyd, 2003) have been proposed to solve this issue. In their model the way to assign social status (called *standing*) differs from the one we have used in the present paper. Players using those strategies do not lose their good *standing* even if they defect from those who have a bad *standing*. With this type of social status, an incentive to retaliate against defectors reappears. Since defecting from those who have a bad standing does not

undermine one's good standing at all, D becomes a fascinating choice here for players whose present opponent has a bad standing. The difference between reputation, which in our model is equivalent to the action in the previous round, and the *standing* lies in the notion of fairness. Players with the standing strategy clearly distinguish retaliatory defection from selfish defection. They find the former fair but the latter unfair. On the other hand, players with discriminating strategies regard both defection as the same. Our results suggest that not only cooperation/defection but also those two kinds of defections should be distinguished if we are to explain the evolution of indirect reciprocity. In fact, Panchanathan and Boyd (2003) showed that with implementation errors in the population those strategies which can distinguish justified defection from unjustified one are evolutionarily stable whereas such strategies as the discriminating strategy cannot. This result agrees with our view. Yet it is still an open problem whether and how such a sort of fairness has evolved.

Recently, Fehr and Gächter (2000, 2002) found that people get more cooperative when altruistic punishment, retaliation which costs the actor but brings nothing to himself, exists. In contrast, Fehr and Rockenbach (2003) found that people get less cooperative when sanctions based on selfish motive exist. These results are totally inconsistent with the rationality hypothesis, which game theory assumes. A player with perfect rationality expects altruistic punishment not to be achieved due to its cost, hence he would keep defecting. On the other hand, he would become more cooperative when selfish sanctions exist, for he expects that kind of punishment to be fulfilled. As a result, game theory predicts the opposite outcomes to the experimental results of theirs. We know little about human fairness. We long for further research.

Acknowledgements

We wish to thank Kazuyuki Aihara, Koh Hashimoto, Yoh Iwasa, Sunao Murashige, Akira Sasaki, and Karl Sigmund for their precious and helpful comments.

Appendix A. Calculation of $W(\mathbf{p}'|\mathbf{p})$

First we will solve recursions (4a)–(4c). Whether $u = u(\mathbf{p})$ equals unity or not affects the results.

(i) When $u \neq 1$,

$$c_n = u^{n-1} \left(p_0 - \frac{v}{1-u} \right) + \frac{v}{1-u},$$

$$c'_n = \begin{cases} u' u^{n-2} \left(p_0 - \frac{v}{1-u} \right) + \left(\frac{v}{1-u} u' + v' \right) & (n \geq 2), \\ p'_0 & (n = 1), \end{cases}$$

$$c''_n = \begin{cases} u' u^{n-2} \left(p_0 - \frac{v}{1-u} \right) + \left\{ u \left(\frac{v}{1-u} u' + v' \right) + v \right\} & (n \geq 3), \\ u p'_0 + v & (n = 2), \\ p_0 & (n = 1), \end{cases} \quad (\text{A.1})$$

where $u = u(\mathbf{p})$, $u' = u(\mathbf{p}')$, $v = v(\mathbf{p})$, and $v' = v(\mathbf{p}')$.

(ii) When $u = 1$,

$$c_n = p_0$$

$$c'_n = \begin{cases} u' p_0 + v' & (n \geq 2), \\ p'_0 & (n = 1), \end{cases}$$

$$c''_n = \begin{cases} u' p_0 + v' & (n \geq 3), \\ p'_0 & (n = 2), \\ p_0 & (n = 1). \end{cases} \quad (\text{A.2})$$

Since $W_n(\mathbf{p}'|\mathbf{p})$ is given by

$$W_n(\mathbf{p}'|\mathbf{p}) = R c'_n c''_n + S c'_n (1 - c''_n) + T (1 - c'_n) c''_n + P (1 - c'_n) (1 - c''_n) \\ = G_1 c'_n c''_n + G_2 c'_n + G_3 c''_n + G_4, \quad (\text{A.3})$$

where

$$G_1 = R - S - T + P, \\ G_2 = S - P, \\ G_3 = T - P, \\ G_4 = P, \quad (\text{A.4})$$

the payoff function $W(\mathbf{p}'|\mathbf{p})$ is given as follows:

$$W(\mathbf{p}'|\mathbf{p}) = G_1 \Gamma_1 + G_2 \Gamma_2 + G_3 \Gamma_3 + G_4 \Gamma_4. \quad (\text{A.5})$$

Gammas above are dependent on the parameters w , and $u = u(\mathbf{p})$.

(a) For $0 \leq w < 1$:

(a.1) When $-1 \leq u < 1$,

$$\Gamma_1 = p_0 p'_0 + w(u' p_0 + v')(u p'_0 + v) + \frac{w^2}{1 - wu^2} (uu' A)^2 \\ + \frac{w^2}{1 - wu} (uu' A) \{ (u + 1) B' + v \} \\ + \frac{w^2}{1 - w} B' (u B' + v),$$

$$\Gamma_2 = p'_0 + \frac{w}{1 - wu} (u' A) + \frac{w}{1 - w} B',$$

$$\Gamma_3 = wu \Gamma_2 + \left(p_0 + \frac{w}{1 - w} v \right),$$

$$\Gamma_4 = \frac{1}{1 - w}, \quad (\text{A.6})$$

where $A = p_0 - (v/(1 - u))$, and $B' = (v/(1 - u)) u' + v'$.

(a.2) When $u = 1$,

$$\begin{aligned}\Gamma_1 &= p_0 p'_0 + w(u'p_0 + v')p'_0 + \frac{w^2}{1-w}(u'p_0 + v')^2, \\ \Gamma_2 &= p'_0 + \frac{w}{1-w}(u'p_0 + v'), \\ \Gamma_3 &= w\Gamma_2 + p_0, \\ \Gamma_4 &= \frac{1}{1-w}.\end{aligned}\quad (\text{A.7})$$

(b) For $w = 1$:

(b.1) When $-1 < u < 1$,

$$\begin{aligned}\Gamma_1 &= B'(uB' + v), \\ \Gamma_2 &= B', \\ \Gamma_3 &= uB' + v, \\ \Gamma_4 &= 1.\end{aligned}\quad (\text{A.8})$$

(b.2) When $u = 1$,

$$\begin{aligned}\Gamma_1 &= (u'p_0 + v')^2, \\ \Gamma_2 &= u'p_0 + v', \\ \Gamma_3 &= u'p_0 + v', \\ \Gamma_4 &= 1.\end{aligned}\quad (\text{A.9})$$

(b.3) When $u = -1$,

$$\begin{aligned}\Gamma_1 &= (u'A)^2 + B'(1 - B'), \\ \Gamma_2 &= B', \\ \Gamma_3 &= 1 - B', \\ \Gamma_4 &= 1.\end{aligned}\quad (\text{A.10})$$

Differentiating Eq. (A.6), we have $f_i(\mathbf{p})$'s in Eq. (12) as follows:

$$\begin{aligned}\left.\frac{\partial \Gamma_1}{\partial p'_0}\right|_{\mathbf{p}'=\mathbf{p}} &= f_0(\mathbf{p}) \\ &= p_0 + wu(up_0 + v) + (1 - q) \\ &\quad \times \left[w(up_0 + v) + \frac{w^2 u A}{1 - wu} u(u + 1) \right. \\ &\quad \left. + \frac{w^2}{1 - w} \frac{1 + u}{1 - u} v \right],\end{aligned}$$

$$\begin{aligned}\left.\frac{\partial \Gamma_1}{\partial p'_1}\right|_{\mathbf{p}'=\mathbf{p}} &= f_1(\mathbf{p}) \\ &= q \left[w(up_0 + v)p_0 + \frac{2w^2 u^3 A^2}{1 - wu^2} + \frac{w^2 u A}{1 - wu} \right. \\ &\quad \times \left\{ \frac{2v}{1 - u} + u(u + 1) \frac{v}{1 - u} \right\} \\ &\quad \left. + \frac{w^2}{1 - w} \frac{1 + u}{1 - u} v \frac{v}{1 - u} \right],\end{aligned}$$

$$\begin{aligned}\left.\frac{\partial \Gamma_1}{\partial p'_2}\right|_{\mathbf{p}'=\mathbf{p}} &= f_2(\mathbf{p}) \\ &= q \left[w(up_0 + v)(1 - p_0) - \frac{2w^2 u^3 A^2}{1 - wu^2} + \frac{w^2 u A}{1 - wu} \right. \\ &\quad \times \left\{ -\frac{2v}{1 - u} + u(u + 1) \left(1 - \frac{v}{1 - u} \right) \right\} \\ &\quad \left. + \frac{w^2}{1 - w} \frac{1 + u}{1 - u} v \left(1 - \frac{v}{1 - u} \right) \right],\end{aligned}\quad (\text{A.11})$$

where $A = p_0 - (v/(1 - u))$.

Appendix B. Calculation of $W(\mathbf{p}'|\phi)$

Of course, $c_1(\mathbf{p})$ is given by

$$c_1(\mathbf{p}) = p_0. \quad (\text{B.1})$$

For $n \geq 2$, we have the following relation:

$$\begin{aligned}c_n(\mathbf{p}) &= \int \{ (1 - q)p_0 + qc_{n-1}(\mathbf{r})p_1 + q(1 - c_{n-1}(\mathbf{r})) \\ &\quad \times p_2 \} \phi(\mathbf{r}) \, d\mathbf{r}.\end{aligned}\quad (\text{B.2})$$

Now, let C_n be the average cooperation rate over the population in round n . It is defined as

$$C_n = \int c_n(\mathbf{p}) \phi(\mathbf{p}) \, d\mathbf{p}. \quad (\text{B.3})$$

Then Eq. (B.2) is rewritten as follows:

$$\begin{aligned}c_n(\mathbf{p}) &= (1 - q)p_0 + qp_1 C_{n-1} + qp_2(1 - C_{n-1}) \\ &= u(\mathbf{p})C_{n-1} + v(\mathbf{p}) \quad (n \geq 2).\end{aligned}\quad (\text{B.4})$$

Calculating the integral of Eq. (B.4) with respect to \mathbf{p} yields

$$\begin{aligned}C_n &= \int c_n(\mathbf{p}) \phi(\mathbf{p}) \, d\mathbf{p} \\ &= \int \{ u(\mathbf{p})C_{n-1} + v(\mathbf{p}) \} \phi(\mathbf{p}) \, d\mathbf{p} \\ &= \left(\int u(\mathbf{p}) \phi(\mathbf{p}) \, d\mathbf{p} \right) C_{n-1} + \int v(\mathbf{p}) \phi(\mathbf{p}) \, d\mathbf{p} \\ &\quad (n \geq 2).\end{aligned}\quad (\text{B.5})$$

Define the average discriminativeness of the population, U , and the average forgivingness of the population, V , as follows:

$$\begin{aligned}U &= \int u(\mathbf{p}) \phi(\mathbf{p}) \, d\mathbf{p} = \int q(p_1 - p_2) \phi(\mathbf{p}) \, d\mathbf{p} \\ &= q(m_1 - m_2),\end{aligned}\quad (\text{B.6a})$$

$$\begin{aligned}V &= \int v(\mathbf{p}) \phi(\mathbf{p}) \, d\mathbf{p} = \int \{ (1 - q)p_0 + qp_2 \} \phi(\mathbf{p}) \, d\mathbf{p} \\ &= (1 - q)m_0 + qm_2.\end{aligned}\quad (\text{B.6b})$$

Then we have the following recursion on C_n :

$$\begin{aligned} C_1 &= m_0, \\ C_n &= UC_{n-1} + V \quad (n \geq 2). \end{aligned} \quad (\text{B.7})$$

This can be solved analytically. When $-1 \leq U < 1$, C_n is given by

$$C_n = U^{n-1} \left(m_0 - \frac{V}{1-U} \right) + \frac{V}{1-U}, \quad (\text{B.8})$$

hence $c_n(\mathbf{p})$ is

$$c_n(\mathbf{p}) = \begin{cases} u(\mathbf{p}) \left\{ U^{n-2} \left(m_0 - \frac{V}{1-U} \right) + \frac{V}{1-U} \right\} + v(\mathbf{p}) & (n \geq 2), \\ p_0 & (n = 1). \end{cases} \quad (\text{B.9})$$

When $U = 1$, $C_n = m_0$, thus

$$c_n(\mathbf{p}) = \begin{cases} u(\mathbf{p})m_0 + v(\mathbf{p}) & (n \geq 2), \\ p_0 & (n = 1) \end{cases} \quad (\text{B.10})$$

holds.

Next we derive $W_n(\mathbf{p}'|\phi)$. In round n , there are four possible outcomes for a \mathbf{p}' strategist; (self, opponent) = (C, C), (C, D), (D, C), and (D, D). We denote each probability by $e_{1,n}(\mathbf{p}')$, $e_{2,n}(\mathbf{p}')$, $e_{3,n}(\mathbf{p}')$, and $e_{4,n}(\mathbf{p}')$, respectively. For $n = 1$, they are $p'_0 m_0$, $p'_0(1 - m_0)$, $(1 - p'_0)m_0$, and $(1 - p'_0)(1 - m_0)$ in this order. For $n \geq 2$, each of them is given by

$$e_{1,n}(\mathbf{p}') = \int \{u(\mathbf{p}')c_{n-1}(\mathbf{r}) + v(\mathbf{p}')\} \cdot \{u(\mathbf{r})c_{n-1}(\mathbf{p}') + v(\mathbf{r})\} \times \phi(\mathbf{r}) \, d\mathbf{r},$$

$$e_{2,n}(\mathbf{p}') = \int \{u(\mathbf{p}')c_{n-1}(\mathbf{r}) + v(\mathbf{p}')\} \cdot [1 - \{u(\mathbf{r})c_{n-1}(\mathbf{p}') + v(\mathbf{r})\}] \phi(\mathbf{r}) \, d\mathbf{r},$$

$$e_{3,n}(\mathbf{p}') = \int [1 - \{u(\mathbf{p}')c_{n-1}(\mathbf{r}) + v(\mathbf{p}')\}] \cdot \{u(\mathbf{r})c_{n-1}(\mathbf{p}') + v(\mathbf{r})\} \phi(\mathbf{r}) \, d\mathbf{r},$$

$$e_{4,n}(\mathbf{p}') = \int [1 - \{u(\mathbf{p}')c_{n-1}(\mathbf{r}) + v(\mathbf{p}')\}] \cdot [1 - \{u(\mathbf{r})c_{n-1}(\mathbf{p}') + v(\mathbf{r})\}] \phi(\mathbf{r}) \, d\mathbf{r}. \quad (\text{B.11})$$

Calculation shows that for $n \geq 2$

$$\begin{aligned} e_{2,n}(\mathbf{p}') &= c_n(\mathbf{p}') - e_{1,n}(\mathbf{p}'), \\ e_{3,n}(\mathbf{p}') &= \{Uc_{n-1}(\mathbf{p}') + V\} - e_{1,n}(\mathbf{p}'), \\ e_{4,n}(\mathbf{p}') &= 1 - c_n(\mathbf{p}') - \{Uc_{n-1}(\mathbf{p}') + V\} + e_{1,n}(\mathbf{p}'). \end{aligned} \quad (\text{B.12})$$

Then we have

$$W_n(\mathbf{p}'|\phi) = \begin{cases} G_1 e_{1,n}(\mathbf{p}') + G_2 c_n(\mathbf{p}') + G_3 \{Uc_{n-1}(\mathbf{p}') + V\} + G_4 & (n \geq 2), \\ G_1 p'_0 m_0 + G_2 p'_0 + G_3 m_0 + G_4 & (n = 1). \end{cases} \quad (\text{B.13})$$

Since $e_{1,n}(\mathbf{p}')$ ($n \geq 2$) is given by

$$e_{1,n}(\mathbf{p}') = c_n(\mathbf{p}') \cdot \{Uc_{n-1}(\mathbf{p}') + V\} + u(\mathbf{p}') [\boldsymbol{\eta}_n \cdot \mathbf{H} \boldsymbol{\xi}_n] \quad (\text{B.14})$$

$[\boldsymbol{\eta}_n \cdot \mathbf{H} \boldsymbol{\xi}_n]$ in Eq. (B.14) represents the inner product of column vectors $\boldsymbol{\eta}_n$ and $\mathbf{H} \boldsymbol{\xi}_n$, where

$$\boldsymbol{\eta}_n = \begin{cases} \begin{pmatrix} 1 - q \\ qC_{n-2} \\ q(1 - C_{n-2}) \end{pmatrix} & (n \geq 3), \\ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} & (n = 2) \end{cases} \quad (\text{B.15})$$

and

$$\boldsymbol{\xi}_n = \begin{pmatrix} 1 - q \\ qc_{n-1}(\mathbf{p}') \\ q(1 - c_{n-1}(\mathbf{p}')) \end{pmatrix} \quad (n \geq 2), \quad (\text{B.16})$$

we have the following result:

Remark B.1. \mathbf{p}' strategist's payoff in the population of ϕ is

$$W(\mathbf{p}'|\phi) = G_1 \Gamma_1 + G_2 \Gamma_2 + G_3 \Gamma_3 + G_4 \Gamma_4. \quad (\text{B.17})$$

Gammas in Eq. (B.17) are given as follows:

(a) For $0 \leq w < 1$:

(a.1) When $-1 \leq U < 1$,

$$\begin{aligned} \Gamma_1 &= m_0 p'_0 + w(u'm_0 + v')(Up'_0 + V) + \frac{w^2}{1 - wU^2} (Uu'E)^2 \\ &\quad + \frac{w^2}{1 - wU} (Uu'E) \{ (U + 1)F' + V \} + \frac{w^2}{1 - w} \\ &\quad \times F'(UF' + V) + \sum_{n=2}^{\infty} w^{n-1} u' [\boldsymbol{\eta}_n \cdot \mathbf{H} \boldsymbol{\xi}_n], \end{aligned} \quad (\text{B.18})$$

$$\Gamma_2 = p'_0 + \frac{w}{1 - wU} (u'E) + \frac{w}{1 - w} F',$$

$$\Gamma_3 = wU\Gamma_2 + \left(m_0 + \frac{w}{1 - w} V \right),$$

$$\Gamma_4 = \frac{1}{1 - w},$$

where $u' = u(\mathbf{p}')$, $v' = v(\mathbf{p}')$, $E = m_0 - (V/(1 - U))$, and $F' = (V/(1 - U)) u' + v'$.

(a.2) When $U = 1$,

$$\begin{aligned}\Gamma_1 &= m_0 p'_0 + w(u'm_0 + v')p'_0 + \frac{w^2}{1-w}(u'm_0 + v')^2 \\ &\quad + \sum_{n=2}^{\infty} w^{n-1} u'[\eta_n \cdot H\xi_n], \\ \Gamma_2 &= p'_0 + \frac{w}{1-w}(u'm_0 + v'), \\ \Gamma_3 &= w\Gamma_2 + m_0, \\ \Gamma_4 &= \frac{1}{1-w}.\end{aligned}\quad (\text{B.19})$$

(b) For $w = 1$:

(b.1) When $-1 < U < 1$,

$$\begin{aligned}\Gamma_1 &= F'(UF' + V) + u'[\eta_* \cdot H\xi_*], \\ \Gamma_2 &= F', \\ \Gamma_3 &= UF' + V, \\ \Gamma_4 &= 1,\end{aligned}\quad (\text{B.20a})$$

where

$$\begin{aligned}\eta_* &= \begin{pmatrix} 1-q \\ q \frac{V}{1-U} \\ q \left(1 - \frac{V}{1-U}\right) \end{pmatrix}, \\ \xi_* &= \begin{pmatrix} 1-q \\ q F' \\ q(1-F') \end{pmatrix}.\end{aligned}\quad (\text{B.20b})$$

(b.2) When $U = 1$,

$$\begin{aligned}\Gamma_1 &= (u'm_0 + v')^2 + u'[\eta_* \cdot H\xi_*], \\ \Gamma_2 &= u'm_0 + v', \\ \Gamma_3 &= u'm_0 + v', \\ \Gamma_4 &= 1,\end{aligned}\quad (\text{B.21a})$$

where

$$\begin{aligned}\eta_* &= \begin{pmatrix} 1-q \\ q m_0 \\ q(1-m_0) \end{pmatrix}, \\ \xi_* &= \begin{pmatrix} 1-q \\ q(u'm_0 + v') \\ q\{1 - (u'm_0 + v')\} \end{pmatrix}.\end{aligned}\quad (\text{B.21b})$$

(b.3) When $U = -1$,

$$\begin{aligned}\Gamma_1 &= (u'E)^2 + F'(1-F') \\ &\quad + u' \cdot \frac{1}{2} [\eta_* \cdot H\xi_* + \eta_{**} \cdot H\xi_{**}], \\ \Gamma_2 &= F', \\ \Gamma_3 &= 1 - F', \\ \Gamma_4 &= 1,\end{aligned}\quad (\text{B.22})$$

where

$$\begin{aligned}\eta_* &= \begin{pmatrix} 1-q \\ q m_0 \\ q(1-m_0) \end{pmatrix}, \\ \xi_* &= \begin{pmatrix} 1-q \\ q(u'm_0 + v') \\ q\{1 - (u'm_0 + v')\} \end{pmatrix},\end{aligned}\quad (\text{B.23a})$$

and

$$\begin{aligned}\eta_{**} &= \begin{pmatrix} 1-q \\ q(1-m_0) \\ q m_0 \end{pmatrix}, \\ \xi_{**} &= \begin{pmatrix} 1-q \\ q(u'(1-m_0) + v') \\ q\{1 - (u'(1-m_0) + v')\} \end{pmatrix}.\end{aligned}\quad (\text{B.23b})$$

This remark shows that the Remark 1 in Section 5 holds.

References

- Alexander, R., 1987. The Biology of Moral Systems. Aldine de Gruyter, New York.
- Axelrod, R., 1984. The Evolution of Cooperation. Penguin, Harmondsworth.
- Axelrod, R., Hamilton, W., 1981. The evolution of cooperation. Science 211, 1390–1396.
- Boyd, R., 1989. Mistakes allow evolutionary stability in the repeated prisoner's dilemma game. J. Theor. Biol. 136, 47–56.
- Boyd, R., Richerson, P., 1992. Punishment allows the evolution of cooperation (or anything else) in sizeable groups. Ethol. Sociobiol. 13, 171–195.
- Clements, K., Stephens, D., 1995. Testing models of non-kin cooperation: mutualism and the prisoner's dilemma. Anim. Behav. 50, 527–535.
- Dieckmann, U., Law, R., 1996. The dynamical theory of coevolution: a derivation from stochastic ecological processes. J. Math. Biol. 34, 579–612.
- Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. The Amer. Econ. Rev. 90, 980–994.
- Fehr, E., Gächter, S., 2002. Altruistic punishment in humans. Nature 415, 137–140.
- Fehr, E., Rockenbach, B., 2003. Detrimental effects of sanctions on human altruism. Nature 422, 137–140.
- Hauert, C., De Monte, S., Hofbauer, J., Sigmund, K., 2002. Volunteering as red queen mechanism for cooperation in public goods games. Science 296, 1129–1132.
- Hofbauer, J., Sigmund, K., 1998. Evolutionary Games and Population Dynamics. Cambridge University Press, Cambridge, MA.
- Iwasa, Y., Pomiankowski, A., Nee, S., 1991. The evolution of costly mate preferences. ii. the 'handicap' principle. Evolution 45, 1431–1442.
- Leimar, O., Hammerstein, P., 2001. Evolution of cooperation through indirect reciprocity. Proc. R. Soc. London B 268, 745–753.

- Lotem, A., Fishman, M., Stone, L., 1999. Evolution of cooperation between individuals. *Nature* 400, 226–227.
- Matsui, A., 1992. Best response dynamics and socially stable strategies. *J. Econ. Theory* 57, 343–362.
- Nowak, M., 1990. Stochastic strategies in the prisoner's dilemma. *Theor. Popul. Biol.* 38, 93–112.
- Nowak, M., Sigmund, K., 1989. Game-dynamical aspects of the prisoner's dilemma. *Appl. Math. Comput.* 30, 191–213.
- Nowak, M., Sigmund, K., 1990. The evolution of stochastic strategies in the prisoner's dilemma. *Acta Appl. Math.* 20, 247–265.
- Nowak, M., Sigmund, K., 1998a. The dynamics of indirect reciprocity. *J. Theor. Biol.* 194, 561–574.
- Nowak, M., Sigmund, K., 1998b. Evolution of indirect reciprocity by image scoring. *Nature* 393, 573–577.
- Page, K., Nowak, M., 2002. Unifying evolutionary dynamics. *J. Theor. Biol.* 219, 93–98 doi:10.1006/jtbi.3112.
- Panchanathan, K., Boyd, R., 2003. A tale of two defectors: the importance of standing for evolution of indirect reciprocity. *J. Theor. Biol.* 224, 115–126 doi:10.1016/S0022-5193(03)00154-1.
- Price, G., 1970. Selection and covariance. *Nature* 227, 520–521.
- Stephens, D., McLinn, C., Stevens, J., 2002. Discounting and reciprocity in an iterated prisoner's dilemma. *Science* 298, 2216–2218.
- Sugden, R., 1986. *The Economics of Rights, Co-operation and Welfare*. Blackwell, Oxford, UK.
- Trivers, R., 1971. The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57.
- Wedekind, C., Milinski, M., 2000. Cooperation through image scoring in humans. *Science* 288, 850–852.
- Wilkinson, G., 1984. Reciprocal food sharing in the vampire bat. *Nature* 308, 181–184.
- Zahavi, A., 1995. Altruism as a handicap—the limitations of kin selection and reciprocity. *J. Avian Biol.* 26 (1), 1–3.