Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Theoretical Biology 239 (2006) 435–444
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

The leading eight: Social norms that can maintain cooperation by indirect reciprocity

Hisashi Ohtsuki*, Yoh Iwasa

Department of Biology, Faculty of Sciences, Kyushu University, Fukuoka 812-8581, Japan

Received 31 May 2005; received in revised form 7 August 2005; accepted 9 August 2005
Available online 19 September 2005

Abstract

The theory of indirect reciprocation explains the evolution of cooperation among unrelated individuals, engaging in one-shot interaction. Using reputation, a player acquires information on who are worth cooperating and who are not. In a previous paper, we formalized the reputation dynamics, a rule to assign a binary reputation (good or bad) to each player when his action, his current reputation, and the opponent's reputation are given. We then examined all the possible reputation dynamics, and found that there exist only eight reputation dynamics named ''leading eight'' that can maintain the ESS with a high level of cooperation, even if errors are included in executing intended cooperation and in reporting the observation to the public. In this paper, we study the nature of these successful social norms. First, we characterize the role of each pivot of the reputation dynamics common to all of the leading eight. We conclude that keys to the success in *indirect* reciprocity are to be nice (maintenance of cooperation among themselves), retaliatory (detection of defectors, punishment, and justification of punishment), apologetic, and forgiving. Second, we prove the two basic properties of the leading eight, which give a quantitative evaluation of the ESS condition and the level of cooperation maintained at the ESS.

Keywords: Cooperation; Reputation; Standing; Justified defection; Leading eight; Social norm

1. Introduction

The evolution of cooperation between unrelated individuals has been a research focus of evolutionary biology. In a classical paper, Trivers (1971) proposed reciprocal altruism (or *direct* reciprocity) that promotes cooperation between a dyad of players interacting repeatedly. This idea was later formalized mathematically as the repeated prisoner's dilemma game by Axelrod and Hamilton (1981) in which a cooperative, retaliatory and forgiving strategy, Tit For Tat, won in the round robin tournament. In contrast, the theory of indirect reciprocation explains cooperation even in a one-shot interaction. Using reputation, a player acquires information on who are worth cooperating and who are not, even if he has no experiences of direct interaction with them. If cheaters were effectively excluded by this mechanism, cooperation could be maintained in the population.

Nowak and Sigmund (1998a, b) examined this scenario mathematically for the first time by introducing the concept of image score. In Nowak and Sigmund (1998a), image score is a binary representation of a player's social reputation and is either *good* or *bad*. When a player gives help to others, his image score becomes *good*. On the other hand, refusal of giving yields *bad* image score to the actor. Under this updating rule, Nowak and Sigmund found that the discriminator strategy, who gives help only to *good* individuals, performed well.

*Corresponding author. Tel.: +81 92 642 2641; fax: +81 92 642 2645.
E-mail address: ohtsuki@bio-math10.biology.kyushu-u.ac.jp (H. Ohtsuki).

0022-5193/$ - see front matter
doi:10.1016/j.jtbi.2005.08.008

However, many subsequent studies have cast doubt on the image-score strategy, especially about the incentive to maintain the discriminator strategy. Discriminators refuse help against a *bad* recipient. Paradoxically this refusal brings about *bad* reputation to the actor. In other words, refusal of help against a bad opponent is not justified in the population dominated by discriminators themselves. Therefore it is beneficial for the actor to cooperate even with *bad* opponents in order to keep his *good* reputation. The population composed of discriminators is invaded and replaced by AllC, which is then followed by the invasion by AllD players. Indeed, Ohtsuki (2004) and Panchanathan and Boyd (2003) have shown the failure of discriminators who use the image-score criterion.

This suggests to us the need of a more sophisticated mechanism than assumed in the image score to judge whether a player is *good* or *bad*. Sugden (1986) had proposed a different rule of assessment, called standing criterion, in which refusal against a *bad* recipient does not undermine the donor's reputation. Leimar and Hammerstein (2001), Panchanathan and Boyd (2003), and Ohtsuki and Iwasa (2004) have shown that the standing strategy, based on the standing criterion, overcomes the issue of unjustified defection and maintains cooperation.

Ohtsuki and Iwasa (2004) examined all the possible ways to assign reputation to a player, once his action, his current reputation, and the opponent's reputation are given. They asked whether cooperative strategies can evolve based on indirect reciprocation even when there is some error included in executing the intended cooperation and in reporting the action of others to the public. Brandt and Sigmund (2004) developed a similar model independently with a small difference in how errors occur, and studied the evolution of cooperation by an individual based model. As a result of exhaustive search, Ohtsuki and Iwasa (2004) found that there are eight reputation dynamics which can maintain a non-trivial ESS with a high level of cooperation. They also showed that these ''leading eight'' are the only reputation dynamics allowing non-trivial ESS when the benefit exceeds the cost only slightly, and that the level of cooperation maintained at the ESS are close to the theoretical maximum (see below for more accurate statement).

In the present paper, we characterize the nature of successful social norms, the leading eight. After explaining the formalism of indirect reciprocity game adopted by Ohtsuki and Iwasa (2004) and Brandt and Sigmund (2004), we discuss the role of the elements that are common to all social norms of the leading eight to illustrate the mechanisms by which the leading eight achieves the high performance. Then we give a formal proof for two defining properties of the social norms of the leading eight.

2. Indirect reciprocity game and the leading eight

2.1. Indirect reciprocity game

Here we explain the formalism of indirect reciprocity game studied by Ohtsuki and Iwasa (2004). We consider a large population of individuals who engage in a game among them. In each generation they play the game for multiple rounds, each with a different player. In each round, players are paired randomly, and play one-shot prisoner's dilemma game. A player who cooperates pays fitness cost c, but the one who defects (i.e. refuses cooperation) pays no cost. A player receives benefit b, if the opponent cooperates. The benefit is larger than the cost ($b > c$). In choosing whether to cooperate or to defect, players can use reputation.

Fig. 1 illustrates events in the model. First, the focal player (player 1) meets player 2 who has been chosen randomly from the population. They both happen to have a good reputation, as indicated by G in the figure. Each player has his own ''behavioral strategy'', and decides whether to cooperate or to defect based on the reputation of the self and the opponent. In Fig. 1, player 1 defects and player 2 cooperates. Based on these actions, each player receives payoff. Here players 1 and 2 receive $+b$ and $-c$, respectively. Then they will change the partner and will never meet again. In Fig. 1, in the next round, player 1 engages in the game with player 3, who has a good reputation, indicated by G. Player 1 had a good reputation in the first round, but he now has a bad reputation in the second round, because he played defection against a good player in the previous round. The rule describing how reputation of a player changes over time is called ''reputation dynamics''. It assigns ''good'' (abbreviated as G) or ''bad'' (B) as his new reputation to the focal player in the next round, according to the action, the previous reputation of the actor, and the reputation of the opponent. Since no two

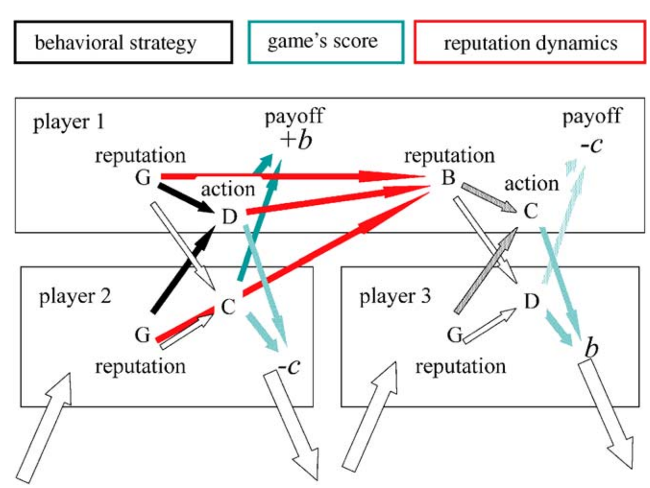

Fig. 1. Scheme of the indirect reciprocity game. See the text for details.

players meet again, a player's past action affects his benefit in the future only through modification of his own reputation. Ohtsuki and Iwasa (2004) asked whether there is any reputation dynamics that encourages the evolution of the cooperative behavioral strategy, rather than AllD, and, if there are some, what those successful reputation dynamics are.

Each player has a binary reputation, either *good* or *bad*, which is known publicly. Whether a player cooperates or defects may in general depend on the reputation of the opponent and that of himself. To express this, we consider behavioral strategy, denoted by p. $p(i, j)$ indicates action (either C or D, indicating cooperation and defection, respectively) when his reputation is i and the opponent's reputation is j (i, j are G or B, indicating good and bad, respectively). There are four different situations concerning the reputation of the self and opponent: both are good (GG), the actor is good whereas the opponent is bad (GB), the actor is bad whereas the opponent is good (BG), and both are bad (BB). For each case, a behavioral strategy prescribes whether to cooperate (C) or not (D). Hence there are $2^4 = 16$ possible behavioral strategies in total. Two simplest examples are AllD which defects (D) for all four cases, and AllC which always cooperates (C). But the other 14 behavioral strategies take different actions depending on the reputation of the self and the opponent.

In addition, the population has a "social norm", which assigns reputation (G or B) to each player. We call this a "reputation dynamics", and denote it by d (Ohtsuki and Iwasa, 2004). $d(i, j, X)$ is either good (G) or bad (B), and indicates the reputation of a player who takes an action X (either C or D), when his own reputation is i and the opponent's reputation is j (i, j = G, B). There are four possible situations (ij = GG, GB, BG, and BB) and for each situation there are two possible actions (X = C and D). A reputation dynamic specifies whether the given action in the given situation is *good* or *bad* for each of $4 \times 2 = 8$ cases. Hence there are $2^8 = 256$ different reputation dynamics in total. For example, the image-score criterion (Nowak and Sigmund, 1998a) is to regard players who cooperate as good and those who defect as bad, hence d (*,*, C) = G and d(*,*, D) = B, in which asterisk (*) is a wild card (i.e. these equations hold if * replaced by G or B). Another famous reputation dynamics is standing (Sugden, 1986), which differs from the image score only in one element: d(GB, D) = G (Ohtsuki and Iwasa, 2004).

Here we also consider a factor not included in Ohtsuki and Iwasa (2004). We assume that the number of rounds experienced by a player in a generation follows a geometric distribution. The expected probability that another round of game is played is given by ω $(0 < \omega < 1)$. The benefit to be obtained in the following round should be discounted by factor ω. When ω is large and close to one, as assumed in Ohtsuki and Iwasa (2004), each player engages in the game over a number of rounds per generation, but in each round with a different opponent randomly chosen from the population. If ω is small, the reciprocation in the future interaction is less likely, and having a good reputation gives a smaller indirect benefit than the case with a large ω.

2.2. The leading eight

Ohtsuki and Iwasa (2004) studied the reputation dynamics that has the following properties: first, in the population dominated by (d, p) players, the behavioral strategy p has a higher fitness than a rare mutant of any of the 15 behavioral strategies p', including AllD. In such a case, (d, p) is called an ESS pair. Second, in the population composed of (d, p) players only, the level of average fitness gain obtained per interaction is high. They studied the model when errors at low frequency occur both in executing the cooperation and in reporting the observed behavior to the public.

As a result of exhaustive examination of all 4096 possible pairs (d, p), only eight pairs are found to show very high performance in both properties (Ohtsuki and Iwasa, 2004). The population dominated by the type (d, p) refuses the invasion of 15 other behavioral strategies including AllD and AllC. The level of cooperation at the ESS is very high, and the average payoff per game is close to the maximum, $b-c$. These properties are shared only for these eight pairs, and we named them "leading eight" (Ohtsuki and Iwasa, 2004). Later we will state these two properties more precisely and prove them.

The leading eight pairs of reputation dynamics and behavioral strategy (d, p) have some elements in common but differ in others (Fig. 2). For example, all the eight reputation dynamics have

$$d(^*, \mathrm{G}, \mathrm{C}) = \mathrm{G} \text{ and } d(^*, \mathrm{G}, \mathrm{D}) = \mathrm{B} \tag{1}$$

where the asterisk (*) indicates a wild card. Eq. (1) implies that a player who cooperates with another player with a good reputation is assigned good in the next round, irrespective of the current reputation. Similarly, the second of Eq. (1) implies that a player who defects against a good opponent becomes to have a bad reputation in the subsequent round. In addition there is one more element common to all the leading eight, as follows:

$$d(\mathrm{G}, \mathrm{B}, \mathrm{D}) = \mathrm{G} \tag{2}$$

This implies that a good player who refuses cooperation to a bad opponent can remain good in the next round. Since each asterisk in Eq. (1) corresponds to two possibilities, Eqs. (1) and (2) combined specify 5 out of 8 elements of the reputation dynamics. However, the

d:	GG	GB	BG	BB
C	G	*	G	*
D	B	G	B	*

p:	C	D	C	**

Fig. 2. The leading eight (Ohtsuki and Iwasa, 2004). Top table: the reputation dynamics d, which specifies the reputation in the next round. Four columns correspond to the different pairs of reputations of the actor and the opponent. G and B indicate good and bad, respectively, and GG, GB, BG and BB are four combinations of reputations of the actor (the first symbol) and the recipient (the second symbol). Left of the table indicates the action of the actor (C and D implies cooperation and defection, respectively). Symbols in the table are the reputation common to all social norms in the leading eight. Asterisk (*) is a wild card, implying that both G and B are included in the element. Bottom table: the behavioral strategy p, which specifies the action of the player. Three elements filled as C and D indicates that those are common to all social norms in the leading eight. The one with double asterisks (**) is either C or D, which is determined according to the social norm that is adopted. It is C if and only if d(BB, C) = G and d(BB, D) = B; and it is D otherwise. See the main text for details.

following three elements are not specified: d(G, B, C), d(B, B, C), and d(B, B, D). Since each of these three can be either G or B, there are $2^3 = 8$ combinations, and all of these can achieve the ESS with a very high level of cooperation if a suitable behavioral strategy is chosen.

Concerning the behavioral strategy, the following three elements out of four are common to all the leading eight (see Fig. 2):

$$p(\mathrm{G,\ G}) = \mathrm{C},\ p(\mathrm{G,\ B}) = \mathrm{D},\ \text{and}\ \ p(\mathrm{B,\ G}) = \mathrm{C}. \tag{3}$$

But the fourth element p(B, B) can either be G or B. However, this cannot be freely chosen. Once we specify the reputation dynamics, then p(B, B) is specified to be G or B for (d, p) to be the ESS pair (see below).

These leading eight are discovered by an exhaustive search in Ohtsuki and Iwasa (2004). Subsequently, we showed that the condition for these eight pairs of d and p to constitute a non-trivial ESS pair under small errors and the level of cooperation at the ESS, indicating that they are desirable social norms. However, they are not the ones deduced from the argument of what the desirable social norms should be. We have not examined the meaning of each element of the leading eight reputation dynamics. In the next section, we attempt to explore the mechanisms by which these leading eight can succeed in maintaining the high level of cooperation.

3. Characterization of the leading eight

3.1. Intuition behind the leading eight

Here we first explain the characteristics common to all of these leading eight social norms and discuss the role of each element in shaping a cooperative society. Later we give a formal proof of two defining properties of the leading eight reputation dynamics.

(1) *Maintenance of cooperation*: First, in the population dominated by a successful pair (d, p), most players should maintain a high reputation (i.e. they are good), and most players cooperate with each other, except for a small fraction caused by unavoidable errors at a low frequency. We request the following two relations:

$$p(\mathrm{GG}) = \mathrm{C}\ \text{and}\ d(\mathrm{GG,\ C}) = \mathrm{G}. \tag{4}$$

The first formula implies that a player with a good reputation should cooperate if the opponent is good. The second equation indicates that the outcome of such an encounter keeps the actor in a good reputation.

(2) *Identification of defectors*: The second requirement is the efficiency of spotting rare defectors using reputation. Suppose that one AllD invades the population dominated by (d, p), in which most players have a good reputation. The AllD player should be immediately labeled as "bad", and in the next round he should be refused cooperation by another (d, p) player. We hence request the following two:

$$d(\mathrm{GG,\ D}) = \mathrm{B}\ \text{and}\ d(\mathrm{BG,\ D}) = \mathrm{B}. \tag{5}$$

These two equations indicate that a player who defects against a good opponent should be regarded as bad in the next round, irrespective of the current reputation. Using an asterisk (*) for a wild card, these can be combined as d(*G,D) = B.

(3) *Punishment and justification of punishment*: When a defector is detected and identified, then other good players should refuse cooperation to him in the next round. Hence a good player who meets with an opponent labeled as "bad" should refuse cooperation.

$$p(\mathrm{GB}) = \mathrm{D}. \tag{6}$$

In addition, a good player who refuses cooperation to a bad opponent should be able to maintain good reputation in the next round. Hence we have

$$d(\mathrm{GB,\ \ D}) = \mathrm{G}, \tag{7}$$

which implies the justification of punishment.

(4) *Apology and forgiveness*: There are errors at a low frequency, either in the execution process or in the way a player's revised reputation is reported. As a result, a player may be given a bad reputation and be refused cooperation once. However after this, the player must have a way to go back to the normal state with a good reputation and to start to cooperate again with other members. If the social norm does not allow such an

immediate forgiveness, one error may result in many events of punishment and refusal of cooperation, which reduces the average level of cooperation in the society. We hence consider the following condition:

$$p(\mathrm{BG}) = \mathrm{C}, \tag{8}$$

which implies that a player who recognized that he has been labeled as "bad", should apologize by playing cooperation. Then after this behavior, his reputation can go back to "good", as expressed by

$$d(\mathrm{BG},\ \mathrm{C}) = \mathrm{G}, \tag{9}$$

which can be called "forgiveness".

Requirements given by Eqs. (4), (5), (7) and (9) specify five elements of reputation dynamics common to all the leading eight. There are three elements, d(GB, C), d(BB, C), and d(BB, D), which remain unspecified. Since these three can be chosen either G or B, there are in total eight ways of filling them — hence the leading eight.

3.2. Rationality of behavioral strategy

For the pair (d, p) to be an ESS, the behavioral strategy p achieves a higher payoff on average than any other 15 behavioral strategies, given that the population is dominated by the players with (d, p). Cooperation (C) is accompanied by a cost, as the actor can save the cost by adopting an alternative option (defection, D). Playing cooperation, hence, must be accompanied by a benefit that is going to be given in future, when interacting with other players. We consider the difference in the expected payoff in the future between two players differing only in the reputation. This benefit of having a higher reputation must exceed the immediate cost of help in a population in which a high level of cooperation is maintained stably.

$$[\text{Benefit of having a higher reputation}] > [\text{Cost of help}]. \tag{10}$$

If this inequality is reversed, then the optimal choice of behavioral strategy is to play defection in any situations ($p(\mathrm{GG}) = p(\mathrm{GB}) = p(\mathrm{BG}) = p(\mathrm{BB}) = \mathrm{D}$), resulting in the population without cooperation. Hence in the population with the cooperation level higher than the one achieved by AllD, Eq. (10) must be satisfied. The cost of cooperation is c. But the benefit of having a higher reputation depends on many things such as the composition of the population, the strategy, or the expected number of times each player interacts per generation. This can be calculated in the population dominated by a particular (d, p), as shown in Appendix A.

Assume Eq. (10) holds, then we can conclude the following:

$$p(ij) = \mathrm{C},\ \text{only if ``}d(i,\ j,\ \mathrm{C}) = \mathrm{G} \text{ and } d(i,\ j,\ \mathrm{D}) = \mathrm{B}\text{''}, \tag{11a}$$

$$p(ij) = \mathrm{D},\ \text{otherwise.} \tag{11b}$$

Eq. (11a) indicates that cooperation is worth playing only when the benefit of having a better reputation exceeds the immediate cost of c, which occurs only when the cooperation brings in good reputation while defection bad reputation. If this is not the case, the optimal choice is to defect, as indicated by Eq. (11b). Once the reputation dynamics is known, the behavioral strategy is fully specified by Eq. (11).

As explained before, there are eight possible ways to fill three unspecified elements of reputation dynamics, but the behavioral strategy corresponding to each of these is specified. Only when cooperation produces a good reputation do players give help.

3.3. General proof of two properties of the leading eight

In deriving the leading eight, Ohtsuki and Iwasa (2004) considered errors in two forms: a player intended to cooperate fails to do so (execution error) and an observer reports the revised reputation of a player incorrectly (assessment error), both of which occur at a small probability of the same order of magnitude as small parameter ε. Definition of the leading eight is stated in the following form:

Property 1. In the population dominated by (d, p) players, the behavioral strategy p has a higher fitness than a rare mutant with any of the 15 behavioral strategies p' different from p. This result holds if an inequality $\omega b - c > 0$ is satisfied, in which the $\omega b - c$ must be greater than 0 by the difference greater than $O(\varepsilon)$, a small term of the same order of magnitude as the error rates.

Property 2. In the population composed of (d, p) players only, the average fitness gain obtained per interaction is close to the possible maximum, $b-c$. To be precise, the mean fitness gain per interaction is $b-c$ minus a small term of $O(\varepsilon)$. Note that this is equivalent to that the mean fitness gain per generation is $(b - c)/(1 - \omega)$ minus $O(\varepsilon)$.

Ohtsuki and Iwasa (2004) discovered the leading eight by exhaustive examination and stated that these properties can be shown for the leading eight strategies, but did not give the proof in general cases. In Appendix A, we give a proof that the reputation dynamics of the leading eight have both of these properties, and conversely, that the reputation dynamics with these properties must be one of the leading eight.

Now we have succeeded in clarifying why the social norms must be one of the leading eight, and why a social norm that is not included in the leading eight fails to maintain cooperation at a high level.

4. Discussion

Indirect reciprocity is one of the major driving forces of the evolution of cooperation in humans (Fehr and Fischbacher, 2003). Several experiments have shown that reciprocation through reputation formation plays an important role in generalized exchange (Bolton et al., 2005; Wedekind and Braithwaite, 2002; Wedekind and Milinski, 2000). In theory, evolutionary biologists as well as social scientists have studied about how indirect reciprocity emerges and works among selfish individuals (Brandt and Sigmund, 2004, 2005; Leimar and Hammerstein, 2001; Nowak and Sigmund, 1998a, b; Ohtsuki and Iwasa, 2004; Panchanathan and Boyd, 2003, 2004; Takagi, 1996).

In a previous paper, we examined exhaustively the ways of assigning reputation to a player in the next round based on the current reputation of the player, that of the opponent, and the action (Ohtsuki and Iwasa, 2004). We found that only eight reputation dynamics can achieve a very high level of cooperation at the equilibrium, but we did not discuss why only these eight reputation dynamics work and not others. In the present paper, we have attempted to identify the mechanisms by which the "leading eight" can maintain a high level of cooperation. We have first characterized these eight reputation dynamics. They have the following common aspects: (1) Cooperation is maintained in the population composed of that type only. (2) When cheaters invade, they should be detected and labeled immediately. (3) Those who are labeled as "bad" are refused cooperation by other members of the population, while those who refused help to them are regarded as good. (4) The player who got a bad reputation by occasional unavoidable errors should apologize, and he will be forgiven and can return to the good reputation again. Those four characteristics give us a clear and intuitive reason why the leading eight are so successful, which we could not find in the previous paper. Second, we have found how evolutionarily stable behavioral strategy is deduced from each reputation dynamics of the leading eight. We have shown that the successful strategy is the one that prescribes cooperation only when giving yields better reputation to the actor. Third, we gave a formal analysis and confirmed the ESS condition for the leading eight and the equilibrium cooperation level in the ESS under small errors. As far as we know, this is the first study that has analytically derived the benefit of having a higher reputation (see Appendix A for its detail). We believe that the method used here is directly applicable to a wide variety of game models on animal behavior, especially in calculating the value of information.

Most theoretical studies on indirect reciprocity focus on the evolutionary stability of a cooperative strategy. One of the characteristics of Ohtsuki and Iwasa (2004) is an explicit consideration of the mean level of cooperation achieved in the ESS. Using this criterion, we can discuss the effect of errors that propagate the population and thereby reduce the fraction of cooperative players at equilibrium. For example in Appendix A, we show that the social norms in the leading eight can achieve the equilibrium population with almost all the members having good reputation even if the small rate of errors are unavoidable in execution and observation processes; whilst other reputation dynamics have the equilibrium including far more members with bad reputation. This can be shown by mathematically examining the dynamical stability near the equilibrium (see Appendix A). If a social norm can maintain a higher level of cooperation than an alternative norm, we can deduce that the former social norm is more likely to be adopted in societies. Hence, we may expect that social norms we face in the world are likely to be those that are able to maintain a high level of cooperation against errors and recurrent invasion of alternative strategies.

In the following, we discuss the relation of our work to other studies.

4.1. Kandori's classical work

In a classical paper of game theory on social norms, Kandori (1992) investigated social norm which can sustain a cooperative and individually rational strategy, through local information processing such as reputation, rumor, or gossip. In Kandori's (1992) formalism, rules of updating reputation may in general depend on (1) the current reputation of the self, (2) the current reputation of the opponent, (3) the action of the self, and (4) the action of the opponent, and hence can be called, following Brandt and Sigmund (2005), "fourth-order assessment". In contrast, Ohtsuki and Iwasa (2004) and Brandt and Sigmund (2004) assume the new reputation to depend on (1), (2), and (3) only, which is "third-order assessment". Takahashi and Mashima (2003) focused on the case with (2) and (3) only, and hence their formalism can be called "second-order assessment", and a discriminator in Nowak and Sigmund (1998a, b) using (2) only is first-order assessment. We may translate the general results obtained by Kandori (1992) into the context of indirect reciprocity game. Kandori has proved that the rule of assigning reputation, "donation to *good* recipients and refusal against *bad* ones, are *good* behavior", is able to sustain cooperation as a sequential equilibrium, which is stricter than Nash equilibrium. Using the formalism discussed in the present paper, Kandori discussed the reputation dynamics given by $d(*\text{G-C}) = d(*\text{B-D}) = \text{G}$ and $d(*\text{G-D}) = d(*\text{B-C}) = \text{B}$, combined with the behavioral strategy: $p(*\text{G}) = \text{C}$ and $p(*\text{B}) = \text{D}$, This rule studied by Kandori (1992) is one of the leading eight in Ohstuki and Iwasa (2004).

Also notable is that Kandori (1992) proved that, under a weak condition, this type of reputation dynamics makes any kind of outcomes an equilibrium in any kind of games, though generally it takes more than one round to impose effective punishment and so there should be more than two reputation. It is interesting that with the suitable reputation dynamics having the concept of punishment and justified defection, everything becomes an equilibrium. This result is similar to the conclusion of Boyd and Richerson (1992), who showed that with the option of punishment everything is possible in a repeated interaction.

4.2. Takahashi and Mashima's reputation dynamics

Based on the second-order assessment, Takahashi and Mashima (2003) studied reputation dynamics which give *good* reputation only when a donor cooperates with a *good* recipient. This reputation dynamics assigns *bad* reputation to any players who happened to meet a *bad* player. With this dynamics, *bad* reputation bearers keep increasing over time. By computer simulation Takahashi and Mashima found that it is robust against invasion by other strategies and that the average cooperation rate is fairly high, about 80%. However, this high giving rate is originated from a short generation time they assumed in their simulation. Takahashi and Mashima (2003) assumed that all the players have good reputation at the beginning of every generation. If the generation time is short, each generation ends before *bad* reputation spreads in the population, as they observed. However if there were more interactions per generation, Takahashi and Mashima's reputation dynamics would not produce a society with a high level of cooperation. That is why this reputation dynamics is not included in the leading eight.

4.3. Private opinion versus social reputation

Brandt and Sigmund (2004) considered a similar scheme of the indirectly reciprocity game as Ohtsuki and Iwasa (2004) (Fig. 1). In the individual-based direct computer simulation, Brandt and Sigmund often found a mixture of multiple strategies instead of pure ESS, as a result of the evolution. One difference between these two studies was a finite number of rounds per generation in Brandt and Sigmund (2004), whilst a very large number of rounds per generation was assumed in Ohtsuki and Iwasa (2004). In the present paper, we relaxed this assumption of Ohtsuki and Iwasa (2004), and showed that qualitatively the same result is obtained for the case the number of rounds each player experiences in a generation is finite. The condition for the leading eight to be ESS is $\omega b > c$, in which the benefit of cooperation by indirect reciprocity is discounted by factor ω, the probability of having a next round.

Another difference, probably more important, was that in Brandt and Sigmund (2004) the evaluation on a player should be done purely based on the evaluator's personal observation of the past action of the player. Therefore players never exchange the information or the views on other members of the community, and hence the reputation (they called score) on the same player may differ between players in the population. In contrast, Ohtsuki and Iwasa (2004) assumed that the information on a player's reputation is exchanged among other members of the population and is shared among the members. Reputation in this model is truly social information, though it may be incorrect due to the error in the observation and reporting processes.

One of the criticism to the latter "social reputation assumption" is that, due to the manipulation of players in favor of himself, the social reputation is not reliable and hence unlikely to be the basis of the cooperation in the society (Brandt and Sigmund, 2004, 2005). Study on the effect of lies and deliberate spread of incorrect or distorted information is a very important and interesting theme (see, for example, Nakamaru and Kawata, 2004). In fact, it is probable that through language the social reputation becomes more and more trustless thus reputation dynamics depending on second or higher order assessment rules become less effective. However, it is also possible that language facilitates the formation of social reputation that is free from lies, because we usually carry out careful evaluation of how trustworthy each piece of information is. If we think of how much information on the action and personality of a particular member of the community are based on the direct observation by ourselves, in contrast to the information obtained from other members of the society, it is clear that the acquisition of the information of others via constantly exchanging views, gossip or rumor plays a very important role in evaluating and judging the character and past behavior of other players in the same society. Thus language may have two mutually opposite effects on the social reputation.

In our analysis we assumed that the population shares the same reputation dynamics, which resulted in the population having social reputation as a consensus. If, however, reputation dynamics, namely moral criteria, are different among players, private opinion might play more important role than social reputation. It is an open question to be examined further in future.

In human society it is often the case that good reputation, once lost, is hard to recover (Herbig et al., 1994; Nicholas and Fournier, 1999). To consider this slow recovery of the reputation, we may introduce the chance of recovery from bad reputation to good reputation, denoted by r, is less than unity. When a player had a bad reputation and took the action that is worth good reputation in reputation dynamics, he can actually recover good reputation only with probability r.

With probability $1-r$ his reputation remains *bad*, though he tried to gain *good* reputation. When it is the case, it takes *bad* reputation bearers on average $1+(1-r)+(1-r)^2+\ldots=1/r$ interactions before recovering *good* reputation even if he always tries. Therefore, under the leading eight the benefit of good reputation is enhanced by a factor of $1/r$, making players more likely to keep their *good* reputation.

Recently Brandt and Sigmund (2005) showed that discriminators and AllCs can coexist stably if age structure is considered.

Axelrod (1984) summarized his study of direct reciprocity by that a good strategy in *direct* reciprocity should be nice, retaliatory, and forgiving. Based on our study of indirect reciprocity, we have found that the keys to the success in *indirect* reciprocity are to be nice (shown by maintenance of cooperation pivots), retaliatory (detection of defectors, punishment, and justification of punishment pivots), apologetic and forgiving.

Acknowledgment

We thank Hannelore Brandt, Michihiro Kandori, Martin Nowak, Karl Sigmund, and Nobuyuki Takahashi for their fruitful comments. H.O. acknowledges the support from JSPS.

Appendix A. Proof of two properties of the leading eight

In the following we give a sketch of the proof that the leading eight reputation dynamics have Properties 1 and 2 stated in the main text, and that they are the only reputation dynamics that satisfy both of them. We start our argument with Property 1.

(1) *Behavioral strategy is the one maximizing the fitness of the player*

This is the same as rational choice condition in the text. We require that $p(i, j)$ should take either C or D, the one that gives a higher payoff to the player. Hence we must compare the benefit of having a higher reputation that makes return a greater reward in the future with the cost. We require that the benefit of having a higher reputation v has to be greater than the immediate cost c: $v>c$. This is necessary for any non-trivial ESSs to exist (see text). Under the assumption of $v>c$, we request Eq. (11) in the main text. This condition specifies the behavioral strategy $p(i, j)$ once we know the reputation dynamics $d(i, j, X)$.

Then the remaining problem is to show that this benefit of having a good reputation exceeds c. To do so, we first start with a few premises that are to be satisfied by successful social rules, including those requested by Property 2.

(2) *All the players should become to cooperate in the monomorphic population in the absence of errors*

In the population composed only of (d, p), starting from any initial condition, all players cooperate and all have a high reputation, except for unavoidable errors at a low frequency. From this, we immediately conclude $p(\mathrm{GG}) = \mathrm{C}$, implying a good player meeting another good player must cooperate. From this, combined with the rational-choice condition (Eq. (11)), we have

$$d(\mathrm{GG},\ \mathrm{C}) = \mathrm{G} \text{ and } d(\mathrm{GG},\ \mathrm{D}) = B. \tag{A.1}$$

Now we consider time change in the fraction of players with good reputation. Let h be a frequency of players with good reputation in the population. In the absence of errors, h should converge to 1 from any initial condition $0<h_0\leqslant 1$. Let $d_{ij} = d(i,j,p(ij))$ be the reputation that a p-strategist is assigned after an interaction in situation $ij = \mathrm{GG,GB,BG,BB}$. Then we have

$$\begin{aligned}\frac{\mathrm{d}h}{\mathrm{d}t} = \{&h^2\cdot\delta(d_{GG}) + h(1-h)(\delta(d_{GB})+\delta(d_{BG}))\\ &+(1-h)^2\cdot\delta(d_{BB})\} - h,\end{aligned} \tag{A.2}$$

where $\delta(G) = 1$ and $\delta(B) = 0$. Define $f(h)$ as the r.h.s. of Eq. (A.2). The stability of the equilibrium $h = 1$ leads to $f(1) = 0$ and $f'(1)\leqslant 0$, which give $\delta(d_{GG}) = 1$ and $\delta(d_{GB})+\delta(d_{BG})\geqslant 1$. When $\delta(d_{GB})+\delta(d_{BG}) = 2$, the global stability holds irrespective of the value of $\delta(d_{BB})$. In contrast, when $\delta(d_{GB})+\delta(d_{BG}) = 1$, we have $f(h) = (1-h)^2\cdot\delta(d_{BB})$, and hence $\delta(d_{BB}) = 1$ is necessary for global stability of $h = 1$. We come to the following conclusions:

(i) $d(\mathrm{GG},\ p(\mathrm{GG})) = \mathrm{G}$, which was already given by Eq. (A.1).

(ii) At least two of the three equations: $d(\mathrm{GB},\ p(\mathrm{GB})) = \mathrm{G}$, $d(\mathrm{BG},\ p(\mathrm{BG})) = \mathrm{G}$, and $d(\mathrm{BB},\ p(\mathrm{BB})) = \mathrm{G}$ must hold.

(3) *Errors of small magnitude reduce the cooperation level only by a small amount*

Ohtsuki and Iwasa (2004) considered errors in two forms: a player intended to cooperate fails to do so (execution error) and an observer reports the revised reputation of a player incorrectly (assessment error), both of which occur at a low rate. These can be expressed as an additional term of small magnitude (say of ε order). The system's behavior when perturbed by a small additional error differs between the following two cases:

(i) When both $d(\mathrm{GB},\ p(\mathrm{GB})) = \mathrm{G}$ and $d(\mathrm{BG},\ p(\mathrm{BG})) = \mathrm{G}$ hold:

The eigenvalue at $h = 1$ is -1. Since time constant is 1, after more than several rounds, h converges to the equilibrium, h^*. Hence small errors of order $O(\varepsilon)$ cause the shift in the equilibrium by $O(\varepsilon)$: the fraction of good people decreases by $1-h^* = O(\varepsilon)$.

(ii) When $d(\mathrm{BB}, p(\mathrm{BB})) = \mathrm{G}$ holds but either $d(\mathrm{GB}, p(\mathrm{GB})) = \mathrm{G}$ or $d(\mathrm{BG}, p(\mathrm{BG})) = \mathrm{G}$ does not hold:

The eigenvalue of the linearized dynamics at $h = 1$ is zero, and $h = 1$ is stable only by the second-order terms: $\mathrm{d}(1-h)/\mathrm{d}t = -(1-h)^2$, hence convergence to the equilibrium is rather slow. Additional error terms of magnitude of ε cause the shift by $1 - h^* = O(\sqrt{\varepsilon})$. The decline of the fraction of "good" players is of order of magnitude of $\sqrt{\varepsilon}$, which is much greater than ε if ε is small. This can be demonstrated in a detailed calculation for each special case of different ways of introducing errors (calculation not shown here).

By comparison between (i) and (ii), the rules satisfying (i) give a higher fraction of "good" people and hence result in a higher cooperation level. Therefore, we conclude that a desirable strategy must satisfy the following:

$$d(\mathrm{GB},\ p(\mathrm{GB})) = \mathrm{G} \text{ and } d(\mathrm{BG},\ p(\mathrm{BG})) = \mathrm{G}. \tag{A.3}$$

(4) *Detecting cheaters and punishment*

Next, we show $p(\mathrm{GB}) = \mathrm{D}$. Assume to the contrary that $p(\mathrm{GB}) = \mathrm{C}$ holds. Then every player, who is supposed to have a good reputation at $h = 1$ equilibrium, gives help even to opponents having a bad reputation. This means that it does not matter at all whether a player has a good reputation or a bad one at this equilibrium, which contradicts $v > c$ condition. Therefore, $p(\mathrm{GB}) = \mathrm{D}$ must hold, which is the condition for punishment pivot in the text.

From $p(\mathrm{GB}) = \mathrm{D}$ together with $d(\mathrm{GB}, p(\mathrm{GB})) = \mathrm{G}$ which was derived in Eq. (A.3), we conclude

$$d(\mathrm{GB},\ \mathrm{D}) = \mathrm{G} \tag{A.4}$$

which implies that the punishment to a bad player (refusal of cooperation against bad players) is justified and the actor in such a situation can maintain a high reputation in the next round.

(5) *Benefit of good reputation*:

Next we examine $d(\mathrm{BG,C})$ and $d(\mathrm{BG, D})$. From $d(\mathrm{BG}, p(\mathrm{BG})) = \mathrm{G}$ obtained in Eq. (A.3), we have the following two possibilities: Case 1, $d(\mathrm{BG, C}) = \mathrm{G}$ and $p(\mathrm{BG}) = \mathrm{C}$, and Case 2, $d(\mathrm{BG, D}) = \mathrm{G}$ and $p(\mathrm{BG}) = \mathrm{D}$. We examine the possibility of $v > c$ in these cases separately:

Case 1: When $d(\mathrm{BG, C}) = \mathrm{G}$ and $p(\mathrm{BG}) = \mathrm{C}$.

Consider two players A_{good} and A_{bad}, whose behavioral strategies are p, and suppose that the reputation of A_{good} is *good* whereas that of A_{bad} is *bad*. In the next interaction (which exists with probability ω), A_{good} cooperates ($p(\mathrm{GG}) = \mathrm{C}$) with his opponent and his opponent also cooperates ($p(\mathrm{GG}) = \mathrm{C}$) with him, resulting in his *good* reputation in the next round ($d(\mathrm{GG,C}) = \mathrm{G}$). The expected payoff is $\omega(b - c + \omega v)$. In contrast, A_{bad} cooperates with his good opponent ($p(\mathrm{BG}) = \mathrm{C}$) but his opponent defects ($p(\mathrm{GB}) = \mathrm{D}$) against him, resulting in his *good* reputation in the next round ($d(\mathrm{BG,C}) = \mathrm{G}$). A_{bad}'s gain is calculated as $\omega \cdot (-c + \omega v)$. The difference in gains between these two gives the benefit of having a higher reputation v, and it is $v = \omega(b - c + \omega v) - \omega(-c + \omega v) = \omega b$. From this, the relation $v > c$ holds when $\omega b > c$.

Case 2: When $d(\mathrm{BG, D}) = \mathrm{G}$ and $p(\mathrm{BG}) = \mathrm{D}$:

Consider two players A_{good} and A_{bad}, whose behavioral strategies are p, and suppose that the reputation of A_{good} is *good* whereas that of A_{bad} is *bad*. In the next interaction (which exists with probability ω), A_{good} cooperates ($p(\mathrm{GG}) = \mathrm{C}$) with his good opponent and his opponent also cooperates ($p(\mathrm{GG}) = \mathrm{C}$) with him, resulting in his new *good* reputation ($d(\mathrm{GG,C}) = \mathrm{G}$). Through this interaction his gain is $\omega(b - c + \omega v)$. In contrast, A_{bad} defects ($p(\mathrm{BG}) = \mathrm{D}$) against his good opponent and his opponent also defects ($p(\mathrm{GB}) = \mathrm{D}$) against him, resulting in his new *good* reputation ($d(\mathrm{BG, D}) = \mathrm{G}$). Therefore his total gain is calculated as $\omega(0 + \omega v)$. The difference in gains between these two cases is equal to the benefit of having a *good* reputation, which is $v = \omega(b - c + \omega v) - \omega(0 + \omega v) = \omega(b - c)$. Hence $v > c$ condition holds if and only if $\omega b > (1 + \omega)c$.

The condition for $v > c$ is satisfied when $\omega b > c$ for the rules in Case 1, but only when $\omega b > (1 + \omega)c$ for the rules in Case 2. The latter is more difficult to satisfy, and hence we request the former condition to be satisfied in a desirable reputation dynamics. Thus we conclude $d(\mathrm{BG, C}) = \mathrm{G}$ and $p(\mathrm{BG}) = \mathrm{C}$. From the rationality of behavioral strategy (Eq. (11)), we also have $d(\mathrm{BG, D}) = \mathrm{B}$.

If the number of interactions within a generation is very large (ω is close to 1), the condition for $v > c$ can be satisfied when $b - c$ is only slightly larger than zero, as shown in Ohtsuki and Iwasa (2004).

(6) *The leading eight*:

The argument above specifies the following five pivots in the reputation dynamics:

$$d(\mathrm{GG},\ \mathrm{C}) = d(\mathrm{BG},\ \mathrm{C}) = d(\mathrm{GB},\ \mathrm{D}) = \mathrm{G}$$

and

$$d(\mathrm{GG},\ \mathrm{D}) = d(\mathrm{BG},\ \mathrm{D}) = \mathrm{B},$$

and the following three pivots of the behavioral strategy:

$$p(\mathrm{GG}) = p(\mathrm{BG}) = \mathrm{C} \text{ and } p(\mathrm{GB}) = \mathrm{D}.$$

Now, we have three more pivots in reputation dynamics: $d(\mathrm{GB,C})$, $d(\mathrm{BB,C})$, and $d(\mathrm{BB,D})$. These three can be chosen freely, and hence there are $2^3 = 8$ possibilities, which are the leading eight social norms (Ohtsuki and Iwasa, 2004). One pivot of behavioral strategy $p(\mathrm{BB})$ remains unspecified, but this is determined by Eq. (11) once both $d(\mathrm{BB,C})$ and $d(\mathrm{BB,D})$ are given.

From these arguments, we can tell that the leading eight satisfies the two properties we requested for desirable reputation dynamics, and that none of the reputation dynamics other than the leading eight can satisfy the conditions we request.

References

Axelrod, R., 1984. The evolution of cooperation. Basic Books, New York, USA.

Axelrod, R., Hamilton, W.D., 1981. The evolution of cooperation. Science 211, 1390–1396.

Bolton, G.E., Katok, E., Ockenfels, A., 2005. Cooperation among strangers with limited information about reputation. J. Public Econ. 89, 1457–1468.

Boyd, R., Richerson, P.J., 1992. Punishment allows the evolution of cooperation (or anything else) in sizable groups. Ethol. Sociobiol. 13, 171–195.

Brandt, B., Sigmund, K., 2004. The logic of reprobation: assessment and action rules for indirect reciprocation. J. Theor. Biol. 231, 475–486.

Brandt, B., Sigmund, K., 2005. Indirect reciprocity, image scoring, and moral hazard. Proc. Natl. Acad. Sci. USA 102, 2666–2670.

Fehr, E., Fischbacher, U., 2003. The nature of altruism. Nature 425, 785–791.

Herbig, P., Milewicz, J., Golden, J., 1994. A model of reputation building and destruction. J. Business Res. 31, 23–31.

Kandori, M., 1992. Social norms and community enforcement. Rev. Econ. Stud. 59, 63–80.

Leimar, O., Hammerstein, P., 2001. Evolution of cooperation through indirect reciprocity. Proc. R. Soc. London B 268, 745–753.

Nakamaru, M., Kawata, M., 2004. Evolution of rumors that discriminate lying defectors. Evol. Ecol. Res. 6, 261–283.

Nicholas, M.W., Fournier, G.M., 1999. Recovering from a bad reputation: changing beliefs about the quality of US autos. Int. J. Ind. Org. 17, 299–318.

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.

Ohtsuki, H., 2004. Reactive strategies in indirect reciprocity. J. Theor. Biol. 227, 299–314.

Ohtsuki, H., Iwasa, Y., 2004. How should we define goodness? reputation dynamics in indirect reciprocity. J. Theor. Biol. 231, 107–120.

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.

Panchanathan, K., Boyd, R., 2004. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. Nature 432, 499–502.

Sugden, R., 1986. The Economics of Rights, Co-operation and Welfare. Blackwell, Oxford, UK.

Takagi, E., 1996. The generalized exchange perspective of the evolution of altruism. In: Liebrand, W.B.G., Messick, D.M. (Eds.), Frontiers in Social Dilemmas Research. Springer, Berlin, pp. 311–336.

Takahashi, N., Mashima, R., 2003. The Emergence of Indirect Reciprocity: Is the Standing Strategy the Answer? Center for the study of cultural and ecological foundations of the mind, Hokkaido University, Japan, Working Paper Series No. 29.

Trivers, R., 1971. The evolution of reciprocal altruism. Q. Rev. Biol. 46, 35–57.

Wedekind, C., Braithwaite, V.A., 2002. The long-term benefits of human generosity in indirect reciprocity. Curr. Biol. 12, 1012–1015.

Wedekind, C., Milinski, M., 2000. Cooperation through image scoring in humans. Science 288, 850–852.