Emergence of cooperation in public goods games

Shun Kurokawa* and Yasuo Ihara

Department of Biological Sciences, The University of Tokyo, Bunkyoku, Tokyo 113-0033, Japan

Evolution of cooperation has been a major issue in evolutionary biology. Cooperation is observed not only in dyadic interactions, but also in social interactions involving more than two individuals. It has been argued that direct reciprocity cannot explain the emergence of cooperation in large groups because the basin of attraction for the ‘cooperative’ equilibrium state shrinks rapidly as the group size increases. However, this argument is based on the analysis of models that consider the deterministic process. More recently, stochastic models of two-player games have been developed and the conditions for natural selection to favour the emergence of cooperation in finite populations have been specified. These conditions have been given as a mathematically simple expression, which is called the one-third law. In this paper, we investigate a stochastic model of $n$-player games and show that natural selection can favour a reciprocator replacing a population of defectors in the $n$-player repeated Prisoner's Dilemma game. We also derive a generalized version of the one-third law (the $(2/[n(n+1)])^{1/(n-1)}$ law). Additionally, contrary to previous studies, the model suggests that the evolution of cooperation in public goods game can be facilitated by larger group size under certain conditions.

Keywords: one-third law; finite populations; cooperation; evolutionary stability; fixation probability; public goods game

1. INTRODUCTION

Understanding the evolution of cooperation in terms of natural selection has been a major challenge in evolutionary biology (Hamilton 1964; Trivers 1971; Axelrod & Hamilton 1981; Axelrod 1984; Nowak 2006). Animals sometimes behave cooperatively even towards unrelated individuals and direct reciprocity has been a key concept to explain this phenomenon (Trivers 1971; Axelrod & Hamilton 1981; Axelrod 1984).

In humans, cooperation is observed not only in dyadic interactions, but also in social interactions involving more than two individuals (Joshi 1987; Boyd & Richerson 1988). Humans often cooperate in public goods games and in situations ranging from family issues to global warming (Kollock 1998; Milinski et al. 2006). It has been argued that direct reciprocity cannot explain the emergence of cooperation in large groups because the basin of attraction for the ‘cooperative’ equilibrium state shrinks rapidly as the group size increases (Boyd & Richerson 1988). However, this argument is based on the analysis of conventional deterministic models, which do not consider the stochastic process in finite populations. Natural populations are always finite and important biological phenomena such as random genetic drift can be studied only in stochastic settings. In the present context, how likely it is that a reciprocator invades and eventually replaces a population of defectors cannot be examined without taking the stochastic process into account.

Nowak et al. (2004) developed a stochastic model of two-player games and specified the conditions required for natural selection to favour the emergence of cooperation in finite populations. Assuming a discrete-time population of fixed size $N$ with one individual replaced at a time, in accordance with the Moran (1958) model, Nowak et al. (2004) showed that the fixation probability of a single ‘tit for tat’ reciprocator (TFT) in a population of defectors (ALLD) can be greater than $1/N$, the fixation probability expected in the absence of natural selection. Furthermore, when both TFT and ALLD are evolutionarily stable in the traditional sense, they demonstrated for sufficiently large population and for sufficiently weak selection that the fixation probability of TFT is larger than $1/N$ if $p^* < 1/3$, where $p^*$ is the frequency of TFT at the unstable equilibrium in the conventional deterministic model (i.e. the ‘one-third’ law). The one-third law has been shown to hold also for a population with discrete non-overlapping generations that follows the Wright–Fisher model (Fisher 1930; Wright 1931) under the same assumptions on population size and selection intensity (Lessard 2005; Imhof & Nowak 2006).

In this paper, we consider a stochastic model of $n$-player games by extending the model of Nowak et al. (2004). Using this model, we obtain the condition under which a single reciprocator replacing a population of defectors is favoured by natural selection in the $n$-player repeated Prisoner's Dilemma game. We also examine the relationship between the fixation probabilities and the basins of attraction of equilibrium states; in other words, we derive a generalized version of the one-third law. In addition, we investigate the effects of the group size on the emergence and maintenance of cooperation in the $n$-player repeated Prisoner's Dilemma game in finite populations.

2. THE GENERAL $n$-PLAYER GAME

Let us consider the situation where groups of $n$ individuals are formed by randomly choosing individuals from a population of size $N$ and a game is played within each of these groups. There are two strategies, A and B, and pay-off...
Table 1. The pay-off matrix of the general n-player game.

<table>
<thead>
<tr>
<th>strategy of the focal individual</th>
<th>number of A individuals among the n−1 opponents</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>a₁ a₂ a₃ ... aₙ₋₁ aₙ</td>
</tr>
<tr>
<td>B</td>
<td>b₁ b₂ b₃ ... bₙ₋₁ bₙ</td>
</tr>
</tbody>
</table>

The expected pay-offs of A and B individuals are given by

\[ F_i = \sum_{k=1}^{n} \left( \frac{i-1}{n-k} \frac{N-i}{k-1} \right) a_k, \quad (2.1) \]

\[ G_i = \sum_{k=1}^{n} \left( \frac{i}{n-k} \frac{N-i-1}{k-1} \right) b_k, \quad (2.2) \]

respectively, where \( i \) represents the number of A individuals in the population \((0 \leq i \leq N)\). Note that \( \binom{i}{j} \) represents a binomial coefficient when \( j \geq t \) and is defined as zero when \( j < t \). The fitness of A and B individuals when there are \( i \) A individuals in the population are, respectively, given by \( f_i = 1 - w + wF_i \) and \( g_i = 1 - w + wG_i \), where \( w \) specifies the contribution of the game to fitness. Most of the results shown in this paper are under the assumption of weak selection \((0 < w \ll 1)\). Population dynamics are formulated as a Moran process with frequency-dependent fitness. At each time step, an individual is chosen for reproduction with the probability proportional to its fitness. One identical offspring is produced to replace another individual chosen for death with the uniform probability \( 1/N \) (Moran 1958). The process has two absorbing states, \( i = 0 \) and \( i = N \). It has been shown (Nowak et al. 2004) that the fixation probability of A, \( \rho_A \), with which a population at state \( i = 1 \) eventually reaches state \( i = N \) is given by

\[ \rho_A = 1 \left( 1 + \frac{1}{N} \sum_{k=1}^{n} \prod_{i=1}^{k-1} \frac{b_i}{a_i} \right). \quad (2.3) \]

The fixation probability of B, \( \rho_B \), that a population at state \( i = n - 1 \) reaches \( i = 0 \) can be similarly obtained.

We find that for the general n-player game with weak selection, the fixation probability of A is given approximately by

\[ \rho_A = \frac{1}{N} \frac{1}{1 - (\alpha N - \beta)w/n(n + 1)}. \quad (2.4) \]

where

\[ \alpha = \sum_{k=1}^{n} (a_k - b_k), \]

\[ \beta = -n^2 b_n + \sum_{k=1}^{n-1} k b_k + \sum_{k=1}^{n} (n + 1 - k) a_k. \]

Hence, \( \rho_A > 1/N \) holds and thus selection favours A replacing B if \( \alpha N > \beta \). In the limit of large population size, \( \rho_A < 1/N \) if

\[ \sum_{k=1}^{n} k a_k < \sum_{k=1}^{n} k b_k. \quad (2.5) \]

The fixation probability of B can be calculated similarly.

It is of interest to ask whether A is more likely to replace B than vice versa. The ratio of the fixation probabilities can be calculated as follows:

\[ \frac{\rho_A}{\rho_B} = 1 + \frac{w}{n} [\gamma N - n(a_1 - b_n)]. \quad (2.6) \]

where

\[ \gamma = \sum_{k=1}^{n} (a_k - b_k). \]

Hence, \( \rho_A > \rho_B \) is equivalent to \( \gamma N > n(a_1 - b_n) \). In the limit of large \( N \), \( \rho_A > \rho_B \) if

\[ \sum_{k=1}^{n} a_k > \sum_{k=1}^{n} b_k. \quad (2.7) \]

Note that both \( \rho_A \) and \( \rho_B \) can be less than \( 1/N \), in which case selection opposes replacement in either direction. It is also possible to find conditions where both \( \rho_A \) and \( \rho_B \) are greater than \( 1/N \) and thus selection favours replacement in either direction. As a special case, conditions (2.5) and (2.7) include the corresponding conditions that have been obtained (Nowak et al. 2004) for the case of \( n = 2 \).

Following Nowak et al. (2004), we call strategy B ESS\(_N\) (evolutionarily stable strategy) if the following two conditions are met: (i) selection opposes A invading B, which means that a single mutant A in a population of B has a lower fitness, and (ii) selection opposes A replacing B, or \( \rho_A < 1/N \). For the general n-player game, the first and second conditions are equivalent to

\[ (N - 1)a_n < (N - n)b_n + (n - 1)b_{n-1} \quad (2.8) \]

and

\[ \alpha N < \beta, \quad (2.9) \]

respectively.

Let us compare the ESS\(_N\) conditions (2.8) and (2.9) with the traditional ESS condition in the conventional deterministic model (e.g. Maynard Smith 1982; Hofbauer & Sigmund 1998). In our notation, the traditional ESS condition for B is \( a_n < b_n \). Consider two special cases. First, for large \( N \), (2.8) and (2.9) reduce to \( a_n < b_n \) and (2.5), respectively. In this case, therefore, ESS\(_N\) requires the traditional condition as well as an additional condition (2.5). That is, the traditional ESS condition is necessary but not sufficient for ESS\(_N\). Second, for the smallest possible population size, \( N = n \), the ESS\(_N\) conditions for B are \( a_n < b_{n-1} \) and

\[ \sum_{k=1}^{n-1} k a_{k+1} < \sum_{k=1}^{n-1} k b_k. \quad (2.10) \]

Hence, the traditional condition is neither necessary nor sufficient for B to be ESS\(_N\). Note that when \( n = 2 \), conditions \( a_n < b_{n-1} \) and (2.10) are equivalent (Nowak et al. 2004), but not so when \( n > 2 \).
3. THE n-PLAYER REPEATED PRISONER’S DILEMMA GAME

Among n-player games of particular interest used to deal with the evolution of cooperation in groups of individuals. Here, we apply the general framework developed above to the n-player repeated Prisoner’s Dilemma game as a version of public goods games. A game consists of m round(s) \((m \geq 1)\), in each of which individuals either cooperate or defect. When some individuals cooperate, all the individuals in the group gain a benefit from it while only the cooperating individuals have to pay a cost. Following earlier works (Joshi 1987; Boyd & Richerson 1988), we consider two strategies: ALLD and TFT. ALLD always defects no matter what other individuals do. TFT cooperates in the first round and then cooperates in each subsequent round if a or most of the \(n-1\) opponents cooperated in the previous round. It can be shown that no strategies other than ALLD can be ESS when \(N=n\) (see the electronic supplementary material). Thus, we assume \(N>n\) in the following.

Let us assume that the pay-offs to cooperating (C) and defecting (D) individuals for a particular round in which there are \(k\) cooperating individuals (\(0 \leq k \leq n\)) are \(V(C|k) = bk/n - c\) and \(V(D|k) = bk/n\), respectively, where \(b\) and \(c\) represent the benefit and cost of cooperation. From the definition of the n-player Prisoner’s Dilemma game, it must be that \(b > c\) or \(b/n > 0\). Let strategies A and B in table 1 represent TFT and ALLD, respectively. According, let \(a_i\) and \(b_i\) represent the total pay-offs of a game (after the \(m\)th round) to a TFT and an ALLD, respectively, when there are \(n-j\) TFTs among the opponents. The total pay-offs are given as follows: when \(j \leq n-a-1\), \(a_i = m[b(n+1-j)/n - c]\) and \(b_i = mb(n-j)/n\); when \(j = n-a\), \(a_i = m[a(n+1)/n - c]\) and \(b_i = ba/n\); and when \(j > n-a+1\), \(a_i = b(n+1-j)/n - c\) and \(b_i = b(n-j)/n\).

In the conventional model, TFT is ESS if \(a_t > b_t\), i.e. \(a = n-1\) and \(m > b(n-1)/(b-c)n\). Hence, TFT\(_{n-1}\) can be traditional ESS when the number of rounds is sufficiently large (Joshi 1987; Boyd & Richerson 1988). However, since ALLD is always traditional ESS (\(a_t < b_t\)), the emergence of cooperation cannot be explained by the conventional approach.

We examine the fixation probabilities of TFT and ALLD in the stochastic model. Since it turns out that selection opposes ALLD invading TFT only if \(a = n-1\), in this paper, we focus on the case of \(a = n-1\). Assuming weak selection, the following results are obtained. First, selection opposes ALLD invading TFT\(_{n-1}\) if

\[
m > \frac{(n-1)[Nb(c-n(b-c-1))]}{n(b-c-1)(N-n)}, \quad (3.1)
\]

and opposes ALLD replacing TFT\(_{n-1}\) if

\[
m > \frac{(n-1)[Nb(c+n)-n(b-c-1)]}{2n(b-c-1)(N-n)}. \quad (3.2)
\]

Condition (3.2) is sufficient for (3.1). Hence, TFT\(_{n-1}\) is ESS\(_N\) when (3.2) is satisfied. The right-hand side of (3.2) gives the critical number of rounds above which TFT\(_{n-1}\) is ESS\(_N\). In the limit of large \(N\), (3.2) becomes

\[m > (n-1)(b/c + n)/[2n(b-c-1)].\]

Second, selection always opposes TFT\(_{n-1}\) invading ALLD while it opposes TFT\(_{n-1}\) replacing ALLD if

\[
m < \frac{(n-1)[(n+1)N-(b/c-1)(N-n)]}{2n(b-c-1)(N-n)} \quad (3.3)
\]

Hence, ALLD is ESS\(_N\) when (3.3) is satisfied. Note that (3.2) and (3.3) hold simultaneously. Condition (3.3) deserves emphasis: when this condition is violated, \(p_{TFT_{n-1}} > 1/N\) is satisfied, which means that natural selection favours a single mutant immigrant to eventually replace a population of defectors. In the limit of large \(N\), (3.3) becomes

\[m < (n-1)(n+2-b/c)/[2(b-c-1)].\]

Third, \(p_{TFT_{n-1}} > p_{ALLD}\) is equivalent to

\[
m > \frac{(n-1)N}{(b-c-1)(N-n)}. \quad (3.4)
\]

The probability that a single mutant TFT\(_{n-1}\) replaces a population of ALLD is, if small, always non-zero. In the long run, TFT\(_{n-1}\) may be more likely to replace ALLD than vice versa. This is exactly what is expected if \(p_{TFT_{n-1}} > p_{ALLD}\). Note that (3.4) is sufficient for (3.2), suggesting that TFT\(_{n-1}\) is ESS\(_N\) whenever \(p_{TFT_{n-1}} > p_{ALLD}\). On the other hand, the right-hand side of inequality (3.3) is always larger than the right-hand side of inequality (3.4). Hence, \(p_{TFT_{n-1}} > p_{ALLD}\) can actually hold even when ALLD is ESS\(_N\). For large \(N\), (3.4) reduces to \(m > (n-1)/(b-c-1)\) (figure 1).

Figure 1 illustrates the dependence of the critical numbers of rounds on the group size for large \(N\). Considering \(b/c\) as constant (figure 1a), the critical number of rounds above which TFT\(_{n-1}\) is ESS\(_N\) increases monotonically with the group size; thus the maintenance of cooperation is less likely in larger groups. Furthermore, the critical number of rounds required to destabilize ALLD increases quadratically and that required for TFT\(_{n-1}\) to replace ALLD increases linearly with the group size. These relationships suggest that larger group size generally has a negative impact on the evolution of cooperation. It should be mentioned, however, that the possible range of values for \(b/c\) increases with the group size, i.e., \(1 < b/c < n\) must always be satisfied. Consequently, sometimes it may be more reasonable to consider \(b/(nc)\), instead of \(b/c\), as constant. Intuitively, \(b/c\) can be regarded as the efficiency of cooperation for an individual when all of the \(n-1\) opponents cooperate. On the other hand, \(b/(nc)\) can be regarded as the efficiency of cooperation when none of the opponents cooperates.

Let us illustrate what the situation with constant \(b/(nc)\) is like. For the sake of explanation, imagine \(n\) unrelated individuals living in the same nest and jointly defending the nest from intruders. Each individual decides whether to watch intruders outside the nest for a time \(1/n\) of a day. While the benefit of avoiding intruders per day, \(b_i\) is constant, the cost of watching per day, \(c\), is inversely proportional to \(n\). Hence, in this case, \(b/(nc)\) can be regarded as constant.

Considering \(b/(nc)\) as constant, a substantially different picture emerges (figure 1b): as is clear when setting \(b/(nc)\) constant in (3.2) in the limit of large \(N\), the critical numbers of rounds above which TFT\(_{n-1}\) is ESS\(_N\) now decreases monotonically with the group size. Similarly, from (3.4), the critical number of rounds required for TFT\(_{n-1}\) to replace ALLD decreases monotonically with the group size. Interestingly,
(iv) TFT

and (ii) TFT

situation when

are equivalent to

1/2 when

b/c=const. Evolution of cooperation becomes more difficult as the group size increases. The parameter value used is b/(nc) = 0.31.

(3.3) suggests that the dependence of the critical number of rounds required to destabilize ALLD on the group size is not monotonic: the required number of rounds is smallest with an intermediate group size.

The latter findings suggest that the evolution, or at least maintenance, of cooperation is more likely when the group size is larger, given that b/(nc) is constant. This is of particular interest; previous studies on the n-player Prisoner’s Dilemma game incorporating direct (Boyd & Richerson 1988) or indirect (Suzuki & Akiyama 2007, 2008) reciprocity have agreed that the evolution of cooperation becomes easier with increasing group size. The parameter value used is b/(nc) = 0.31.

The frequency of TFT$_{n-1}$ at this equilibrium, $p^*$, is given by

$$p^* = \left( \frac{(n-b/c)/[n(b/c-1)(n-1)]}{N} \right)^{1/(n-1)}.$$

TFT$_{n-1}$ will be lost even when it is ESS unless its initial frequency exceeds $p^*$. Since $p^*$ increases rapidly with $n$, it has been argued that evolution of cooperation is unlikely with sizable group (Boyd & Richerson 1988). In our stochastic model, for large $N$, $p_{TFT_{n-1}} > 1/N$ and $p_{ALLD} > 1/N$ are equivalent to

$$p^* < \left( \frac{2}{n(n+1)} \right)^{1/(n-1)}.$$

Figure 2. The relationship between the fixation probabilities ($p_{TFT_{n-1}}$, $p_{ALLD}$) and the frequency of TFT$_{n-1}$ at the unstable equilibrium ($p^*$). (3.5) holds, then $Np_{TFT_{n-1}} > 1 > Np_{ALLD}$, i.e. selection favours TFT$_{n-1}$ replacing ALLD and opposes ALLD replacing TFT$_{n-1}$. If (3.6) holds, then $Np_{ALLD} > 1 > Np_{TFT_{n-1}}$, i.e. selection favours ALLD and opposes TFT$_{n-1}$. If neither (3.5) nor (3.6) is satisfied, then both $Np_{TFT_{n-1}}$ and $Np_{ALLD}$ are less than 1, i.e. selection opposes the fixation of either strategy. We also find that (3.7) is equivalent to $p_{TFT_{n-1}} > p_{ALLD}$. All these relationships hold for large population size. Although ALLD is evolutionarily stable against invasion by TFT$_{n-1}$ in the deterministic model, in our stochastic model, the probability that a single mutant of TFT$_{n-1}$ takes over an ALLD population can exceed 1/N.

and

$$p^* > \left( \frac{2}{n+1} \right)^{1/(n-1)},$$

respectively. Note that these inequalities give the one-third law (Nowak et al. 2004; Ohtsuki et al. 2007) when $n=2$. Similarly, for large $N$, $p_{TFT_{n-1}} > p_{ALLD}$ is equivalent to

$$p^* < \left( \frac{1}{n} \right)^{1/(n-1)},$$

which reduces to $p^* < 1/2$ when $n=2$. This argument does not necessarily apply to general n-player games (figure 2).

Figure 3 compares the values of $p^*$ across varying $n$ with the thresholds of $p^*$ above which $p_{TFT_{n-1}} > 1/N$, $p_{TFT_{n-1}} < p_{ALLD}$ or $p_{TFT_{n-1}} < 1/N$ holds. Considering $b/c$ as constant (figure 3a), the basin of attraction for the cooperative equilibrium state shrinks (i.e. $p^*$ becomes larger) as the group size increases in the deterministic model. In accordance with this observation, $p_{TFT_{n-1}} > p_{ALLD}$ does not hold true (i.e. the solid line is above the dashed line in figure 3a) when the group size is larger than a critical value in the stochastic model. Hence, both the deterministic and stochastic models suggest that larger group size has a negative impact on the evolution of cooperation. Considering $b/(nc)$ as constant (figure 3b), however, even though the basin of attraction for the cooperative equilibrium state shrinks as the group size increases in the deterministic model (except when $n$ is small), $p_{TFT_{n-1}} > p_{ALLD}$ does hold true in the stochastic model when the group size exceeds a critical value. Hence, our stochastic model reveals what cannot be found by analysing the deterministic model: large group size can facilitate the evolution of cooperation, given that $b/(nc)$ is constant.

4. DISCUSSION

Our model specifies the fixation probability of a strategy in n-player games and reveals that a TFT$_{n-1}$ replacing a
Hence, constant (r threshold above which size. (cooperation becomes more difficult with increasing group (dotted line), among all the children. While this experimental design cooperative children were tripled and equally shared groups. In their experiment, candy bars donated by groups cooperating significantly less than those in small size affects the level of cooperation, with children in large classmates in a public goods game and found that group school children decide whether to cooperate with their relationship between (Nowak et al. 2004; Imhof & Nowak 2006). In the general n-player game, however, the one-third law does not hold (figure 4a). Similarly, for n = 2, A > B holds if and only if the Malthusian fitness of A is positive when p = 1/2. However, in general n-player games, this relationship does not necessarily hold true (figure 4b).

Let us conclude with two more examples of n-player games. First, for the pay-off matrix in figure 4c, choosing strategy B will bring a higher pay-off, no matter what the opponents do. Obviously, B is traditional ESS while A is not. This is a dilemma (but not the Prisoner’s Dilemma), because it tends to be that everyone chooses B though one can gain more if everyone chooses A. By contrast, our analysis reveals that A can be ESS and B cannot be ESS for small N. Second, it can be shown that for n = 2 or 3, no pay-off matrix exists for which both A and B are traditional ESS but neither A nor B is ESS (see the electronic supplementary material). For n ≥ 4, however, a pay-off matrix exists for which both A and B are traditional ESS but neither A nor B is ESS. This discussion holds for weak selection and large population size (figure 4d).

We thank M. Seki and J. Y. Wakano for their comments. This research is supported by the Japan Ministry of Education, Culture, Sports, Science and Technology, Grant-in-Aid for Young Scientists (B), 18770217.

REFERENCES


