Solo pool-punishers defeat free-riders through individual learning in compulsory public goods games

Mitsuru Sano（sano@info.human.nagoya-u.ac.jp）
Kenta Nagasawa, and Masahiro Nagao
〔Nagoya University〕

強制参加公共財ゲームにおいて個別学習を通して協力的なプール罰戦略がフリーライダー戦略に勝利する

佐野 充、長澤 健太、長尾 征洋
名古屋大学 大学院環境学研究科

要約

社会は、利他的で協力的な相互作用で成り立っている。しかし、他者を収奪し自己の利益を最大化する利己的な者が、自然選択では優位である。この協力ジレンマの機構を解明するために、多くの研究が平均得点を基礎として進化ゲーム理論により報告されてきた。しかし、これらの研究は次の深刻な問題を持っている。平均得点を算出するには、すべての公共財ゲームですべてのプレーヤーのすべての得点をすべてのプレーヤーが知らねばならず、それは現実には不可能である。それ故、平均得点ではなく、個人的な学習によって知り得た得点でプレーヤーが行動することを基礎として、どのように協力社会に至るのかを遷移確率を用いて計算する必要がある。本論文で、我々はプール罰をする協力的な人が公共財を増やす小さな乗数 r においてさえもフリーライダーに打ち勝ち、誰の助けも借りずに協力社会を築くことを理論的に示す。これには、フリーライダーが受ける罰の量がプール罰をする人が支払うコスト量よりも大きいことが必要である。また、より小さい集団やより高いゲーム参加率が協力社会をより成立させやすいことも理論的に証明する。

Key words

cooperation, public goods game, individual learning, pool-punishment, participation rate

1. Introduction

Cooperation is a fundamental human behavior (Axelrod & Hamilton, 1981; Boyd & Mathew, 2007; Clutton-Brock, 2009; Dugatkin, 1997; Fehr & Fischbacher, 2003; Hill, 2002; Kaplan, Hill, Lancaster, & Hurtado, 2000; Nowak, 2006; Nowak, 2006; Trivers, 1971) with the potential to improve individual wealth, but it remains a fragile strategy. Individuals who do not contribute but who exploit public goods fare better than those who contribute and pay costs; defectors who are called free-riders therefore obtain a greater payoff. If strategies that are more successful spread, then cooperation will vanish from the population, and public goods will simultaneously disappear. This propagation of the defectors' behavior can drive a population into a "tragedy of the commons" (Hardin, 1968). Thus, clarifying the mechanisms underlying this cooperation dilemma in society is of great importance to many fields of research.

Some prior reports describe how costly punishment might yield a solution to the problem of the cooperation dilemma (Boyd, Gintis, & Bowles, 2003; Boyd & Richerson, 1992; Brandt, Hauert & Sigmund, 2006; Fehr & Gächter, 2002; Fehr & Gächter, 2000; Fowler, 2005; Hauert, Traulsen, Brandt, Sigmund, & Nowak, 2007; Ostrom, Walker, & Gardner, 1992; Rand, Dreber, Ellingsen, Fudenberg, & Nowak, 2009; Rockenbach & Milinski, 2006; Sigmund, Silva, Traulsen, & Hauert, 2010; Sigmund & Hauert, 2001; Silva, Hauert, Traulsen, & Sigmund, 2010; Yamagishi, 1986). A stable cooperative society, however, does not form from only the three strategies of cooperators, defectors, and pool-punishers, as shown in Appendix I with a stochastic model. Cooperators contribute but do not punish. Defectors do not contribute to the public good but exploit the contributions of the other participants. Pool-punishers contribute to the public good and punish all participants who do not contribute to the punishment mechanism. When one punisher invades a state occupied by defectors, he gains a meager payoff from public goods but must contribute to both public goods and the punishment pool. In contrast, although the defectors receive dispersed punishment by the punisher, they gain their payoff from public goods. Therefore, the defector payoff is greater than that of the punisher: the punisher cannot invade the state which the defectors occupy. Consequently, a cooperative society does not form.

Some previous studies (Brandt et al., 2006; Hauert et al., 2007; Silva et al., 2010) developed a voluntary public goods game based on the average payoffs and gave the necessary conditions for natural selection to favor the emergence of cooperation in finite populations. Sigmund et al. also reported that pool-punishers will invade and take over with the four strategies of cooperation, defection, pool-punishment, and loner (Sigmund et al., 2010). These studies rely on the assumptions that players can decide voluntarily whether to participate in the joint enterprise and that loners who do not participate can obtain an income independently of the other players' decisions. But, the loners live alone, getting no payoffs from anyone else. They incur a cost for choosing to be alone. The assumption is therefore illogical. Accordingly, the payoff of the loner can be lower than that of the community in which all residents are defectors, where a cooperative society does not emerge.

2. Problem and Purpose

Now we must point out a severe problem in these studies based on the average payoffs. Estimating the average values requires sufficient knowledge of the payoffs for all players in all PGGs. Such omniscience is clearly difficult to achieve in practice. People make estimates every day based on insufficient knowledge. Players, therefore, make their decisions based on the payoffs received in games played before. Here we examine in compulsory PGGs how pool-punishers overcome defectors in societies comprising those following the three strategies of cooperation, defection, and pool-punishment through individual learning using knowledge of the payoffs from previous encounters, not based on the average payoffs.

3. Model and method

In the compulsory PGG considered here, M denotes the population size with variable compositions of X, Y, and V ($M = X + Y + V$) as the quantities of cooperators, defectors, and pool-punishers. The method is based on a straightforward application of evolutionary game theory to PGGs for finite populations of fixed size (M). Random samples of N individuals chosen from M individuals play a compulsory PGG ($M \geq N$). If $N \geq 2$ individuals participate in the interaction, then each can decide whether to contribute a fixed amount, $c > 0$, to the common-pool, and whether to contribute a fixed amount, $G > 0$, to the punishment-pool. c is multiplied by a factor of r ($N > r > 1$) and is divided among all players. Each player obtains $rc - c$ if all players contribute, whereas their payoffs are 0 if all self-interested players contribute nothing. Punishers specify and sanction players, through imposing penalties, based on the latters' behaviors in a round. The payoffs for the players are the following, where x, y, and v respectively denote the numbers of participants of X, Y, and V ($N = x + y + v$). For x-players, the payoffs are $rc\dfrac{x+v}{N} - c - Bv$ in the presence of y-players and $rc\dfrac{x+v}{N} - c$ in the absence of y-players. For y-players, the payoff is $rc\dfrac{x+v}{N} - Bv$. For v-players, the payoff is $rc\dfrac{x+v}{N} - c - G$. The PGG continues until all M individuals in a generation have played. Individual learning is executed as follows. Two players, i and j, who act respectively as a student and a teacher, are chosen randomly. Student i adopts the strategy of teacher j with probability $\dfrac{1}{1+\exp[-s(p_j - p_i)]}$ for $s \to \infty$, where payoffs are obtained from previous games (Blume, 1993; McFadden, 1981; Sigmund et al., 2010; Szabo & Toke, 1998; Traulsen, Nowak, & Pacheco, 2006). This process is repeated for two or more episodes of learning. Once the learning is completed, the strategies are changed according to the probabilities. The strategy of a randomly chosen player is changed by mutation. Each generation is established in this way. Evolution proceeds over many generations.

4. Results and discussion

The appearance of a pool-punisher in a state occupied by defectors is considered. Two combinations in PGGs are considered. The first combination comprises a single pool-punisher and N-1 defectors. The pool-punisher's payoff is $\dfrac{rc}{N} - c - G$. The defector's payoff is $\dfrac{rc}{N} - B$. The second combination comprises only defectors, whose payoff is zero. The payoffs to players depend on their partners in these PGGs and might differ from the average. Moreover, in this setting, learning leading to the preferential copying of successful strategies is dependent on the payoffs used.

Given a case of individual learning in which two chosen individuals have knowledge of each other's payoffs from previous games, and assuming that a pool-punisher with a maximum payoff (P_{Vmax}) is designated as a teacher and that a defector with a minimum payoff (P_{Ymin}) is designated as a student, then if the value of P_{Vmax} is higher than that of P_{Ymin}, defectors will imitate pool-punishers. The number of pool-punishers will increase if such a case occurs repeatedly by chance. Accordingly, invasion by a single pool-punisher requires that $P_{\mathrm{Vmax}} > P_{\mathrm{Ymin}}$. The case is described by the equation $B > c + G$,[(1)] where the punishment (B) that the defector receives is greater than the cost ($c + G$) to the pool-punisher.

Next considering the case of i ($i \geq 2$) pool-punishers and $M - i$ defectors, based on the condition $B > c + G$, the maximum payoff for the pool-punishers is higher than the minimum payoff for the defectors. Therefore, the defectors will imitate the pool-punishers and the pool-punishers thereby establish stable cooperation which defectors cannot invade. In addition, the transition probabilities with these payoffs resemble those with the average payoffs for the conversion of all cooperators to all defectors and vice versa, and for the conversion of all cooperators to all pool-punishers and vice versa. This demonstrates that pool-punishers can establish a cooperative society under the condition $B > c + G$.

A stationary distribution is computed using a transition matrix for the payoffs from previous games through individual learning. The fixation probabilities ρ_{XV} and ρ_{YV} tend to differ from the average payoffs, although some are the same: $\rho_{YX} = 0$, $\rho_{VX} = 0$, and $\rho_{VY} = 0$ under the condition $B > c + G$. The transition matrix therefore reduces to the following:

$$\begin{bmatrix} 1-\frac{\rho_{XY}}{2}-\frac{\rho_{XV}}{2} & \frac{\rho_{XY}}{2} & \frac{\rho_{XV}}{2} \\ 0 & 1-\frac{\rho_{YV}}{2} & \rho_{YV} \\ 0 & 0 & 1 \end{bmatrix}.$$

In this matrix, $\rho_{XY} \neq 0$ and $\rho_{YV} \neq 0$. The stationary distribution for (X, Y, V) is therefore given as (0, 0, 1), meaning that the pool-punishers eventually prevail. That is, under the condition $B > c + G$, pool-punishers can establish a cooperative society stably. Moreover, pool-punishers prevail in the four-strategy case that includes peer-punishers in addition to these other three strategies. The constraint $B > c + G$ means that the defector's punishment is greater than the pool-punisher's cost. This inequality corresponds to a social rule within a policing system whereby the authorities impose an additional penalty on a tax cheat. Our society accepts this restriction as a rational rule.

We should note the following point when a pool-punisher appears in a state occupied by defectors. In a case where a single pool-punisher overcomes $N - 1$ defectors, the cost G is converted to the total punishment $B(N - 1)$. This conversion equates to a high-performance punishment mechanism. Examples of such practices might be found in primitive religions, white magic, oracles, shamanism, and social norms (Benson, 1989; Eliade, 1972; Malinowski, 1961; Weber, 1976; Sugiyama, 1991). However, further research is necessary to clarify this situation.

The long-term frequencies of the three strategies can be examined using numerical simulations for a case in which the two individuals chosen as the student and the teacher know each other's payoffs from their previous games through individual learning. For the case where $M = 20$ and $N = 5$, Figure 1 shows that a single pool-punisher who appears repeatedly by mutation eventually overcomes defectors in some generation under the condition $B > c + G$, even when the value of r is close to 1 (Figure 2). However, when the condition $B > c + G$ is not satisfied, defectors prevail as shown in Figure 3.

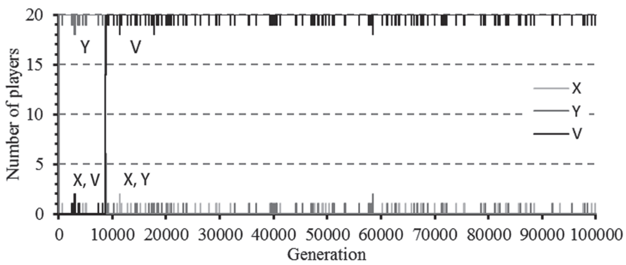

Figure 1: Evolution of competition in populations consisting of X, Y, and V under condition $B > c + G$. Parameters are $M = 20$, $N = 5$, $r = 3$, $c = 1$, $G = 0.5$, $B = 1.6$, and $\mu = 10^{-4}$. Updating occurred by strong imitation ($s \to \infty$), i.e., a student with a lower average payoff always imitated a teacher with a higher average payoff. The initial populations were set as $X = 0$, $Y = 20$, and $V = 0$.

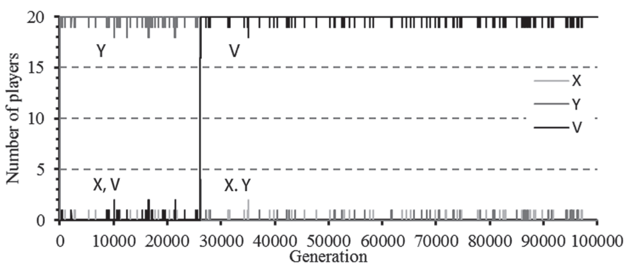

Figure 2: Evolution of competition in populations consisting of X, Y, and V under condition $B > c + G$. Parameters are the same as those in Figure 1, except $c = 1.1$.

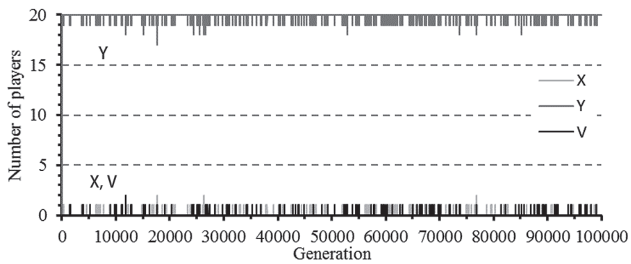

Figure 3: Evolution of competition in populations consisting of X, Y, and V under condition $B < c + G$. Parameters are the same as those in Figure 1, except $B = 1.4$.

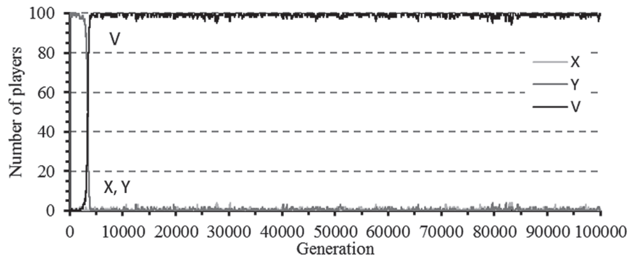

Figure 4: Evolution of competition in populations consisting of X, Y, and V with a larger population and higher participation rate. Parameters are $M = 100$, $N = 20$, $r = 3$, $c = 1$, $G = 0.5$, $B = 1.6$, and $\mu = 10^{-4}$. Initial populations were set as $X = 0$, $Y = 100$, and $V = 0$.

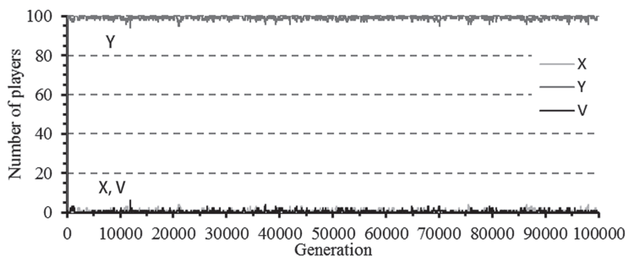

Figure 5: Evolution of competition in populations consisting of X, Y, and V with a larger population and lower participation rate. Parameters are $M = 100$, $N = 5$, $r = 3$, $c = 1$, $G = 0.5$, $B = 1.6$, and $\mu = 10^{-4}$. Initial populations were set as $X = 0$, $Y = 100$, and $V = 0$.

Two simulation results with $N = 20$ and $N = 5$ in $M = 100$ are shown respectively as Figures 4 and 5. In $N = 20$ (Figure 4), a single pool-punisher easily overcomes defectors under the condition $B > c + G$. Nevertheless, as presented in Figure 5 for $N = 5$, a pool-punisher cannot overcome defectors with a lower participation rate even if the restriction is satisfied. This inability demonstrates that establishing a cooperative society depends on the participation ratio. This result agrees well with a sociologist's observation (Olson, 1965; Putnam, 1993). In contrast, prior works based on the average payoffs reported that establishing a cooperative society is independent of the participation rate, which is inconsistent with the observation.

We next specifically examine the probability of fixation when moving from a state with a single pool-punisher and $M - 1$ defectors to a state with M pool-punishers under the condition $B > c + G$. This relates to the time period for establishing a cooperative society.

In a case with one individual learning episode per generation, Nowak (2006) reported that the fixation probability x_I is the equation $x_I = \dfrac{1}{1+\sum_{j=i}^{N-1}\prod_{k=i}^{j}\gamma_k}$ when moving from a state with a single pool-punisher and $M - 1$ defectors to that with M pool-punishers. Here, $\gamma_k = \dfrac{\beta_k}{\alpha_k}$ and α_k and β_k respectively denote the probabilities of transitions from k to $k + 1$ and from k to $k - 1$. In a case with two episodes of individual learning per generation, the fixation probabilities are expressed as follows:

$$\begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_M \end{pmatrix} = P\begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_M \end{pmatrix}.$$

Here, P is expressed by the matrix at the bottom of this page:

Calculating the values of α_k and β_k requires knowledge of the order of the magnitudes of the pool-punisher's payoffs P_V and the defector's payoffs P_Y. Although the values are dependent on the parameters c, r, G, and B, their order is the following:

$$\ldots > P_{YN} > \ldots > P_{V1} > P_{YN\text{-}1} > P_{YN\text{-}2} > \ldots > P_{Y2} > P_{Y1}.$$

Here, $P_{YN\text{-}1}$ and P_{V1} respectively denote the defector's and pool-punisher's payoffs obtained by $N - 1$ defectors and a single pool-punisher in their PGG

For $N = 5$, $r = 3$, $c = 1$, $G = 0.5$, and $B = 1.6$, the magnitudes have the following order:

$$P_{V5} > P_{V4} > P_{V3} > P_{Y5} > P_{V2} > P_{V1} > P_{Y4} > P_{Y3} > P_{Y2} > P_{Y1}.$$

We can thereby obtain the following equations:

$$\begin{aligned}\alpha_k = {} & \frac{N_{Y5}}{M}\cdot\frac{N_{V5}+N_{V4}+N_{V3}}{M-1} + \frac{N_{Y4}+N_{Y3}+N_{Y2}+N_{Y1}}{M}\cdot\frac{N_{V5}+N_{V4}+N_{V3}}{M-1} \\ & + \frac{N_{Y4}+N_{Y3}+N_{Y2}+N_{Y1}}{M}\cdot\frac{N_{V2}+N_{V1}}{M-1} \\ = {} & \frac{N_{Y5}+N_{Y4}+N_{Y3}+N_{Y2}+N_{Y1}}{M}\cdot\frac{N_{V5}+N_{V4}+N_{V3}}{M-1} \\ & + \frac{N_{Y4}+N_{Y3}+N_{Y2}+N_{Y1}}{M}\cdot\frac{N_{V2}+N_{V1}}{M-1}.\end{aligned}$$

$$\beta_k = \frac{N_{V2}+N_{V1}}{M}\cdot\frac{N_{Y5}}{M-1}$$

Here, N_{Vh} and $N_{YN\text{-}h}$ respectively denote the numbers of pool-punishers with a payoff of P_{Vh} and defectors with a payoff of $P_{YN\text{-}h}$ in a population of M individuals. After a sufficient number of games, they are expressed as shown below.

$$N_{Vh} = M\cdot\frac{\binom{V}{h}\binom{M-V}{N-h}}{\binom{M}{N}}\cdot\frac{h}{N}.$$

$$N_{YN-h} = M\cdot\frac{\binom{V}{h}\binom{M-V}{N-h}}{\binom{M}{N}}\cdot\frac{N-h}{N}.$$

$$\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\
\beta_{1_1} & 1-\alpha_{1_1}-\beta_{1_1} & \alpha_{1_1} & \alpha_{2_1} & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\
\beta_{2_2} & \beta_{1_2} & 1-\alpha_{2_2}-\alpha_{1_2}-\beta_{1_2}-\beta_{2_2} & \alpha_{1_2} & \alpha_{2_2} & \cdots & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & \cdots & \beta_{2_{M-2}} & \beta_{1_{M-2}} & 1-\alpha_{2_{M-2}}-\alpha_{1_{M-2}}-\beta_{1_{M-2}}-\beta_{2_{M-2}} & \alpha_{1_{M-2}} & \alpha_{2_{M-2}} \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & \beta_{2_{M-1}} & \beta_{1_{M-1}} & 1-\alpha_{2_{M-1}}-\alpha_{1_{M-1}}-\beta_{1_{M-1}}-\beta_{2_{M-1}} & \alpha_{1_{M-1}} \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1
\end{pmatrix}$$

where $\alpha_{2_k} = \alpha_k\cdot\alpha_k$ $\alpha_{1_k}, = 2\cdot\alpha_k\cdot(1-\alpha_k-\beta_k)$, $\beta_{1_k} = 2\cdot\beta_k\cdot(1-\alpha_k-\beta_k)$, and $\beta_{2_k} = \beta_k\cdot\beta_k$.

Therefore, we can derive the following equations:

$$\alpha_k = \frac{M}{(M-1)N^2\binom{M}{N}^2}\left(\sum_{N-h=1}^{5}\binom{k}{h}\binom{M-k}{N-h}(N-h)\cdot\sum_{h=3}^{5}\binom{k}{h}\binom{M-k}{N-h}h + \sum_{N-h=1}^{4}\binom{k}{h}\binom{M-k}{N-h}(N-h)\cdot\sum_{h=1}^{2}\binom{k}{h}\binom{M-k}{N-h}h\right).$$

$$\beta_k = \frac{M}{(M-1)N^2\binom{M}{N}^2}\left(\sum_{h=1}^{2}\binom{k}{h}\binom{M-k}{N-h}\cdot h\cdot\binom{k}{0}\binom{M-k}{5}\cdot N\right).$$

For the case in which $N = 5$, $r = 1.1$, $c = 1$, $G = 0.5$, and $B = 1.6$, the order of the magnitudes is

$$P_{Y5} > P_{V5} > P_{V4} > P_{V3} > P_{V2} > P_{V1} > P_{Y4} > P_{Y3} > P_{Y2} > P_{Y1}$$

and we can thereby obtain the following equations:

$$\alpha_k = \frac{N_{Y4}+N_{Y3}+N_{Y2}+N_{Y1}}{M}\cdot\frac{N_{V5}+N_{V4}+N_{V3}+N_{V2}+N_{V1}}{M-1}.$$

$$\beta_k = \frac{N_{V5}+N_{V4}+N_{V3}+N_{V2}+N_{V1}}{M}\cdot\frac{N_{Y5}}{M-1}.$$

Table 1 presents results for a case with a low mutation rate and two episodes of individual learning during one generation. Fixation probabilities decrease concomitantly with increased M and decreased N, i.e., the fixation probability decreases with a decrease in the participation rate in PGG. A probability of 0.001 is obtained for five participants in a population of 40 individuals. The emergence of a cooperative society requires that, on average, one pool-punisher arises per 1,000 events. Because pool-punishers appear repeatedly by mutation and because the cooperative society is stable once it has been established, this value is not regarded as unrealistic. Dunbar reported that establishing a stable cooperative society required 30-50 participants in a population of about 150 individuals (Dunbar, 1993). In the current case, the calculated probabilities are 0.03-0.18, which easily establishes cooperation.(2) It is noteworthy that the calculated probabilities are greater when the population size is smaller and the participation rate is higher. The observation of larger probabilities in smaller groups agrees
with sociological findings (Olson, 1965). Our results are also consistent with Putnam's observation that the participation rate is highly correlated with the emergence of cooperation (Putnam, 1993). Notably, the mutation rate is related to the emergence of a pool-punisher in a defector's society. Higher values shorten the period necessary to establish a cooperative society. Moreover, if pool-punishers or cooperators who appear by mutation in a defector's society are not chosen as students in the learning, they persist in the succeeding generation, leading to a higher probability than that calculated here and relaxing the requirement that $B > c + G$.

5. Conclusion

Through individual learning, we examined how pool-punishers overcome defectors to establish stable cooperation within a society without help from either non-participants or players using other strategies. The main results are presented below. A pool-punisher can establish stable cooperation easily in a society, provided that the defector's punishment is greater than the pool-punisher's cost, which is consistent with our social rules. Higher fixation probabilities were found to be obtained with a smaller population size and a higher participation rate. This study shed

Table 1: Fixation probabilities moving from a single pool-punisher to all pool-punishers with two episodes of individual learning during a generation. The probability shows that a smaller population and a higher participation rate higher values. The other parameters; r = 3; c = 1; G = 0.5; B = 1.6.

	M=10	M=20	M=30	M=40	M=50	M=60	M=70	M=80	M=90	M=100
N=20			0.63	0.41	0.26	0.17	0.11	0.072	0.047	0.031
N=10		0.39	0.16	0.064	0.026	0.010	$4x10^{-3}$	$2x10^{-3}$	$7x10^{-4}$	$3x10^{-4}$
N=9	0.89	0.32	0.11	0.042	0.015	$6x10^{-3}$	$2x10^{-3}$	$7x10^{-4}$	$3x10^{-4}$	$1x10^{-4}$
N=8	0.78	0.24	0.076	0.024	$8x10^{-3}$	$2x10^{-3}$	$7x10^{-4}$	$2x10^{-4}$	$8x10^{-5}$	$3x10^{-5}$
N=7	0.65	0.17	0.045	0.012	$3x10^{-3}$	$8x10^{-4}$	$2x10^{-4}$	$6x10^{-5}$	$2x10^{-5}$	$4x10^{-6}$
N=6	0.51	0.10	0.021	$4x10^{-3}$	$9x10^{-4}$	$2x10^{-4}$	$4x10^{-5}$	$8x10^{-6}$	$2x10^{-6}$	$3x10^{-7}$
N=5	0.36	0.053	$8x10^{-3}$	$1x10^{-3}$	$2x10^{-4}$	$3x10^{-5}$	$6x10^{-6}$	$9x10^{-7}$	$1x10^{-7}$	$2x10^{-8}$
N=4	0.21	0.017	$1x10^{-3}$	$1x10^{-4}$	$1x10^{-5}$	$1x10^{-6}$	$1x10^{-7}$	$9x10^{-9}$	$8x10^{-10}$	$7x10^{-11}$
N=3	0.12	0.011	$1x10^{-3}$	$1x10^{-4}$	$2x10^{-5}$	$2x10^{-6}$	$3x10^{-7}$	$4x10^{-8}$	$5x10^{-9}$	$7x10^{-10}$

light on the processes of self-organization in human societies. Our future research will examine under what conditions cooperation can be established by pool-punishers through individual learning even in the presents of anti-social punishment.

Notes

(1) The defectors obtain two payoffs: $\frac{rc}{N} - B$ and 0. When $\frac{rc}{N} - B \leq 0$, $P_{\mathrm{Ymin}} = \frac{rc}{N} - B$. Assuming $P_{\mathrm{Vmax}} > P_{\mathrm{Ymin}}$, we obtain the condition $B > c + G$, in which a pool-punisher can invade a defector's society. By contrast, when $\frac{rc}{N} - B \geq$, $P_{\mathrm{Ymin}} = 0$. The condition $\frac{rc}{N} - c - G > 0$ can be derived from $P_{\mathrm{Vmax}} > P_{\mathrm{Ymin}}$. These findings are not incompatible because the dilemma condition demands that $c > \frac{rc}{N}$ and that the value of G be positive. Consequently, we obtain the condition $B > c + G$.

(2) The calculated probabilities are 0.18, 0.094, and 0.032 for $N = 50$, 40, and 30, respectively, in $M = 150$. The other parameters are $r = 3$, $c = 1$, $G = 0.5$, and $B = 1.6$.

Acknowledgments

We thank Mr. T. Nii for useful discussions related to social and individual learning.

References

Axelrod, R. & Hamilton,W. D. (1981). The evolution of cooperation. *Science*, 211, 1390-1396.

Benson, B. L. (1989). Enforcement of private property rights in primitive societies: Law without government. *The Journal of Libertarian Studies*, 9, 1-26.

Blume, L. E. (1993). The statistical-mechanics of strategic interaction. *Games and Economic Behavior*, 5, 387-424.

Boyd, R. & Mathew, S. (2007). Behavior - A narrow road to cooperation. *Science*, 316, 1858-1859.

Boyd, R., Gintis, H., Bowles, S., & Richerson, P. J. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 3531-3535.

Boyd, R. & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, 13, 171-195.

Brandt, H., Hauert, C., & Sigmund, K. (2006). Punishing and abstaining for public goods. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 495-497.

Clutton-Brock, T. (2009). Cooperation between non-kin in animal societies. *Nature*, 462, 51-57.

Dugatkin, L. A. (1997). *Cooperation Among Animals: An Evolutionary Perspective,* Princeton, NJ: Oxford University Press.

Dunbar, R. I. M. (1993). Coevolution of neocortical size, group-size and language in humans. *Behavioral and Brain Sciences*, 16, 681-735.

Eliade, M. (1972). *Shamanism: Archaic Techniques of Ecstasy*, Princeton, NJ: Princeton University Press.

Fehr, E. & Fischbacher, U. (2003). The nature of human altruism. *Nature*, 425, 785-791.

Fehr, E. & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415, 137–140.

Fehr, E. & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *The American Economic Review*, 90, 980-994.

Fowler, J. H. (2005). Altruistic punishment and the origin of cooperation. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 7047-7049.

Hardin, G. (1968). The tragedy of the commons. *Science*, 162, 1243-1248.

Hauert, C., Traulsen, A., Brandt, H., Sigmund, K., & Nowak, M. A. (2007). Via freedom to coercion. *Science*, 316, 1905-1907.

Hill, K. (2002). Altruistic cooperation during foraging by the ache, and the evolved human predisposition to cooperate. *Human Nature*, 13, 105-128.

Kaplan, H., Hill, J., Lancaster, J., & Hurtado, A. M., (2000). A theory of human life history evolution: Diet, intelligence, and longevity. *Evolutionary Anthropology*, 9, 156-185.

Malinowski, B. (1961). *Crime and Custom in Savage Society*, London: Routledge & Kegan Paul, Ltd.

McFadden, D. (1981). *Structural Analysis of Discrete Data and Econometric Applications*. Cambridge, MA: MIT Press.

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, 314, 1560-1563.

Nowak, M. A. (2006). *Evolutionary Dynamics: Exploring the Equations of Life*, Cambridge, MA: Harvard University Press.

Olson, M. (1965). *The Logic of Collective Action*, Cambridge, MA: Harvard University Press.

Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants With and Without a Sword. *The American Political Science Review*, 86, 404-417.

Putnam, R. D. (1993). *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton, NJ: Princeton University Press.

Rand, D. G., Dreber, A., Ellingsen, T., Fudenberg, D., & Nowak, M. A. (2009). Positive Interactions Promote Public Cooperation. *Science*, 325, 1272-1275.

Rockenbach, B. & Milinski, M. (2006). The efficient interaction of indirect reciprocity and costly punishment. *Nature*, 444, 718-723.

Sigmund, K., Silva, H. D., Traulsen, A., & Hauert, C. (2010). Social learning promotes institutions for governing the commons. *Nature*, 466, 861-863.

Sigmund, K., Hauert, C., & Nowak, M. A. (2001). Reward and punishment. *Proceedings of the National Academy of Sci-*

ences of the United States of America, 98, 10757-10762.

Silva, H. D., Hauert, C., Traulsen, A., & Sigmund, K. (2010). Freedom, enforcement, and the social dilemma of strong altruism. *Journal of Evolutionary Economics*, 20, 203-217.

Sugiyama, H. (1991). The Political Power and Punishment in Ancient Japan. *Waseda Law Review*, 67, 1-23.

Szabo, G. & Toke, C. (1998). Evolutionary prisoner's dilemma game on a square lattice. *Physical Review E*, 58, 69-73.

Traulsen, A., Nowak, M. A., & Pacheco, J. M. (2006). Stochastic dynamics of invasion and fixation. *Physical Review E*, 74, 011909.

Trivers, R. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46, 35-57.

Weber, M. (1976). *Wirtschaft und Gesellschaft, Grundriss der verstehenden Soziologie*, Mohr, Tübingen.

Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology*, 51, 110-116.

(Received November 28, 2014; accepted December 10, 2014)

Appendix I

Stationary distributions with average payoffs in social learning:

The model PGG comprises a finite population of players with the three strategies of cooperation, defection, and pool-punishment (Sigmund et al., 2010). Public goods are distributed based on altruism. Cooperators contribute to public goods but not to the punishment-pool, which is used to penalize free-riders and second-order free-riders. As a consequence of overlooking the free-riding, cooperators receive a second-order penalty from an agent of the punishment-pool (such as a police officer) delegated by its members. Defectors contribute to neither public goods nor the punishment-pool. Moreover, they receive a penalty because of their free-riding. Pool-punishers contribute both to the public goods and to the punishment-pool used to penalize both cooperators and defectors. The numbers of cooperators, defectors, and pool-punishers are denoted X, Y, and V ($M = X + Y + V$). Random samples of N individuals play the PGG, and x, y, and v respectively denote the quantities of the X, Y, and V participants in a game ($N = \mathbf{x} + y + v$). The average payoffs for the players follow those reported elsewhere in the literature (Sigmund et al., (2010), specifically,

$$P_X = rc\frac{M-Y-1}{M-1} - c - \frac{B(N-1)V}{M-1}.$$

$$P_Y = rc\frac{M-Y}{M-1} - \frac{B(N-1)V}{M-1}.$$

$$P_V = rc\frac{M-Y-1}{M-1} - c - G.$$

For a small probability of mutation μ, the embedded Markov chain describing the transitions between All_X, All_Y, and All_V is given as shown below.

$$\begin{bmatrix} 1-\frac{\rho_{XY}}{2}-\frac{\rho_{XV}}{2} & \frac{\rho_{XY}}{2} & \frac{\rho_{XV}}{2} \\ \frac{\rho_{YX}}{2} & 1-\frac{\rho_{YX}}{2}+\frac{\rho_{YV}}{2} & \frac{\rho_{YV}}{2} \\ \frac{\rho_{VX}}{2} & \frac{\rho_{VY}}{2} & 1-\frac{\rho_{VX}}{2}-\frac{\rho_{VY}}{2} \end{bmatrix}.$$

For the average payoffs, this transition matrix reduces to the following.

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The stationary distribution for (X, Y, V) is given as $(0, 1 - \frac{V}{M}, \frac{V}{M})$. The population is dependent upon an initial distribution with coexisting defectors and pool-punishers. For the three strategies, a cooperative society was not established.

Appendix II

Table 1b: Fixation probabilities moving from a single pool-punisher to all pool-punishers with one episode of individual learning during a generation. The other parameters were r = 3, c = 1, G = 0.5 and B = 1.6. For the values in parentheses, the other parameters were r = 1.1, c = 1, G = 0.5 and B = 1.6.

	M =10	M=20	M=30	M=40	M=50	M=60	M=70	M=80	M=90	M=100
N=20			0.63	0.40	0.26	0.17	0.11	0.072	0.047	0.031
N=10		0.39	0.16	0.064	0.026	0.010	4×10^{-3}	2×10^{-3}	7×10^{-4}	3×10^{-4}
N=9	0.89	0.32	0.11	0.042	0.015	5×10^{-3}	2×10^{-3}	7×10^{-4}	3×10^{-4}	1×10^{-4}
N=8	0.77	0.24	0.076	0.024	8×10^{-3}	2×10^{-3}	7×10^{-4}	2×10^{-4}	8×10^{-5}	3×10^{-5}
N=7	0.65	0.17	0.045	0.012	3×10^{-3}	8×10^{-4}	2×10^{-4}	6×10^{-5}	2×10^{-5}	4×10^{-6}
N=6	0.51	0.10	0.021	4×10^{-3}	9×10^{-4}	2×10^{-4}	4×10^{-5}	8×10^{-6}	2×10^{-6}	3×10^{-7}
N=5	0.36 (0.36)	0.054 (0.050)	8×10^{-3} (7×10^{-3})	1×10^{-3} (1×10^{-3})	2×10^{-4} (1×10^{-4})	3×10^{-5} (2×10^{-5})	6×10^{-6} (3×10^{-6})	9×10^{-7} (4×10^{-7})	2×10^{-7} (5×10^{-8})	2×10^{-8} (7×10^{-9})
N=4	0.21	0.017	2×10^{-3}	1×10^{-4}	1×10^{-5}	1×10^{-6}	1×10^{-7}	9×10^{-9}	8×10^{-10}	7×10^{-11}
N=3	0.12	0.011	1×10^{-3}	1×10^{-4}	2×10^{-5}	2×10^{-6}	3×10^{-7}	4×10^{-8}	5×10^{-9}	7×10^{-10}

Table 1c: Fixation probabilities moving from a single pool-punisher to all pool-punishers with three episodes of individual learning during one generation. The other parameters were r = 3, c = 1, G = 0.5 and B = 1.6.

	M=10	M=20	M=30	M=40	M=50	M=60	M=70	M=80	M=90	M=100
N=20			0.63	0.41	0.26	0.17	0.11	0.072	0.047	0.031
N=10		0.39	0.16	0.064	0.026	0.010	4×10^{-3}	2×10^{-3}	7×10^{-4}	3×10^{-4}
N=9	0.89	0.32	0.12	0.042	0.015	6×10^{-3}	2×10^{-3}	7×10^{-4}	3×10^{-4}	1×10^{-4}
N=8	0.78	0.24	0.076	0.024	8×10^{-3}	2×10^{-3}	7×10^{-4}	2×10^{-4}	8×10^{-5}	2×10^{-5}
N=7	0.65	0.17	0.044	0.012	3×10^{-3}	8×10^{-4}	2×10^{-4}	6×10^{-5}	2×10^{-5}	4×10^{-6}
N=6	0.51	0.10	0.021	4×10^{-3}	9×10^{-4}	2×10^{-4}	4×10^{-5}	8×10^{-6}	2×10^{-6}	3×10^{-7}
N=5	0.37	0.053	8×10^{-3}	1×10^{-3}	2×10^{-4}	3×10^{-5}	5×10^{-6}	9×10^{-7}	1×10^{-7}	2×10^{-8}
N=4	0.21	0.017	1×10^{-3}	1×10^{-4}	1×10^{-5}	1×10^{-6}	1×10^{-7}	8×10^{-9}	8×10^{-10}	7×10^{-11}
N=3	0.12	0.011	1×10^{-3}	1×10^{-4}	2×10^{-5}	2×10^{-6}	3×10^{-7}	4×10^{-8}	5×10^{-9}	7×10^{-10}

Table 1d: Fixation probabilities moving from a single pool-punisher to all pool-punishers with five episodes of individual learning during one generation. The other parameters were r = 3, c = 1, G = 0.5 and B = 1.6.

	M=10	M=20	M=30	M=40	M=50	M=60	M=70	M=80	M=90	M=100
N=20			0.67	0.47	0.33	0.23	0.16	0.11	0.079	0.055
N=10		0.42	0.19	0.081	0.035	0.015	7×10^{-3}	3×10^{-3}	1×10^{-3}	5×10^{-4}
N=9	0.91	0.35	0.14	0.053	0.020	8×10^{-3}	3×10^{-3}	1×10^{-3}	5×10^{-4}	2×10^{-4}
N=8	0.80	0.27	0.090	0.030	0.010	3×10^{-3}	1×10^{-3}	4×10^{-4}	1×10^{-4}	4×10^{-5}
N=7	0.68	0.19	0.0453	0.015	4×10^{-3}	1×10^{-3}	3×10^{-4}	9×10^{-5}	3×10^{-5}	7×10^{-6}
N=6	0.54	0.12	0.025	5×10^{-3}	1×10^{-3}	3×10^{-4}	6×10^{-5}	1×10^{-5}	3×10^{-6}	6×10^{-7}
N=5	0.39	0.059	0.010	2×10^{-3}	3×10^{-4}	5×10^{-5}	8×10^{-6}	1×10^{-6}	2×10^{-7}	4×10^{-8}
N=4	0.22	0.018	2×10^{-3}	2×10^{-4}	1×10^{-5}	1×10^{-6}	1×10^{-7}	1×10^{-8}	1×10^{-9}	1×10^{-10}
N=3	0.12	0.012	1×10^{-3}	2×10^{-4}	2×10^{-5}	3×10^{-6}	4×10^{-7}	6×10^{-8}	8×10^{-9}	1×10^{-9}

Abstract

Human societies are organized around cooperative and altruistic interactions. Natural selection, however, favors selfish and strong individuals who maximize their own resources at the expense of others. Although many previous studies with average payoffs have developed mechanisms for resolving the cooperation dilemma, they have a severe problem. Estimating the average values requires sufficient knowledge of the payoffs for all players in all public goods games (PGGs), which is difficult to achieve in practice. People make estimates every day based on insufficient knowledge. The transition probabilities ought to be therefore calculated based on known payoffs rather than on the average. Through this individual learning, we show that pool-punishers can overcome free-riders to establish stable cooperation within a society without help from non-participants or players using other strategies, even with a small multiplier (r) of public goods. This scenario requires the punishment that the free-riders receive to be greater than the cost to the pool-punishers. We also demonstrate that smaller population sizes and higher participation rates engender greater fixation probabilities for cooperation.