Journal of Information Processing Vol.24 No.3 534–539 (May 2016)

[DOI: 10.2197/ipsjjip.24.534]

Regular Paper

Study of the Evolution of Cooperation Based on an Alternative Notion of Punishment "Sanction with Jealousy"

Tetsushi Ohdaira[1,a)]

Received: July 12, 2015, Accepted: February 8, 2016

Abstract: The effectiveness of punishment that a player pays certain costs and punishes an uncooperative player is currently discussed in the field of the study of cooperation under non-kin relationships. The discussions of the effectiveness of punishment are based on either negative or positive point of view. Contrary to these previous discussions, this study proposes a novel model introducing an alternative notion of punishment "sanction with jealousy". The degree of sanction is proportional to the payoff of the sanctioning player. The condition for sanction to occur reflects jealousy in that each player sanctions their neighbor players when their payoff is smaller than the payoff of their neighbor players. Utilizing this model, the author investigates whether the introduction of the sanction with jealousy improves both the number of players having the strategy of cooperation and the average payoff of all players or not. In addition, the author organizes the new findings from this investigation in comparison with these previous discussions.

Keywords: agent designs, architectures, and theories, agent-based simulation, agent-based system development, multi-agent games, imperfect information games

1. Introduction

Currently, the study of cooperation under non-kin relationships mainly discusses the effectiveness of punishment, and these discussions are based on either negative [1], [2], [3], [4], [5] or positive [6], [7], [8], [9], [10], [11], [12], [13], [14] point of view. Now, punishment means that a player paying certain costs charges an uncooperative player with the loss of payoff. Briefly explaining these recent discussions, on the one hand, in the prisoner's dilemma game, Dreber et al. [2] point out that the introduction of punishment reduces the average payoff of all players. Furthermore, in the public goods game that is the multiplayer extension of the prisoner's dilemma game, Rand and Nowak [5] propose that punishment is an egocentric tool to protect a player itself because through natural selection, punishment does not increase the number of players having the strategy of cooperation (=cooperators), and promotes antisocial behavior, i.e., the retaliation like the punishment of a player having the strategy of defection (=defector) on a cooperator.

On the other hand, Garcia and Traulsen [13] reveal that Rand and Nowak [5] only cover very limited case where there are some players who abstain from the collective action. They show that even if the antisocial behavior is possible, cooperators who punish only defectors can prosper enough when these abstaining players are isolated. Perc and Szolnoki [14] introduce adaptive punishment that enables a player to vary the degree to which to perform punishment in response to the degree of the success of cooperation. They show that the adaptive punishment activates the reciprocity on the spatial relationships, and facilitates cooperation.

Here, contrary to these previous discussions, this study proposes an alternative notion of punishment, "sanction with jealousy". This notion, as expressed in its name, realizes the sanction by humans from jealousy because each player sanctions their neighbor players when their payoff is smaller than the payoff of their neighbor players. Comparing the sanction with jealousy with peer-punishment [6] and pool-punishment [11], in the sanction with jealousy, sanctioning costs and the degree of sanction are not constant, but dynamically change because they are proportional to the payoff of the sanctioning player. The sanctioning player is also damaged a little because sanctioning activity imposes some stress on the sanctioning player. Fehr and Schmidt's inequity aversion [15] also considers the difference of payoff between sanctioning and sanctioned players regarding the cost and the degree of punishment. However, this study differs from Fehr and Schmidt's study because they deal with the public goods game rather than the prisoner's dilemma game, and a player can punish not related but all other players. This paper investigates whether the introduction of the sanction with jealousy induces the increase and the maintenance of cooperative players (=the evolution of cooperation) and also the improvement in the average payoff of all players or not. The author also considers the previous discussions on the effectiveness of punishment and organizes the new findings from this investigation.

1 Institute of Information and Media, Aoyama Gakuin University, Shibuya, Tokyo 150–8366, Japan

a) odaira@aim.aoyama.ac.jp

2. Model

This study is basically based on Nowak and May's spatial prisoner's dilemma game [16], and introduces the sanction with jealousy. Each player has their strategy of defection or cooperation, matches their neighbor players, and gains the resulting payoff. When we express the total number of players as N, two players of the match as players i and j ($i \neq j$, $1 \leq i, j \leq N$), their strategy as $s(i)$ and $s(j)$, and their payoff as $p(i)$ and $p(j)$, $p(i)$ is expressed as the following Eq. (1) by utilizing the payoff matrix A of Eq. (2). Note that $s(i)$ and $s(j)$ are expressed as (1 0) or (0 1) by unit vectors. The former is the strategy of cooperation and the latter is the strategy of defection. $O(i)$ represents the collection of the opponents of player i.

$$p(i) = \sum_{j \in O(i)} s(i) A s(j)^T \quad (i \neq j,\ 1 \leq i,\ j \leq N) \tag{1}$$

$$A = \begin{pmatrix} 1 & 0 \\ b & 0 \end{pmatrix} \quad (1 < b \leq 2) \tag{2}$$

As shown in **Fig. 1**, this study defines neighbor players around each player as the one-dimensional regular lattice [17]. A node of the lattice exhibits a player, and the total number of neighbor players around each player ($e(i)$) is the same among all players. Figure 1 shows the initial state; 10% of all players enclosed in squares are defectors, and others are cooperators. This initial percentage of the number of defectors follows the Nowak and May's study (see Fig. 1 of Ref. [16]).

Next the author describes the implementation of an alternative notion of punishment "sanction with jealousy". Dreber et al. [2] treat the punishment of a player on another player as the strategy of game. On the other hand, in this study, no player has the strategy of costly punishment but does have the option of sanction with jealousy. Player i sanctions player j when player i recognizes that player j in $O(i)$ has the larger payoff than the payoff of player i. However, because player i cannot sanction player j without damage, player i also gets injured a little. Player i, depending on the probability of $1/n(i)$ ($n(i) > 0$), decides whether he/she sanctions player j or not in ascending order regarding j (this reason is described later in this section). The function of this probability can be regarded as the heterogeneous sensitivity to the difference of payoff like Iwasa and Lee's graduated punishment [18]. Note that $n(i)$ indicates the total number of players in $O(i)$ that satisfies the condition $p(i) < p(j)$. We can express this scenario with the following Eq. (3) utilizing the payoff of sanctioning player i ($p(i)'$) and sanctioned player j ($p(j)'$). Note that r ($0 < r < 1$) is the coefficient of sanction, and of course player i cannot sanction player j when $p(i)(1 - rn(i)) < 0$. Therefore, $p(i)'$ and $p(j)'$ cannot be negative.

$$\begin{aligned} p(i)' &= p(i) - rp(i) \\ p(j)' &= p(j) - rp(i) \end{aligned} \tag{3}$$

The decrease in the payoff due to sanctioning the opponent and sanctioned by the opponent is calculated independently, and eventually $p(i)' = 0$ in the case of $p(i)' < 0$. After the payoff of all players changes due to all sanctioning and sanctioned activities, as the following Eq. (4), player i chooses the strategy of player $j_{\max}$ in $i \cup O(i)$ for the strategy of player i of the matches of the next generation. When more than one player has the same maximum payoff, player i randomly chooses the strategy of one of them. Each strategy of all players is synchronously updated.

$$\begin{aligned} s(i)' &= s(j_{\max})\ j_{\max} \in i \cup O(i) \\ p(j_{\max})' &= \max(p' \in i \cup O(i)) \end{aligned} \tag{4}$$

This study also utilizes the lattice of the scale-free topology of $N = 100$, $\bar{e} = 4$ (the average of $e(i)$) for the sensitivity analysis. Note that $\bar{e}$ is expressed as the following Eq. (5).

$$\bar{e} = \frac{1}{N} \sum_{1 \leq i \leq N} e(i) \tag{5}$$

The construction of the lattice of the scale-free topology follows the method by Barabási and Albert [19], i.e., starting with

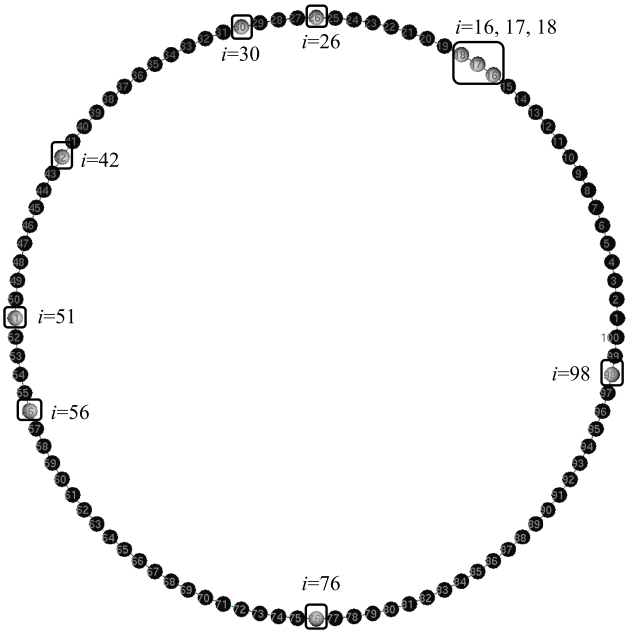

Fig. 1 This is the illustration of neighbor players around each player as the one-dimensional regular lattice [17]. The total number of players $N = 100$, and the total number of neighbor players around each player ($e(i)$) is the same among all players. In the initial state, 10% of all players enclosed in squares are defectors.

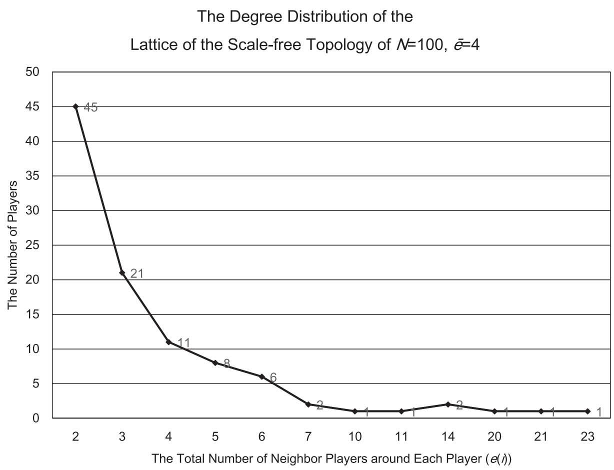

Fig. 2 This figure depicts the degree distribution that indicates how many players with a certain number of neighbor players around each player ($e(i)$) exist in the lattice of the scale-free topology when $N = 100$, $\bar{e} = 4$.

$m_0 = 5$ players of a complete graph, at every time step we add a new player with $m = 2$ edges that link the new player to m different players already present in the lattice until the total number of players (N) reaches 100. A new player will be connected to player u depending on the connectivity k_u of that player, so that $P_u(k_u) = k_u / \Sigma_v k_v$ (v: the number of players already present in the lattice). Therefore, players of small i turn into high degree players. Note that **Fig. 2** indicates the degree distribution of the constructed lattice of the scale-free topology that indicates how many players with a certain number of neighbor players around each player ($e(i)$) exist in the lattice. In order to sanction those high degree players with large payoff in the lattice of the scale-free topology, player i sanctions player j in $O(i)$ in ascending order regarding j with the probability $1/n(i)$ as noted before.

3. Results

This study employs three values of the average total number of neighbor players around each player ($\overline{e}$), i.e., $\overline{e} = 4$, 8 and 16 based on Santos and Pacheco's study [20]. The total number of players (N) is 100 because they describe that their results are robust even in the small community of 100 players. The parameter b in the payoff matrix A is set to 1.5. This is the medium value in the range of b ($1 < b \leq 2$) [16]. As mentioned before, 10% of all players are defectors, and others are cooperators in the initial state. The coefficient of sanction (r) is 0.15 because the result of $r = 0.15$ indicates the stable evolution of cooperation, i.e., standard deviation (SD) of the number of cooperators is the smallest value among the results of $r = 0.05$, 0.1, 0.15 and 0.2 in the case of $\overline{e} = 4$. The following results of the number of cooperators, the number of defectors, and the average payoff of all players up to 300 generations are the average of 20 independent simulation runs, and basically have error bars of SD.

As shown by Santos and Pacheco [20], when the initial ratio of the number of defectors to the number of cooperators equals one to one, a small b (≤ 1.175) is necessary for the evolution of cooperation (please refer to the following discussion as well). The aim of this study is to show that the sanction with jealousy has an effect on the evolution of cooperation in the case of large b where defectors are advantageous. To this end, the initial percentage of the number of defectors should have a small value. In addition, Masuda and Aihara [21] set the initial percentage of the number of defectors at 2, and show that the small-world topology is the optimal structure when we consider the speed at which cooperation evolves. Therefore, the initial percentage of the number of defectors of 10 in this study is not actually a small value when discussing the evolution of cooperation.

Firstly, the author shows both results with and without sanction regarding $\overline{e} = 4$ (**Fig. 3**). We find that in the case with sanction, the number of cooperators is quite large in comparison with the case without sanction. Of course, the average payoff of all play-

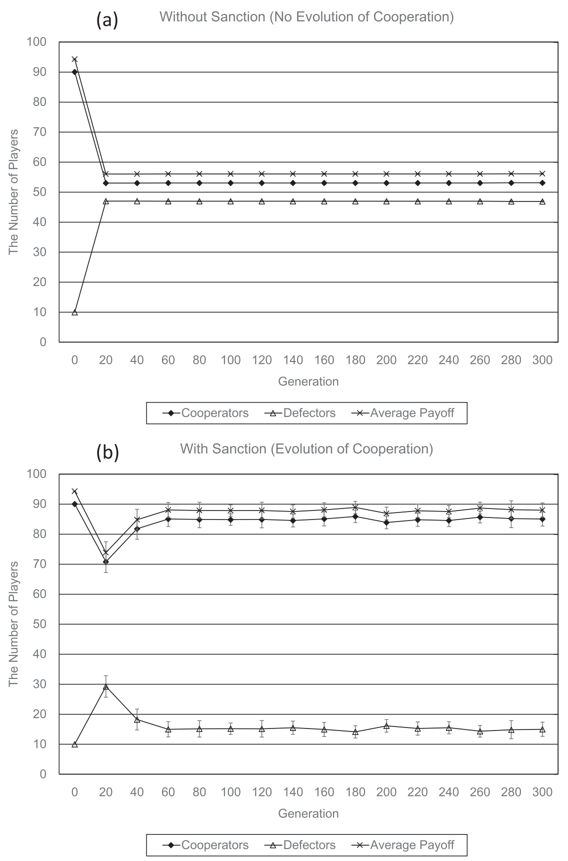

Fig. 3 This figure shows the number of cooperators, the number of defectors, and the average payoff of all players within 300 generations regarding both cases without sanction (a) and with sanction (b) when $\overline{e} = 4$. The average payoff of all players in the 300 generation is 2.24 in the case without sanction, while it is 3.52 in the case with sanction. Note that error bars are SD (standard deviation).

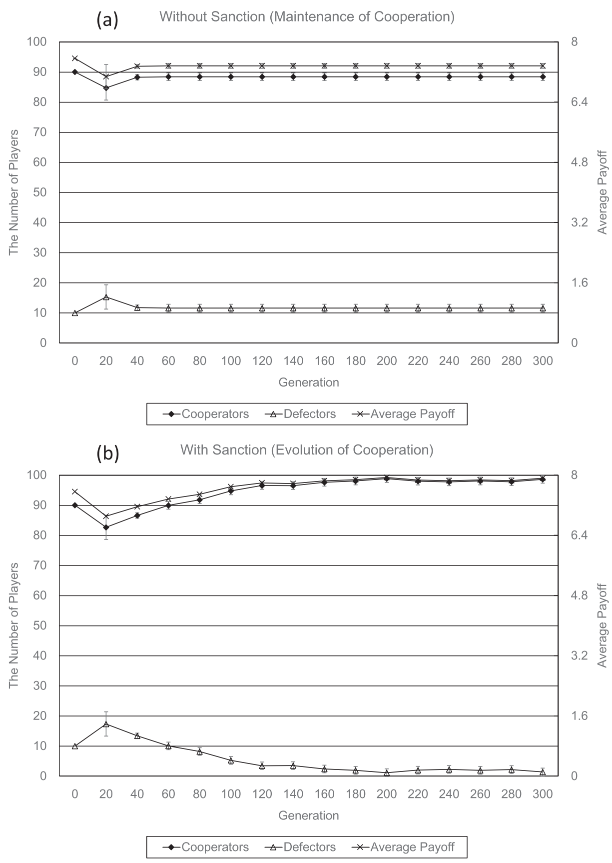

Fig. 4 This figure shows the number of cooperators, the number of defectors, and the average payoff of all players within 300 generations regarding both cases without sanction (a) and with sanction (b) when $\overline{e} = 8$. The average payoff of all players in the 300 generation is 7.36 in the case without sanction, while it is 7.92 in the case with sanction. Note that error bars are SD (standard deviation).

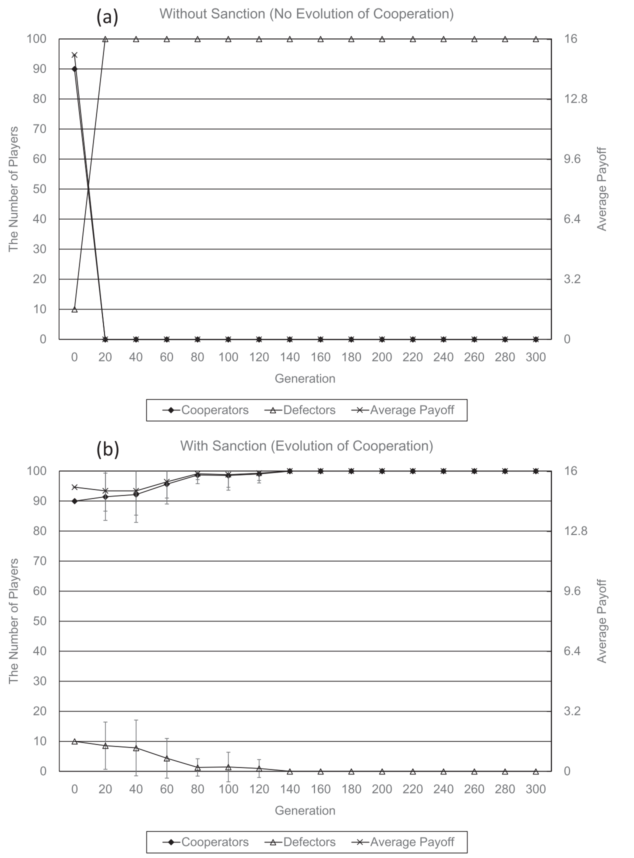

Fig. 5 This figure shows the number of cooperators, the number of defectors, and the average payoff of all players within 300 generations regarding both cases without sanction (a) and with sanction (b) when $\overline{e} = 16$. In the case with sanction, the average payoff of all players in the 300 generation is a lot larger than that without sanction in all simulation runs. Note that error bars are SD (standard deviation).

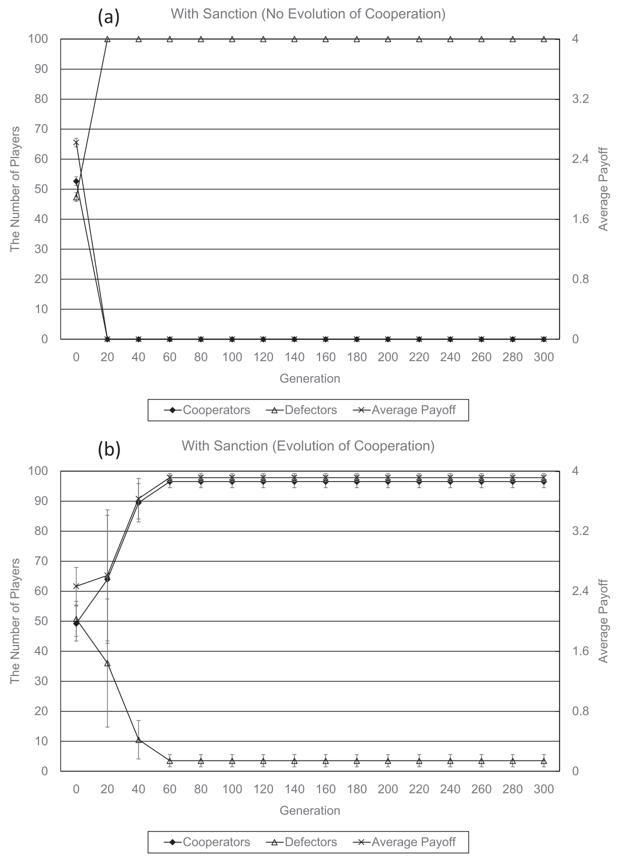

Fig. 6 This figure shows the number of cooperators, the number of defectors, and the average payoff of all players within 300 generations regarding the case with sanction when $N = 100$, $\overline{e} = 4$, and the initial percentage of the number of randomly scattered defectors is 50% for each simulation run. Graph (a) shows the case where the number of defectors increases (3/20 simulation runs), while Graph (b) shows the case where the number of cooperators increases (17/20 simulation runs). Note that error bars are SD (standard deviation).

ers in the 300 generation with sanction is higher than that without sanction (2.24 without sanction vs. 3.52 with sanction).

Secondly, **Fig. 4** indicates both results with and without sanction regarding $\overline{e} = 8$. In the case without sanction, each number of cooperators and defectors shows almost no change within 300 generations. On the other hand, in the case with sanction, the number of cooperators gradually increases and eventually reaches near 100 in the 300 generation, while the number of defectors gradually decreases and eventually reaches around 0 in the 300 generation. In addition, the average payoff of all players in the 300 generation with sanction is slightly large in comparison with that without sanction (7.36 without sanction vs. 7.92 with sanction).

Finally, **Fig. 5** exhibits both results with and without sanction regarding $\overline{e} = 16$. The result without sanction indicates that all players quickly become defectors. On the other hand, in the case with sanction, all players become cooperators in all simulation runs within 140 generations, and in this case, of course, the average payoff of all players in the 300 generation is a lot larger than that without sanction.

As for the sensitivity analysis, when $N = 100$, $\overline{e} = 4$, the initial percentage of the number of defectors is 50%, and those defectors are randomly scattered in the lattice for each simulation run, the averaged percentage of the number of defectors is finally under 4% in 17/20 simulation runs regarding the case with sanction (see **Fig. 6**).

When $N = 1000$, $\overline{e} = 4$, the initial percentage of the number of defectors is 10%, and those defectors are randomly scattered in the lattice for each simulation run (see **Fig. 7**), the time series change of the number of cooperators, the number of defectors, and the average payoff of all players indicates nearly the same trend as Fig. 3 ($N = 100$, $\overline{e} = 4$) regarding both cases without sanction and with sanction. This result means that the size of the lattice does not affect those values, and follows Santos and Pacheco's study [20] that shows the robustness of their results by exhibiting each result of $N = 1000$ (large) and $N = 100$ (small).

Figure 8 shows the result of the lattice of the scale-free topology [19] where $N = 100$, $\overline{e} = 4$, and the initial percentage of the number of defectors is 10%, i.e., the distribution of defectors is the same as Fig. 1. As Santos and Pacheco [20] show that the lattice of the scale-free topology has an effect on the emergence of cooperation, the result without sanction of the lattice of the scale-free topology maintains the number of cooperators and keeps the number of defectors low. However, the result with sanction of the lattice of the scale-free topology apparently exhibits a larger number of cooperators and fewer defectors in time series. Therefore, the introduction of the sanction with jealousy activates the effect of the lattice of the scale-free topology on the maintenance of the number of cooperators and the suppression of

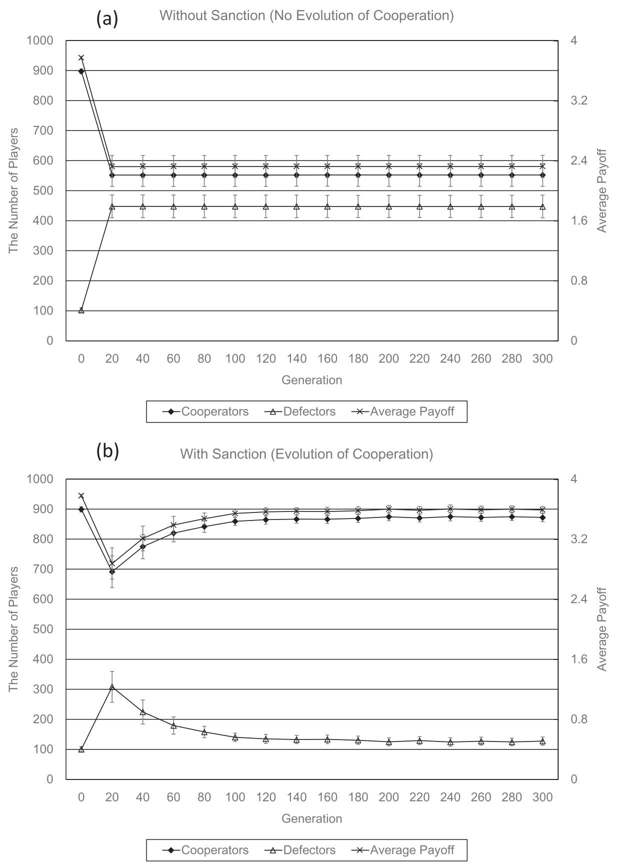

Fig. 7 This figure shows the number of cooperators, the number of defectors, and the average payoff of all players within 300 generations regarding both cases without sanction (a) and with sanction (b) when $N = 1000$, $\overline{e} = 4$, and the initial percentage of the number of randomly scattered defectors is 10% for each simulation run. Note that error bars are SD (standard deviation). The time series change of the number of cooperators, the number of defectors, and the average payoff of all players indicates nearly the same trend as Fig. 3 ($N = 100$, $\overline{e} = 4$).

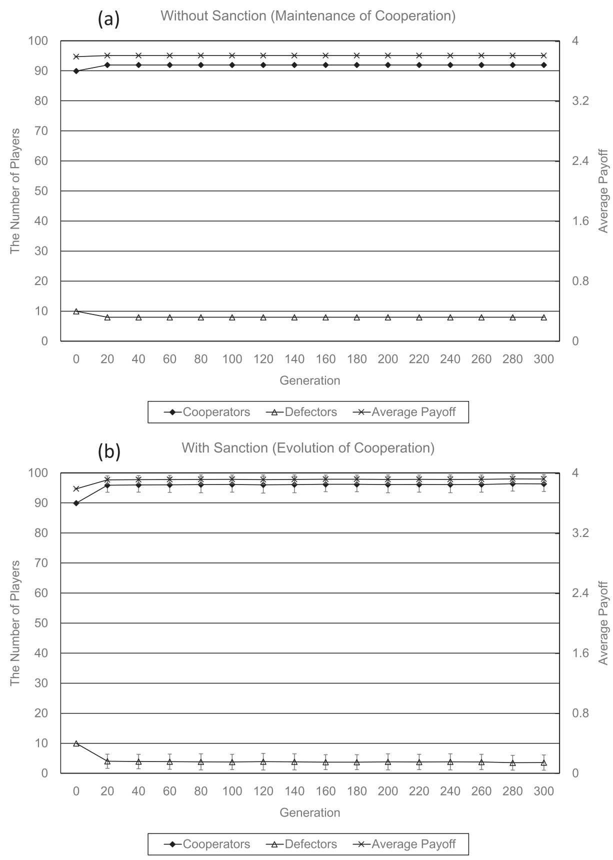

Fig. 8 This figure shows the number of cooperators, the number of defectors, and the average payoff of all players within 300 generations regarding both cases without sanction (a) and with sanction (b) when $N = 100$, $\overline{e} = 4$, the lattice is the scale-free topology, and the initial percentage of the number of defectors is 10%, i.e., the distribution of defectors is the same as Fig. 1. Note that error bars are SD (standard deviation). The result without sanction (a) maintains the number of cooperators and keeps the number of defectors low. However, the result with sanction (b) apparently exhibits a larger number of cooperators and a lower number of defectors in time series.

the number of defectors. In addition, when we utilize the completely random rewired regular lattice of $N = 100$, $\overline{e} = 4$, and the initial percentage of the number of randomly scattered defectors is 50% for each simulation run, defectors prevail in the lattice in 18/20 simulation runs in the case without sanction, while in the case with sanction, they become advantageous in 10/20 simulation runs. The introduction of the sanction with jealousy apparently suppresses the advantage of defectors. These results of the sensitivity analysis show that the sanction with jealousy tends to have a boosting effect on the evolution of cooperation in various parameters.

4. Discussion

As described in the introduction, Dreber et al. [2] propose that punishment causes the loss of the average payoff of all players. Rand and Nowak [5] also propose that punishment is an egocentric tool to protect a player itself because punishment promotes antisocial behavior through natural selection. Furthermore, Nowak [22] proposes that punishment is not a mechanism for the evolution of cooperation but enhances the level of cooperation emerging from other mechanisms (e.g., indirect reciprocity, group selection, or network reciprocity). As this study shows in the results, however, the sanction with jealousy suppresses an egocentric behavior and is effective in both the evolution of cooperation and the rise of the average payoff of all players. The results on the regular lattice in the case with sanction (Figs. 3 (b), 4 (b), and 5 (b)) indicate that the larger $\overline{e}$ induces a larger number of cooperators and a lower number of defectors. This is because the larger $\overline{e}$ leads to the higher payoff of each player, i.e., the stronger power of sanction of each player. The sanction with jealousy is the myopic punishment because it is effective in each player and their neighbor players, reduces the advantage of defectors, and smooths the difference of payoff among players.

In the case without sanction, the number of cooperators is approximately equal to the number of defectors in the case of $\overline{e} = 4$, the number of cooperators does not change too much in the case of $\overline{e} = 8$, and defectors apparently dominate the lattice in the case of $\overline{e} = 16$. This trend indicates that some increase in $\overline{e}$ enhances the construction of the clusters of cooperators that prevent the invasion of defectors, while the excessive increase in $\overline{e}$ induces the situation where defectors easily invade many more clusters of cooperators.

Contrary to the Santos and Pacheco's study [20], when $N = 100$, the initial ratio of defectors is 50%, and those defectors are randomly scattered in the lattice for each simulation run, in the case of $\bar{e} = 4$ and $b = 1.175$, the introduction of the sanction with jealousy leads almost all players to cooperators in 18/20 simulation runs. Furthermore, in the case of $\bar{e} = 16$ and $b = 1.1$, the sanction with jealousy realizes almost total cooperation in 14/20 simulation runs. These results also exhibit that the sanction with jealousy has an effect on the evolution of cooperation under the condition where defectors easily prevail in the lattice.

The reason why the author includes probabilistic factor in the sanction with jealousy is to avoid overpunishing that is not necessary to fix cooperation [23]. Overpunishing means that the payoff of defectors is reduced not by a fixed penalty, but by a sanction proportional to the number of punishers. In fact, in the case of $N = 100$, $\bar{e} = 4$, and the same initial percentage of the number of defectors as Fig. 1, when player i respectively sanctions player j in $O(i)$ that satisfies the condition $p(i) < p(j)$, i.e., each player sanctions all other related players, in the 300 generation, the average number of cooperators is 54, the average number of defectors is 46, and the average payoff of all players is 2.28 for 20 simulation runs. Those values are much smaller than each value indicated in Fig. 3 (b).

Of course, the author considers the improvement of the sanction with jealousy. For instance, when the difference of payoff between each player and their opponents is large, they sanction their opponents at the high probability. Furthermore, the author plans to consider the different coefficient r between sanctioning and sanctioned players. For example, it may be difficult for sanctioning players to reduce the high payoff of sanctioned players because they may have a kind of power due to their high payoff. In addition, it is necessary to investigate whether players who sanction with jealousy can survive through evolution (through "survival of the fittest") when others do not employ such behavioral patterns.

5. Conclusion

This study presents both the negative [1], [2], [3], [4], [5] and the positive [6], [7], [8], [9], [10], [11], [12], [13], [14] discussions regarding the effectiveness of punishment. Contrary to these previous discussions, the author proposes an alternative notion of punishment "sanction with jealousy" that these previous discussions do not cover. The results of this study reveal that not only the sanction with jealousy facilitates cooperation, but also it solves the flaw of punishment that the discussions of negative point of view [2], [5] designate. In the future works, the author continues to improve the sanction with jealousy as noted in the discussion.

Acknowledgments The author truly thanks Prof. Takao Terano for giving the author the opportunity to discuss this study, and the anonymous referees for their constructive comments and suggestions.

References

[1] Sigmund, K., Hauert, C. and Nowak, M.A.: Reward and punishment, *Proc. Natl. Acad. Sci. U.S.A.*, Vol.98, pp.10757–10762 (2001).

[2] Dreber, A., Rand, D.G., Fudenberg, D. and Nowak, M.A.: Winners don't punish, *Nature*, Vol.452, pp.348–351 (2008).

[3] Herrmann, B., Thöni, C. and Gächter, S.: Antisocial punishment across societies, *Science*, Vol.319, pp.1362–1367 (2008).

[4] Ohtsuki, H., Iwasa, Y. and Nowak, M.A.: Indirect reciprocity provides only a narrow margin of efficiency for costly punishment, *Nature*, Vol.457, pp.79–82 (2009).

[5] Rand, D.G. and Nowak, M.A.: The evolution of antisocial punishment in optional public goods games, *Nature Communications*, Vol.2, 434 (2011).

[6] Fehr, E. and Gächter, S.: Altruistic punishment in humans, *Nature*, Vol.415, pp.137–140 (2002).

[7] Fehr, E. and Fischbacher, U.: The nature of human altruism, *Nature*, Vol.425, pp.785–791 (2003).

[8] Fowler, J.H.: Altruistic punishment and the origin of cooperation, *Proc. Natl. Acad. Sci. U.S.A.*, Vol.102, pp.7047–7049 (2005).

[9] O'Gorman, R., Henrich, J. and Van Vugt, M.: Constraining free riding in public goods games: Designated solitary punishers can sustain human cooperation, *Proc. R. Soc. B*, Vol.276, pp.323–329 (2009).

[10] Rankin, D.J., Santos, M.D. and Wedekind, C.: The evolutionary significance of costly punishment is still to be demonstrated, *Proc. Natl. Acad. Sci. U.S.A.*, Vol.106, E135 (2009).

[11] Sigmund, K., Silva, H.D., Traulsen, A. and Hauert, C.: Social learning promotes institutions for governing the commons, *Nature*, Vol.466, pp.861–863 (2010).

[12] Szolnoki, A., Szabó, G. and Czakó, L.: Competition of individual and institutional punishments in spatial public goods games, *Phys. Rev. E*, Vol.84, 046106 (2011).

[13] Garcia, J. and Traulsen, A.: Leaving the loners alone: Evolution of cooperation in the presence of antisocial punishment, *J. Theor. Biol.*, Vol.307, pp.168–173 (2012).

[14] Perc, M. and Szolnoki, A.: Self-organization of punishment in structured populations, *New J. Phys.*, Vol.14, 043013 (2012).

[15] Fehr, E. and Schmidt, K.M.: A theory of fairness, competition, and cooperation, *Quart. Jour. Econ.*, Vol.114, pp.817–868 (1999).

[16] Nowak, M.A. and May, R.M.: Evolutionary games and spatial chaos, *Nature*, Vol.359, pp.826–829 (1992).

[17] Watts, D.J. and Strogatz, S.H.: Collective dynamics of 'small-world' networks, *Nature*, Vol.393, pp.440–442 (1998).

[18] Iwasa, Y. and Lee, J.-H.: Graduated punishment is efficient in resource management if people are heterogeneous, *J. Theor. Biol.*, Vol.333, pp.117–125 (2013).

[19] Barabási, A.L. and Albert, R.: Emergence of scaling in random networks, *Science*, Vol.286, pp.509–512 (1999).

[20] Santos, F.C. and Pacheco, J.M.: Scale-free networks provide a unifying framework for the emergence of cooperation, *Phys. Rev. Lett.*, Vol.95, 098104 (2005).

[21] Masuda, N. and Aihara, K.: Spatial prisoner's dilemma optimally played in small-world networks, *Phys. Lett. A*, Vol.313, pp.55–61 (2003).

[22] Nowak, M.A.: Five rules for the evolution of cooperation, *Science*, Vol.314, pp.1560–1563 (2006).

[23] Dercole, F., De Carli, M., Della Rossa, F. and Papadopoulos, A.V.: Overpunishing is not necessary to fix cooperation in voluntary public goods games, *J. Theor. Biol.*, Vol.326, pp.70–81 (2013).

Tetsushi Ohdaira (M.E., University of Tokyo, 2003, Ph.D. in Media and Governance, Keio University, 2012) is an assistant professor at the Institute of Information and Media, Aoyama Gakuin University. His interest includes cooperation and altruistic behavior in conflicts of interest, especially in the prisoner's dilemma game. He attempts to approach this issue utilizing the agent-based simulation. He is also interested in information visualization especially regarding the network science.

