Network reciprocity by coexisting learning and teaching strategies

Jun Tanimoto,[1] Markus Brede,[2] and Atsuo Yamauchi[1]
[1]*Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Kasuga-koen, Kasuga-shi, Fukuoka, 816-8580, Japan*
[2]*Commonwealth Scientific and Industrial Research Organisation, Marine and Atmospheric Research, Commonwealth Scientific and Industrial Research Organisation, Center for Complex System Science, F C Pye Laboratory, GPO Box 3023, Clunies Ross 2601, Australia*
(Received 27 July 2011; revised manuscript received 6 February 2012; published 21 March 2012)

We propose a network reciprocity model in which an agent probabilistically adopts learning or teaching strategies. In the learning adaptation mechanism, an agent may copy a neighbor's strategy through Fermi pairwise comparison. The teaching adaptation mechanism involves an agent imposing its strategy on a neighbor. Our simulations reveal that the reciprocity is significantly affected by the frequency with which learning and teaching agents coexist in a network and by the structure of the network itself.

DOI: 10.1103/PhysRevE.85.032101 PACS number(s): 02.50.Le, 89.65.−s, 87.23.Kg

I. INTRODUCTION

In most previous network game models, strategy adaptation is a process where a focal agent copies a neighbor's strategy, i.e., either cooperates (C) or defects (D). Copying is akin to learning a neighbor's methods. Szolnoki and Szabo [1] proposed a strategy adaptation called teaching in which an agent tries to impose its C or D strategy on a neighbor. This then inspired scholars to investigate learning as well as teaching [1–9]. In human social systems, learning and teaching are common activities that enable adaptation strategies to spread throughout a population. Note that a situation in which everyone unanimously adopts either learning or teaching is unrealistic. In addition, sometimes people learn from others and try to impose what they learned on others.

This study investigates the effects of learning and teaching agents coexisting on a network involving a prisoner dilemma (PD) game and an agent chooses its pairwise opponent in a probabilistically skewed manner [10] instead of in a random fashion.

II. MODEL CONSTRUCTION

At every time step, agents on a time-constant network play PD games with their neighbors and synchronously update their strategy based on their accumulated payoffs. In this model we regard accumulated payoff as a fitness for strategy adaptation instead of payoff per an interaction [11]. In a PD game, a player receives a reward R for mutual C's and a punishment P for mutual D's. If one player chooses C and the other chooses D, the latter obtains a temptation payoff T and the former is labeled a saint S. Without loss of generality, we can define a PD game space by presuming $R = 1$ and $P = 0$ as follows:

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix} = \begin{pmatrix} 1 & -D_r \\ 1 + D_g & 0 \end{pmatrix}, \tag{1}$$

where $D_g = T - R$ and $D_r = P - S$ imply a chicken-type and stag-hunt dilemma, respectively [12]. We limit the PD game class by assuming $0 \leqslant D_g \leqslant 1$ and $0 \leqslant D_r \leqslant 1$.

When updating its strategy, an agent selects an opponent for pairwise comparison by random selection or skewed selection. In the latter method, focal agent x selects one neighbor y based on the following probability [10]:

$$p_y^{\mathrm{PW\ for}\ x} = \frac{\exp(w\Pi_y)}{\sum_{i\in\{\Omega_x\}} \exp(w\Pi_i)}, \tag{2}$$

where Ω_x is the focal agent's neighbor set and Π is the accumulated payoff of each agent. The parameter w implies that a positive (negative) value favors a situation where a neighbor with a relatively large (small) payoff would be selected. After choosing an opponent for pairwise comparison, the focal agent adopts either a learning or a teaching strategy, which are defined in the following.

A. Learning model

Focal agent x may or may not copy the strategy of a pairwise opponent y based on a Fermi pairwise process. The decision is based on the difference in payoff between its strategy and its neighbor's strategy:

$$p_{\mathrm{copy}}^{x\leftarrow y} = \frac{1}{1+\exp\left(\frac{\Pi_x-\Pi_y}{\kappa}\right)}.$$

Again, we emphasize that learning is a common concept [13,14].

B. Teaching model

Focal agent x may or may not choose to impose its strategy on a pairwise opponent y based on a Fermi pairwise process. The decision is based on the payoff difference

$$p_{\mathrm{impose}}^{x\rightarrow y} = \frac{1}{1+\exp\left(\frac{-\Pi_x+\Pi_y}{\kappa}\right)}.$$

If several neighbors try to impose their strategy on a certain agent simultaneously, each imposing trial that is randomly ordered among those teaching neighbors would be considered one after another. We assume that $\kappa = 0.2$ throughout our simulations, based on previous studies [15].

At the beginning of each simulation episode, we randomly designate half of the total number of agents N on a network as cooperative and half as defectors. We assume that $1 - \lambda$ of the total agents adopt learning and λ adopt teaching. At each time step we randomly determine who adopts learning or teaching to meet the coexisting frequency λ. Accordingly, an agent's adaptation choice, whether it is learning or teaching, is never fixed. When skewed pairwise opponent selection is assumed, we assign $w = 1$ for learning and $w = -1$ for teaching agents.

We presumed synchronous strategy updating even if learning and teaching agents coexist. Assume agent A has teaching neighbors B, C, and D. When all three neighbors attempt to impose their strategies on A simultaneously, A copies from the last agent among B, C, and D, which are randomly ordered. At first, A is imposed on to copy D's strategy. Next, C tries to impose its strategy on A. Finally, A adopts the strategy of B, who is the last agent in the random order. Assume another example in which learner A has learner B and teacher C as neighbors. If A wants to learn from B and C wants to teach A, then A will experience the later event: learning from B and thrusting by C, which are randomly ordered.

Each simulation runs as follows. Initially, an equal percentage of strategies is distributed randomly among the players allocated on different network vertices. Then several generations are run until the frequency of cooperation reaches a certain quasiequilibrium where the difference between the average cooperation fraction for the last 100 generations and that for the previous 100 generations becomes less than 1% of the average cooperation fraction for the last 100 generations. If the frequency of cooperation continues fluctuating, we obtain the frequency of cooperation for the last 100 generations over a 6000-generation run. We conduct this procedure at 11×11 points of the PD area ($0 \leqslant D_g \leqslant 1$ and $0 \leqslant D_r \leqslant 1$) 300 times to draw those 121 ($=11 \times 11$) ensemble averages. We investigate two types of network structures: the Barabási-Albert (BA) scale-free (SF) network [16] and the two-dimensional (2D) lattice with the assumptions $N = 4900$ and $\langle k \rangle = 8$. To run 300 realizations, we prepared 20 BA SF networks derived from 20 random seeds. This implies that each BA SF network runs 15 realizations starting from different randomly distributed initial allocations. Concerning the underlying network topology, we adopted the 2D lattice and BA SF network as representative graphs for the homogeneous and heterogeneous networks, although the graph should be chosen more carefully [17]. For example, networks with a different assortment of coefficients might be interesting to investigate despite having the same degree distribution.

III. RESULTS AND DISCUSSION

Figure 1 shows cooperative fractions with different teaching frequencies covering an entire PD area with an assumed BA SF network. The cooperation decays as the number of teaching agents increases, while skewed pairwise opponent selection substantially enhances cooperation.

We specifically define three subclasses of PD to evaluate cooperation within $0 \leqslant D_g \leqslant 1$ and $0 \leqslant D_r \leqslant 1$. The first is the average cooperation fraction covering the entire PD area (APD) and derived from the data of all 121 points. The second is the average of game structures featuring $D_g = D_r$, which is the donor and recipient (DR) game, derived from the 11 point average. The last is the average collected from the region of $D_r = 0$, which consists of boundary games between the PD and chicken games without the stag-hunt dilemma (BPDC), also derived from the 11 point average. Many previous studies postulate the BPDC for representing the PD because it can be characterized by the single dilemma parameter D_g. However,

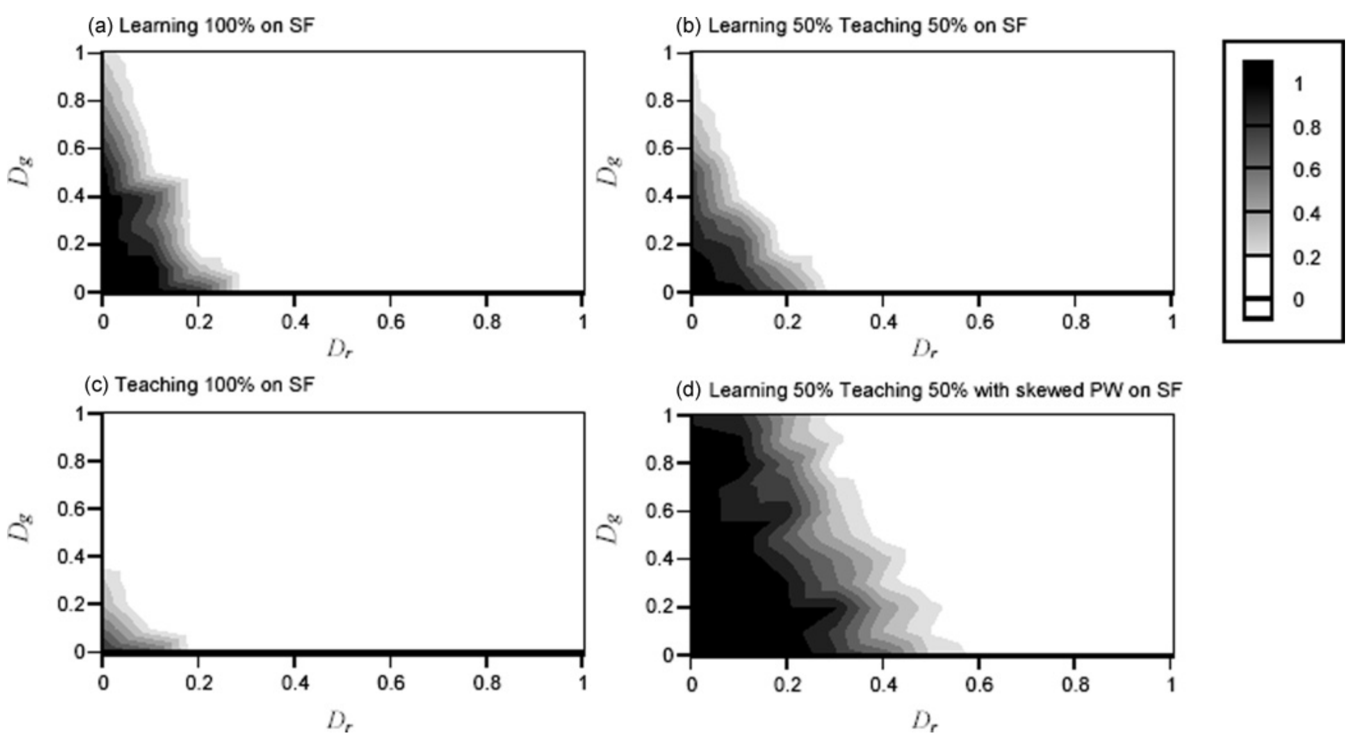

FIG. 1. Average cooperation fraction within $0 \leqslant D_r \leqslant 1$ and $0 \leqslant D_g \leqslant 1$ obtained from 100 independent simulation episodes. We assumed $N = 4900$ and BA SF networks with an average degree of $\langle k \rangle = 8$. Results with (a) $\lambda = 0$, (b) $\lambda = 0.5$, and (c) $\lambda = 1$ based on random pairwise opponent selection. (d)$\lambda = 0.5$ with skewed pairwise opponent selection. (a) Customary network reciprocity on SF networks with $\langle k \rangle = 8$ is implied, when Fermi pairwise synchronous updating is assumed [16].

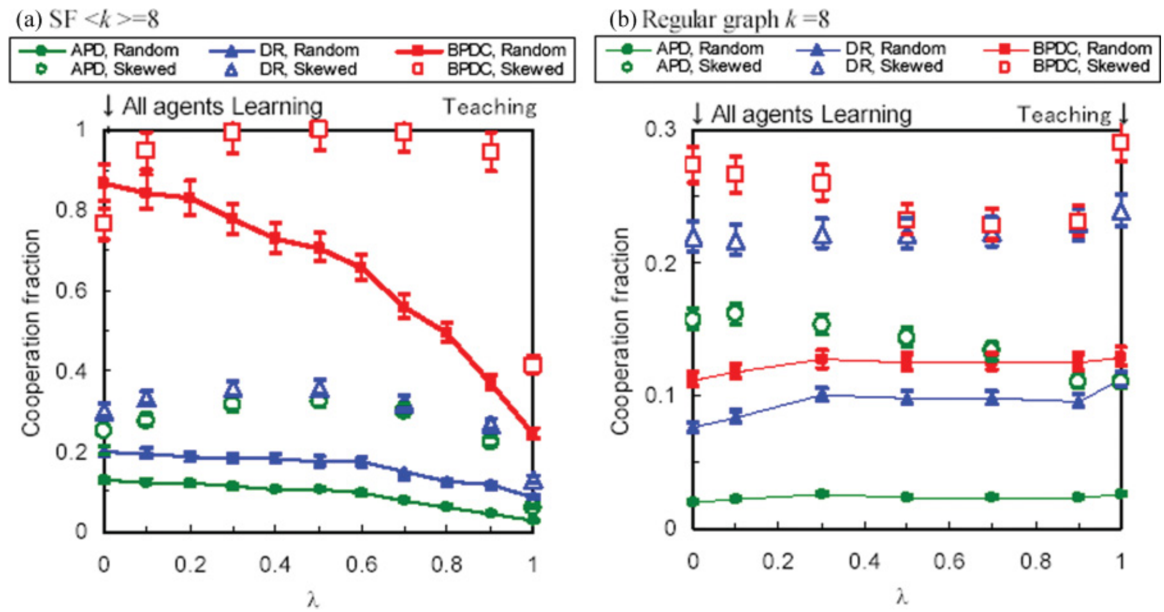

FIG. 2. (Color online) Average cooperation fractions of APD, DR, and BPDC games with random (closed symbols) and skewed (open symbols) pairwise opponent selection on SF networks assuming (a) $\langle k \rangle = 8$ and $N = 4900$ and (b) regular network with $k = 8$ and $N = 4900$, obtained from 300 realizations. Error bars indicate standard deviation of 300 realizations.

as indicated by Yamauchi *et al*. [18], the network reciprocity [19] of SF networks would be overestimated if we consider only the cooperation fraction for the BPDC. Therefore, we show the average cooperation fractions of both the APD and DR subclasses as well as the BPDC subclass in Fig. 2. Those three average cooperation fractions can, at a glance, give a holistic evaluation as indices of how each of the assumed models enhances cooperation, covering a range of different dilemma strengths. Several physicists have used r_c for this purpose, which means a threshold dilemma strength absorbed by the entire defection phase when increasing the dilemma strength. We rely on the average cooperation fraction this time because r_c for the BPDC cannot be observed in $0 \leqslant D_g \leqslant 1$, as shown in Fig. 1(d).

With regard to the results of random pairwise selection, it is noteworthy that cooperation in the SF case degrades rapidly as λ increases, whereas on a regular network, cooperation shows a slightly opposite tendency. The underlying reason might be that when a high teaching frequency is assumed for a SF network, the hub agents, having a higher degree than others, are more exposed to a neighbor's strategy. A higher degree implies more frequent opportunities to be imposed on by neighbors. Thus, if this hub agent is a hub-C agent and suffers the imposition of D instead of C, the C cluster it leads could be damaged. However, on the regular graph, this drawback for the teaching adaptation can work in the opposite direction. Let us assume the two extreme cases of $\lambda = 0$ (all agents adopt learning) and $\lambda = 1$ (all agents adopt teaching). In the former case, every agent has a chance to update its strategy at each time step, irrespective of whether it actually does so. However, in the latter case, a considerable number of agents might have no chance to update their strategy because other agents are imposed on by several neighbors simultaneously. This fact enhances cooperation only on homogeneous networks.

Skewed pairwise opponent selection enhances more cooperation than random pairwise opponent selection, irrespective of the underlying network topology. Again, we should remember that when random pairwise opponent selection occurs on a SF network, the $\lambda = 0$ case is superior to the $\lambda = 0.5$ case. Nevertheless, in the case of skewed pairwise opponent selection, the most cooperative situation is realized when one half of the agents adopt teaching and the other half adopt learning. One possible reason for this is that the assumption of $w = 1$ for learning agents and $w = -1$ for teaching agents stimulates cooperation because both imitate stronger agents and killing (overwriting) weaker agents seems to diffuse cooperation among defectors. Alternatively, this can be justified as follows. In the case of $\lambda = 0$ learning agents always tend to learn from hub agents and in the case of $\lambda = 1$ teaching agents tend to teach lower-degree agents. This amplifies cooperation up to some λ. If λ becomes too large it becomes hard to remove defectors who have been initially allocated to hub nodes.

To confirm our hypothesis, we observed the initial 100 time steps in a specific simulation episode to examine whether the game's early evolutionary stages crucially affect whether cooperators survive, perish, or coexist with defectors. Those episodes shown Fig. 3 are typical among 300 realizations except for episodes showing an all-defector equilibrium. When a dilemma is imposed on any simulation that starts with half the agents who offer C being dispersed randomly on a network, the initial C clusters are immediately invaded by defectors. Therefore, the cooperation fraction declines. If network reciprocity were effective, cooperators who form C clusters could endure this initial D invasion and thereafter would be able to extend C clusters among defectors. Figure 3 shows a time series for the following: (i) the fraction of agents whose strategy is not updated (they are teachers and their teaching neighbors never select those focal teachers to

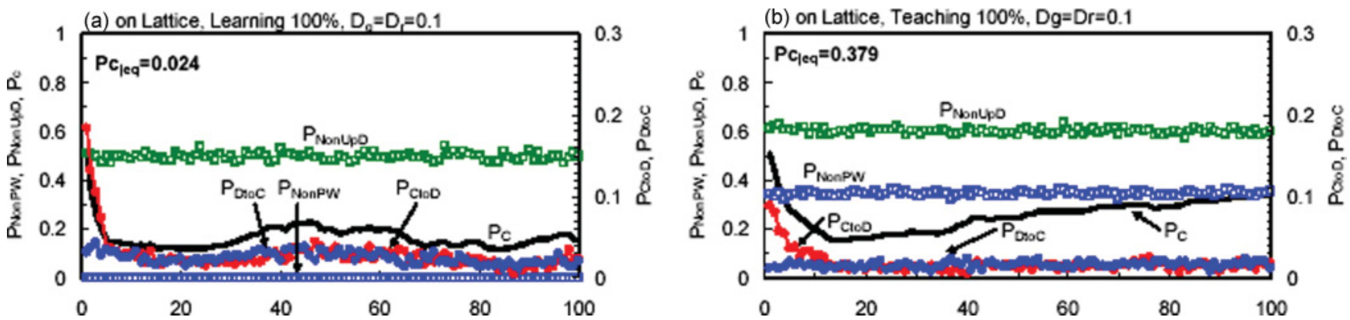

FIG. 3. (Color online) Time series reflecting (i) the agent fraction that does not update (no learning and no teaching by neighbors) ($P_{\text{non-PW}}$), (ii) the agent fraction that retains its strategy (trying to learn or be taught but not adopting the neighbor's strategy) ($P_{\text{non-up D}}$), (iii) the agent fraction that changes strategy from D to C ($P_{\text{D to C}}$), (iv) the agent fraction that switches strategy from C to D ($P_{\text{C to D}}$), and (v) the cooperation fraction during the initial 100 time steps of a specific simulation episode when $D_g = D_r = 0.1$ (DR game). In (a) all agents adopt learning and in (b) all agents adopt teaching on a regular network ($k = 8$) with random pairwise opponent selection. The cooperation fraction at equilibrium for each episode is denoted by $P_{\text{c|eq}}$.

impose on) ($P_{\text{non-PW}}$), (ii) the fraction of agents who retain their previous strategy [i.e., (a) if the focal agent is a learner, it decides not to copy from one of its neighbors and (b) if it has teaching neighbors, they select the focal agent as an opponent to impose on, but can fail to impose on it because of the pairwise comparisons] ($P_{\text{non-up D}}$), (iii) the fraction of agents who switch strategy from D to C ($P_{\text{D to C}}$), (iv) the agent fraction that switches from C to D ($P_{\text{C to D}}$), and (v) the cooperation fraction (P_{c}) when all agents adopt learning and all agents adopt teaching on a regular network. Strictly speaking, there are two more groups other than (i)–(iv), i.e., $P_{\text{C to C}}$ and $P_{\text{D to D}}$. Both of these are the fractions of agents who copy their own strategy (C or D) from one of their neighbors through either learning or teaching. Thus the total summation of $P_{\text{non-PW}}$, $P_{\text{non-up D}}$, $P_{\text{D to C}}$, $P_{\text{C to D}}$, $P_{\text{D to D}}$, and $P_{\text{C to C}}$ is equal to 1. Obviously, the 100% teaching case has a larger $P_{\text{non-PW}}$ and $P_{\text{non-up D}}$ than the 100% learning case. In other words, teaching could retard the spread of *a strategy throughout the population* by increasing the number of agents that are insensitive to strategy updating. This is consistent with the fact that the 100% teaching case has fewer agents who revise their current strategy than the 100% learning case. This situation prevents defectors from invading C clusters. To that end, this particular episode of 100% teaching realizes greater cooperative equilibrium than the 100% learning episode. In summary, larger teaching frequencies decrease the speed of a strategy spread in the initial time steps, thus making it easier for clusters of cooperators to survive the initial onslaught of defectors.

Although the figures are not shown, we confirmed that how slowly a defection spreads on a SF network is no longer attributable to a higher cooperation at equilibrium. On a SF network, the defectors' invasion of C clusters is not so severe *vis à vis* the regular network case because more effective network reciprocity is derived from its heterogeneity. Therefore, the cooperation fraction is determined by how quickly cooperation spreads among surrounding defectors after the initial defectors' invasion.

IV. CONCLUSION

The extent of network reciprocity differs greatly, depending on the frequency with which learning or teaching is adopted and on the underlying network structure. On heterogeneous networks, teaching tends to degrade cooperation because an instance in which plural neighbors try to overwrite a hub agent's strategy implies that a hub-C agent rarely survives the defectors' initial attack. However, on regular networks, teaching slightly enhances cooperation.

[1] A. Szolnoki and G. Szabo, EPL **77**, 30004 (2007).
[2] A. Szolnoki and M. Perc, New J. Phys. **10**, 043036 (2008).
[3] M. Perc, A. Szolnoki, and G. Szabo, Phys. Rev. E **78**, 066101 (2008).
[4] G. Szabo and A. Szolnoki, Phys. Rev. E **79**, 016106 (2009).
[5] A. Szolnoki and M. Perc, Eur. Phys. J. B **67**, 337 (2009).
[6] A. Szolnoki, M. Perc, G. Szabo, and H.-U. Stark, Phys. Rev. E **80**, 021901 (2009).
[7] M. Droz, M. Szwabinski, and G. Szabo, Eur. Phys. J. B **71**, 589 (2009).
[8] G. Zschaler, A. Traulsen, and T. Gross, New J. Phys. **12**, 093015 (2010).
[9] A. Szolnoki, Z. Wang, J. Wang, and X. Zhu, Phys. Rev. E **82**, 036110 (2010).
[10] Z. Wang and M. Perc, Phys. Rev. E **82**, 021115 (2010).
[11] S. Weber and M. Porto, Eur. Phys. J. B **69**, 599 (2009).
[12] J. Tanimoto and H. Sagara, BioSystems **90**, 105 (2007).
[13] G. Szabo and C. Toke, Phys. Rev. E **58**, 69 (1998).
[14] C. Hauert and G. Szabó, Am. J. Phys. **73**, 405 (2005).
[15] J. Ren, W. X. Wang, and F. Qi, Phys. Rev. E **75**, 045101(R) (2007).
[16] A. L. Barabási and R. Albert, Science **286**, 509 (1999).
[17] A. Pusch, S. Weber, and M. Porto, Phys. Rev. E **77**, 036120 (2008).
[18] A. Yamauchi, J. Tanimoto, and A. Hagishima, BioSystems **103**, 85 (2011).
[19] M. A. Nowak, Science **314**, 1560 (2006).