Journal of Theoretical Biology 361 (2014) 111–123

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Selfish punishment with avoiding mechanism can alleviate both first-order and second-order social dilemma

Pengbi Cui [a,b], Zhi-Xi Wu [a,b,*]

[a] Institute of Computational Physics and Complex Systems, Lanzhou University, Lanzhou 730000, China
[b] Key Laboratory for Magnetism and Magnetic Materials of the Ministry of Education, Lanzhou University, Lanzhou 730000, China

HIGHLIGHTS

- We introduce altruistic and selfish punishment to evolutionary prisoner's dilemma game.
- We derive theoretical predictions by an extended pair approximation method.
- The co-existing strategy phases are analyzed by means of interaction webs.
- Selfish punishment can alleviate both first- and second-order social dilemma.

ARTICLE INFO

Article history:
Received 9 April 2014
Received in revised form
26 June 2014
Accepted 19 July 2014
Available online 1 August 2014

Keywords:
Evolutionary game theory
Punishment
Identifying probability
Pair approximation

ABSTRACT

Punishment, especially selfish punishment, has recently been identified as a potent promoter in sustaining or even enhancing the cooperation among unrelated individuals. However, without other key mechanisms, the first-order social dilemma and second-order social dilemma are still two enduring conundrums in biology and the social sciences even with the presence of punishment. In the present study, we investigate a spatial evolutionary four-strategy prisoner's dilemma game model with avoiding mechanism, where the four strategies are cooperation, defection, altruistic and selfish punishment. By introducing the low level of random mutation of strategies, we demonstrate that the presence of selfish punishment with avoiding mechanism can alleviate the two kinds of social dilemmas for various parametrizations. In addition, we propose an extended pair approximation method, whose solutions can essentially estimate the dynamical behaviors and final evolutionary frequencies of the four strategies. At last, considering the analogy between our model and the classical Lotka-Volterra system, we introduce interaction webs based on the spatial replicator dynamics and the transformed payoff matrix to qualitatively characterize the emergent co-exist strategy phases, and its validity are supported by extensive simulations.

1. Introduction

For centuries, the scales of most of human economic communities have expanded dramatically from kin-based work-shops to intensively large-scale cooperative groups in which selfish individuals frequently cooperate with other genetically unrelated ones. Kinds of mechanisms or rules have been developed to explain or support the existence of such cooperative behaviors (Nowak, 2006). In these range of rules, kin selection is merely applied to small kinship groups (Hamilton, 1971; Foster et al., 2006). Direct reciprocity (Axelrod, 1984) can explain the emergence of cooperation between unrelated individuals or even between members of different species, but it is limited to the repeated encounters between the same two individuals. In the context of indirect reciprocity (Nowak, 2006; Nowak and Sigmund, 1998b, 1998a), randomly chosen pairwise encounters where two individuals do not have to meet again are admissible. However, it can only promote cooperation on condition with sufficient reputation (Alexander, 1987; Haley and Fessler, 2005) that drives this deed. Additionally, network reciprocity (Lieberman et al., 2005; Nowak, 2006; Ohtsuki et al., 2006) and social diversity (Santos et al., 2008) are only established in the population that is not well-mixed, i.e., their operations rely heavily on the hierarchical structure of populations. The significance of migration for the emergence and persistence of cooperation has also been highlighted by the previous studies (Helbing and Yu, 2009; Yang et al., 2010). Nevertheless, in real life, the cost of migration may be very high, and the information about the destinations may also

* Corresponding author at: Institute of Computational Physics and Complex Systems, Lanzhou University, Lanzhou 730000, China.
E-mail address: wuzhx@lzu.edu.cn (Z.-X. Wu).

http://dx.doi.org/10.1016/j.jtbi.2014.07.021
0022-5193/

be insufficient and limited (Borjas, 1989; Drinkwater et al., 2003; Martin, 2012; Buesser et al., 2013). For these reasons, punishment turns out to be a key role in sustaining the cooperation as strangers frequently engage in interest transactions in large-scale institutions (Boyd and Richerson, 1992; Panchanathan and Boyd, 2004).

Furthermore, altruistic punishment has been used as a paradigm to promote cooperation in large populations consisting of selfish unrelated or faceless individuals (Fehr and Gächter, 2002; Fowler and Christakis, 2010). However, it is less likely to become a robust strategy (Perc, 2012) owing to the extra expenses for the cost to punish defectors. Only recently, a few works have suggested another way that altruism may be maintained by the defectors though punishing other defectors, known as selfish punishment (Wilson and O'Gorman, 2006; Nakamaru and Iwasa, 2006; Eldakar et al., 2007). The concept of selfish punishment was originally suggested by an empirical experiment on humans demonstrating that individuals most likely to punish other defectors themselves are most tempted to defect (Wilson and O'Gorman, 2006). This experiment actually implies that a certain part of defectors prefer to punish other defectors for themselves rather the public welfares in some situations. Combing with reality, it is possible that certain groups of individuals, especially tricky cheaters, can take some steps such as lying to punishers to avoid the sanctions in the presence of communication (Serra-Garcia et al., 2013; Sheremeta and Shields, 2013; Belot et al., 2012), exemplified by the proverb ‘a thief crying “Stop thief”’. These cheaters use punishment as an evasion, which can be regarded as the alternative form of selfish punishment. However, to our knowledge, this actual important mechanism has received relatively little attention in evolutionary game theory. Further studies are still necessary.

In order to further explore how this selfish punishment works, we design a model involving cooperators (C), defectors (D), cooperative punishers (CP, i.e., altruistic punishers), and defective defectors (DP, i.e., selfish punishers) with avoiding mechanism. Differing from the previous models (Helbing et al., 2010a, 2010c; Szolnoki et al., 2011; Nakamaru and Iwasa, 2006; Panchanathan and Boyd, 2004; Fehr and Gächter, 2002; Fowler and Christakis, 2010) with respect to punishment, our model is performed in the context of prisoner's dilemma game (PDG) along with a low level of random mutation. In detail, the sanctions from punishment are always considered to be costly (Helbing et al., 2010a, 2010c; Szolnoki et al., 2011). Similarly, both defective punishers and cooperative punishers sanction defectors with a punishment fine at a personal punishment cost in our model, without loss of generality and rationality. Unlike previous studies on the evolution of altruism with punishment (Helbing et al., 2010a, 2010c; Szolnoki et al., 2011; Nakamaru and Iwasa, 2006; Panchanathan and Boyd, 2004; Fehr and Gächter, 2002; Fowler and Christakis, 2010), we add the avoiding behaviors represented by the identifying probability to defective punishers, who punish not only other defectors but also the ones with the same strategy. Moreover, we propose an extended pair approximation for the time evolution of the four strategies, allowing us to track the dynamics features and stationary states of the system. At last, motivated by the works on ecological interaction networks (Knebel et al., 2013), we introduce interaction webs to qualitatively understand the stable coexistence and extinction of different strategies.

It is worth noting that there are two kinds of social dilemmas in the model. One is the conventional social dilemma – PDG, namely the first-order social dilemma in which the free riders such as defectors can earn higher personal profits than cooperators whereas the well-being of the population depend only on the level of cooperation. The other is the second-order social dilemma, where punishers carry out punishment which reduces their fitnesses relative to those second-order free riders (including pure cooperators) who do not punish (Nakamaru and Iwasa, 2006; Fehr and Rockenbach, 2003; Fowler, 2005; Sigmund and De Silva, 2010). In this paper, it will be demonstrated that the presence of selfish punishment with avoiding mechanism can help the individuals out of the two dilemmas.

2. Model

We consider a spatially structured population where each player is fastened on one site of a square lattice of size $N=L\times L$ with periodic boundaries. Each player (i) adopts strategy $s_i\in\{C,D,CP,DP\}$. Initially, the four strategies (C, D, CP, and DP) are distributed randomly and uniformly over the lattice sites.

In each iteration, each player (i) firstly plays the PDG with its four nearest neighbors in addition to itself to accumulate its original overall payoff $P^o_{s_i}$ without punishment. We have found that the situation where cooperators and altruistic punishers in addition to selfish punishers coexisting stably will not be fulfilled for various parametrizations if self-interaction (that the players can play the game with themselves) is excluded. Without self-interaction, the positive role of selfish punishment on the evolution of cooperation is weakened. We thus introduce the self-interaction into the current model. In line with the definition of PDG, each player gets the *reward* R if both choose to cooperate (C, CP) with each other, or the payoff P if both defect (D, DP). A cooperator or cooperative punisher gets the *sucker's* payoff S against a defector or defective punisher, who gets the *temptation* to defect T in such circumstance. We have checked that none of our findings for $T=b(b>1)$, $R=1$, and $P=S=0$ are essentially changed if we instead set $P=\varepsilon$ where $\varepsilon$is positive but significantly below unity. For the sake of simplicity, we just use the parametrization $T=b(b>1)$, $R=1$, and $P=S=0$, which is also commonly adopted in many studies (Nowak and May, 1992; Szabó and Tőke, 1998; Santos and Pacheco, 2005; Gómez-Gardeñes et al., 2007).

Secondly, the punishment is executed, i.e., the payoff $P^o_{s_i}$ may be modified as the remaining payoff $P^m_{s_i}$ by subtracting punishment costs and/or punishment fines. In reality, the cheaters in the face of punishers have a strong incentive to hide their identities after a defection so as to escape the punishment, causing information asymmetry. Considering this fact, we make an assumption that the states of non-cooperative (D and DP) individuals are unobservable to other punishers in our model. Consequently, there are three cases as follows: (i) To each punisher (s_i=CP or DP), it just selects one target j randomly from the non-cooperative neighbors (s_j=D or DP, $j\in\Gamma_i$ where Γ_i denotes the set of neighbours of player i) to identify the target either successfully (for D) or probably (for DP at a probability γ). Then the punisher i will impose a fine β on the exposed target at a personal cost α if its original payoff is sufficient for punishment ($P^o_{s_i}>\alpha$), or else it will do nothing. It indicates that defective punishers (DP) can still avoid being punished with probability $1-\gamma$ even though they are selected. (ii) Correspondingly, the selected non-cooperative player j will be either absolutely sanctioned if s_j=D or successfully punished with a probability γ when s_j=DP, so that its payoff is reduced by β. Instead, the unselected ones are capable of escaping the punishment. (iii) $P^m_C=P^o_C$ if s_i=C. A run for punishment over the whole lattice is performed in a random fashion in which each punisher has and only has one chance to punish.

Next, each player i chooses one of its four nearest neighbors at random and imitates the strategy of the chosen co-player j with a probability (Szabó and Tőke, 1998)

$$W_{s_i\to s_j}(P^m_{s_j}-P^m_{s_i})=1/\{1+\exp[-(P^m_{s_j}-P^m_{s_i})/\kappa]\},\tag{1}$$

where the remaining payoff of j ($P^m_{s_j}$) are also acquired in the same way mentioned above. κ=0.1 is a noise parameter describing uncertainty of strategy adoption. Over one whole Monte Carlo step (MCS), all the individuals perform an attempt for strategy updating simultaneously.