

Indirect punishment can outperform direct punishment in promoting cooperation

温, 雨佳

九州大学大学院総合理工学府総合理工学専攻機械・システム理工学メジャー

<https://hdl.handle.net/2324/7358386>

出版情報 : Kyushu University, 2024, 修士, 修士
バージョン :
権利関係 :



Indirect punishment can outperform direct punishment in promoting cooperation

Complex Social and Environmental Systems Laboratory WEN YUJIA

Introduction

Costly punishment is a key mechanism for maintaining cooperation and exists in both human and animal societies. It comes in two forms: direct punishment, where individuals punish those who directly betray them, and indirect punishment, it involves withholding cooperation rather than imposing direct penalties. Despite its importance, the effectiveness of different punishment forms in promoting cooperation remains an open question. Indirect punishment has received less attention from an evolutionary perspective [1, 2]. To address this, we examined both types of punishment in a spatial prisoner's dilemma game on a grid lattice network. We introduced a novel indirect punishment mechanism, allowing individuals to punish second-order neighbors—the neighbors of their neighbors. This approach evaluates the impact of punishment under identical costs and penalties.

Model

Cooperation can be spontaneously maintained through network reciprocity within a structured population, as cooperative clusters resist the invasion of defectors [3]. However, an increase in dilemma strength can undermine this mechanism, as defectors gain a significant payoff advantage, enabling them to infiltrate cooperative clusters. To investigate the effectiveness of punishment in promoting cooperation, we focus on representative low ($r = 0.05$) and high dilemma strength ($r = 0.2$), corresponding to scenarios where network reciprocity can either effectively maintain cooperation or fail to do so in the absence of punishment.

	CN/CP	DN/DP
CN/CP	(1,1)	(-r,1+r)
DN/DP	(1+r,-r)	(0,0)

(a) The first-stage of interaction

	Punisher's accumulated cost	Defector's accumulated fine
Punish	$-\gamma \cdot n_{D_2nd}$	$-\beta \cdot n_{P_2nd}$
Not punish	0	$-\beta \cdot n_{P_2nd}$

(b) The second-stage of interaction

Table 1. Two-stage payoff matrix for a focal player interacting with opponents, with the player's payoff being the sum of both stages.

Considering a population of size N placed on a grid lattice network with periodic boundaries, in which players play two-stage prisoner's dilemma game (PDG) with the eight direct neighbors (i.e., Moore neighborhood) [56]. In the first stage, players make decisions regarding whether to cooperate; in the second stage, players decide whether to impose punishment.

We employ a Monte Carlo asynchronous simulation where each time step consists of N updates. During each update, a focal player is randomly selected to revise their strategy, ensuring that, on average, each player updates their strategy once per time step. The focal player i updates strategy through pairwise imitation, mimics the strategy of a randomly selected direct neighbor j with a Fermi probability:

$$W = \frac{1}{1 + \exp\{(\Pi_i - \Pi_j)\kappa\}} \quad (1)$$

where Π_x represents the total payoff of player x , which derives from the interactions with all direct neighbors, along with the associated punishment imposed and incurred costs (if possible). The parameter κ denotes the imitation strength; as $\kappa \rightarrow 0$, imitation becomes random, while as $\kappa \rightarrow +\infty$, imitation is driven by payoff differences. In this study, we set $\kappa = 10$ to investigate a strong imitation scenario that imitation dependent on payoff differences. The network sizes N range from 100^2 to 400^2 in simulation. The final results are averaged from over 50 independent simulations, with each simulation producing average values from the last 2000-time steps out of a total exceeding 1×10^5 time steps.

Results

Interestingly, although indirect punishment does not increase the payoffs of direct neighbors and only punish second-order neighbors, it can promote cooperation more

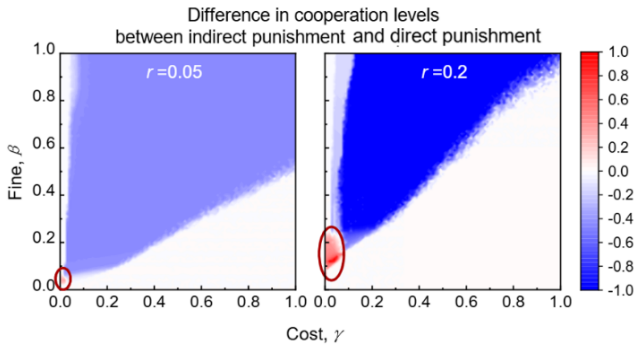


Fig 2. Indirect punishment can promote cooperation more effectively than direct punishment under conditions of low-costs and low-fines (marked with dark red circles).

effectively than direct punishment. A narrow punishment parameter region characterized by low-costs and low-fines enables indirect punishment to outperform direct punishment. In contrast, in the high-cost and high-fines region, direct punishment effectively establishes and sustains cooperation within this region. Similar phenomena are observed under high dilemma strength, where the parameter region in which indirect punishment is more effective slightly expands compared to scenarios with low dilemma strength.

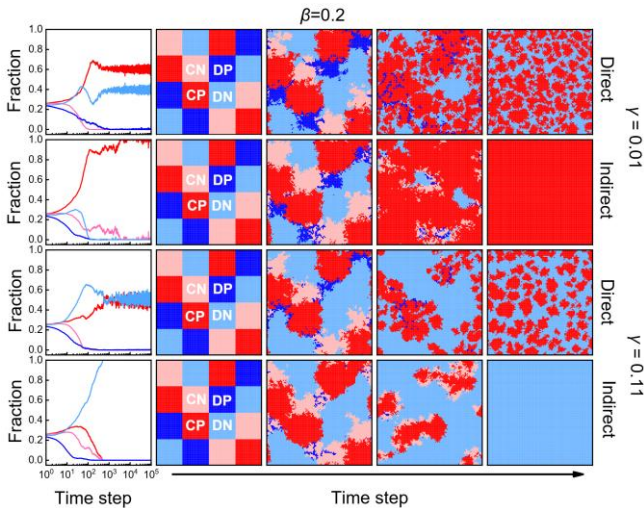


Fig 3. Indirect punishment promotes the expansion of CP clusters when the punishment cost is low, but lead to their collapse when the cost becomes high.

Indirect punishment behavior can lead to defectors being

punished more without significantly increasing the total cost for punishers when the punishment cost is low, making them more effective than direct punishment in promoting cooperation. However, high punishment costs result in a greater total punishment cost in the case of indirect punishment, which diminishes the payoff advantage for punishers, making them less effective than direct punishment.

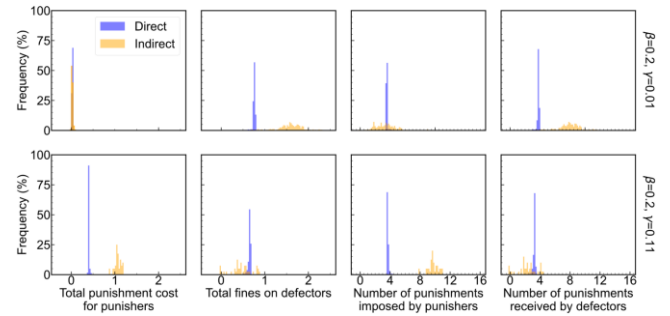


Fig 4. Probability distribution of total punishment costs and times on direct and indirect punishment.

Conclusion

we uncovered a narrow region where indirect punishment is more effective than direct punishment in structured populations. However, under other conditions, indirect punishment does not demonstrate greater effectiveness in promoting cooperation, and even be less effective than direct punishment. Taken together, these findings provide novel insights into the comparative effectiveness of different punishment forms and highlight the importance of indirect punishment in understanding the evolution of cooperation.

Reference

- [1] B.L, N.N, R.B. Direct and indirect punishment among strangers in the field. 2014;111(45):15924–15927.
- [2] U.A, S.A, R.A, Cason TN. Indirect punishment and generosity toward strangers. 2009;326(5960):1701–1704.
- [3] Nowak MA. Five rules for the evolution of cooperation. science. 2006;314(5805):1560–1563.