



Evolution of cooperation driven by incremental learning



Pei Li, Haibin Duan*

State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, PR China
 Science and Technology on Aircraft Control Laboratory, School of Automation Science and Electronic Engineering, Beihang University,
 Beijing 100191, PR China

HIGHLIGHTS

- The decision making process is formulated as an incremental learning rule.
- Strategies of players are updated according to self-learning and social-learning.
- Evolution of cooperation in the PD and the SD is inspected.
- We quantify the macroscopic features using six cluster characteristics.
- The time evolution course is examined to analyze the evolutionary results.

ARTICLE INFO

Article history:

Received 4 April 2014
 Received in revised form 26 September 2014
 Available online 14 October 2014

Keywords:

Evolution of cooperation
 Evolutionary games
 Incremental learning
 The prisoner's dilemma game
 The snowdrift game

ABSTRACT

It has been shown that the details of microscopic rules in structured populations can have a crucial impact on the ultimate outcome in evolutionary games. So alternative formulations of strategies and their revision processes exploring how strategies are actually adopted and spread within the interaction network need to be studied. In the present work, we formulate the strategy update rule as an incremental learning process, wherein knowledge is refreshed according to one's own experience learned from the past (self-learning) and that gained from social interaction (social-learning). More precisely, we propose a continuous version of strategy update rules, by introducing the willingness to cooperate W , to better capture the flexibility of decision making behavior. Importantly, the newly gained knowledge including self-learning and social learning is weighted by the parameter ω , establishing a strategy update rule involving innovative element. Moreover, we quantify the macroscopic features of the emerging patterns to inspect the underlying mechanisms of the evolutionary process using six cluster characteristics. In order to further support our results, we examine the time evolution course for these characteristics. Our results might provide insights for understanding cooperative behaviors and have several important implications for understanding how individuals adjust their strategies under real-life conditions.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Social dilemmas constitute situations that individuals take altruistic actions to contribute to collective wellbeing although selfish strategies may lead to better results for their own prosperity and success [1–4]. Generally, mutually cooperative behavior, which is ubiquitous throughout the history of evolution, is regarded as an alternative way for understanding this

* Corresponding author at: State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, PR China.
 E-mail address: hbduan@buaa.edu.cn (H. Duan).

paradoxical outcome [5–14]. That is, cooperation could emerge spontaneously among individuals who care only about their own benefits, resulting in satisfactory social welfare. The emergence and persistence of cooperation have been addressed and investigated by means of game theoretic models by researchers from a broad range of disciplines, from sociology to biology, ecology, economics, and mathematics [15,16]. Among them the prisoner's dilemma game and the snowdrift game have received great attention in both experimental and theoretical studies, and have been adopted as metaphors of behaviors in biological and social systems [17–22]. In particular, it is well accepted that spatial structure facilitates the evolution of cooperation, known as spatial reciprocity [2,4,23–28]. This mechanism allows the coexistence of cooperators and defectors even in the prisoner's dilemma game, and the cooperation dominant state, in some cases.

It is well accepted that the population dynamics in structured games, stemming from the microscopic dynamic or strategy update rules that define how successful strategies spread, have a crucial impact on evolutionary outcomes [29,30]. So it is important to explore the potential dynamic update rules underlying the evolutionary process, which is crucial for understanding the emergence of cooperation. An extremely wide variety of models have been proposed in preceding works, as reviewed comprehensively in Ref. [29], such as those based on replication [31,32], imitation [9,27], and learning [1,33], just to name a few. Among the learning strategies, imitation based processes account for a large proportion of these rules, in which condition a player usually follows the strategy of one of his fellow players. Typically, the players are assumed to adopt strategies from more successful neighbors with certain probabilities since it seems reasonable for the players to imitate his successful neighbors to expect higher payoffs. More importantly, an intensity of selection has been introduced into the imitation process to stress the stochastic strategy adoption phenomenon with errors [34]. Under this condition, an inferior strategy is possible to replace a more profitable one, as is often seen in the society or nature. Another representative strategy update rule is aspiration-based preference learning [12,30,35], under which strategies judged to have resulted in satisfactory payoffs are more likely to be selected and those leading to unsatisfactory payoffs tend to be abandoned, namely *win-stay-lose-shift*.

Although extensive research has been conducted regarding strategy update rules, how strategies are actually adopted and spread within the population is still unknown. Indeed, the important aspect of aspiration lies in that players learn from past experience in an adaptive and myopic fashion [12,29]. However, aspiration puts less emphasis on the interactive nature of decision-making even though players may act based on observed experience of peers. On the other hand, players do not necessarily follow the “*imitate-the-best*” rule and imitation rules have been questioned by recent experimental results for spatial games [36]. Nonetheless, these imitation dynamics could give some useful inspirations for understanding the high levels of cooperation characterized by the underlying dynamics. According to social psychology, individuals' cognitive processes are heavily influenced by those around them and individuals renew knowledge to promote their survival probability through social interactions [33,37–39]. In other words, learning capabilities originate from social interactions, suggesting that social sharing of information provides great evolutionary advantages and offer alternatives to solve complicated social problems.

Considering the above several facts, alternative formulations of strategies and their revision processes exploring how strategies are actually adopted and spread within the interaction network need to be further explored. In this paper, we formulate the strategy update rule as an incremental learning process to address the aforementioned problems, wherein knowledge is refreshed according to one's own experience learned from the past (self-learning) and that gained from social interaction (social-learning) and individuals learn from new information without forgetting prior knowledge. As for the strategies, it is conventional to take two competing strategies, namely cooperation and defection, as responses to social dilemmas. Yet the decision-making process during interaction with others in real-life situations is often sophisticated and much less clear-cut [40,41]. Consequently, we propose a continuous version of strategy update rules, by introducing the willingness to cooperate W , to better capture the flexibility of decision-making behavior. Importantly, the newly gained knowledge including self-learning and social learning is weighted by the parameter ω , establishing a strategy update rule involving innovative element. We should note that the strategy update rules alter the willingness to cooperate rather than the strategies themselves in the context of our work. In this way, we wish to further inspect these microscopic dynamics that describe how individuals adopt and change their strategies. For further details regarding the studied evolutionary game, the underlying microscopic dynamics driven by incremental learning, we refer to the Methods section.

In the present work, we study the evolution of cooperation driven by incremental learning on regular lattices in the prisoner's dilemma game and the snowdrift game. To inspect the underlying mechanisms of the evolutionary process, we elaborated on this subject further by quantifying the macroscopic features of the emerging patterns using six cluster characteristics. In order to further support our results, we examine the time evolution course for these five cluster characteristics. Our results highlight the underlying mechanisms of cooperative behavior and have several important implications for understanding how strategies are actually adopted and spread.

2. Methods

2.1. Games and evolutionary dynamics

Two social dilemmas, the prisoner's dilemma game and the snowdrift game have attracted significant efforts, emerging as two promising metaphors to explore the persistence of cooperation [18,21,22,25]. Notably, variation in payoff ranking for these two games, although trivial, induces a considerable change in the evolutionary dynamics [17]. In well-mixed populations, cooperation based on the snowdrift game seems easy to survive but unlikely for the prisoner's game, whereas

the odds are reversed in spatial extensions, that is, spatial structure is believed to promote cooperation in the prisoner's dilemma game but not necessarily in the snowdrift game.

In the prisoner's dilemma game and the snowdrift game, two players simultaneously decide whether to cooperate or defect during an interaction, which is characterized by the following payoffs [27,40]: mutual cooperation yields the reward R , mutual defection leads to punishment $P = 0$, and the mixed choice gives the cooperator the sucker's payoff S and the defector the temptation $T = b$. In the prisoner's dilemma game, cooperation leads to a benefit b to the opposing player, but incurs a cost c to the cooperator, thus $R = b - c$ and $S = -c$. However, in the snowdrift game, two cooperators share the costs with $R = b - c/2$ and the cooperator will bear the entire costs when competing with a defector, then $S = b - c$. Note that the benefit b exceeds the cost of cooperation c in both cases $b > c > 0$, which results in the characteristic payoffs ranking $T > R > P > S$ in the prisoner's dilemma game and $T > R > S > P$ in the snowdrift game [18]. For simplicity, the payoff matrices for these two games can be rescaled such that the evolutionary process depends only on a single parameter [21]. In this way, for the prisoner's dilemma game, we have

$$\begin{bmatrix} 1 & 0 \\ 1+u & u \end{bmatrix} \quad (1)$$

where $u = c/b$ denotes the cost to benefit ratio of cooperation and for the snowdrift game:

$$\begin{bmatrix} 1 & 1-v \\ 1+v & 0 \end{bmatrix} \quad (2)$$

where $v = c/(2b - c)$ indicates the cost to benefit ratio of mutual cooperation and $b > c$ constraints u and v in the interval $[0, 1]$.

Consider a spatially structured population where each individual is confined to sites on a two-dimensional $100 * 100$ regular lattices with periodic boundary conditions. Each individual engages in pairwise interaction with its four nearest neighbors and gains accumulative payoffs of all the interactions. Initially, the strategies of all the players are designated randomly from uniformly distributed values of W in the interval of $[0, 1]$, where W determines the willingness to cooperate for a player. Players choose to cooperate with a probability that is proportional to their willingness to cooperate W when engaged in games. For example, a player is likely to cooperate with probability 1 when $W = 1$. Individuals learn to change their willingness to cooperate W adaptively, hence constituting a continuous version of the games. Moreover, evolution of the willingness to cooperate W occurs in accordance with the Monte Carlo simulation procedure consisting of the following elementary steps. First, a randomly selected player x acquires its accumulative payoff p_x by playing the game with all its k_x neighbors. Subsequently, neighbors of x also acquire their payoffs through interactions with all their neighbors. Last then, player x updates its willingness to cooperate according to the following equation:

$$\begin{aligned} W_x(t+1) &= W_x(t) + \Delta W_x(t+1) \\ \Delta W_x(t+1) &= \omega [W_x^* - W_x(t)] + (1-\omega) [W_n^*(t) - W_x(t)] \end{aligned} \quad (3)$$

where W_x^* is the most profitable strategy of player x in all its history actions, whereas $W_n^*(t)$ denotes the most successful strategy of its four neighbors at time t . Importantly, the parameter $\omega \in [0, 1]$ takes into consideration of two factors, one's own experience learned from the past and the knowledge gained from social interaction. The case $\omega = 0$ implies that the player will surely adopt the strategy that yields the highest payoff. On the other hand, the case $\omega = 1$ indicates that it will definitely follow the best performing strategy among its neighbors, which recovers the "imitate the best" updating rule. To this end, an intermediate value of ω provides a tradeoff between these two significant factors, balancing the memory and sharing of social information in the learning process. Different from that in proceeding literature, our strategy update rule involves learning and innovative element, making it more appropriate to capture the important aspects of cognitive behavior.

2.2. Characteristics of emerging patterns

To figure out the potential relationship between microscopic strategy update rules and the macroscopic emerging patterns, here we inspect six characteristics of cooperator aggregations: cluster number N_C , and cluster shape SH_C , the number of clusters N_C , average cluster size S_C , size of the largest cluster $S_{largest}$ and number of isolated cooperators $N_{isolated}$. The shape of cluster i can be derived from the following equation [42]:

$$SH_{ci} = \frac{2l_{cc} - O_{CD}}{2l_{cc} + O_{CD}} \quad (4)$$

where l_{cc} is the number of C-C links within the cluster i and O_{CD} denotes the number of C-D links connecting the cooperators in cluster i with the surrounding defectors. Intuitively, compact clusters of cooperators have more C-C links than C-D links, reflected by $SH_i > 0$. Conversely, for filament like clusters there are fewer C-C links than C-D links with $SH_i < 0$, indicating negative assortment among cooperators. Note that cluster shape SH_C is obtained by averaging over all SH_{ci} and weighted by the size of cluster i to reveal the overall features of the emerging patterns.

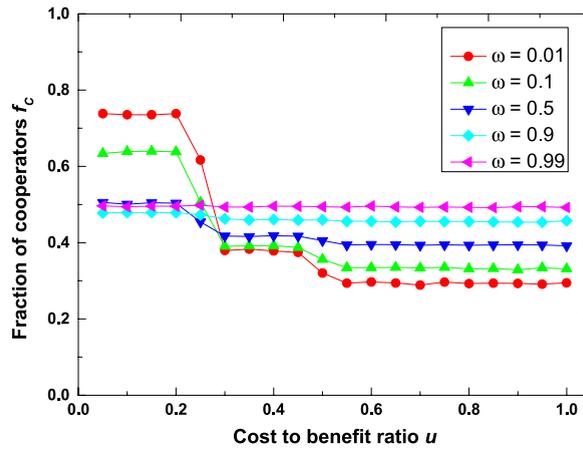


Fig. 1. The fraction of cooperators in dependence on the cost to benefit ratio for different values of ω driven by incremental learning for the prisoner's dilemma game.

3. Results and discussions

Simulations are carried out for a population of size $N = 100 \times 100$ on a two-dimensional regular lattice with periodic boundary conditions. Initially, each player on the regular lattice chooses to be a cooperator or defector with a probability in proportion to its randomly initialized willingness to cooperate W . Then all the players update their willingness to cooperate according to our proposed incremental learning rules in an asynchronous manner. Specially, each individual can adjust its strategy once on average during a full Monte Carlo step, which consists of repeating the three elementary steps mentioned previously 10,000 times. The stationary fraction of cooperators was determined within 10,000 full Monte Carlo steps after sufficiently long transients were discarded. All experiments are carried out on a personal computer with Intel Core Duo CPU T6600, 4 GB memory and Windows 7 under compiler VS 2010. It is also worth mentioning that varying the network size does not qualitatively change our reported results. Importantly, to eliminate the effects of intrinsic stochastic characteristics of the game dynamic, final results regarding the density of cooperators, the number of clusters, cluster size and the cluster shape were averaged over 20 realizations for each set of parameter values. Besides, the final fraction of cooperators under a specific set of parameters was the averaged value of last 500 Monte Carlo steps to guarantee appropriate accuracy.

It has been reported that cooperation is enhanced by forming compact clusters in structured populations in the prisoner's dilemma, often referred to as spatial reciprocity or R reciprocity. However, in contrast to the prisoner's dilemma, spatial structure is generally detrimental to cooperation in the snowdrift game. This fundamental difference arises from distinct payoff structures of these two games, wherein it is rational for the players to adopt strategies opposite to their opponents in the snowdrift game. Specially, evolution of cooperation in the snowdrift game deserves to be addressed to explore the underlying mechanisms that promote the social welfare, although persistence of cooperation is not a problem in this game. In addition, the snowdrift game has also often been adopted as a metaphor for dilemma situations in biological and social systems. So here we investigate the evolution of cooperation driven by incremental learning to explore how strategies are actually adopted and spread within the population.

The evolutionary results driven by incremental learning are depicted in Figs. 1 and 2. Generally, the fraction of cooperators under this update rule is consistent with that produced by other learning rules. As expected, the equilibrium proportion of cooperators decreases with the increase of cost to benefit ratio in both games. We have also noticed that the cooperators or defectors are never able to dominate the whole population across the whole span of parameters in both games, which can be attributed to the stochastic feature of our learning rule. In other words, the cooperators and defectors coexist during the evolutionary process, in accordance with many real world scenarios. Apparently, it recovers the condition that individuals adjust their willingness to cooperate according to social learning (imitate actions from their successful neighbors) when $\omega \rightarrow 0$, whereas individuals adopt strategies based on memory-based learning (imitate their best past action) when $\omega \rightarrow 1$.

For the prisoner's dilemma game, low ω strongly supports the cooperation level for small u , up to $u = 0.25$, whereas low ω is detrimental to the maintenance of cooperation for large u (Fig. 1). At this point, we argue that for $\omega \rightarrow 0$, when players adjust their willingness to cooperate according to the knowledge gained from social learning, cooperation becomes the dominant strategy. In particular, for $\omega \rightarrow 1$, when players learning from their past experience, players stick to their past best actions, thus become immune to invasion of defectors for high temptation to defect. The spatial reciprocity effect deteriorates since players' learning capability is limited to their original state when $\omega \rightarrow 1$. Consequently, the players are constrained to the initial freezing state, which perfectly explains why cooperators are able to avoid being occupied by defectors for large u . Importantly, two regimes can be distinguished: the fraction of cooperators declines dramatically at the threshold of $u = 0.25$ and different values of ω have a reversed effect on the evolutionary results when $u > 0.25$. Moreover, the fraction of cooperators remains almost the same with a certain range of u , unlike the popular games that do not have innovative elements in strategy update rules [27].

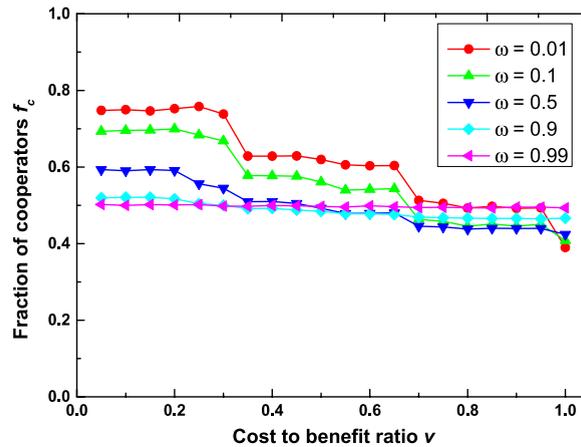


Fig. 2. The fraction of cooperators in dependence on the cost to benefit ratio for different values of ω driven by incremental learning for the snowdrift game.

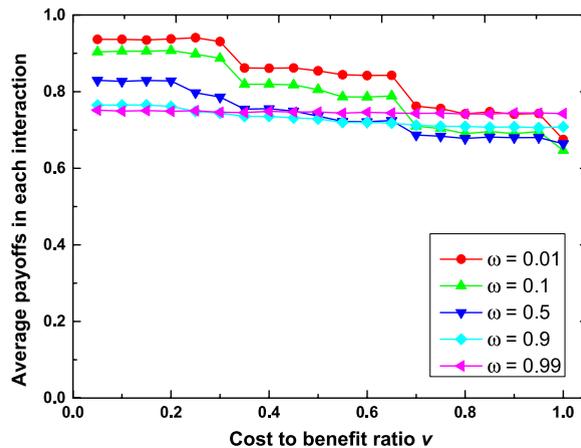


Fig. 3. Average payoffs in each interaction in dependence on the cost to benefit ratio for different values of ω driven by incremental learning for the snowdrift game.

As for the snowdrift game, it has been realized that it is more profitable to adopt alternating coordinating strategy than to act as cooperators constantly, which is also referred to as *ST* reciprocity [43–45]. Importantly, *ST* reciprocity is much more meaningful for evolution in the context of the snowdrift game than *R* reciprocity in the prisoner's dilemma game [44]. In what follows, we analyze the fraction of cooperators as well as the payoff levels in the snowdrift game driven our learning rule (Figs. 2 and 3). The network maintains a cooperator dominant state when $\omega = 0.01$ for most v in the snowdrift game (Fig. 2), suggesting that more emphasis on social interaction with neighbors favors cooperation. In particular, cooperation levels drop more slowly than that of the prisoner's dilemma game with increasing v for a specific ω . We have also noticed that the average payoff shown in Fig. 3 is in good agreement with the fraction of cooperators in Fig. 2. In contrast to the prisoner's dilemma, it is relatively robust against dilemma strength since the cooperation level and average payoff are maintained to some extent for larger cost to beneficial ratio v . In analogy to the prisoner's dilemma game, we observed a qualitatively similar impact of the incremental learning, that is, two regimes can be figured out although it is inconspicuous. From this point of view, adopting strategies from successful peers facilitates the cooperation for small cost to benefit ratio, exhibiting evolutionary advantage over that learning from their past experience.

Further inspection of the spatial distribution of strategies provides us some intuition in regard to how our incremental learning rule affects the outcome and stationary equilibrium of the evolutionary game. For the sake of convenience, in the following discussion, we denote the players as cooperators when their willingness to cooperate $W > 0.5$, or as defectors when $W < 0.5$. In the prisoner's dilemma game, with increasing ω , the clusters of defectors emerge and split the giant cooperator cluster into small islands for $u = 0.2$ (Fig. 4(A)). However, the situation is reversed when $u = 0.6$, that is, all the players tend to defect for small ω and cooperators arise along with increasing ω (Fig. 4(B)). Here we choose the parameter $v = 0.2$ and $v = 1$ for the snowdrift game, at which we can easily recognize two regimes qualitatively identical to that in the prisoner's dilemma game for different ω (Fig. 4(C), (D)). Note that the snowdrift game provides better environments for evolution of cooperation than that in the prisoner's dilemma game when $u = v = 0.2$, as evidenced by spatial patterns arise

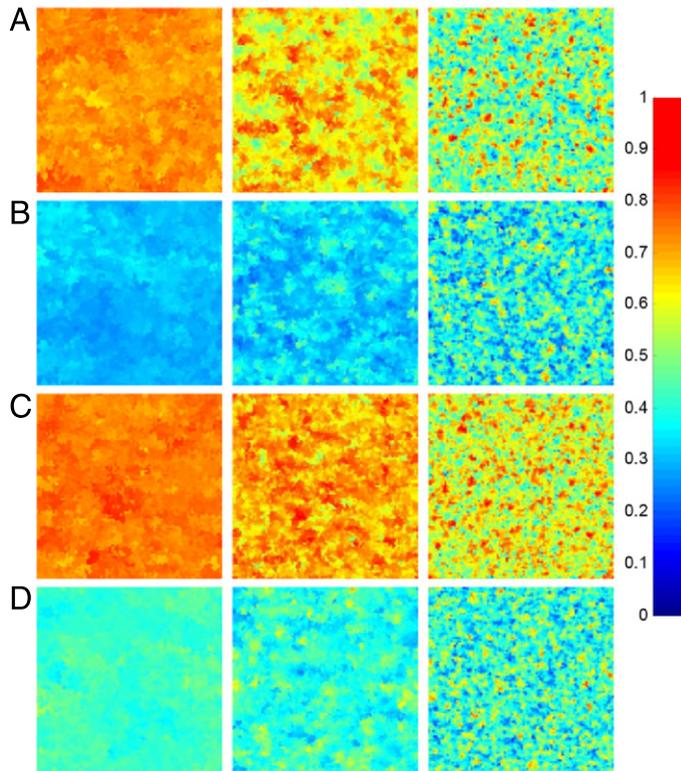


Fig. 4. Characteristic spatial distribution of color-coded strategies for the prisoner's dilemma game and the snowdrift game, as obtained for different values of u (v for the snowdrift game) and ω . From left to right, the values of ω are 0.01, 0.1, 0.5, respectively. (A) $u = 0.2$ for the prisoner's dilemma game. (B) $u = 0.6$ for the prisoner's dilemma game. (C) $v = 0.2$ for the snowdrift game. (D) $v = 1$ for the snowdrift game. The willingness to cooperate W is color-coded as follows: blue represents full defector and red represents full cooperators. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

from microscopic processes. Also worth highlighting is that, by taking into account the willingness to cooperate, compact clusters emerge in the snowdrift game rather than filament-like clusters reported in many literatures [17,21].

In what follows, we quantify the macroscopic features of the emerging patterns to inspect what are the mechanisms using six cluster characteristics: the fraction of cooperators f_c , cluster shape SH_c , the number of clusters N_c , the cluster size S_c , size of the largest cluster $S_{largest}$ and number of isolated cooperators $N_{isolated}$ (Fig. 5). Interestingly, curves showing the cluster shape are in good agreement with those showing the fraction of cooperators, irrespective of the type of games and the value of ω , indicating that the cluster shape quantifies the qualitative descriptions of various clusters perfectly. In the prisoner's dilemma game, the cluster size and cluster count exhibit two phases: the sharp drop in cluster size is accompanied by a sharp increase in the number of clusters near the threshold $u = 0.25$, which is in consistent with our observation that the fraction of cooperators undergo a transition at about $u = 0.25$, for both $\omega = 0.01$ and $\omega = 0.5$.

When $\omega = 0.01$, size of largest cluster is close to the network size for $u < 0.25$. Combined with cluster shape shown in Fig. 5(A), it is reasonable to infer that cooperators grow into a compact cluster, allowing them to minimize the cluster surface and hence reduce exploitation by defectors, evidenced by the relatively high level of cooperators. Increasing u gives rise to smaller cluster sizes, negative assortment among cooperators and more isolated cooperators, that is, defectors succeed in splitting up the cluster of cooperators into smaller components. In this situation, exploitation seems more appealing to the defectors and cooperators are easy to be invaded by the defectors. Note that the number of clusters can be more than an order of magnitude larger when $\omega = 0.5$ than that when $\omega = 0.01$ for small u . Moreover, the cluster size decreases much more slowly as u increases and reaches a stationary level of 3.8 for $\omega = 0.5$.

Next, we analyze the spatial configurations of cooperation in the snowdrift game for different values of v to verify the impact of incremental learning on evolution of cooperation (Fig. 6). In line with our previous observations, the snowdrift game provides better playground for the evolution of cooperation. The system remains a high level of cooperation across different v and declines more slowly with respect to that of the prisoner's dilemma game. Moreover, the decrease of cooperators has experienced several stages rather than the clear-cut two phases. When focused on the size of the largest cluster, we would find that the largest cluster survives for a wider range of v in increasingly hostile conditions before it split into small islands when $\omega = 0.01$. As expected, the fraction of cooperators remains almost the same for $\omega = 0.5$, owing to its less learning capabilities from social interaction. Altogether, it is worth pointing out that the macroscopic pattern of cooperation remains qualitatively the same for both types of games.

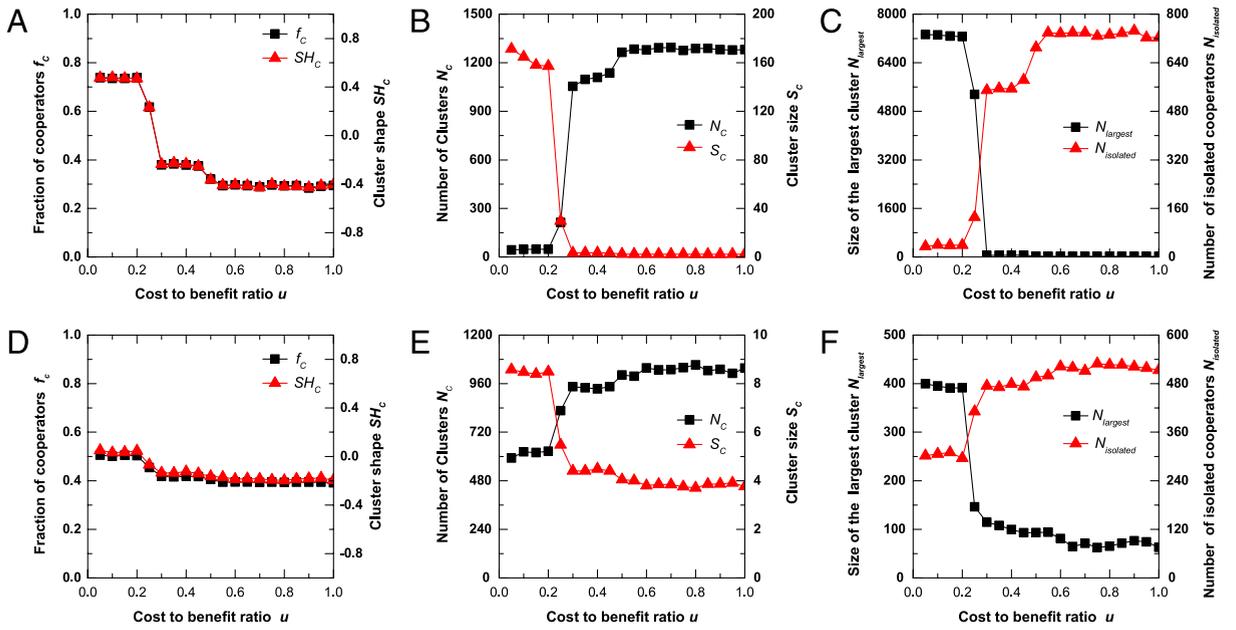


Fig. 5. Macroscopic characteristic of evolutionary dynamics of cooperation in dependence on different u in prisoners' dilemma game: the fraction of cooperators, cluster shape, cluster number, cluster size, size of the largest cluster and number of isolated cooperators, top panels for $\omega = 0.01$ and bottom panels for $\omega = 0.5$.

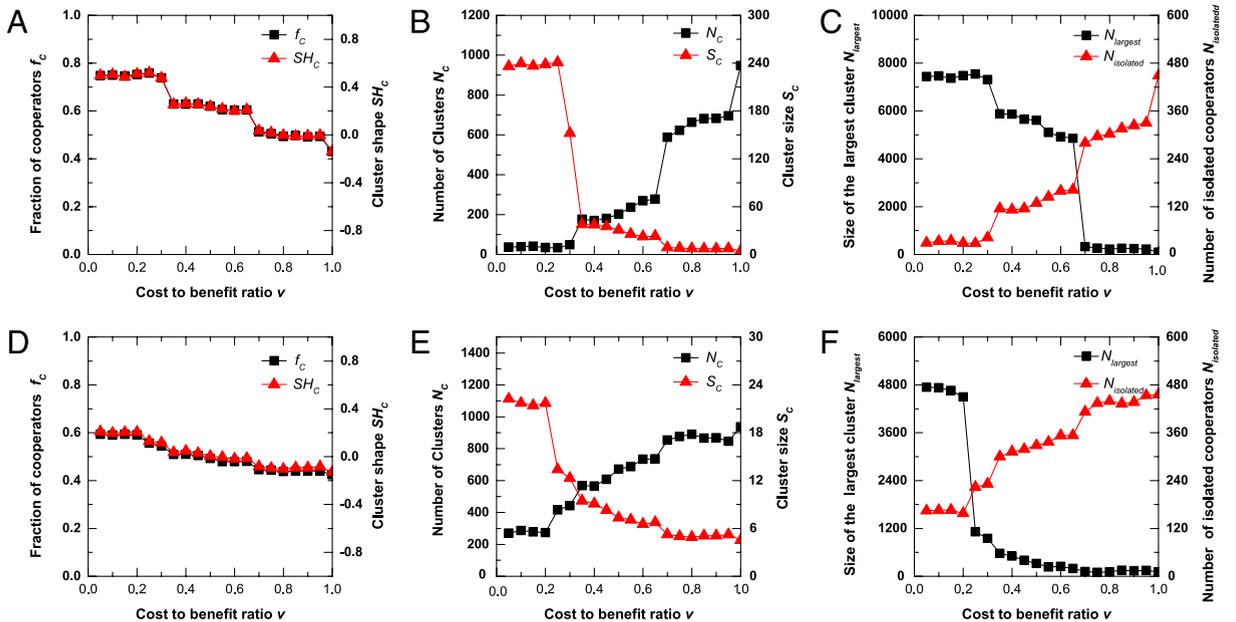


Fig. 6. Macroscopic characteristic of evolutionary dynamics of cooperation in dependence on different v in the snowdrift game: the fraction of cooperators, cluster shape, cluster number, cluster size, size of the largest cluster and number of isolated cooperators, top panels for $\omega = 0.01$ and bottom panels for $\omega = 0.5$.

In order to further support our results, we examine the time evolution course for the fraction of cooperators and five other cluster characteristics for these two games. Initially, cooperators and defectors are distributed uniformly on the regular lattice, as reflected by the cluster shape in Fig. 7. Then cooperators learn to optimize their strategies gradually through social interactions. In this way, cooperators are able to avoid exploitation from defectors by forming compact clusters. During this period, isolated cooperators are easily to be eliminated by defectors, constituting the so called enduring period [21,46]. Correspondingly, in case of the snowdrift game, similar trends can also be observed despite that the enduring period is less obvious. The absence of the decrease of cooperators in the enduring period may be owing to the microscopic dynamic determined by the inherent payoff structure of the snowdrift game. Subsequently, expanding period comes since cooperators

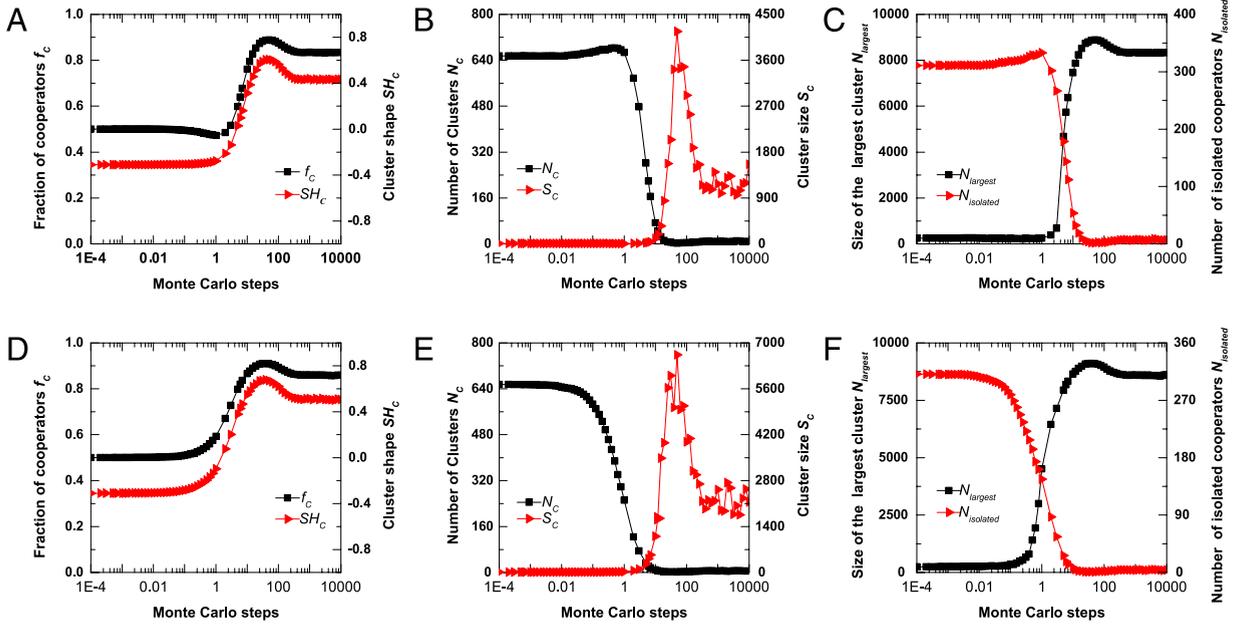


Fig. 7. Time course depicting the evolution of cooperation in the prisoner's dilemma game (top panels) and the snowdrift game (bottom panels) for $\omega = 0.01$ and $u = 0.2$. From left to right (A–C, D–F), depicted in the figure are the fraction of cooperators, cluster shape, the number of clusters, cluster size, size of the largest cluster and number of isolated cooperators.

successfully survive and expand their areas by forming cooperator clusters. It is interesting that the second local equilibrium of the system is along with a decrease of the fraction of cooperators. To explain this unusual phenomenon, we start by considering the following case that the willingness to cooperate of a player is between its personal best strategy W_x^* and its neighborhood's best strategy W_n^* . Suppose that the distance between W_x and W_n^* is further than the one between W_x and W_x^* . In this case, the player will adjust its strategy and move towards W_n^* . However, the distance between W_x and W_n^* will increase as moving towards W_n^* . To this end, the oscillation would occur and decay convergence. So the unusual phenomenon may be attributed to memory-based learning included in the strategy update rule driven by incremental learning.

4. Conclusions

It has been shown that evolutionary outcomes of spatial games depend crucially on the details of the microscopic population dynamics defined by strategy update rules. A considerable body of literature has proposed various models to explore the potential dynamic update rules underlying the evolutionary process, among which imitation-based and aspiration-based dynamics are two relevant learning methods. Importantly, we note that most rules are non-innovative, in other words, they cannot introduce new strategies not currently present in the population. We have also noticed that individual inertia is a ubiquitous in social or biological systems, which refers to the phenomenon that individuals are resistant to change their current state. According to social psychology, individuals update their strategies from one's own experience learned from the past and the knowledge gained from social interaction. Motivated by the above several facts, we formulated the strategy update rule as an incremental learning process, incorporating the self-learning and social learning aspects of decision-making into the population dynamics. Specially, we consider a continuous alternative of strategy update rules to capture the flexibility of behavior, by introducing the willingness to cooperate, since decision-making process in real-life situations is often sophisticated and much less clear-cut. In this way, individuals are able to choose strategies according to their willingness to cooperate rather than to choose only between the two extremes and take arbitrary values as strategies in the interval $[0, 1]$, allowing us to investigate the evolution of cooperation in more realistic conditions. It is also worth pointing out that the newly gained knowledge including self-learning and social learning is weighted by the parameter ω , establishing a strategy update rule involving innovative element.

Due to the differences between the prisoner's dilemma game and the snowdrift game, we have taken these two games as candidates to inspect the evolution of cooperation driven by incremental learning on regular lattices. Importantly, two regimes can be recognized from the fraction of cooperators in both games for different values of u , irrespective of ω , as evidenced by spatial patterns arise from microscopic processes. To inspect what are the underlying mechanisms of the evolutionary process, we elaborated on this subject further by quantifying the macroscopic features of the emerging patterns using six cluster characteristics: the fraction of cooperators, cluster shape, the number of clusters, the cluster size, size of the largest cluster and number of isolated cooperators. We have noticed that the macroscopic pattern of cooperation remains qualitatively the same for both types of games, in good agreement with previous results of stationary states. In order to

further support our results, we examine the time evolution course for these five cluster characteristics for these two games. It is interesting that the second local equilibrium of the system is along with a decrease of the fraction of cooperators during the time evolution, which can be attributed to the intrinsic stochastic property of incremental learning. Combined with the above analysis, it is reasonable to conclude that these cluster characteristics we adopted provide good measurements for the evolution of cooperation in both games. Our results highlight the underlying mechanisms of cooperative behavior and have several important implications for understanding how strategies are actually adopted and spread.

Acknowledgments

This work was partially supported by Natural Science Foundation of China (NSFC) (Grant Nos. 61333004 and 61273054), Top-Notch Young Talents Program of China, and Aeronautical Foundation of China (Grant No. 2013585104).

References

- [1] M.W. Macy, A. Flache, Learning dynamics in social dilemmas, *Proc. Natl. Acad. Sci. USA* 99 (2002) 7229–7236.
- [2] F.C. Santos, J.M. Pacheco, T. Lenaerts, Evolutionary dynamics of social dilemmas in structured heterogeneous populations, *Proc. Natl. Acad. Sci. USA* 103 (2006) 3490–3494.
- [3] A. Szolnoki, M. Perc, Promoting cooperation in social dilemmas via simple coevolutionary rules, *Eur. Phys. J. B* 67 (2009) 337–344.
- [4] F.L. Pinheiro, F.C. Santos, J.M. Pacheco, How selection pressure changes the nature of social dilemmas in structured populations, *New J. Phys.* 14 (2012) 073035.
- [5] R. Axelrod, W.D. Hamilton, The evolution of cooperation, *Science* 211 (1981) 1390–1396.
- [6] D. Challet, Y.-C. Zhang, Emergence of cooperation and organization in an evolutionary game, *Physica A* 246 (1997) 407–418.
- [7] A. Sanchez, J.A. Cuesta, Altruism may arise from individual selection, *J. Theoret. Biol.* 235 (2005) 233–240.
- [8] J. Gómez-Gardeñes, M. Campillo, L. Floría, Y. Moreno, Dynamical organization of cooperation in complex topologies, *Phys. Rev. Lett.* 98 (2007) 108103.
- [9] A. Szolnoki, M. Perc, Z. Danku, Making new connections towards cooperation in the prisoner's dilemma game, *Europhys. Lett.* 84 (2008) 50007.
- [10] F.C. Santos, M.D. Santos, J.M. Pacheco, Social diversity promotes the emergence of cooperation in public goods games, *Nature* 454 (2008) 213–216.
- [11] C. Xu, H. Duan, Artificial bee colony (ABC) optimized edge potential function (EPF) approach to target recognition for low-altitude aircraft, *Pattern Recognit. Lett.* 31 (2010) 1759–1772.
- [12] Y. Liu, X. Chen, L. Wang, B. Li, W. Zhang, H. Wang, Aspiration-based learning promotes cooperation in spatial prisoner's dilemma games, *Europhys. Lett* 94 (2011) 60002.
- [13] J. Gómez-Gardeñes, I. Reinares, A. Arenas, L.M. Floría, Evolution of cooperation in multiplex networks, *Sci. Rep.* 2 (2012) 00620.
- [14] P. Li, H. Duan, Robustness of cooperation on scale-free networks in the evolutionary prisoner's dilemma game, *Europhys. Lett* 105 (2014) 48003.
- [15] K. Shigaki, S. Kokubo, J. Tanimoto, A. Hagishima, N. Ikegaya, Does copy-resistance enhance cooperation in spatial prisoner's dilemma? *Europhys. Lett* 98 (2012) 40008.
- [16] Z. Wang, A. Szolnoki, M. Perc, Evolution of public cooperation on interdependent networks: the impact of biased utility functions, *Europhys. Lett* 97 (2012) 48001.
- [17] C. Hauert, M. Doebeli, Spatial structure often inhibits the evolution of cooperation in the snowdrift game, *Nature* 428 (2004) 643–646.
- [18] M. Doebeli, C. Hauert, Models of cooperation based on the prisoner's dilemma and the snowdrift game, *Ecol. Lett.* 8 (2005) 748–766.
- [19] M. Perc, Evolution of cooperation on scale-free networks subject to error and attack, *New J. Phys.* 11 (2009) 033027.
- [20] C.P. Roca, J.A. Cuesta, A. Sánchez, Evolutionary game theory: temporal and spatial effects beyond replicator dynamics, *Phys. Life Rev.* 6 (2009) 208–249.
- [21] F. Fu, M.A. Nowak, C. Hauert, Invasion and expansion of cooperators in lattice populations: prisoner's dilemma vs. snowdrift games, *J. Theoret. Biol.* 266 (2010) 358–366.
- [22] Z. Wang, C.-Y. Xia, S. Meloni, C.-S. Zhou, Y. Moreno, Impact of social punishment on cooperative behavior in complex networks, *Sci. Rep.* 3 (2013) 03055.
- [23] M.A. Nowak, R.M. May, Evolutionary games and spatial chaos, *Nature* 359 (1992) 826–829.
- [24] M.A. Nowak, Five rules for the evolution of cooperation, *Science* 314 (2006) 1560–1563.
- [25] P. Langer, M.A. Nowak, C. Hauert, Spatial invasion of cooperation, *J. Theoret. Biol.* 250 (2008) 634–641.
- [26] C.P. Roca, J.A. Cuesta, A. Sánchez, Effect of spatial structure on the evolution of cooperation, *Phys. Rev. E* 80 (2009) 046106.
- [27] C. Zhang, J. Zhang, G. Xie, L. Wang, M. Perc, Evolution of interactions and cooperation in the spatial prisoner's dilemma game, *PLoS ONE* 6 (2011) e26724.
- [28] M. Perc, J. Gómez-Gardeñes, A. Szolnoki, L.M. Floría, Y. Moreno, Evolutionary dynamics of group interactions on structured populations: a review, *J. R. Soc. Interface* 10 (2013) 20120997.
- [29] G. Szabó, G. Fáth, Evolutionary games on graphs, *Phys. Rep.* 446 (2007) 97–216.
- [30] J. Du, B. Wu, P.M. Altrock, L. Wang, Aspiration dynamics of multi-player games in finite populations, *J. R. Soc. Interface* 11 (2014) 20140077.
- [31] J. Bendor, P. Swistak, Types of evolutionary stability and the problem of cooperation, *Proc. Natl. Acad. Sci. USA* 92 (1995) 3596–3600.
- [32] J. Hofbauer, K. Sigmund, *Evolutionary Games and Population Dynamics*, Cambridge University Press, 1998.
- [33] S. Wang, M.S. Szalay, C. Zhang, P. Csermely, Learning and innovative elements of strategy adoption rules expand cooperative network topologies, *PLoS ONE* 3 (2008) e1917.
- [34] G. Szabó, J. Vukov, A. Szolnoki, Phase diagrams for an evolutionary prisoner's dilemma game on two-dimensional lattices, *Phys. Rev. E* 72 (2005) 047107.
- [35] M. Perc, Z. Wang, Heterogeneous aspirations promote cooperation in the prisoner's dilemma game, *PLoS ONE* 5 (2010) e15117.
- [36] C.P. Roca, D. Helbing, Emergence of social cohesion in a model society of greedy, mobile individuals, *Proc. Natl. Acad. Sci. USA* 108 (2011) 11370–11374.
- [37] R. Poli, J. Kennedy, T. Blackwell, Particle swarm optimization, *Swarm. Intell.* 1 (2007) 33–57.
- [38] D. Marchiori, M. Warglien, Predicting human interactive learning by regret-driven neural networks, *Science* 319 (2008) 1111–1113.
- [39] A. Traulsen, D. Semmann, R.D. Sommerfeld, H.-J. Krambeck, M. Milinski, Human strategy updating in evolutionary games, *Proc. Natl. Acad. Sci. USA* 107 (2010) 2962–2966.
- [40] X. Chen, Y. Liu, Y. Zhou, L. Wang, M. Perc, Adaptive and bounded investment returns promote cooperation in spatial public goods games, *PLoS ONE* 7 (2012) e36895.
- [41] Y. Zhang, F. Fu, T. Wu, G. Xie, L. Wang, A tale of two contribution mechanisms for nonlinear public goods, *Sci. Rep.* 3 (2013) 02021.
- [42] K. Shigaki, Z. Wang, J. Tanimoto, E. Fukuda, Effect of initial fraction of cooperators on cooperative behavior in evolutionary prisoner's dilemma game, *PLoS ONE* 8 (2013) e76942.
- [43] J. Tanimoto, H. Sagara, A study on emergence of alternating reciprocity in a 2×2 game with 2-length memory strategy, *BioSystems* 90 (2007) 728–737.
- [44] M. Wakiyama, J. Tanimoto, Reciprocity phase in various 2×2 games by agents equipped with two-memory length strategy encouraged by grouping for interaction and adaptation, *BioSystems* 103 (2011) 93–104.
- [45] J. Tanimoto, What initially brought about communications? *BioSystems* 92 (2008) 82–90.
- [46] Z. Wang, S. Kokubo, J. Tanimoto, E. Fukuda, K. Shigaki, Insight into the so-called spatial reciprocity, *Phys. Rev. E* 88 (2013) 042145.