Special Agents Can Promote Cooperation in the Population

Xin Wang, Jing Han*, Huawei Han

Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China

Abstract

Cooperation is ubiquitous in our real life but everyone would like to maximize her own profits. How does cooperation occur in the group of self-interested agents without centralized control? Furthermore, in a hostile scenario, for example, cooperation is unlikely to emerge. Is there any mechanism to promote cooperation if populations are given and play rules are not allowed to change? In this paper, numerical experiments show that complete population interaction is unfriendly to cooperation in the finite but end-unknown Repeated Prisoner's Dilemma (RPD). Then a mechanism called soft control is proposed to promote cooperation. According to the basic idea of soft control, a number of special agents are introduced to intervene in the evolution of cooperation. They comply with play rules in the original group so that they are always treated as normal agents. For our purpose, these special agents have their own strategies and share knowledge. The capability of the mechanism is studied under different settings. We find that soft control can promote cooperation and is robust to noise. Meanwhile simulation results demonstrate the applicability of the mechanism in other scenarios. Besides, the analytical proof also illustrates the effectiveness of soft control and validates simulation results. As a way of intervention in collective behaviors, soft control provides a possible direction for the study of reciprocal behaviors.

Introduction

Since Darwin's evolutionary theory, researchers have been long puzzled by the problem that why there exists wide cooperation among species [1–3]. As the paradigm of studying reciprocal behaviors, the Prisoner's Dilemma has been abstracted to depict many biological processes [4–8], and it raises a question to us, how to sustain cooperation in the group of self-interested agents without centralized control.

As known, in a single shot of two-agent Prisoner's Dilemma, mutual defection is the only equilibrium. With the number of agents increasing, it becomes unfriendly to cooperation either [9]. A large amount of theoretical work have studied assorted scenarios where cooperation can emerge. They can be mainly divided into three categories. First, the "catalysis" to sustain cooperation is studied. A specific proportion of "Tit for Tat" (TFT) in the population is crucial to the emergence of cooperation but the strategy of "Pavlov" is the last laughter [10,11]. Punishment is considered as an important way to support cooperative behaviors and studied in spatial public goods game [12,13], indirect reciprocity [14–16], group selection [17,18] or other scenarios [19–22]. Besides costly punishment, reward can also promote cooperation [23]. Second, extra abilities or characteristics are provided to agents. The tag mechanism where an agent's decision depends not only on its play strategy but also on arbitrary tags associated with the agents can make it easy for populations to reach reciprocal cooperation [24,25]. The mobility of an agent who interacts with its local neighbors also increases the capability

of cooperation to emerge [26–29]. Third, introducing the topological structure in games, e.g. the lattice, the random graph or the scale-free network, has been proved to be an effective way to support cooperation because local interaction provides an opportunity for cooperators to cluster, grow and resist against the invasion by defectors [30–38]. Different samplings of interaction partners have effect on the cooperation level [39]. In addition, the introduction of coevolutionary rules combining the evolution of play strategies and other properties is beneficial to the prevalence of cooperation [40–49].

In this paper, our purpose is not to study which scenario can favor cooperation, but to propose a mechanism called soft control [50,51] to promote cooperation in the unfavorable scenario. Moreover sometimes original populations and play rules are not allowed to alter because any change may incur high cost. Thus it is natural to ask how to promote cooperation under this circumstance. According to the basic idea of soft control, a number of special agents called shills are added to the original group to intervene in the evolution of cooperation. These shills pose as normal agents by conforming to play rules, thus they are always treated as normal agents by truly normal ones. The difference is that a shill has its own strategy and it can recognize other shills. This allows shills to share their knowledge of interacting with normal agents and take appropriate action in games based on knowledge. We think that this assumption is reasonable in some scenario of real life. Consider e-commerce: in order to publicize products, some sellers may employ a number of shills to compliment products in web media. Those shills recognize each

other if the seller informs them, but ordinary consumers cannot differentiate. The preliminary result of soft control to promote cooperation in the particular scenario can be found in [52].

In following parts we study the performance of soft control under different settings by numerical experiments, which include: (1) the short-term vs. long-term RPD; (2) noise-free vs. noisy interaction; (3) complete vs. incomplete population interaction. For (1), our purpose is to show how the mechanism takes effect upon a wide range of the finitely RPD. Here note that albeit the finitely RPD is considered in this paper, we always assume that the time period of games is unknown to all players (i.e. normal agents and shills). So under this circumstance the finitely RPD is usually considered as the infinitely RPD. There are many theoretical studies based on the finitely RPD [1,53–56]. For (2), the sensitivity of soft control to noise is presented where noise derives from mistakes to take the opposite action. The motivation of this experiment is to check whether the mechanism is robust to noise because any single-bit error in action between two TFT agents will destroy cooperation. It is called "cascade of curse". For (3), we derive our main results in this paper on considering complete population interaction, i.e. every player plays with all others. A case of incomplete interaction is studied to demonstrate the applicability of soft control in other scenarios. In addition, we also give the analytical proof of the effectiveness under complete interaction (see in Appendix S1) to validate and complement simulation results.

Methods

In the Prisoner's Dilemma, both players make their choices simultaneously, cooperation (C) or defection (D). Their payoffs depend on which action they choose. The payoff matrix considered in this paper is written as the following form:

$$
\begin{array}{cc}
 & \begin{array}{cc} C & \quad D \end{array} \\
\begin{array}{c} C \\ D \end{array} & \begin{pmatrix} (R,R) & (S,T) \\ (T,S) & (P,P) \end{pmatrix}
\end{array}
$$

where parameters in the matrix satisfy $T > R > P > S$ and $R > (T+S)/2$.

According to the above matrix, note that in a single stage of the game D is the best option for a player regardless of its opponent's choice. As a result both obtain P points. But if they had cooperated with each other, they would have received higher payoffs, R points. This is the dilemma between individual and collective rationality. Meanwhile playing C continuously is better than doing C and D alternatively for the study of reciprocal behaviors in the RPD.

Basic model: populations and play rules

Consider the mixed reactive strategy [57] for each normal agent, which is described as $(y,p,q)\in\mathbb{R}[0,1]^3$, where y is the probability of cooperating on the first stage, p and q are the conditional probability of taking cooperation on the current stage responding to the opponent's last move defection and cooperation respectively. The space of reactive strategies displays plentiful phenomena and has rich analytical results [10,24,54,57–59]. It can also describe the fundamental element in decision makings, i.e. the mapping from stimulus to response, similar to if-then rule.

Let N_A denote as the number of normal agents, β as the time period of games (e.g. $\beta = 10$ means the 10-stage RPD) and t as the index of generations. Let $\mathcal{A} = \{1,2,\ldots,N_A\}$. We assume the number of the population to be constant in each generation.

This paper mainly studies complete population interaction, i.e. each normal agent interacts with all others (in simulations, the incomplete interaction cases are also provided). In t generation ($t \geq 0$), every pair of agents play the β-stage RPD once. The pairing order is random, but it does not influence an agent's payoff. Agent i ($i\in\mathcal{A}$) updates its payoff after each RPD game. Let $f_{ij}(t)$ denote as the total payoff agent i receives from playing with agent j for the β-stage RPD. Then agent i's total payoff $f_i(t) = \sum_{j\in\mathcal{A}\backslash\{i\}} f_{ij}(t)$. At the end of a generation, all agents self-reproduce. The expected number of agent i's offspring in $t+1$ generation, denoted as $E\{\#_i(t+1)\}$, is calculated as follows:

$$
E\{\#_i(t+1)\} = \frac{f_i(t)}{\sum_{k\in\mathcal{A}} f_k(t)} N_A \qquad \forall i\in\mathcal{A} \qquad (1)
$$

As we will show in the simulation section, cooperation is impossible to emerge in the self-organized group of normal agents.

Model with soft control: shills and their strategies

To promote cooperation, a number of shills are added to the original group. Let N_S denote as the number of shills. Again, we assume the number of the population $N = N_A + N_S$ to be constant in each generation. Let $\mathcal{P} = \{1,2,\ldots,N\}$. As mentioned above, shills are treated as normal agents by conforming to play rules. Meanwhile it is assumed that shills know nothing about normal agents' strategies, but they can remember and share the action sequence of normal agents playing with shills in the current generation. With this knowledge, they can estimate the level of cooperativity of normal agents and take suitable action. The simplest way of estimating a normal agent's strategy based on the action sequence is to calculate the frequency of cooperation. And then a shill uses it to decide appropriate reaction: to cooperate if the normal agent has high frequency of cooperation, otherwise to defect. This is what we called Frequency-based Tit for Tat (F-TFT). Note that F-TFT is a different form of strategy from normal agents' reactive strategy (y,p,q). But this is allowed in soft control because shills can use their own strategies as long as they conform to play rules in the original group. We utilize F-TFT as a shill's strategy in the following part.

In each generation, all shills share knowledge (m_i, n_i) for normal agent i ($i\in\mathcal{A}$), where m_i is how many stages agent i has interacted with shills so far and n_i is the number of cooperation in m_i stages. At the beginning of each generation m_i and n_i are initialized as 0. Then a shill with F-TFT uses (m_i, n_i) to make decisions: if $m_i = 0$, the shill cooperates; otherwise it cooperates with the probability n_i/m_i. After a stage, m_i is increased by 1, and n_i is increased by 1 if agent i cooperates at that stage. Because each shill can access (m_i, n_i), F-TFT is always based on the history of shills interacting with agent i so far. Therefore at the end of each generation, $m_i = \beta \cdot N_S$ and n_i is the total number of cooperation that agent i takes while playing with shills.

For any $k\in\mathcal{P}$, player k's total payoff $f_k(t) = \sum_{j\in\mathcal{P}\backslash\{k\}} f_{kj}(t)$ where $f_{kj}(t)$ is the total payoff player k receives from interacting with player j. Rewrite Eq. (1) as below:

$$
E\{\#_i(t+1)\} = \frac{f_i(t)}{\sum_{k\in\mathcal{P}} f_k(t)} N \qquad \forall i\in\mathcal{A} \qquad (2)
$$

According to Eq. (2), note that the role of shills is to intervene in a normal agent's payoff through interacting with it. In fact, shills

playing F-TFT reward cooperative behaviors and punish defective ones so as to promote cooperation.

Results

The performance of soft control is studied through a series of numerical experiments. Simulation results presented in figures are averaged on 100 independent random realizations where $N_A = 500$ and a normal agent's strategy is uniformly generated in $\mathbb{R}[0,1]^3$. Besides, we assume that any player can be influenced by noise to take the opposite action with the probability p_n in each stage. In experiments let $R = 3$, $T = 5$, $S = 0$ and $P = 1$ [1]. But our analytical proof (see in Appendix S1) illustrates the effectiveness of soft control under complete interaction for arbitrary R,S,T,P which satisfy $T > R > P > S$ and $R > (T + S)/2$.

Survival of the fittest

Actually Eq. (2) reflects the idea of "survival of the fittest", i.e. the more payoff one player gets, the more offspring it reproduces. Because shills are assumed to pose as normal agents, we first study the case that shills are also subject to "survival of the fittest". In this scenario, we define the frequency of cooperation f_c as the fraction of cooperation taken by players (i.e. normal agents and shills) in all games of one generation.

The simulation results (Fig. 1) demonstrate that no matter in the short-term ($\beta = 10$) or long-term ($\beta = 100$) RPD, even though there is a small proportion (not less than 5% in the figure) of shills in the population, they will become the majority at last. Thus f_c mainly derives from shills' action. So the cooperation level can be high since shills like to cooperate when the opponent cooperates. Soft control seems effective in this sense. But it is mainly due to the fact

that shills win the game of "survival of the fittest" and replace normal agents. This is not so fair since shills get more information than normal agents. So we restrict the number of shills N_S to be constant in following parts of simulations to see how soft control works. Therefore, f_c is defined as the fraction of cooperation taken by normal agents in all games of one generation.

Evolution of f_c and strategies

Fig. 2 demonstrates the performance of soft control with various N_S. When $N_S = 0$, normal agents with smaller p and q (i.e. less likely to cooperate when the opponent defects or cooperates in the last move respectively) get more payoff, which leads to the prevalence of defection. When defection prevails, p is more important than q on determining a normal agent's payoff. So the red line in Fig. 2 (A) fits to the red line in Fig. 2 (C) well. Comparatively when $N_S = 450$, there are sufficient shills to make normal agents with larger q get more payoff by cooperating with them. Thus cooperation is beneficial such that cooperation dominates defection. Interestingly note that when $N_S = 150$, f_c has a first decrease and then increases. The reason is that although cooperation is sustained by shills all the time, in the first period the number of shills is not large enough to ensure cooperation more profitable, which leads to the dominance of defection. But later, defection is no longer advantageous. On one hand defection is not supported by shills; on the other hand, playing defection only receives P points rather than T points in most interaction due to the prevalence of defection. But by contrast cooperation is more beneficial because it is supported by shills. Consequently f_c increases after the first period.

Above results indicate that after adding shills, cooperation is promoted. In the following part, we study soft control under other
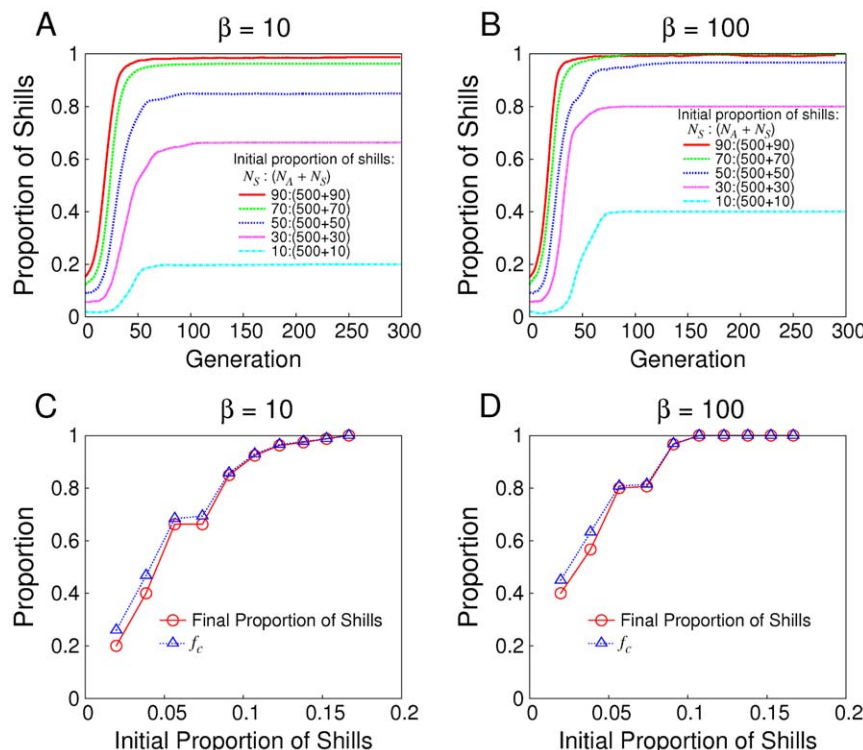

Figure 1. Shills are subject to survival of the fittest. (A) & (B) how the proportion of shills changes with different initializations when β is 10 and 100 respectively. (C) & (D) the relationship between the proportion of shills and f_c on $t = 300$ with different initializations when β is 10 and 100 respectively.
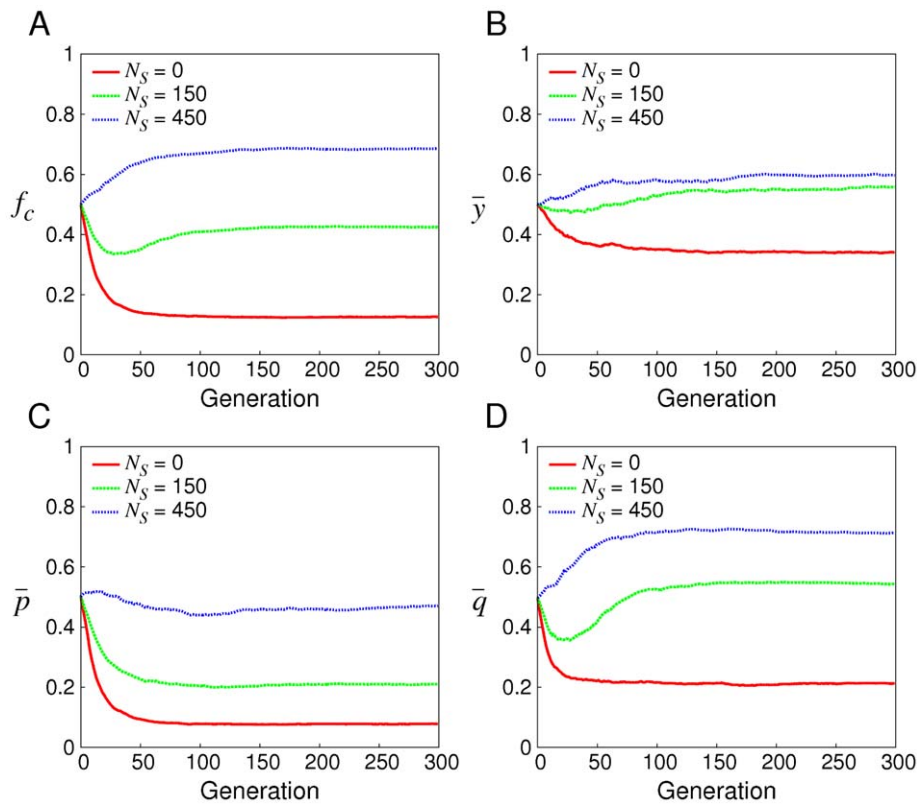doi:10.1371/journal.pone.0029182.g001

Figure 2. The evolution of f_c and strategies with different N_S. When $p_n = 0$ and $\beta = 10$, the evolution of f_c, $\bar{y}$, $\bar{p}$ and $\bar{q}$ are demonstrated as N_S varies, where $\bar{y} = \sum_{i=1}^{N_A} y_i / N_A$, $\bar{p} = \sum_{i=1}^{N_A} p_i / N_A$ and $\bar{q} = \sum_{i=1}^{N_A} q_i / N_A$.
doi:10.1371/journal.pone.0029182.g002

settings, which include the short-term $(\beta = 10)$ vs. long-term $(\beta = 100)$ RPD, noise-free $(p_n = 0)$ vs. noisy $(p_n = 0.05)$ interaction, and sharing vs. non-sharing knowledge.

Different settings

Simulation results (Fig. 3 (A) & (B)) illustrate the robustness of the mechanism to noise. We find that soft control is slightly sensitive to noise. It is because the strategy F-TFT is on a basis of shared knowledge but noise causes shills' knowledge to be inaccurate. Also shills' own action is subject to noise. But mixed reactive strategies contain randomness, so noise in the interaction does not have a significant impact on the performance. In the meantime, we find that soft control is still efficacious to promote cooperation no matter in the short-term or long-term RPD. At this point, soft control is robust.

In order to evaluate the importance of knowledge on soft control, we compare the difference between sharing knowledge and non-sharing knowledge among shills for both the short-term and long-term RPD (Fig. 3 (C) & (D)). For the short-term RPD, sharing knowledge is better. Otherwise a shill does not have enough knowledge to estimate accurately the cooperativity of normal agents. In this situation, shills need to help each other, so sharing knowledge is crucial. However for the long-term RPD, this difference is no longer evident. It is because $\beta = 100$ is sufficient for a shill to estimate its opponents even without knowledge providing from other shills. Thus sharing knowledge is not essential in this case. As a whole, sharing knowledge is rudimentary for the short-term RPD while it becomes dispensable for the long-term RPD.

Additionally, note that there is an inversely proportional relationship between β and N_S. That is, to attain a given f_c, the required N_S decreases as β grows. The reason is that, for smaller β there have to be more shills to accumulate enough knowledge to estimate a normal agent accurately. Therefore as long as β is sufficiently large, theoretically one shill can promote cooperation of the group.

Incomplete population interaction

Above discussions are made in the complete population interaction case. But in real world systems it is not always like that. We should also consider how soft control works in the case of incomplete interaction, that is, players can interact with a proportion of the population. This proportion is denoted by $\alpha \in \mathbb{R}[0,1]$ called the interaction locality (in the case of complete interaction, the proportion α is equal to 1). In one generation, player i $(i \in \mathcal{P})$ is chosen at random and then it selects another one from $\mathcal{F}_i$ to play the β-stage RPD once, where $\mathcal{F}_i$ denotes as the set of players that player i has never interacted with in the current generation. For normal agents, because they have no knowledge of others, their selection is random. But for shills, they can share knowledge and make full use of it. In this case, each shill k $(k \in \mathcal{P} \backslash \mathcal{A})$ keeps its own knowledge (m_i^k, n_i^k) for normal agent i where $i \in \mathcal{A}$. Shill k prefers to choose normal agents whose cooperative level (judged by n_i^k / m_i^k, according to its knowledge) is higher than a threshold, $\delta \in \mathbb{R}[0,1]$, called the selection level. The set of these "qualified normal agents" is denoted as $\mathcal{G}_k$. Shill k randomly selects a normal agent from $\mathcal{F}_k \cap \mathcal{G}_k$ if not empty; otherwise it chooses from $\mathcal{F}_k$ at random. After interacting with a normal agent, shill k shares its knowledge with a proportion of other randomly chosen shills. This proportion is called the share proportion, denoted as p_s. Above selection and interaction
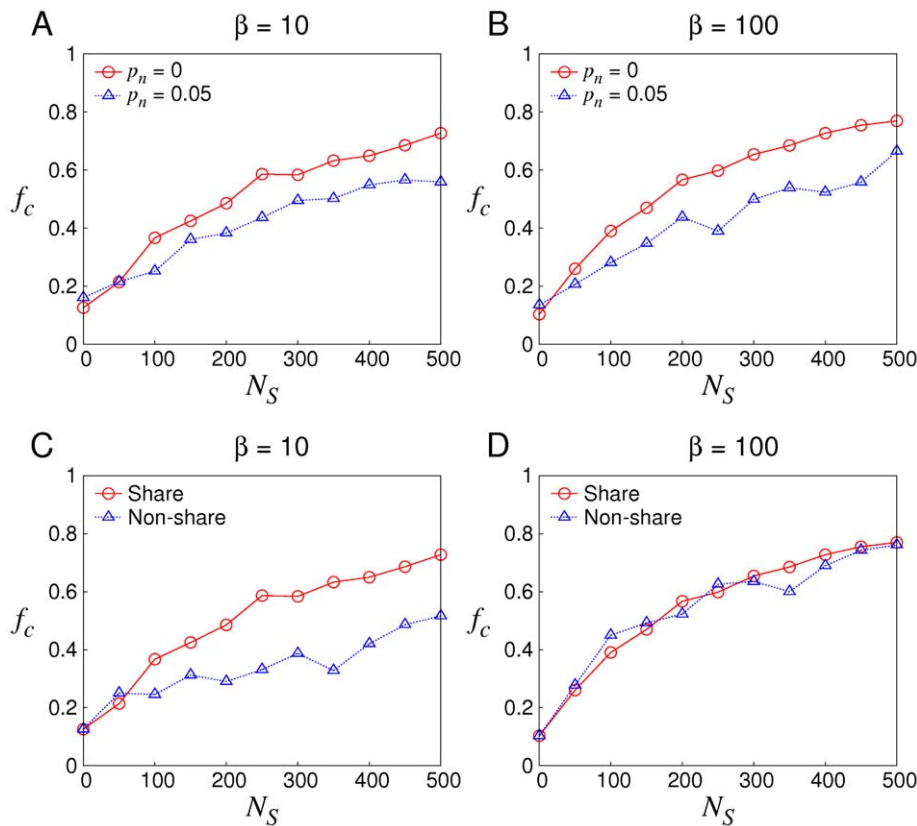
Figure 3. The performance of soft control under different settings. f_c on $t = 300$ is shown because mostly after 100th generation, f_c is convergent. (A) & (B) the sensitivity of the mechanism to noise when β is 10 and 100 respectively. (C) & (D) the importance of sharing knowledge when $p_n = 0$, $\beta = 10$ and $p_n = 0$, $\beta = 100$ respectively.
doi:10.1371/journal.pone.0029182.g003

processes repeat until on average each player interacts with $\alpha \cdot N$ players. Then they reproduce offspring based on Eq. (2).

Fig. 4 (A) illustrates the efficacy of soft control under incomplete interaction. And it also demonstrates an inversely proportional relationship between β and N_S. Note that compared to the case of complete interaction, cooperation is sustained for much smaller β and N_S because shills' knowledge are used to choose opponents from normal agents as well. This dramatically enhances the performance of soft control.

In addition, the parametric sensitivity of soft control is studied for different values of δ, p_s and α. Fig. 4 (B) shows that for δ, there is a tradeoff in the selection scheme of shills, which is similar to the relationship between exploration and exploitation. When δ is large, there are few normal agents getting benefits from shills; whereas when δ is small, shills' selection is almost random such that cooperative behaviors cannot get more rewards than defective ones. As a result playing cooperation is not advantageous. We also find that even though a shill shares its knowledge with a small proportion of other shills, soft control is still effective. Thus shills do not need to share with all others to promote cooperation (Fig. 4 (C)). It is noted that in Fig. 4 (D) as α increases, incomplete interaction degenerates into complete interaction gradually such that shills lose the advantage on selection. Thus for a given f_c, the required N_S is proportional to α.

Mutation

We know that randomness derives not only from the strategy *per se* and noise in the interaction, but sometimes from strategy

reproduction. So in the case of incomplete interaction, we investigate how soft control performs if randomness exists in strategy reproduction. Here y,p,q in a normal agent's strategy are represented as 10-bit binary string apiece. During reproduction, each bit in the string mutates from 0 to 1 or from 1 to 0 with the probability p_m which is called the mutation probability. In Fig. 5 (A)–(D), we find that soft control can still promote cooperation when the order of magnitude of p_m is no larger than 10^{-2}. Meanwhile note that with the increase of p_m, the capability of the mechanism becomes worse. This is due to the fact that for larger p_m, offspring are more different from their predecessor. As a result, any possible equilibrium becomes unstable any longer. Hence as long as p_m is not very large, cooperation is always promoted by adding shills. In Fig. 5 (E) & (F) it can be found that rare mutation (p_m is not larger than 10^{-3}) in reproduction is beneficial to increase the capability of soft control. In fact, small p_m can increase the diversity of the strategy space such that there would be a possibility to incorporate higher cooperativity of normal agents while it does not destroy the established equilibrium.

Discussion

In this paper we propose a mechanism called soft control to promote cooperation. For a group of agents playing the finitely but end-unknown RPD, the self-organized evolution of the population does not favor cooperation. However simulation results show that cooperation is promoted after introducing shills without violating play rules in the original group. Meanwhile the
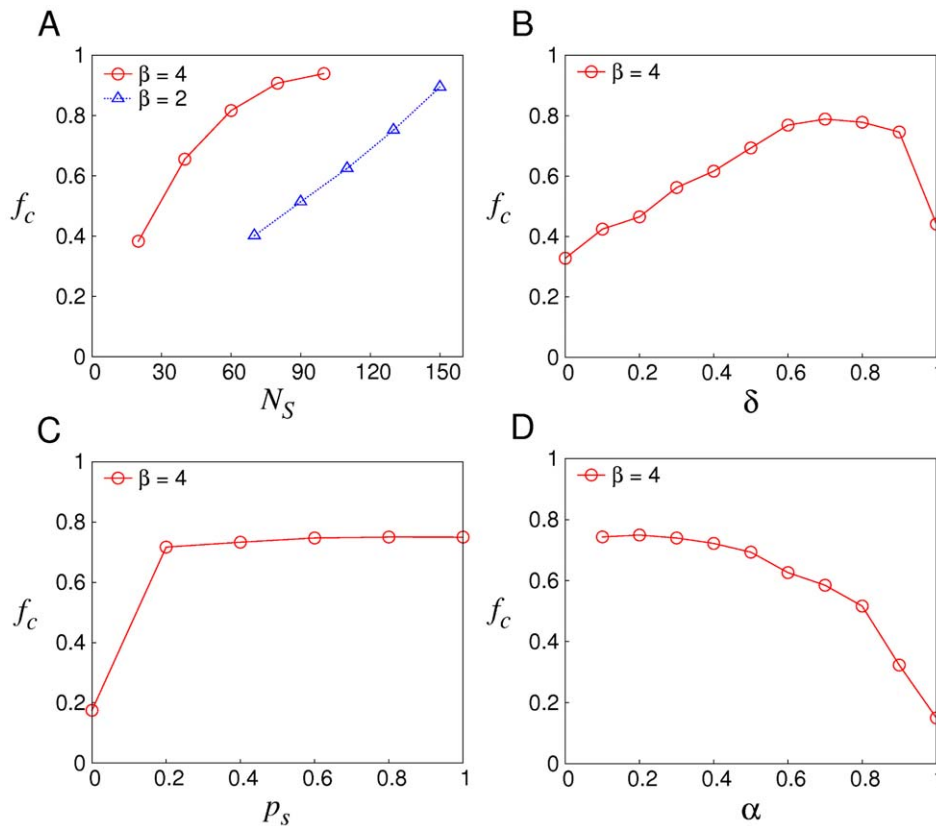
Figure 4. Effectiveness of soft control and its parametric sensitivity under incomplete population interaction. In this case if there is no extra declaration, parameters are $\alpha = 0.2$, $\delta = 0.9$, $p_n = 0$ and $p_s = 1$. f_c on $t = 300$ is demonstrated. (A) the relationship between f_c and N_S when β is 2 and 4 respectively. (B)(C)(D) the sensitivity to δ, p_s and α where $N_S = 50$ and $\beta = 4$.
doi:10.1371/journal.pone.0029182.g004

performance of soft control is studied under different settings, which include short-term vs. long-term RPD, noise-free vs. noisy, and complete vs. incomplete population interaction. We find that the mechanism is slightly sensitive to noise but still effective. At this point, soft control is robust to noise. In the short-term RPD, sharing knowledge is essential to shills while it becomes unimportant in the long-term RPD. Cooperation can be promoted by shills in both complete and incomplete interaction case. Yet with selection based on knowledge, it is more efficient for shills to promote cooperation in the incomplete interaction case than in the complete interaction case. We find that shills also perform well in both the complete and the incomplete interaction case even with mutation in strategy reproduction, and rare mutation is beneficial to increase the capability of soft control. Our results demonstrate that, to achieve a given cooperation level, the required number of shills is inversely proportional to the time period of games, but proportional to the interaction locality. In addition the effectiveness of soft control under complete interaction is proven analytically in the appendix.

There are several literatures relevant to the intervention in individual behaviors [60,61]. In [60], authors investigate how a teacher guides a learner to cooperate in the Prisoner's Dilemma. The role of the teacher is similar to that of a shill, but they focus on the learning scheme in 2-agent games. Authors in [61] explore the effect of three different kinds of special agents, namely radicals, revolutionaries and reactionaries, on the transition of regimes. It has dramatically changed the transition time. Different from shills, those agents utilize strategies without feedback knowledge.

In our study it is required that shills should pose as normal agents by complying with play rules. This is the main point of soft control, to keep play rules in the original group unchanged. What is more, it has additional reasons in this paper: if shills are treated as special agents, a normal agent may behave differently on interacting with shills and other normal agents. It may pretend to be a cooperator in order to get benefits from shills, but act as a defector to exploit other normal agents. In this situation mutual defection is still the only consequence. Therefore we stress the importance of a shill being treated as a normal agent by truly normal ones.

This paper is the first step to study soft control in the well-mixed population, and more extensions deserve our further efforts to explore. It is interesting to study soft control based on other strategy sets besides reactive strategies, such as deterministic finite automata [56,62], look-up table [1], Turing machines [63] and neural networks [64]. In our further study, we will also investigate whether F-TFT is the best strategy for shills and the properties of the best strategy for the specific scenario.

We will extend soft control to structured populations and study the influence of different spatial structures (e.g. the regular network, the random network and the scale-free network) on the mechanism. In fact, the network topology appears in many real-world systems. The models with the spatial structure display different properties (such as pattern formation and diffusion [31,65]) from the mean-field type model. Consider soft control in the case of structured populations: except the number and the strategy of shills, we also need to decide which nodes (normal agents) these shills should link to and how many links there are for
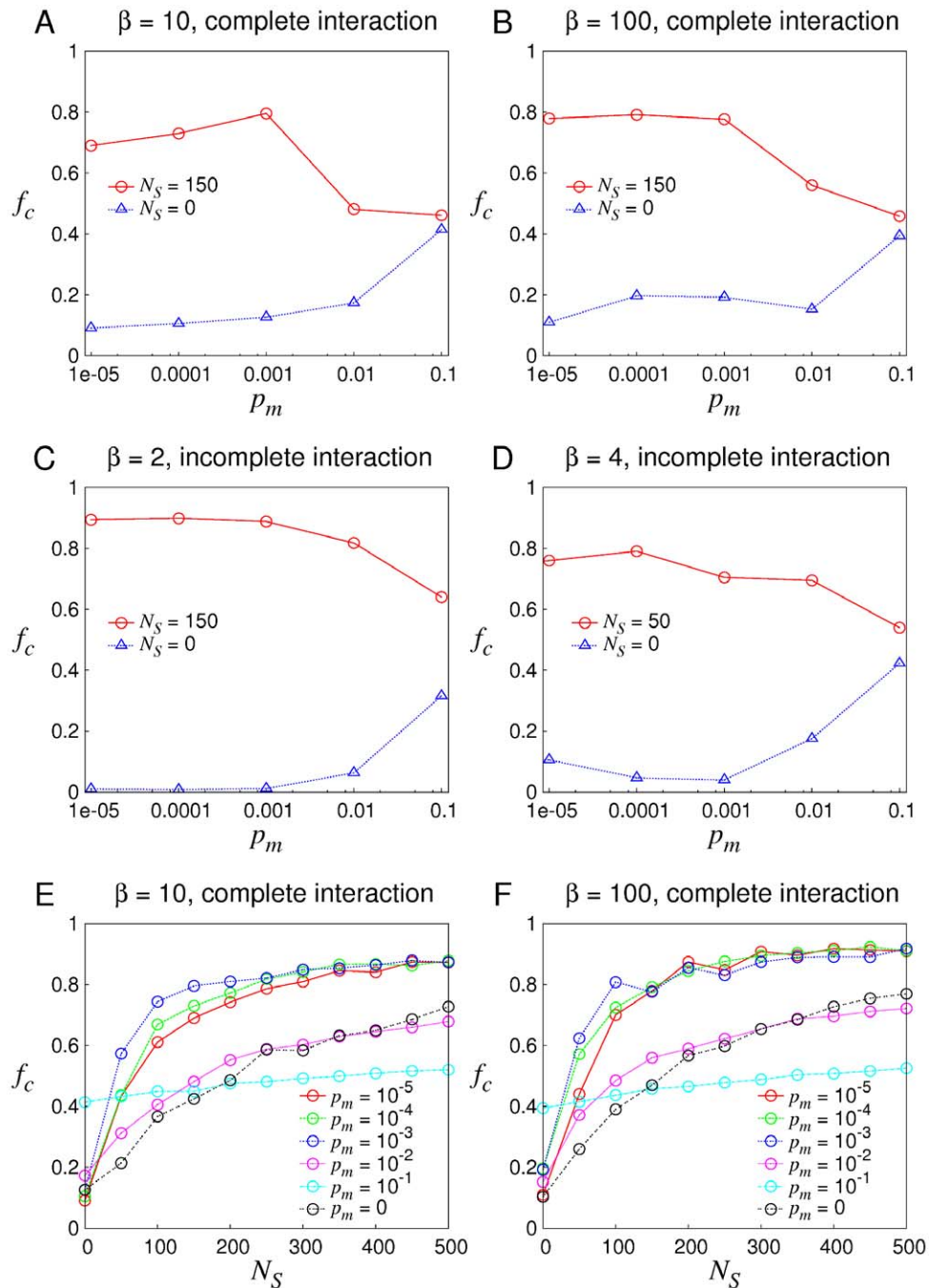
Figure 5. The effect of mutation in the reproduction. (A)–(D), the efficacy of soft control is demonstrated for p_m from 10^{-5} to 10^{-1}, where (A) & (B) are under complete interaction with $p_n=0$ and (C) & (D) are under incomplete interaction with the parameters $\alpha=0.2$, $\delta=0.9$, $p_n=0$ and $p_s=1$. (E) & (F) under complete interaction, f_c varies with the increase of N_S for different scales of p_m.
doi:10.1371/journal.pone.0029182.g005

each shill. Different networks might need different linking schemes. We know that in many networks, some nodes (such as hubs, nodes with high centrality, etc.) have more impact than the others on the overall performance. So it is crucial for shills to select important nodes to affect. The linking scheme will influence the performance of soft control. Notice that the importance of a node is also related to the dynamics of the system. So there might not exist a general heuristics of node selection for all systems. But some common principles might be discovered. On the other hand,

adding links will increase cost in some systems. The trade-off between the performance and the cost will be another important topic of soft control.

Soft control can be viewed as a way of intervention in collective behaviors. It does not focus on how to re-design play rules of every agent for the desired purpose, but on how to induce our desired collective behaviors without changing play rules. At this point, soft control provides a possible direction for the study of reciprocal behaviors and it may be applied to other scenarios like Public

Goods Game and Fashion Game, and to hinder the spread of panic/rumor in crowd or to control dynamical behaviors of other systems. Additionally it is necessary to study the applicability and limitation of soft control. Inspired by control theory [66], we will define and analyze the controllability of soft control in a general framework, i.e. to search for conditions that soft control can lead the system to the expected behavior. We believe that the controllability will relate to the jointly connectivity (or alike) of the system, which indicates every normal agent should be affected by shills directly or indirectly. We also believe that there will be a critical value of shill numbers or impact strength (which varies in different systems) to achieve the soft-control goal. Research following this line will provide a deep insight to soft control.

Supporting Information

Appendix S1 The proof of the effectiveness under complete population interaction.
(PDF)

Author Contributions

Conceived and designed the experiments: XW JH. Performed the experiments: XW. Analyzed the data: XW JH. Wrote the paper: XW JH. Contributed mathematical proof: XW HH.

References

1. Axelrod R (1984) The evolution of cooperation. New York: Basic Books.
2. Nowak MA (2006) Evolutionary dynamics. Cambridge, MA: Harvard University Press.
3. Sigmund K (2010) The calculus of selfishness. PrincetonNJ: Princeton University Press.
4. Lombardo MP (1985) Mutual restraint in tree swallows: a test of the tit-for-tat model of reciprocity. Science 227: 1363–1365.
5. Milinski M (1987) Tit-for-tat in sticklebacks and the evolution of cooperation. Nature 325: 433–435.
6. Stephens DW, McLinn CM, Stevens JR (2002) Discounting and reciprocity in an iterated prisoner's dilemma. Science 298: 2216–2218.
7. Greig D, Travisano M (2004) The prisoner's dilemma and polymorphism in yeast suc genes. Proceedings of the Royal Society of London (series B) 271(3): 525–526.
8. Krams I, Krama T, Igaune K, Mänd R (2008) Experimental evidence of reciprocal altruism in the pied flycatcher. Behavioral Ecology and Sociobiology 62: 599–605.
9. Yao X, Darwen PJ (1998) An experimental study of n-person iterated prisoner's dilemma games. Informatica 18(4): 435–450.
10. Nowak M, Sigmund K (1992) Tit for tat in heterogeneous populations. Nature 355: 250–252.
11. Nowak M, Sigmund K (1993) A strategy of win-stay, lost-shift that outperforms tit-for-tat in the prisoner's dilemma game. Nature 364: 56–58.
12. Helbing D, Szolnoki A, Perc M, Szabó G (2010) Defector-accelerated cooperativeness and punishment in public goods games with mutations. Physical Review E 81: 057104.
13. Szolnoki A, Szabó G, Perc M (2011) Phase diagrams for the spatial public goods game with poolpunishment. Physical Review E 83: 036101.
14. Sigmund K, Hauert C, Nowak MA (2001) Reward and punishment. Proceedings of National Academy of Sciences USA 98: 10757–10762.
15. Ohtsuki H, Iwasa Y, Nowak MA (2009) Indirect reciprocity provides only a narrow margin of efficiency for costly punishment. Nature 457: 79–82.
16. Santos MD, Rankin DJ, Wedekind C (2011) The evolution of punishment through reputation. Proceedings of the Royal Society of London (series B) 278: 371–377.
17. Boyd R, Gintis H, Bowles S, Richerson PJ (2003) The evolution of altruistic punishment. Proceedings of National Academy of Sciences USA 100: 3531–3535.
18. Bowles S, Gintis H (2004) The evolution of strong reciprocity: cooperation in heterogeneous populations. Theoretical Population Biology 65(1): 17–28.
19. Fehr E, Gächter S (2002) Altruistic punishment in humans. Nature 415: 137–140.
20. Fehr E, Rockenbach B (2003) Detrimental effects of sanctions on human altruism. Nature 422: 137–140.
21. Gürerk Ö, Irlenbusch B, Rockenbach B (2006) The competitive advantage of sanctioning institutions. Science 316: 108–111.
22. Rand DG, Armao JJ, Nakamaru M, Ohtsuki H (2010) Anti-social punishment can prevent the co-evolution of punishment and cooperation. Journal of Theoretical Biology 265(4): 624–632.
23. Szolnoki A, Perc M (2010) Reward and cooperation in the spatial public goods game. Europhysics Letters 92: 38003.
24. Riolo RL (1997) The effects and evolution of tag-mediated selection of partners in populations playing the iterated prisoner's dilemma. In: Proceedings of the 17th International Conference on Genetic Algorithm. pp 378–385.
25. Riolo RL, Cohen MD, Axelrod R (2001) Evolution of cooperation without reciprocity. Nature 414: 441–443.
26. Vainstern MH, Silva ATC, Arenzon JJ (2007) Does mobility decrease cooperation? Journal of Theoretical Biology 244: 722–728.
27. Droz M, Szwabinski J, Szabó G (2009) Motion of influential players can support cooperation in prisoner's dilemma. European Physical Journal B 71: 579–585.
28. Helbing D, Yu W (2009) The outbreak of cooperation among success-driven individuals under noisy conditions. Proceedings of the National Academy of Sciences, USA 106: 3680–3685.
29. Chen Z, Gao JX, Cai YZ, Xu XM (2011) Evolution of cooperation among mobile agents. Physica A 390(9): 1615–1622.
30. Axelrod R, Dion D (1988) The further evolution of cooperation. Science 242: 1385–1390.
31. Nowak M, May R (1992) Evolutionary games and spatial chaos. Nature 359: 826–829.
32. Nowak MA, Bonhoeffer S, May RM (1994) Spatial games and the maintenance of cooperation. Proceedings of National Academy of Sciences USA 91: 4877–4881.
33. Szabó G, Töke C (1998) Evolutionary prisoner's dilemma game on a square lattice. Physical Review E 58: 69–73.
34. Lieberman E, Hauert C, Nowak M (2005) Evolutionary dynamics on graphs. Nature 433: 312–316.
35. Santos FC, Rodrigues JF, Pacheco JM (2006) Graph topology plays a determinant role in the evolution of cooperation. Proceedings of the Royal Society of London (series B) 273: 51–55.
36. Ohtsuki H, Nowak M (2007) Direct reciprocity on graphs. Journal of Theoretical Biology 247: 462–470.
37. Szabó G, Fáth G (2007) Evolutionary games on graphs. Physics Reports 446: 97–216.
38. Lozano S, Arenas A, Sánchez A (2008) Mesoscopic structure conditions the emergence of cooperation on social networks. PLoS ONE 3(4): e1892.
39. Woelfing B, Traulsen A (2009) Stochastic sampling of interaction partners versus deterministic payoff assignment. Journal of Theoretical Biology 257: 689–695.
40. Hammond RA, Axelrod R (2006) Evolution of contingent altruism when cooperation is expensive. Theoretical Population Biology 69: 333–338.
41. Fu F, Hauert C, Nowak MA, Wang L (2008) Reputation-based partner choice promotes cooperation in social networks. Physical Review E 78: 026117.
42. Szolnoki A, Perc M, Danku Z (2008) Making new connections towards cooperation in the prisoner's dilemma game. Europhysics Letters 84: 50007.
43. Segbroeck SV, Santos FC, Lenaerts T, Pacheco JM (2009) Reacting differently to adverse ties promotes cooperation in social networks. Physical Review Letters 102: 058105.
44. Moyano LG, Sánchez A (2009) Evolving learning rules and emergence of cooperation in spatial prisoners dilemma. Journal of Theoretical Biology 259: 84–95.
45. Szolnoki A, Perc M (2009) Resolving social dilemma on evolving random networks. Europhysics Letters 86: 30007.
46. Szolnoki A, Perc M (2009) Emergence of multilevel selection in the prisoner's dilemma game on coevolving random networks. New Journal of Physics 11: 093033.
47. Szolnoki A, Perc M (2009) Promoting cooperation in social dilemmas via simple coevolutionary rules. European Physical Journal B 67: 337–344.
48. Perc M, Szolnoki A (2010) Coevolutionary games — a mini review. Biosystems 99: 109–125.
49. Cardillo A, Gómez-Gardeñes J, Vilone D, Sánchez A (2010) Co-evolution of strategies and update rules in the prisoner's dilemma game on complex networks. New Journal of Physics 12(10): 103034.
50. Han J, Guo L, Li M (2005) Guiding a group of locally interacting autonomous. In: Proceedings of the 24th Chinese Control Conference. pp 184–187.
51. Han J, Guo L, Li M (2006) Soft control on collective behavior of a group of autonomous agents by a shill agent. Journal of Systems Science and Complexity 19: 54–62.

52. Wang X, Han J (2011) Special agents can promote cooperation. In: Proceedings of the 30th Chinese Control Conference. pp 5764–5768.

53. Fogel DB (1993) Evolving behaviors in the iterated prisoners dilemma. Evolutionary Computation 1(1): 77–97.

54. Cressman R (1996) Evolutionary stability in the finitely repeated prisoner's dilemma game. Journal of Economic Theory 68(1): 234–248.

55. Sandholm TW, Crites RH (1996) Multiagent reinforcement learning in the iterated prisoners dilemma. Biosystems 37(1): 147–166.

56. Miller JH (1996) The coevolution of automata in the repeated prisoner's dilemma. Journal of Economic Behavior and Organization 29: 87–112.

57. Nowak M, Sigmund K (1989) Oscillations in the evolution of reciprocity. Journal of Theoretical Biology 137: 21–26.

58. Nowak M (1990) An evolutionarily stable strategy may be inaccessible. Journal of Theoretical Biology 142: 237–241.

59. Ohtsuki H (2004) Reactive strategies in indirect reciprocity. Journal of Theoretical Biology 227(3): 299–314.

60. Brafman RI, Tennenholtz M (1996) On partially controlled multi-agent systems. Journal of Artificial Intelligence Research 4: 477–507.

61. Axtell RL, Chakravarty S. Radicals revolutionaries and reactionaries in a multi-agent model of class norms.

62. Rubinstein A (1986) Finite automata play the repeated prisoner's dilemma. Journal of Economic Theory 39: 83–96.

63. Binmore K (1987) Modeling rational players: part i. Economics and Philosophy 3: 179–214.

64. Darwen PJ, Yao X (2001) Why more choices cause less cooperation in iterated prisoners dilemma. In: Proceedings of the 2001 IEEE Congress on Evolutionary Computation. pp 987–994.

65. Cattaneo G, Dennunzio A, Farina F (2006) A full cellular automaton to simulate predator-prey systems. Lecture Notes in Computer Science 4173: 446–451.

66. Ogata K (1997) Modern Control Engineering. Upper Saddle RiverNJ: Prentice Hall.