Evolution of Cooperation Among Autonomous Agents in the Presence of Public Information

Srishti Patil
Supervisor: M.S. Santhanam

January, 2021

1 Introduction

In games like Prisoner's Dilemma and the Public Goods game, defection is theoretically the dominant strategy. Therefore, it is hard to sustain cooperation[1,2] in the classical repeated Prisoner's Dilemma game in a well-mixed population. This is often at odds with reality where cooperation is ubiquitous, and thus understanding the emergence and persistence of altruistic behaviour in this setting poses an interesting challenge. Evolutionary Game Theory provides a mathematical framework to explore this problem. Previous work on the topic[3,6] has focused on spatial reciprocity[4,5] and nearest neighbour interactions among agents. The relevant results have been summarized in Figure 1. These are reproduced from the paper by C. Hauert and G. Szabó.[3]

In other related works, incentivisation such as social rewarding[7–10] and punishment,[11–14] is seen to be an effective method to promote cooperation. Other mechanisms like memory effect among agents,[15] incomplete information,[16] external forcing[17] have proven effecitve. The aspect of public information is relatively unexplored and extremely relevant in such systems since the end-goal is to promote collective good.

The aim of this project is to understand the effect of public information[18] and reputation[19] in finitely repeated spatial games (mainly Prisoner's Dilemma). In Section 2, three new models are proposed - *Pdist*, *fcT* and *Rep*. Results are shown and discussed in Section 3. Section 4 explores in brief, an application of one of the models on a real-world problem. Further details about the models and results are listed in Appendix A

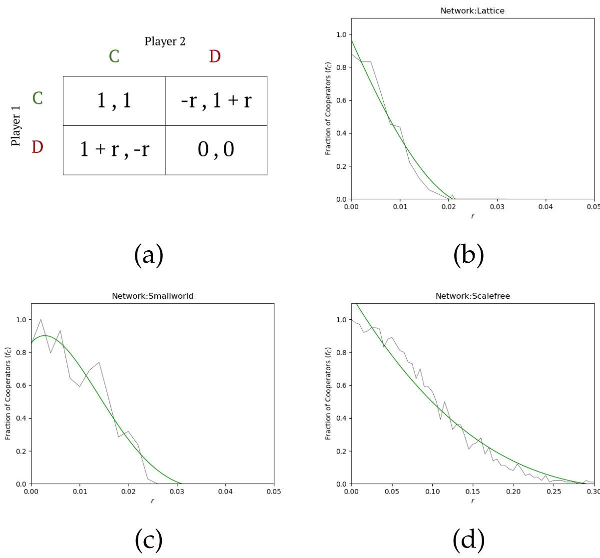


Figure 1: *(a)* The payoff matrix for Prisoner's Dilemma Game. r is the cost to benefit ratio. *(b)* f_C vs r plot for repeated prisoner's dilemma on a regular lattice network. f_C vanishes around $x \approx 0.02$ *(c)* f_C vs r plot for repeated prisoner's dilemma on a smallworld network. f_C vanishes around $x \approx 0.03$ *(d)* f_C vs r plot for repeated prisoner's dilemma on a scalefree network. f_C vanishes around $x \approx 0.3$

2 Models

Three models have been introduces as mechanisms to understand and enhance cooperation among agents playing the Prisoner's Dilemma game on different network topologies (Regular Lattice, Scalefree, Small-World). Each node represents an agent and each agent has two actions (strategies) - Cooperation and Defection.

Every iteration (Monte Carlo Step) of the game allows, on an average, one agent (agent i) to update its strategy. The rule of updation depends on the model being used. If the agents update their strategies by imitating a randomly chosen nearest neighbour (agent j) with a Fermi-Dirac like transition probability function (eq 1) depending on the difference in payoffs (as seen in the existing literature),[20]

$$W(s_i \leftarrow s_j) = \frac{1}{1+\exp[(p_i - p_j)/k]} \tag{1}$$

we get the results shown in Figure 1 (replicated results from[3]). p_i is the payoff of agent i in the current round and k is the noise parameter (here, taken to be 0.1). It can be observed that cooperation in the system dies out for low values of parameter r (ratio of cost to benefit of cooperation).

2.1 Probability Distribution Model (PDist)

Instead of assigning strategies to agents at the start of the game, we assign for each agent, a probability of adoption to each strategy. At each time step (MCS) all agents update this probability distribution (one by one).

For agent i, the probabiliity of adopting strategy s at time step t ($P_t^s(i)$) is given by equation 2, where s is the strategy that was adopted in the previous timestep. ΔP can take a range of values, both positive and negative (see A.1)

$$\begin{aligned} P_t^s(i) &= P_{t-1}^s(i) + \Delta P \\ P_t^{-s}(i) &= P_t^s(i) \end{aligned} \tag{2}$$

In this model, the focal agent increases the probability of adoption of the curent strategy (s) if the payoff is greater than the average payoff in the agent's local neighbourhood.

2.2 fc-Threshold Model (fcT)

To introduce public information to the system, we consider f_C - the fraction of cooperators in the population. If this information is available to the focal agent (this event occurs with probability P_{f_C}), then the updated probability distribution depends on the difference between f_C and the focal agent's personal threshold$_{f_\mathrm{C}}$ ($\in [0,1]$).

$$\begin{aligned} P_t^C &= P_{t-1}^C (2 - P_{t-1}^C)^a \\ P_t^D &= 1 - P_t^C \end{aligned} \tag{3}$$

where $a = (f_\mathrm{C} - \text{threshold}_{f_\mathrm{C}})$ if $f_\mathrm{C} > \text{threshold}_{f_\mathrm{C}}$ and 0 otherwise.

If the focal agent does not have access to f_C, then the current strategy (s) is updated according to equation 4. a' depends on $p_i - p_{\text{avg}}$ (see A.2 for more details).

$$\begin{aligned} P_t^s &= P_{t-1}^s (2 - P_{t-1}^s)^{a'} \\ P_t^{-s} &= 1 - P_t^s \end{aligned} \tag{4}$$

Here, the parameter P_{f_C} is a measure of the availablility of public information and the parameter threshold$_{f_\mathrm{C}}$ (which is a randomly chosen value between [0,1]) indicates the "tendency" of the agent to use that information. This structure tries to model external influence and personal biases separately. The availibility of external information promotes cooperation as the knowledge of f_C gives an indication of how safe it is to cooperate, given that defection is the dominant strategy and cooperation is the risky one from a rational agent's perspective.

2.3 Reputation Model (Rep)

This model introduces reputation as a social reward mechanism. Every agent has a reputation value ($\in [0,1]$) that increases every time the agent chooses to cooperate (see A.3). At each timestep, the focal agent (randomly chosen) will imitate a nearest neighbour based on either the difference in payoffs (the original imitaiton model) or the difference in reputations. The parameter p decides the probability of choosing imitation based on reputations.

Public information is in the form of f_R - the fraction of population with a reputation above 0.9. All agents have access to this information. Similar to the *fcT* model, each agent has a "tendency" to use the information denoted by threshold$_{f_\mathrm{R}}$. If f_R is greater than this personal threshold, then the transition probabililty can be given by equation 5.

$$\begin{aligned} P_t^C &= \frac{1}{1+\exp[-\text{reputation}_{\textit{self}}/k_{r_1}]} \\ P_t^D &= 1 - P_t^C \end{aligned} \tag{5}$$

If f_R is less than the personal threshold, then the agent resorts to local neighbourhood information (see Figure 2). It will pick the agent with maximum reputation and imitate its strategy with transition probability (eq 6).

$$W(s_i \leftarrow s_j) = \frac{1}{1+\exp[\text{difference}/k_{r_2}]} \tag{6}$$

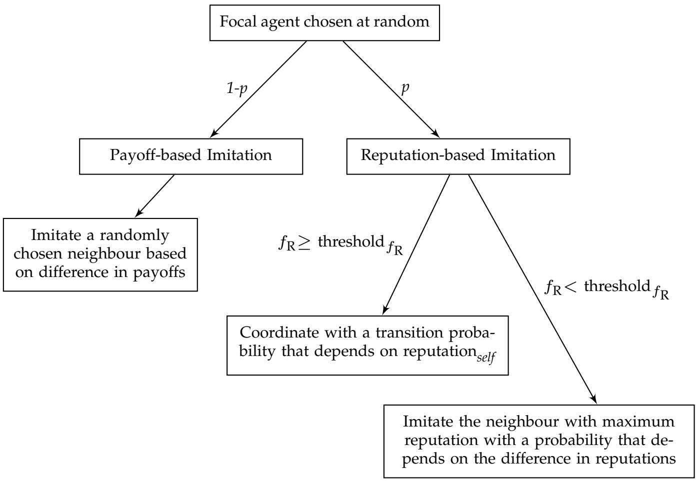


Figure 2: Flowchart for model *Rep*

$$\text{difference} = \text{reputation}_{self} - \text{reputation}_{max}$$

Therefore, if the maximum reputation (reputation$_{max}$) in the focal agent's neighbourhood is greater than the focal's reputation (reputation$_{self}$) then the probability of imitating that agent is high. Since reputation is an indication of how much an agent has cooperated in the past, it makes sense to "trust" that agent and imitate its strategy (s_j).

3 Results

All three models *(PDist, fcT and Rep)* enhance cooperation among agents. Figure 3 summarises the results for all the models discussed in the last section. We vary the parameter r across a range and observe if cooperative behaviour persists.

For model *PDist* (3a, 3b and 3c), we can see that a low fraction of cooperators survives for even large values of r. For values of parameter ΔP_{max} greater than 0.05, f_C vanishes at low r. For scale-free networks, a higher fraction survives. A possible reason for higher values of f_C at lower values of ΔP_{max} may be that for small increments or decrements in the probability distribution allow the system to reach an equilibrium.

Model *fcT* (3d, 3e, 3f) shows a drastic increase in the equilibrium value of f_C. For values of parameter $P_{f_C} > 0.7$, f_C is very close to 1 and stays that way for high values of r. To see the effect of public information, we can compare the blue curves ($P_{f_C} = 0$) with the rest.

Results for model *Rep* show that increase in probability of choosing reputation-based imitaiton (p) leads to an increase in f_C. 3g, 3h and 3i do not have any factor of public information. On choosing reputation-based imitation, the focal agent simply looks at its nearest neighbours. 3j, 3k and 3l show the effect of public information for a fixed value of $p = 0.3$.

The last row in Figure 3 shows how reputation of all the agents evolves with time. For $p = 0$, the values are spread out and have a lower average. For higher values of p, all the agents reach a reputation very close to 1. In general, we observe that introduction of a reward mechanism (reputation) and public information (in models *fcT* and *Rep*) promotes cooperation among agents.

4 Application in Opinion Dynamics

Social dilemma games like Prisoner's Dilemma and Public Goods game are inherently asymmetric in that one of the possible strategies is clearly more benefecial for the collective good. In opinion dynamics models (using EGT), the strategies (opinions) have an equal footing. By modifying the *Rep* model, we try to model the effect of public information and reputation on populations of agents that can choose between two opinions (A and B).

A randomly chosen focal agent, in each timestep, will imitate a nearest neighbour either based on the difference in payoffs or depending on the parameter x (eq 7). Imitation based on payoffs should promote the formation of clusters of agents with the same strategy (similar to what we see in[3]). Reputation increases in magnitude if an agent adopts the same strategy (sticks to it). It is a reward mechanism that favours one strategy over the other depending on the history of the focal agent. Reputation values range from -1 to 1. Negative values correspond to strategy B and positive to startegy A. See appendix A.4 for details about the transition function and how it depends on x. s and o are the reputations of the focal agent and the agent chosen for imitation respectively.

$$x = \text{sign}(s.o)\left|\frac{o}{s}\right| \tag{7}$$

Public information in this model is the fraction of agents with reputation above 0.9 (case I - $f_R(A)$) or below -0.9 (case II - $f_R(B)$).To observe the effect of public information, we introduce a new parameter P_A which is the probability of agents having access

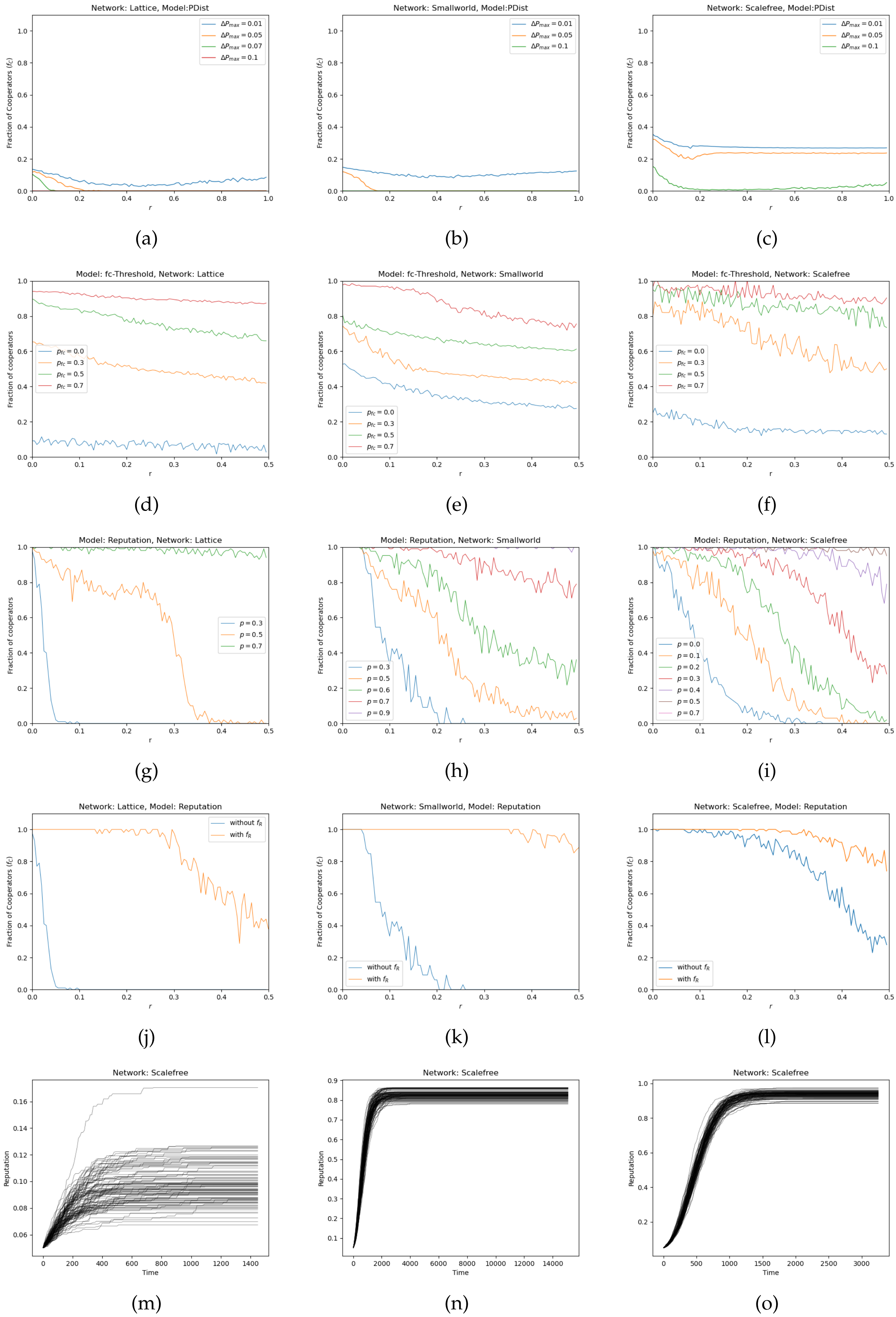


Figure 3: Row 1: Model *PDist*, Row 2: Model *fcT*, Row 3: Model *Rep* (varying parameter p), Row 4: Model *Rep* (effect of public information f_R), Row 5: Evolution of agent reputations with time (for scalefree network). All plots are generated by averaging over 100 runs on a population size of 100 agents. For smallworld networks, $Q = 0.6$. Note that due to technological and time constraints, 3j and 3k have been run 10 and 35 times respectively.

to $f_R(A)$. Agents who do not have access to it have access to $f_R(B)$ (see Figure 4). This parameter is a loose indicator of how well an opinion has been advertised to the masses i.e. the "popularity" of a strategy in the population. We vary this parameter to see how knowing about a strategy's popularity in the population affects the fraction of agents that adopt it in equilibrium.

Once the focal agent has access to this information, it will adopt the next strategy based on its personal threshold (similar to the one in model *Rep*, see A.4) and the sign of its reputation (startegy A if positive, B if negative). The higher the personal threshold, the higher is the probability of adopting the strategy.

The results for simulations on a regular lattice and a scalefree network are given in Figure 5. The flat orange line indicates that without public information playing a role, the effects of social reward mechanisms like reputation cancel out and we get equal number of agents adopting each strategy. Variation in the parameter P_A shows that with increase in public knowledge of a startegy, agents are more likely to adopt that strategy in the particular case where both the strategies are equal otherwise.

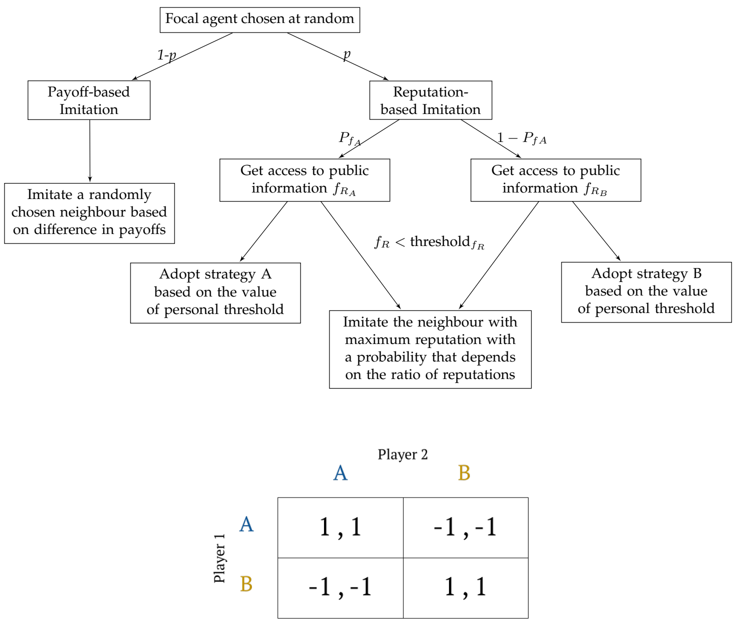


Figure 4: Flowchart and payoff matrix for the opinion dynamics model. An agent gets positive payoff for interactions with same strategy neighbours. This may lead to formation of clusters.

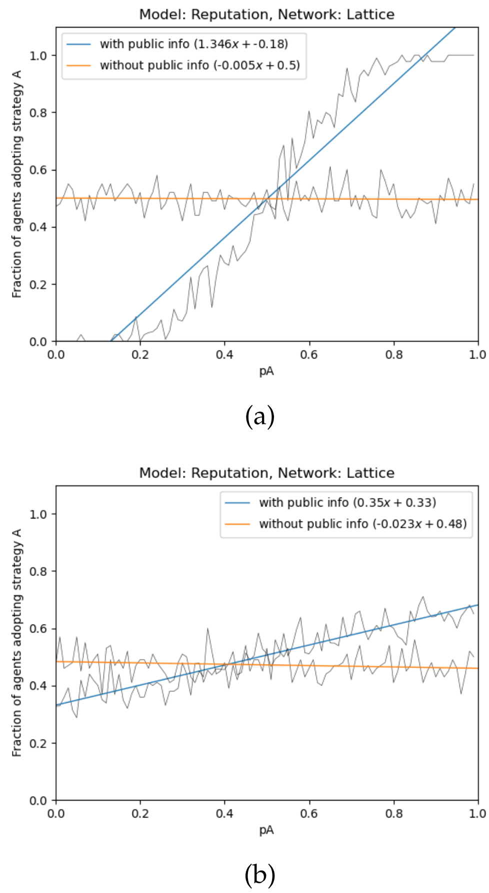


Figure 5: *(a)* Network: Lattice. Low values of P_A means that public information (f_R) related to strategy B is more likely to be used by the agents. High values of P_A show a very significant increase in the fraction of population adopting strategy A. *(b)* Network: Scalefree.

5 Conclusion

In this project, three new models that promote the emergence and persistence of cooperation in a structured population were explored. The effect of reputation as a social reward mechanism and public information as a trust mechanism on networked populations was observed. There was a significant increase in the equilibrium values of f_C in all the models. We applied one of the models (*Rep*) to a real-world situation. It was observed that increase in popularity of an opinion and its effective dissemination among the agents increases the fraction of agents that adopt that strategy (under equilibrium conditions). Further work on the topic can test how clusters are formed in the population and how the fraction of agents adopting a particular strategy would change with time.

A Appendix

A.1 Model *PDist*

For the *PDist* model, ΔP varies as a function of the difference between the focal agent's payoff and the average nearest neighbor payoff, p_{avg}. The value of k_{pd} for all the simulations is taken to be 1.5.

$$\Delta P = \Delta P_{\max}\left[\frac{2}{1+\exp[(p_i - p_{\text{avg}})/k_{pd}]} - 1\right] \tag{8}$$

A.2 Model *fcT*

When the focal agent does not have access to public information (f_{C}) then it has to use local information i.e. nearest neighbour payoffs. In equation 4, a' is given by

$$a' = \begin{cases} \dfrac{a'_{\max}}{1+\exp[p_i - p_{\text{avg}}/k_{fc}]} & p_i > p_{\text{avg}} \\ 0 & p_i \leq p_{\text{avg}} \end{cases} \tag{9}$$

p_{avg} is the average nearest neighbour payoff. The value of a' and subsequently P_t^s is high if the difference in the focal agent's payoff and p_{avg} is high. For all the simulations, the parameter values are taken to be as follows: $a'_{\max} = 0.2$, $k_{fc} = 1.5$.

A.3 Model *Rep*

We use the usual Fermi-Dirac like probability function for this model. Updates based on public information take into account the reputation of the agent so that agents with a history of cooperation keep cooperating in the future. Updates based on local information take into account the difference in reputation as agents are likely to be influenced by their nearest neighbours, especially by those who have a higher reputation.

Every time an agent chooses to cooperate, its reputation is increased. Reputation before and after increase are denoted by rep_i and rep_f.

$$\text{rep}_f = \text{rep}_i(2 - \text{rep}_i) \tag{10}$$

For all the simulations, the parameter values are taken to be as follows: $k_{r_1} = 0.25$, $k_{r_2} = 0.05$, $k = 0.1$ (for payoff-based imitation).

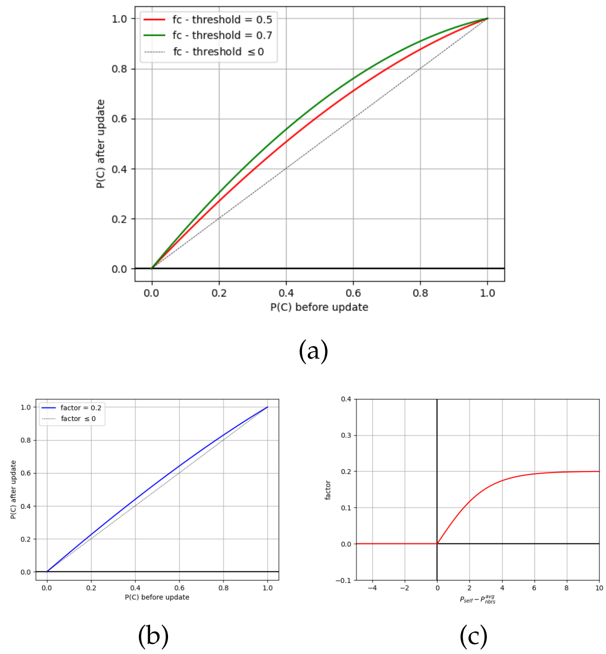


Figure 6: *(a)* Shows P_t^C before and after updation. The curves are for different values of ($f_{\text{C}}-\text{threshold}_{f_{\text{C}}}$). *(b)* show how P_t^C evolves in the absence of public information. *(c)* is the variation of factor a' with the difference in payoffs.

A.4 Opinion Dynamics Model (Futher Details)

The transition probability function for imitation based on reputation is given by equation 11.

$$W(s_i \leftarrow s_j) = \begin{cases} \dfrac{1}{1+\exp[\text{0.5-x}/k_{op_1}]} & x \geq 0 \\ \\ \dfrac{0.9}{1+\exp[\text{x+5}/k_{op_2}]} & x < 0 \end{cases} \tag{11}$$

We choose this particular functional form because it satisfies the following conditions -

1. Reaches maximum at $+1$: this is so that any agent that encounters another agent with similar reputation will imitate it with a high probability.

2. Value at -1 is very low: this is so that the probability of imitating an agent with reputation equal in magnitude but opposite in sign is negligible

3. Probability is high for large negative values: this is to model the influence of high-reputation agents on low reputation individuals with an opposite sign

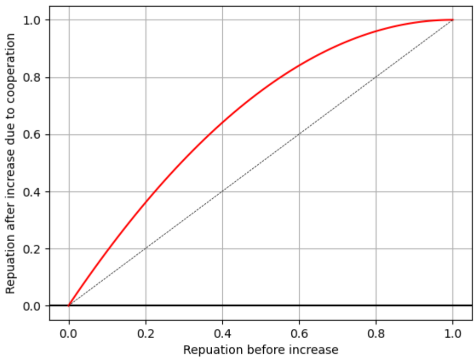
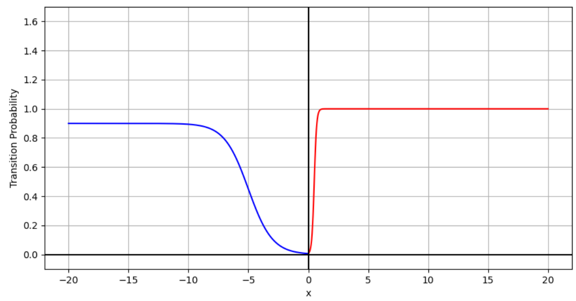


Figure 7: The graph shows the value of reputation before and after updation. It starts at 0 and tapers off near 1. The initial reputation for all agents is set to be 0.05

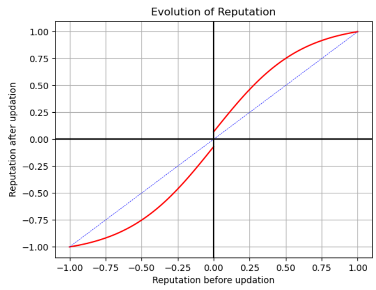
If the updation of strategy is based in public information (i.e. if f_R is greater than the personal threshold), then the probability of adopting a given strategy is given by equation 12.

$$P_t^s = \frac{1}{1 + \exp[-\text{threshold}_{f_R}/k_{op3}]} \qquad (12)$$
$$P_t^{-s} = 1 - P_t^s$$

In the above equation, s is A if the focal agent's reputation is positive and B if it is negative. For all the simulations, the parameter values are taken to be as follows: $k_{op_1} = 0.1$, $k_{op_2} = 1$, $k_{op_3} = 0.25$.



(a)



(b)

Figure 8: *(a)* is the transition probability function for reputation based imitation. We can see that it follows the 3 conditions listed. *(b)* shows how the reputation of an agent evolves. Reputation goes in only one direction (either positive or negative).

References

[1] Rational Cooperation in Finitely Repeated Prisoner's Dilemma. *Economics and Politics*.

[2] Garrett Hardin. The tragedy of the commons. *Journal of Natural Resources Policy Research*, 1(3):243–253, 2009.

[3] Christoph Hauert and György Szabó. Game theory and physics. *American Journal of Physics*, 73(5):405–414, 2005.

[4] Qi Su, Aming Li, Long Wang, and H Eugene Stanley. Spatial reciprocity in the evolution of cooperation. *Proceedings of the Royal Society B: Biological Sciences*, 286(1900):20190041, 2019.

[5] Peng Lu and Xiaoping Zheng. Social Stratification and Cooperative Behavior in Spatial Prisoners' Dilemma Games. *PLOS ONE*, 10(7):e0131005, 2015.

[6] Hirokazu Shirado and Nicholas A. Christakis. Locally noisy autonomous agents improve global human coordination in network experiments. *Nature*, 545(7654):370–374, 2017.

[7] Raúl Jiménez, Haydee Lugo, José A. Cuesta, and Angel Sánchez. Emergence and resilience of cooperation in the spatial prisoner's dilemma via a reward mechanism. *Journal of Theoretical Biology*, 250(3):475–483, 2008.

[8] Yu'e Wu, Shuhua Chang, Zhipeng Zhang, and Zhenghong Deng. Impact of Social Reward on the Evolution of the Cooperation Behavior in Complex Networks. *Scientific Reports*, 7(1):41076, 2017.

[9] Christoph Hauert. Replicator dynamics of reward & reputation in public goods games. *Journal of Theoretical Biology*, 267(1):22–28, 2010.

[10] Miguel dos Santos and Jorge Peña. Antisocial rewarding in structured populations. *Scientific Reports*, 7(1):6212, 2017.

[11] Xianjia Wang and Shaojie Lv. The public goods game with shared punishment cost in well-mixed and structured populations. *Journal of Theoretical Biology*, 476:36–43, 2019.

[12] The Anh Han and Tom Lenaerts. A synergy of costly punishment and commitment in cooperation dilemmas. *Adaptive Behavior*, 24(4):237–248, 2016.

[13] Chunyan Zhang, Yuying Zhu, Zengqiang Chen, and Jianlei Zhang. Punishment in the form of shared cost promotes altruism in the cooperative dilemma games. *Journal of Theoretical Biology*, 420:128–134, 2017.

[14] Fabio Dercole, Marco De Carli, Fabio Della Rossa, and Alessandro V. Papadopoulos. Overpunishing is not necessary to fix cooperation in voluntary public goods games. *Journal of Theoretical Biology*, 326:70–81, 2013.

[15] Wenwen Lu, Juan Wang, and Chengyi Xia. Role of memory effect in the evolution of cooperation based on spatial prisoner's dilemma game. *Physics Letters A*, 382(42-43):3058–3063, 2018.

[16] Xiuqin Deng and Jiadi Deng. A Study of Prisoner's Dilemma Game Model with Incomplete Information. *Mathematical Problems in Engineering*, 2015:1–10, 2015.

[17] Keke Huang, Yichi Zhang, Yonggang Li, Chunhua Yang, and Zhen Wang. Effects of external forcing on evolutionary games in complex networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28(9):093108, 2018.

[18] Hongsen Luo, Ying Gao, and Fulei Shi. An Evolutionary Game Analysis on Public Information Communication between the Government and the Public in China. *Complexity*, 2021:1–11, 2021.

[19] Zhen Wang, Lin Wang, Zi-Yu Yin, and Cheng-Yi Xia. Inferring Reputation Promotes the Evolution of Cooperation in Spatial Social Dilemma Games. *PLoS ONE*, 7(7):e40218, 2012.

[20] György Szabó and Gábor Fáth. Evolutionary games on graphs. *Physics Reports*, 446(4-6):97–216, 2007.