Complexity in Social Sciences

Achieving global cooperation in social networks through peer punishment

Alvaro Cabrejas Egea
1565937

May 4, 2016

Abstract

Achieving widespread cooperation while addressing global issues is one of the most severe problems in today's society. Once such a state has been reached it is usually stable, since global strategies maximise the expected average benefit for all the members of the system. In this paper we argue how allowing playing agents to exert only local punishment, this is, punish those defectors which are in direct contact with them, can trigger a cascade of collaboration leading to global cooperation. We conclude by analysing how the international community could use such approaches to take on global challenges such as climate change, public resources, investments and governmental corruption.

Keywords: Cooperation, Social Dilemma, Game Theory

1 Introduction

1.1 Scope

In the last decades we, as a species, are facing a kind of social dilemmas which differ tremendously from the typical issues societies had to deal with during history. Now, we are encountering problems that require that all human groups act as one, since the lack of action of one of them would jeopardise the efforts of the rest. Examples of these dilemmas would be industrial emissions, global warming, public investments, resources exploitation and even the treatment of corruption. The extension of this issues has now shifted from regional and national scales, where they could be addressed by relatively small actors, to international and worldwide scales that demand coordination from all participants in order to be addressed. This changes what is needed from the decision-taking groups,which are usually geographically confined to their own nation-state and cannot cover all locations involved and so, making collaboration between all the involved actors becomes a must. The setup of this problem is a classic example of the "tragedy of the commons", as popularised by Garret Hardin through the metaphor of herdsmen with access to common pasture land: each can always prosper individually by adding another head of cattle to his herd, but eventually this leads to overgrazing and ruin for all [1] [2].

1.2 Agent Based Modelling

In this essay we will explore, using Agent Based Modelling and game theory, how collaboration can be achieved spontaneously in a system where all players act selfishly and have a heterogeneous but general predisposition to exploit the system for their benefit, if the rules which they abide are oriented in the correct way. We will also explore how these dynamics are affected by information travel times through the community. Agent Based Modelling (ABM) is a dynamic modelling technique which is composed of:

1. Agents: The individual members of the model. They will be characterised by their traits and will act independently from each other, i.e. Bees of a hive.

2. Environment: It will depend on the model and encases the boundary conditions that define the interactions, essentially who can interact with who and under which conditions, i.e. One hive or a set of them.

3. Rules: They will define how interactions occur during the simulation, what the agents will need and do, i.e. Units of food needed per day in order to survive.

ABMs are advantageous because of three main characteristics. Firstly they can capture emergent phenomena, that seems to have no particular rule or force that would make it occur but none the less does happen. Secondly it provides a natural description of a system with rules and decision making that is easily understood and can be backed up with real life. Thirdly ABMs are flexible which is because the rules and decision making can be changed and adapted for different situations [3].

1.3 Game Theory

Game theory is "the study of mathematical models of conflict and cooperation between intelligent rational decision-makers" [4]. These decision-makers will be represented by the agents in a model, choosing what is most beneficial for them (in this case cooperate or defect) at a given time in order to maximise their reward or "payoff". In the scope of this essay we will understand punishment as negative payoff. In game theory, the tragedy of the commons is understood as a Nash equilibrium [5], where the game collapses to a stationary state in which no rational agent cooperates, despite cooperation would be the strategy which would maximise collective payoff if adopted by all players. Note that achieving a Nash equilibrium does not imply that everyone is benefited by the current state, but instead that all players are playing their best selfish strategies *given what the remaining players are doing.*

1.4 Real life experiments

One side of this problem that has not been tackled often is that different players will have different predispositions to cooperate and even different understanding of the game.

A group dynamic game used by Human Resources departments starts by splitting the group in two rooms preventing any communication, then they are asked to maximise the game's reward, without further explanations about whose reward. After explaining the rules, the participants must play the "Iterated Prisoner's Dilemma". Typical behaviour observed includes the two groups starting to adjust their strategies so they can defect in the most effective way to maximise the difference in accumulated reward by the end of the game [6]. It follows that normally both teams understand this reward maximisation as referred to *their* reward, since by being split in two rooms it is assumed that they are split into two different teams, while they are effectively one. The action of splitting makes the participants neglect that the only way to maximise their rewards is to actively cooperate with the other team to force a draw in which both teams will have achieved the maximum score. This can be easily related with social dilemmas in which, even if all players share the same environment or resource, they effectively identify themselves as part of two antagonistic factions.

2 Previous models

2.1 Iterated Prisoner's Dilemma

Previous ABM studies have dealt with this kind of interactions, specially related to those problems classified as social dilemmas - situations in which there is a benefit for everyone that follows from cooperation, but also a personal gain opportunity and temptation to defect. It started with different approaches to the "Iterated Prisoner's Dilemma", where each agent can adjust their strategies between rounds and conditional cooperation can be the best option for the individual as long as the players can remember each other [7] [8]. Different experiments and tournaments

The Prisoners' Dilemma

		Prisoner A Choices	
		Stay Silent	*Confess and Betray*
Prisoner B Choices	*Stay Silent*	Each serves one month in jail	Prisoner A goes free Prisoner B serves full year in jail
	Confess and Betray	Prisoner A serves full year in jail Prisoner B goes free	Each serves three months in jail

Figure 1: Examples the choices and payoffs in a single round of the Prisoner's Dilemma.

ended up showing that the best kind of strategy in this situation would be so called "Tit-for-tat with forgiveness" [9]. This strategy while not unbeatable, outperforms those not designed specifically to counter it. It operates based on the following set of rules:

1. The agent always collaborates, unless the other agent defects first.

2. The agent will take revenge for every defection it receives.

3. The agent forgives easily once it had its revenge.

4. There is a small probability of forgiving a defector by cooperating anyway. This prevents getting trapped in long cycles of retaliation.

The consistent victories obtained by this algorithm are based on (despite its name) wide cooperation and a degree of naiveness. It shows how collaborating and benefiting all sides on the process also maximises the benefit for all parties involved and takes a clever approach on the fact that a record score of any kind is not less valuable because it was achieved by more than one player in a draw, since it still provides the maximum aggregated payoff out of all possible alternatives.

The "Iterated Prisoners Dilemma", however, does not contain a big enough number of players, and the rules are too simple to be able to extrapolate strategies to more complex, real-world games, in which a simple winning strategy can become useless based on changes on the environment or other distant players, where our own actions are not powerful enough to force changes in those factors combined. Nevertheless, it makes a powerful show of how widespread cooperation tends to lead to higher global payoffs.

2.2 Public Goods Game

A better representation is the "Public Goods Game" [10]. Each player can choose how much "money" they put in a common pot which multiplies the global amount by some factor (bigger than one and smaller than the number of players) and redistributes the resulting amount equally between all players. Here, the maximum score can only be obtained if all players put all their money at every iteration. This situation, as the "Iterated Prisoner's Dilemma", generates a draw since all players will cooperate at every step and their earnings will be the same, resulting in all of them "winning" or all of them "losing". However, an individual player can "win" (be the one with the most accumulated wealth) by being the one that places the least aggregated amount of money in the pot over all rounds, since they will benefit of the "Public Goods" as much as any of the other players, but on top of that they will be able to keep their wealth. A side effect of this approach is that the total amount of money in the pot will be smaller than it could be if all players used all their money, so the total gain for all the players will be lower overall. Many experiments with humans have shown that cooperation can be enhanced by allowing the players to punish defectors, despite the punishers incurring a cost for doing so [11].

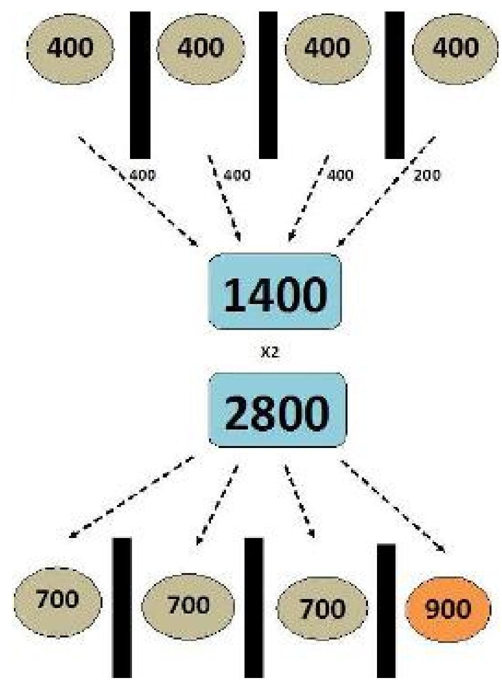


Figure 2: Schematic example of one round of the public goods game.

This model correctly accounts for the appearance of different parties with conflicting interests, making a better representation of the complexity of the real world, but the fact that individual agents can choose not only to cooperate or defect, but to which degree they do, makes impossible to assess the effect of local collaboration as a trigger to eliminate widespread defection i.e. An agent that does not pay all of its taxes is still defecting in the eyes of its peers, even if is collaborating to some degree in a system level.

2.3 Targeted Punishment

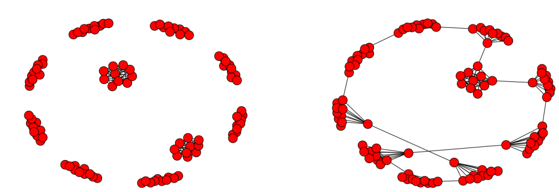
This model represents better the effect of targeted punishment among peers in the emergence of cooperation from a situation of widespread defection [12]. It has a set of players that can choose between cooperation and defection, but each player will have a different predisposition towards cooperation. Once the simulation begins, each agent will stoachastically choose to cooperate or defect based on a transition probability function controlled by two parameters describing the level of the punishment exerted by the collaborating agents towards the defectors and the rationality of the agents. For low rationality they will all behave randomly, neglecting their environment and for high rationality they will all deterministically choose what is best for them in the current situation. The model proposes different ways of punishment, from the general case for all defectors to specific targeted punishments only to certain members of the defecting group. It finds that with different punishment strategies, a small population of collaborators can shift a population of defectors into a state global cooperation. However, it is limited in the way that it allows any agent to punish any other, having all of them the same power or influence and infinite reach.

3 Methodology

3.1 Social topology, Caveman's Network

In order to adapt this punishment model to networks topologically equivalent to real-world social networks we decided to use a variation of the "Caveman Network" (CN) called a "Relaxed Caveman Network" (RCN).

In these, each agent is represented as a vertex, a line between two vertex represents a connection between those agents, in this case it means they can influence each other. Agents will play, collaborate or defect and punish each other when they find defectors among those connected to them (in their communities). This connectivity pattern successfully accounts that influence is not possible at all levels regardless of an individual's position, it becomes easier when one tries to influence those closer to them. In order to better approximate the network to reality, random versions of the RCN were used, using different probabilities for the rewiring parameter (see Appendix). Two different values have been chosen, such that the resulting structures have, on average, the same average path lengths (distance) as societies as it was theorised in the era pre-social networks [13] $length_1 \simeq 6$ and post-social networks [14] $length_2 \simeq 3.4$.



(a) (b)

Figure 3: Examples of Caveman Networks with 20 caves of 10 members: Completely connected isolated caves in which all members know each other (a) and a Relaxed Caveman Network in which there is always one link between different caves (b).

3.2 Peer Punishment Model

For the punishment modelling, it was set so each collaborating agent will equally split their punishment between those defectors in direct contact with them. This results in a great majority of intra-group punishment, only having meaningful inter-group punishment once the group has achieved consensus and all of the punishing power of those agents having inter-group connections is concentrated in one direction in order to "convince" their counterpart in the other group. Once the counterpart has changed their strategy towards cooperation, this can trigger a cascade of changes in the new group, sequentially extending to other groups.

Since the predispositions towards cooperation are randomised, the model also allows for kernels of defection-predisposed agents. If the community has a big external connectivity, it will be possible through external multi-front punishment to force them to collaborate. However, if there are few external connections, even the most rational agents will choose to defect as the only possible way not to be "taken advantage of" by the other defecting agents. We believe that this behaviour correctly represents those areas where crime or corruption are widespread, and individuals have to choose between imitation to lower their loss or abiding the rules and being penalised both by the system and their own neighbours. [15]

4 Discussion

Two sets of simulations were run for the model in order to experiment on the effect of the reduction of the average path length between individuals in the society, one was set to resemble the typical "six degrees of separation" setup while the other was set to model the current digital "three point four degrees of separation". The average cooperation between all the agents was measured at the end of each experiment, obtaining an average cooperation value for each combination of rationality and punishment in each of the two models, this was averaged over several runs.

4.1 The effect of punishment

As it can be seen from the figures 4, 5 and 6 we find the exact same behaviour regarding punishment and rationality as it was found before in reference [12]. However it is interesting to see how the difference in the network's average path length has a noticeable effect in the convergence precisely in the areas where the original model suffered from more reluctance to cooperation, those with mid-high punishment and ra-

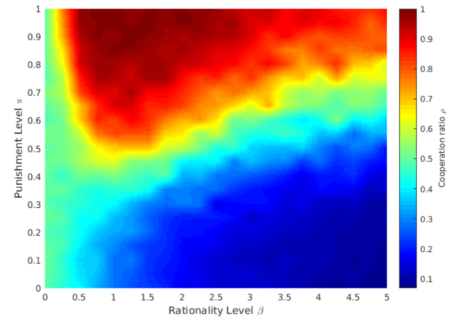


Figure 4: Results for the model with average path length 6

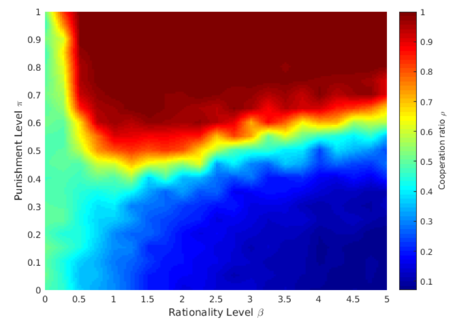


Figure 5: Results for the model with average path length 3.4

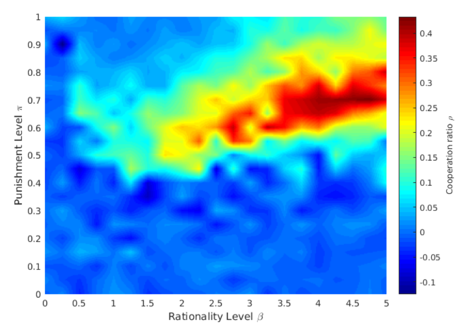


Figure 6: Results from the second model minus the results of the first model. Positive values indicate more cooperation in a more connected world.

tionality, where the agents behave in an arguably rational way but the punishment is not yet high enough for an almost completely selfish agent to collaborate. Nevertheless, note that the whole collaborating region does it in a more intense way in the model with shorter distances.

4.2 Real World Implications

Specially interesting is the information provided in the figure 6. Here we can see how it is easier to achieve and sustain global cooperation when the paths are shorter and the information travels faster within the system. This increased speed of the information and punishment causes a global phenomenon in the whole community driving the system towards a more communal state.

This circumstance can be easily identified with the last twenty years of human history, where the advent and popularisation of the internet and social media has deeply changed the way humans relate to each other and how the information travels.

Nowadays, totally blocking a geographical area out of the information flow seems a near impossible, usually causing situations such as the "Streisand effect" [16], when trying to prevent the spread of some data, actually makes that data much more appealing to the general public, being almost certain that it will be distributed more widely than if no action had been taken.

For the near future, if the tendency stays the same, it is foreseeable that the area in which is beneficial to achieve global cooperation will expand as the average "informational distance" keeps reducing between populations. This alone can arguably trigger deep changes in a society that is currently adapted to a well-defined form of capitalism based on oligopolies, but in which the benefits of cooperation become more and more evident even for the most selfish individuals as informational technology develops. Well documented examples of this tendency can be found in the newly founded companies based on "collaborative economy" such as Uber, Blablacar or open source development platforms.

Appendices

.1 Agents and Runs

The model was run in an undirected RCN modular network using 20 equal communities of 10 agents each. Each parametric point was iterated for 2500 updates and averaged over 5 repetitions.

.2 Networks

A CN starts with a set of N fully connected groups or "caves" in which all the M members share an edge with all other members as seen in Figure 3a. Within the model, each group represents a family, village or closely related nations or states. Then we rewire one edge from each group to the next one, allowing intergroupal interaction with a lesser intensity than the intra-group interactions as seen in Figure 3b. This process generates pseudo-deterministic networks in which the average degree will be the constant, but the average path length will suffer small variations depending on which nodes are chosen for rewiring. This networks are characterised by having the "small world property" of high clustering coefficient, but the characteristic path lengths are too long to represent real-world networks.

In order to reduce the average path length we conducted the experiments in a RCN. In order to create an RCN we start with a CN of the desired N and M and then we randomly rewire edges with probability P. For this essay P was chosen such that the typical path lengths closely resemble those in real-world networks. For the model with 20 groups of 10 agents with $length_1 \simeq 6$ we chose $P_1 = 0.03$, and with $length_2 \simeq 3.4$ we chose $P_2 = 0.14$.

.3 Model Parameters

The rationality parameter β is cycled between the values of $\beta_{min} = 0$ to $\beta_{max} = 5$. We obtain a random behaviour for β_{min} and totally deterministic behaviour when $\beta \to \inf$, β_{max} was set to its value in order to make the stochastic transition function parametrically symmetric with respect to the punishment.

The punishment parameter π is cycled between the values of $\pi_{min} = 0$ and $\pi_{max} = 1$. Each collaborating agent has, in a timestep a punishing capacity of π that will be applied to all of its neighbouring defectors, splitting it evenly. Following this, a collaborator i with D defecting neighbours under a punishment parameter π, will exert a punishment $p_i = \pi/D$ to each one of them.

.4 Cooperation Predisposition

For an agent i, its predisposition collaboration H_i is a combination of the innate predisposition h_i and the punishment it would receive in the case of defection p_i. The innate predisposition is set to obtain an heterogeneous population such that for the sequence of players $i = 1, 2, ..., N$ their respective predispositions will be $h_i = -(i-2)/(N-2)$, which will be assigned randomly through the network, resulting in a global predisposition for an agent i of $H_i = p_i + h_i = \pi/D - (i-2)/(N-2)$.

.5 Stochastic Flipping

The stochastic transition function P defines the collaboration probabilities of an agent i which has a predisposition H_i under a rationality parameter of β. These form of the function coincides with the transition probabilities for the spins in an Ising model or for the neurons in a Hopfield Neural Network. For an agent i, its collaboration probability at any given timestep is $P_i = 1/2[tanh(\beta H_i) + 1]$.

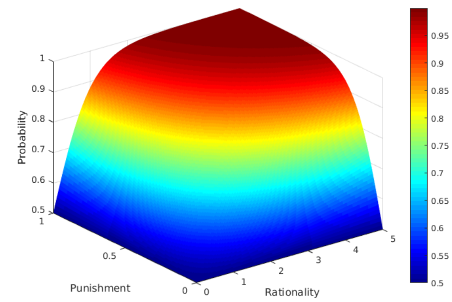


Figure 7: Surface plot of the stochastic flip function

References

[1] G. Hardin, "The Tragedy of the Commons," *Science*, vol. 162, no. 3859, pp. 1243–1248, 1968.

[2] B. Lloyd, "The Commons revisited: The tragedy continues," *Energy Policy*, vol. 35, no. 11, pp. 5806–5818, 2007.

[3] J. M. Epstein, R. L. Axtell, and T. a. Swanke, *Growing Artificial Societies: Social Science from the Bottom Up*. Brookings Institution Press, 1996.

[4] R. B. Myerson, "Game Theory: Analysis of Conflict," *Cambridge: Mass, Harvard University*, 1991.

[5] J. F. Nash, "Equilibrium Points in n-Person Games," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 36, no. 1, pp. 48–49, 1950.

[6] S. Kiesler, L. Sproull, and K. Waters, "A prisoner's dilemma experiment on cooperation with people and human-like computers," *Journal of personality and social psychology*, vol. 70, no. 1, pp. 47–65, 1996.

[7] R. Axelrod and W. D. Hamilton, "The Evolution of Cooperation.," 1981.

[8] J. Grujić, C. Gracia-Lázaro, M. Milinski, D. Semmann, A. Traulsen, J. a. Cuesta,

Y. Moreno, and A. Sánchez, "A comparative analysis of spatial Prisoner's Dilemma experiments: Conditional cooperation and payoff irrelevance.," *Scientific reports*, vol. 4, p. 4615, 2014.

[9] R. Axelrod, "Effective Choice in the Prisoner' s Dilemma," *The Journal of Conflict Resolution*, vol. 24, no. 1, pp. 3–25, 1980.

[10] C. Gracia-Lazaro, J. Gomez-Gardenes, L. M. Floria, and Y. Moreno, "Intergroup information exchange drives cooperation in the public goods game," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 90, no. 4, pp. 1–8, 2014.

[11] S. Gächter and B. Herrmann, "Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment.," *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, vol. 364, no. 1518, pp. 791–806, 2009.

[12] S. Johnson, "Escaping the tragedy of the commons through targeted punishment," *Royal Society Open Science*, 2015.

[13] S. Milgram, "The small world problem," *Psychology Today1*, vol. 2, no. 1, pp. 61–66, 1967.

[14] J. Ugander, B. Karrer, L. Backstrom, C. Marlow, and P. Alto, "The Anatomy of the Facebook Social Graph," *arXiv.org*, vol. abs/1111.4, 2011.

[15] S. L. Myers, "The economics of crime in the urban ghetto," *The Review of Black Political Economy*, vol. 9, no. 1, pp. 43–59, 1978.

[16] K. O'Grady, "Exercising right to be forgotten' could spark Streisand effect'," 2014.