Evolution of Cooperation

5 and 7

Abstract

Nature has produced cooperating communities in virtually every environment independently. This suggests that there are very simple, elegant rules that sustain evolutionarily stable cooperating populations. The goal of this paper is to shed light on some of the simplest known models that yield such cooperation. With the advent of the internet a new arena for human interaction was born. The full potential of cooperation in this new frontier has not been realized; the hope is to tame this wilderness into a safe haven where cooperation can thrive.

1 Introduction

The Prisoner's Dilemma game illuminates the inherent conflict between cooperation and selfishness. Many interactions between individuals in nature can be modeled by an instance of Prisoner's Dilemma. We illustrate this claim and introduce some terminology by considering a specific example. In his book "The Selfish Gene," Richard Dawkins writes about a bird which is susceptible to catching a deadly disease from a particular tick [2]. The parasitized bird, known as the *recipient* since it is to receive a favour, needs a fellow bird, the potential *donor*, to preen the top of its head. This incurs a *cost*, c, to the donor's *fitness*, while providing a *benefit*, $b > c$, to the recipient. In evolutionary game theory, an agent's utility is referred to as fitness and directly corresponds to the reproductive health of the agent. Assume the donor *cooperates* and preens the recipient's head, he might expect his fellow bird to return the favour. The bird formerly known as the recipient might become the donor, and vice versa. If the new donor decides not to cooperate, i.e. the donor *defects*, the outcome of the interaction would have been $+b$ to the initial recipient's fitness and $-c$ to the initial donor. Moreover, if both birds cooperate, they each benefit $b - c$, while if they both defect, they both ultimately meet their doom.

Because this game was not played simultaneously, it is not quite an instance of Prisoner's Dilemma, yet it entails many similar consequences. In particular, a society composed solely of cooperators will indulge in the highest possible average fitness, while a society of defectors would suffer the lowest [6]. More importantly, defecting is the only *evolutionarily stable strategy* (ESS); a population consisting purely of defectors is unaffected by the occurrence of mutant cooperators, while such a pure population of cooperators will eventually turn into defectors if a mutation introduces any. This is reminiscent of the “Tragedy of the Commons” observed in a multi-player version of Prisoner’s Dilemma.

The “Tragedy of the Commons” demonstrates the unfortunate tendency for individuals to behave selfishly and collectively harm the populace. Cooperation is not inherently a dominant strategy. In a society where all agents adopt the “defect” strategy, everyone suffers. In such a society, cooperation on behalf of an individual is strictly disadvantageous. However, cooperation may be achievable, as a stable strategy, when initially there is an even distribution amongst strategies. In order for the cooperative strategies to remain beneficial, agents must be motivated to participate in *altruistic* actions. An altruistic action is one that demands a small cost from the instigator and results in a large benefit to some recipient. If an agent can expect a mutual exchange, *reciprocity*, of altruistic actions then selfless behaviour might be justifiable.

In an environment comprised of random interactions between individuals, cooperation may be attained if defection can have future repercussions. If the interaction between two randomly paired agents has the potential to involve “infinite” rounds then the threat of *punishment* can lead to cooperation. *Direct reciprocity* refers to the altruistic/cooperative actions motivated in a single pairwise match. Within the context of a society, gossip can provide information about the reputation of agents. Agents may have the option to shun individuals that are known to be habitual defectors; shunning of known defectors can be a socially acceptable act. *Indirect reciprocity* refers to the altruistic/cooperative actions motivated by gossip.

It seems that cooperation is abundant in the world of insects, birds, mammals, and especially human societies. Evidently, Dawkins’ generic example above must be hiding underlying mechanisms that favour cooperation and altruism, even for self-interested agents (which, under Darwin’s theory of natural selection, all living creatures are). The present paper explores theories that attempt to explain the evolution of cooperation within the framework of natural selection. The next section will present the mechanism of reciprocity and punishment. The following section will report experimental results from simulations. The rest of the paper is dedicated to a discussion on the merits of these mechanisms in explaining the abundance of cooperation.

2 Models

The central theme at the foundation of Darwin’s theory of Evolution is “survival of the fittest” or “natural selection”; the idea that Nature mercilessly selects only the fittest genes to survive for generations is one that is well accepted today. Self-serving genes survive genetic propagation since they are dedicated to promoting the fitness of the individual and the gene; the fittest genes are appropriately coined “selfish genes” by Dawkins.

The selfish propagation of genes might make it difficult to rationalize the evolution of cooperation in human society. Selfishness is not understood in the traditional sense of refusing aid to others, rather, a selfish gene promotes actions that best serve the interests of the carrier; those interests may include cooperation. The Prisoner's Dilemma game can be understood as a model of many real-world encounters; an agent often has the choice between cooperating, in the hopes of a greater social reward or defecting to guarantee minimal sacrifice on behalf of the agent. A generic Prisoner's Dilemma game is described in Figure 1. In order to better understand the relationship between cooperation and natural selection the Prisoner's Dilemma game is often used to model the interaction of agents.

	C	D
C	$b-c$	$-d-c$
D	$b+d$	0

Figure 1: A generic repeated Prisoner's Dilemma game; each player has the options of cooperate (C) or defect (D). Note that only player 1's utilities are shown since the game is symmetric. The constant b represents the benefit to player 1 when player 2 cooperates, constant c represents the cost of cooperation, and constant d represents the difference in reward associated with defection.

Selfish genes are not limited to promoting the best interests of an individual. A selfish gene can be concerned with the propagation of similar genes. In this setting, any altruistic act is counter-intuitive. The confusion arises because nature does not select individuals but rather it selects genes, genes that are present in multiple individuals. One's genes are more likely to be present in a sibling than in anyone else. This is the main idea behind the concept of *kin selection.* One might risk one's life if such a risk were to save more than two siblings, or more than eight cousins, and so on. In a mathematical language, this translates to

$$r > \frac{c}{b} \tag{1}$$

where r is the *relatedness*, i.e. $\frac{1}{2}$ for parents, children, and siblings, $\frac{1}{8}$ for cousins, etc. From the point of view of the genes, the risk is selfishly beneficial on expectation, while from an individual's perspective, a fatal risk is taken for virtually no personal benefit; it would appear as if an altruistic act were taking place.

The preceding illustrates a mechanism that promotes cooperation without violating natural selection. Nowak presents four more so-called rules for the evolution of cooperation [6]. We shall not delve into network reciprocity and group selection, and will instead focus on the idea of indirect reciprocity.

2.1 Indirect Reciprocity

The parable of the birds and the ticks in the introductory section was an example of *direct* reciprocity. Indeed, the recipient in the first interaction returns the favour to the very same donor. The phrase "You scratch my back, I'll scratch yours" is often used to explain such an exchange. However, it is often difficult or even impossible to return the favour directly; the donor might not need anything, the recipient might not have anything to give, or by the time the donor needs the favour returned, he might not remember who the recipient was. This is increasingly relevant in human exchanges where e-commerce and the internet are quickly replacing local shops and face-to-face interactions.

The notion of *indirect reciprocity* becomes important. In this setting, a donor will only help a recipient (cooperate) if the latter is likely to help others. Once again, this seemingly altruistic act is designed to increase one's odds of receiving a favour in the future. Mathematically, the indirect reciprocity rule can be written as

$$q > \frac{c}{b} \tag{2}$$

where q is the probability that the recipient would help others. On its own, q is a nebulous quantity, because it is hard to measure. However, the addition of *reputation*, or *image scores*, to the model helps the agents approximate q [7]. As the name suggests, the donors are provided with a score that informs them on the past actions of the recipient, not unlike how a credit score informs a bank that an individual is more or less likely to make timely payments. Depending on the model, the agents might have perfect information or they might have only an approximation of other agents' pasts. Regardless, this knowledge allows agents to make informed decisions on whether the recipient is worth cooperating with.

Such a society is usually modeled by giving every agent i an integer score s_i and an integer strategy k_i. An agent's score increments by 1 every time it cooperates, and decrements by 1 after every defection. Meanwhile, an agent i's strategy k_i is a, usually fixed, integer that dictates whether said agent will cooperate. The rule is: cooperate with agent j if $s_j \geq k_i$, defect otherwise. In this model, the higher one's strategy, the more one defects, and vice versa; agents with strategies close to 0 are called *discriminators*. Note that for more realistic models of large systems, one can associate a level of uncertainty to s as we shall see in our review of experiments [7].

The introduction of quantities s and k to the system provides a way for certain members of a society (discriminators) to punish agents that do not help others (defectors). This idea that *punishment* is necessary for long-term cooperation seems reasonable and is, not surprisingly, common in studies of cooperation. This section was an example of punishment arising as a consequence of implementing (2) in a model. There are, however, some models that introduce punishment more explicitly, and this is the topic of the next section.

2.2 Costly Punishment

Punishment, in and of itself, is a fairly generic concept. Applications of punishment may result in many variants, all of which reflect the same sentiment through potentially different means. In general, punishment is not directly beneficial to the instigator, and the cost associated with punishment should reflect this restriction. In this way, the cost accurately represents reality. Furthermore, the cost of punishment also deters agents from mindlessly punishing others. In the example of indirect reciprocity, punishing someone costs 1 unit of score which, in turn, decreases the punisher's future expected utility because fewer people are likely to cooperate with the agent.

Costly punishment is a means to explicitly specify the action for reprimanding another agent. In general, costly punishment should be expensive to produce; it involves going out of one's way to inconvenience or harm another agent. Costly punishment is modeled with a modified version of Prisoner's Dilemma. Prisoner's Dilemma is generally describable by an instantiation of Figure 1. The translation of Prisoner's Dilemma to a game allowing for costly punishment is shown in Figure 2.

	C	D	P
C	$b-c$	$-d-c$	$-\beta-c$
D	$b+d$	0	$-\beta+d$
P	$b-\alpha$	$-d-\alpha$	$-\beta-\alpha$

Figure 2: A generic repeated Prisoner's Dilemma game where costly punishment is an option; each player has the options of cooperate (C), defect (D), or punish (P). Note that only player 1's utilities are shown since the game is symmetric. Constants α, β, c, and d are chosen such that the reduced table describes a Prisoner's Dilemma game [3].

2.3 Public Goods with Punishment

Another model studies an entirely different game and is worth mentioning here. The public goods experiment splits agents into groups of three. Agents' utilities are initialized to a prescribed amount at time $t = 0$ and agents build utility over consecutive rounds. At the beginning of a round each agent in a group can invest utility in a central pot. At the end of a round each agent gains 0.5 util for each unit of utility (util) in the pot. Clearly then, if everyone invests in the pot, everyone gains. However, this is not an equilibrium strategy. Indeed, knowing that the two other agents will cooperate (i.e. invest in the pot), if the third agent cooperates, for every 1 util he only gains 0.5 util yielding a net loss of 0.5. Therefore, a mechanism for punishment is introduced to examine whether cooperation is induced. In this specific case, punishing another agent costs the punisher 1 util, while costing the punished agent 3 utils.

3 Experimental Results

In this section, we present results obtained with various models designed to explain the emergence of evolutionarily stable cooperation.

3.1 Indirect reciprocity with Nowak and Sigmund

We now consider the results of Nowak and Sigmund concerning indirect reciprocity by image scoring [7]. They use the image scores s that keep track of each agent's cooperation and defection history, and the integer strategies $k \in [-5, 6]$. If $k = -5$ the individual's strategy is to unconditionally cooperate. If $k = +6$ the individual's strategy is to unconditionally defect. A strategy where $k = 0$ is a discriminatory strategy that provides assistance to individuals only if they are in good standing. They report four sets of results and we shall relay them in the following sections.

3.1.1 Image scoring

This first experiment is the basic implementation of the model described in Section 2.1. The computer initializes 100 agents with random strategies k and scores $s = 0$. Every generation 125 pairs of agents are chosen at random to interact according to the $s_j \geq k_i$ cooperation rule. At the end of the generation, agents reproduce proportionally to their scores; the next generation of agents inherit their parent's strategy k exactly but their scores are reset to $s = 0$.

In this simple experiment, it took 166 generations for all non-zero strategies to become extinct. In other words, after 166 generations, all agents in the system are perfect discriminators.

3.1.2 Mutations

In the basic model above, once a strategy is extinct, it can never reappear in the system. In the next experiment, Nowak and Sigmund introduce mutations, which is very reasonable. In particular, they assign a probability of 0.1% to a mutation, where an agent in the next generation is assigned a strategy at random. After 10 million generations, the results were still in favour of cooperation – roughly a 70/30 split in favour of cooperation. Although the discriminators ($k \in \{-2, -1, 0\}$) are still a majority, the mutations seem to sustain a population of cooperators ($k \in \{-5, -4, -3\}$) and defectors ($k \in \{4, 5, 6\}$), with the latter taking advantage of the cooperators.

3.1.3 Uncertainty

In the previous two experiments, every agent in the system observes every interaction with perfect information. For large systems, especially in nature, this can be unrealistic. Therefore, in the next experiment, each interaction is observed by the participants and only 10 randomly chosen agents, all others have no knowledge of the interaction and do not update the donor's score. This was tested on system sizes of 20, 50, and 100.

As expected, for small systems, discriminators are still the most abundant population. Meanwhile, the larger the system the more abundant defectors become, until at a population size of 100, defectors are by far the most abundant. This is not at all surprising, since the number of observers is fixed at 10. Indeed, at the mathematic limit of increasing system size, this should asymptotically approach a system without indirect reciprocity, which we know favours defectors.

3.2 Costly Punishment with Dreber

The work of Dreber et al. analyzes the practicality of attributing cooperation in society to costly punishment [3]. A series of expectedly short-run experiments analyze the actions and strategies of human players in a repeated Prisoner's Dilemma game where costly punishment is an option. A generic repeated Prisoner's Dilemma game with costly punishment is shown in Figure 2. The results of the experiments of [3] indicate that costly punishment did not improve the average payoff, in fact, it provides no benefit for the group. Furthermore, individual agents should not be motivated to partake in costly punishment since the highest scoring agents never punished and agents who participated in costly punishment received the lowest average payoff. Interestingly, Dreber et al. argue that, in practice, punishment is not a complete explanation of cooperation. Rather, punishment could be seen as as means of coercion, of forcing a weaker agent into submission, or of establishing a hierarchy.

3.3 Public Goods with Gächter

Gächter et al. implemented the Public Goods game with 207 participants, given 20 utils to start, who were split in groups. There were 4 different types of groups: groups that played the game *with* punishment for 10 and 50 turns, called P10 and P50, respectively, and control groups that played the game *without* punishment for 10 and 50 turns, called N10 and N50 [4]. They report interesting results which we interpret in the following.

First and foremost, they note a very significant increase in average payoff between N50 and P50 in favour of the latter. Meanwhile, the opposite holds for the shorter N10 and P10 experiments. The specific end-time for which one beats the other depends on the somewhat arbitrary numbers chosen for the experiment, so it is not of particular interest here. However, it is interesting to see that the longer the experiment runs, the larger the difference in average payoff.

Both N experiments exhibited very smooth evolutions of average payoffs relative to the P experiments. In fact, one can safely speculate that large dips in average payoffs are most likely due to punishments, since that is the only action that siphons a net utility out of the system. Furthermore, these are always followed by large increases in payoff, often resulting in a higher average payoff than before the punishment. These results definitely suggest that there is a social benefit to costly punishment.

Surprisingly, the P50 experiment exhibits a significant dip (over 20%) on the final turn. This is obviously because all agents know when the game ends and would most likely not happen in an infinite game or one where the end is dictated by some random event. However, it is strange that participants would punish on the final turn.

3.4 Phenotypic Defectors

Punishment alone has been shown to be an insufficient motivation for the evolution of cooperation. Even with the addition of a social environment, namely the representation of gossip, cooperation does not achieve a stable equilibrium. The experiment of [5] models the interaction of individuals when there exist *phenotypic defectors* in the population. A phenotypic defector is one that is incapable of providing assistance to others even if they have a predisposition to help; the handicapped, the sick, and the young can be examples of phenotypic defectors.

Lotem et al. model a complex system of interactions. Individuals interact randomly in a generation and build one score according to the success of their strategy and another describing their image. Each individual is prescribed a genetic strategy that discriminates between other agents according to their image score. The propagation of genetic scores to the next generation is dependent on the success of the genetic strategies; only fit individuals reproduce.

Throughout a series of experiments the proportion of phenotypic defectors and the cost of cooperation varied. In the absence of phenotypic defectors where the cost of cooperation was high, defectors quickly became the dominant population. As the cost of cooperation was decreased, cooperation became a successful strategy, but strategies tended towards unconditional cooperation permitting the periodic invasion of defectors. With the addition of phenotypic defectors, discriminatory cooperation became a stable strategy. Interestingly, in a situation where the cost of cooperation was relatively low and the population of phenotypic defectors was substantial (20%) the dominant strategy was one of forgiveness and cooperation ($k = -1$).

4 Discussion

Cooperation among individuals in nature is ubiquitous. From small animal communities to immense insect colonies, one easily finds many examples of evolutionarily stable cooperation. However, the simplest known game to model interactions predicts defection as the only pure strategy ESS. Therefore, either more complicated mixed strategies need to be explored, or an entirely different model should be studied to explain the evolution of cooperation.

Nowak and Sigmund suggest that the evolution of cooperation in nature is directly related to, among other things, reputation [8]. This would explain the refinement of the senses that allow individuals to recognize others, for instance an acute sense of smell in dogs, or extremely developed eye sight in humans. In social settings, reputation is strongly linked to face-to-face interactions. Other settings, such as online exchanges, rely on more literal scoring, such a Facebook Likes, Google +1's, Yelp reviews, etc. The results by Nowak and Sigmund over the years have been very compelling and in the present paper, we have reported some of their results that show how adding a simple rule for indirect reciprocity to the cooperation strategy improves the outcome greatly in perfect information settings and small groups. Depending on the specific model that was implemented, cooperators were more or less abundant, however, in most models cooperators were present in the evolutionarily stable population. (The one model where they were almost extinct was for large groups with imperfect information; this setting approaches a model with no indirect reciprocity, so this was not surprising.) For large groups, where indirect reciprocity without perfect information fails, they have proposed intuitive rules such as network reciprocity and group selection [6]. Axelrod contributed a novel approach to studying cooperation [1]; modeling the evolutionary success of strategies provides a computationally efficient means of exploring the space of strategies (including unconditional defection, unconditional cooperation, and conditional cooperation) and settling on an equilibrium.

While Nowak and Sigmund pioneered the use of image scores in the context of an evolutionary model [7], Lotem et al. harnessed the idea of image scores and introduced a subset of the population to represent individuals incapable of contributing to society [5]. The addition of these so-called phenotypic defectors represents an interesting aspect of society that may seem inconsequential at first glance but proves to be significant to the dynamic of a repeated Prisoner's Dilemma game. The model of [7] is blighted with periodic instability; with the addition of phenotypic defectors to the model, a stable equilibrium is achievable. The phenotypic defectors of [5] are made to be indistinguishable from selfish defectors and, for that reason, would be punished just as harshly for their failure to cooperate. Such a simplified depiction of phenotypic defectors may be inconsistent with reality since individuals incapable of contributing to society are more frequently provided assistance.

Costly punishment is popularly proposed as a mechanism to encourage participation in cooperation. Dreber et al. [3] provide compelling reasons to doubt the rationale that costly punishment is the basis for the evolution of cooperation. Dreber et al. conducted carefully structured experiments where costly punishment was the unique factor that distinguished the control group from the test group. The results indicated that neither did costly punishment improve the average payoff nor did the fittest (most successful) agents participate in costly punishment. Since costly punishment is analyzed in cases where it is the only factor, one can conclude that though costly punishment is not sufficient for the evolution of cooperation it may still be essential. The work of Dreber et al. is limited to costly punishment in the context of direct reciprocity; explorations in the application of costly punishment to settings allowing for indirect reciprocity may yield fruitful results.

Gächter et al. studied a completely different model, involving interactions between three individuals: the Public Goods game. Punishment proved to be extremely beneficial to social welfare as long as the experiment lasted long enough. Interestingly, this experiment was done with actual human participants, so it incorporates complex human psychology. Unfortunately, the experiment does not seem to be repeated over "generations", so the results cannot be convincingly associated to evolutionarily stable solutions. However, this would be difficult to implement in this case since the experiment is not computer-simulated.

5 Conclusion

The evolution of cooperation in nature can be interpreted as the result of playing a seemingly infinite run of the game of survival or gene-propagation. Prisoner's Dilemma is a simple and widely accepted model for this game. On its own, Prisoner's Dilemma is considered insufficient to model the evolution of cooperation in Nature, however, it provides a fruitful base for explorations in this domain. Techniques such as image scoring, indirect reciprocity, costly punishment, and phenotypic defectors have proved useful in improving a model of cooperation. We explored a variety of models and experiments, each implementing a subset of these techniques. Though each model provides insight into the evolution of cooperation, as yet, no model provides a complete explanation of the evolution and stability of cooperation. Each model presents an extremely simplified depiction of reality and though simple models are ideal, none accurately explain the reality of Nature. The key to an accurate model of the evolution of cooperation may lie in a complex combination of these fundamental techniques.

References

[1] R. Axelrod. *The Evolution of Cooperation.* Basic Books, 1984.

[2] R. Dawkins. *The Selfish Gene.* Oxford University Press, 1976.

[3] A. Dreber, D. G. Rand, D. Fudenberg, and M. A. Nowak. Winners don't punish. *Nature*, 452(1785):348–351, 2008.

[4] S. Gächter, E. Renner, and M. Sefton. The Long-Run Benefits of Punishment. *Science*, 322(5907):1510, 2008.

[5] A. Lotem, M. A. Fishman, and L. Stone. Evolution of cooperation between individuals. *Nature*, 400:226–227, 1999.

[6] M. A. Nowak. Five rules for the Evolution of Cooperation. *Science*, 314(5805):1560–1563, 2006.

[7] M. A. Nowak and K. Sigmund. Evolution of indirect reciprocity by image scoring. *Nature*, 393:573–577, 1998.

[8] M. A. Nowak and K. Sigmund. Evolution of indirect reciprocity. *Nature*, 437:1291–1298, 2005.