

# Emergence of Cooperative Societies in Evolutionary Games Using a Social Orientation Model

Kan-Leung Cheng

## Abstract

We utilize evolutionary game theory to study the evolution of cooperative societies and the behaviors of individual agents (i.e., players) in such societies. We present a novel player model based upon empirical evidence from the social and behavioral sciences stating that: (1) an individual's behavior may often be motivated not only by self-interest but also by the consequences for others, and (2) individuals vary in their interpersonal social tendencies, which reflect stable personal orientations that influence their choices. Alongside the formal player model we provide an analysis that considers possible interactions between different types of individuals and identifies five general steady-state behavioral patterns. We present evolutionary simulations that ratify previous findings on evolution of cooperation, and provide new insights on the evolutionary process of cooperative behavior in a society as well as on the emergence of cooperative societies. Our main experimental result demonstrates that in contrast to previous common knowledge, increasing mutual reward or mutual punishment in the Prisoner's dilemma game *does not* result in the same type of cooperative society: while increasing reward does increase the society's cooperativeness level, increasing mutual punishment does *not*.

## 1 Introduction

Evolution of cooperation has been studied for years, most notably starting from the seminal work of Smith [19] and Axelrod [2]. The underlying question can be summarized briefly as follows: why and how does cooperative behavior evolve in a Darwinian society where the *survival of the fittest* is the prominent behavioral rule? Existing research on this question typically utilize normal-form games that are highly simplified models of social dilemmas. For example, Figure 1 presents the famous Prisoner's Dilemma (PD) game, where two players are both faced with a decision to either cooperate (C) or defect (D). If the game is played once, then defecting will provide a higher payoff regardless of whether the other player cooperates or defects. However, if the game is played repeatedly for an unknown number of times, cooperative behavior in an individual might emerge to increase accumulated payoffs (see [9] for an overview).

Prisoner's Dilemma		Player 2	
		Cooperate (C)	Defect (D)
Player 1	Cooperate (C)	(3, 3)	(0, 5)
	Defect (D)	(5, 0)	(1, 1)

Figure 1: The Prisoner's Dilemma game.

We consider *how cooperative societies emerge, given varying social tendencies of the individuals*. Traditional game theory typically uses the *rationality assumption*, i.e., the assumption that human behavior is purely rational, self-maximizing behavior [13]. However, this assumption has received wide criticism from the behavioral science literature (e.g. [11]). For example, the *Social Value Orientation* (SVO) theory [5, 4] conjectures based on many empirical studies, that the social choices people make depend, among other things, on other people; in particular on the stable personality differences between individuals.

In this paper, we describe a new player model and a formalism based on the SVO Theory. Our formalism captures the notion of varying and persistent social orientations exhibited in human behavior and enables a player to reason about the relationships between the player's social orientation and how that player develops strategies in a repeated game. We present theoretical results showing how players with different social tendencies interact in a class of 2x2 symmetric games. Our analysis identifies five general *steady state* behavioral patterns that can be explained in terms of the players' varying social orientation values.

Our experiments based on evolutionary simulations in the *Iterated Prisoner's Dilemma* (IPD) demonstrated the effects of social orientations on the evolution of cooperative behavior in individual players and on the emergence of a cooperative society. One set of experiments showed that prosocial tendency increases with increasing reward or with decreasing temptation, thus confirming previous intuitions from [14].

In our experiments, we also found out that there are scenarios in which the definition of cooperativeness of previous studies might lead to erroneous conclusions. Previous works on the evolution of cooperation typically used the average payoff of the society as a measure of its cooperativeness: i.e., the higher the average payoff is, the more cooperative the society is thought to be [15, 16, 7]. However, our results showed that while increasing the value of rewards or punishments results in a similar increase in the average payoff, this **does not** result in the emergence of the same kind of society. In particular in our experiments, increasing mutual reward typically resulted in a cooperative society, but increasing mutual punishment resulted in a divided society that includes two distinct clusters: one of highly selfish players and the other of highly cooperative players.

Our contributions can be summarized as follows:

1. A new player model based upon the well-founded Social Value Orientation theory. This player model provides a more accurate description of human decision making in 2x2 games than the self-maximizing model [4].

2. An analytical work that identifies five general *steady state* behavioral patterns.
3. New experiments that, in contrast to old findings, shows the differences in the emergent societies when playing the iterated prisoner’s dilemma game with different values.

## 2 Background

*Iterated Prisoner’s Dilemma (IPD)*, an iterated variant of the prisoner’s dilemma that is played repeatedly an unknown number of iterations, has been the most common model for social dilemmas between two players and has been often used in order to study the evolution of cooperation. Since Axelrod’s IPD tournament [3] that focused on generating winning strategies, there has been a large body of research on the various aspects of the basic IPD model: varying payoffs, number of players, and various population and structural dynamics (see an excellent overview in [9]).

Traditionally, the majority of the works in game theory follow the classical *rationality assumption* (a.k.a the *economic man* assumption in social and behavioral sciences), either implicitly or explicitly. On the other hand, experiments in social and behavioral sciences show that humans rarely follow this assumption. For instance, consider the *Ultimatum Game*, in which two players interact to decide how to divide a sum of money that is given to them. The first player proposes how to divide the sum between the two players, and the second player can either accept or reject this proposal. If the second player rejects, neither player receives anything. If the second player accepts, the money is split according to the proposal. Existing experiments in this game show that the offers that are issued or accepted are closer to a “fair” division of the money (\$50 for each) than the “rational” choice [11].

Indeed, it is widely accepted in social and behavioral sciences that players explicitly take into account the outcome of the other players’ actions when considering their course of action. Moreover, the choices people make depend, among other things, on stable personality differences in the manner in which they approach interdependent others. This observation can be traced back to the seminal work by Messick and McClintock [5] in which they presented a motivational theory of choice behavior that considers both players’ payoffs in game situations. This theory was later denoted as the *Social Value Orientation* theory, that has since developed into a class of theorems (see [4] for an excellent review).

Most of the SVO based studies typically recognize two opposing social value orientations: a proself and prosocial orientation. A *proself orientation* is one that gives higher consideration to its own payoff, while a *prosocial orientation* gives higher regards to the payoff of the agents he is interacting with. The social orientation of a player is not an absolute value; it describes a spectrum of possible behaviors, in which one end of the spectrum denotes proself behavior

2x2 symmetric game		Player 2	
		Cooperate (C)	Defect (D)
Player 1	Cooperate (C)	$(R, R)$	$(S, T)$
	Defect (D)	$(T, S)$	$(P, P)$

Figure 2: Generalized form of 2x2 symmetric games.

and the other end denotes prosocial behavior.

Note that in contrast to the diversity of the SVO theory, the traditional rationality assumption dictates that all individuals are proself, without any difference between one and another. As most social or psychological traits, the claim that SVO is a fundamental personality trait is supported by both biological and sociological findings [1]. Biological support also can be found, among others, in Van Lange’s work [20] showing that the basic form of SVO is visible early in life as part of a child’s temperament.

Over the years, there has been significant advances on social dilemmas and various aspects of the social value orientations since the seminal work in [5]. For example, Parks and Rumble [17] showed that different aspects of the Tit-for-tat strategy have different effects on the cooperation rates of individuals with different SVO values. In addition, there were several other research questions that considered some relaxation of the rationality assumption in their solution, for instance in [6] the authors presented a computational model that allows for achieving fairness in multi-agent systems. Their computational model uses the *Homo Equalis* utility function that has been shown to adequately describe human behavior in several games.

In the subsequent sections, we describe a new formalism based on the SVO theory, for studying evolution of cooperation. Previous works on the evolution of cooperation typically used the average payoff of the population as a measure of the cooperativeness [15, 16, 7]. In contrast, our formalism focuses on the social value orientations of the players and uses the average social values of the society as a new measure of cooperativeness.

### 3 Our Model

We consider normal-form games for studying social interactions between individuals. Figure 2 shows a generalized form of the payoff matrix for such games, where various constraints on the payoffs can be used to define different classes of social dilemmas. In this paper, we only focus on symmetric games where  $T > S$ . Note that this assumption is not restrictive: many well-known games satisfy this condition, including the well-known Prisoner’s Dilemma, Chicken Game, and Stag-Hunt [18].

We start by defining the *social-orientation space* of the two players in a game, namely Player  $i$  and Player  $j$ . The social-orientation space of a game can be viewed as a two-dimensional Euclidean space, as illustrated in Figure 3 [12].

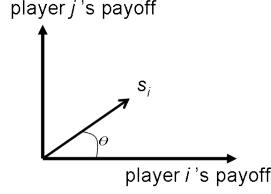


Figure 3: An Illustration of the social value orientation space. The  $x$  and  $y$  axes show the accumulated total payoff for Players  $i$  and  $j$ , respectively.

The  $x$ -axis represents the accumulated total payoff of Player  $i$  and the  $y$ -axis represents that of Player  $j$ .

The *social orientation* of Player  $i$  is a unit vector  $\hat{s}_i$  such that  $\hat{s}_i$ 's initial point is at the origin of the social-orientation space. We represent  $\hat{s}_i$  by the angle,  $\theta_i$  between  $\hat{s}_i$  and the  $x$ -axis of the social-orientation space. Intuitively, the *social orientation* of a player is a model of its tendency to adopt a prosocial or proself behavior in a game.

For example, when  $\theta_i = 0$  then Player  $i$  acts as a proself individual. If  $\theta_i = \pi/4$ , then this means that player is fair, i.e., it acts to balance the accumulated total payoffs of two players. When  $\theta_i = \pi/2$ , the player is purely prosocial, i.e., it never attempts to maximize its own payoff, but rather it tries to increase the payoff of the other player.

We define a player's *game model* at any time point  $t$  in the repeated game as a vector:

$$\vec{g} = \langle p_i, p_j \rangle$$

where  $p_i$  is the accumulated total payoff that Player  $i$  receives from the start of the game to the point  $t$  and  $p_j$  is that of Player  $j$ . Note that both players hold the same game model that describes their accumulated total payoff, and it is the only variable state that players remember.

Suppose one of the players, say Player  $i$ , takes an action,  $C$  or  $D$ , in the game. Let  $\vec{g}$  be the current game model. The expected amount of change in the game model  $\vec{g}$  after Player  $i$  takes the action  $C$  or  $D$  is as follows:

$$E[\text{effect of } C] = \vec{p}_C = \left\langle \frac{R+S}{2}, \frac{R+T}{2} \right\rangle$$

$$E[\text{effect of } D] = \vec{p}_D = \left\langle \frac{T+P}{2}, \frac{S+P}{2} \right\rangle$$

The above definition assumes that Player  $i$ 's model of Player  $j$  is of a random player. In other words, Player  $i$  does not have any background knowledge about the other player and it cannot store and learn from the other player's actions.

The *expected utility vector for an action*  $a \in \{C, D\}$  is

$$\vec{g} + \vec{p}_a,$$

where  $\vec{g}$  is the game model and  $\vec{p}_a$  is the expected amount of change in the game model by doing  $a$ .

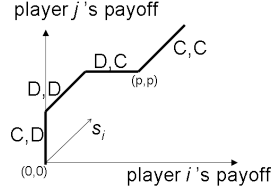


Figure 4: An example reaction of a fair player (Player  $i$ )

During the course of the game, each player aims to bring the game model closer to its social-orientation vector,  $\hat{s}_i$ . In other words, each player aims to change the world to conform to its preference and social-orientation. The differences between the orientations of the players create the tensions in their social interactions – hence the social dilemmas. Note that with the traditional rationality assumption, the players will try to do utility-maximization on their own payoff. In other words, the theta equals to zero and social orientation equals to  $\langle 1, 0 \rangle$ .

Let  $\vec{g} = \langle p_i, p_j \rangle$  be the current game model. The objective of each player is to minimize the in-between angle  $\alpha$  between its own social-orientation vector  $\hat{s}_i$  and the expected utility vector  $\vec{g} + \vec{p}_a$ . The angle  $\alpha$  is computed as follows:

$$\text{For cooperation: } \cos \alpha_C = \frac{\hat{s}_i \cdot (\vec{g} + \vec{p}_C)}{|\vec{g} + \vec{p}_C|}.$$

$$\text{For defection: } \cos \alpha_D = \frac{\hat{s}_i \cdot (\vec{g} + \vec{p}_D)}{|\vec{g} + \vec{p}_D|}.$$

Thus, each player will choose an action  $a$  such that

$$a = \operatorname{argmax}_{a \in A} \cos \alpha_a.$$

**Example 1.** Consider the well-known Iterated Prisoner's Dilemma (IPD) game as depicted in Figure 1. Figure 4 shows how a fair player (i.e.,  $\theta = \pi/4$ ) interact with another player in an IPD game. For IPD,  $\vec{p}_C = \langle \frac{3}{2}, 4 \rangle$ ,  $\vec{p}_D = \langle 3, \frac{1}{2} \rangle$ . At the beginning,  $\vec{g} = \langle 0, 0 \rangle$ ,  $\alpha_C$  is smaller than  $\alpha_D$ , therefore, the fair player will cooperate at the first iteration. When the other player defects at the first iteration, the utility vector becomes  $\langle 0, 5 \rangle$ , and  $\alpha_D$  is smaller than  $\alpha_C$ . Therefore, the fair player defects at the second iteration. When the other player defects again, both of them get 1 and the utility vector becomes  $\langle 1, 6 \rangle$ .  $\alpha_D$  is smaller than  $\alpha_C$  again, therefore, the fair player defects until the other cooperates at some point after which the payoffs of both players becomes balanced, i.e.,  $\vec{g} = \langle p, p \rangle$  for some  $p$ . In other words, a fair player in IPD game behaves exactly the same as the well-known Tit-For-Tat (TFT) strategy.  $\square$

## 4 Analysis

We are interested in the dynamics of a player's behaviors (strategies), based on its social-orientation, over the course of its interaction with another player in a repeated game. This section presents an exhaustive analysis of such dynamics based on the model described earlier.

We use the definition for the social-orientation of a player in the following form. Let the two players be  $i$  and  $j$ , and their social-orientation angles be  $\theta_i$  and  $\theta_j$ , respectively. We define the *preference ratios* for each player as  $r_i = \frac{\cos \theta_i}{\cos \theta_i + \sin \theta_i}$  and  $r_j = \frac{\cos \theta_j}{\cos \theta_j + \sin \theta_j}$ .

We define the following ratios for each action,  $C$  and  $D$ , in the 2x2 symmetric game:  $r_C = \frac{R+S}{(R+S)+(R+T)}$  and  $r_D = \frac{T+P}{(S+P)+(T+P)}$ . Intuitively, these ratios describe the expected share of payoff that the first player will get by choosing  $C$  and  $D$ , respectively. We assume that  $S < T$ , therefore,  $r_C < 0.5 < r_D$ . In other words, a cooperation by Player  $i$  results in Player  $i$  receiving a smaller share of the payoff and thus Player  $j$  will be receiving a larger share in the game.

Let  $\vec{g} = \langle p_i, p_j \rangle$  be the current game model. We define the current ratios for each player as  $g_i = \frac{p_i}{p_i + p_j}$  and  $g_j = \frac{p_j}{p_i + p_j}$ . In a steady state, Player  $i$  cooperates whenever he is satisfied with his current ratio, i.e., his current ratio is greater than or equal to his preference ratio (i.e.,  $g_i \geq r_i$ ), or defects otherwise. Without loss of generality, we assume  $r_i \leq r_j$ . There are five possible cases in steady state:

**Theorem 1.** *If  $r_j \geq r_i > 0.5$  (i.e., prosel), both players always defect and get  $P$  at each game in steady state.*

*Proof.* When  $g_i \geq r_i$ , Player  $i$  cooperates while Player  $j$  defects, so  $g_i$  moves toward  $\frac{S}{S+T} < r_i$ . When  $g_i < 1 - r_j$ , Player  $i$  defects while Player  $j$  cooperates, so  $g_i$  moves toward  $\frac{T}{S+T} > 1 - r_j$ . Otherwise, both players defect and  $g_i$  moves toward  $\frac{P}{P+P} = 0.5$ .  $\square$

For example, in an IPD game, let  $r_i = 0.6$  and  $r_j = 0.7$ . This means that Player  $i$  will always aim to get a share of 60% of the total payoff, while Player  $j$  will aim to get a share of 70% of the total payoff. Therefore, both will never be satisfied and will constantly defect to get a payoff of 1.

**Theorem 2.** *If  $r_i \leq r_j \leq 0.5$  (i.e., prosocial), both players always cooperate and get  $R$  at each game in steady state.*

*Proof.* When  $g_i < r_i$ , Player  $i$  defects while Player  $j$  cooperates, so  $g_i$  moves toward  $\frac{T}{S+T} > r_i$ . When  $g_i \geq 1 - r_j$ , Player  $i$  cooperates while Player  $j$  defects, so  $g_i$  moves toward  $\frac{S}{S+T} < 1 - r_j$ . Otherwise, both players cooperate and  $g_i$  moves toward  $\frac{R}{R+R} = 0.5$ .  $\square$

For example, in an IPD game, let  $r_i = 0.3$  and  $r_j = 0.4$ . This means that Player  $i$  will aim to get a share of 30% of the total payoff, while Player  $j$  will aim to get a share of 40% the total payoff. As such, they will both be easily satisfied, and therefore always cooperate and get the rewards, i.e., 3.

**Theorem 3.** *If  $r_i < 0.5$  and  $r_i + r_j = 1$ , there are two cases:*

- *when  $r_j > \frac{T}{S+T}$ , Player  $i$  gets  $S$  while Player  $j$  gets  $T$ ;*
- *otherwise, Player  $i$  gets  $r_i(T + S)$ , and Player  $j$  gets  $(1 - r_i)(T + S)$ .*

*Proof.* The first case above immediately follows from the fact that when  $r_j > \frac{T}{S+T}$ , we will have the repeated sequence of Cooperate-Defect actions in all interaction traces.

The proof for the second case is as follows. When  $g_i < r_i$ , Player  $i$  defects while Player  $j$  cooperates, so  $g_i$  moves toward  $\frac{T}{S+T} > r_i$ . When  $g_i \geq r_i$ , Player  $i$  cooperates while Player  $j$  defects, so  $g_i$  moves toward  $\frac{S}{S+T} < r_i$ . In steady state, they interact in a way that the ratio  $r_i$  (and  $r_j$  as well) is achieved, so Player  $i$  gets  $r_i(T + S)$  while Player  $j$  gets  $T + S - r_i(T + S)$ .  $\square$

For example, in an IPD game, let  $r_i = 0.4$  and  $r_j = 0.6$ . This means that Player  $i$  will always aim to get a share of 40% of the total payoff, while Player  $j$  will aim to get a share of 60% of the total payoff. Therefore, they will try to grasp the share *alternatively*. In a steady state, Player  $i$  gets 2 and Player  $j$  gets 3 at each game on average.

**Theorem 4.** *If  $r_j > 1 - r_i > 0.5$ , there are two cases:*

- *when  $r_j > \frac{T}{S+T}$ , Player  $i$  gets  $S$  while Player  $j$  gets  $T$ ;*
- *otherwise, Player  $i$  gets  $\bar{p}_i = \frac{SP-PT}{(P-T)-(P-S)\frac{1-r_i}{r_i}}$ , and Player  $j$  gets  $\bar{p}_j = \bar{p}_i \frac{1-r_i}{r_i}$ .*

*Proof.* The proof for the first case is the same that of Theorem 3 above. The proof of the second case is as follows. When  $g_i < r_i$ , both players defect and get  $P$ . When  $g_i \geq r_i$ , Player  $i$  cooperates and gets  $S$  and Player  $j$  defects and gets  $T$ . In steady state, they will get  $(S, T)$  or  $(P, P)$  in each game in a way that  $r_i$  is achieved. Let  $n_{DD}$  be the portion of the games resulted in  $DD$ , Player  $i$  gets  $\bar{p}_i$  and Player  $j$  gets  $\bar{p}_j$  where  $\bar{p}_i = r_i(\bar{p}_i + \bar{p}_j)$ ,  $\bar{p}_i = Pn_{DD} + S(1 - n_{DD})$  and  $\bar{p}_j = Pn_{DD} + T(1 - n_{DD})$ . Solving them, we can obtain the above formula.  $\square$

For example, in an IPD game, let  $r_i = 0.4$  and  $r_j > 0.6$ . Now we are in a situation where there is lack of resources (as  $r_i + r_j > 1$ ) and Player  $j$  is more proself than Player  $i$ . As such, Player  $j$  will always defect, while Player  $i$  will sometimes cooperate, and sometimes defect. In a steady state, Player  $i$  will get  $\frac{10}{11}$  and Player  $j$  will get  $\frac{15}{11}$  at each game on average.

**Theorem 5.** *If  $r_i < 1 - r_j < 0.5$ , there are two cases:*

- *when  $r_i < \frac{S}{S+T}$ , Player  $i$  gets  $S$  while Player  $j$  gets  $T$ ;*
- *otherwise, Player  $i$  gets  $\bar{p}_i = \bar{p}_j \frac{1-r_j}{r_j}$  and Player  $j$  gets  $\bar{p}_j = \frac{TR-RS}{(R-S)-(R-T)\frac{1-r_j}{r_j}}$ .*



*Proof.* The proof for the first case is the same that of Theorem 3 above. The proof of the second case is as follows. When  $g_j < r_j$ , Player  $i$  cooperates and gets  $S$  and Player  $j$  defects and gets  $T$ . When  $g_j \geq r_j$ , both players cooperate and get  $R$ . In steady state, they will get  $(S, T)$  or  $(R, R)$  in each game in a way that  $r_j$  is achieved. Let  $n_{CC}$  be the portion of the games resulted in  $CC$ , Player  $i$  gets  $\bar{p}_i$  and Player  $j$  gets  $\bar{p}_j$  where  $\bar{p}_j = r_j(\bar{p}_i + \bar{p}_j)$ ,  $\bar{p}_i = Rn_{CC} + S(1 - n_{CC})$  and  $\bar{p}_j = Rn_{CC} + T(1 - n_{CC})$ . Solving them, we can obtain the above formula.  $\square$

For example, in an IPD game, let  $r_i < 0.4$  and  $r_j = 0.6$ . Then, as resources are plentiful (as  $r_i + r_j < 1$ ) and Player  $i$  is more prosocial than Player  $j$ , Player  $i$  will always cooperate, while Player  $j$  will sometimes cooperate, and sometimes defect. In a steady state, Player  $i$  will get  $\frac{30}{13}$  and Player  $j$  will get  $\frac{45}{13}$  at each game on average.

## 5 Experiments

We have performed several experiments in order to investigate the emergence of cooperative populations. These experiments involve evolutionary simulations on a society of players and the simulations are designed based on social orientations of individuals, as described below.

We used the replicator dynamics for evolutionary simulations [19]. We used the well-known “infinite population” setup for initializing the population as described in [15, 16, 7]. We randomly generated 10 theta values from the interval  $[0, \pi/2]$  and assumed the size of a group with a particular  $\theta$  value constitutes 10% of the entire population.

In each generation, the players engaged in pairwise encounters, resulting in a payoff for each of the players that is equal to the sum of the payoffs from each individual round of the game. The expected values of the score of a player in a pairwise game in steady state are described in the previous section. After each generation, each player had a number of offspring that is proportional to its expected total payoff. Each offspring had the same social-orientation value  $\theta$  as its parent. If the frequency of a group of players with a particular  $\theta$  value drops below a threshold 0.001 then the group is discarded from the population.

On average, in every 100 generations, a small amount (frequency of 0.01) of new randomly generated mutant players are introduced into the population. Each simulation was performed for 10,000 generations, resulting a total of about 100 mutant strategies.

### 5.1 Prisoner’s Dilemma with Constant Payoffs

Figure 5 shows the average population  $\theta$  and average population payoff of an evolutionary process of over  $10^5$  generations in the Prisoner’s Dilemma (PD) game. Here, the average payoff varies between  $P$  and  $R$ , which correspond to full defection and full cooperation, respectively. At the beginning of the evolutionary simulation, cooperative players in the population which have high

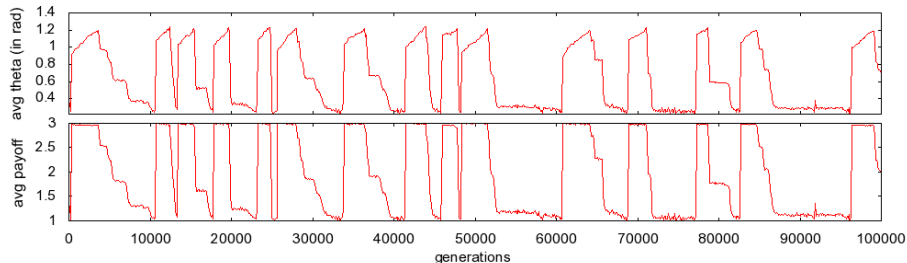


Figure 5: An evolutionary simulation of IPD. The top graph shows the average theta per generation. The bottom graph shows the average payoff per generation.

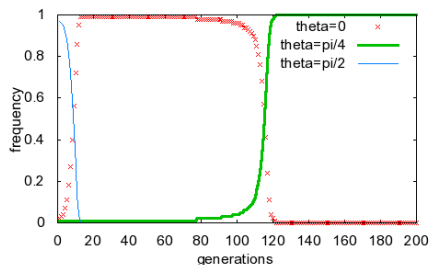


Figure 6: Invasion of fair player.

$\theta$  values dominate the population quickly. Then, prosself players (with low  $\theta$  value) emerge gradually to dominate after about 4000 generations. At around the 10000<sup>th</sup> generation, cooperative players *suddenly* regain the majority of the population. This wave-like behavior between cooperation and defection (prey-predator cycle), is a widely known phenomenon in repeated PD games, that was observed under various conditions [16, 7]. As can be seen in Figure 5, similar behavior also emerges in our experiments with players modeled by their social orientation values.

By examining the evolutionary traces, we found that this phenomenon is caused by mutant players introduced in the population with  $\theta_{mutant} \approx \pi/4$ . These fair players avoided the exploitation of prosself players with  $\theta_{proself} < \pi/4$ , and at the same time cooperated with the other cooperative players with  $\theta_{cooperative} > \pi/4$ . In other words, the mutant players were using strategies similar to Tit-For-Tat.

Figure 6 illustrates the change of population frequencies of three types of players (selfish, altruistic and fair) without mutation. At the beginning, altruistic players dominate the population. The theta value of an altruistic player is  $\pi/2$ ; i.e., it will always cooperate. Therefore, it can be easily exploited and invaded by a selfish players ( $\theta = 0$ ). When the altruistic players are extinct, selfish players can also be invaded by a group of fair players who will cooperate

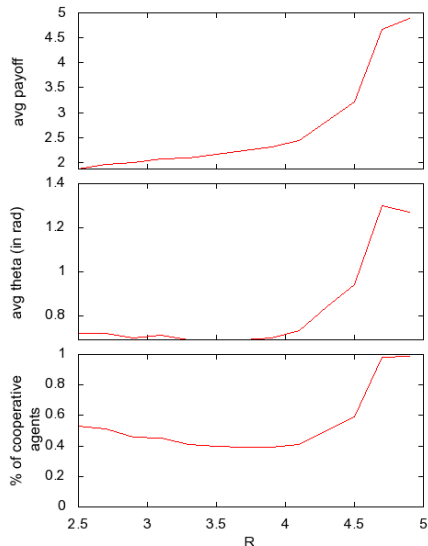


Figure 7: Top graph: effect of  $R$  on average payoff. Middle graph: effect of  $R$  on average theta. Bottom graph: effect of  $R$  on the percentage of cooperative agents.

among themselves. This evolutionary pattern is similar to the one that emerges in the classical rational agent model [8].

Our results shown in Figure 5 also suggested that after the fair player beats the selfish player, the population enters a random drift period. Due to the random mutation, the average theta of the population increases slowly to a point at which there are many highly-cooperative players. Then, mutations introduce selfish players into the populations and their numbers grow quickly until they dominate the entire population. This pattern repeats at least until  $10^7$  generations. This ratifies previous findings on evolutionary cycles of cooperation and defection [10], which shows that our model based social orientations is capable of explaining those findings.

## 5.2 Prisoner's Dilemma with Varying payoffs

We also investigated the effects that different matrix values have on the result of the evolutionary process and the resulting cooperative societies. In these experiments, we varied one of the entries in the PD game matrix while keeping the others constant with their original values as well as keeping the preference relations in the PD matrix, i.e.,  $S < P < R < T$  and  $2R > S + T$ , so that the game will still be a PD game. For each matrix generated in this way, we ran 20 evolutionary simulations with  $10^5$  generations in each run with a total of about 1000 mutant strategies.

Figure 7 shows the effect of varying  $R$  on the average theta and average payoff

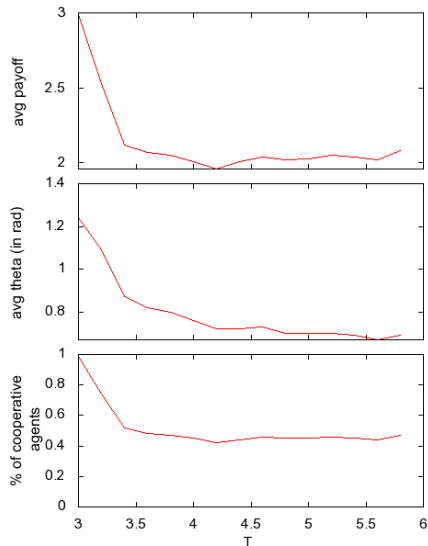


Figure 8: Top graph: effect of  $T$  on average payoff. Middle graph: effect of  $T$  on average theta. Bottom graph: effect of  $T$  on the percentage of cooperative agents.

of the population. We report the average of the data after 1000 generations because often the majority of the groups of players did not emerge before that. Increasing  $R$  provides added incentive to cooperate. Therefore, both the theta and average payoff increase with  $R$ . Note that the payoff almost reaches the maximum (i.e.,  $R$ ) after  $R = 4.7$ , i.e., it becomes always full cooperation when  $R$  is large enough. The bottom graph shows the effect of  $R$  on the percentage of *cooperative* agents which is defined by the portion of agents whose  $\theta$  is greater than the  $\pi/4$  (i.e., the  $\theta$  of a fair player).

Figure 8 shows the effect of  $T$  on the average payoff and the cooperativeness of the population. These results suggest that increasing  $T$  will lead to increase in the incentive to defect. In any situation that can be modeled by a 2x2 game similar to Prisoners' Dilemma, which shows that there is a degradation in the cooperation level. Therefore, both  $\theta$  and payoff decrease when  $T$  increases.

Figure 9 shows the effect of  $P$  on the average payoff and the cooperativeness of the population. In general, the average payoff increases when  $P$  increases. However, unlike the case for  $R$  or  $T$ , the average of  $\theta$  drops sharply when  $P$  is very large compared to  $R$ . These results suggest that increasing  $P$  will lead to an increase in the average payoff, but not increase the cooperativeness of the population. In other words, using our model we are able to notice that there is no one-to-one correlation between the observed average payoff and the society's cooperativeness level. In this case, using previously suggested models one could mistakenly reason that increasing  $P$  and  $R$  has the same effect on the society, while with our new model the difference in the true cooperativeness of

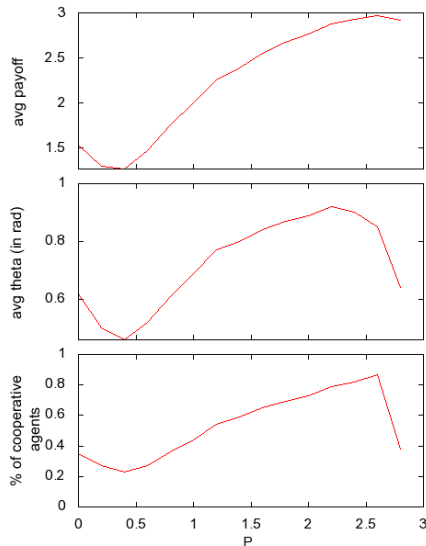


Figure 9: Top graph: effect of  $P$  on average payoff. Middle graph: effect of  $P$  on average theta. Bottom graph: effect of  $P$  on the percentage of cooperative agents.

the society is apparent by looking at the *theta* values of its individuals.

## 6 Conclusions

We have described a formal model that combines game-theoretical analyses for cooperation in repeated 2x2 symmetric games (where  $S < T$ ) with insights from social and behavioral sciences. Our model is not claimed to be the most accurate account of social orientations; rather, it is a simple model that takes the first step in the above direction. Unlike existing models, this formalism captures the notion of prosocial vs. proself orientations exhibited in human behavior and explicitly provides an abstract representation for how a player develops its strategies in repeated games.

We have presented theorems showing how players with different social tendencies interact. Our theorems identify five general steady-state behavioral patterns, that can be explained in terms of the players social orientation values. We have also performed an experimental evaluation of our model using evolutionary simulations in the well-known IPD game. The results of the experiments demonstrated that our model captures the well known behavior patterns in IPD. Furthermore, it allows modeling richer behavior patterns since it does not depend on the particular game matrix.

When we varied the payoffs in the game matrix while keeping the preference relations intact in the PD game, one set of experiments showed that prosocial

tendency increases when the reward (i.e.,  $R$ ) of the game increases or when the temptation (i.e.,  $T$ ) decreases. Another set of experiments identified a class of scenarios in which the evolution simulations produced a population that is not socially-oriented toward cooperation, whereas the average payoff of the population is still high. This result is contrary to the implicit assumption of all previous works that considered cooperative populations, that the high-payoff was assumed to be an indicator for cooperativeness. Our experiment showed that social orientations in a population could be a more realistic representations of the cooperativeness of the entire population.

## References

- [1] Wing Tung Au and Jessica Yuk-ye Kwong. Measurements and effects of social value orientation in social dilemmas: A review. In Ramzi Suleiman, David V. Budescu, Ilan Fischer, and David M. Messick, editors, *Contemporary Approaches to Social Dilemma Research*, pages 71–98. Cambridge University Press, Cambridge, UK, 2004.
- [2] R. Axelrod and WD Hamilton. The evolution of cooperation. *Science*, 211(4489):1390, 1981.
- [3] Robert Axelrod. *The Evolution of Cooperation*. Basic Books, New York, 1984.
- [4] Bogaert, Sandy, Boone, Christophe, Declerck, and Carolyn. Social value orientation and cooperation in social dilemmas: A review and conceptual model. *British Journal of Social Psychology*, 47(3):453–480, September 2008.
- [5] Charles G. McClintock David M. Messick. Motivational bases of choice in experimental games. *Experimental Social Psychology*, 1(4):1–25, 1968.
- [6] Steven de Jong, Karl Tuyls, and Katja Verbeeck. Artificial agents learning human fairness. In *AAMAS*, pages 863–870, 2008.
- [7] Anders Eriksson and Kristian Lindgren. Evolution of strategies in repeated stochastic games with full information of the payoff matrix. In *GECCO*, pages 853–859, 2001.
- [8] Jack Hirshleifer and Juan Carlos Martinez Coll. What strategies can support the evolutionary emergence of cooperation? *Journal of Conflict Resolution*, 32(2):367–398, June 1988.
- [9] Robert Hoffmann. Twenty years on: The evolution of cooperation revisited. *J. Artificial Societies and Social Simulation*, 3(2), 2000.
- [10] L.A. Imhof, D. Fudenberg, and M.A. Nowak. Evolutionary cycles of cooperation and defection. *Proceedings of the National Academy of Sciences*, 102(31):10797, 2005.

- [11] Daniel Kahneman, Jack Knetsch, and Richard H. Thaler. Fairness and the assumptions of economics. *Journal of Business*, 59(4):S285–300, 1986.
- [12] Charles G. McClintock and Scott T. Allison. Social value orientation and helping behavior. *Journal of Applied Social Psychology*, 19(4):353 – 362, 1989.
- [13] John Von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.
- [14] M.A. Nowak. Five rules for the evolution of cooperation. *Science*, 314(5805):1560, 2006.
- [15] M.A. Nowak and K. Sigmund. Tit for tat in heterogeneous populations. *Nature*, 355(6357):250–253, 1992.
- [16] Martin Nowak and Karl Sigmund. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner’s dilemma game. *Nature*, 364(6432):56–58, July 1993.
- [17] Craig D. Parks and Ann C. Rumble. Elements of reciprocity and social value orientation. *Personality and Social Psychology Bulletin*, 27(10):1301–1309, 2001.
- [18] A. Rapoport. *Two-Person Game Theory. The Essential Ideas*. The University of Michigan Press, Ann Arbor, 1966.
- [19] John Maynard Smith. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge, UK, 1982.
- [20] P.A.M. Van Lange, E. De Bruin, W. Otten, and J.A. Joireman. Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. *Journal of personality and social psychology*, 73(4):733–746, 1997.