Emergence of Cooperative Societies in Evolutionary Games

Kan-Leung Cheng, Inon Zuckerman, Ugur Kuter, and Dana Nau
Department of Computer Science, Institute for Advanced Computer Studies, and
Institute of Systems Research
University of Maryland, College Park, Maryland 20742, USA
klcheng@cs.umd.edu, inon@cs.umd.edu, ukuter@cs.umd.edu, nau@cs.umd.edu

ABSTRACT
We utilize evolutionary game theory to study the evolution of cooperative societies and the behaviors of individual agents (i.e., players) in it. 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 the others, and (2) individuals vary in their interpersonal social tendencies, which reflect stable personal orientations that influence their choices. Alongside the formal model we provide an analysis that considers possible interactions between different types of individuals.

We present series of 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 demonstrate that in contrast to previous common knowledge, increasing the value of mutual Reward or increasing mutual Punishment in the Prisoner’s dilemma game do not result in the same type of cooperative society; while increasing P does result in a more cooperative society, increasing P does not.

Categories and Subject Descriptors
I.2.11 [Distributed Artificial Intelligence]; J.4 [Social and Behavioral Sciences]

General Terms
Theory, Human Factors

Keywords
Evolutionary Game Theory, Social Orientations

1. INTRODUCTION
The question of how cooperation evolves between individuals has been studied for years, most notably starting from the seminal work of Axelrod [2]. Existing works on evolution of cooperation typically focused on social dilemmas using normal form games (see [7] for an overview). 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 what the other player does. However, if the game is played repeatedly, unknown amount of times, cooperative behavior might emerge to increase accumulated payoffs.

In this paper, we consider how cooperative societies emerge, given the social tendencies of the individuals. In contrast to traditional studies on game-theoretical model, which typically assume that human behavior is purely motivated by rational self-interest [12], we describe a formalism for modeling the relationship between a player’s social orientation and the player’s strategy. Our formalism is based on the well-founded Social Value Orientation (SVO) theory [4, 3]. The SVO theory is widely accepted based on many empirical studies and conjectures that the social choices people make depend, among other things, on stable personality differences in the situations they interact with the others.

Our formalism captures the notion of social orientations exhibited in human behavior and provides an abstract formal representation for how a player develops its strategies in a repeated game. We present theoretical results showing how players with different social tendencies interact in a class of 2x2 symmetric games. This analysis identifies five general steady state behavioral patterns that can be explained in terms of the players’ varying social orientation values.

We present experiments using evolutionary-game simulations, demonstrating the effects of social orientations on evolution of cooperative behavior in individual players and the emergence of a cooperative society. In one set of experiments, pro-social tendency increases with increasing reward (i.e., R) in the game or with decreasing temptation (i.e., T), confirming the intuitions in the previous works [13].

Previous works on cooperative societies typically used the average payoff of the society as a measure of its cooperativeness [14, 15, 5]: i.e., the higher the average payoff is, the

<table>
<thead>
<tr>
<th>Prisoner’s Dilemma</th>
<th>Player 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Player 1</td>
<td>Cooperate (C)</td>
</tr>
<tr>
<td>Cooperate (C)</td>
<td>(3, 3)</td>
</tr>
<tr>
<td>Defect (D)</td>
<td>(5, 0)</td>
</tr>
</tbody>
</table>

Figure 1: The Prisoner’s Dilemma game.
more cooperative the society is thought to be. However, our experiments also showed that there are scenarios in which the conclusions of the previous works do not hold: one set of experiments showed that the society is not a cooperative one, whereas the average payoff of the society is still high.

2. OUR MODEL

We consider normal-form games for studying social interactions between people or between societies. Among the well-known examples of such games are the Prisoner’s Dilemma (PD), Chicken Game, and Stag-Hunt [16].

Figure 2 shows a generalized form of the payoff matrix for such games, where various constraints on the outcomes can be used to define different classes of social dilemmas. In this paper, we only focus on symmetric games where $T > S$. This means that when a player defects, he/she is expected to get a larger share of the total payoff (we will explain it in details in the Analysis section). 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.

Players’ Social Orientation.

Most studies on Social Value Orientation (SVO) based human behavior typically identify two opposing social-value orientations: pro-self and pro-social. A person with the former gives higher consideration to his/her own payoff, while a pro-social person gives higher regards to the payoff of the agents he/she is interacting with. The social orientation is not an absolute value: it is a range in which at one extreme we have complete pro-self behavior, and at the other extreme complete pro-social behavior. Humans differ from one another by being placed at different locations on that range. In contrast to this diversity, the traditional rationality assumption in game theory dictates that all individuals are completely pro-self, without any difference between one another.

Based on these results from behavioral sciences, 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 [11]. 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 $s_i$ such that $s_i$’s initial point is at the origin of the social-orientation space. We represent $s_i$ by the angle, $\theta_i$, between $s_i$ 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 pro-social or pro-self behavior in the game.

For example, when $\theta_i = 0$ then Player $i$ acts as a pure individualistic. 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 pro-social, i.e., it never attempts to maximize its own payoff, but rather it tries to increase the payoff of the other player.

Players’ Models of Repeated Games.

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

$$\vec{g} = (p_i, p_j)$$

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:

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

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

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 $a$. We only consider the set of action whose expected allocation is Pareto optimal, i.e., its expected utility vector is not dominated by other’s expected utility vector.

Players’ Behavioral and Strategical Tendencies.

During the game, each player aims to bring the direction of game model closer to its social-orientation vector, $s_i$. 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 agents will try to do utility-maximization on its own payoff. In other words, the theta equals to zero and social orientation equals to $\langle 1, 0 \rangle$.

Let $\vec{g} = (p_i, p_j)$ 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 $\bar{g} + \bar{p}_o$. The angle $\alpha$ is computed as follows:

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

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

Thus, each player will choose an action $a$ such that $a = \arg\max_{a \in A} \cos \alpha_a$.

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., $\theta = \pi/4$) interacts with another player in an IPD game. For IPD, $\bar{p}_C = (3, 2, 4)$, $\bar{p}_D = (\frac{3}{2}, 4)$. At the beginning, $\bar{g} = (0, 0)$, $\alpha_C$ 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 $(0, 5)$, 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 $(1, 6)$, $\alpha_D$ 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., $\bar{g} = (p, p)$ for some $p$. In other words, a fair player in IPD behaves exactly the same as the well-known Tit-For-Tat (TFT) strategy.

3. 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 the game. This section presents an exhaustive analysis of such dynamics based on our model described earlier.

In the following analysis, 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 $r_i = \cos \theta_i + \sin \theta_i$ and $r_j = \cos \theta_j + \sin \theta_j$.

We define the following ratios for each action, $C$ and $D$, in the 2x2 symmetric game: $r_{C} = \frac{T + P}{(S + T) + (T + P)}$ and $r_{D} = \frac{S + T}{(S + T) + (T + P)}$. Intuitively, these ratios describe the expected share of payoff 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 $\bar{g} = (p_i, p_j)$ 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 steady state, Player $i$ cooperates whenever he/she is satisfied with his/her current ratio, i.e., his/her current ratio is greater than or equal to his/her preference ratio (i.e., $g_i \geq r_i$), 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., pro-self), then they both always defect and get $P$ at each game in steady state.

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

For example, in a PD 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., pro-social), then they both 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 $g_j$ moves toward $\frac{S}{S+T} < 1 - r_j$. Otherwise, both players cooperate and $g_i$ moves toward $\frac{P}{S+T} = 0.5$.

For example, in a PD 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% of 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 $r_i + r_j = 1$, then there are two cases:

- when $r_j \geq \frac{T}{S+T}$, 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 $g_j$ 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 $T + S - r_i(T + S)$.

For example, in a PD 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 \( r_j > 1 - r_i > 0.5 \), then there are two cases:

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

Proof. The proof for the first case is just the same as that of Theorem 3 above. The proof for 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.

For example, in a PD 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 pro-self 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 \( \frac{1}{n} \) and Player \( j \) will get \( \frac{1}{m} \) at each game on average.

Theorem 5. If \( r_i < 1 - r_j < 0.5 \), then there are two cases:

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

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 cooperate and get \( S \) and Player \( j \) defects and gets \( T \). When \( g_i \geq r_i \), both players cooperate and get \( R \). In steady state, they will get \((S, T)\) or \((R, R)\) in each game in a way that \( r_i \) 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 \( \bar{p}_i = r_i(\bar{p}_i + \bar{p}_j) \), \( \bar{p}_i = Rn_{CC} + S(1 - n_{CC}) \) and \( \bar{p}_j = Rn_{CC} + T(1 - n_{CC}) \). Solving them, we can obtain the above formula.

For example, in a PD game, let \( r_i = 0.6 \) and \( r_j < 0.4 \). Then, as resources are plentiful (as \( r_i + r_j < 1 \)) and Player \( i \) is more pro-social 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 \( \frac{1}{n} \) and Player \( j \) will get \( \frac{1}{m} \) at each game on average.

4. Experiments

We have performed several experiments using evolutionary simulations in order to investigate emergence of cooperative populations based on social orientations of individuals. We used replicator dynamics for evolutionary simulations [17]. We used the well-known “infinite population” setup for initializing the population as described in [14, 15, 5] for replicator dynamics in evolutionary simulations. 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 the rounds in 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 was discarded from the population.

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

4.1 Prisoner’s Dilemma

Figure 5 shows the average population \( \theta \) and average population payoff of an evolutionary process of over \( 10^4 \) generations for 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 who have higher \( \theta \) values dominate the population quickly. Then, selfish players (with low \( \theta \) value) emerge gradually to dominate after about 4000 generations. At around the 10000th 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 in the classic rational agent model [15, 5]. 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 some mutants introduced in the population with \( \theta_{mutant} \approx \pi/4 \). These fair players avoided the exploitation of selfish players with \( \theta_{selfish} < \pi/4 \), 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 \( \pi/2 \);
...this evolutionary pattern is similar to the one that emerges in the classical rational agent model [6].

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 [8].

4.2 Variations of Prisoner’s Dilemma

We also investigated the effect of changing the values in the matrix in the evaluation of 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 is still PD. For each matrix generated in this way, we ran 20 evolutionary simulations with 10^7 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 of the population. We report the average of the data after 1000 generations because a majority of a group of players did not emerge in the first 1000 evolutionary simulations. 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 θ is greater than the π/4 (i.e., the θ 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 of the incentive to defect. In any situation that can be modeled by a 2x2 game similar to Prisoners’ Dilemma, that means that the degradation in the cooperation level. Therefore, both θ 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 θ drops sharply when P is very large compared to R. These results suggest that increasing P will lead to increase in the average payoff, but not 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 the society is apparent by looking at the theta values.

5. RELATED WORK

In order to study the evolution of cooperation, an iterated variance of the game is often used in which the game is played repeatedly an unknown amount of times. Since Axelrod’s Iterated Prisoner’s Dilemma (IPD) tournament [1] that looked for the winning strategy, there was a large amount of work that researched various aspects of the model [7].

In addition to the work cited in previous sections of this paper, several other approaches in social and behavioral science have been investigated to model how humans’ social tendencies and the effects of those tendencies on the individual’s choices and actions. Many experiments have shown that humans explicitly take the outcome of the other player into account when considering taking a cooperative action or an individual one. For instance, many experiments in the well-researched ultimatum 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 [10].

Since the seminal work by Messick and McClintock [4], the Social Value Orientation (SVO) theory has developed into a class of theories that extends the original work in different ways and uses different names, including social values, interpersonal orientation, social orientation and others (see [3] for a good review).

The validity of SVO based theorems was shown both in laboratories and field studies that indicate that pro-social
generally cooperate more and show greater concern for the effect of their actions on the well being of others and on the environmental in general. For examples we refer to [11] that shows that pro-social students were more willing to contribute time to help others, and to [9] that shows that pro-social tend to take more pro-environmental and collective policies than self-interest actions.

Previous works on the evolution of cooperation typically used the average payoff of the population as a measure of the cooperativeness [14, 15, 5]. Our definition is based on social orientations of players: i.e., higher average theta implies more cooperative society.

6. CONCLUSIONS AND FUTURE WORK

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 pro-social vs. pro-self 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 PD game. The results of the experiments demonstrated that our model captures the known behavior patterns in PD. 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 pro-social 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 conclusions of all previous works that considered cooperative populations: in the previous works, the high-payoff was assumed to be an indicator for cooperativeness, whereas our experiment showed that social orientations in a population could be more realistic representations of the cooperativeness of the entire population.

In the near future, we intend to do an extensive evaluation of our approach. We also plan to generalize our model and analysis to repeated heterogeneous games, where different generations may play different games. We believe the social orientation of the players in such situations will provide insights on how they decide on their strategies and how/if cooperation evolves.
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.

Acknowledgments.
This work was supported in part by NGA grant HM1582-10-1-0011, ARO grant W911NF0920072 and AFOSR grant FA95500610405. The opinions expressed in this paper are those of the authors and do not necessarily reflect the opinions of the funders.

7. REFERENCES