Predicting Agent Strategy Mix of Evolving Populations

Sabyasachi Saha
Department of Math & Computer Science
The University of Tulsa
sabyasachi-saha@utulsa.edu

Sandip Sen
Department of Math & Computer Science
The University of Tulsa
sandip-sen@utulsa.edu

ABSTRACT

We study agent societies where self-interested agents interact repeatedly over extended time periods. In particular, we are interested in environments where agents can form mutually beneficial relationships by exchanging help but an agent would rather receive help than give it. Evolutionary tournaments with competing help-giving strategies can model scenarios where agents periodically adopt strategies that are outperforming others in the population. Such experiments, however, can be computationally costly and hence it is difficult to prescribe a rational strategy choice given environmental conditions like task mix, strategy distribution in the population, etc. A preferred approach, pursued in this paper, is to analytically capture the dynamics of the strategy mix in the population under an evolutionary tournament. Such an analytical model can be used to predict the evolutionarily dominant strategy, the rational strategy choice.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence

General Terms

Experimentation,Performance

Keywords

Cooperation, reciprocity, agents, adaptation

1. INTRODUCTION

With the burgeoning of agent based electronic commerce, recommender systems, personal assistant agents, etc. it is becoming increasingly clear that agent systems must interact with a variety of information sources in an open, heterogeneous environment. One of the key factors for successful agent based systems (ABSs) of the future would be the capability to interact with other ABSs and humans in different social and role contexts and over extended periods of time. Research in societal aspects of agent strategies has been relatively scarce [9]. Whereas economic models can provide a basis for structuring agent interactions [14], other non-monetary approaches [1, 2, 3, 4, 5, 8, 11] may provide effective solutions in certain situations. We assume that typical real-world environments abound in *cooperation possibilities*: situations where one agent can help another agent by sharing work such that the helping cost of the helper is less than the cost saving of the helped agent. Whereas previous work have identified agent strategies that allow agents to take advantage of cooperation possibilities in their environments, there is no existing formal framework to characterize the performance of different competing strategies given environmental parameters. We develop an analytical model to predict the dynamics of such systems assuming agents adopt strategies based on their past performance in the environment.

Sen *et al.* [10, 11] have presented strategies that promote cooperation among homogeneous groups and can resist exploitation by malevolent agents in heterogeneous groups. Such strategies can lead to both improved local performance for individual agents and effective global performance for the entire system. A restrictive assumption in this line of work has been that agents have fixed strategies. For example, they have assumed that agents with specified strategies interact repeatedly over a sustained period of time and their effectiveness is calculated as a function of the total cost incurred to complete all assigned tasks. The resultant performance reflects cost incurred for local tasks, cost incurred to help other agents with their tasks, and savings obtained from others when help was received.

A more realistic scenario would be to give an agent the freedom of choosing from one of several of these strategies and to change its strategy as dictated by the environmental conditions. An agent may be inclined to adopt a strategy if agents using that strategy is observed to be performing better than others. Such a strategy adoption method leads to an evolutionary process with a dynamically changing group composition of agent strategies [12, 13]. In this paper, we present a mathematical model of the dynamics of the agent population. Our goal is to identify the dominant regions for different strategies based on environmental conditions like initial population composition, number of tasks agents need to accomplish, the strategy selection criteria, etc.

We consider a problem domain where each of the agents are assigned some tasks. The cost of executing a task can be reduced or eliminated if help is obtained from another agent. An agent is an *expert* of a task type if it requires less cost than others to accomplish tasks of that type. Hence, non-experts in a task type will benefit by receiving help for tasks of that type from corresponding experts. Experts in differing task types can form mutually beneficial relationships by exchanging help. After all agents have finished processing their assigned tasks, their performances are tallied. This comprises one evaluation period, or *generation*, of the current agent strategies. The strategies adopted by the agents in the next evaluation period is determined by a performance-proportionate scheme where the probability with which an agent adopts a strategy increases with the average performance of agents employing that strategy in the most recent evaluation period. Thus, it is likely that more agents are produced with strategies that generated above-average performance. As a result, if a strategy produces better performance in one evaluation period compared to other strategies, we are likely to see more individuals adopting that strategy in the next evaluation period. This generational scheme is semantically equivalent to every agent periodically selecting its strategy based on the current relative performance of the set of available strategies. This generational approach is akin to work on identifying "evolutionary stable strategy" [6].

The goal of this paper is to identify the dominant strategies under different environmental conditions including initial population composition, the frequencies of the tasks assigned to the agents and the selection criterion used for population evolution. We present a mathematical analysis of the dynamics of the agent population. Using this model we can predict the strategy that will eventually dominate the population given the initial configuration.

It is evident that if each agent was to perform only one task, i.e., the number of interactions between two agents were at most one, agents who seek but do not provide help (selfish agents) will outperform those who offer and return help (reciprocative agents). On the other hand, if the agent group were completely stable, i.e., the agents had an infinite, or very large, number of tasks to perform and hence interacted with each other infinitely often, the reciprocative strategy will dominate the selfish strategy as after some time only reciprocative agents will receive the benefit of help from each other. The switch in dominance happens at an intermediate value of the number of tasks per agent, and is dependent on other environmental factors like initial group composition. The proposed mathematical analysis and predictive model is to identify the regions where selfish or reciprocative agents are dominant. It is also interesting to investigate if there are situations where mixed populations of reciprocative and selfish agents are sustained through the evolutionary process. The significance of such a predictive model is that given the initial population configuration and number of tasks, an agent can predict the evolutionarily dominant strategy and hence adopt the same.

2. ADAPTATION VIA RECIPROCITY

A significant body of work by mathematical biologists or economists on the evolution of altruistic strategy deals with the idealized problem called Prisoner's dilemma [2] or some other repetitive, symmetrical, and identical 'games'. To consider a well-known study in this area, Axelrod demonstrates that a simple, deterministic reciprocal scheme or the *tit-for-tat* strategy is quite robust and efficient in maximizing local utility [2]. Sen argues that the assumptions underlying the simple reciprocative strategy, repeated play of identical games, simultaneous moves, etc. are untenable in most real-life situations [10]. Though Sen's formulation has the right motivations, it is limited by the constraint of not allowing agents to change strategies based on experience.

The evaluation framework used by Axelrod considers an evolving population composition by allowing propagation of more successful strategies and elimination of unsuccessful ones. In this paper, we evaluate the variants of exploitative and reciprocative strategies suggested by Sen [11] in a generational framework as used by Axelrod [2] which can also be used to model agents adopting strategies based on their past performance. This allows us to see what strategies emerge to be dominant or are evolutionarily stable.

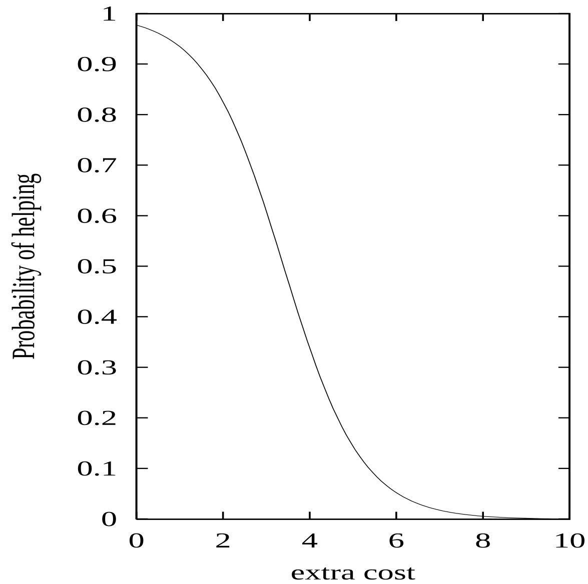

Figure 1: Probability distribution for accepting request for cooperation.

In our evolutionary framework we study two agent types with fixed strategies: those who always help (philanthrops) and those who never help (selfish). We also used Sen's reciprocative agents, described next, which use a probability function involving balance of past exchanges to determine whether or not to help another agent [10].

3. PROBABILISTIC RECIPROCITY

Each agent is assigned to carry out T tasks. The mth task assigned to the ith agent, t_{im}, will cost it C_{ij} if the mth task is of type j. However, if agent k carried out this task together with its own tasks, the cost incurred for task j by agent k is C_j^{kl} (no cost is incurred by agent i), where agent k is doing tasks of type l. If $C_{ij} > C_j^{kl}$, there exists a cooperation possibility as agent k can help agent i save C_{ij} by incurring a cost of only C_j^{kl}.

S_{ik} and W_{ik} are respectively the cumulative savings obtained from and extra cost incurred by agent i from agent k over all of their previous exchanges. Also, $B_{ik} = S_{ik} - W_{ik}$ is the balance of these exchanges (note that, in general, $B_{ik} \neq -B_{ki}$).

Sen [10] proposes a probabilistic decision mechanism that satisfies a set of criteria for choosing when to honor a request for help that was described at the end of the previous section. The probability that agent k will carry out task t_{ij} for agent i while it is carrying out its task t_{kl} is given by:

$$Pr(i,k,j,l) = \frac{1}{1 + \exp^{\frac{C_j^{kl} - \beta * C_{avg}^k - B_{ki}}{\tau}}}, \qquad (1)$$

where C_{avg}^k is the average cost of tasks performed by agent k, and β and τ are constants. This is a sigmoidal probability function (not a probability distribution) where the probability of helping increases as the balance increases and is more for less costly tasks. A sample probability distribution is presented in Figure 1. β can be set to a low value to move the probability curve left (less inclined to cooperate) or to a high value to move the curve to the right (more inclined to cooperate). Initially, $B_{ki} = 0$ for all i and k. At this point the probability that an agent will help another agent by incurring an extra cost of $\beta * C^k_{avg}$ is 0.5. τ can be used to control the steepness of the curve. For a very steep curve approximating a step function, an agent will almost always accept cooperation requests with extra cost less than $\beta * C^k_{avg}$, but will rarely accept cooperation requests with an extra cost greater than that value. In this paper we plan to include the following philanthropic, selfish and reciprocative agent types [10]:

Philanthropic agents: Agents that always honor a cooperation request irrespective of past experience.

Selfish agents: Agents who ask for help but never return favors. Selfish agents can thrive on the benevolence of philanthropic agents.

Reciprocative agents: Agents that use the probabilistic reciprocity scheme described above.

We also study a variant of the reciprocative strategy [11]:

Earned-Trust based reciprocative agents: While evaluating a request for help, these agents consider balances of only those agents with whom they themselves have favorable balances. In place of using B_{ki} in Equation 1, a conservatively trusting reciprocative agent k uses $\sum_{j \neq i \wedge B_{kj} > 0} B_{ji}$ while calculating the probability of helping agent i. This behavior is required to counter false balance reporting by exploitative agents.

4. AGENT POPULATION DYNAMICS MODEL

In this section, we present the mathematical analysis of the agent population dynamics. Using this model, an agent can predict the population configuration after each time period given its environmental configuration and without doing any experimentation or exploration in the domain. We consider the *proportion of selfish*, s, and *proportion of philanthropic*, p, agents as the independent variables. These two variables completely determine the initial strategy distribution in the population, because the sum of the proportions of the three different agent types must be one. Given initial strategy distribution in the population and number of tasks in the domain, T, we want to we want to calculate the evolutionarily stable and dominant strategy.

Let there be N agents in the environment. In the initial population, $\langle p_s, p_p, p_r \rangle$ is the proportion of the selfish, philanthropic and reciprocative agents. From the discussion at the end of Section 1, the reciprocative strategies are likely to be dominant for highly stable environments (large T) and selfish strategies are expected to be dominant in very dynamic groups (small T). Our goal then is to identify a decision surface that separates the regions of dominance of selfish and reciprocative strategies in the three-dimensional space defined by p_s, p_p, and T. The surface will have the property that the reciprocative strategy will be evolutionary dominant for any configuration corresponding to a point lying above this surface, i.e., for a higher value of T.

In this paper, we have considered two types of tasks: *type1* and *type2*. The proportion of task types $\langle tp_1, tp_2 \rangle$ is assumed to be equal, *i.e.* $tp_1 = tp_2 = 0.5$. Given the proportion of the initial population, one can find out the number of different types of agents. $N_{r,l}$, $N_{p,l}$ and $N_{s,l}$ are the number of *reciprocative, philanthropic* and *selfish* agents respectively, which are expert in task type $l = 1, 2$, where $N_{r,l} = N * p_r * tp_l$, $N_{p,l} = N * p_p * tp_l$ and $N_{s,l} = N * p_s * tp_l$. To predict if there exists any evolutionary dominant strategy, we have to predict the expected population configuration and their performances in each period given this initial configuration of the agent population and the total number of tasks per agent.

$P(i, k, j, l)$ is the probability that agent k, if it is asked, will help agent i for a particular task t of type j when agent k is expert in tasks type l. This $P(i, k, j, l)$ is defined as,

$$
\begin{aligned}
P(i,k,j,l) &= \frac{1}{1+\exp^{\frac{C_j^{kl} - \beta * C^k_{avg} - B_{ki}}{\tau}}}, \text{if k is reci \&d } j = l \\
&= 1, \text{if k is philanthropic \& } j = l \\
&= 0, \text{otherwise.}
\end{aligned}
$$

B_{ki} is defined as $B_{ki} = S_{ki} - W_{ki}$ for all i, k, where S_{ki} and W_{ki} are cumulative savings from and extra cost incurred by agent k from agent i. Initially, $S_{ki} = W_{ki} = 0$ and hence $B_{ki} = 0$.

We calculate $P_1(i, k, j, l)$, the probability that for a task of type j in the task distribution of agent i, agent k, an expert in task type l, will be the one to help agent i. This event corresponds to the situation that all the agents asked before agent k will refuse to help and agent k will help.

$$
\begin{aligned}
P_1(i,k,j,l) &= \sum_{a=1,\neq i}^{N_{r,l}} [Pr(L_{k,a} \cap R(i,a,j,l)) * P(i,k,j,l)], \\
&\quad \text{if k is reciprocative} \\
&= \sum_{a=1,\neq i}^{N_{r,l}} [Pr(L^p_{k,a} \cap R^p(i,a,j,l))], \\
&\quad \text{if k is philanthropic} \\
&= 0, \text{if k is selfish} \qquad (2)
\end{aligned}
$$

where $L_{k,a}$ is the event that k is selected as the ath among the $N_{r,l}$ reciprocative agents that are expert in tasks of type l, i.e., after $a-1$ reciprocative agents expert in task type l. $R(i, a, j, l)$ is the event that all those $a - 1$ agents refuse to help agent i for tasks of type j. So,

$$Pr(L_{k,a} \cap R(i,a,j,l)) = Pr(L_{k,a}) * Pr(R(i,a,j,l)|L_{k,a}), \quad (3)$$

where

$$Pr(L_{k,a}) = \frac{\begin{pmatrix} N_{r,l} - 1 \\ a - 1 \end{pmatrix}}{\begin{pmatrix} N_{r,l} + N_{p,l} \\ a \end{pmatrix}} \quad (4)$$

and as agent decisions are independent,

$$Pr(R(i,a,j,l)|L_{k,a}) = \prod_{t=1}^{a-1} (1 - P(i, a_t, j, l)), \quad (5)$$

where a_t is the tth agent selected for seeking help. Since all agents are starting with the same probabilities to help another agent, it is immaterial in which order or who exactly are the agents that are selected before agent j is selected.

The probabilities $Pr(L^p_{k,a})$ are similar to the probability $Pr(L_{k,a})$ except that the numerator in the expression for the former contains $N_{r,l}$ instead of $N_{r,l} - 1$ (this is because for a philanthrop, all the reciprocatives may have already been asked). Also, the probability $Pr(R^p(i,a,j,l)|L^p_{k,a})$ is similar to the probability $Pr(R(i,a,j,l)| L_{k,a})$ except that the expression for the former uses a instead of $a - 1$ range of the product (for similar reasons as above).

Let us now consider the expected change of balance between two reciprocative agents i and k for a particular task of type l. One can compute the expected savings and expected spending of agent i for agent k by

$$S_{ik} = S_{ik} + \sum_{l=1}^{2} P_1(i, l, k, l) * C_{i,l} * pt_l,$$

$$W_{ik} = W_{ik} + \sum_{l=1}^{2} P_1(k,l,i,l) * C_{i,l} * pt_l$$

and $B_{ik} = S_{ik} - W_{ik}$. Using these B_{ik} values we can again find out the probability of helping for the next task. So, we can calculate the performance of an agent as the expected net wealth it will generate after processing all the assigned tasks. The expected net wealth generated by an agent is the total of the expected balance with the other agents.

At the end of each evaluation period, i.e., after every agent completes all the tasks they are assigned and those that it too on to help other agents, the performances are tallied. At this point, new agent strategy assignments are made as follows: for each agent i, two agents are selected randomly from the population without replacement. Then, of these two selected agents, the strategy of the one with higher performance is adopted by agent i[1].

This will determine our new expected population with $\langle p_s, p_p, p_r \rangle$ as the proportion tuple. This proportion will be approximated as the probability of an agent choosing the corresponding strategies. We now calculate the expected probability that an agent will adopt reciprocative strategy *i.e.* p_r. It is defined as,

$$\begin{aligned} p_r &= Pr(reci, reci) + Pr(reci, self) * Pr(reci \geq self) \\ &\quad + Pr(reci, philan) * P(reci \geq philan), \end{aligned}$$

where

$$Pr(reci, reci) = \frac{\begin{pmatrix} \lfloor N * p_r \rfloor \\ 2 \end{pmatrix}}{\begin{pmatrix} N \\ 2 \end{pmatrix}},$$

$$Pr(reci, oth) = \frac{\begin{pmatrix} \lfloor N * p_r \rfloor \\ 1 \end{pmatrix} * \begin{pmatrix} \lfloor N * p_{oth} \rfloor \\ 1 \end{pmatrix}}{\begin{pmatrix} N \\ 2 \end{pmatrix}}$$

where, oth = selfish or philanthropic. Then the expected value of $Pr(reci \geq oth)$ can be found from the previously calculated expected performance. Similarly, we can find out p_p and p_s values and determine the expected new strategy distribution in the population in the next generation. We repeat this process until the agent population becomes homogeneous, i.e., all agents use the same strategy. There may exist situations where one single strategy may not be evolutionary dominant. In such situations, we repeat this process up to a finite horizon.

So, for a particular initial agent population and number of tasks available to each agent one can find out the dynamics of expected population configuration and the performances of different agents. One can identify if there exists any evolutionarily dominant strategy. Therefore an agent can choose a utility-maximizing strategy if it was cognizant of the environmental factors like number of tasks per agent, how many evolution periods it wants to stay and the starting strategy distribution in the population.

5. ANALYSIS OF THE POPULATION DYNAMICS

The decision surface that separate the preferred strategies in the space defined by initial proportions of selfish and philanthrop agents are shaped by the following propositions

[1]Selection of the best from a set of randomly selected candidates is known as *tournament selection* in the genetic algorithms literature [7].

Proposition 1: *If philanthropic agents are present in the population, then the probability that an agent (selfish or reciprocative) will receive help from at least one of the agents in the population is 1.*

Proposition 2: *If both philanthropic and selfish agents are present in the population, the philanthropic agents cannot dominate the selfish agents.*

Proposition 3: *In the presence of philanthropic agents, reciprocative agents cannot dominate the selfish agents.*

Since the philanthropic agents help others without considering their help-giving behavior, others benefit at their expense and it is impossible for them to dominate the other strategies, as stated in Proposition 2. Moreover, their presence help the exploitative agents dominate the environment. Since, the selfish agents always exploit the philanthropic agents without helping back, in their presence the performance of the philanthropic agents rapidly deteriorate and hence agents using that strategy switches to other strategies. Also, as stated in Proposition 3, in the presence of philanthropic agents, a reciprocative agent cannot perform better than the selfish agents. Only after the philanthropic strategy become *extinct*, does the reciprocatives have any chance of dominating the selfish agents. The remaining critical consideration is the help-exchange that takes place in a group consisting only of selfish and reciprocative agents. Within each evaluation period, the reciprocatives initially help the selfish agents and expects help from them. It takes some time for them to recognize the selfish agents and stop helping them. Once the reciprocative agents stop helping the selfish agents, the selfish needs to do all of the work by themselves and incur more cost compared to the reciprocatives who continue to exchange help among themselves. But, if the number of tasks within an evaluation period is small, the reciprocatives will not be able to compensate the initial cost incurred for helping the selfish agents.

To demonstrate the expected evolution of the population over a single evolutionary run, we plot the proportion of the strategies, calculated by our analytical model, over different evaluation periods in Figure 2. The environmental parameters are as follows: $N = 200$, $\beta = 4.0$, $\tau = 0.9, p_p = 0.3, p_s = p_r = 0.35$. In the figure, we present three different plots corresponding to task numbers from ranges which leads to three distinct evolutionary patterns.

In the left plot of Figure 2, T has a high value of 200. After two evaluation periods, the floor value of the expected number of philanthrops drop to zero. In the presence of philanthrop agents, the expected performance of the selfish agents was better than that of reciprocative agents. But after the philanthrops die off, the primary source of help for selfish has been driven to extinction. At this point, the reciprocatives are expected to outperform the selfish agents and increasingly more agents are expected to adopt reciprocative strategy. After about six evaluation periods the expected number of selfish agents drop down to zero leaving reciprocatives as the absolutely dominant strategy.

In the middle plot of Figure 2, T has a very low value of 5 which will not allow the reciprocatives to compensate for the initial cost it incurred by helping the selfish agents before they were recognized as such. So, the selfish agents will outperform the reciprocative agents and become the dominant strategy when tasks in the domain is very low. From the figure we see that in contrast to the previous case, the percentage of selfish agents in the population monotonically grows in the population even after the philantrhops die off. After only a few periods, the reciprocatives also become extinct, leaving the selfish as the absolutely dominant strategy.

In the right plot of Figure 2, an intermediate value of $T = 50$ is used. In a very interesting development, we find that neither the selfish and reciprocatives becomes evolutionary dominant strat-

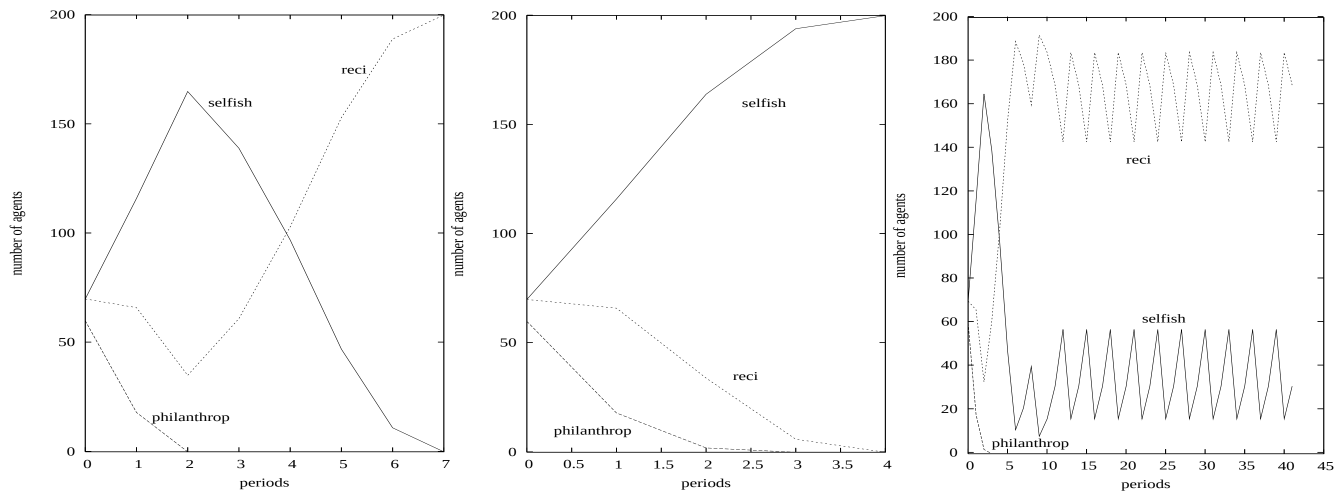

Figure 2: Expected number of agents of different types over evaluation periods. $p_s = p_r = 0.35,\ p_p = 0.3,\ N = 200,\ T =$ **200 (left), 5 (middle) and 50 (right) respectively.**

egy in this environment. Rather we have a periodically oscillating mixed population. The oscillation takes place because, with a small enough number of selfish, the number of reciprocatives is large enough for a selfish agent to distribute its exploitation among the large number of reciprocatives. Thus a given reciprocative does not get have enough interactions with the selfish to unmistakably identify them. This allows selfish agents to outperform the reciprocatives and hence increase their proportion in the next generation. Now, however, there is not enough reciprocatives for each selfish agent to benefit from. As a result, they cannot match the performance of reciprocatives and hence their proportion decreases in the next evolutionary period. This oscillation continues after the short initial period when the philanthrops die off and the reciprocatives grow to become the majority in the population. The oscillation in the population, described above, continues indefinitely and no one strategy is evolutionarily stable for such intermediate values of T.

p_s and p_p are the independent variables in Figure 3, with values varying from 0.1 to 0.8 in steps of 0.1. For each initial population distribution, we find the minimum number of tasks required for the reciprocative agents to become evolutionary dominant. For any point above the resulting decision surface, i.e., for a higher T for the same initial population distribution, the reciprocative strategy is the evolutionarily dominant strategy. The maximally dark, flat, region corresponds to invalid populations, *e.g.* the point $\langle 0.8, 0.5, *\rangle$ does not exist as the agent proportions must sum to ≤ 1.

From Propositions 2 and 3, we see that the selfish agents will benefit in the presence of philanthropic agents and grow in proportion over generations as long as the philanthrops do not become extinct. The evolutionarily dominant strategy is thereafter determined by the distribution of selfish and reciprocative agents when the philanthropic agents become *extinct*. To better understand the population dynamics after the philanthrops are eliminated, we analyze performance of agents when only reciprocative and selfish agents are present in different proportions in the population. We present in Figure 4, for three T values, the expected average payoff received by these strategies, calculated from our analytical model, for different percentage of selfish agents with no philanthrop agents in the environment. Note that the range of p_s values for which the selfish agents have better performance than the reciprocatives increases as T decreases.

It is interesting to note that for small and intermediate values of

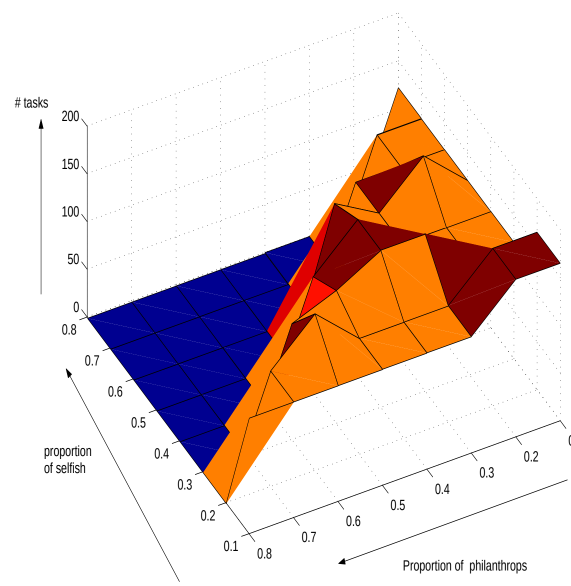

Figure 3: For a given initial population distribution, and for number of tasks equal or more than the decision surface, reciprocative strategy will be evolutionarily dominant.

T, the payoff curves for reciprocative and selfish agents intersect. Such points of intersection correspond to population configurations with mixed strategy equilibrium. Since the average performance of the two different types are same, their numbers will not change in the subsequent periods. Such equilibria are then fundamentally different from the population oscillations we have observed in Figure 2.

We now use these plots to provide further explanation of results presented in Figure 2. From the left plot of Figure 2 ($T = 200$), we see that $p_s \approx 0.8$ when philanthrops die off. But from the right plot of Figure 4, i.e., for $T = 200$, we see that the reciprocatives dominate the selfish for $p_s \approx 0.8$, and hence the selfish proportion will decrease over generations until reciprocatives become evolutionarily dominant. From the middle plot of Figure 2 ($T = 5$),

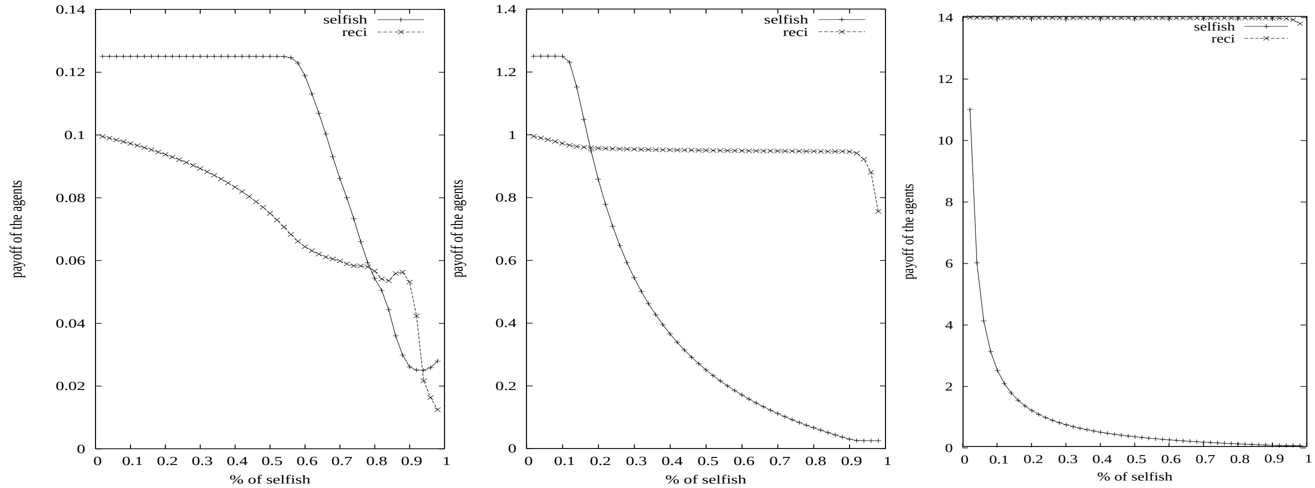

Figure 4: Expected payoff of selfish and reciprocative agents. $T = 5$ (left), 50 (middle) and 200 (right).

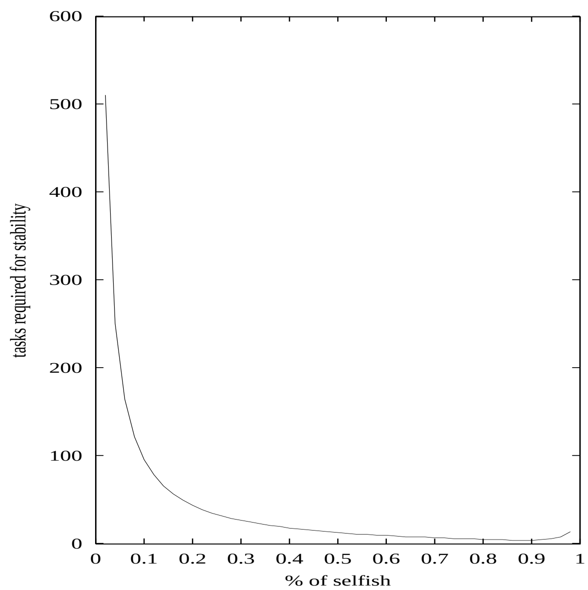

Figure 5: T value for which expected performance of reciprocatives and selfish are equal.

we see that $p_s \approx 0.95$ when philanthrops die off. From the left plot of Figure 4, i.e., for $T = 5$, we see that the selfish performs better than the reciprocatives at $p_s = 0.95$, and hence the selfish proportion improve to become evolutionarily dominant. Note that a slightly lower selfish percentage at the extinction of philanthrops might lead to oscillatory population distributions.

From the right plot of Figure 2 ($T = 50$), we see that when philanthrops die off, p_s is just more than 0.8. From the middle plot of Figure 4, i.e., for $T = 50$, we see that the reciprocative outperform the selfish and will therefore increase in proportion. What follows is more interesting: with falling selfish percentage, their performance improves. In the oscillatory phase of this plot, the selfish proportion varies between approximately 0.1 and 0.3 in successive generations. The corresponding points lie on different sides of the equilibrium point of equal payoff for the two strategies in the middle plot of Figure 4. Compared to the reciprocatives, the selfish payoff is lower when $p_s = 0.3$ and higher when $p_s = 0.1$. This explains the oscillating population under a fitness proportionate selection scheme.

To gain a further perspective on the relative dominance of the two strategies, we plot, in Figure 5, the number of tasks where selfish and reciprocative agents are expected to have identical performance. The plot corresponds to an exponentially decaying function. We see that a small number of selfish agents can exploit the reciprocatives for an extended period of time. As each reciproactive agent individually needs several interactions to recognize selfish behavior, a selfish agent can, in the presence of many reciprocatives, get helped with a large number of tasks. Additionally, oscillatory mixed strategies with small number of selfish agents is possible even for relatively large T values.

This problem of residual selfish agents, or free riders, is addressed by the *earned trust based* (ETB) reciprocity model. This is a reputation based model where an agent, deciding on whether or not to provide help to another agent, considers the feedback of other agents with whom it has had good balance of help exchange. In Figure 6, we present plots similar to Figure 4 except here ETB reciprocative agents are used. We see that the ETB agents almost always outperform the selfish agents except when the number of selfish is very high and T is very small. As ETB reciprocative agents share experiences, they can recognize the selfish even when it has little interaction with a selfish. The only exception where selfish wins out is when their proportion is very high, and the corresponding extensive exploitation by many selfish agents in the first few tasks, before ETB can recognize them, cannot be compensated by ETB as there are only a few tasks to be performed. But with T values of only as little as 15, the ETB reciprocatives start completely outperforming the selfish agents. We also do not observe any oscillatory strategy with small number of residual selfish agents in the population with ETB agents at these T values.

6. CONCLUSION AND FUTURE WORK

In this paper, we have presented an analytical model for predicting the mix of different strategy distributions under an evolutionary scheme. Related population evolution models have been extensively studied in population biology literature with stylized, simple interactions with little or no history or memory. We are not aware of studies like the one presented here, which tries to predict evolution of strategy distributions of rational agents with complex

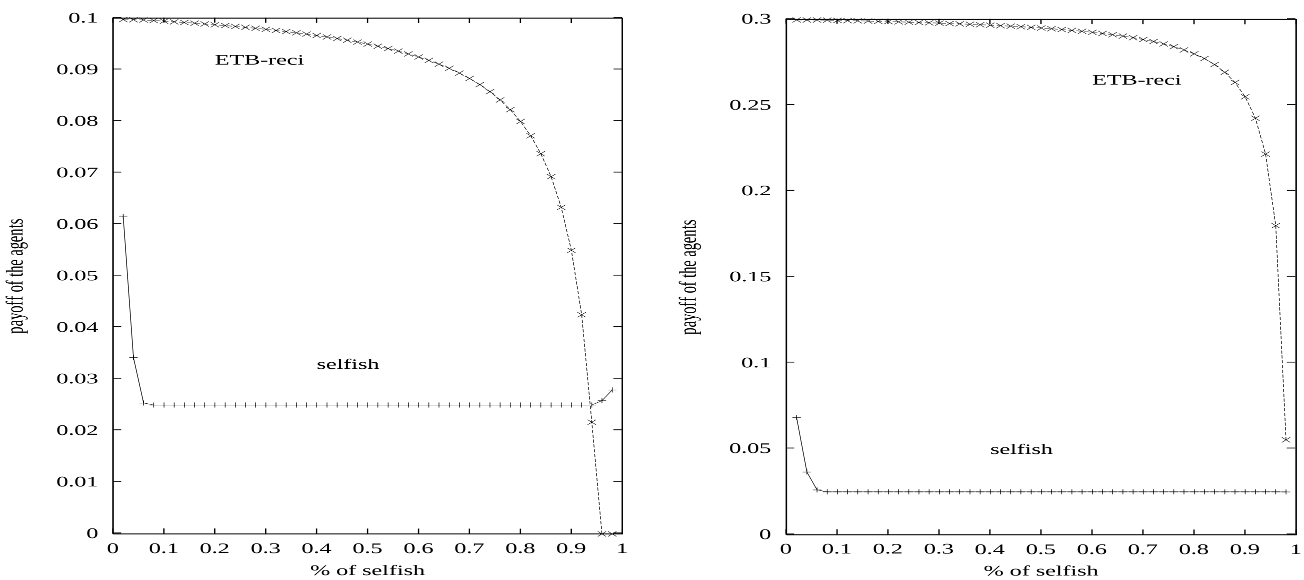

Figure 6: Expected payoff of selfish and ETB reciprocative agents. $T = 5$ **(left) and 15 (right).**

decision functions and interaction histories. We believe that such models will greatly enhance our capacity of designing, evaluating, implementing, and understanding future agent societies.

Such an evolutionary scenario captures the dynamic of agents periodically adopting strategies that have been providing higher payoff in the current environment. Our goal has been to identify the evolutionarily dominant strategy given the starting strategy distribution and the number of tasks to be performed per iteration before agents reconsider changing their strategies. Our analytical model helps us predict both the population dynamics after any given number of strategy adoption decisions and the evolutionarily dominant strategy given the environmental conditions. More importantly, such predictive analysis allows us to construct a decision surface using which a rational agent can choose the most beneficial strategy for the long run given the initial strategy profile in the population and the assigned task load.

The current analytical model to predict evolution of strategy mixes can be augmented in multiple ways. The current model captures synchronous strategy adoption decisions by the entire population made at discrete time intervals. One extension can be to maintain the synchronous strategy adoption but by only a fraction of the population. A more radical change will be to allow asynchronous strategy adoption decisions by the population members. Alternative strategy adoption schemes, for example physical neighborhood based sampling schemes that do not require global performance data, can also be studied.

Acknowledgments: This work has been supported in part by an NSF award IIS-0209208.

7. REFERENCES

[1] A. A. Armstrong and E. H. Durfee. Mixing and memory: Emergent cooperation in an information marketplace. In *Proceedings of the Third International Conference on Multiagent Systems*, pages 34–41, Los Alamitos, CA, 1998. IEEE Computer Society.

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

[3] R. Azoulay-Schwartz and S. Kraus. Stable strategies for sharing information among agents. In *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence*, pages 1128–1134, 2001.

[4] C. Castelfranchi, R. Conte, and M. Paolucci. Normative reputation and the costs of compliance. *Journal of Artificial Societies and Social Simulation*, 1(3), 1998.

[5] A. Cesta and M. Miceli. Help under risky conditions: Robustness of the social attitude and system performance. In *Proceedings of the Second International Conference on Multiagent Systems*, pages 18–25, Menlo Park, CA, 1996. AAAI Press.

[6] T. Day and P. D. Taylor. The evolution of temporal patterns of selfishness, altruism, and group cohesion. *The American Naturalist*, 152(1):102–113, July 1998.

[7] K. Deb and D. Goldberg. A comparative analysis of selection schemes used in genetic algorithms. In G. J. Rawlins, editor, *Foundations of Genetic Algorithms*, pages 69–93, San Mateo, CA, 1991. Morgan Kaufman.

[8] A. Glass and B. Grosz. Socially conscious decision-making. *Autonomous Agents and Multi-Agent Systems*, 6(3):317–339, May 2003.

[9] M. Nowak, R. May, and K. Sigmund. The arithmetics of mutual help. *Scientific American*, 272(6):76–81, June 1995.

[10] S. Sen. Reciprocity: a foundational principle for promoting cooperative behavior among self-interested agents. In *Proceedings of the Second International Conference on Multiagent Systems*, pages 315–321, Menlo Park, CA, 1996. AAAI Press.

[11] S. Sen. Believing others: Pros and cons. *Artificial Intelligence*, 142(2):179–203, 2002.

[12] S. Sen and P. S. Dutta. The evolution and stability of cooperative traits. In *Proceedings of the First Intenational Joint Conference on Autonomous Agents and Multiagent Systems*, pages 1114–1120, New York, NY, 2002. ACM Pres.

[13] K. Tuyls, K. Verbeeck, and T. Lenaerts. A selection-mutation model for q-learning in multi-agent systems. In *Proceedings of the Second Intenational Joint Conference on Autonomous*

Agents and Multiagent Systems, pages 693–700, 2003.
[14] M. P. Wellman. A market-oriented programming environment and its application to distributed multicommodity flow problems. *Journal of Artificial Intelligence Research*, 1(1):1–22, 1993.

APPENDIX

Proof of Proposition 1: In this environment an agent asks every other agent sequentially for help until it receives help. If a philanthropic agent is present in the environment, agents will eventually ask it for help if turned down by other agents. If asked, a philanthropic will always help. So, if a philanthropic agent is present in the environment, other agents either receive help before asking it or they will ask and receive help from it. Hence the probability of receiving help is 1.

Alternatively, using Equation 3, 4 and 5, we can show that, $\sum_{k\in(\mathcal{P}\cup\mathcal{R})} P_1(i, k, j, l) = 1$, provided $|\mathcal{P}| \geq 1$, where $\mathcal{P}$ and $\mathcal{R}$, respectively, are the set of philanthropic agents and reciprocative agents.

Proof of Proposition 2: In Proposition 1 we have shown that in the presence of philanthropic agents all agents are guaranteed to receive help for each of their tasks. Since, the selfish agents never help any other agent, it will not incur any cost for helping others. The philanthropic agents, on the other hand, may incur the cost of helping other agents including the selfish agents. Hence, in a mixed group philanthropic agents cannot outperform, and hence dominate, selfish agents.

Proof of Proposition 3: As shown in Proposition 1, in the presence of philanthropic agents all the agents will receive help. Since, the selfish agents never help any other agent, they will not incur any cost for helping others. The reciprocative agents, however, will incur some cost of helping as they will help the philanthropic agents and other reciprocative agents with positive probability. Initially, i.e., before identifying them as selfish, the reciprocative agents will also help the selfish agents with positive probability. Hence, in a mixed group containing philanthropic, selfish and reciprocative agents, the reciprocative agents cannot outperform, and hence dominate, selfish agents.