Reciprocity: a foundational principle for promoting cooperative behavior among self-interested agents

Sandip Sen
Department of Mathematical & Computer Sciences,
The University of Tulsa
Phone: (918) 631-2985, FAX: (918) 631-3077
e-mail: sandip@kolkata.mcs.utulsa.edu

Abstract

If participating agents in a multiagent system can be assumed to be cooperative in nature, coordination mechanisms can be used that will realize desirable system performance. Such assumptions, however, are untenable in open systems. Agent designers have to design agents and agent environments with the understanding that participating agents will act to serve their self-interests instead of working towards group goals. We investigate the choice of interaction strategies and environmental characteristics that will make the best self-interested actions to be cooperative in nature. We analyze the inadequacy of traditional deterministic reciprocity mechanisms to promote cooperative behavior with a fair distribution of the workload. A probabilistic reciprocity mechanism is introduced and shown to generate stable and cooperative behavior among a group of self-interested agents. The resultant system exhibits close to optimal throughput with a fair distribution of the workload among the participating agents.

Introduction

Researchers involved in the design of intelligent agents that will interact with other agents in an open, distributed system are faced with the challenge of modeling other agents and their behavior (Weiß & Sen 1996). If one can can assume that all agents will be cooperative in nature, efficient mechanisms can be developed to take advantage of mutual cooperation. These will lead to improved global as well as individual performance. But, in an open system, assumptions about cooperative agents or system-wide common goals are hard to justify. More often, we will find different agents have different goals and motivations and no real inclination to help another agent achieve its objectives. Agents, therefore, need to adapt their behaviors depending on the nature or characteristics of the other agents in the environment.

Mechanisms for adaptation that use a lot of information and require complex processing of that information consume significant computational resources (Booker 1988; Watkins 1989). We are interested in developing adaptive mechanisms that are simple and impose little cognitive burden on the agents. Also, whereas the above and other researchers are interested in developing strategies for adapting to the environment of the agent (Kephart, Hogg, & Huberman 1989; Weiß & Sen 1996), we are particularly interested in developing mechanisms for adapting to other agents in a group.

In this paper, we assume agents to be self-motivated in their interactions with other agents, and that the interacting agents are uniquely identifiable. An agent may help others in performing assigned tasks. We plan to develop a criteria for an agent to decide to help or not to help another agent when the latter requests for help. The decision criteria should be such that it allows an agent to perform effectively in the long run. This means that to be effective, an agent must be able to adapt its behavior depending on the behavior of other agents in the environment.

We investigate a simple decision mechanism using the principle of *reciprocity*, which means that agents help others who have helped them in the past or can help them in the future. In this paper, we use a multiagent domain where agents can exchange their tasks. We show that agents can use the principle of reciprocity to effectively adapt to the environment (for our discussion, the nature of the other agents determine the environment).

Reciprocity as an adaptive mechanism

The evolution of cooperative behavior among a group of self-interested agents have received considerable attention among researchers in the social sciences and economics community. Researchers in the social sciences have focused on the nature of altruism and the cause for its evolution and sustenance in groups of animals (Krebs 1970; Schmitz 1993; Trivers 1972). Mathematical biologist and economists have tried to explain the rationality of altruistic behavior in groups of self-interested agents by proposing various fitness models that analyze the success of altruistic individuals and more importantly the evolution of genetic traits supporting altruistic behavior (Dugatkin *et al.* 1994; Nee 1989; Nowak, May, & Sigmund 1995). Our goal in this paper is not to model altruistic behavior in animals; so we do not address the issues raised in the social science literature on this topic. Our purpose is to propose mechanisms by which cooperation can be encouraged and established in groups of self-interested agents. To this end, we have to compare and contrast and build upon the work reported by game theorists and economists on this topic. Space limitations do not permit a thorough review of the literature. Hence, we first identify a common trait in most of this body of work that we have surveyed, identify some underlying problems with the common trait, and then motivate how our proposed approach addressess these problems.

Most of the work by mathematical biologists or economists on the evolution of altruistic behavior deals with the idealized problem called Prisoner's dilemma (Rapoport 1989) or some other repetitive, symmetrical, and identical 'games'. Some objections have already been raised to using such sanitized, abstract games for understanding the evolution of complex phenomena like reciprocal altruism (Boyd 1988). In the following we analyze in some detail one of the often-cited work that share the typical assumptions made by economists and mathematical biologists, and then present our own set of suggestions for relaxing the restrictive assumptions made in that work.

In a seminal piece of work Robert Axelrod has shown how stable cooperative behavior can arise in self-interested agents when they adopt a reciprocative attitude towards each other (Axelrod 1984). The basic assumptions in this work include the following: agents are interested in maximizing individual utilities and are not pre-disposed to help each other; agents in a group repeatedly interact over an extended period of time; all interactions are identical (they are playing the same "game" again and again); agents can individually identify other agents and maintain a history of interactions with other agents; individual agents do not change their behavioral strategy over time; composition of agent groups change infrequently and the changes are minimal (only a few agent leaves and joins a group at a time). Using primarily simulated games, and, to a lesser extent, theoretical analysis, Axelrod convincingly argues for the effectiveness of simple behavioral rules for a variety of agent interactions. Specifically, he shows that a simple, deterministic reciprocal scheme of cooperating with another agent who has cooperated in the previous interaction (this strategy, for obvious reasons, is referred to as the *tit-for-tat* strategy), is quite robust and efficient in maximizing local utility. Whereas such a behavioral strategy can be exploited by strategies designed for that purpose, in general, the tit-for-tat strategy fairs well against a wide variety of other strategies. Two properties of the tit-for-tat strategy deserve special mention:

- if all agents use this strategy, system performance is optimal,
- it is stable against invasion by agents using other strategies (i.e., if an agent using another strategy is introduced into a group of tit-for-tat agents, the former cannot obtain greater utility than that obtained by tit-for-tat agents).

Though Axelrod's work is interesting and convincing, we believe that the assumptions used in his work makes the results inapplicable in a number of domains of practical interest. We now analyze some of this critical assumptions, identifying how they are violated in domains of practical interest, and motivate the need for an alternative framework for reciprocal behavior (we believe the term *reciprocal behavior*, as compared to the term *altruistic behavior*, more appropriately summarizes the motivation and mechanism that we use) that avoids these unrealistic assumptions:

Initial decision: Since tit-for-tat uses the history of one interaction, the first decision is crucial. Axelrod assumes that such agents start of cooperating, which leads to everybody cooperating forever thereafter. If agents start off by not cooperating, then the same tit-for-tat strategy will never produce cooperative action. Either of the above assumptions about initial decisions are equally meaningful for the tit-for-tat strategy.

Symmetrical interactions: Axelrod assumes that every interaction is perfectly symmetrical. This implies that if two agents cooperate in any interaction, both incur the same cost and benefit. In real-life interactions, more often than not in any one interaction one agent incurs the cost and the other incurs the benefit. While individual interacts are asymmetrical, averaging over an ensemble of interactions can put one agent as many times in the position of the benefactor as in the position of the beneficiary. Because of this, an agent has to decide whether to help another agent or not in an interaction by considering past history and future expectations of interactions.

Repetition of identical scenarios: The same situation recurs very infrequently in real-life. More often than not, either the parties involved or the environmental conditions that have an impact on the agent decisions, are at least slightly different. Even if an identical situation recurs one or a few times, it is highly unlikely to be repeated again and again as assumed by the game-playing framework used by Axelrod. As such, in real-life situations, agent decisions will be affected by other factors not addressed in the above-mentioned body of work.

Lack of a measure of work: Since all interactions are assumed to be identical, there is no need to measure the cost of cooperation. Real life scenarios present differing circumstances which need to be compared based on some common metric. For example, consider a scenario where time is the cost metric of cooperation. Suppose that A helped B by picking up a set of photographs that B had dropped off to a local store for developing; this act of cooperation cost A 5 minutes.

Now, A asks B to drive him/her to the nearest airport which will cost B 2 hours. Should B honor such a request? The simple tit-for-tat mechanism will suggest that B cooperates, but that may not be the best choice. Lets take the example a little further. What if A keeps on repeating similar requests before any situation arises where A may be of help to B. Just because A had helped B the last time it was asked to help, should B keep on continuing to help A? The most straightforward application of the tit-for-tat strategy would suggest just that (we can always modify it by saying one cooperative action would be reciprocated by exactly one cooperative action, but that still does not address the question of comparing the cost of cooperation). The point is that there is no mechanism for comparing past favors and future expectations in the tit-for-tat strategy. It was not designed for scenarios in which individual cooperation acts benefits one party while the other incurs a cost.

Hence, the simple reciprocative strategy is not the most appropriate strategy to use in most real-life situations because most of the underlying assumptions that motivate its use are violated in these situations. Our proposal is for agents to use a reciprocity-based interaction scheme that is based on more realistic assumptions. More specifically, we believe that a probabilistic, rather than a deterministic reciprocity scheme is more suitable for real-life problems. Such a scheme should have at least the following properties:

- allow agents to initiate cooperative relationships (this implies that it should be able to handle asymmetrical interactions),
- use a mechanism to compare cooperation costs,
- allow agents to be inclined to help someone with whom it has a favorable balance of help (have received more help than have offered help),
- be able to flexibly adjust inclination to cooperate based on current work-load (e.g., more inclined to cooperate when less busy, etc.).

Probabilistic reciprocity

We assume a multiagent system with N agents. Each agent is assigned to carry out T tasks. The jth task assigned to the ith agent is t_{ij}, and if agent k carried out this task independently of other tasks, the cost incurred is C^k_{ij}. However, if agent k carried out this task together with its own task t_{kl}, the cost incurred for task t_{ij} is C^{kl}_{ij}. Also, the cost incurred by agent k to carry out its own task t_{kl} while carrying out task t_{ij} for agent i is C^{kij}_{kl}. In this paper, we allow an agent to carry out a task for another agent only in conjunction with another of its own tasks.

We now identify the scopes for cooperation. If an agent, k, can carry out the task of another agent, i, with a lower cost than the cost incurred by the agent who has been assigned that task ($C^i_{ij} > C^{kl}_{ij}$), the first agent can cooperate with the second agent by carrying out this task. If agent k decides to help agent i, then it incurs an extra cost of C^{kl}_{ij} but agent i saves a cost of C^i_{ij}. The obvious question is why should one agent incur any extra cost for another agent. If we consider only one such decision, cooperation makes little sense. If, however, we look at a collection of such decisions, then reciprocal cooperation makes perfect sense. Simple reciprocity means that an agent k will help another agent i, if the latter has helped the former in the past. But simple reciprocity by itself does not promote cooperative behavior. This is because, no one is motivated to take the first cooperative action, and hence nobody ever cooperates!

In practice, reciprocity also involves a predictive mechanism. An agent helps another agent, if it expects to receive some benefit from the latter in the future. Developing a domain-independent predictive model is a very difficult problem. In absence of such a general predictive mechanism, we propose a much simpler but equally effective stochastic choice mechanism to circumvent the problem of simple reciprocity. In the following, we propose a probabilistic decision mechanism that satisfies the set of criteria for choosing when to honor a request for help that we described at the end of the previous section. It should be noted that the probability function used here is only a representative function that we have found to be very effective in promoting cooperation among self-interested agents. No claim is hereby made regarding the uniqueness or optimality of the proposed probability mechanism.

We will define S_{ik} and W_{ik} as respectively the savings obtained from and extra cost incurred by agent i from agent k over all of their previous exchanges. Also, let $B_{ik} = S_{ik} - W_{ik}$ be the balance of these exchanges ($B_{ik} = -B_{ki}$). We now present the probability that agent k will carry out task t_{ij} for agent i while it is carrying out its task t_{kl}. This probability is calculated as:

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

where C^k_{avg} is the average cost of tasks performed by agent k (this can be computed on-line or preset), and β and τ are constants. This gives a sigmoidal probability distribution in which the probability of helping increases as the balance increase and is more for less costly tasks. We include the C_{avg} term because while calculating the probability of helping, relative cost should be more important than absolute cost (if the average cost of an agent is 1000, incurring an extra cost of 1000 is less likely than incurring an extra cost of 10). Due to the stochastic nature of decision-making some initial requests for cooperation will be granted whereas others will be denied. This will break the deadlock that prevented simple reciprocity from providing the desired system behavior[1].

[1] Our probabilistic scheme is different from a simple, de-

We present a sample probability distribution in Figure 1. The constants β and τ can be used to make agents more or less inclined to cooperate. The factor β can be used to move the probability curve left (more inclined to cooperate) or right (less inclined to cooperate). At the onset of the experiments B_{ki} is 0 for all i and k. At this point there is a 0.5 probability that an agent will help another agent by incurring an extra cost of $\beta * C^k_{avg}$ (we assume that the average cost incurred is known; an approximate measure is sufficient for our calculations). The factor 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. Similar analyses of the effects of β and τ can be made for any cooperation decision after agents have experienced a number of exchanges. In essence, β and τ can be used to choose a cooperation level (Goldman & Rosenschein 1994) for the agents at the onset of the experiments. The level of cooperation or the inclination to help another agent dynamically changes with problem solving experience.

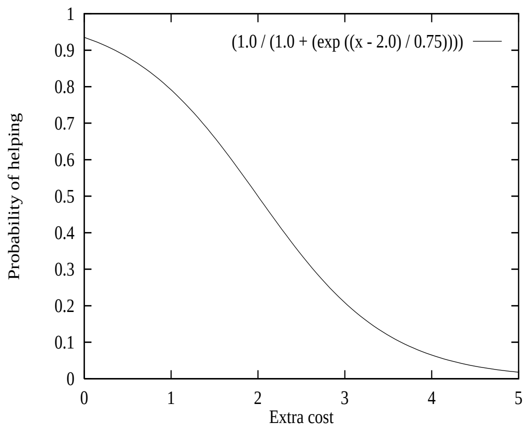

Figure 1: Probability distribution for accepting request for cooperation.

A package delivery problem

In this section, we present a simple package delivery problem which we will use to demonstrate the effectiveness of our proposed mechanism to allow an agent to adapt its environment. We assume N agents, each of which is assigned to deliver T packets. All the packets are located in a centralized depot. The packet destinations are located on one of R different radial fins, and at a distance between 1 and D from the depot. Agents can only move towards or away from the depot following one of the fins; they cannot move directly between fins. On arriving at the depot, an agent is assigned the next packet it is to deliver. At this point, it checks if other agents currently located in the depot are going to deliver along the same radial fin. If so, it asks the other agent to deliver this packet.

The cost of an agent to deliver one of its packets individually is double the distance of the delivery point from the depot. If it carries another package to help another agent, it incurs one unit of extra cost per unit distance traveled when it is carrying its own packet and this extra packet. In addition, if it is overshooting its own destination to help the other agent, an additional cost measured as double the distance between the destination of its packet and the destination of the other agent's packet is incurred. Suppose agent X is carrying one of its deliveries to a location (1,2) (a location (x, y) means a point at a distance y units from the depot on radial fin number x). It is concurrently carrying a packet for agent Y to be delivered at location (1,3) and a packet for agent Z to be delivered at location (1,4). Then the extra cost is 2 units for the first, second and third unit distances traveled, and 1 unit for going from (1,3) to (1,4), and two units to come back from (1,4) to (1,2) for a total of 9 units; 5.5 units are charged to agent Z and 3.5 units are charged to agent Y.

We impose the following limitations on agents helping other agents: 1) An agent will request for help only if the cost incurred by the helping agent is less than the savings obtained by the helped agent. 2) Though an agent can help several agents at the same time, it can carry at most one packet for each of these other agents at the same time.

Experimental results

In this section, we present experimental results on the package delivery problem with agents using the reciprocity mechanism described in the 'Probabilistic Reciprocity' section to decide whether or not to honor a request for cooperation from another agent. We vary the number of agents and the number of packets to be delivered by each agent to show the effects of different environmental conditions. The other parameters for the experiments are as follows: $R = 4$, $D = 3$, $\tau = 0.75$, and $\beta = 0.5$. Each of our experiments are run on 10 different randomly generated data sets, where a data set consist of an ordered assignment of package deliveries to agents. All the agents are assigned the same number of deliveries. The evaluation metric is the average cost incurred by the agents to complete all the deliveries.

We used this domain to also investigate the effects of agent characteristics on overall system performance.

terministic tit-for-tat strategy, e.g., agent k may decide to help agent i even if the later had refused help in the previous time-step. The decision is based only on the balance, not on when requests for help where accepted or denied.

We experimented with the following types of agents:

Philanthropic agents: Agents who will always accept a request for cooperation. Philanthropic agents will produce the best system performance. To aid this process, we impose the restriction that if two philanthropic agent are assigned deliveries on the same fin, the one going further away from the depot takes over the delivery of the agent who is going a shorter distance. In this way, the system incurs minimal extra cost.

Selfish agents: Agents who will request for cooperation but never accept a cooperation request. Selfish agents can benefit in the presence of philanthropic agents by exploiting their benevolence.

Reciprocative agents: Agents that uses the balance of cost and savings to stochastically decide whether to accept a given request for cooperation.

Individual agents: Agents who deliver their assigned packets without looking for help from others. They will also not accept any cooperation requests.

We expect the individual and the philanthropic agents to provide the two extremes of system performance. The individual agents should travel on the average the longest distance to complete their deliveries (because no one is helping them), whereas the philanthropic agents should travel the least. We expect reciprocative agent behaviors to lie in between. The frequency of occurrence of cooperation possibilities should determine which of the two ends of the spectrum is occupied by the reciprocative agents. We want to find out if selfish agents can profit at the expense of reciprocative agents. It would seem that reciprocative agents should perform better because with sufficient interactions they become philanthropic towards each other, a possibility denied to the selfish agents.

For the first set of experiments we chose the number of agents, N, as 100 and varied the number of deliveries per agent from 100 to 500 in increments of 100. Different experiments were performed on homogeneous sets of individual, reciprocative, and philanthropic agents. Results from these set of experiments are presented in Figure 2. As expected, the performance of the individual agents was the worst, and the philanthropic agents were the best. The interesting thing is that the performance of the reciprocative agent is almost identical to that of philanthropic agents. That is, when a reciprocative agent is placed in a group of other reciprocative agents it adapts over time to behave like a philanthropic agent, and this adaptation benefits everybody. This is a significant result because we have been able to show that under proper environmental conditions (frequent interactions with possibilities of cooperation), self-motivated behavior based on reciprocity can produce mutually cooperative behavior that leads to near-optimal system performance. In addition, with more packages to deliver, the savings in distance traversed is more with reciprocative and philanthropic agents over individual agents. The ratio of corresponding points on the two curves should be the same, however, as it is determined by the probability of another agent being able to help one agent with its delivery. For the package delivery problem this probability is largely determined by the number of radial fins, R, the maximum distance traversed from the depot, D, and the number of agents, N.

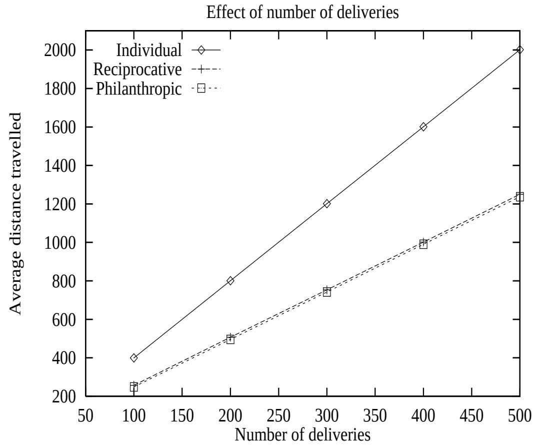

Figure 2: Average distance traversed by each agent to complete all deliveries.

We also performed a similar set of experiments by fixing the number of deliveries per agent at 500 and varying the number of agents from 25 to 50 to 75 to 100. Results from these set of experiments are presented in Figure 3. As above, the performance of the individual agents was the worst, and the philanthropic agents was the best (approximately one-third savings is obtained). The performance of the reciprocative agents was very close to that of the philanthropic agents, and it improved with more agents (with more agents there is more scope of cooperation). Relational agents perform less efficiently than philanthropic agents as occasionally they turn down globally beneficial cooperation requests that will affect local problem solving (involve incurring additional cost for an agent with whom there is a already a large negative balance).

The next set of experiments were designed to find out the effects of selfish agents in a group of reciprocative agents. We expected that selfish agents should be able to obtain some help from reciprocative agents, and hence would perform better than individual agents. But they would not be able to match the performance of reciprocative agents. For these set of experiments, we fixed the number of agents at 100 and the number of deliveries at 500. We varied the percentage of selfish agents in the group. Results are presented in Figure 4, which also contains the results from individual and philanthropic agent groups for comparison purposes. Our intuitions regarding the relative performance of

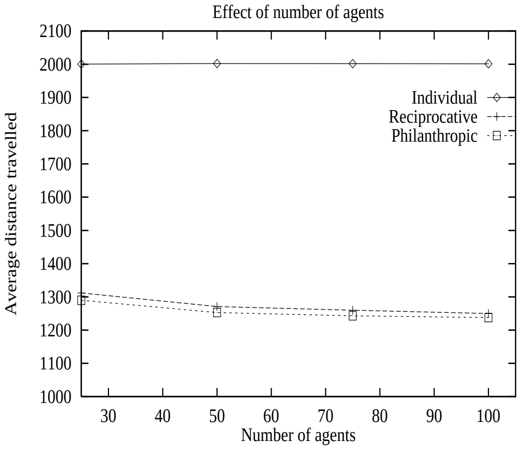

Figure 3: Average distance traversed by each agent to complete all deliveries.

the agents are corroborated by the figure. The average performance of the group, obviously, lies in between the performance of the selfish and reciprocative agents, and moves closer to the performance of the selfish agent as the percentage of the latter is increased. It appears that the selfish agents are able to exploit the reciprocative agents to improve their performance significantly over individual agents. This is because there are many reciprocative agents and they do not share their balance information with other reciprocative agents. If reciprocative agent would broadcast the continuous denial of acceptance request by a selfish agent, the latter would not be able to exploit other reciprocative agents. But this scheme requires more "cooperation" between reciprocative agents, and has not been further studied. Since reciprocative agents incur extra cost for selfish agents without being reciprocated, their performance is noticeably worse than the performance of philanthropic agents. On further analysis of the experimental data we found that the use of reciprocity allows the reciprocative agents to adopt their behavior such that after sufficient number of interactions they learn to reject requests for help from the selfish agents, while at the same time acting "philanthropically" towards other reciprocative agents. The presence of selfish agents, however, can lower the performance of the whole group.

To find out more about the relative performance of selfish and reciprocative agents, we ran a further set of experiments in which we varied the number of deliveries while keeping the number of agents fixed at 100 of which 25 agents were selfish in nature. Results from these set of experiments are presented in Figure 5. A noteworthy result was that with few deliveries to make, selfish agents outperformed reciprocative agents. This can be explained by the fact that the number of reciprocative agents were large enough compared to the

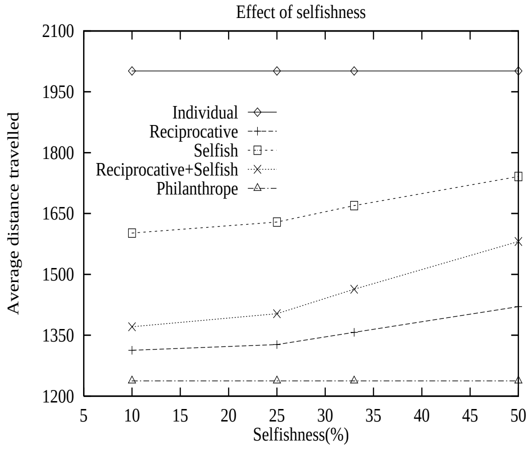

Figure 4: Average distance traversed by each agent to complete all deliveries as the percentage of selfish agent in a group of reciprocative agents is varied. The individual and the philanthropic agent results do not contain selfish agents and are presented for comparison.

number of deliveries, and this allowed selfish agents to exploit reciprocative agents for most of its deliveries. This in turn affected the performance of the reciprocative agents, as they could not recover from the extra cost incurred to help these selfish agents. With sufficient deliveries to make, however, reciprocative agents emerged to be clear winners. This lends further credence to our claim that in the long run it is beneficial for an agent to be reciprocative rather than selfish.

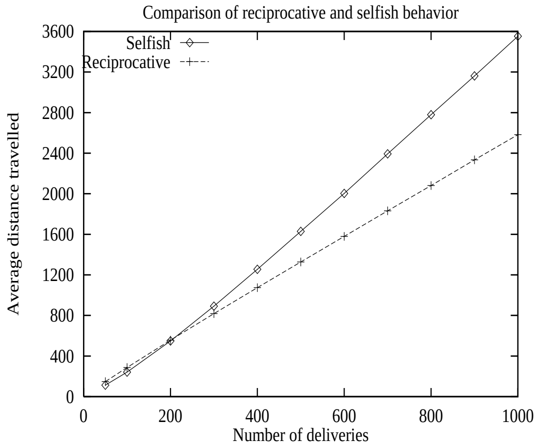

Figure 5: Average distance traversed by each agent to complete all deliveries with different number of deliveries.

In the last set of experiments, we investigated the relative performance of tit-for-tat and reciprocative agents both in homogeneous groups and when pitted against selfish agents. For each experiment there were 100 agents in a group, each of which were to deliver 500 packages. The percentage of selfish agents in the group were varied from 0 to 50. The major findings from these set of experiments were as follows:

- The average cost incurred by tit-for-tat agents was slightly less (the difference is less than 5% of the cost incurred) than that incurred by reciprocative agents. For homogeneous groups, this is because in some cases reciprocative agents will refuse to help because of the outstanding balance with the agent requesting for help. Tit-for-tat agents will continue to help in this situations. If we modify the tit-for-tat strategy to reciprocate one cooperative action with exactly one cooperative action, then their performance will deteriorate. For heterogeneous groups, reciprocative agents may help selfish agents more than tit-for-tat agents. For example, a tit-for-tat agent will stop helping a selfish agent the first time it is refused help by the latter. This may happen before the corresponding selfish agent has requested help from that tit-for-tat agent. A reciprocative agent in place of the tit-for-tat agent will still help the selfish agent according to its probability calculation (the balance is still 0 and from the reciprocative agents' point of view it is as if they have not interacted at all; this suggests a possible improvement of the reciprocative strategy: each denial of request will be used to decrement its balance with the other agent).
- Though tit-for-tat is a stable strategy given the criterion for stability that we have used (i.e., selfish agents perform worse than tit-for-tat agents), it may not necessarily be attractive to all agents. This is because the variance of the work performed by different agents in the group is high. For homogeneous groups, the variance of the cost incurred by tit-for-tat agents is much higher than the corresponding measure for reciprocative agents (see Figure 6). This means that though a group of tit-for-tat agents perform well on the average, some people work more while the others reap the benefit. In real life, we do not expect such a group to be stable! A group of reciprocative agents, on the other hand, provide a more equitable distribution of workload, even if agents incur slightly more cost on the average. Ironically, if the percentage of selfish agents increase in the group, the variance of work of the tit-for-tat agents decrease as they help fewer agents. At the same time the variance of work of the reciprocative agents increases as different agents have different history of interactions with more selfish agents.

Analysis

We now provide a coarse analysis of why reciprocative agents outperform selfish agents. The only work savings obtained by selfish agents come from exploiting reciprocative agents. Sooner or later a selfish agent will have realized all such benefits, and no further interactions will bring it any savings. After such a point is reached, selfish agents will incur the same cost as individual agents as they have to deliver all their packages themselves. The reciprocative agents, on the other hand, can benefit from prolonged interaction with other reciprocative agents. Since cooperative actions are reciprocated, they can continue to benefit off each other.

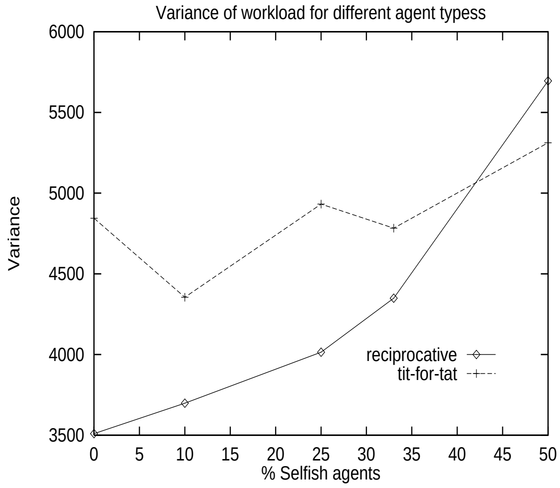

Figure 6: Variance of workload of tit-for-tat and reciprocative agents with different percentage of selfish agents in the group.

Let us analyze the amount of savings a selfish agent can reap from a reciprocative agent. From Equation 1 this can be calculated as

$$\gamma = \int_0^\infty x \,.\, \frac{1}{1 + \exp^{\frac{x - \beta * C_{avg}}{\tau}}} \, dx,$$

where C_{avg} is the average cost incurred in delivering a packet by the reciprocative agent. The expected total savings obtained by a selfish agent is then $N(1-f)\gamma$ where N is the total number of agents and f is the fraction of selfish agents in the group. If the probability of interaction of any two agents at the depot is p, and we assume that on the average half the time one reciprocative agent will help another, then the total savings obtained by the reciprocative agents for D deliveries is $\mathcal{S} = \frac{pD(v-c)}{2}$, where c is the average cost incurred in helping someone and v is the average cost of delivering a package on its own (which is the same as the savings obtained when another agent delivers this packet). Therefore, when $\mathcal{S} > N(1-f)\gamma$, reciprocative agents are expected to perform better than selfish agents. This happens when the number of deliveries are large and the savings obtained by the helped agent is large compared to the cost incurred by the helping agent.

We now briefly highlight some of the other assumptions that contribute to the success of the reciprocative agents:

- The major assumption is that cooperation is always beneficial for the group. In practical situations, if agent A owes agent B a favor, agent B may delegate one of its tasks to agent A even though it can itself do it more efficiently than agent A, i.e., the savings obtained by agent B is less than the cost incurred by agent A.
- We assume all agents have the same capability and evaluation metric. The latter in particular is a critical assumption. It means if agent A though it incurred a cost x while helping agent B, the latter concurs. Though there is nothing in our model from preventing the evaluation metrics differ from agent to agent, we believe that large discrepancies in evaluation metric will prevent sustained cooperation.

Conclusions

In this paper, we have shown that self-motivated behavior can evolve cooperation among a group of autonomous agents. Under appropriate environmental conditions, such a group of agents can also achieve near-optimal global performance. This can be achieved by using reciprocity as an aid to adaptation to other agents. This allows agents to realize scopes for cooperation while avoiding wasting efforts on helping unresponsive agents. This is a significant result because in an open, distributed environment, an autonomous agent is likely to face a multitude of agents with different design philosophies and attitudes. Assuming benevolent or cooperative agents is impractical in these situations. Our analysis and experiments show that agents can use reciprocal behavior to adapt to the environment, and improve individual performance. Since reciprocating behavior produces better performance in the long run over selfish or exploitative behavior, it is to the best interest of all agents to be reciprocative. Our results hold for domains where cooperation always leads to aggregate gains for the group.

We have presented a coarse analysis explaining when the reciprocative agents will outperform selfish agents. We are currently working on a more detailed analysis on this issue and we plan to present theoretical predictions and experimental verifications from this analysis.

We plan to relax the requirements of all cooperation being beneficial ffor the group. Currently, an agent receives help from the first person (from an ordered list) that agrees to help. We plan to study the performance of the mechanism when the agent considers all the offers for help and chooses to take help from the agent with which its got the most negative balance. We also plan to investigate more complex and realistic domains such as distributed monitoring, distributed information gathering, etc. to further evaluate the strengths and limitations of our proposed mechanism.

Acknowledgments : This work has been supported, in part, by an NSF grant IRI-9410180.

References

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

Booker, L. B. 1988. Classifier systems that learn internal world models. *Machine Learning* 3:161–192.

Boyd, R. 1988. Is the repeated prisoner's dilemma a good model of reciprocal altruism? *Ethol. Sociobiology* 9:211–222.

Dugatkin, L.; Wilson, D.; III, L. F.; and Wilkens, R. 1994. Altruism tit for tat and 'outlaw' genes. *Evolutionary Ecology* 8:431–437.

Goldman, C., and Rosenschein, J. S. 1994. Emergent coordination through the use of cooperative state-changing rules. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, 408–413.

Kephart, J. O.; Hogg, T.; and Huberman, B. A. 1989. Dynamics of computational ecosystems: Implications for DAI. In Huhns, M. N., and Gasser, L., eds., *Distributed Artificial Intelligence*, volume 2 of *Research Notes in Artificial Intelligence.* Pitman.

Krebs, D. 1970. Altruism – an examination of the concept and a review of the literature. *Psychological Bulletin* 73(4):258–302.

Nee, S. 1989. Does hamilton's rule describe the evolution of reciprocal altruism? *Journal of Theoretical Biology* 141:81–91.

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

Rapoport, A. 1989. Prisoner's dilemma. In Eatwell, J.; Milgate, M.; and Newman, P., eds., *The New Palgrave: Game Theory.* London: Macmillan. 199–204.

Schmitz, D. 1993. Reasons for altruism. *Social Philosophy & Policy* 10(1):52–68.

Trivers, R. 1972. The evolution of reciprocal altruism. *Quarterly Review of Biology* 46:35–57.

Watkins, C. 1989. *Learning from Delayed Rewards.* Ph.D. Dissertation, King's College, Cambridge University.

Weiß, G., and Sen, S., eds. 1996. *Adaptation and Learning in Multi–Agent Systems.* Lecture Notes in Artificial Intelligence. Berlin: Springer Verlag.