Limiting Deception in Groups of Social Agents

Anish Biswas, Sandip Sen & Sandip Debnath
600 South College Avenue
Department of Mathematical & Computer Sciences
University of Tulsa
Tulsa, OK 74104-3189, USA
e-mail:{abiswas,sandip,debnath}@euler.mcs.utulsa.edu

Abbreviated title: Limiting deception

Abstract

In open environments there is no central control over agent behaviors. On the contrary, agents in such systems can be assumed to be primarily driven by self interests. Under the assumption that agents remain in the system for significant time periods, or that the agent composition changes only slowly, we have previously presented a prescriptive strategy for promoting and sustaining cooperation in the group. This strategy was shown to improve both individual and group performance in the long run. Our prescribed strategy has been an adaptive, probabilistic, reciprocity-based policy for deciding which other individual to cooperate with. In this paper we investigate two mechanisms to limit exploitation of the reciprocative strategy by deceptive agents : 1) a penalty factor for declining requests for help, and 2) a cutoff limit on outstanding balance of help. We evaluate the relative effectiveness of these mechanisms for augmenting robustness of agent behaviors without adversely affecting performance of homogeneous groups of reciprocative agents.

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 [Bra97, CAC94, CAC99, HS97]. The ABSs of the future will be situated in a social context, playing a variety of roles in different relationships and problem solving situations.

Research in societal aspects of agent behaviors, unfortunately, has been relatively scarce. Whereas economic models may provide a basis for structuring agent interactions [Wel93], other approaches inspired by non-monetary relationships [Axe84] may provide more effective social relationships in certain situations. We have been interested in agent strategies for interactions with other agents that can promote cooperation in groups.

We have developed and analyzed probabilistic reciprocity schemes as strategies to be used by self-interested agents to decide on whether or not to help other agents [SS95, Sen96, SB98]. The goal of this work has been to identify procedures and environments under which self-interested agents may find it beneficial to help others. We claim that if the group composition changes only slowly, and there is sustained interaction between the agents reciprocative agents would be able to outperform exploitative agents. Probabilistic reciprocity strategies are considerably more sophisticated than deterministic reciprocity schemes like tit-for-tat [Axe84, CM96] and avoid major problems associated with the latter schemes [Sen96].

Our experiments under a variety of environmental conditions, group composition, work estimate difference, etc. have shown that under prolonged interaction the probabilistic reciprocity strategy produces close to optimal individual and group performance. Additionally, this strategy is stable against selfish intruders, i.e., in the long run, selfish agents perform worse than reciprocative agents in a mixed group.

Our past experience showed that though the reciprocative agents outperformed the selfish agents in the long run, they would be outclassed if agent lifetimes were restricted to only a few interactions. This losing out in the short run can be attributed to many factors including:

- Our previous design of reciprocative agents used only a balance of help to decide whether or not to honor a request for help. This balance was unaffected if the other agent refused to help. Refusal to reciprocate, however, is a clear indicator of selfish behavior and can be used as an early detection mechanism.

- Our assumption of cooperation possibilities allows for differential exploitation. In a given interaction, the savings obtained by the helped agent is more than the cost incurred by the helping agent. As a result, a smart exploitative agent can help a little and then ask for a lot more help. Such differential exploitation is more difficult to detect and resist than the purely selfish agents described above.

- The deception or the selfish nature of an agent takes advantage of the initial helpful advances of the reciprocative agent. When the cumulative help given to one agent increases above a threshold a reciprocative agent stops helping that agent. Since, agent do not share their opinion about one agent with others, a given selfish agent can exploit each reciprocative agent in turn.

In this paper, we evaluate two auxiliary schemes to improve the robustness of reciprocative agents against exploitative agents:

- a penalty mechanism by which the probability of helping another agent is reduced every time that agent denies a request for help,

- a maximum limit of balance of work with another agent.

Both of these schemes reduces or restricts the amount of exploitation possible by another agent. They can also eliminate some cooperation possibilities between like-minded agents.

The latter is an undesirable, albeit mandatory, side-effect. In this paper, we investigate if the benefits of these auxiliary schemes do outweigh their adverses side-effects. Once again, the tradeoff involved is the sacrificing of homogeneous group performance to the improvement of robustness in the face of deception and exploitation.

The rest of the paper is organized as follows: Section 2 presents a well-known deterministic reciprocity scheme and analyzes its shortcomings; Section 3 describes our probabilistic reciprocity mechanisms that overcomes the above-mentioned shortcomings and is more suited for real-life applications; Section 4 presents the experimental results with probabilistic reciprocity with and without penalties and exploitation bounds) summarizes the major findings; Section 5 presents the main conclusions of the paper and outlines ongoing and future work.

2 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 [Bad93, dVZ94, GB66, Kre70, Sch93, Tri72]. 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 [HMS98].

Most of the work by mathematical biologists or economists on the evolution of altruistic behavior deals with the idealized problem called Prisoner's dilemma [Rap89] 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 [Boy88]. 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 [Axe84]. Specifically, he shows that a simple, deterministic reciprocal scheme of co-operating 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.

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. In real-life situations, a particular help-giving interaction between two agents often means one agent helps and incurs a cost while the other receives help and obtains a savings in cost or effort. Such interactions are necessarily asymmetrical in nature in contrast to the symmetrical formulation of games like prisoner's dilemma. Another key restrictive feature of Axelrod's experiment with the iterated prisoner's dilemma game is that identical scenarios are repeated. This is not likely in real life as every interaction is different from others. The assumption of repetition of identical scenarios enable Axelrod to work with strategies that do not compare different interactions. In real life, history of interaction will have to capture not only the outcomes, but also the context in which a certain outcome was produced. Also, there has to be a means to compare two different scenarios or two help-giving actions of different magnitude. This requires the use of some measure of work or cost involved in help-giving. Such a metric will allow systematic evaluation of different scenarios under different interaction histories.

Based on these observations, we believe that a simple tit-for-tat like deterministic strategy is not adequate for more realistic agent domains. We now identify the desirable features of a

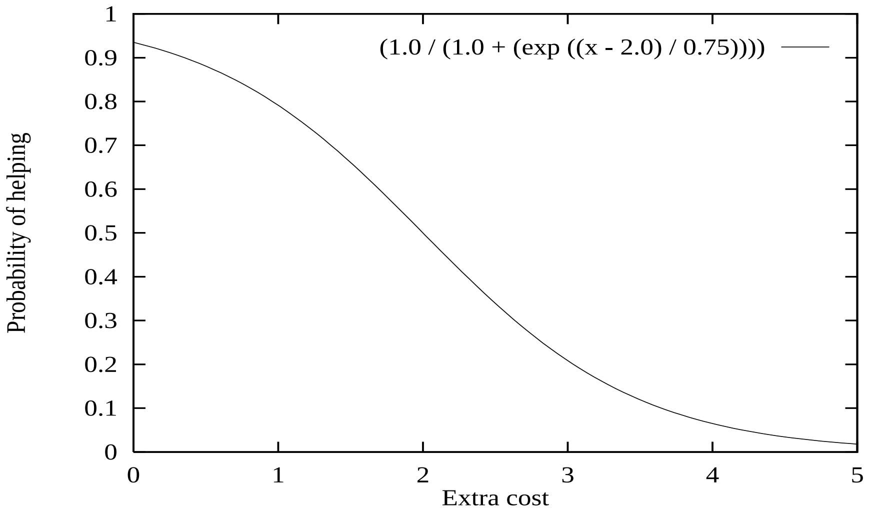

Figure 1: Probability function for accepting request for cooperation.

behavioral strategy that will be suitable for open environments: a risk attitude that allows the agent to initiate help-giving to a new agent but quickly shun it if requests for help are rejected repeatedly; ability to compare cooperation costs across different scenarios; ability to adjust help-giving beahvior based on local work-load.

3 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 costs it C_{ij}. If agent k carried out this task together with its own task t_{kl}, the cost incurred for task t_{ij} is C_{ij}^{kl}.

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 ($C_{ij} > C_{ij}^{kl}$), 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_{ij}^{kl} but agent i saves a cost of C_{ij}^{i}.

We now 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. 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 $B_{ik} = W_{ik} - S_{ik}$ be the balance of these exchanges. 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:

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

where C_{avg}^{k} is the average cost of tasks performed by agent k, and β and τ are constants. This gives a sigmoidal probability function in which the probability of helping increases as the balance increases and is more for less costly tasks.

We present a sample probability function in Figure 1. The constant 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 $\beta * C_{avg}^{k}$. The constant τ can be used to control the steepness of the curve. In essence, β and τ can be used to choose a cooperation level [GR94] for the agents. The level of cooperation or the inclination to help another agent dynamically changes with problem solving experience.

4 Experimental results

In the simple package delivery problem that we have used for experimentally evaluating strategies, we assume there are 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 F 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 any other agents are currently located in the depot. If so, it can ask those agents 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.

In this section, we present experimental results on the package delivery problem with agents using the reciprocity mechanism described in Section 3 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 and strategic 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 have experimented with the following types of agents in the package delivery domain:

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.

In homogeneous groups, we expect individual agents to incur the most cost as they shun each other and the philanthropic agents to perform the best as they help each other at every opportunity. From our previous work, we know that the average performance of homogeneous groups of reciprocative agents is almost identical to that of the average performance of a group of philanthropic agents [Sen96]. The performance of reciprocative agents do deteriorate from a homogeneous group to groups containing other exploitative, selfish agents. The positive result is that even in such heterogeneous groups, the reciprocative agents clearly outperforms the selfish agents in the long run. As such, adaptive reciprocative agents prove to be effective strategies for both promoting and sustaining cooperation in agent groups.

4.1 Resisting exploitation of reciprocative agents

Even though reciprocative agents outperform selfish agents in the long term in a heterogeneous group, their own performance suffers because they waste some of their efforts in helping selfish agents. Mechanisms that quickly identify selfish agents, i.e., agents from whom no benefit can be obtained, can help alleviate this problem. The key problem is that without willing to initiate a relationship by incurring some up-front costs, even cooperative relationships cannot be established. We have investigated two mechanisms by which reciprocative agents can limit potential losses to exploiters:

Penalty: In this scheme, each refusal to help will incur a penalty that is used to reduce the balance of help with that agent.

Balance limit or bounds: In this scheme, the balance of help with another agent is not allowed to cross a prespecified limit.

Each of these methods can be used by a single agent to prevent exploitation by another agent.

4.1.1 Bounding exploitation

If agent A offers help to agent B, then the mutual balance between A and B, from A's point of view, increases [1]. So, at any point of time, a positive outstanding balance between one agent and another denote unreciprocated help. One can control exploitation, by imposing a limit on this outstanding balance. If honoring a request will require doing an amount of work that will put the balance with the requesting agent over the limit, then that request for help would be denied. We have experimentally evaluated the effect of bounds on the robustness of reciprocative strategy. As before we used 100 agents in total with the percentage of selfish agents varied between 10 and 50. Results from these simulations are presented in Figure 2. The first to note is that in the heterogeneous group with reciprocative agents with bound(RB) the selfish agents are performing worse than in with original reciprocative agents. We also notice that, because of less exploitation, the reciprocative agents with bounds (RB) perform better than reciprocative agents without such bounds in similar group compositions.

If we consider the curves for two different bounds (see Figure 2), we find that with the more restrictive (lower) bound, selfish agents are more severely affected. Also, such severe bounds does not significantly affect the performance of the reciprocative agents themselves.

[1] For ease of exposition we will use this interpretation of balance. In the actual formulation however, balance of exchanges decreases with help and becomes negative with no reciprocation (refer to Section 3).

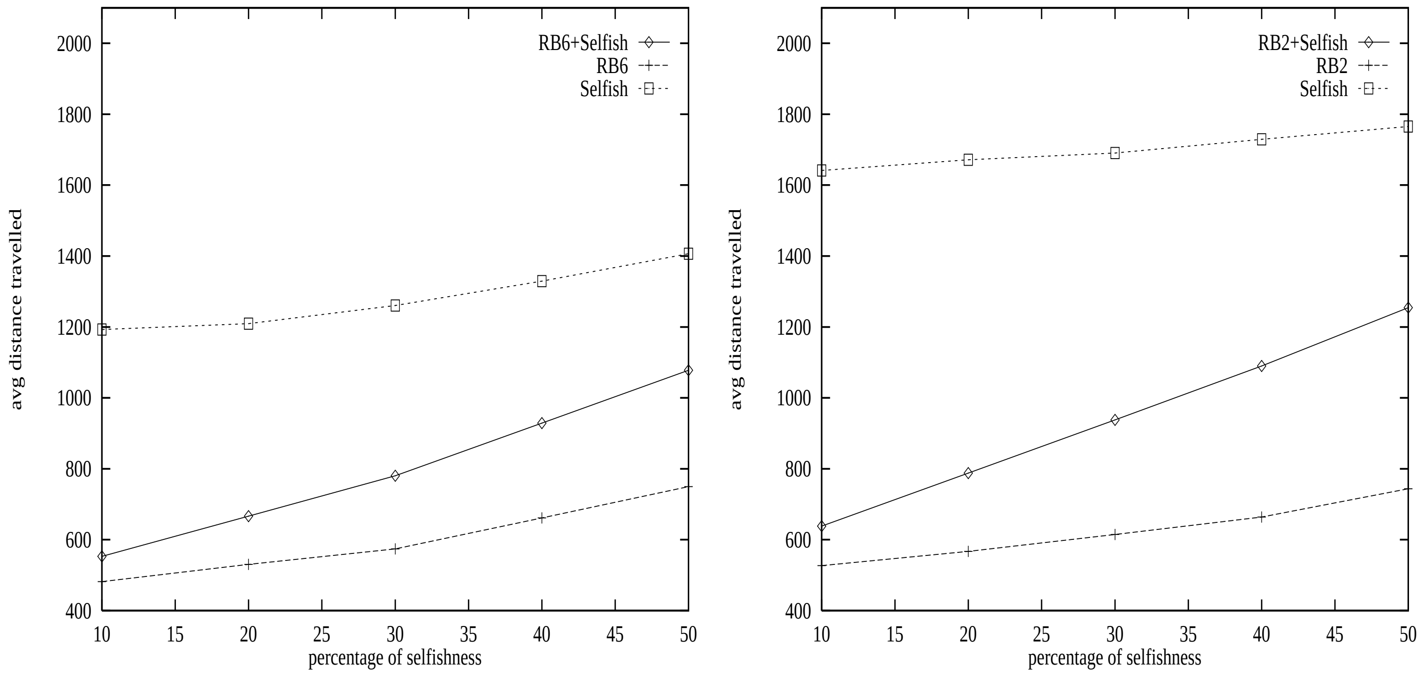

Figure 2: Left: Performance of reciprocative(R) strategy with bound(B) value of 6 in the presence of selfish agents. Right: Performance of reciprocative(R) strategy with bound(B) value of 2 in the presence of selfish agents.

A bound of 2 (which is of $\beta * C^k_{avg}$ for the parameter values chosen in the experiment) is found to be a very effective deterrent to exploitation without affecting cooperation possibilities. This is what we had ideally hoped for.

4.1.2 Taking exception to denials

We can also improve performance in th presence of selfish agents by adding a penalty term to the mutual balance when agent B declines a request for help from agent A. In that case agent A, increases their mutual balance. By this method, the outstanding balance with selfish agents will become large and the probability of helping such an agent will rapidly decrease. But, this penalty may also affect the performance in a homogeneous group as sometime agents have to deny requests for help for other non-exploitative reasons (e.g., if it was already overburdened with work). We again experimented with 100 agents and varied the number of selfish agents in the group from 10 to 50. In figure 3, we observe that a low penalty value,e.g., 0.25, is enough to all but eliminate the exploitative effects of the selfish agents who are left to do all of their own tasks (their performance become very similar to individual agents). As an added incentive, the reciprocative agents with these small penalty values do outperform the original reciprocative agents under similar group compositions. This is because the penalty factor severely limits exploitation by selfish agents. However, higher penalty values do adversely affect the performance of the reciprocative(R) agents(for penalty value 1 in figure 3). Reciprocative agents with penalty value of 1, for example, perform worse than original reciprocative (R) agents.

We have also compared the performances of the improved strategies, incorporating either a penalty factor or exploitation bound, in homogeneous groups with the original reciprocity mechanism. Results are presented in figure 4. We experimented with 100 agents and varied the number of tasks from $T = 500$ to $T = 2500$. We observe that the deterioration of performances of the modified strategies are minor and negligible for this range of workload. As a result we can claim that either approach to limiting exploitation is more robust in the face of exploitation without sacrificing the capability of promoting cooperative behavior in homogeneous groups.

5 Conclusions and Future Work

We have evaluated two augmentations of our adaptive mixed strategy to promote resistance to exploitation. The first scheme adds a penalty factor to the mutual balance between the agents to decrease the probability of helping an agent who has declined a request. We have also investigated the effect of limiting the maximum balance of help with another agent. With a reasonable choice of parameters, both these mechanisms significantly improve resistance to exploitation without noticeably decreasing cooperation potential between similar agents.

As mentioned in the 4 section, the penalty and balance limit based methods are designed for individual reciprocative agents and helps an agent limit exploitation by another agent. Two further enhancements can be examined:

Group labeling: Agents pool together the overall balance of other agents. An agent with steadily increasing or significant outstanding balance with other agents can be identified by everyone as being a selfish or exploitative agent. Such agents can then be shunned by reciprocative agents. This scheme, however, can be easily manipulated by deceitful agents who can manipulate the common balances to their advantage. A variation on this method would be for agents to exchange to broadcast their balances. A reciprocative agent can then form their estimates about another agent by summing the balances for that agent received from other agents. Though this scheme is susceptible to the same manipulation as the previous one, a variant of this can be more useful. While forming estimates of an agent, the reciprocative agent can consider the balances of only those agents that it trusts. A first approximation of trust can be obtained by considering only agents with which the reciprocative agent has a positive balance, i.e., it has received more help than it has given. Since a reciprocative agent does not stand to gain by lying to hurt another agent, this scheme should enable reciprocative agents to jointly identify selfish agents within a few interactions. Similarly, number of refusals to help can also be pooled together to further limit the exploitation possibilities by selfish agents.

Individual summary balance: A reciprocative agent can use an adaptive mixed strategy even if it cannot identify other agents uniquely. It can use its total outstanding balance with all agents to decide on whether or not to help another agent. Such a policy is necessarily non-discriminatory, i.e., it treats all agents identically irrespective of who

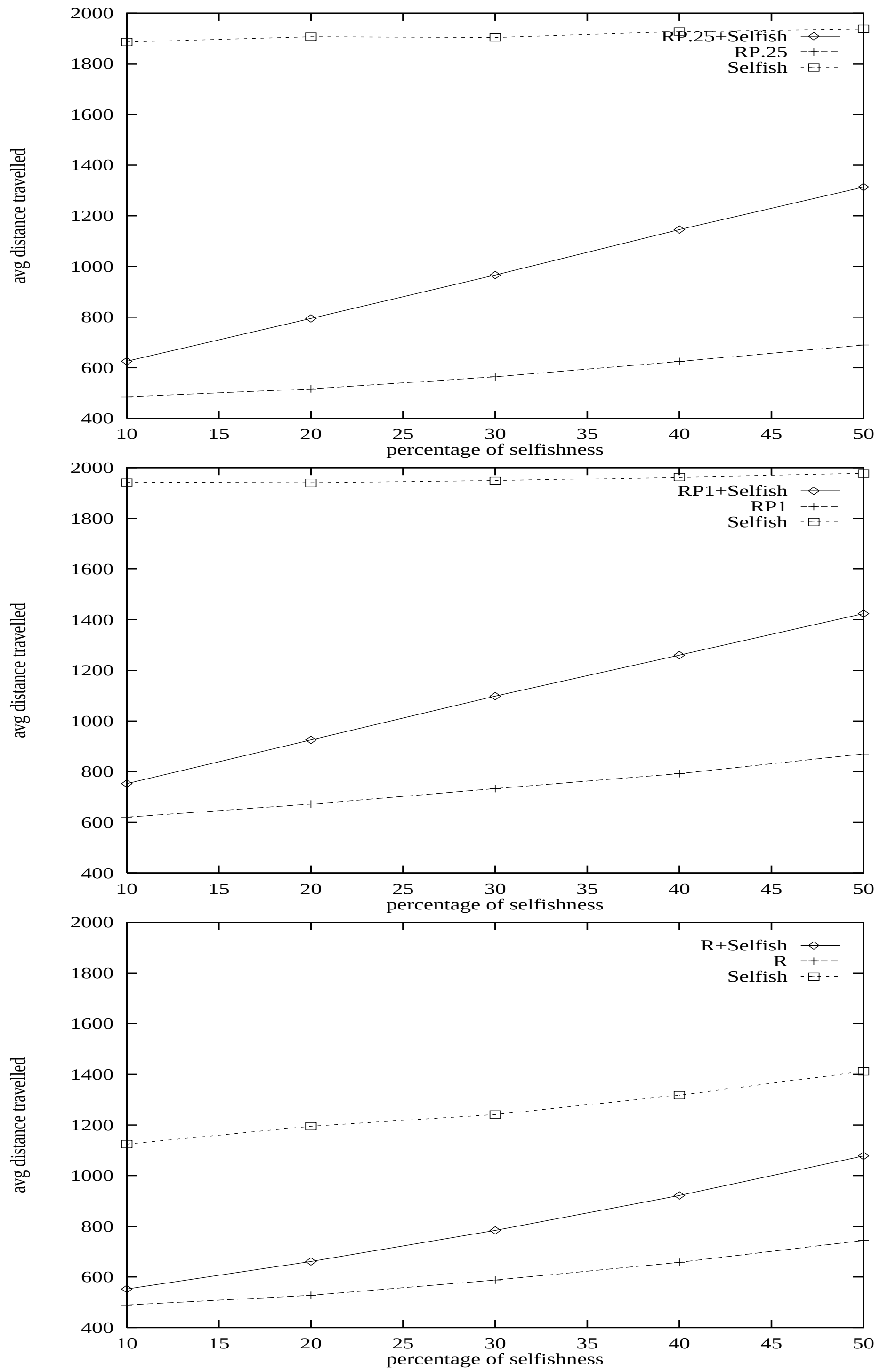

Figure 3: Top: Performance of reciprocative(R) strategy with penalty(P) value(0.25) in the presence of selfish agents.Middle: Performance of reciprocative(R) strategy with penalty(P) value(0.5) in the presence of selfish agents. Bottom: Performance of reciprocative(R) strategy without penalty in the presence of selfish agents.

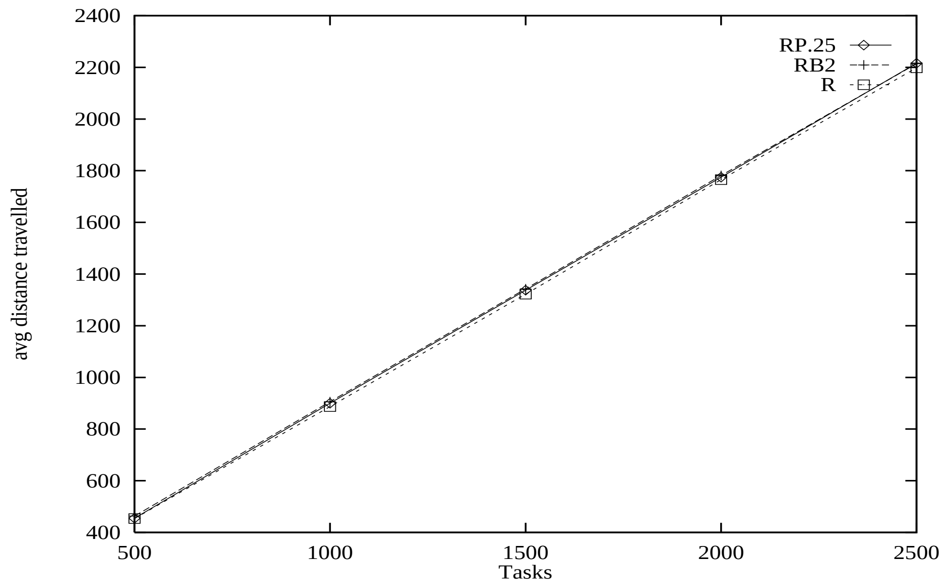

Figure 4: Performance in homogeneous group with the augmentations

is asking for help. Though this strategy is exploitable by selfish agents, the level of exploitation is fixed and can be set. Thus this strategy is likely to fare better than the non-adaptive mixed strategy used in this paper. An interesting aspect of this strategy is that an agent using it will appear benevolent only when others appear so to it. In a sense this is a probabilistic "tit-for-tat" to the society. Ironically, selfish agents would be able to exploit such agents more when they constitute a smaller fraction of the population.

A restrictive assumption in the current line of work is that agents are assumed to have fixed behaviors, i.e., each agent uses one of several pre-specified behaviors over their entire lifetime. A more realistic scenario would be for an agent to have the freedom of choosing from one of several of these behaviors and to change its behavior as and when it deems appropriate[2]. An agent may be prompted to adopt a behavior if agents using that behavior is seen to be performing better than others. Such a behavior adoption method leads to an

[2]It should be clarified that the agents in this study are adaptive nature. What we mean by fixed behavior is that for an agent the adaptation process is fixed.

evolutionary process with a dynamically changing composition of agent group behaviors. It is not clear *a priori* if behaviors that produce greater returns if all agents were forced to use the same behavior over their lifetime would emerge as the dominant behavior in a group where agents change behaviors regularly based on limited-term performance.

We plan to experiment with a related framework that shares the same motivation of identifying behaviors that emerge based on the performance of different agents in mixed groups. Our experimental framework will consist of a population of agents with the initial population containing representatives of different behaviors in specified proportions. Each of these agents will then be assigned some tasks. The cost of executing a task can be reduced or eliminated if help is obtained from another agent. After all agents have finished processing their assigned tasks, their relative performance will be tallied. The next generation of the population will be created by a performance-proportionate scheme which produces more agents with behaviors that produced above-average performance and eliminates some of the agents that produced below-average performance. Over time, the entire population will have the tendency to become homogeneous in terms of adopting the "winning" behavior. We plan to investigate the emergence of such "evolutionarily stable strategy" (ESS) [DT98].

Acknowledgements: This work has been supported, in part, by an NSF CAREER Award: IIS-9702672.

References

[Axe84] Robert Axelrod. *The Evolution of Cooperation.* Basic Books, 1984.

[Bad93] N.K. Badhwar. Altruism versus self-interest: sometimes a false dichotomy. *Social Philosophy & Policy*, 10(1):90–117, 1993.

[Boy88] R. Boyd. Is the repeated prisoner's dilemma a good model of reciprocal altruism? *Ethol. Sociobiology*, 9:211–222, 1988.

[Bra97] Jeffrey M. Bradshaw. *Software Agents.* AAAI Press/The MIT Press, Menlo Park, CA, 1997.

[CAC94] Communications of the ACM, july 1994, volume 37, number 7, 1994. Special Issue on Intelligent Agents.

[CAC99] Communications of the ACM, march 1999, volume 42, number 3, 1999. Special Issue on Multiagent Systems on the Net and Agents in E-commerce.

[CM96] Amedeo Cesta and Maria 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.

[DT98] Troy Day and Peter D. Taylor. The evolution of temporal patterns of selfishness, altruism, and group cohesion. *The American Naturalist*, 152(1):102–113, July 1998.

[dVZ94] H. de Vos and E. Zeggelink. The emergence of reciprocal altruism and group-living: an object-oriented simulation model of human social evolution. *Social Science Information*, 33(3):493–517, 1994.

[GB66] R.E. Goranson and L. Berkowitz. Reciprocity and responsibility reactions to prior help. *Journal of Personality and Soial Psychology*, 3(2):227–232, 1966.

[GR94] Claudia Goldman and Jeffrey S. Rosenschein. Emergent coordination through the use of cooperative state-changing rules. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, pages 408–413, 1994.

[HMS98] Elizabeth Hoffman, Kevin A. McCabe, and Vernon L. Smith. Behavioral foundations of reciprocity: experimental economics and evolutionary psychology. *Economic Inquiry*, 36(3):335–352, July 1998.

[HS97] Michael N. Huhns and Munindar P. Singh. *Readings in Agents*. Morgan Kaufmann, San Francisco, CA, 1997.

[Kre70] D.L. Krebs. Altruism – an examination of the concept and a review of the literature. *Psychological Bulletin*, 73(4):258–302, 1970.

[Rap89] A. Rapoport. Prisoner's dilemma. In J. Eatwell, M. Milgate, and P. Newman, editors, *The New Palgrave: Game Theory*, pages 199–204. Macmillan, London, 1989.

[SB98] Sandip Sen and Anish Biswas. Effects of misconception on reciprocative agents. In *Proceedings of the Second International Conference on Autonomous Agents*, pages 430–435, New York: NY, 1998. ACM Press.

[Sch93] D. Schmitz. Reasons for altruism. *Social Philosophy & Policy*, 10(1):52–68, 1993.

[Sen96] Sandip 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.

[SS95] Mahendra Sekaran and Sandip Sen. To help or not to help. In *Seventeenth Annual Conference of the Cognitive Science Society*, pages 736–741, Hillsdale, NJ, 1995. Lawrence Erlbaum Associates, Publishers.

[Tri72] R. Trivers. The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46:35–57, 1972.

[Wel93] Michael 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.