Emergence of Stable Coalitions via Negotiation of Task Exchanges

Partha Sarathi Dutta and Sandip Sen
Mathematical & Computer Sciences Department
University of Tulsa
Tulsa, OK 74104
partha@euler.mcs.utulsa.edu, sandip@kolkata.mcs.utulsa.edu

Abstract

Autonomous agents interacting in an open world can be considered to be primarily driven by self interests. In this paper, we evaluate the hypotheses that self-interested agents with complementary expertise can learn to recognize cooperation possibilities and develop stable, mutually beneficial coalitions that is resistant to exploitation by malevolent agents. Previous work in this area has prescribed a strategy of reciprocal behavior for promoting and sustaining cooperation among self-interested agents. That work had considered only the task completion time as the cost metric. To represent more realistic domains, we expand the cost metric by using both the time of delivery and quality of work. In contrast to the previous work, we use heterogeneous agents with varying expertise for different job types. This necessitates the incorporation of the novel aspect of learning about other's capabilities within the reciprocity framework.

Introduction

Agent-based systems are an important aspect of real world applications like electronic commerce, recommender systems and personal assistants. Agents deployed in these applications often interact in an open environment with other agents or humans (Bradshaw 1997; CACM July 1994 issue 1994; CACM March 1999 issue 1999; Huhns & Singh 1997). The interactions involve cooperation, collaboration or competition for resources to achieve the specified goals of these agents. With increase in the complexity of agent interactions, the behavioral characteristics of agents acting in a group should be studied in detail and suitable interaction strategies developed that optimize system performance.

We have been interested in agent strategies for interactions with other agents that can promote cooperation in groups of self-interested agents. Our approach is different from other researchers who have designed effective social laws that can be imposed on agents (Shoham & Tennenholtz 1992). 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 saved by the helped agent. As agent system designers we can define rules of interactions to increase the likelihood of cooperation possibilities. We prescribe reciprocative behavior as a stable sustainable strategy that creates cooperative groups in the society. This behavior not only sustains group formation in a homogeneous society of self-interested agents, but also helps to ward off exploitative tendencies of selfish agents in the society (Biswas, Sen, & Debnath 2000). This strategy of reciprocal behavior becomes more stable if the helping agent takes into account the opinion of all other agents before extending any favor (Sen, Biswas, & Debnath 2000).

A restriction of the previous work on reciprocity was the incorporation of only a single cost metric, time, used by the agents. In real-life scenarios multiple objectives like time, quality, dependability, etc. will be involved when an agent evaluates the benefit of interacting with another agent. As a first step to handling such a scenario, we expand on the set of cost metrics by including both time and quality in an agent's evaluation scheme. The measures of *time* and *quality* of a work need clarification. The *time* attribute refers to the absolute time units required for completing a particular task, and the *quality* attribute is a measure of the effectiveness of executing a task. These values will depend on the expertise level of agents on different task various task types.

A second restriction in the previous work was the explicit assumption that all agents had the same capabilities i.e. agents were homogeneous in task expertise. We assumed this to focus on the help-giving behaviors of the agents. Having established a basic competence of probabilistic reciprocity based agents in recognizing mutually cooperative relationships and effectively neutralizing exploitative behavior (Sen, Biswas, & Debnath 2000), we now turn our attention to the interesting and practical aspect of variance in agent expertise. We assume that different agents have different skill sets which make them more effective in accomplishing some tasks compared to others. We require agents to learn the capabilities of themselves and others through repeated task performance and interaction.

The goal of this work is to evaluate whether self-interested agents can learn to recognize agents with complementary expertise and develop a self-sustaining relationship through exchange of help. This can be described as an augmentation of the basic probabilistic reciprocity model with the concept of *selecting a partner*. Additionally, such help exchange inclinations must be resistant to exploitation by malevolent agents who do not reciprocate help-giving actions. Our hypothesis is that when combined with an appropriate learning scheme, probabilistic reciprocity based strategies will enable the development of stable, mutually beneficial coalitions of self-interested agents with complementary skill sets.

Reciprocity

Our probability-based reciprocal behavior is designed to enable an agent make decisions about agreeing to or refusing a request for help from another agent. The probability that agent k, having task l, will help agent i to do task j is given by

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

where C_{avg}^{k} is the average cost of tasks performed by agent k; B_{ki} is the net balance that agent k has with i, due to the previous helps given to and received from agent i; OP_i is the balance that agent i has with other agents excluding agent k; C_{ij}^{k} is the cost of agent k to do the task j of agent i; β and τ are the only two constant parameters used in the function, where β is used to set the cost an agent is ready to incur to help an unknown agent with the hope of initiating a long-term cooperative relationship and τ is used to set the shape of the probability curve. This is a sigmoidal probability function where the probability of helping increases as the balance increases 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.

In this paper we aim at extending the basic probabilistic reciprocity framework. We evaluate the importance of "quality of performance" as another cost metric and the applicability of learning other agents' capabilities in developing stable, cooperative coalitions among self-interested agents.

Problem domain

We evaluate our hypothesis using simulations in a job completion problem domain. In this domain each of N agents are assigned T jobs. There are K job types and each agent has expertise in exactly one of these K job types. An agent who is an "expert" in a particular job type can do the job in less time and with higher quality than other job types. We draw the time and quality of performance from a normal distribution with a set mean and a standard deviation. We have two different values of the mean - "high" and "low". For a task type in which an agent is expert, the time required to complete the task is computed from the distribution using the "low" mean value (the agent completes the task in which it is an expert, in less time). We draw the quality of performance of an expert using the "high" mean value (experts complete with high quality). For performance measure of a non-expert, however, we use the "high" and "low" mean values for computing the time and quality respectively. The standard deviation chosen is the same for both experts and non-experts. The jobs can be finished at any one of F different machines. Each agent is assigned the same number of jobs at the start of the simulation. At this point the agents ask for help from one another. When help is given, the agent who is helped updates its estimate of the helper agent's expertise, i.e., the time and quality of performance of the helper agent. With more interactions, therefore, agents develop better models of each other. This biases their help asking behavior - for a given task type, an agent is more likely to ask help from those agents who are expected to produce higher quality results in less time. The simulation ends when every agent has finished all of their assigned jobs.

Partner selection and coalition formation

We have incorporated simple reinforcement learning schemes to update agent performance estimates. Agents also have estimates of their own abilities to do the different job types. Estimates are of two types: *time estimate*, which reflects the possible time of completion of the job, and *quality estimate*, which reflects the possible performance level of an agent to do that job. Agents also keep estimates of every other agents' abilities.

The time and quality estimates are used by the agents to compute the cost of a particular task delivery. In the current formulation, the cost of completing a task is directly proportional to the time and inversely proportional to the quality of performance. It is intuitive that the cost incurred for a task delivery increases with more time required for the delivery and decreases with increasing quality of performance. For a given task type, an expert takes less time to produce higher quality work than a non-expert. Hence, the design of the cost function in the above manner ensures that the cost of completing a task of a certain type is less for an expert in that type and compared to a non-expert. When asking for help, agents compute the cost $C1$, incurred by itself to do that task. The estimated cost $C2$ that the prospective helping agent incurs for that work is also computed. Help is obtained only if $C2 < C1$. This condition corresponds to a "cooperation possibility".

Initially, agents have neutral estimates about their own abilities and that of other agents. To obtain accurate estimates about their own abilities, agents must themselves perform jobs of different types. When an agent performs a task, it requires a certain time and achieves a certain quality of performance. These values are used by the agents to measure their performance. When an agent helps another, the helped agent updates its estimate of the helper using the time and quality of performance values that the helper agent requires to complete the particular type of job. The simple reinforcement that we use to update the time and quality estimates is given below:

$$t_{ij}^{n+1} \leftarrow (1-\alpha)t_{ij}^{n} + t_{ij},$$

where t_{ij}^{n} is the time taken by agent i to do task j on the n^{th} time instance.

We use a similar update policy for the quality of performance (q_{ij}). More interactions between the same pair of agents provide each with the opportunity of developing more accurate models of one another. This information is used to select an agent for asking help. As a consequence of developing the reinforced model of others' abilities, an agent i who specializes in job type T_1 when given a job of type T_2, will decide in favor of asking help from an agent j specializing in job type T_2, rather than agents who specialize in other job types. The knowledge that agent j is capable of delivering the job with lesser time and with higher quality comes from the reinforced models that the asking agent has developed over time. As a consequence, agents with complementary expertise interact with each other more often than agents with similar expertise. Thus, given sufficient interaction opportunities and cooperation possibilities, agent coalitions develop where self-interested agents have complementary skill sets.

Agent types

The behavior of reciprocative agents is determined by the level of available information to the agents about other agents. We mentioned that agent A can receive help from agent B only if the cost incurred by A to do that task is more than incurred by B for helping A. However, in deciding whether to help or not to help, the reciprocative agents can ask *some* other agents for their *opinion* about the agent asking help. The agents from whom a help-giving agent x, asks for opinion are only those with whom x has a favorable balance of help. This is a reasonable policy, to believe in those who have *earned trust* by their interactions. The opinion received by x from another agent y is the cumulative balance that y has with the agent seeking help. An agent who has a "good" record of help giving behavior is more likely to be helped. Reciprocative agents who decide to help or not based on such collaborative opinion are called *earned trust based reciprocative agents*.

Selfish agents by definition do not extend help under any circumstances. The selfish agents that we use in our simulations are characterized by the additional property of *lying* when they are asked for their opinion. They lie about the balance they hold at any time with the agent who is currently waiting to receive help.If agent x is asking agent y whether or not to help agent z, then y reports a negated balance with z if y is a liar.This means that liars "bad-mouth" helpful agents.

Experimental results

To examine our claim that the ability to estimate other agents' expertise enhance the performance of agents and boost coalition formation among agents with complementary expertise, we present the following experimental results. In our experimental setups we have used 3 different task types, i.e., each agent having one of 3 possible expertise levels. The "high" and "low" mean values used to define the distributions from which we draw time and quality of performance of experts and non-experts are fixed at 10.0 and 1.0 respectively, the standard deviation being 0.05. The difference among the mean values, together with the low standard deviation, ensure that the cost of task completion is substantially lower for an expert than a non-expert. Making the mean values closer and increasing the standard deviation would, in effect, make the distinction in agent expertise imprecise. Thus, the population heterogeneity that we want to introduce and study the nature of help giving among agents with distinct expertise levels, would blend into an approximately homogeneous set. Besides, making the mean values closer and increasing the standard deviation would require more intelligent strategies for selecting partners and agents who can be requested for opinion about a help-seeking agent. In our current work we are focusing on the emergence of stable cooperative partnerships assuming the agent population is segregated into ordinal types based on expertise. The issues of handling close performance means and larger standard deviation are addressed in a current work of Sen et. al. (Sen & Sajja 2002).

We have considered two task cost metrics, time of completion and quality of performance. Our first experimental setup focuses on examining the relative performances of the reciprocative and selfish agents on those two cost metrics. We record the average time taken and the average quality of performance of the reciprocative and selfish agents for completing all tasks. We ran experiments taking a total of 100 agents with equal number of reciprocatives and selfish agents. The tasks per agent was set at 100, 200, 400, 800 and 1000. The average time to complete and the average quality of performance to complete all tasks for each agent type were recorded. Figures 1 and 2 show the results.

Figure 1 shows the relative performance of individual lying selfish and earned-trust based reciprocative agents in terms of the average time spent in performing all tasks. We see from the graphs that selfish agents outperform reciprocatives (take lesser time to complete tasks) until the per agent task reaches a value of 300. As the task value is increased, reciprocatives prove to dominate the performance of selfish agents. As noted earlier, the increase in the number of tasks allow the reciprocatives to have more interactions with other agents. This, in turn, helps in their learning process of identifying potential helpful partners, and in particular, those with non-selfish behavior and having complementary expertise and detect the exploiters. This, when coupled with the strategy of earned trust that the reciprocatives employ, i.e., rely on the opinion of only those agents who have been helpful in the past, leads to dominating performances in most cases.

Figure 2 shows the quality of performance of "individual lying" selfish and "earned trust based" reciprocatives. The graphs follow a pattern similar to that in figure 1. Reciprocatives, initially fail to extract the advantage by learning other agents' models since the number of tasks is not large enough for learning to produce good models and using them to form fruitful partnerships. With increasing tasks, agents interact with agents more often which help them better identify others' expertise and learn more robust models of other agents. Hence, with increasing tasks reciprocatives can get their tasks completed at a higher quality by obtaining help from partners of complementary expertise in emerging coalitions.

In the results presented so far we have attempted to verify our claim that agents do learn to identify partners with complementary expertise by examining and explaining the con-

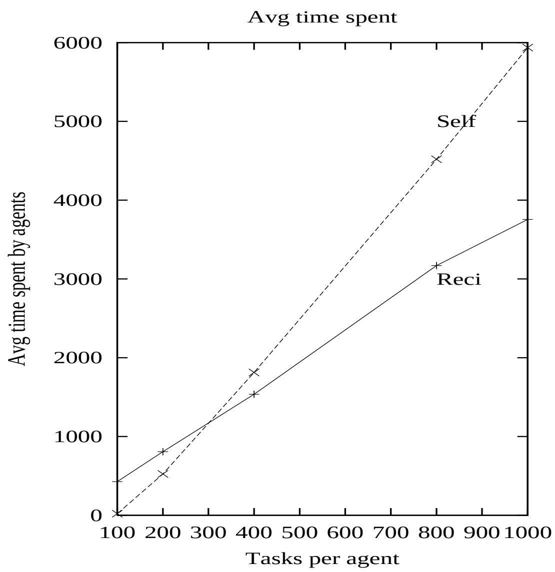

Figure 1: Average time spent by agents.

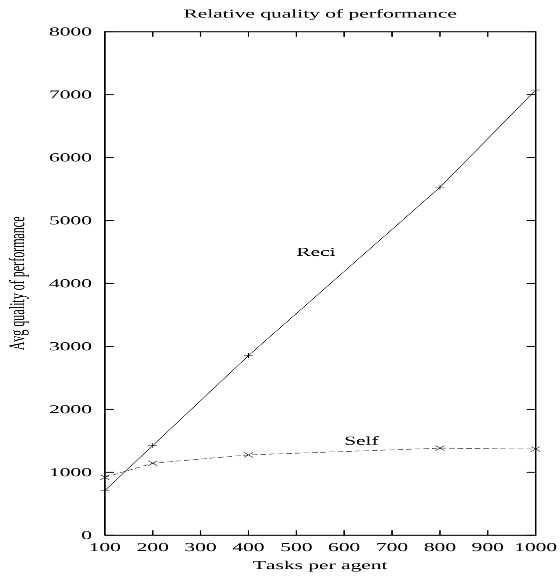

Figure 2: Average quality of performance of agents.

sequences or external effects of the strategies. In our next experimental setup, we show the results that fully establishes the fact. We present, in this experimental setup, information regarding the group/coalition formation among agents. Stability of a coalition is determined by the frequency of interaction and the net savings earned by such coalitions.

Table 1 shows the number of helps (interaction frequency) and total savings generated by the different groups or coalitions of agents of different expertise in a population of size 30 of earned trust based reciprocatives with 2000 tasks per agent. The interaction frequency of a "group"/"coalition" (like $(1, 1)$) is defined as the total number of interactions between all agents with expertise in task type 1. Since there are 3 task types and we are considering coalitions formed of pairs of interacting agents (of similar or complementary expertise), there can be a total of 6 groups that are enumerated in the table. The "savings" of such a group is the total savings earned over the interactions that define the group size. From table 1, we find that both the interaction frequencies and savings earned by the groups of agents having complementary expertise are larger than the corresponding values of groups of similar agents. This corroborates our hypothesis that when agents are endowed with the ability to learn about others' expertise, they develop mutually beneficial coalitions with agents of complementary expertise. The benefit is particularly obtained in groups of heterogeneous agents.

We assume that a task requires one expertise level to be completed. Also, an agent with the required expertise can complete an entire task on its own. We hypothesize that, mutually beneficial cooperative coalitions would still develop among agents with complementary expertise even after relaxing the above constraints. If a task requires k expertise levels ($k > 1$), we expect coalitions to develop where k agents, of different expertise, cooperate to complete a task.

Table 1: Group interaction information of earned trust based reciprocative

Group	Number of helps	Saving
(1,1)	4470	2987.03
(2,2)	4409	3023.38
(3,3)	4584	3560.69
(1,2)	12946	130703.34
(1,3)	12744	128684.23
(2,3)	12976	135528.07

In the previous sets of experiments we have shown that reciprocatives perform better on both the dimensions, time and quality, that define task cost. Savings earned by an agent gives in a nutshell how well it performed in accomplishing its assigned tasks. So, in our last set of experiments we studied the variations in average savings earned by both selfish and reciprocative agents after each agent has completed all of its jobs. We used a mix of individual lying selfish and earned-trust based reciprocative agents. Figures 3 and 4 show the results obtained by varying the percentage of selfish agents in the population under two different values of "tasks per agent", 20 and 40 respectively. The percentage of selfish agents in the population was varied from 50 to 80% in steps of 10%. In both the figures we find reciprocative agents outperform selfish agents, because the average savings earned by the reciprocative agents are more than those of the selfish agents for all selfish percentages in the population. In both cases the savings earned by reciprocatives decrease with increase in the percentage of selfish population. For the same percentage of selfish population, however, the savings earned by the reciprocatives are higher when the per-agent tasks are higher. From these observations, we can conclude that, reciprocative agents learn to identify poten-

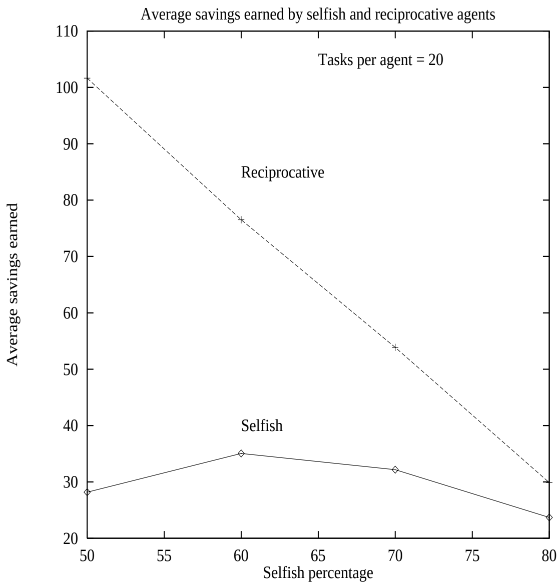

Figure 3: Average savings of agents with 20 tasks per agent.

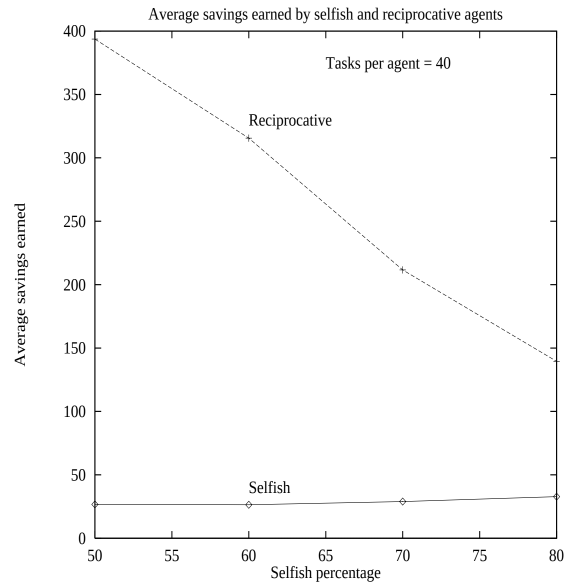

Figure 4: Average savings of agents with 40 tasks per agent.

tial helpful agents and their expertise which enables them to enter into mutually beneficial coalitions. Savings earned, reflect the effectiveness and stability of agent coalitions. Lower values of savings earned by the selfish agents indicate that the reciprocative agents have effectively shunned the exploitative selfish agents. The performance of reciprocative agents, however, decreases with increase in the number of the selfish agents. This is due to the fewer reciprocative agents with which stable mutually beneficial coalitions can be formed. With increase in the task level, however, the savings of reciprocative agents for the same number of selfishes in the population are more because a higher number of per agent tasks allows agents to receive help from few reciprocative agents for a much longer time.

Related work

Using agent coalitions is an interesting and much explored approach to solve complex, distributed tasks. Assigning groups of agents to do a task or multiple tasks has the advantage of complementary individual agent expertise being used to complete different parts of the global problem.

However, work in the area of coalition formation in agent societies has focused on cooperative agents (Shehory & Kraus 1998). A related work on coalitions in agent societies takes self-interested agents into account (Lerman & Shehory 2000). But it does not consider the possible heterogeneity in performance of a task between different agents. Our work is different from these because it takes into consideration self-interested agents that have different performance levels for different task types. We also have a learning parameter in our agents which is used to identify other's capabilities. This, in turn, favors those with complementary expertise when asking for help. Mutual understanding of capabilities evolve cooperative groups in the society of self-interested agents. Learning of cooperative behavior has also been addressed in (Denzinger & Kordt 2000). However, there the learning uses an off-line learning module which generates situation-action pairs using GA. To determine the fitness of the evolved situation-action pairs during off-line learning an agent-model is required. Our work uses only on-line learning and agents do not have to store a priori models of other agents. However, in an attempt to learn behaviors of all other agents, our approach can become time consuming in situations where number of agents increase exponentially. Multiple inter-dependencies among contracts are solved by partially ordering the negotiation process in a multi-agent supply chain domain in (Zhang & Lesser 2002). In contrast, we focus on the generation of stable cooperative partnerships among agents with different capabilities instead of developing smart negotiation strategies. In (Brooks, Durfee, & Armstrong 2000), a similar problem of agents locating other agents with complementary abilities at "congregations" has been discussed. There, strategies are studied that maximizes the chance of agents, that can maintain mutually beneficial relationships, interacting among each other, increasing, thereby, local agent utility and enhancing global performance of the system.

Conclusions and future directions

Adoption of reciprocal behavior in a society of self-interested agents generates incentives that promote cooperative behavior. In this paper we hypothesized that in an agent society where agents differ in their expertise, incorporation of a learning component will help generate cooperative groups of agents with complementary expertise. Our simulations confirmed this hypotheses. These groups are mutually beneficial to the agents because they save more time and achieve higher quality of performance by giving the tasks to other agents who have the required expertise and accomplish their jobs with better quality. A simple cost function is used that reflects the effects of time of completion and quality of performance for a task on the task cost. Such cooperative group formation is stable since the exploitative agents are effectively shunned.

One assumption in this work has been that agents have fixed behaviors. A more realistic scenario would be for an agent to have the freedom of changing its behavior when it deems appropriate. Such behavior adoption leads to an evolutionary process with a dynamically changing composition of agent group behaviors. We plan to investigate whether such evolving agent behavior might create a society of agents with the same behavior, which could be the optimal behavior to adopt.

References

Biswas, A.; Sen, S.; and Debnath, S. 2000. Limiting deception in groups of social agents. *Applied Artificial Intelligence, Special Issue on Deception, Fraud, and Trust in Agent Societies* 14(8):785–797.

Bradshaw, J. M. 1997. *Software Agents*. Menlo Park, CA: AAAI Press/The MIT Press.

Brooks, C. H.; Durfee, E. H.; and Armstrong, A. 2000. An introduction to congregation in multiagent systems. In *Proceedings of the Fourth International Conference on Multi-Agent Systems*, 79–86.

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

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

Denzinger, J., and Kordt, M. 2000. Evolutionary online learning of cooperative behavior with situation-action-pairs. In *Proceedings of Fourth International Conference on MultiAgent Systems, ICMAS 2000*, 103–110. Los Alamitos, CA: IEEE Computer Society.

Huhns, M. N., and Singh, M. P. 1997. *Readings in Agents*. San Francisco, CA: Morgan Kaufmann.

Lerman, K., and Shehory, O. 2000. Coalition formation for large-scale electronic markets. In *Proceedings of the Fourth International Conference on Multi-Agent Systems*, 167–174.

Sen, S., and Sajja, N. 2002. Robustness of reputation-based trust. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS 2002) (to appear)*.

Sen, S.; Biswas, A.; and Debnath, S. 2000. Believing others: Pros and cons. In *Proceedings of the Fourth International Conference on Multiagent Systems*, 279–285. Los Alamitos, CA: IEEE Computer Society.

Shehory, O., and Kraus, S. 1998. Methods for task allocation via agent coalition formation. *Artificial Intelligence Journal* 101(1-2):165–200.

Shoham, Y., and Tennenholtz, M. 1992. On the synthesis of useful social laws for artificial agent societies (preliminary report). In *Proceedings of the National Conference on Artificial Intelligence*, 276–281.

Zhang, X., and Lesser, V. 2002. Multi-linked negotiation in multi-agent system. In *To appear in the proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2002) (to appear)*.