Helping Based on Future Expectations

Sabyasachi Saha
Department of Mathematics & Computer Science
The University of Tulsa
sahasa@ens.utulsa.edu

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

Partha Sarathi Dutta*
Department of Mathematics & Computer Science
The University of Tulsa
partha@ens.utulsa.edu

ABSTRACT

Autonomous agents interacting in an open world can be considered to be primarily driven by self interests. Previous work in this area has prescribed a strategy of reciprocal behavior, based on past interactions, for promoting and sustaining cooperation among such self-interested agents. Here we present a new mechanism where agents base their decisions both on historical data as well as on future interaction expectations. A decision mechanism is presented that compares current helping cost with expected future savings from interaction with the agent requesting help. We experiment with heterogeneous agents that have varying expertise for different job types. We evaluate the effect of both change of agent expertise and distribution of task types on subsequent agent relationships. The reciprocity mechanism based on future expectations is found to be robust and flexible in adjusting to the environmental dynamics.

Categories and Subject Descriptors

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

General Terms

Experimentation

Keywords

Cooperation, reciprocity, agents, adaptation

1. INTRODUCTION

Agent-based systems provide a useful framework for developing 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 [2, 7]. 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 improve or optimize system performance.

We have been interested in interaction strategies of 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 [12]. 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. The development of cooperative relationships leading to exchanges of help can improve both agent and system-level performances (the latter through minimizing resource consumption, increasing throughput, etc.). We have designed suitable interaction strategies that take advantage of cooperation possibilities in the environment. Additionally, we identified that reciprocal cooperation among learning agents with different expertise for different task types generated stable, mutually beneficial groups among agents with complementary expertise [5]. In our previous work on agents forming reciprocal groups based on identifying others' expertise, the decision mechanism of agents was solely dependent on past interaction history [10]. An agent, while deciding whether or not to honor a help request from another agent, was more likely to help if it had a net profit from help exchanges with that agent in the past. This is an effective decision mechanism if the probabilities of arrival of different task types are time-invariant and the agents are assumed to have fixed expertise levels over different task types for the duration of the experiments. The agents do not look into the future, and hence, a change in the pattern of task arrival or an agent acquiring new expertise, that might happen in the future, was not accounted for in the decision procedure. This inflexibility can result in performance degradation of agents in more dynamic domains.

In this paper, we modify the decision mechanism used by reciprocative agents to incorporate an outlook for the future. We endow our agents with an expectation on the future behavior of the task arrival pattern and the ability to adapt to expertise changes of other agents. In our framework, a given agent is expert in only one task type. The cost of performing a task depends on the expertise of the agent in that task

*Currently at University of Southampton, UK

AAMAS'03, July 14–18, 2003, Melbourne, Australia.

type and has two component metrics: the time to complete the task and the quality of performance. An agent bases the decision of helping another agent on the expected utility of its action by using expectations of future interactions with the help-seeking agent. We posit that such expected utility based decision mechanism is effective in generating significant cost savings under time varying task arrival distributions. We evaluate the effectiveness of this new decision mechanism for enabling the agents to identify others' expertise and forming stable, mutually beneficial groups with agents having complementary capabilities. We also investigate the effects of agents dynamically adopting new behaviors. It is observed that the expected utility based strategy endows an agent to identify such changes and reform its relationships with agent groups.

2. EXPECTED UTILITY BASED HELPING DECISIONS

We assume a set of $\mathcal{A}$ agents executing tasks from a set Υ. Let $\mathcal{H}$ denote the interaction histories of the agents. $\mathcal{H}$ is an ordered list of tuples where each tuple is of the form $\langle i, j, x, t, c_i, c_j, help\rangle$ where the components are respectively the agent requesting help for a task, the agent being asked for help, the task type, the time instance, the cost of performing the task to the requesting agent, the cost of performing the task to the agent being asked for help, and whether or not j helped i. Let $\mathcal{H}_{i,j} \subseteq \mathcal{H}$ be the part of the history that contains interactions between agents i and j only. Let H denote the space of all possible histories. Our goal is to derive a decision procedure $\mathcal{F} : \mathcal{A} \times \mathcal{A} \times \Upsilon \times H \rightarrow Yes/No$ that maps a request from an agent to another agent to a boolean decision based on the task type involved and the interaction history of these two agents.

We introduce an expected utility based decision mechanism used by the reciprocative agents to decide whether or not to honor a request for help from another agent. When requested for help, an agent, using this decision mechanism, estimates the utility of agreeing to the request by evaluating its chance of obtaining help from the asking agent in future. An agent, being self-interested, has the objective of earning more savings by receiving help than cost incurred by helping others in the long run. When an agent using this strategy decides whether or not to provide help, it uses a statistical summary of its past interactions with the requesting agent as a metric for evaluating its expected interaction pattern with the latter in future. Using this information, it evaluates the difference between the expected benefit and the expected cost it might incur for that agent by helping it in the future. In the following, we present the expected utility based decision mechanism that agent m uses to evaluate a help request by another agent o for helping with task type τ. The expected utility of agent m for interacting with agent o at time $\mathcal{T}$ and future time steps, $E_{\mathcal{T}}(m, o, \tau)$, is defined as:

$$E_{\mathcal{T}}(m,o,\tau) = \sum_{t=\mathcal{T}}^{\infty} \gamma^{t-\mathcal{T}} [\sum_{x\in\Upsilon} (D_m^t(x) \mathrm{Pr}_{m,o}^t(x) cost_m(x)) - \sum_{x\in\Upsilon} (D_m^t(x) \mathrm{Pr}_{o,m}^t(x) cost_m(x))] - cost_m(\tau), \quad (1)$$

where $cost_i(x)$ is the expected cost that i incurs doing a task of type x, $cost_m(\tau)$ is the cost to be incurred by agent m to help agent o in the current time instance, γ is the time discount, and Υ is the set of different task types. We assume that an agent is expert in only one of the possible task types. The evaluation of the expected utility of agent m helping agent o considers all possible interactions in future and for all task types. In equation 1, $D_m^t(x)$ is the expected future distribution of task types that agent m will receive at time instance t. We define $\mathrm{Pr}_{i,j}^t(x)$ as the probability that agent i will receive help from agent j at time step t, given it has a task of type x.

The term $\sum_{t=\mathcal{T}}^{\infty} \gamma^{t-\mathcal{T}} \sum_{x\in\Upsilon} D_m^t(x) \mathrm{Pr}_{m,o}^t(x) cost_m(x)$ represents the time discounted (with discount factor γ) expected savings of m by receiving helps from o in future. We assume that when an agent is helped by another agent, the helped agent incurs no cost for the task. Hence, when an agent m is helped with task type x, its savings is $cost_m(x)$, the cost it would have incurred to complete the same task on its own. We use an infinite time horizon and increasingly discount the impact of estimates for future interactions by the factor $\gamma^{t-\mathcal{T}}$, where $0 < \gamma < 1$, and t refers to the time period. The sum of the terms $-cost_m(\tau)$ and $\sum_{t=\mathcal{T}}^{\infty} \gamma^{t-\mathcal{T}} \sum_{x\in\Upsilon} D_m^t(x) \mathrm{Pr}_{o,m}^t\, cost_m(x)$ is the net expected cost that can be incurred by m for (a) helping on the current time instance and (b) incurring helping cost for o in the future. Thus, $E_{\mathcal{T}}(m, o, \tau)$ gives the net time-discounted future expected benefit that agent m has for interacting with agent o.

We note that the task distributions values, $D_m^t(x)$ in the future and the help giving probabilities $\mathrm{Pr}_{i,j}^t(x)$ will not be known to an agent. As an approximation, we estimate these values by the corresponding observed values over the last time period. For example, we use the ratio of the count of the number of times j helped i with task type x to the count of the number of tasks of type x received by i in the last time period to approximate $\mathrm{Pr}_{i,j}^t(x)$. We assume that the near past is a predictor of the near future, and accordingly estimate the probabilities of future task arrivals based on observed counts only in the last time period and not over all previous time periods. With this approximation, the $E(m, o)$ term given in Equation 1 can be rewritten as:

$$E_{\mathcal{T}}(m,o,\tau) = \sum_{t=\mathcal{T}}^{\infty} \gamma^{t-\mathcal{T}} [\sum_{x\in\Upsilon} (D_m^{\mathcal{T}}(x) \mathrm{Pr}_{m,o}(x) cost_m(x)) - \sum_{x\in\Upsilon} (D_m^{\mathcal{T}}(x) \mathrm{Pr}_{o,m}(x) cost_m(x))] - cost_m(\tau). \quad (2)$$

Our suggested decision mechanism $\mathcal{F}$ will evaluate a help request from agent o to agent m for a task x_0 of type τ at time $\mathcal{T}$ given the history of interactions, $\mathcal{H}_{m,o}^{\mathcal{T}}$ and $\mathcal{H}_{o,m}^{\mathcal{T}}$, between these two agents. The history is used to first calculate $\mathrm{Pr}_{m,o}(x)$ and $\mathrm{Pr}_{o,m}(x)$ values and the distribution of task arrivals. It then calculates the expected utility of agent m for interacting with agent o in the current and future time steps, $E_{\mathcal{T}}(m, o.tau)$. Our prescription is for agent m to help agent o in the current time if $E(m, o)$ is greater than the expected cost of an agent processing its next assigned task without receiving help from others, i.e.,

$$\mathcal{F}(o,m,\tau,\mathcal{H}_{m,o}^{\mathcal{T}} \cup \mathcal{H}_{o,m}^{\mathcal{T}}) = \text{Yes, if } E_{\mathcal{T}}(m,o,\tau) + \sum_{x\in\Upsilon} D_m(x) cost_m(x) > 0, \quad = \text{No, otherwise,} \quad (3)$$

where the summation term is the expected cost of agent m

for doing a task in the next time instance. This summation term represents the initial bias or willingness for an agent m to help another agent o incurring a risk of not being reciprocated. As initial probability values are all zero, the agent will not help another without this initial bias.

Each agent also maintains its own as well as others' estimates of costs for doing tasks of each type. It updates the estimates using a simple reinforcement learning mechanism that is outlined in Section *Problem domain*.

3. PROBLEM DOMAIN

We evaluate our hypothesis using simulations in a job completion problem domain. In this domain each of N agents are assigned m jobs. There are Υ job types and each agent has expertise in exactly one of these job types. Each job requires a finite time t and a quality of performance q to be completed. The total cost of finishing a job is t/q, where t is the time taken to complete the task and q is a quality measure indicating how well the task was completed. An agent who is an "expert" in a particular job type can do jobs of that type in less time and with higher quality, and therefore at lower cost, than other job types.

In our simulation, we generate the time and quality of performance from a normal distribution with a preset mean and a standard deviation. We use 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, i.e., 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 i.e. experts produce higher quality task completions. 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 of performance is the same for both experts and non-experts. Each agent is assigned the same number of tasks at each time period of our simulation that runs for a total of τ time periods. Once tasks are assigned, the agents ask for help from one another. 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 task is also computed. Help is obtained only if $C2 < C1$. This condition corresponds to a "cooperation possibility". Agents 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.

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 agent's capability of the relevant task type using the time taken and quality produced by the helper agent. The reinforcement scheme that we use to update the time and quality estimates, after $n + 1$ observations, is given by

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

where t_{ij} is the time taken by agent i to do task j on the n^{th} interaction between the two agents for this task type, and α is a learning parameter with values in the interval $(0, 1]$. We use a similar update policy for the quality of performance q_{ij}.

The agents complete all the assigned jobs for one time period and then receive their assignments for the next time period. The simulation runs for a fixed number of time periods. We have used a population of reciprocative agents only, all using the decision mechanism of equation 3. Our goal, in this paper, is to identify the suitability of such a strategy in generating savings for agents and allowing them to form mutually beneficial groups with agents of complementary capabilities and adapt to dynamically changing environment.

4. EXPERIMENTAL RESULTS

In this section we describe the results of our simulations that evaluate the performance of agent groups, the effects of changing agent behaviors and time-varying task generation on group performance.

In our experiments, we have considered three task types $\{0, 1, 2\}$. When an agent helps another agent, the helping agent incurs a cost by which it increases its "balance" with the helped agent. The helped agent, having saved some cost, decreases its balance with the helping agent by the cost it saved. We designate the total balance in a group (x, y) to be the sum of the balances between two agents such that one agent has expertise in task type x and the other has expertise in task type y. We use the total balance of a group as a metric of the effectiveness of the group. We also use the total savings earned by agents of a group, by receiving help from other members of the group, as a metric indicating the performance of a member of that group. In the following, we first describe our results under conditions where the task generation is time-invariant and agents do not change behavior. Next, we report results under dynamic conditions where agents detect changes in behavior of other agents and maintain future expected distributions of task types by observing task arrival trends over the last time interval. At the end we describe results with a much larger group of agents.

The primary hypotheses we are trying to evaluate are the following:

Opportunistic help-giving behavior: We expect our expected utility based help-giving agents to exchange help with agents of complementary expertise. Hence an agent with expertise in a particular task type would have significantly more negative balances with agents with expertise in other task types than with agents who are experts in the same task type.

Adaptability to change in agent expertise: If other agents change expertise, our expected utility based help-giving agent with an expertise in task type x should be able to respond in the following time interval and either increase, decrease, or maintain the level of its willingness to help that agent depending on whether that agent changed its expertise from task type x to $y \neq x$, from $y \neq x$ to x, or from $y \neq x$ to $z \neq x$ respectively.

Adaptability to change in task distribution: If the distribution from which the arriving tasks are drawn changes, our agent should adaptively modify its help-giving behavior to others. For example, if an agent is expert in task type x, and over time the frequency of task type $y \neq x$ increases (decreases), then this agent would be more (less) inclined to help agents with expertise in task type y.

Scalability of reciprocative behavior: We believe that this help-giving strategy will scale well with larger agent population provided that population is relatively stable. This implies that as the agent population increases, the pattern of help-exchange will not significantly differ provided the number tasks per time interval is increased proportionately.

4.1 Group performance under static conditions

In this set of experiments we studied group performance in terms of balances generated and the savings earned among agent members of a group. We assumed time-invariant task generation and static agent expertise. Thus, the probability of a task type arriving at any instant of time is fixed and agents do not change expertise. We used a total of 6 agents, with equal number of agents having each of the three expertise. The simulations were carried out with a total of 100 tasks per agent per time interval, for a total 10 time interval[1]. The discount factor γ was set at 0.25. Figures 1 and 2 show the results averaged over 10 runs.

Figure 1 shows the total savings earned per time interval by different agent groups over time. We have shown only the savings earned by groups formed by agents with different expertise, i.e., groups $(1, 2)$, $(0, 2)$ and $(0, 1)$. The savings earned by agents with similar expertise are much less compared to savings earned by groups of agents with heterogeneous expertise, and hence, they are excluded from the figure due to scaling issues. From the figure we notice that agents use the first time period to recognize the benefit of forming groups with agents with complementary expertise. In the following time periods, agents can use performance estimates from previous time period to effectively give and obtain help from other agents. The total savings of such groups of agents with complementary capabilities, therefore, converge to a steady state value. Thus, agents form stable relationships by identifying those with complementary capabilities and forming mutually beneficial groups.

Figure 2 shows the total, cumulative balance that an agent with expertise in task type 2 has with all other agents of different expertise. This agent does not have any balance with other agents with a matching expertise (agents with expertise in task type 2). This indicates that, it does not receive help from agents of its own type. Its balances with those having complementary expertise, in task types 0 and

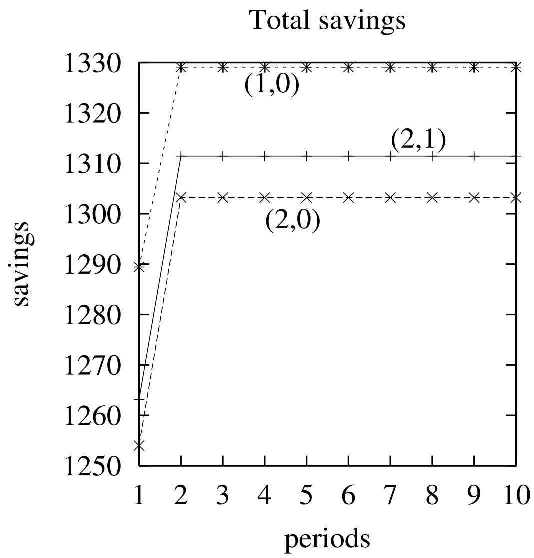

Figure 1: Total savings by different heterogeneous groups per time interval (static condition).

1, however, continually decrease over successive time periods. Note that a negative balance implies more help received than given. This shows how an agent benefits by receiving help from others with complementary expertise. The result corroborates the observation (in Figure 1) of the greater benefit generated by groups of agents with dissimilar expertise by explaining the phenomenon on a per-agent scale.

4.2 Group performance with varying task and agent behavior

In this set of experiments we remove the restriction of static task generation and non-changing agent expertise used in the previous experiments. First, we demonstrate the effectiveness of the expected utility based decision mechanism in enabling an agent to identify changing agent expertise and to respond by altering its relationships with those agents. Second, we show that with time varying agent expertise and task generation patterns, our agents can effectively adopt to the dynamic environment and shift their alliance with specific groups by gradually decreasing/increasing their interactions with them.

Figures 3 and 4 shows respectively the total group savings earned per time period by the groups of agents with heterogeneous capabilities and the cumulative balance obtained by an agent expert in a given task type from other agents of different expertise when some agents change their task expertise over time. A total of 6 agents and 100 tasks per agent per time interval were used. The simulation was conducted for 10 time intervals. A time discount factor γ of 0.25 is used. In this experiment, at time period 3 one agent switches expertise from task type 1 to 2. From Figure 3 we see that this change impacts the group savings earned by the groups $(0, 2)$ and $(0, 1)$ but not that of $(1, 2)$. Since there is one more agent with expertise in task type 2, the group $(0, 2)$ earns more total savings than the previous time period due to the additional help received from the agent who has changed its expertise. For similar reasons, group

[1]Though we work with finite time periods, we assume that the agents are not aware *a priori* of the number of tasks to process and thus will not be doing strategic reasoning based on the remaining number of tasks. It is well recognized in game-theoretic literature that to promote stable group interactions, it is preferred that agents use an infinite time horizon in their reasoning process. This will also be the case when agents have no knowledge about when their interactions will end. Also, agents are still more interested in receiving higher utility in the near future compared to the more distant future. These assumptions justify the use of the time discount parameter γ.

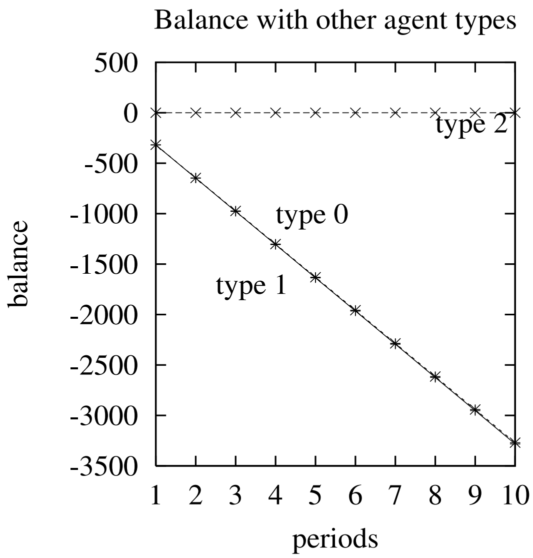

Figure 2: Cumulative balance of an agent with others (static condition).

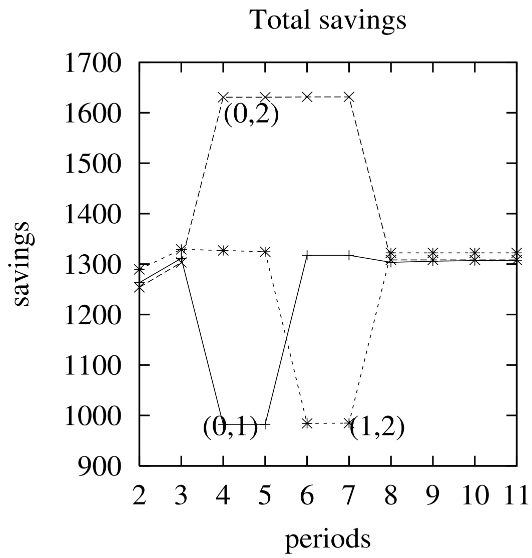

Figure 3: Total savings per time period of heterogeneous groups (time varying agent expertise).

$(0, 1)$ earns lesser than the previous time period because of the reduction of the size of agent population of experts in type 1 by one agent. The group $(1, 2)$ maintains their previous cumulative savings since there has not been any effective change in their composition. We observe a similar phenomenon at time period 5 when an agent of expertise in task type 2 switches to expertise in type 0. At period 7, again, we observe the effect of an agent of expertise type 0 switching to expertise type 1. We notice that group savings adapt flexibly to changes in agent expertise. This indicates that agents are able to quickly identify the switching of expertise by other agents and accordingly alter their affinity for or reluctance to interaction with corresponding groups. For example, at time instance 3, the agents identify the individual that has switched from expertise task type 1 to task type 2 and hence, are able to earn more savings as a group (group $(0, 2)$) by benefiting from the new expertise of the changed agent.

In Figure 4, we study the total balance that an agent x with an initial expertise in task type 1 has with other agents of different expertise. Agent x, initially having expertise in task type 1, had a zero balance with other agents of its own expertise in the first time period. It had a favorable balance with agents having complementary expertise (task types 0 and 2). At time period 2, x switched its expertise to task type 2. Agents with expertise in task types 0 and 1 are able to identify this alteration and the balances of x with agents of expertise in types 0 and 1 improve. At time period 4, another agent y switches expertise from task type 2 to task type 0. This increases the population of agents with expertise in task type 0 and reduces that of task type 2. Hence, agent x, who is now expert in task type 2, maintains a lower balance with agents of type 0 and higher with those of type 2 as there are less number of agent experts in that task type. At time period 6, there is another switch in expertise by an agent z from task type 0 to task type 1. Thus, the population of agents with expertise in type 1 increases while that with expertise in task type 0 reduces. The agent set with one of x's complementary expertise, task type 1, increases while the set with the other complementary expertise, task type 0, reduces. This results in reduced balances for agent x (still an expert in type 2) with the agents having expertise in task type 1 and increased balances with those having expertise in task type 0. Its balances with agents expert in task type 2 remain unchanged. This sequence of alterations in agent expertise and consequent change in group behavior for agent x implies that our agents, using the expected utility based strategy, are capable of responding effectively to such dynamic situations.

In the next experiment we studied the effects of changing the task generation pattern and having agents dynamically adopt new expertise on the help-giving behavior of the agents. Figure 5 shows the cumulative balance that an agent x with initial expertise in task type 1 has with all other agents of different expertise over successive time periods. Before time period 2, x had zero balance with others of similar expertise and favorable, i.e., negative balances with those having complementary capabilities. At time instant 2, x switches expertise to task type 2. We observe that its behavior with the other agents changes: its balance with agents of type 2 remains constant from thereon and those with expertise in task types 0 and 1 decrease. This qualifies our claim that the agents can recognize a change in agent expertise and alter their behavior by entering into new groups (x enters into stable relationship with agents of task type 1) or refraining from others (x refraining from the previously beneficial grouping with experts in task type 2). At time instants 4 and 6, we observe similar switching of expertise by other agents and corresponding shift in behavior by our agent x who continues to be an expert in type 2.

So far the description of events in Figure 5 resembles that of Figure 4. In Figure 5, we also observe the effect of altering the distribution of task arrival by decreasing the probability of occurrence of task types 1 by a constant percentage (2%

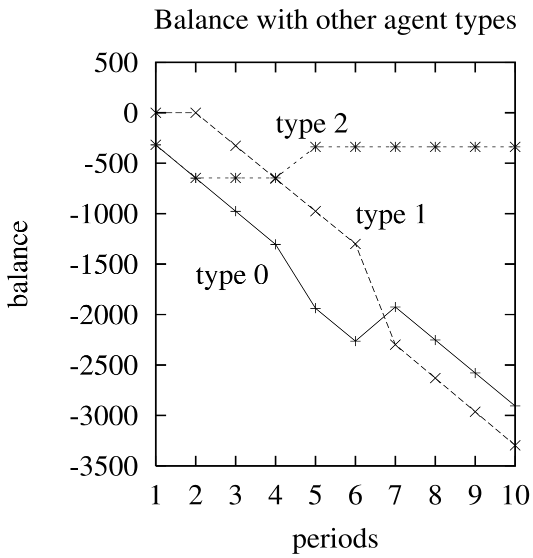

Figure 4: Cumulative balance of an agent with others (time varying agent expertise).

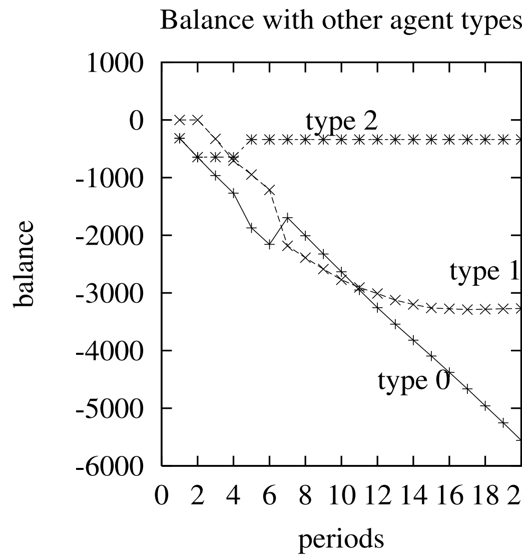

Figure 5: Cumulative balance of an agent with others (time varying task generation and agent expertise).

decrement) over successive time intervals and increasing the probabilities of occurrences of the other two task types by half the amount. As specified before, our agents estimate future expected distributions of task types by recording the arrival rates of different task types over the immediately previous time interval. The changing arrival rates of different task types are observed by the agents and they use this information to calculate the parameter $D_i^T(x)$ in Equation 2. We find that, beyond time instant 6, the balance earned by x with other expert in task type 1 slowly levels off. The balance of x with agents expert in task type 0, however, continues to increase. The results from this experiment clearly demonstrate that agents with an expectation of future task distributions are able to effectively adapt help-giving behavior under time-varying task distributions.

4.3 Scaling up

Experiments in the previous section were performed with a small number of agents. To find out the scaling up property of the expected utility based help-giving behavior we experimented with several larger agent groups. We now present results from experiments with a group size of 60, a group that is an order of magnitude larger than those in the experiments reported above.

In the first set of experiments, we start with an equal distribution of agent expertise over the three task types, i.e., initially there are 20 experts in each task type. We then decrease the number of experts in task type 1 by 2 in each time step and these agents become experts in task type 2. We stop this decrease when there are only 4 agents left with expertise in task type 1. At this point there are 36 agents expert in task type 2. In figure 6 we plot the balance of an agent expert in task type 0 with its average balance with experts in different task types. The balance of the agent with others expert in task type 0 is negligible. It is interesting to note that as the number of agents expert in a task type decreases (increases), the average balance of our agent, of complementary expertise, with those agents become more (less) negative. This is shown by the more negative value of the average balance with agents with task type 1 compared to task type 2. The two curves have opposite curvatures up to the point when there are no more expertise changes.

This observation can be explained from the fact that as there are less expert agents in task type 1, an agent with expertise in task type 0 is more likely to need help more frequently from any one agent expert in task type 1. Correspondingly this agent will also be more inclined to help agents with expertise in task type 1. This observation is supported by Figure 7 where we plot the expected utilities of an agent expert in task type 0 for helping agents with different task expertise. We note that the expected utility of the agent for interacting with a type 1 agent increases and that with a type 2 agent decreases over time. This is because there are increasingly more agents of the latter type than the former. The utilities become stable once the change of expertise stops.

We also plot, in Figure 8, the total balances of an agent with expertise in task type 0 with all the agent experts of different task types. From the figure we find that the total balance with agents expert in task type 2 dominate that of the total balance with agents expert in task type 1. This observation, combined with the previous analysis, means that even though less help is obtained from any one agent expert in task type 2 compared to an agent expert in task type 1, the total saving from experts in task type 2 is significantly larger than the total saving from experts in task type 1. This suggests that as experts in a given task type decreases, others do not get enough help for that particular task type. This is because the remaining few experts do not find it useful to honor all the help requests they receive. After providing some help, their expected utility calculation will prevent them from providing more help as that would be derogatory to their self-interest.

To round up the experiments, we ran a further set of ex-

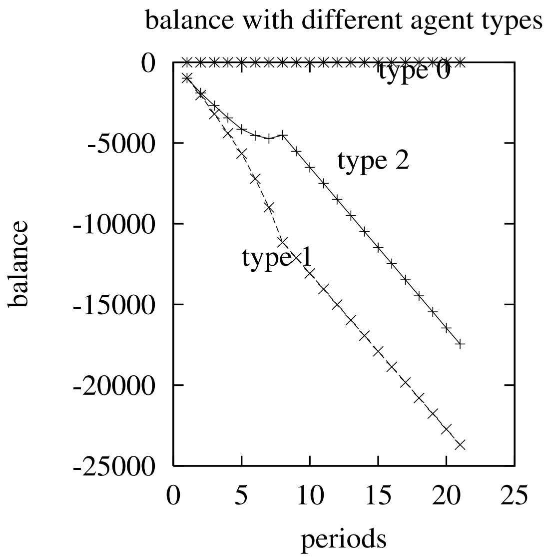

Figure 6: Balance of an agent with others (time varying agent expertise).

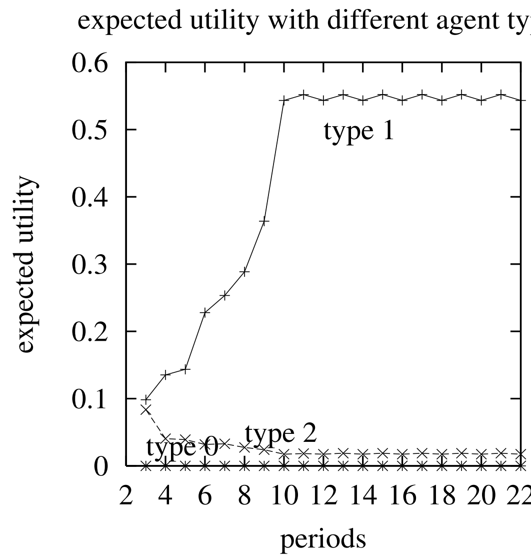

Figure 7: Expected utility of an agent for interacting with others (time varying agent expertise).

periments where in addition to the change of agent expertise, as in Figure 6, we also decreased the frequency of arrival of task type 1 and proportionately increased frequency of task types 0 and 2. This situation is similar to the results presented in Figure 5 with experiment run on a smaller group of agents. The results from this set of experiments is presented in Figure 9 and follow a similar pattern to Figure 5. We find that though initially a decrease in the number of experts in task type 1 makes our expert in task type 0 more inclined to exchange help those agents, over time such exchanges reduce significantly as fewer and fewer tasks of type 1 arrive. So, after a while, more help is exchanged with experts in task type 2 even though they are far more in number. This is because there is also a much larger number of tasks of type 2 and hence it proves beneficial to be more helpful to those agents.

The above set of experiments lend credence to each of our hypotheses about our expected utility based help-giving behavior including (a) responsiveness to agents with complementary expertise, (b) adaptability to changing agent expertise and task distribution, and (c) scale up to larger agent groups.

5. RELATED WORK

Complex problem solving involving distributed tasks often rely on the formation of effective agent groups. 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.

Some research in the area of coalition formation in agent societies has focused on cooperative agents [11]. A related work on coalition in agent societies takes self-interested agents into account [8]. 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. Learning of cooperative behavior has also been addressed in [4]. However, there the learning uses an off-line learning module which generates situation-action pairs using a genetic algorithm. To determine the fitness of the evolved situation-action pairs during off-line learning an agent-model is required. Our work uses only online learning and agents do not have to store a priori models of other agents.

A fair number of approaches have been developed that use reasoning mechanisms that consider exchange of help or social reputations in deciding how to interact with other agents. Our own work has concentrated on the use of a probabilistic decision mechanism for deciding to help based on past balance with the requesting agent [9]. Castelfranchi, Conte, and Paolucci use normative reputation [3] to enhance the performance of agents that comply with social norms. In the SPIRE framework developed by Grosz and collaborators [6, 13], performance in a group is improved when agents reason about the effects of withdrawing from social commitments on their reputation.

Our approach of using expected utility based decision mechanism is different from recent work on utility based strategies for sharing information [1] because we take into consideration dynamically changing environments and endow our agents with the ability to maintain future expectations of such dynamics to adaptively alter group behaviors.

6. CONCLUSIONS

We hypothesized that cooperative behavior based on future expected utility is a robust decision mechanism for agents to develop stable, mutually beneficial groups containing agents with complementary capabilities. This type of decision mechanism benefits agents by allowing them to quickly respond to changing behaviors of other agents and the environmental conditions. We simulated artificial environments and conducted experiments both under static conditions and where agents dynamically adopt new strate-

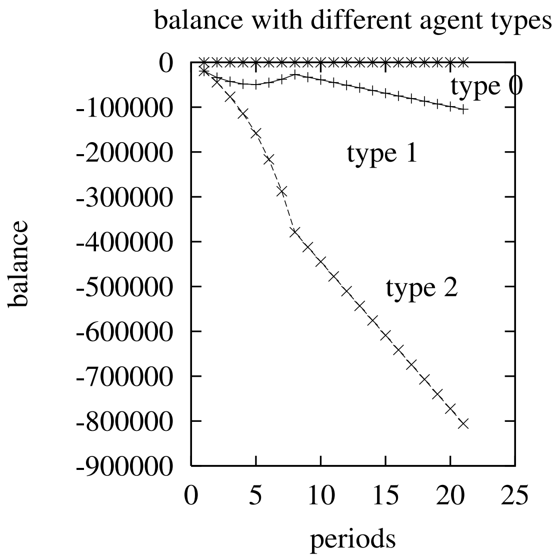

Figure 8: Total balance of an agent with each group of experts (time varying agent expertise).

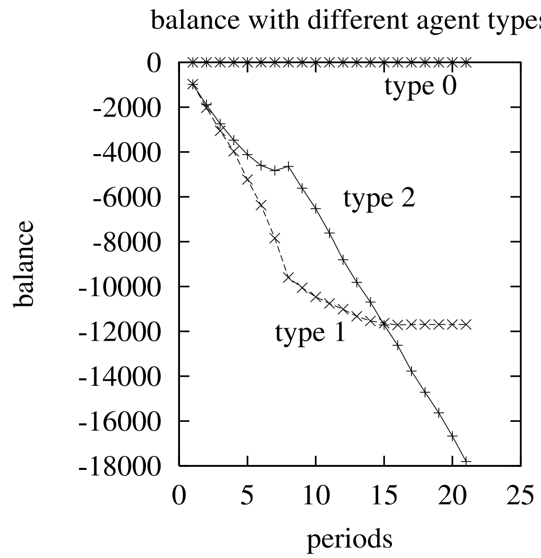

Figure 9: Balance of an agent with others (time varying agent expertise and task distribution).

gies or task arrival distributions change. Our experiments established that the expected utility based strategy allows the agents to form mutually beneficial relationships with agents of complementary qualities and to respond quickly and effectively to dynamic environments. We also demonstrated successful scale-up to agent groups that are larger by an order of magnitude.

We have restricted our simulations using agents employing the utility-based decision strategy. We plan to introduce other competitive decision mechanisms to determine the relative effectiveness of different agent interaction strategies under dynamic agent behavioral and environmental settings. Evaluating scenarios where agents have the freedom of using different decision mechanisms can produce interesting insight about evolving agent strategies. We also plan to investigate the effects of task failures and performance variation on the stability and strength of the relationship between agents of complementary expertise.

Acknowledgments: This work has been supported in part by an NSF CAREER award IIS-9702672.

7. REFERENCES

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

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

[3] Cristiano Castelfranchi, Rosaria Conte, and Mario Paolucci. Normative reputation and the costs of compliance. *Journal of Artificial Societies and Social Simulation*, 1(3), 1998.

[4] Joerg Denzinger and Micheal Kordt. Evolutionary on-line learning of cooperative behavior with situation-action-pairs. In *Proceedings of Fourth International Conference on MultiAgent Systems, ICMAS 2000*, pages 103–110, Los Alamitos, CA, 2000. IEEE Computer Society.

[5] Partha Sarathi Dutta and Sandip Sen. Identifying partners and sustenance of stable, effective coalitions. In *Proceedings of Fifth International Conference on Autonomous Agents, 2001*, pages 23–24, New York, NY, 2001. ACM Press.

[6] Alyssa Glass and Barbara Grosz. Socially conscious decision-making. In *Proceedings of the Fourth International Conference on Autonomous Agents*, pages 217–224, New York, NY, 2000. ACM Press.

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

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

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

[10] Sandip Sen, Anisha Biswas, and Sandip Debnath. Believing others: Pros and cons. In *Proceedings of the Fourth International Conference on Multiagent Systems*, pages 279–285, Los Alamitos, CA, 2000. IEEE Computer Society.

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

[12] Yoav Shoham and Moshe Tennenholtz. On the synthesis of useful social laws for artificial agent societies (preliminary report). In *Proceedings of the National Conference on Artificial Intelligence*, pages 276–281, San Jose, California, July 1992.

[13] David G. Sullivan, Barbara Grosz, and Sarit Kraus. Intention reconciliation by collaborative agents. In *Proceedings of the Fourth International Conference on Multiagent Systems*, pages 293–300, Los Alamitos, CA, 2000. IEEE Computer Society.