

Expertise Selection by Predicting Mutually Beneficial Partnerships

Dipyaman Banerjee, Sabyasachi Saha and Sandip Sen

Department of Mathematics & Computer Science

The University of Tulsa

{banerjdi,sahasa,sandip}@ens.utulsa.edu

Abstract

We consider a domain where each agent is an expert in a particular task type and can ask other, expert agents, to perform tasks for which it is not an expert. Agents are self-interested and respond favorably to requests for help only if the requesting agent is estimated to provide reciprocal benefits. It has been shown that self-interested agents can develop mutually beneficial cooperative relations with other like minded agents of complementary expertise in such domains. Previous work in the area presented a mechanism for forming coalition based on previous interaction history and expected future interactions. One constraint of that work was the assumption of fixed agent expertise. In a dynamic environment with continuously varying task distributions, however, agents will have to change their area of expertise to increase profitability and maintain competitiveness. The agent maintains a successful coalition till it is profitable. In this paper, we present an adaptive mechanism for choosing task expertise that estimates the likelihood of forming beneficial coalition with agents of complementary expertise allowing the agent to improve its utility. We augment this decision mechanism with an adaptive exploration strategy to improve robustness. Based on the experimental results, the new adaptive mechanism is shown to be more effective and responsive to the changes in the environment than other non-adaptive strategies.

Introduction

Autonomous agents interacting in an open world can be considered to be primarily driven by self interests. We believe, however, 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 (Dutta & Sen 2001) leading to exchanges of help can improve both agent and system-level performances (the latter through minimizing resource consumption, increasing throughput, etc.). Prior research has shown that these self-interested agents can develop mutually beneficial cooperative relations with other like minded agents of complementary expertise depending on historical data as well as on the expectation of future interaction (Saha, Sen, & Dutta 2003). In this research, the agents learn, over extended interactions, to form

coalition with those agents from which it can expect future savings. This also enables agents to recognize beneficial coalition and to avoid exploitation by the malevolent agents. But, in that framework, agent expertise were fixed. However, in a dynamic environment if an agent finds that it can improve its utility by changing its expertise then it should become an expert in a new task type and form more profitable coalition.

We consider a society of self interested agents situated in an environment where they are required to perform different types of tasks. Each agent is an expert in a particular task type. At a particular instance each of these agents is given a set of tasks to accomplish. The task distribution is assumed to be global, but is subject to change over time. An agent can transfer a task in which it is not an expert to an agent who is an expert in that type of task if the latter agrees to help. An agent helps another agent based on a decision mechanism which considers historical evidence as well as future cooperation possibilities (Saha, Sen, & Dutta 2003). In a dynamic environment, an agent's performance vary depending on the expertise distribution of the population, helping attitude of the other agents and the current task distribution. Periodically, an agent decides, based on its past interactions and expected future interactions with different agents, whether changing expertise can yield more revenue. An agent, however, incurs some cost to change its expertise.

In this paper, we design an integrated decision mechanism used by reciprocative agents that determine their help-giving behavior as well as selection of new expertise to adopt. We endow our agents with an expectation on the future task distributions and the ability to adapt its help-giving behavior to expertise changes of other agents. Also, as the distribution of tasks can change over time, we require these self-interested agents to periodically re-evaluate its performance and relationships with other agents and decide whether to adapt some different expertise in order to maximize future performance. To select a new expertise an agent estimates the cooperation possibilities in the future, based on which it calculates the expected utility for each type of expertise. Since estimations based on past interaction history with other agents loose fidelity over time, selective exploration by sampling different expertise is needed to increase reliability. The agent then switches its expertise to that task type which is expected to realize maximum utility. Such adaption is es-

sential to survive in a competitive environment and we posit that such an expected utility based decision mechanism is effective in generating significant cost savings under time varying task arrival distributions.

The expertise adaptation strategy recommends that an agent should change its expertise if the expected additional future savings generated by changing expertise in the future is more than the cost for changing expertise. The primary impediment to such adaptation is the paucity of information about the profitability of assuming other expertise. Let us assume a scenario having three types of tasks: T1, T2, and T3. An adaptive agent who is an expert in T1, finds it difficult to estimate the other agent's help-giving behavior towards it if it were to change its expertise to task type T2 or T3. To circumvent this problem we incorporate an initial exploratory phase, wherein an agent iteratively adopts each type of expertise for a certain period to obtain an estimation of the profitability of different task expertise under the initial task distribution. Though the information collected in this initial phase can be used to choose initial expertise, the premise of such choice becomes obsolete with time as the task distribution changes. A possible solution to this problem is to repeat the exploration cycle at regular intervals. With randomly changing task distribution this scheme allows agents to adapt expertise but suffers from the cost of regular exploration, as exploration requires the agent to try out all strategies for some length of time and some of these strategies have poor payoffs. For environments where task distribution changes have some predictable patterns an agent may be able to extrapolate past experience to compute preferred expertise and thus save on significant fraction of exploration costs. We present a predictive strategy where the agent tries to extrapolate future trends from its previous exploration results. The deviation between the predicted value and actual exploration results is used to tune the frequency of explorations. Small deviations signal prediction accuracy and hence the agent explore at longer intervals. In effect the agent skips unnecessary explorations for static or monotonically changing environments. Large deviations between prediction and observed performance, however, signifies abrupt changes in the environment and the agent explores more frequently to respond effectively to these fluctuations. In this paper, we experiment with different agent population configuration and task distributions and show the benefit of our prescribed mechanism for adaptive expertise selection.

Reciprocal behavior in agent societies

In the literature of the social sciences and economics, the adaptation of a group of self-interested agents is dealt with a great importance. The social sciences researchers analyze the nature of altruism and the cause for its evolution and sustenance in animal groups (Axelrod 1984). Mathematical biologists and economists evaluate the rationality of altruistic behavior in groups of self-interested agents by proposing fitness models that analyze the success of altruistic individuals and the evolution of altruistic genetic traits (Dugatkin *et al.* 1994; Nowak, May, & Sigmund 1995). We do not intend to model altruistic behavior in animals or humans and hence do not address the issues raised in the social science or

experimental economics literature on this topic (Hoffman, McCabe, & Smith 1998). A significant body of work by mathematical biologists or economists on the evolution of altruistic behavior deals with the idealized problem called Prisoner's dilemma (Rapoport 1989) or some other repetitive, symmetrical, and identical 'games'. To consider a well-known study in this area, Axelrod demonstrates that a simple, deterministic reciprocal scheme or the *tit-for-tat* strategy is quite robust and efficient in maximizing local utility (Axelrod 1984). Sen criticizes the simple reciprocative strategy is not the most appropriate strategy to use in most real-life situations because most of the underlying assumptions that motivate its use are violated in these situations (Sen 1996). The evaluation framework used by Axelrod considers an evolving population composition by allowing propagation of more successful behaviors and elimination of unsuccessful ones. Sen *et. al.*, showed (Sen & Dutta 2002) what behaviors emerge to be dominant or are evolutionarily stable. In this paper, we show for a given task distribution and agent population how the agents adapt to new expertise in order to reach the stable population that produce maximum utility to each of the self-interested agents and in turn increase the profit of the entire system.

Decision procedure for changing expertise

Here, we have considered a task completion domain where each agent is assigned a set of tasks to accomplish. We assume a set of \mathcal{A} agents executing tasks from a set Υ and the set of task types is given by Γ . 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 $\langle i, j, x, t, c_i, c_j, \text{help} \rangle$ 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 . In our model, an agent asks for help from other agents in the group when it needs to accomplish a task in which it is not an expert. For simplicity, we have assumed that it asks for help randomly (without replacement) from the other agents. 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. In our model, given the task set Υ and the expertise of the different agents in \mathcal{A} , a self-interested agent $ag \in \mathcal{A}$ needs to take two different types of decisions. First, whether to honor a help request asked by another agent and second, after all the tasks are accomplished in one iteration, whether it will change its expertise and become an expert in some other task type that maximizes its expected future savings.

We have discussed the first decision procedure in (Saha, Sen, & Dutta 2003) as:

$\mathcal{F} : \mathcal{A} \times \mathcal{A} \times \Upsilon \times H \rightarrow \text{Yes/No}$ 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 presented 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, estimates the utility of agreeing to the request by evaluating its chance of ob-

taining help from the asking agent in future. The agent 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 savings from the asking agent and the expected cost it might incur for that agent by helping it in the future. The agent helps only if its expected benefit exceeds the estimated risk over all future interactions with the requesting agent.

In this paper we introduce a future expected utility maximizing decision mechanism for the second type of decisions for the agents (*i.e.* whether to change expertise). An adaptive agent takes this decision after each iteration *i.e.* after it accomplishes all tasks assigned to him. We present the decision function as: $\mathcal{G} : \mathcal{A} \times \Upsilon \times H \times T \rightarrow \Gamma$. This function yields the task types in which a given agent will change its expertise to maximize future utility given the task set, interaction history, and current time.

In the following, we present the expected utility based decision mechanism that agent m uses to choose a task type in which to become an expert.

$$\mathcal{G}(m, \tau, \delta, T) = \arg \max_{\beta \in \Gamma} \sum_{a \in \mathcal{A}} \mathcal{K}(m, a, T, \beta, \delta) - \text{chcost}(\tau, \beta), \quad (1)$$

The function $\text{chcost}(\tau, \beta)$ represents the cost of changing expertise and is 0 if $\tau = \beta$. Summation of $\mathcal{K}(m, a, \beta)$ over all a in the set of agents \mathcal{A} gives the total future expected utility of an agent m if its expertise is in task type β in the next iterations. Taking maximum of β over Υ , \mathcal{G} returns that task type which yields maximum future utility over extended period of action.

$$\begin{aligned} \mathcal{K}(m, a, T, \beta, \delta) = & \sum_{t=T}^{\infty} \gamma^{t-T} \left[\sum_{x \in \Gamma} (D_m^t(x) \text{Pr}_{m,a}^1(x, \beta, \delta) \text{cost}_m(x)) - \right. \\ & \left. \sum_{x \in \Gamma} (D_a^t(x) \text{Pr}_{a,m}^2(x, \beta, \delta) \text{cost}_m(x)) \right], \end{aligned} \quad (2)$$

where $\text{cost}_i(x)$ is the expected cost that i incurs doing a task of type x , γ is the time discount, and Γ is the set of different task types. τ is the current expertise of the deciding agent and $\beta \in \Gamma$ is another task type that the agent is considering changing its expertise to. In equation 2, $\mathcal{K}(m, a, T, \beta, \delta)$ is the expected future utility, or gain, of the agent m from agent a , starting from time T , if the agent m 's expertise is β and given the interaction history δ . This evaluation of the expected utility of agent m helping agent a considers all possible interactions in future and for all task types. In equation 2, $D_m^t(x)$ is the expected future distribution of task types that agent m will receive at time instance t . We define $\text{Pr}_{i,j}^1(x, \beta, \delta)$ as the probability that agent i will receive help from agent j if agent i is of expertise β , given it has a task of type x and $\text{Pr}_{i,j}^2(x, \beta, \delta)$ as the probability that agent i will receive help from agent j if the agent j is of expertise β , given it has a task of type x . Both these probabilities are dependent on the interaction history δ . Therefore, the

first sum in equation 2, represents the time discounted (with discount factor γ) expected savings of m by receiving helps from a in future if agent i is of expertise β . Similarly, the second sum in equation 2, represents the time discounted expected cost incurred by agent m for helping agent a in future if agent i is of expertise β . Hence $\mathcal{K}(m, a, T, \beta, \delta)$ is the expected future utility (or gain) of the agent m from agent a if the agent m 's expertise is β . When an agent m is helped with task type x , it incurs no cost and hence its savings is $\text{cost}_m(x)$, its own cost for doing that task.

So, after each iteration, the adaptive agent evaluates the maximum utility generating expert types based on earlier interactions and expected future expectations. Now, in a practical scenario we may not know the probabilities used in equation 2. To counter this problem the adaptive agent will need to perform exploration at regular intervals. We presented results in the experimental section using this scheme. However, as pointed out earlier this strategy seems to be somewhat naive as it explore unnecessarily even for static or uniformly changing task distribution. During exploration, each expertise is adopted for a full iteration and hence the agent performance drops when adopting expertise with low utility. To overcome this problem our agent tries to predict the future utility values from past exploration information. The degree of accuracy of prediction then governs the frequency of exploration. If prediction is accurate, the agent will explore less often and hence can improve performance by delaying exploring expertise with low payoff.

At any time t let $e_n^t(\beta)$ be the time instance where the n th exploration of expertise β before time t took place, with $e_1^t(\beta)$ be the time for the immediately preceding exploration. To accommodate the possibility of a changing environment, we use an extrapolation term, based on the last two exploration experiences, in the calculation of the currently preferred expertise, *i.e.* we redefine the \mathcal{G} function as:

$$\mathcal{G}(m, \tau, \delta, T) = \arg \max_{\beta \in \Gamma} \mathcal{E}(m, T, \beta, \delta) - \text{chcost}(\tau, \beta), \quad (3)$$

where the extrapolated expected benefit for changing to expertise β is

$$\mathcal{E}(m, T, \beta, \delta) = \mathcal{S}(m, T, \beta, \delta) + \mathcal{P}(m, T, \beta, \delta),$$

where the \mathcal{S} function represents the total expected benefit from all agents given the information from the last exploration

$$\mathcal{S}(m, T, \beta, \delta) = \sum_{a \in \mathcal{A}} \mathcal{K}(m, a, T, \beta, \delta) \quad (4)$$

and the \mathcal{P} function incorporates an average predicted trend from the last two explorations to account for steady drifts in the environment

$$\mathcal{P}(m, T, \beta, \delta) = \frac{\mathcal{S}(m, e_1^T(\beta), \beta, \delta) - \mathcal{S}(m, e_2^T(\beta), \beta, \delta)}{e_1^T(\beta) - e_2^T(\beta)}. \quad (5)$$

Just after the exploration of an expertise, we calculate the time for the next exploration of that expertise. The gap in exploration should be more if our last exploration results were consistent with the immediately preceding expectations and

vice versa. The time for the next exploration of expertise β is calculated as:

$$\mathcal{N}(m, \mathcal{T}, \beta, \delta) = (e_1^{\mathcal{T}}(\beta) - e_2^{\mathcal{T}}(\beta))(1 + e^{-\lambda \mathcal{D}(m, \mathcal{T}, \beta, \delta)}) + e_1^{\mathcal{T}}(\beta),$$

where λ is a tuning constant and the function \mathcal{D} measures the fractional difference between the expected and the explored values of the benefit of expertise β

$$D(m, \mathcal{T}, \beta, \delta) = \frac{|S(m, e_1^{\mathcal{T}}(\beta), \beta, \delta) - \mathcal{E}(m, e_1^{\mathcal{T}}(\beta) - 1, \beta, \delta)|}{\mathcal{E}(m, e_1^{\mathcal{T}}(\beta) - 1, \beta, \delta)}. \quad (6)$$

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.

Experimental results

We present the performance of our adaptive agent under different dynamic scenarios. We compare the performance of the three strategies:

Basic Reciprocative (BR): These are reciprocative agents that do not change expertise. This is also referred to as the non-adaptive strategy.

Adaptive Reciprocative (AR): Reciprocative agents that adapt their strategies based on the information obtained from exploration at fixed intervals. These agents explore every N iterations (we have used $N=10$ in our experiments). This is also referred to as the non-predictive adaptive strategy.

Predictive Adaptive Reciprocative (PAR): Adaptive reciprocative agents who vary their exploration frequency based on prediction mismatches. These agents adjust the time to next exploration based on past prediction errors.

In our experiments, all agents in each group is of the BR type. Only one agent in each group is either of type AR or of type PAR.

We also evaluate the effect of expertise distribution of the agent population on the performance of the adaptive agent. In this paper we have experimented with the task completion domain where agents are required to complete assigned tasks of different types. In our experiments, we have considered three task types $\{0, 1, 2\}$ and each run consisted of 40 instances of task allocations to agents. Unless otherwise mentioned, the number of experts in different task types are the same.

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. Hence, more negative balance indicates better performance.

In the first experiment, we consider a total of 10 agents in the society. At the start of each iteration, 400 tasks are assigned to each agent. In this experiment we have used

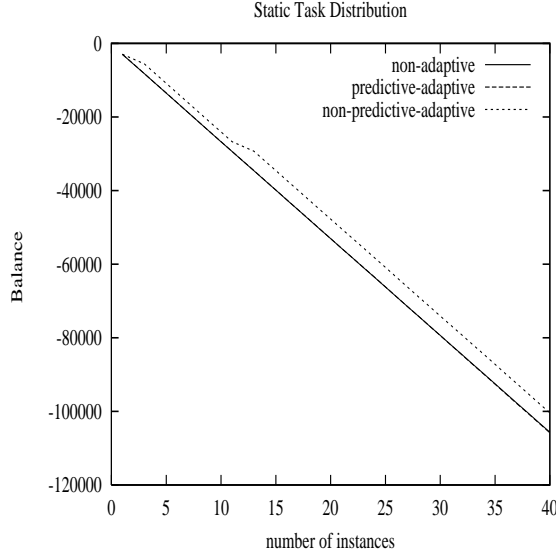


Figure 1: Performance of BR, AR, and PAR agents under static task distribution.

an asymmetric but static task distribution. We consider $\{200, 150, 50\}$ as the number of tasks of three types that is received by each agent at the beginning of each iteration. It is seen that under these conditions adaptive agents (both predictive and non-predictive) will change their expertise and become an expert in task type 2. This is an important observation. This means that the adaptive agents try to be an expert of a task which is less frequent in the environment. The reason here is, since all the agents are reciprocative, it finds helping agents for its non-expert tasks and incurs more profit (as it does not incur any cost for accomplishing that task) and on the other hand in return it needs to help in less number of tasks as tasks of its expertise is the least frequent task in the task distribution. So, the adaptive agent takes this decision to change its expertise to type 2 it earns maximum future utility. So the adaptive strategy make most out of the environment which suits its self interested perspective. Figure 1 also shows that the predictive agent performs better than his non-predictive peer. This happens as predictive agent captures the static nature of the environment and skips unnecessary explorations whereas the latter fails to do so and thus falls short. However, we must understand that non-adaptive agents who are expert in the most profitable task type from the beginning performs equally well in this static condition.

In the second experiment, we changed the task arrival distribution after each iteration. The agent population configuration is same as the last experiment. In the first iteration, number of $\{0, 1, 2\}$ types of tasks are $\{133, 133, 134\}$. We introduced a seasonal variation in the task distribution. We increased the number of tasks of type 1 monotonically for the first 15 instances and decrease that of task type 2 keeping the total number of tasks constant. After 15 instances we reverse the scenario. Having this seasonal, regular fluctuation

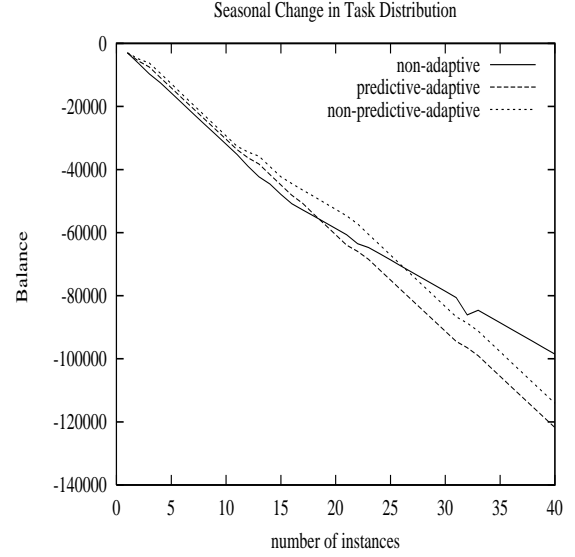


Figure 2: Performance of BR, AR, and PAR agents when task distribution is changing seasonally.

in the environment the predictive agent outperforms the other two agents by a large margin. Figure 2 shows that the performance of non-adaptive agents drop after instance number 15 due to its incapability to adapt to the changes. The predictive agent adapts successfully and also extrapolates its past exploration results to save exploration costs. The non-predictive agent, although adapting, lags behind the predictive agent as it cannot track the regularity of the change in the environment.

In the next experiment we varied the task distribution randomly. The number of experts in each task type is kept identical to previous experiments. The experimental results when plotted in Figure 3 confirms our hypothesis that in case of random changes in task distribution, our predictive agent performs only as effectively as the non-predictive agent and hence the two plot converges. The PAR agents cannot predict accurately and hence end up exploring as often as the AR agent. We also observe, however, that their performances are much superior to that of the basic reciprocative agents who could not respond to such abrupt fluctuation in the environment.

In the next experiment, we have shown the effect of population configuration on the adaptive strategy. In the earlier two experiments we used equal distribution of agent expertise and changed the task distribution. In this experiment, we kept the task distribution static but allow some agents to randomly change their expertise. This results in a dynamically changing population configuration and thus provides scope for adaption. We see from the results plotted in figure 4 that in this case also the predictive agent adapts successfully to this other kind of environmental dynamics and outperforms the non-adaptive agents.

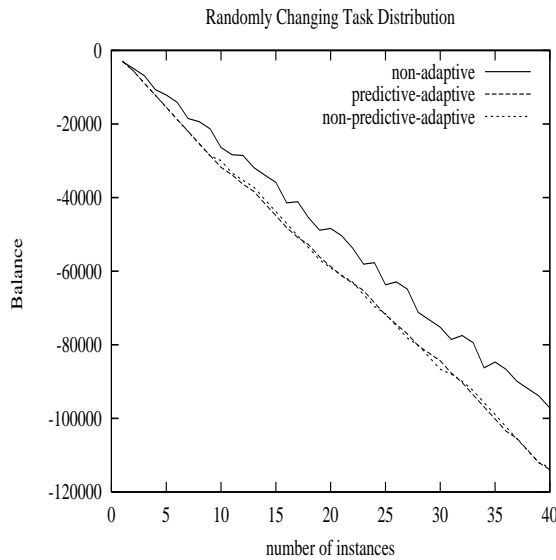


Figure 3: Performance of BR, AR, and PAR agents when task distribution is changing randomly.

Conclusions and future work

In this paper we have studied the performance enhancement by the adaptive expertise-selection strategy in more realistic scenarios where agents can change their behavior if such a change is perceived to be beneficial. We have analyzed the effect of changes in the task distribution and the effect of initial population on the success of adaptive expertise selection. In the experiments, we have shown how the adaptive agents dominate the non-adaptive ones. We have seen that agents change their expertise in direction where it requires to do less number of help but receives more help from the other agents for tasks in which they are not experts. In this paper, we have considered only reciprocative agents and as a result the agents find it more beneficial to be in that group which consists of more experts. Because then it needs to help less number of people. We believe that the condition will change if there are some selfish agents in the domain. The benefit of adaptation will be magnified over longer time periods and hence such adaptive mechanisms are particularly well-suited for use in semi-stable long-lasting social networks.

We have also developed a predictive adaptive agent that saves on exploration cost by tracking environmental changes. This agents are particularly effective when task distribution changes or changes of expertise distribution in the population follows a regular pattern. Experimental confirmation lends support to the effectiveness of this predictive adaptive strategy.

The mechanism proposed in this paper relies on historical interaction with other agents. The richer the history and the more stable such interactions, the more the benefit that can possibly accrue from using this adaptive decision procedure. As such the mechanisms of reciprocity is particularly well-suited for long-term interaction in agent groups that though open, are semi-stable in nature, i.e., agents are free to join

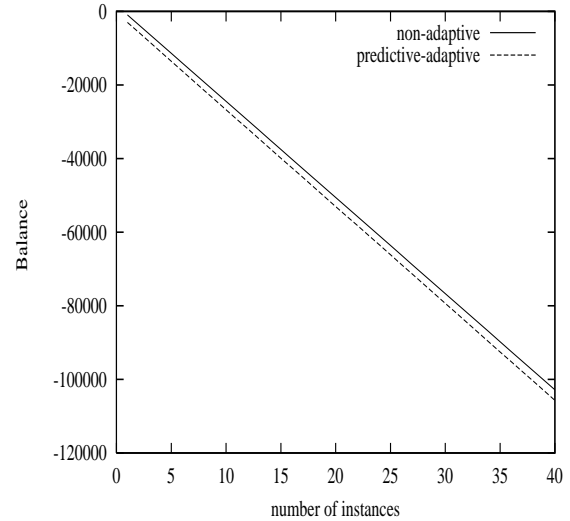


Figure 4: Performance of BR, AR, and PAR agents when population configuration is changing.

and leave such groups, but the statistical characteristics describing group composition changes only slowly over time.

One future goal is to analytically capture the dynamics of the evolution of agent population. It will be interesting to see the performance of the adaptive agents when the agent population changes continually after each iteration. We have also plan to see the effect of other adaptive agents in the performance of an adaptive agent when there is more number of adaptive agents in the environment. We are also planning to see the performance of the different agents under different population compositions.

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

References

- Axelrod, R. 1984. *The Evolution of Cooperation*. Basic Books.
- Dugatkin, L.; Wilson, D.; III, L. F.; and Wilkens, R. 1994. Altruism tit for tat and 'outlaw' genes. *Evolutionary Ecology* 8:431–437.
- Dutta, P. S., and Sen, S. 2001. Identifying partners and sustenance of stable, effective coalitions. In *Proceedings of Fifth International Conference on Autonomous Agents, 2001*, 23–24. New York, NY: ACM Press.
- Hoffman, E.; McCabe, K. A.; and Smith, V. L. 1998. Behavioral foundations of reciprocity: experimental economics and evolutionary psychology. *Economic Inquiry* 36(3):335–352.
- Nowak, M.; May, R.; and Sigmund, K. 1995. The arithmetics of mutual help. *Scientific American* 272(6):76–81.

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

Saha, S.; Sen, S.; and Dutta, P. S. 2003. Helping based on future expectations. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multi-agent Systems*, 289–296. New York, NY: ACM Press.

Sen, S., and Dutta, P. S. 2002. The evolution and stability of cooperative traits. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-agent Systems*, 1114–1120. New York, NY: ACM Press.

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