Agent Decision-Making in Open Mixed Networks

Ya'akov Gal[a,b], Barbara Grosz[b], Sarit Kraus[c,d], Avi Pfeffer[e], Stuart Shieber[b]

[a]*Department of Information Systems Engineering, Ben-Gurion University of the Negev*
[b]*School of Engineering and Applied Sciences, Harvard University*
[c]*Computer Science Department, Bar Ilan University*
[d]*Institute for Advanced Computer Studies, University of Maryland*
[e]*Charles River Analytics*

Abstract

Computer systems increasingly carry out tasks in mixed networks, that is in group settings in which they interact both with other computer systems and with people. Participants in these heterogeneous human-computer groups vary in their capabilities, goals, and strategies; they may cooperate, collaborate, or compete. The presence of people in mixed networks raises challenges for the design and the evaluation of decision-making strategies for computer agents. This paper describes several new decision-making models that represent, learn and adapt to various social attributes that influence people's decision-making and presents a novel approach to evaluating such models. It identifies a range of social attributes in an open-network setting that influence people's decision-making and thus affect the performance of computer-agent strategies, and establishes the importance of learning and adaptation to the success of such strategies. The settings vary in the capabilities, goals, and strategies that people bring into their interactions. The studies deploy a configurable system called Colored Trails (CT) that generates a family of games. CT is an abstract, conceptually simple but highly versatile game in which players negotiate and exchange resources to enable them to achieve their individual or group goals. It provides a realistic analogue to multi-agent task domains, while not requiring extensive domain modeling. It is less abstract than payoff matrices, and people exhibit less strategic and more helpful behavior in CT than in the identical payoff matrix decision-making context. By not requiring extensive domain modeling, CT enables agent researchers to focus their attention on strategy design, and it provides an environment in which the influence of social factors can be better isolated and studied.

1. Introduction

Computer systems are increasingly being deployed in group settings in which they interact with people to carry out tasks [3, 46, 42, 44, 28]. To operate effectively in such settings, computer agents need capabilities for making decisions and negotiating with other participants—both people and computer-based agents—about the procurement and allocation of resources necessary to complete their tasks. For example, in a civil disaster like an earthquake, rescue personnel and equipment are dispersed geographically and may be under the jurisdiction of various dispatchers. Dispatchers might depend on computer agents to allocate these limited resources to affected locations quickly and to alert them about changing environmental conditions, such as wind speed and traffic. In turn, computer agents might depend on people to provide up-to-the-minute information about the availability of personnel and equipment. In another realm, in some electronic auction settings, both people and computer agents (representing groups or individuals) might participate not only to acquire items of value, but also to exchange information about the reliability of others.[1]

First response and e-commerce are two different kinds of examples of open mixed networks. By "open" we mean that the autonomous agents in the network may be designed by or represent different individuals or organizations. By "mixed" we mean that the participants of the network may be computer agents or people. Computer agents operating in open, mixed networks may support people in their work (e.g., collaborative human-computer interfaces [47, 3]), serve as proxies for people or institutions (e.g., electronic commerce [25, 44]), or interact with other agents to carry out tasks for which they are responsible (e.g., robots in rescue operations [46, 36]). These examples exhibit several key characteristics of mixed network settings: (1) the participants are both human and computer-based; (2) they depend on each other to make decisions; (3) they may need to exchange resources and information; (4) they have different, complementary roles.

Open mixed network settings present a range of challenges for agent designers. First, the participants in these networks—whether people or com-

---

[1]One example of an existing application like this are sniper agents that bid on e-bay as proxies for their human users.

puter agents—are loosely coupled and not under the control of any single entity. Agent designers are unlikely to know a priori the strategies that people or agents designed by others will adopt, and they cannot force others' agents to adopt a particular strategy. Second, people's decision-making behavior in group settings does not follow the strategies of classical economic or game theoretic models, but is affected by such social and psychological factors as cognitive biases, social preferences, and framing effects [12, 7, 4]. It is difficult to measure the effects of such factors directly, and preferences are hard to elicit explicitly from people [9, 35]. Third, agents may differ in their goals and plans, so agent designers need to develop strategies that are flexibly able to accommodate different levels of cooperation or competitiveness. For these reasons, it is at best challenging, and at worst, impossible, to construct effective agent strategies purely analytically.

An alternative approach is to learn and evaluate agent strategies empirically. However, past empirical investigations of computer agent strategies such as the Trading Agent Competition [1] and RoboCup soccer [2] have typically required a fully specified domain model. The need for extensive modeling of domain specific knowledge in such settings makes it difficult to distinguish among possible causes of agents' failures and successes, such as the way agents model the specifics of the domain or the way they make decisions more generally.

On the other hand, completely abstract settings such as the payoff matrices or decision trees traditionally used in studies in the behavioral sciences collapse the structure of a domain into a list of choices that does not capture the essential relationships among tasks, goals and resources.[2] Such relationships often play an important role in decision making.

The investigations in this paper were done in an environment that represents an intermediate approach. They use the CT (Colored Trails) system [20] which provides an analogue to the ways in which goals, tasks and resources interact in real-world settings, but abstracts away the complexities of real-world domains. CT supports comparisons of the performance of different computational strategies for interacting in groups comprising people and computer agents as well as solely computer agents.[3]

---

[2]For a comprehensive account of behavioral economics experiments in decision-making, see Camerer [7].

[3]CT is open-source software, which has been made available for download at http://www.eecs.harvard.edu/ai/ct.

This paper presents several new decision-making models that represent, learn and adapt to various social attributes of negotiation in open, mixed-network settings. We consider in particular, social factors that influence possible negotiation deals (e.g., joint benefit and inequality of outcome), traits of individual negotiators (e.g., altruism, trustworthiness, helpfulness) and group structure (e.g., solidarity, hierarchy). Our results show that (1) people exhibit more helpful behavior and increase their social welfare in CT settings than in payoff-matrix types of settings; and (2) computer agents that model and learn the social factors that influence human negotiation strategies can outperform traditional game-theoretic equilibria strategies when interacting with people and other computer agents in mixed networks.

The contributions of the paper are four-fold: it presents new multi-agent decision-making models, ones that are able to learn and adapt to the social attributes that affect behavior in open mixed networks; it presents a novel approach to evaluating such models; it shows empirically that agents using these models outperform traditional game-theoretic equilibria strategies. Lastly, it describes CT more completely than before as a new environment for investigating the design and performance for negotiation strategies in open-mixed networks. It integrates earlier reports of initial CT studies [17, 16, 48] and describes a broader range of experimental investigations which demonstrate the flexibility of the CT infrastructure to support different agent-design studies.

The purpose of the work reported in this paper was not to design "plug-and-play" strategies for specific applications such as first response or electronic commerce. Rather, the studies we describe show empirically that agents will be better able to negotiate with people if they take into account social factors. In this respect, our results relate to recent work in the social sciences that point to societal and cultural factors that people "bring into the game", as influencing the way they behave in negotiation settings [41, 8]. Our studies differ from these in providing and evaluating computational models for decision-making in these settings. In particular, they identify the influence of factors that have not been addressed in past human-computer decision-making studies and show the influence of these factors on agent-strategy performance.

The next section of this paper describes CT and the ways in which it corresponds to real-world task settings; it compares the CT environment with alternative test-bed environments used in multi-agent system and discusses related work. Section 3 presents a study that establishes that people make

4

different negotiation choices when interacting in a CT game than when doing so in the payoff-matrix settings used in the behavioral sciences. Section 4 describes a model for learning the different types of social factors that affect people's behavior in a simple negotiation setting of complete information. Sections 5 and 6 compare the performance of different computational strategies for adapting to other agents' decision-making in a repeated negotiation setting. Section 7 discusses the implications of this work for the design of agent-strategies in open, mixed networks and presents several open research questions.

2. The Colored Trails Game

Colored Trails (CT) is a game played by two or more participants on a board of colored squares. The CT system is highly configurable, allowing for the specification of games that reflect a wide variety of task environments and decision-making situations. The basic CT board includes players' icons and goal squares, but configurations may also include traps or other features needed to model a task situation. Players are typically allocated a set of chips of colors chosen from the same palette as the board squares. To move a piece into an adjacent square a player must turn in a chip of the same color as the square. The board state may be dynamic, for instance with goals moving or traps appearing and disappearing. A player's performance in CT is determined by a scoring function which is a parameter of the game configuration. The score can be defined to depend on such factors as a player's distance from its goal-square when the game ends, the number of moves made, the number of chips the player possesses at the end of the game, or the number of goals achieved. In addition, to represent incentives for cooperation or social good, an individual player's score can be made to depend on the performance of other players.

In the canonical use of CT, paths through the board represent doing a task, with each square in the path corresponding to a subtask. Chips represent resources needed for doing the tasks. Typically, at least one player does not have the chips needed to reach its goal. The heart of the game is players' abilities to negotiate over these resources. Chips may be exchanged by the players, and the conditions of exchange may be varied to model different decision-making situations.
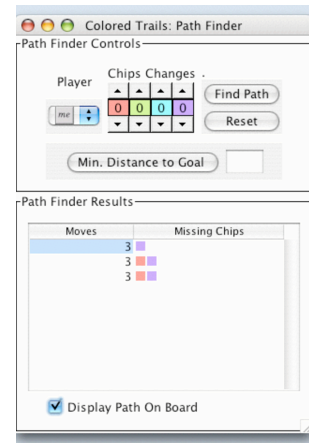
Humans play CT using a graphical user interface, while computer agents use a set of application program interfaces (API). Snapshots of the CT GUI
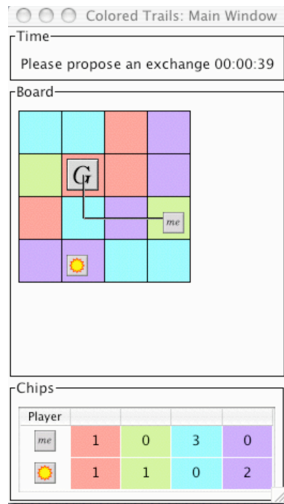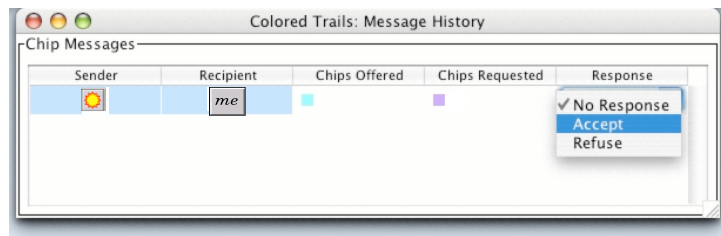
(a) Main Game Panel (shown at the onset of the game)
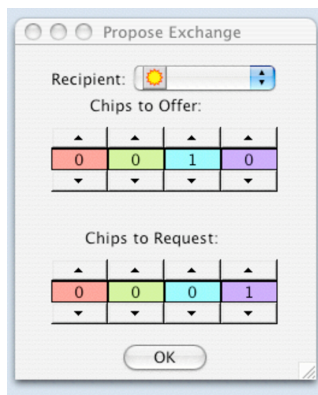
(b) Proposal Panel (used by players to make offers)

(c) Decision Aid Panel (used by players to reason about possible paths to the goal)

(d) Message History Panel (shows past offers and responses)

Figure 1: Snapshots of CT GUI

for one of the canonical games used in this paper is shown in Figure 1. The Main Game panel (Figure 1a) includes the board game, the goal square, represented by an icon displaying the letter G, and two icons, "me" and "sun", representing the location of the two players on the board at the start of the game.[4] The bottom part of the Main Game panel, titled "Chips", shows the chip distributions for the players. In the game shown here, neither player has sufficient chips to get to the goal square. The Decision Aid panel (Figure 1c) provides decision support tools to be used by players during the game. It displays a list of possible paths to the goal for players and the chips required to fulfill each path. Players can view this information for the chips that they currently possess, or for any hypothetical chip set for each of the players. For example, the "me" player is lacking a purple chip to get to the goal in the path that is highlighted in the Main Game panel (this is the shortest possible path for the "me" player to get to the goal). Similarly, the "sun" player is lacking a cyan chip to get to the goal (using its shortest path to move two squares up). The Proposal Panel (Figure 1b) can be used by players to negotiate exchanges. In the example shown here, the "me" player has offered to give one cyan chip to the "sun" player in return for one purple chip. This offer is displayed in the Message History Panel (Figure 1d), and also includes the response of the "sun" player, which accepted the offer.[5]

2.1. Analogy with Task Settings

CT provides a realistic analog to task settings, highlighting the interactions among goals, tasks required to achieve these goals, and resources needed for completing tasks. Chips correspond to agent capabilities and skills required to fulfill tasks. Different squares on the board represent different types of tasks. A player's possession of a chip of a certain color corresponds to having the skill available for use at a time. Not all players possess chips in all colors, much as different agents vary in their capabilities. Traversing a path through the board corresponds to performing a complex task whose constituents are the individual tasks represented by the colors of each square.

CT is parametrized in ways that allow for increasing complexity along various dimensions that influence the performance of different strategies and

---

[4]A "me" icon on the main board panel is used to enable a human player to easily distinguish between the player's own icon and those of other players.

[5]Note that the panel is displayed from the point of view of the "sun" player, and therefore a "me" icon is displayed as the recipient of the offer.

algorithms for decision making. It allows for specification of different reward structures, enabling examination of such trade-offs as the one between the performance of the group as a whole and the outcome of an individual. It also allows to examine the cost-benefits of collaboration-supporting actions.

Game parameters such as the number of players, the size of the board, and the number of chips may be set to vary the task complexity. The amount of information available to agents can be controlled by varying the information revealed on the board and about players' chips. Uncertainty can be introduced by changing features on the board and chips during play.

Two kinds of inter-dependencies among players can be varied in CT: task and score dependence. A *task dependence* arises whenever players lack the chips they need to reach their goals and must depend on other players to supply those chips. A *score dependence* arises when players' scores depend on each other's performance. The degree to which a player's score depends on the performance of other players may be used to distinguish collaborative teamwork from situations in which group members act independently. For example, a fully cooperative setting can be modeled in CT by setting the score for each player as the sum of all players' scores.

To illustrate the task analogy, we present an example of the way CT corresponds to the rescue domain described in Section 1. Players in CT correspond to fire-engine dispatchers. There are different ways to assign players to teams, such as by representing a dispatcher's affiliation with other dispatchers in their geographical vicinity by designating a shared goal square. A goal square might represent a mission to accomplish, such as rescuing people from areas afflicted with fire and smoke or assisting other rescue teams in different locations. Paths on the board represent the completion of tasks, such as clearing safe passage in a building or bringing special equipment to aid in the rescue. Chips represent resources, such as fire-engines, ladders and personnel. Negotiation over these resources by the dispatchers is needed to have an efficient deployment of resources. Players' scores in this game may be set to depend solely on their individual performance as dispatchers or may include, at some level, the score of their teammates.

2.2. Related Work: Decision-Making Environments

A variety of simulation systems have been developed to enable evaluation and comparison between computational strategies for decision-making in particular domains. Prime examples of such test-beds are (1) RoboCupRescue, which simulates first-response strategies that integrate disaster information,

planning, and human interfaces [30]; (2) the Trading Agent Competition (TAC), that facilitates the study of bidding strategies for the procurement of resources as well as supply-chain management [1], (3) the Agent Reputation Test-bed (ART), for studying reputation and trust in environments of varying agent-capabilities and reliability [14], and (4) the Electric Elves system for immersing agent technology in scheduling tasks in organizations [43].

CT is distinguished from these systems in several ways. First, these test-beds require that significant domain knowledge be represented in computer agents (e.g., modeling stock prices prior to bidding, estimating the number of people in danger in a burning building). In contrast, CT allows agent designers to focus on such general properties of interactions between humans and computers such as people's perception of the usefulness of an interruption request for information [26] and the way people form teams to carry out tasks in strategic environments [50]. This abstraction also has advantages for experiments that involve people: First, human subjects need not be domain experts. Second, CT enables the specification of different reward structures, allowing system-designers to vary the importance of different decision-making factors such as the performance of others or the group as a whole to the outcome of an individual. As a result, CT is novel in addressing the need to have better ways of evaluating computer agent strategies in contexts in which systems are participants in group activities that include people.

Although CT abstracts away from such domain specific details as fire engines and evacuation routes, it provides a task-like context in which decisions are made. This task context means that decisions are presented less abstractly than payoff matrices or trees, the canonical forms used to present outcomes in behavioral economic experiments. These canonical forms explicitly specify the payoffs to all players for each potential strategy. Real-world decision-making seldom presents choices this starkly. CT immerses people in an environment in which underlying relationships among tasks, goals and resources matter. It thus places decision-making in a more real context.

3. The Effects of Decision-Presentation on Human Negotiation Behavior

This section provides empirical evidence that people's behavior is significantly more cooperative when they negotiate using CT than when they are presented with payoff-only representations of the sort common to studies in behavioral economics. Recent work in the social sciences has demonstrated

that people cooperate despite the lack of direct short-term benefits from co-operative behavior [11, 39]. This section shows the use of CT induces such cooperative behavior and thus it provides support for CT as being the right kind of test-bed with which investigate decision-making in open-mixed net-works.

3.1. Empirical Design

We presented subjects with identical multi-agent decision-making problems in two conditions and measured their behavior and outcomes. In one condition, called the "task condition", people made decisions in the context of a CT game. In the other, the "table condition", they were given the decision in the context of a payoff-only table representation.

For the task condition, we used a 2-player CT setting that varied a configuration of 4x4 boards, the chip allocations for each player, and the placement of the goal square. The study comprised a one-shot negotiation setting for which one player was designated the proposer and the other was designated the responder. Players had full view of the board and each others' chips. The proposer player could make of offer to exchange a particular subset of its chips for some subset of the responder's chips. The responder player could accept or reject the proposer's offer. If no offer was made (there was a 3-minute deadline for proposers to make offers), or if the offer was declined, then both players were left with their initial allocation of chips. If the offer was accepted, the chip exchanges were enforced by the game controller. Following this interaction, both players' icons were automatically moved as close as possible to the goal square given the chips in their possession and their computed score.

The scoring function for players depended solely on their own performance: 100 points bonus for reaching the goal (otherwise, 50 points bonus); 5 points for each chip left in a player's possession at the end of the game; 10 points deducted for any square in the shortest path between a player's final position in the game and the goal-square. These parameters were chosen so that getting to the goal was by far the most important component, but if a player could not get to the goal it was preferable to get as close to the goal as possible, rather than hoard chips.

An example of one of the boards used in the study is given in Figure 1a. In this example, neither player can get to the goal by using the original chip allocation. The games used in the study all involved at least one of the players having insufficient chips to get to the goal, but not necessarily both.

You are a proposer.

Proposal Phase. Time left: 162

Please choose the default outcome or one of the alternative offers in the table. If you do not choose an offer by the time limit, the result will be the default outcome.

Default outcome:

	You	Other
○	180	65

Possible offers:

Sorted by
● Original Order
○ Your Benefit Descending
○ Other Benefit Descending

	You	Other		You	Other		You	Other		You	Other		You	Other		You	Other		You	Other		You	Other
○	180	170	○	190	55	○	75	170	○	55	190	○	200	40	○	180	170	○	75	170	○	190	55
○	200	40	○	190	55	○	55	190	○	190	55	○	65	180	○	65	180	○	75	170	○	45	200
○	200	45	○	190	50	○	65	180	○	85	160	○	190	160	○	180	170	○	75	170	○	55	190
○	210	35	○	170	180	○	65	180	○	190	160	○	55	190	○	180	170	○	55	190	○	170	75
○	160	190	○	190	55	○	45	200	○	170	75	○	85	160	○	180	170	○	45	200	○	190	55
○	55	190	○	75	170	○	200	45	○	190	55	○	65	180	○	200	45	○	170	180	○	65	180
○	55	190	○	190	160	○	210	30	○	180	170	○	210	30	○	190	55	○	200	150	○	210	35
○	190	55	○	190	50	○	200	150	○	200	45	○	160	190	○	200	45	○	65	180	○	210	35
○	45	200	○	65	180	○	95	150	○	220	20	○	170	180	○	190	55	○	170	75	○	65	180
○	65	180	○	75	170	○	170	75	○	85	160	○	75	170	○	75	170	○	190	55	○	170	180
○	200	45	○	150	200	○	210	30	○	55	190	○	85	160	○	65	180	○	170	75	○	160	190
○	65	180	○	170	180	○	170	180	○	180	170	○	190	160	○	190	55	○	160	190	○	190	55
○	200	45	○	75	170	○	45	200	○	170	180	○	200	45	○	170	180	○	35	210			
○	170	75	○	180	60	○	190	55	○	65	180	○	170	180	○	200	40	○	190	160			
○	75	170	○	190	50	○	160	190	○	200	40	○	85	160	○	200	45	○	180	170			
○	180	170	○	180	170	○	75	170	○	200	45	○	160	85	○	190	160	○	180	170			
○	55	190	○	190	55	○	170	180	○	65	180	○	55	190	○	190	160	○	55	190			

Propose Selected Outcome

Figure 2: Snapshot of an interaction in the table condition. Each entry in the matrix lists a pair of payoffs representing the score to the proposer player (left) and the responder player (right).

The table condition consisted of a payoff matrix representing potential offers that could be selected by the proposer player in a CT game. Each offer was represented as a pair of payoffs for the proposer and the responder players. Figure 2 shows a snapshot of a game in this representation as seen from the point of view of a proposer player. Each entry in the table represents an offer, and selecting one of the entries corresponds to choosing the offer that was associated with the payoffs inside the entry. One of the entries contained the no-negotiation alternative score and was presented to players as the default outcome of the interaction.

The score that each player received if no offer was made was identical to the score each player received if the offer was rejected by the responder. We refer to this score as the "no-negotiation alternative" score and refer to the score that each player received for an offer that was accepted by the responder as the "proposed outcome" score.

A total of 32 subjects participated in the experiment. They were equally divided between the two conditions. Participants in the task condition interacted with each other using the CT environment, whereas those in the

table condition interacted with each other using the payoff matrix representation. In both conditions, participants were compensated in a manner that depended solely on their individual scores, aggregated over all rounds of interaction. To prevent confounding effects on their behavior, participants were randomly divided into the two condition groups and seated in two rooms, such that no participant interacted with another participant seated in the same room. Participants only interacted with others in their condition group and were not provided any information about each other. We trained each group of subjects and tested their proficiency using a pre-study questionnaire.[6]

For each CT round that was played in the task condition, an equivalent round was played in the table condition. Each entry in the table listed a payoff pair representing the payoffs to the proposer and responder player for a possible exchange. For example, the payoff matrix shown in Figure 2 is equivalent to the CT game shown in Figure 1(a).

We use the term "table proposers" and "task proposers" to refer to the participants that were designated with the proposer role in the table and task condition respectively (and similarly for responder players). We use the term "offer benefit" to refer to the difference between the proposed outcome for an offer and the no-negotiation alternative score for the game. We measured people's behavior in the experiment using two features. The degree to which proposers were *helpful* to responders was measured in terms of the average offer benefit they proposed to responders. Similarly, the degree to which proposers were *selfish* was measured in terms of the average offer benefit they proposed to themselves. These features are not independent. For example, proposers can exhibit both a degree of selfishness and a degree of helpfulness based on the average benefit of their offers.

3.2. Analysis of Proposer Behavior

Table 1 presents the average offer benefit to participants in the task and table conditions for proposers and responders. Table proposers offered significantly more benefit to themselves than did task proposers (t-test $p < 0.05$). Also, they offered significantly less benefit to table responders than task proposers offered to task responders (t-test $p < 0.01$). Thus, proposers were more likely to engage in helpful behavior, in the task setting, and less likely to engage in selfish behavior.

---

[6]Subjects that did not score full points on the questionnaire were given a show-up fee and did not participate in the experiment.

Table 1: Average Benefit of Exchange

	Offer Benefit to	
	Proposer	Responder
Task	82.3	**47.6**
Table	**98**	36

Another result arising from the table is that proposers offered significantly more to themselves than they did to responders in both conditions (t-test $p < 0.05$). However, the difference between the benefit for proposer and responders was significantly larger in the table condition than in the task condition (t-test $p < 0.05$). We can thus conclude that although proposers were competitive in both conditions, they were significantly more competitive in the table condition than in the table condition.

Table 2: Average Benefit for Accepted Exchanges

	Proposer	Responder	Total
Task	79.5	**56.4**	135.9
Table	**85.6**	40.7	126.3

Table 2 shows the exchange benefit to proposers and responders averaged over all accepted proposals, as well as the total accumulated benefit in each condition. The benefit to responders from accepted proposals was significantly higher in the task condition than in the table condition, and conversely for the proposers (t-test $p < 0.05$). Thus, task responders outperformed table responders, whereas table proposers outperformed task proposers.

Our results also found that the CT task setting had a positive effect on the combined performance of participants. As the rightmost column shows, the total performance (combined proposers and responders scores) was higher in the task condition than in the table condition (t-test $p < 0.1$). The benefit for accepted exchanges is a measurement of performance, because the outcome of each round of interaction was fully determined by the action of the responder. Although this result was not significant at the $p < 0.05$ confidence level, it highlights that CT did not only have an effect on people's helpful behavior tendencies, but also that this helpful behavior was beneficial overall for the participants.

3.3. Comparison with Nash-Equilibrium Strategies

This section compares the offers that were made in the CT and table conditions with the offers dictated by the exchange corresponding to the Nash equilibrium strategy. We use the term *NE exchange* of a round to refer to the exchange prescribed by the Nash equilibrium strategy profile for the round. This exchange offers the maximum benefit for the proposer, out of the set of all of the exchanges that offer non-negative benefits to the responder. For the CT scenario used in our experiments, the Nash equilibrium offer is the one that maximizes the proposer's benefit out of the set of all possible exchanges that offered any benefit, however small, to the responder. Thus, the NE exchange is a competitive offer, which is more beneficial to proposers than to responders.

Because proposers were more competitive in the table setting, we expected table proposers to be more likely to offer NE exchanges than task proposers. We found that the number of NE offers made in the table condition (57) was significantly greater than the number of NE offers made in the task condition (13), (chi-square $t < 0.01$). To compare the extent to which the exchanges made by proposers in the two conditions differed from the NE exchange, we plotted the average benefit offered by NE exchanges and by proposed exchanges for both task and table conditions, as shown in Figure 3.
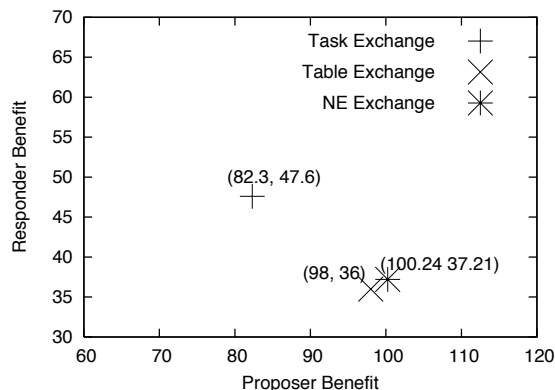
Figure 3: Benefit from Proposed Exchanges vs. NE Exchanges

The difference between the average benefit to responders from the NE offer and the average proposed exchange was close to zero in the table condition, and large and positive in the task condition (t-test $p < 0.05$). Similarly, the difference between the benefit to proposers from the NE offer and the

average proposed exchange was close to zero in the table condition, and large and negative in the task condition (t-test $p < 0.05$). The Euclidean distance between the points representing the NE benefit and the proposed exchange was significantly larger in the task condition than in the table condition. In fact, there was no statistically significant difference between proposed exchanges in the table condition and NE offers; participants who make decisions in the table condition were more likely to follow the equilibrium choice.

These results align with recent work in the social sciences that has found differences between people's behavior in field studies and their behavior in controlled laboratory experiments [21]. However, they differ from classical findings which show that people do not adhere to traditional game theoretic equilibria in one-shot games such as prisoners' dilemma [45]. One possible explanation for the adherence to game theoretic exchanges in the table condition may be the fact that subjects were guaranteed the no-negotiation alternative score in each game (unlike the canonical behavioral economics experiments, in which the outcome for no negotiation was zero). The guarantee of a positive payoff to proposers could have made them less likely to fear rejection by the responders. Thus they offered exchanges that were more selfish (more game-theoretic) as compared to prior studies.

3.4. The Context Hypothesis and Related Approaches

The results of the experiment described above support the hypothesis that the way in which a decision is represented influences people's negotiation behavior. While considerable amount of research in the social sciences has studied the effects of framing on behavior this is the first work that has explored the particular role played by task settings. In particular, the result that people are more helpful and less competitive in the task-like CT setting than when deciding in a payoff-only context, suggests that task settings have a positive effect on the way people (and in turn, the agents that interact with people) behave. However, there are several alternative explanations of this behavior. In this section, we rule out two obvious alternatives.

First, the competitive behavior exhibited by table proposers might be attributable to the fact that the specific payoffs for each possible strategy was explicitly presented to them, whereas proposers in the CT-task setting were not given the payoffs explicitly. To rule out this hypothesis, we ran an additional experiment in the CT game context. In this experiment, we used a CT setting like the one used in the original task context, with one difference: Subjects were able to access an extended version of the path-finder panel

that allowed them to view the payoffs for potential paths to the goal. This setting preserved the task environment, but also allowed for explicit display of the payoffs for both players associated with agreements. We used the same set of games that were used in the original experiment. The results showed that there was no significant difference in the average benefit allocated to proposers and responders in this intermediate representation than in the CT-task condition (t-test $t(29) = 2.34, p < 0.05$). Thus, we can preclude the lack of explicitly presented payoff information in the CT context as an explanation of the different proposers' behavior.

Second, the different proposer behaviors might be attributable to differences in the cognitive demand on subjects making decisions in two different settings. In both of these settings, the number of possible offers for proposer players in each game was large, and bounded by $2^8 = 256$. Typically, there were between thirty and forty offers in each game associated with a distinct benefit to both proposer and responder players. All of these strategies were explicitly presented to table proposers, while the task proposers needed to infer the possible strategies (and payoffs) by reasoning about the task setting. Both conditions potentially required large cognitive effort that could confound players' reasoning about the degree of helpful behavior to exhibit. For the original set of experiments, we provided decision support tools for each decision representation setting, that were designed to mitigate these potential cognitive load problems. In the CT game, subjects could use the Decision Aid panel, shown in Figure 1c, to query the system for suggestions about the best paths to take given any hypothetical chip distribution and the chips required for these paths. In the table condition, subjects were able to sort the table by their own benefit or by benefit to the other player. Thus, in both cases, we reduced the effects of the decision-making complexity on the behavior exhibited by subjects.

Thus, the task context provided by CT remains the best explanation of the more helpful, less competitive behavior exhibited by subjects in that condition. Although we have not yet established how many of the elements essential for studying decision-making trade-offs of actual open, mixed network application settings are reflected in the decision-making environment provided by CT, these results show that the CT environment differs for non-trivial reasons from payoff-only settings and in ways that make decision-making in CT closer to the types of task settings that occur in the real world.

Lastly, we note that this study is distinguished from work on collabora-

tive user interfaces and communication protocols for facilitating tasks such as preference acquisition in e-commerce [35, 25], turn-taking [10] and diagram generation [5]. Many aspects of interfaces affect the way people interact; In designing both the task interface and the table interface we attempted to create the most natural representations while restricting the modalities to a graphic display. The fact that our results did not change even when the payoffs were available to subjects using CT implies that the difference in behavior should not be attributed to the payoffs themselves, but to other aspects relating to the context that CT provides, such as the explicit presentation of tasks, goals and resources. We do not mean to imply that the difference in behavior can be attributed solely to one aspect or the other, such as the communication protocol.

4. Learning Social Preferences in Negotiation

This section describes an investigation of the hypothesis that computer agents that model and learn the influence of social factors on people's negotiation strategies will outperform agents using traditional equilibrium strategies. It defines three particular social factors—aggregate benefit, advantage of outcome, advantage of trade—which together with individual benefit (the sole characteristic taken into account by traditional equilibrium strategies) are used to characterize an offer. It then describes an algorithm for learning these factors from data of people's performance in a 2-player CT setting that highlights the task-resource relationship, and presents the results of empirical studies comparing this learned model to equilibrium strategy approaches. The CT setting used in this study was identical to the one described in Section 3.

4.1. A Model for Learning Social Preferences

This section describes a model that can be used to enable a computer agent playing the role of the proposer to make offers that take into account social factors that influence people's decisions when they are responders in this CT scenario. The model addresses three essential challenges. First, people vary in the extent to which they are affected by social factors when they negotiate. For example, an offer that is rejected by a competitive responder might be accepted by a responder who is more altruistic. In addition, people sometimes make mistakes, so they may not behave consistently and thus may at times deviate from the model of their typical behavior. Third,

people's behavior has been shown to depend on their inter-dependence. For example, the extent of generosity reflected in a proposer's offer may depend on whether the proposer needs chips from the responder to get to the goal.

The formal decision-making model used in this study is defined as follows: Let k represent a CT game, associated with a set $\mathbf{C}^k$ of possible proposals. For each proposal $c_j^k \in \mathbf{C}^k$, let NN_P^k and NN_R^k denote the no-negotiation alternative scores for the proposer and responder. These are the scores that the proposer and responder would receive if no agreement is reached in the game, and the players use the chips they were allocated at the onset of the game to move towards the goal-square. Let $PO_P^{k,j}$ and $PO_R^{k,j}$ denote the proposed outcome scores for the proposer and responder for proposal c_j^k.

For each proposal c_j^k of game k, we represent four social factors $\mathbf{x}_j^k = \{x_{j,1}^k, \ldots, x_{j,4}^k\}$ that affect the behavior of the responder agent.

- Individual benefit $\quad x_{j,1}^k = PO_R^{k,j} - NN_R^k$

- Aggregate benefit $\quad x_{j,2}^k = (PO_P^{k,j} + PO_R^{k,j}) - (NN_P^k + NN_R^k)$

- Advantage of outcome $\quad x_{j,3}^k = PO_R^{k,j} - PO_P^{k,j}$

- Advantage of trade $\quad x_{j,4}^k = (PO_R^{k,j} - NN_P^k) - (PO_R^{k,j} - NN_P^k)$

The individual benefit of an offer measures the extent to which the offer is beneficial to the responder. The aggregate benefit of an offer measures the extent to which the offer is beneficial to both players. The other two social factors represent different potential sources of inequality. The advantage of outcome for an offer measures inequality in the proposed outcome scores for the offer, without reference to the no-negotiation alternative scores. The advantage of trade for an offer measures inequality in the benefits to the agents of the offer. We illustrate these factors using the offer in the game shown in Figure 1, in which the "me" player (the proposer) offers one cyan chip to the "sun" player (the responder) in return for one purple chip. This exchange provides both players with the chips they need to get to the goal.[7]

---

[7]The exchange provides a higher individual benefit to the "sun" player than to the "me" player, according to the scoring function. This is because the "sun" player is initially located closer to the goal square than the "me" player, and has the same number of four chips at the onset of the game. Therefore it will be left with more chips than the "me" player if it accepts the exchange and moves towards the goal.

It would allow the "sun" player to choose a shortest path of length two to the goal. This player would have two chips left in its possession, and earn a score of 120 points. Similarly, the offer would allow the "me" player to choose a shortest path of length three to the goal (the path that is outlined in Figure 1a). This player would have one chip left in its possession and earn a score of 110 points. The no-negotiation alternative score for the "me" player in the game is 40 points, while the no-negotiation alternative score for the "sun" player in the game is 50 points. Therefore the exchange above provides an individual benefit to the "sun" player of 70 points, an aggregate benefit of (70+60)=130 points, and an advantage of outcome and of trade of 10 points.

To enable learning of people's decision-making strategies in this CT game, we introduce the notion of "responder types" and use them to represent ways that responders may differ in the extent to which the various social factors affect their decision-making. Each type represents a particular weighting of these factors, reflecting the different ways in which responders make decisions. There is a finite set of responder types $\mathbf{T}$, and a prior probability distribution over types, denoted by $P(\mathbf{T})$. The weight $w_{i,l}$ denotes the weighting of social preference x_l for responder type t_i. These weights measure the relative importance for the responder of each of the social preferences.

Let k denote a CT game instance. Given a proposal c_j^k, possible social preferences $\mathbf{x}_j^k$ and responder type t_i that is selected from $P(\mathbf{T})$, we define a social utility u_i for the responder as a weighted sum

$$u_i(\mathbf{x}_j^k) = \sum_{l=1}^{4} w_{i,l} \cdot x_{j,l}^k$$

Let r^k denote the response to exchange c_j^k. A responder that always follows its social utility function u_i would agree to any exchange c_j^k such that $u_i(\mathbf{x}_j^k) > 0$. The probability of acceptance of an exchange c_j^k at game k by a responder of type t_i is defined by the following sigmoid function

$$P(r^k = \text{accept} \mid \mathbf{x}_j^k, t_i) = \frac{1}{1 + e^{-u_i(\mathbf{x}_j^k)}} \tag{1}$$

There are two advantages to using this function to model the probability of acceptance using this function. First, it captures the fact that people may make mistakes with respect to their utility function. Accepting a proposal is more likely when the utility u_i is a large positive number than a small positive

number. In particular, the probability of acceptance converges to one as the utility becomes large and positive, and to zero as the utility becomes large and negative. When the utility is close to zero, the decision is less clear-cut, and the probability of acceptance is close to 0.5, meaning that mistakes, or decisions that are inconsistent with the utility function are more likely to be made by the responder. Second, the mathematical properties of the sigmoid function facilitate the derivation of the learning rules, as we will soon describe. The expected utility to the proposer for an exchange c_j^k of type t_i can be written as

$$
\begin{aligned}
EU_P(c_j^k \mid t_i) = P(r^k = \text{accept} \mid \mathbf{x}_j^k, t_i) \cdot PO_P^{k,j} + \\
(1 - P(r^k = \text{accept} \mid \mathbf{x}_j^k, t_i) \cdot NN_P^k)
\end{aligned}
\tag{2}
$$

The proposer does not get to observe the type of the responder it is interacting with at each game. As a result, the best-response strategy for the proposer, as specified by Equation 3, will equal the unique proposal that maximizes the proposer's expected score in the game for all the different types of possible responders.

$$
c_*^k = \text{argmax}_{c_j^k \in \mathbf{C}} \sum_{t_i \in \mathbf{T}} P(t_i) \cdot EU_P(c_j^k \mid t_i)
\tag{3}
$$

4.2. Learning a Mixture Model

To use Equation 3 in the model for a proposer agent's offer strategy, it is necessary to learn the distributions over the type space of responders and the likelihood that each type agrees to a particular offer. Let $\mathbf{D} = \{d^{(1)}, \ldots, d^{(N)}\}$ denote a set of instances, each instance representing a single CT game. The proposal c_j^k in CT game k comprises a (possibly empty) subset of the chips of the proposer to be exchanged with some (possibly empty) subset of the chips of the responder. We define a data instance $d^{(k)}$ to consist of the following features: a CT game k, a set $\mathbf{C}^k$ of possible offers in game k, the chosen proposal $c_*^k \in \mathbf{C}^k$, the set of social preferences $\mathbf{x}_*^k = \{x_{*,1}^k, \ldots, x_{*,4}^k\}$ associated with the chosen proposal, and the response r^k.

The goal of the learning task is to use the data set $\mathbf{D}$ to estimate the distribution over the responder types and the strategy of the responder for each type. As specified by Equation 1, the strategy for the responder depends on the social preference weights $w_1, \ldots, w_4$. The likelihood of the response at game k is defined as $P(r^k \mid \mathbf{x}_*^k, t_i)$. Let err_i^k denote an error term for

predicting whether the proposal c_*^k at game k is accepted by responder type t_i, defined as follows:

$$err_i^k = 1 - P(r^k \mid \mathbf{x}_*^k, t_i) \tag{4}$$

This term comes from assuming that if the model was a perfect predictor of the responder, it would have predicted the true response r^k with probability 1. The difference between 1 and the $P(r^k \mid \mathbf{x}_*^k, t_i)$ factor is the degree of error for the model, given that the responder is using block t_i.

We can minimize err_i^k by taking the derivative of this function with respect to each weight w_j. Employing a gradient descent technique, we get the following update rule for each weight w_j, where α is the learning rate.

$$w_j = w_j + \alpha \sum_{k \in \mathbf{D}} x_j^k \big(1 - P(r^k \mid x_*^k t_i)\big). \tag{5}$$

However, the responder type t_i in a CT game is unknown. Therefore we define a mixture model that is comprised of a probability distribution $P(\mathbf{T})$ over the type space, and an associated set of weights $w_{i,1}, \ldots, w_{i,4}$ for each type t_i. We sum over the type space to compute err^k, the expected error of the model at each instance k for all types

$$err^k = \sum_{t_i \in \mathbf{T}} P(t_i \mid r^k, \mathbf{x}_*^k) \cdot err_i^k \tag{6}$$

where $p(t_i \mid r^k, \mathbf{x}_*^k)$ is the posterior probability of type t_i given response r^k and the set of social preferences $\mathbf{x}_i^k$. This can be computed using Bayes rule. The degree to which a training instance k can be used to learn the weights $w_{i,1}, \ldots, w_{i,4}$ is proportional to the probability that each type t_i generated instance k. We get the following update rule for weight $w_{i,j}$ of type t_i.

$$w_{i,j} = w_{i,j} + \alpha \cdot P(t_i \mid r^k, \mathbf{x}_*^k) \sum_{k \in \mathbf{D}} x_j^k \big(1 - P(r^k \mid x_*^k t_i)\big) \tag{7}$$

Computing the new values of the $P(\mathbf{T})$ parameters is performed using an on-line version of the EM algorithm [38], a standard stochastic learning paradigm for estimating distributions in the presence of missing information. We compute the expected sufficient statistics for each type t_i using the current parameter settings and normalize to get a new distribution over the type space.

$$P(t) = \frac{1}{Z} \sum_{k \in D} P(t_i \mid r^k, \mathbf{x}_*^k)$$

where Z is a normalizing factor.

4.3. Empirical Study

In this section we compare the performance of an agent that used a social preference model to negotiate with people to agents using traditional equilibrium strategies. A total of 42 subjects participated in the experiment, 32 in the data collection phase and ten in the evaluation phase.

The CT games we used were randomly generated, but filtered to meet the following conditions: (1) at least one of the players could reach the goal after trading with the other player; (2) it was not the case that both players could reach the goal without trading. In this way, the games we trained and tested on characterized a wide range of task dependency relationships between players. For each board we used in the study, we recorded the board and chip settings, as well as the actions of both agents. We ran two instances of the data collection phase, each with different subjects, collecting 192 games in total. The data obtained in this phase was then used to learn a model of human play.

We learned separate models for one, two and three possible types of responders, referred to as $M1$, $M2$ and $M3$, respectively. The parameters in these models represented different convergence points of the learning algorithm in the parameter space of possible values for the social preferences. For all models, we used random initial values for the distribution over responder types. For $M1$ we also set random values for all of the social preference weights. For $M2$ and $M3$ we assigned each responder type initial values that highlighted certain social preferences by giving them significantly higher initial value than others. Specifically, in $M2$, one of the responder types highlighted advantage-of-outcome and advantage-of-trade, while the other highlighted aggregate benefit. In $M3$, responder types highlighted advantage-of-outcome, aggregate benefit, and advantage-of-trade separately. We ran each model on the data from the data collection phase.

To keep from over-fitting the data, we stopped the learning process after no improvement had been recorded on a held-out validation set. This occurred after about 20 epochs. We obtained the following posterior parameter values for each model. Model $M1$, which had a single responder type, learned social preference weights $(7.00, 5.42, 0.40, 4.00)$ for individual benefit, aggregate benefit, advantage of outcome and advantage of trade, respectively. This model thus describes a responder who cares about both players' outcome, but also likes to do better than the proposer. In $M2$, the distribution over responder types was $(0.36, 0.63)$ and the social preference weights were $(3.00, 5.13, 4.61, 0.46)$ and $(3.13, 4.95, 0.47, 3.30)$ for each type

respectively. This model thus describes two partially altruistic responders, both of whom have high weights for social welfare, while still caring about their own benefit. One of the types cares more about advantage of outcome, and the other type cares more about advantage of trade. In $M3$, the distribution over responder types assigned minuscule probability for the third type, and resembled $M2$ in all other parameter values. We therefore decided to use $M2$ in the evaluation phase.

Ten people participated as subjects in the evaluation study, which compared the performance of three different computer agents. The computer agents were always in the proposer role. Each of these agents was capable of mapping any CT game position to some proposed exchange. The agent using the social preference model, denoted SP, proposed the exchange matching the best-response strategy of Equation 3. The second agent, denoted NE, proposed the exchange corresponding to the Nash equilibrium strategy for the proposer role. A proposer that uses a Nash equilibrium approach will make the offer that maximizes the proposer's benefit out of the set of all possible exchanges that offer non-negative benefit to the responder.

$$c_*^k = \text{argmax}_{c_j | (PO_R^{k,j} - NN_R^k) > 0} (PO_P^{k,j} - NN_P^k)$$

The third agent, denoted NB, proposed the exchange corresponding to the Nash bargaining strategy [37] for the proposer.[8]

The subjects were divided into two groups with the members of each group playing four rounds in a round-robin fashion. At each round, four concurrent games were played by each group. One of the (human) subjects and the three computer agents were designated as proposers; the other four (human) subjects were designated as responders. In each round, the game settings—including board layout, start and goal positions and initial chip distributions—were the same for all of the games played by members of the same group. Through the round-robin, each of the (human) subjects played each of the computer agents at least once as well as some other human player(s). No two people played each other more than once, but people could play the same computer agent more than once.

Participants were given the same instructions and tutorial as in the data collection experiment. To make the conditions in both collection and evalua-

---

[8]The Nash Bargaining solution maximizes the product of the agents' utilities and is always Pareto optimal.

Table 3: Total score achieved by agents

Model	Total Reward	Proposals Accepted	Proposals Declined	No Offers
SP	**2880**	**16**	5	0
NE	2100	13	8	0
NB	2400	14	2	5
HU	2440	**16**	1	4

tion phases identical, we did not tell subjects they would be playing computer agents as well as people.[9] We recorded the settings of each game, as well as players' proposals and responses. Altogether we collected 21 rounds, where each round consisted of four games with different proposers, as explained above.

4.4. Results

The evaluation criterion for the SP agent was the total accumulated reward received by the agent as compared to the agents that used other models to guide their play. We also evaluated whether the offers that the SP agent made were accepted more often by people than the offers made by other agents.

Table 3 presents the results of the evaluation phase for each of the agents used in the experiment. It lists the total monetary reward, the number of proposals accepted, the number of proposals rejected, and the number of times no offer was proposed.

The computer proposer labeled NE always proposed the exchange that corresponded to the proposer's strategy in the Nash equilibrium of each CT game. In essence, this resulted in offering the best exchange for the proposer, out of the set of all of the exchanges that are not worse off for the responder. Many of the exchanges proposed by this agent were declined. We hypothesize this is because they were not judged as fair by the responder. This result closely follows the findings of behavioral game theory. The performance of NE was the worst of the four.

---

[9]Approval was obtained from the Harvard Human Subjects Committee for this procedure.

The computer proposer labeled NB always proposed the exchange that corresponded to the proposer's strategy in the Nash Bargaining profile. In particular, this always corresponded to an exchange that was Pareto optimal with respect to the set of feasible allocations, that is, the proposed outcome was not worse off than the no-negotiation alternative for both agents. This exchange consistently offered more to the responder than the NE agent did for the same game, when the board and chip distribution enabled it to do so. Because NB tended to offer quite favorable deals to the responder, they were accepted more than the other computer agents, provided that an offer was made. This player did not make any offer for those games in which there was no Pareto optimal offer.

The results in the row labeled HU in the table combine the monetary rewards for all of the human agents. The identity of the human proposer was different at each round. With one exception, human offers were accepted whenever they were made.

The computer proposer that followed our expected utility model, labeled SP, achieved a significantly higher reward than the NE, NB and HU (t-test comparison for mean reward was $p < .05, p < .05, p < .1$, respectively). SP and HU had the highest number of accepted proposals. Interestingly, the SP agent proposed the same offer as the human proposer in 4 of the games, whereas the Nash equilibrium agent did not match a human proposal in any game, and the Nash bargaining agent matched human proposals in 2 games.

The distinguishing behavior of the SP agent is illuminated by the examples in Tables 4 and 5 in which the NNA heading gives the no-negotiation alternative scores. Table 4 presents a round in which SP proposed an exchange which was accepted; the (human) responder was altruistic in this case. This example is an interesting illustration of generalization in the learning systems. There was only one observation in the learning phase in which a responder altruistically agreed to such an exchange. In the evaluation, the SP agent proposed exchanges that were asking such a favor from a respondent who was much better off four times, and these offers were consistently accepted by the (human) subjects. The ability of the SP agent to make the right kind of trade-offs — when to ask for favors and when not to ask for favors from the responder — was a contributing factor to its success.

Table 5 displays an example in which the proposed outcome of the exchange proposed by NE, while beneficial for the responder, was lower than the exchange proposed by SP. The NE exchange was rejected, while the SP exchange was accepted. This difference seems to indicate that the re-

Table 4: An example of a proposal made by the SP agent

Model	Proposer Score	Responder Score
NNA	45	170
SP	70	150

Table 5: An example of the offers made by all agents in a single game

Model	Proposer Score	Responder Score	Accepted
NNA	75	150	
SP	170	170	Yes
NE	180	160	No
NB	150	190	Yes
HU	170	170	Yes

sponders in this game cared about the equality of outcomes. Note that in this exchange, the SP exchange and the exchange proposed by the human were equal.

4.5. Related Work: Learning from People

The results reported in this section relate to several different strands of prior research that address, in a variety of contexts, the importance of learning for agents working with people. Past works in AI have used heuristics, equilibrium strategies, and opponent modeling approaches toward building computer agents that negotiate with people. For a recent comprehensive review, see Lin and Kraus [33]. Within repeated negotiation scenarios, Kraus et al. [32] modeled human bilateral negotiations in a simulated diplomatic crisis characterized by time constraints and deadlines in settings of complete information. They adapted equilibrium strategies to people's behavior using simple heuristics, such as considering certain non-optimal actions. Jonker et al. [24] designed computer strategies that involve the use of concession strategies to avoid impasses in the negotiation. Byde et al. [6] constructed

agents that bargain with people in a market setting by modeling the likelihood of acceptance of a deal and allowing agents to renege on their offers. Traum et al. [49] constructed agents for the training of individuals to develop leadership qualities and interviewing capabilities. Recent approaches have used learning techniques to model the extent to which people exhibit different social preferences when they accept offers in multiple interaction scenarios [40, 34].

Our work is also related to approaches for elicitation of users' preferences for automated negotiation [35, 9]. In these works, people were asked to disclose information about their preferences over various negotiation deals, and evaluate various acquisition models that trade-off the importance of different features of an agreement to learn a parsimonious representation of users' preferences. However, the participants of open, mixed networks directly negotiate in task settings and it is not possible to ask people about their preferences directly. Indeed, people are generally reluctant to reveal their goals and preferences when they negotiate in task settings [18]. Thus we took a different approach, namely to infer people's utility functions from their negotiation behavior.

Lastly, this work is also related prior work include adaptation to user preferences for improved human-computer interface design or computer-supported collaborative agents that infer users' goals and intentions [23]. These user models inform agent-design for applications such as office assistants [5], and scheduling and meeting management [43]. Our work differs from these in its focus on learning people's decision-making in strategic—rather than collaborative— settings, and adapting to the social factors that guide their play.

5. Repeated Negotiation in Settings of Full Information

This section describes a novel agent design for a mixed-network setting that extends the one described in the previous section to include multiple negotiation rounds and non-binding agreements. Such settings characterize many types of real-world scenarios, including diplomatic relations, trade agreements and contract negotiations, but have not been considered before in the design of agents that interact with people. These settings pose additional challenges to the design of effective computer negotiation strategies. In particular, when agreements are not enforceable, negotiators may initially behave reliably, but eventually take advantage of others that trust them.

27

The study described in this section focuses on negotiation between computer agents and humans under conditions of full information, while Section 6 focuses on negotiation solely between computer agents under conditions of incomplete information. Together, they demonstrate CT's flexibility in allowing to configure more complex negotiation protocols that support repeated negotiation and allow to control the extent to which participants abide to agreements. To date, *all work on human-computer negotiation assumes that agreements are binding*, and uses constrained negotiation protocols that do not allow people to reveal information or to argue their positions when they negotiate. An exception is an agent proposed by Kraus and Lehmann [31] that was developed for a specific domain, that of diplomacy, and which did not reason about behavioral traits of participants.

We hypothesized that to be successful in settings where agreements are non-binding, agents need to (1) model behavioral traits that affect the negotiation behavior of other participants; and (2) adapt its negotiation strategy to the way these traits change over time. To this end, we designed an agent that explicitly modeled the extent to which its negotiation partners were reliable and helpful. We begin the section by describing the CT setting, then outline the strategy used by the agent and provide an empirical evaluation of this strategy when negotiating with other people.

5.1. The CT Setting

The CT setting described in Section 3.1 was extended as follows: First, players were allowed to renege on their agreements in part or in full. This was done by introducing a transfer phase which immediately followed the negotiation phase, and required chips to be transferred manually. In contrast, in the CT setting described in Section 3.1 the chip transfer was automatic. Second, the negotiation protocol included a recurring sequence of negotiation interactions that allowed players to make and agree to offers, transfer chips and move about the board.

This study used two different board configurations. In all of these boards there was a single distinct path from each player's initial location to its goal square. One of the board configurations exhibited a symmetric dependency relationship between players: Neither player could reach the goal given its initial chip allocation, and there existed at least one exchange such that both players could reach the goal. We referred to players in this game as *task co-dependent*. The other board type exhibited an asymmetric dependency relationship between players: One of the players, referred to as *task*

28

independent, possessed the chips it needed to reach the goal, while the other player, referred to as *task dependent*, required chips from the task independent player to get to the goal. An example of the task co-dependent board is shown in Figure 4. In this game both "me" and "O" players are missing three chips to get to the goal. The relevant path from the point of view of the "me" player is outlined.

(a) Board Panel

(b) Chip Display Panel (showing the chips in the possession of both participants)
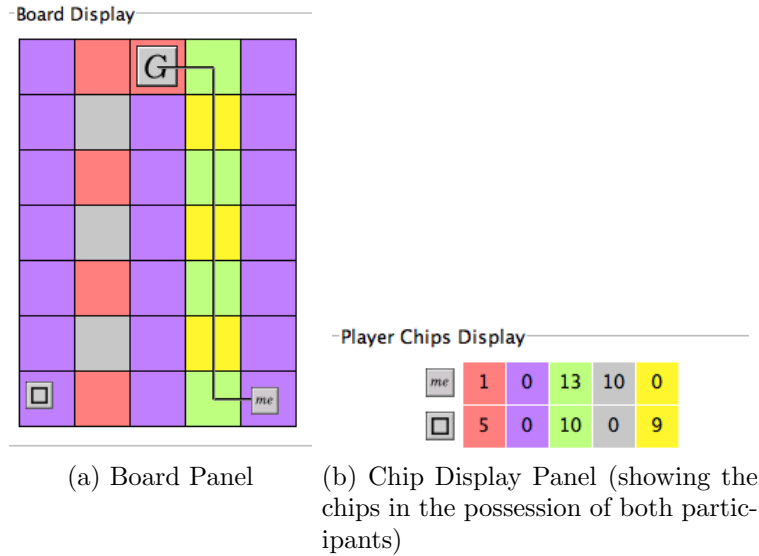
Figure 4: Snapshots of CT GUI

At the onset of the game, one of the players was randomly allocated the role of proposer, while the other was given the role of responder. The interaction proceeded in a recurring sequence of communication, transfer and movement phases. In the *communication phase* the proposer could make an offer to the responder, who could accept or reject the offer. In the *transfer phase*, both players could choose chips to transfer to the other player. The transfer action was done simultaneously, such that neither player could see what the other player transferred until the end of the phase. In particular, players were not required to fulfil agreements: A player could choose to transfer more chips than were agreed, or any subset of the chips that were agreed, including transferring no chips at all. In the *movement phase*, players could manually move their icons on the board across one square by surrendering a chip in the color of that square. At the end of the movement phase, a new communication phase began. The players alternated roles, such that

the previous proposer was designated as a responder, and vice versa.

These phases repeated until the game ended, which occurred when one of the following conditions held: (1) at least one of the participants reached the goal square; or, (2) at least one of the participants did not move for three consecutive movement phases. At this point, both participants were automatically moved as close as possible to the goal square using the chips in their possession and their score was computed as follows: 100 points bonus for getting to the goal square, 5 points bonus for any chip left in a player's possession; 10 points penalty for each square left in the path from a player's final possession and the goal square.

5.2. The Personality-Based Agent

The agent that we constructed for this setting, called the Personality Based (PURB) agent, modeled other participants in terms of two behavioral traits: helpfulness, and reliability.[10] The *helpfulness* measure of participants, denoted h, represented the extent to which they shared resources with their negotiation partners through initiating and agreeing to proposals. This was computed as the percentage of proposals in the game in which participants offered more chips to their negotiation partners than they requested for themselves. The *reliability* measure of participants, denoted r, represented the extent to which they kept commitments with their negotiation partners. This was computed as the ratio between the number of chips transferred by participants and the number of chips they actually offered, averaged over all proposals in the game. Together, we refer to the pair (h, r) as the *cooperativeness* measure of a participant.

5.3. Social Utility Function

The PURB agent used a social utility function to negotiate with people which was a weighted combination of the several features. For the remainder of this section we will use the term "agent" to refer to the PURB agent, and "person" to refer to its negotiation partner.

1. The expected future score for the PURB agent. This score was estimated using a heuristic function that estimated the benefit to i agent from a potential exchange. It depended on the probability that PURB will get to the goal given that proposal O is fulfilled at a state s which

---

[10]We use the term "participants" to refer to both people and computer agents.

comprises a board game, the positions of both participants on the board, and the chips in their possession. We denote this probability as $P(G \mid s, O)$, and define the expected future score to i as

$$(P(G \mid s, O) \cdot 100) + (1 - P(G \mid s, O) \cdot 10 \cdot d) + c \cdot 5$$

where 100 is the number of bonus points to get to the goal according to the CT scoring function; d is the Manhattan distance of the agent from its final position on the board and the goal square, given that the agreement was fulfilled the p; 10 is the number of penalty points for each square in the distance from the final position of PURB and the goal square; c is the number of chips left in the player's possession after it advances to the goal using the shortest possible path, and 5 is the number of points awarded to the player for each chips left in its possession at the end of the game. The probability $P(G \mid s, O)$ to get to the goal at state s given proposal O was estimated as the ratio between the number of chips that the other participant delivered to the PURB agent, and the number of chips that the PURB agent was missing to get to the goal at state s given that O was fulfilled.

2. The expected future score for the other participant (computed in the same way as for the PURB agent).

3. The cooperativeness of the other participant (in terms of helpfulness and reliability).

4. The *perceived* cooperativeness of the other participant. This feature represented the PURB agent's model of the other participant's beliefs about its own reliability and helpfulness.

The weights for the features of the social utility function were set by hand, and depended on the dependency relationships between participants as well as their cooperative measures. Generally, as the other participant increased its cooperativeness measures, the weighting in the social utility function for PURB that was associated with the score of the other participant were increased. This was to provide an incentive to the PURB agent to be more generous when its negotiation partner was cooperative. Each time an agreement was reached and transfers were made in the game, the PURB agent updated the helpfulness and reliability measures of both agents. Using this

social utility allows the PURB agent to vary its strategy based on its estimate of the other participant's cooperativeness measure. For example, if the reliability of the other participant was high, this would increase the social utility of actions that favour the other participant.[11]

5.4. Rules of Behavior

The second component of PURB's decision-making paradigm was a set of rules that narrowed the search space of possible actions to be considered by the agent when using its social utility. These rules depended on aspects relating to the state of the game (e.g., the number of chips each agent had, whether a participant can independently reach the goal). At each step of the game, the PURB agent used its social utility function to choose the best action out of the set of possible actions that were constrained by the rules. The rules were designed such that the PURB agent begins by acting reliably, and adapts over time to the individual measure of cooperativeness that is exhibited by its negotiation partner. These rules are based in part on a decision-making model designed to adapt to people's negotiation behavior in different cultures [19, to appear].

To enable to specify a finite set of rules for different measures of reliability and helpfulness, the possible values that these traits can take were divided into three equal intervals representing low, medium or high measures. For example, low reliability measures ranged from 0 to $\frac{1}{3}$. We then defined the cooperativeness of an agent to depend on the extent to which it was reliable and helpful. Specifically, we defined the cooperativeness of a participant to be *high* when it exhibited high helpfulness and high reliability measures, or high helpfulness and medium helpfulness measures; the cooperativeness measure of a participant was *medium* when it exhibited medium reliability and medium helpfulness measures, or medium reliability and high helpfulness measures; the cooperativeness measure of a participant was *low* when it exhibited low reliability measures (regardless of its helpfulness measure) or medium reliability measures and low helpfulness measure. These values were tuned by hand on games that were not considered in the evaluation.

---

[11]Although this strategy may resemble the principle behind the Tit-for-Tat paradigm, depending on PURB's model, its strategies can be more nuanced. For example, depending on the dependency relationships that hold in the game, the PURB agent may offer a generous exchange in response to a selfish offer made by a person.

We now list the set of rules used by the PURB agent in combination with its social utility function:

a) *Making Proposals* The PURB agent generated a subset of possible offers and non-deterministically chose any proposal out of the subset that provided a maximal benefit (within an epsilon interval) according to its social utility function.

Before outlining the rules by which the set of possible proposals were generated, we will introduce the following notation: We say an agent i is "stronger" than agent j if i is able to reach the goal independently of j, or if it requires less chips to reach the goal than j. Let $O_{i=j}$ represent the number of proposals in which agent i asks for as many chips as it receives; $O_{i>j}$ represents the set of proposals in which i asks for more chips than it receives; $O_{j>i}$ represents the set of proposals in which i asks for less chips than it receives.

Offers were generated by PURB in a way that considered which participant was stronger than the other. Let i denote the PURB agent and j denote the other participant. When participants were co-dependent, the set of possible offers i considered included those offers that favoured the stronger agent. If i was stronger than j, then the set $O_{i>j}$ was considered (i.e., i requested from j more chips than i proposed to j) And conversely for the case in which j was stronger than i. In both cases, the set $O_{i=j}$ was also generated and considered (i asks for as many chips as it sends).

In the other dependency roles, the offers that were generated depended on the cooperativeness measure of j:

1) When the cooperativeness of j was high or medium, then if i was stronger than j, then the set of possible offers that i considered included $O_{i>j}$. This is because that when the reliability of j was high, there was a higher likelihood that j would keep its commitments, and thus the set of possible exchanges for i included exchanges that were highly favorable to i. However, if j was stronger than i, then offers were chosen from the set $O_{j>i}$. This was because that i wished to minimize the chances that j would reject its offers given that j did not need i to get to the goal.

2) When the cooperativeness of j was low, then offers were chosen from the set $O_{i>j}$, regardless of which agent was stronger. This was because

33

i did not expect j to fulfil its agreements, and thus it proposed offers that were less beneficial to j.

b) *Accepting Proposals* As a responder, the PURB agent accepted an offer if it was more advantageous to it than the offer it would make as a proposer in the same game state, or if accepting the offer was necessary to prevent the game from terminating. To state this formally, let $u_i(O, \textsf{accept} \mid s)$ denote the social utility for i from an offer O made by j at state s. Let O' denote the offer that agent i would make at state s according to the rules in (a). Agent i accepted an offer O if $u_i(O, \textsf{accept} \mid s) \geq u_i(O', \textsf{accept} \mid s)$. In addition, i would accept any proposal that prevented the game from ending, which occurs when the following conditions hold: (1) the chips in the possession of agent i do not allow it to move on the board at state s; (2) the offer O_j allows agent i to move; and (3) if i rejects the offer, the limit for dormant turns will be reached and the game would end.

c) *Transferring Chips* These rules specify the extent to which the PURB agent fulfilled its agreements in the game. This behavior directly depended on its model of the other's reliability:

If the reliability of j was *low*, it was likely that the other participants would not fulfil its agreement. Therefore i did not send any of its promised chips. However, if both agents were task dependent, and the agreement resulted in *both* agents becoming task independent then i sent half of the promised chips with a given probability, because it was not certain that j would fulfil the agreement.

If the reliability of j was *high*, then i sent all of the promised chips.

If the reliability of j was *medium*, then the extent to to which i was reliable depended on the dependency relationships in the game:

1) If j was task dependent, and the agreement resulted in j becoming task *independent*, then i sent the largest set of chips such that j remained task dependent.

2) If the exchange resulted in the PURB agent becoming task independent, and j remaining task dependent, then the PURB agent sent all of the promised chips, or two thirds of its promised chips, depending on its confidence level of j's reliability measure being medium. This confidence level depended on the number of times in which the PURB

34

	Co-Dep.	Task Ind.	Task Dep.	Total
PURB agent vs. People	$(n = 28)$ **163**	$(n = 12)$ 187	$(n = 18)$ 82	$(n = 58)$ 143
People vs. People	$(n = 50)$ 131	$(n = 30)$ 181	$(n = 30)$ 102	$(n = 110)$ 136

Table 6: Performance for different Dependency Conditions

    agent interacted with j and j exhibited a medium measure of reliability.

3) If both agents were task dependent, and the agreement resulted in *both* agents becoming task independent, then i sent all of the promised chips.

Combining the PURB agent's social utility function with these rules allows it to adapt its negotiation behavior to that of the other participant.

5.5. Empirical Methodology and Results

    Fifty-four human subjects participated in the study drawn from a pool of undergraduate computer science students at Bar Ilan University. Each participant was given an identical 30 minute tutorial on Colored Trails that did not disclose the actual boards used in the study. Each participant was seated in front of a terminal for the duration of the study, and could not speak to any of the other participants.

    There were two conditions in the study. In the first condition, people played other people while in the second condition, people played the PURB agent. In all, there were 110 games that were played by people, and 58 games that were played by the PURB agent and people. Subjects were not told whether they would be playing a person or a computer agent.[12]

    Unless otherwise stated, the following results compare the behavior of the PURB agent playing people with that of people playing people. We list the number of observations and means for each result. Significance of results were confirmed for $p < 0.05$ using parametric statistical tests. Table 6 shows the average performance for the PURB agent and people, measured as the average score obtained in all games played.

---

[12]The exact wording given to subjects were "you may be interacting with a computer agent or with a person".

5.5.1. Analysis of Performance

As shown by the "total" column in the table, the PURB agent was able to negotiate as well as people: There was no statistically significant difference between its total average performance (143 points) and people's performance (136 points). However, there were distinct difference in performance for different dependency relationships between players. When players were co-dependent, the PURB agent significantly outperformed people (163 points versus 131 points). People outperformed the PURB agent in the task dependent condition (102 versus 82 points), but the difference was not statistically significant. There was also no significant difference in performance between the PURB agent and people in the task independent condition. In addition (and not shown in the table), in the co-dependent setting, the PURB agent was able to reach the goal 90% of the time, significantly more often than people, who reached the goal 67% of the time. There was no significant difference between people and the PURB agent in the extent to which the goal was reached when one of the players was task independent and the other player was task dependent.

5.5.2. Analysis of Behavior

The purpose of this section is to analyze the extent to which both people and the PURB agent fulfilled their commitments in negotiation. For any two participants i and j, let C_i denote the set of chips in possession of i at round n in the game. Let $O = (O_i, O_j)$ denote a proposal at round n, where $O_i \subseteq C_i$ was the set of chips that i agreed to send to j. Let $O_i^* \subseteq C_i$ be the set of chips actually sent by i following the agreement. (And similarly define C_j, O_j, and O_j^*.) Let $r_i(\{C_i \cup O_j\})$ denote the score to player i in the case that j sent all of its promised chips O_j, and i did not send any of its chips. We assume that the score is computed using the scoring function for the CT game that is described in Section 5.1. We refer to this value as the score that was *promised by j* in proposal O. The factor $r_i(\{C_i \cup O_j^*\})$ denotes the score to player i given the chips O_j^* that j actually delivered to i. We refer to this as the actual score to i given the chips that j transferred. The *actual reliability* of j given proposal O at round n is the ratio between the promised score to i from the chips O_j in proposal O and the actual score to i given the chips that O_j^* chose to transfer. This is computed as $\frac{r_i(\{O_i \cup O_j^*\})}{r_i\{C_i \cup O_j\}}$. Note that this score-based measure of reliability is different than the chip-based measure of reliability that was used by the PURB agent (Section 5).

	Co-Dep.	Task Ind.	Task Dep.	Total
PURB agent vs. People	(**0.79**, 0.27)	(**0.92**, 0.62)	(**0.74**, 0.62)	(**0.81**, **0.38**)
People vs. People	(0.53, 0.56)	(0.63, 0.58)	(0.41, 0.58)	(0.49, 0.53)

Table 7: Actual Reliability (left) and probability of acceptance (right) Measures

	Co-Dep.	Task Ind.	Task Dep.	Total
PURB agent vs. People	(**74, 86**)	(79, 23)	(10, 81)	(54.33, 63.33)
People vs. People	(40, 52))	(66, 15)	(14, 89)	(40, 52)

Table 8: other-benefit (left) and self-benefit (right) Measures

Table 7 compares the actual reliability and probability of acceptance for the PURB agent playing other people, and people playing other people. As shown in the Table, for all dependency relationships the PURB agent was consistently more reliable than people. Also shown by the table is that the PURB agent accepted offers significantly less often than did people (38% vs. 53%). In addition there was a significant correlation of 0.42 between the PURB agent's reliability and performance, while there was no such effect for people. (This result is not shown in the table). Thus, although the performance of the PURB agent was similar to people in score, the negotiation strategy used by the PURB agent was different from that of people in several ways. First, the PURB agent was significantly more reliable than people; and second, the PURB agent was less likely to accept offers than people.

We also compared the types of offers made by people and those made by the PURB agent, as well as how often these offers were accepted. Table 8 shows the average benefit from all proposals in a game that are associated with the proposing player (self-benefit), and the responding player (other-benefit). These results show that when both players were co-dependent, the PURB agent made offers that *Pareto dominated* (these offers were significantly more beneficial to both participants) the offers made by people.

Although the PURB agent we designed was able to negotiate as well as people, our analysis showed that its negotiation strategy differed from people: When it was task co-dependent its offers Pareto dominated those of people. In this condition, people exhibited medium reliability, and the PURB agent responded to this by generating offers that benefited itself more than people, as described in its strategy. Interestingly, the proposals made by the PURB

agent also offered more to people than did proposals made by people (74 versus 52 points). On average, the PURB agent was less likely to accept offers than people; also, it was more reliable than people in all conditions, and there was a significant correlation between the reliability of the PURB agent and its performance.

We conclude this section by describing an example that highlights the way the PURB agent was able to adapt to the different behavioral traits. The example is taken from two games in which the PURB agent was task independent, and its negotiation partner was task dependent. In both of these games, there was an identical offer made by the human participant that would allow both players to reach the goal. This example consisted of the person asking 1 chip from the PURB agent and offering 3 chips in return. The proposal was accepted by the PURB agent in both games. In the first game, the PURB agent chose to fulfil the agreement and sent the promised chip. However, in the second game, the PURB agent chose not to fulfil the agreement and did not send the promised chip. This was because the reliability of the person in the first game was considerably lower than the reliability of the person in the second game.

6. Repeated Negotiation in Settings of Incomplete Information

In this section, we describe a variant of the PURB agent that was designed to negotiate in a setting that modified the one in Section 5 as follows: (1) the participants were agent-based, some of which were designed by human subjects; (2) the participants in the negotiation lack pertinent information about each other's resources when they make decisions; and (3) the negotiation included multi-party interactions of four participants. The relevance of this mixed-network setting to negotiations of the type that may occur in the real world is in its inclusion of agents playing unknown negotiation strategies. Such negotiations are already commonplace in electronic commerce applications such as ebay, where people can construct automatic bidders to serve as proxies and represent themselves in the bidding process. The object of the study was to build an agent that is able to negotiate proficiently in this setting, despite its uncertainty about others' resources as well as their negotiation strategy.

Because our setting included numerous negotiation partners and limited information, the agent-design we describe in this section modified the PURB design in two ways: First, it was endowed with a separate adaptation mecha-

nism for each negotiation partner, and second, it estimated the cooperativeness of the negotiation partners without knowing the actual benefit associated with potential actions in the game.

6.1. Experimental Design

We modified the CT repeated interaction setting described in Section 5 so that each participant could observe its own chips, but not the chips of the other participants. Three types of agents were used in the study. The first type of agents were *Peer-Designed* (PD) agents, created by graduate-level computer science students at Bar Ilan University who were not given any explicit instructions beyond designing an agent that will represent themselves in a negotiation.

The second type of agent was a *Constant Personality* (CP) agent. This agent used a technique to estimate the cooperativeness of the other participants that was similar to the PURB agent described in Section 5.

The third type of agent was a *Varying Personality* (VP) agent. This agent was a variant of PURB that extended its utility function described in Section 5 to adopt a separate measure of cooperativeness for different levels of cooperativeness exhibited by the other participants. Table 9 specifies a match between cooperativeness measures exhibited by the VP agent given the cooperativeness measures exhibited by its negotiation partners. The matching process was done empirically, using a held-out test-set of PD agents that were not used again in the evaluation process. Both VP and CP agents were adaptive: they changed their behavior as a function of their estimate of others' cooperativeness, given the history of their observations. However, the VP agent adopted a unique measure of cooperativeness for each player, whereas the measure of cooperativeness for the CP agent was constant.

A series of repeated games were played between the VP agent and with the other agents in the systems. Each agent's final outcome was the aggregate of its scores in all of the games it participated in. For purpose of analysis, we classified PD and CP agents as either "cooperative" or "non-cooperative" according to the following: Cooperative CP agents were those that engaged in helpful exchanges more than 50% of the time and reneged on their commitments less than 20% of the time. We expected cooperative agents to realize opportunities for exchange with each other more often than non-cooperative agents and to exceed them in performance, as measured by the score in the game. We also expected that in some cases, non-cooperative agents would be able to take advantage of the vulnerability of those cooperative agents

Perceived Cooperativeness	Matched Cooperativeness
LL	LL
LM	LM
LH	LM
MM	LM
MH	MM
HM	HM
HH	MM

Table 9: Matching scheme used by VP agent to adapt its cooperativeness to the cooperativeness of its negotiation partners

	VP agent	PD and CP agents
Cooperative	170.6	114.8
Non-Cooperative	142.5	98.2

Table 10: Average performance of VP agent against cooperative/non-cooperative agents (3 repeated games)

that allow themselves to be exploited. Additionally, we hypothesized that the VP agent would be able to identify and reciprocate cooperative agents more quickly than CP or PD agents, while staying clear of agents that are non-cooperative. As a result, the VP agent would perform better than all other agents in the game.

We ran 5,040 games in our experiment, played in 1,080 rounds of three consecutive games each. The board games we used in each round varied the task dependency relationships between players. There were 4 players in each game, consisting of a VP agent, two CP agents, and one of the PD agents. The boards used were generated to span all possible task dependency role (dependent, co-dependent). Table 10 presents the average score for the VP agent when playing against cooperative and non-cooperative agents across all games. The scores reported in the table sum over the other players in the game.

As shown by the table, the average score achieved by the VP agent was significantly higher than all other agents, regardless of their level of cooperativeness. Also, the VP agent's score when playing against cooperative agents (170.6) was higher than its score when playing against non-cooperative agents

Exchange Type	Cooperative agents	Non-Cooperative agents
Reciprocal	60%	25%
Idle	20%	39%

Table 11: Percentage of exchange types proposed by VP agent

(142.5). Cooperative agents also benefited from cooperating with the VP agent: their performance was significantly higher than their non-cooperative counterparts (114.8 vs. 98.2).

In addition, the VP agent engaged in cooperative exchanges with cooperative agents significantly more often than the other agents, while the amount of time the VP agent remained idle when dealing with non-cooperative agents was longer than the amount of time other agents remained idle.

To conclude, in repeated game settings, the VP agent, which conditioned its own personality based on its estimate of others, outperformed all of the other agents in the system. Some types of agents escaped identification in intermediate rounds, resulting in an increase in their scores. However, the general performance of the VP agent was not affected. It performed better than all of the other agents in all of the games we played, and increased its score from game to game during the final rounds of the experiment. This shows that adopting a separate measure of cooperativeness for different agents is key to agents' success in task settings of incomplete information.

6.2. Related Work: Adapting to Agents' Behavioral Traits

The results reported in this section relate to recent approaches for learning to adapt to others' negotiation behavior in repeated interactions under uncertainty[13]. One strand of research involves learning and adapting to agents' behavioral traits. Zhang et al. [51] explored the trade-off between selfishness and helpfulness in environments in which agents were uncertain about the helpful nature of others in the system. They showed that although realizing every opportunity for cooperation was impossible, selfish agents do better than helpful agents as the rate of uncertainty in the system grows. Hogg and Jennings [22] proposed a model in which agents' utilities were a

---

[13]Other works have considered negotiation settings that allow reputation and revelation mechanisms, which does not directly relate to our work.

weighted summation of each others' expected outcomes. By learning these weights from observations, agents changed their measure of helpfulness over time. When all agents in the system were adaptive, high exploration rates led agents to seek out new negotiation opportunities and increased the overall social welfare of the group. All these models allowed agents to change their measure of helpfulness over time as a function of their model of others, and investigated the effects of this behavior on agents' cooperation and system performance. However, these works assumed that players fulfil their commitments following agreements, and did not model the behavioral types of others.

7. Conclusion and Future Work

This paper has presented studies which demonstrate the importance of learning and of incorporating social factors into agents' decision-making models when they operate in open, mixed networks. The studies focused on the design of computer agents for negotiation with each other, or with people, in settings that varied the availability of information, the negotiation protocol and the interdependence's that held between agents. For settings of complete information, our findings showed that

- Computer agents can successfully learn the extent to which different people are affected by social factors, and that this improves their decision-making.

- Agents that learn can outperform computational agents using traditional game- and decision-theoretic strategies.

For settings in which complete information about agents' resources is not available, or in which agents can renege on agreements, we have shown that

- Computer agents that adapt their level of cooperation to the varying degree of helpful behavior exhibited by others outperform computer agents that do not adapt.

- This adaptation process also facilities successful interaction among computer agents and people.

These studies used different configurations of the CT test-bed, a framework designed to investigate decision-making in such networks. The advantage of using CT is that it presents decisions to people in a context that

mirrors task settings in the real world. People are more likely to engage in cooperative behavior in CT environments than when in the context of traditional representations such as payoff matrices.

Additional work by the co-authors and others further demonstrate the usefulness of the CT framework in investigating a variety of questions about decision-making in group settings. In particular, the Colored Trails system has been used to construct and evaluate computational models of human reciprocity [15], to investigate the role of gender and social relationships in people's negotiation behavior [29], the way people reason about belief hierarchies in negotiation [13] and the way people respond to interruptions from computers [27]. CT has proved to support the rapid prototyping and ease of analysis of different kinds of decision-making models that is made possible when using CT. It has also been used as a pedagogical tool in for teaching agent design to students in courses at Harvard, Ben-Gurion and Bar-Ilan Universities.

8. Acknowledgments

References

[1] R. Arunachalam and N. M. Sadeh. The supply chain trading agent competition. *Electronic Commerce Research and Applications*, 4:63–81, 2005.

[2] M. Asada, P. Stone, H. Kitano, A. Drogoul, D. Duhaut, M. Veloso, H. Asamas, and S. Suzuki. The RoboCup physical agent challenge: Goals and protocols for phase I. *Robocup-97: Robot Soccer World Cup I*, pages 42–61, 1998.

[3] T. Babaian, B. J. Grosz, and S. M. Shieber. A writer's collaborative assistant. In *Intelligent User Interfaces Conference*, pages 7–14, 2002.

[4] M. Bazerman. *Judgment in Managerial Decision Making.* Wiley Publishers, 2001.

[5] S. R. Bocionek. Agent systems that negotiate and learn. *International Journal of Human-Computer Studies*, 42:265–288, 1995.

[6] A. Byde, M. Yearworth, K.Y. Chen, C. Bartolini, and N. Vulkan. Autona: A system for automated multiple 1-1 negotiation. In *Proceedings of the 4th ACM conference on Electronic commerce*, pages 198–199, 2003.

[7] C. F. Camerer. *Behavioral Game Theory. Experiments in Strategic Interaction.* Princeton University Press, 2003.

[8] J-C. Cardenas, T. K. Ahn, and Elinor Ostrom. Communication and co-operation in a common-pool resource dilemma: A field experiment. In Steffen Huck, editor, *Advances in Understanding Strategic Behaviour: Game Theory, Experiments, and Bounded Rationality: Essays in Honour of Werner Gth.* Palgrave, 2004.

[9] J.J. Castro-Schez, N.R. Jennings, X. Luo, and N.R. Shadbolt. Acquiring domain knowledge for negotiating agents: a case of study. *International Journal of Human-Computer Studies*, 61(1):3–31, 2004.

[10] A. Chan, K. MacLean, and J. McGrenere. Designing haptic icons to support collaborative turn-taking. *International Journal of Human-Computer Studies*, 66:333–355, 2008.

[11] A. Dreber, D. Rand, D. Fudenberg, and M. Nowak. Winners dont punish. *Nature*, 452(7185):348–351, 2008.

[12] A. Falk and U. Fischbacher. A theory of reciprocity. *Games and Economic Behvior*, 54(2):293–315, 2006.

[13] S. G. Ficici and A. Pfeffer. Simultaneously modeling humans' preferences and their beliefs about others' preferences. In *Proc. 7th International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS)*, 2008.

[14] K. Fullam, T. Klos, G. Mueller, J. Sabater, A. Schlosser, Z. Topol, K. S. Barber, J. Rosenschein, L. Vercouter, and M. Voss. The agent reputation and trust (ART) testbed competition. Technical report, University of Texas at Austin, 2004.

[15] Y. Gal and A. Pfeffer. Modeling reciprocity in human bilateral nego-
tiation. In *Proc. 22nd National Conference on Artificial Intelligence
(AAAI)*, 2007.

[16] Y. Gal, A. Pfeffer, F. Marzo, and B. Grosz. Learning social preferences
in games. In *Proc. 19th National Conference on Artificial Intelligence
(AAAI)*, 2004.

[17] Y. Gal, B. Grosz, S. Shieber, A. Pfeffer, and A. Allain. The effects of
task contexts on the decision-making of people and computers. In *Proce.
of the Sixth International Interdisciplinary Conference on Modeling and
Using Context, Roskilde University, Denmark*, 2007.

[18] Y. Gal, S. Dsouza, P. Pasquier, I. Rahwan, and S. Abdallah. The effects
of goal revelation on computer-mediated negotiation. In *Proceedings of
the Annual meeting of the Cognitive Science Society (CogSci), Amster-
dam, The Netherlands*, 2009.

[19] Y. Gal, S. Kraus, M. J. Gelfand, Y. Blumberg, and E. Salmon. An
adaptive agent for negotiating with people in different cultures. *ACM
Transactions on Intelligent Systems and Technology*, 2010.

[20] B. Grosz, S. Kraus, S. Talman, and B. Stossel. The influence of social
dependencies on decision-making. Initial investigations with a new game.
In *Proc. 3rd International Joint Conference on Autonomous Agents and
Multi-agent Systems (AAMAS)*, 2004.

[21] J. Henrich, R. Boyd, S. Bowles, C. F. Camerer, H. Gintis, R. McElreath,
and E. Fehr. In search of Homo economicus: Experiments in 15 small-
scale societies. *American Economic Review*, 91(2):73–79, 2001.

[22] L. M. J. Hogg and N. R. Jennings. Socially intelligent reasoning for
autonomous agents. *IEEE Trans on Systems, Man and Cybernetics -
Part A 31(5)*, pages 381–399, 2001.

[23] E. Horvitz. Principles of mixed-initiative user interfaces. In *Proc. of
ACM SIGCHI Conference on Human Factors in Computing Systems
(CHI)*, 1999.

[24] C. M. Jonker, V. Robu, and J. Treur. An agent architecture for multi-
attribute negotiation using incomplete preference information. In *Proc.*

6th International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS), 2007.

[25] E. Kamar, E. Horvitz, and C. Meek. Mobile Opportunistic Commerce: Mechanisms, Architecture, and Application. In *Proc. 7th International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS)*, 2008.

[26] E. Kamar, Y. Gal, and B. J. Grosz. Incorporating helpful behavior into collaborative planning. In *Proc. 8th International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS)*, 2009.

[27] E. Kamar, Y. Gal, and B. J. Grosz. Modeling User Perception of Interaction Opportunities for Effective Teamwork. In *IEEE Conference on Social Computing*, pages 271–277. IEEE Computer Society, 2009.

[28] R. Katz and S. Kraus. Efficient agents for cliff edge environments with a large set of decision options. In *Proc. 5th International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS)*, 2006.

[29] R. Katz, Y. Amichai-Hamburger, E. Manisterski, and S. Kraus. Different orientations of males and females in computer-mediated negotiations. *Computers in Human Behavior*, 24(2):516–534, 2008.

[30] H. Kitano. RoboCup rescue. a grand challenge for multi-agent systems. In *Fourth International Conference on Multi-Agent Systems (ICMAS)*, 2000.

[31] S. Kraus and D. Lehmann. Designing and building a negotiating automated agent. *Computational Intelligence*, 11(1):132–171, 1995.

[32] S. Kraus, P. Hoz-Weiss, J. Wilkenfeld, D.R. Andersen, and A. Pate. Resolving crises through automated bilateral negotiations. *Artificial Intelligence*, 172(1):1–18, 2008.

[33] R. Lin and S. Kraus. Can automated agents proficiently negotiate with humans? *Communications of the ACM*, 53(1), 2010.

[34] R. Lin, S. Kraus, J. Wilkenfeld, and J. Barry. Negotiating with bounded rational agents in environments with incomplete information using an automated agent. *Artificial Intelligence Journal*, 172(6-7):823–851, 2008.

[35] X. Luo, N. R. Jennings, and N. Shadbolt. Acquiring user tradeoff strategies and preferences for negotiating agents: A default-then-adjust method. *International Journal of Human-Computer Studies*, 64:304–321, 2006.

[36] R. R. Murphy. Humanrobot interaction in rescue robotics. *IEEE Transactions On Systems, Man, and CyberneticsPart C: Applications and Reviews*, 34(2), May 2004.

[37] J. Nash. The bargaining problem. *Econometrica*, 18:155–162, 1950.

[38] R. M. Neal and G. E. Hinton. A view of the EM algorithm that justifies incremental, sparse, and other variants. *Learning in graphical models*, 89:355–368, 1998.

[39] M.A. Nowak. Five rules for the evolution of cooperation. *Science*, 314 (5805):1560, 2006.

[40] Y. Oshrat, R. Lin, and S. Kraus. Facing the challenge of human-agent negotiations via effective general opponent modeling. In *Proc. 8th International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS)*, 2009.

[41] E. Ostrom, R. Gardner, and J. Walker. *Rules, games, and common-pool resources*. Univ of Michigan Press, 1994.

[42] M.E. Pollack. Intelligent technology for an aging population: The use of AI to assist elders with cognitive impairment. *AI Magazine*, 26(9), 2006.

[43] D.V. Pynadath, M. Tambe, Y. Arens, H. Chalupsky, Y. Gil, C. Knoblock, H. Lee, K. Lerman, J. Oh, S. Ramachandran, et al. Electric elves: Immersing an agent organization in a human organization. In *Proceedings of the AAAI Fall Symposium on Socially Intelligent Agents*, 2000.

[44] D. Rajarshi, J. E. Hanson, J. O. Kephart, and G. Tesauro. Agent-human interactions in the continuous double auction. In *Proc. 17th International Joint Conference on Artificial Intelligence (IJCAI)*, 2001.

[45] D. Sally. Conversation and cooperation in social dilemmas. *Rationality and Society*, 7(1):58–92, 1995.

[46] N. Schurr, P. Patil, F. Pighin, and M. Tambe. Using multiagent teams to improve the training of incident commanders. In *Proc. 5th International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS)*, 2006.

[47] S. M. Shieber. A call for collaborative interfaces. *Computing Surveys*, 28A (electronic), 1996. URL `http://www.eecs.harvard.edu/~shieber/Biblio/Papers/collab-interface-stmt.html`.

[48] S. Talman, Y. Gal, M. Hadad, and S. Kraus. Adapting to agents' personalities in negotiation. In *Proc. 4th International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS)*, 2005.

[49] D. Traum, J. Rickel, J. Gratch, and S. Marsella. Negotiation over tasks in hybrid human-agent teams for simulation-based training. In *Proc. 2nd International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS)*, 2003.

[50] A. van Wissen. Team formation among self-interested human actors. Master's thesis, Utrecht University, 2009.

[51] X. Zhang, V. Lesser, and T. Wagner. Integrative negotiation in complex organizational agent systems (reserch abstract). In *Proc. 1st International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS)*, 2002.