Modeling User Perception of Interaction Opportunities for Effective Interruption Management

Ece Kamar, Ya'akov Gal and Barbara J. Grosz
School of Engineering and Applied Sciences
Harvard University, Cambridge, MA 02138 USA
{kamar, gal, grosz}@eecs.harvard.edu

***Abstract*—This paper proposes a new model for interruption management, one which takes into account the costs and benefits to both people and computer agents engaging in a collaborative activity. The model aims to maximize the efficiency of the collaboration by better estimating the utility of interruption outcomes and by reasoning about a possible mismatch between a computer agent's computation of this utility and a person's perception of it. The paper formalizes interruptions as a multi-agent decision making problem and shows how to decouple it into smaller, interacting sub-problems that can be analyzed independently and efficiently. We show how to combine the sub-problems to compute the utility of interruption correctly. This process is evaluated empirically using an abstract two-player game that is analogous to collaborative settings in which computer agents and people interact, and one partner can interrupt the other in order to obtain needed information. Results show that the magnitude of the benefit of interruption is a major factor that influences the likelihood that people will accept interruption requests. When the benefit of the interruption is not clear-cut, people's perceived type of the partner (whether it was a human or computer) significantly affected their perceptions of the usefulness of interruptions. These results imply that system designers need to consider the possible benefit of interruptions to collaborative human-computer teams as well as the way that this benefit is perceived by people.**

I. Introduction

Interruptions are important for effective collaborative work, because agents often possess information required by others on their team. This need to get information from another agent arises in homogeneous multi-agent environments as well as heterogeneous groups that include people and computer agents. Individuals collaborating as part of such groups benefit from their interactions with each other, gaining from sharing information as well as by coordinating their activities. For example, a (human) driver may see changes in weather conditions that affect route selection while an automated navigation system may not have access to this information, but may need it to identify the best route. This need for interaction also arises in groups involving a single computer and user. Consider a writer's collaborative assistant that autonomously searches for bibliographical and citation information [1]. The system primes the user for key-words, searches for bibliographical data, and asks the user to choose the appropriate citations from this list, or to refine the search. However, interruptions are inherently disruptive. If they are not managed and timed properly, they may negatively affect the emotional state and awareness of the user and may reduce the overall task performance of the user and the system [2]. For example, a navigation system that continuously asks drivers for weather conditions may distract their attention from the road, and a writer's assistant that asks whether to cite each paper that meets the user's keywords will disrupt the user's writing process. Thus, it is crucial for computer agents to manage interruptions appropriately for working efficiently with people and with other computer agents.

Our research aims to enable the design of efficient interfaces, ones which maximize the likelihood that valuable interruptions will be accepted. We propose a new model for interruption management which synthesizes techniques from decision theory and computer science, but adapts them to collaborative contexts. Our model takes into account the costs and benefits to all participants, both people and computer systems, so that decisions to interrupt are based on the collaborative benefit to the group. Unlike previous models of interruption management, this model also reasons about the possible mismatch between a computer's estimate of the utility of an interruption and a person's perception of it. It focuses on determining the factors that influence people's perception of interruptions, and their tendency to accept or reject them when they are generated by a computer system.

Our study is focused on fast-paced collaborative settings in which agents are distributed, conditions may be rapidly changing, actions are occurring at a fast pace, and decisions must be made within tightly constrained time frames with uncertainty and partial information. We model this setting as a multi-agent decision-making problem, and introduce a novel method for decoupling this problem into individual, simpler sub-problems which can be analyzed independently. We show how to integrate between the solutions to these constituent sub-problems in order to correctly compute the value of an interruption for the participants in the global multi-agent problem. These computations are used to inform a computer agent for managing interruptions with people and is compared with the way people actually perceive interruptions in an empirical collaborative setting.

The investigations deploy a domain-independent, abstract game which we designed to provide an analogue of human-computer interactions in collaborative fast-paced settings in which participants have access to different sources of information, the environment is uncertain, and one of the players can choose to interrupt the other player to obtain information. This collaborative game provides a test-bed in which to identify the factors that affect people's decision making. It is built on the Colored Trails (CT) infrastructure which has been used previously as a research test-bed for a variety of multi-agent decision-making problems [3].

We investigated the effect of three factors on human perception of the usefulness of interruption requests in a set of experiments which vary: the magnitude of the interruption utility, the timing of interruptions, and the perceived type of the partner (a human or a computer agent). The results revealed that the magnitude of the utility of interruption is the major factor affecting the likelihood that people will accept interruptions. Interruptions offering significant benefit to the collaboration are consistently accepted regardless of the type of partner and the timing of the interruption. The results also indicate that the perceived partner type and cost of the interruption to the subject affect people's perception of interruptions when the utility gain is less clear-cut.

This paper makes four key technical contributions: (1) it defines a model of interruption that considers the costs and benefits to both the interrupting agent and the person with whom that agent is interacting; (2) it defines an efficient method for estimating the utility of interactions in uncertain multi-agent environments; (3) it describes a new, multi-agent game for investigating interruption decisions in fast-paced domains; (4) it empirically investigates people's decisions about interruptions to reveal the major factors influencing those decisions.

II. Related Work

A key aspect of reasoning about interruptions in collaborative settings is the ability to accurately estimate the costs and benefits of the interruption to all parties so that the outcome of the interruption positively affects group task outcomes. Previous work on adjustable autonomy identifies the points at which it is most suitable to initiate interactions with a person, but does so without relating this decision to a user's mental state or the task being performed. Interruptions are driven solely by system needs and managed based on benefit to the system [4]. Prior work on interruption management addressed user needs, and has focused mostly on the effect an interruption has on a person's cognitive state, rather than the benefit of interruptions to collaborative activities [5]. Few models have combined these two aspects into an integrated decision making mechanism [6], and none have done so in the kinds of fast-paced domains we consider.

While there has been significant work on mixed-initiative system design, there has been little empirical work on how people perceive interruption utilities and make interruption decisions in human-computer interaction settings. Avrahami et al. [7]. investigated the differences between a person's self report of interruptibility and other people's predictions about that person's interruptibility. However, this work considered face to face human interaction, rather than human-computer interaction. Gluck et al. [8], focused on designing notification methods to increase human perception of utility whereas Bunt et al. [9] showed that displaying system rationale to people may increase person trust to computer systems. Nevertheless, little attention has been paid to the possible discrepancy between a computer agent's calculation of the utility of an interruption and a person's perception of its usefulness. The failure to estimate this accurately may cause a person to reject a valuable interruption, and thus lead to a missed opportunity to improve team performance, thereby turning what could have been a beneficial interruption into a performance degrading disturbance.

III. The Interruption Game

This section describes the design of a special-purpose game for investigating the interruption management problem in a setting that is not tied to a specific domain. The "interruption game" involves two players, referred to as the "principal" and the "agent". Each player needs to complete an individual task. The game is collaborative in that the score for each player depends on its own performance as well as the performance of the other player. The players share a joint score function that is the cumulative score of both players. The agent player lacks critical information about its task, which is known to the principal player. Thus, the agent has incentive to request this information from the principal by initiating an interruption, and the principal has reason to provide the information in order to improve the joint performance. Providing this information is costly for both players. When the agent deliberates about whether to ask the information, or when the principal deliberates about whether to reveal the information to the agent, the players need to weigh the trade-offs associated with the potential benefit to the agent player with the detriments to their individual performance in the game.

The game is played on a board of 6x6 squares. Each player is allocated a starting position and a goal position on the board. The game comprises a fixed, known number of rounds. At each round, players advance on the board by moving to an adjacent square. The goal of each player advances stochastically on the board according to a Gaussian probability distribution centered at the current position of the player. [1] Players earn 10 points in the game each time they move to the square on which their assigned goal is located and the goals are reassigned to a random position on the board. Players can see the positions of both players and the goal location of the other player, but they differ in their ability to see their own goal location: The principal can see the location of its goal throughout the game, while the agent can see the location of its goal at the onset of the game, but not during consecutive rounds.

At any round, the agent can choose to interrupt the principal and request the current location of its goal. The principal is free to accept or reject an interruption request. If the principal rejects the interruption request, the players continue their individual play. If the interruption is accepted by the principal agent, the location of the agent's goal in the current round (but not in consecutive rounds) is automatically revealed to

[1]The movement of the goal is restricted in that it does not move closer to the position of the player.

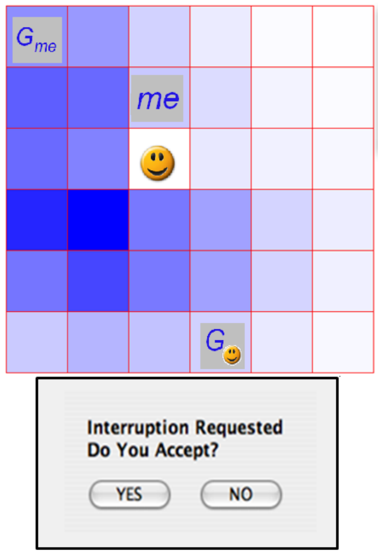

Fig. 1. Game Screen-shot: *me* is the principal player, smiley is the agent player, G_{me} is the principal's goal, G_{smiley} is the agent's goal. The degree to which each square is shaded represents the agent's uncertainty about its goal. Dark squares imply higher certainty.

the agent. There is a joint cost for revealing this information to the agent, in that both participants will not be able to move for one round. The agent cannot interrupt the principal more than once.

The rules of the game also require the agent to provide the principal with its belief about the location of its goal. This information is available to the principal when it needs to decide whether to accept an interruption. A snapshot of the game from the perspective of the principal player is shown in Figure 1. The rules of the game provide incentives to players for reaching their goals as quickly as possible and interruptions initiated by the agent are critical determinants of players' performance. The agent's uncertainty about the location of its own goal increases over time, and its performance depends on successfully querying the principal and obtaining the correct position of its goal.

An interruption is potentially beneficial for the individual performance of the agent, who can use this information to direct its movement, but not for the individual performance of the principal, who has complete information about the world. The success of both players in the game depends on the agent's ability to estimate the collaborative value of interruption at each point in the game, and use that information to choose when to interrupt the principal.

The interruption game is not meant to model or specify a complete domain or application. Its purpose is to provide a simple setting in which to study the factors that influence interruption management in collaborative settings. It is analogous to the types of interactions that occur in fast-paced collaborative settings described in Section I. For example, the principal player in the interruption game may represent the user of a collaborative system for writing an academic paper, and the agent may represent the collaborative assistant responsible for obtaining bibliographical data. While both of the participants share a common goal of completing a document, each of them must work independently to complete its individual task, such as composing paragraphs or searching for bibliographical information. This aspect is represented in the interruption game by assigning an individual goal for each player. The movement of these goals on the board corresponds to the dynamic nature of these tasks. For example, the user may not know the next paragraph to write, and the system may have uncertainty about the search results for new bibliographical data. The agent's lack of information about its own goal location in the game corresponds to the uncertainty of a system about the preferences and intentions of its user, such as which bibliographical information to include in the paper. The ability to query the user for keywords and to choose among different bibliographies provides the system with valuable guidance and direction but may impede the performance of both participants on their individual tasks. This is because the system needs to suspend its search for bibliographical data when it queries the user, and the user may be distracted by the query. This dynamic cost of interruption represents the costs incurred to both users and computer agents due to task switching and task recovery for initiating and responding to an interruption.

IV. Modeling Interruption Opportunities

In this section we show that the interruption game can be formalized as a multi-agent decision-making problem under uncertainty, and provide efficient methods for computing players' estimates for the benefit of interruptions in the game. These models are not meant to predict how people play the interruption game, or respond to interruption requests in general. Rather, they provide a way to compute a theoretical baseline, which is a fully rational computational estimate for the value of interruption in the game. In the empirical section, we use this baseline to enable empirical analysis of human behavior in mixed-initiative settings. These analysis allow the study of the efficacy of these models when they are used by computers to interact with people in the game under various experimental conditions.

The interruption game can be modeled as a Decentralized Markov Decision Process (Dec-MDP) [10], a formalism for multi-agent planning that captures the collaborative nature of the interruption game and its associated uncertainty. A Dec-MDP includes a set of states with associated transition probabilities, a set of actions and observations for each agent, and a joint reward function. A solution of a Dec-MDP is an optimal joint policy for all agents that is represented as a mapping from states to actions.

To model the interruption game, the state space of the Dec-MDP will combine all of the information relating to the tasks of both players, including their positions on the board, the positions of their goals, the current round and the belief of the agent about its own goal position. The solution of Dec-MDP will assign a policy to the agent that initiates interruption requests when they are expected to result in a benefit to both players according to the joint reward function, and will assign a policy for the principal to accept interruption requests with actual positive benefit. In this way, the representation is able to capture the collaborative and stochastic aspects of the interruption game. Unfortunately, finding optimal solutions to Dec-MDPs is NEXP-complete [11]. The size of state space makes it infeasible to compute the complete joint policy for both players in the interruption game.

However, our goal is not to exhaustively compute optimal policies in the interruption game, but to be able to generate interruptions when they are perceived to be beneficial to the collaboration. We hypothesized that such interruptions would be likely to be accepted by people. To do so, we will exploit an important characteristic of this game: When players are not making or replying to interruption requests, they are performing their individual tasks, and each player only needs to consider its individual score in the game. In this case, the two tasks are essentially independent, and they can be solved separately. Because the agent can only interrupt the principal once, the expected utility of an interruption can be computed efficiently, because an interruption request will render the two tasks independent from the interruption moment until the end of the game. When the principal is interrupted by the agent, the principal does not need to consider future opportunities of interruptions as the tasks of both players are completely independent of each other for game turns following the interruption request. Thanks to this decoupling, at each turn, the policy for the agent is to interrupt and request information from the principal when it is deemed beneficial for both participants. We shall now detail how to solve the individual tasks for both participants in the game, and how to combine these solutions in order to devise strategies for interruption management in the game.

A. Computing a Policy for the Principal

The principal has complete information about the game, and its task can be modeled as a Markov Decision Process (MDP). Let B denote the set of board positions; $|B|$ denotes the size of the game board; $p \in B$, $g \in B$ are the positions of the principal and its goal respectively; $m \in A$ is a movement action of the player; R is the reward associated with reaching the goal; $P(g', p, g)$ is the probability of the goal position moving from position g to position g' when the player is in position p. The state space of the MDP includes every possible position of the principal and its goal at each round. We denote $S_P^h = \langle p, g \rangle$ to be the state at round h and induce a state transition function T that assigns a probability to reaching state S_P^{h+1} from S_P^h given action m. T can be directly derived from P.

The reward function assigns the score in the game for reaching the goal if an action transitions a player to its goal square, and 0 otherwise.

$$R(S_P^h, m) = \begin{cases} 10 & \text{if } p + m = g \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Let Π_P^* denote the optimal policy for the principal agent in the game. The value $V^{\Pi_P^*}(S^h)$ of this policy at state S_P^h maximizes the reward at state S_P^h for an action m and future states given the transition probability function,

$$V^{\Pi_P^*}(S_P^h) = max_m[R(S_P^h, m) + \sum_{S^{h+1}} T(S^{h+1}, m, S_P^h) \cdot V^{\Pi_P^*}(S_P^{h+1})] \tag{2}$$

We compute the optimal policy and its value using ExpectiMax search. In this process we grow a tree with two types of nodes, decision nodes and chance nodes. There is a decision node for each state in the MDP, and each child of a decision node is labeled with a movement action for the principal. Chance nodes represent moves of nature, and each child of a chance node is labeled with a possible movement of the goal of the principal, and is assigned a probability according to the transition function. When traversing the tree, we recursively compute a value for each chance node that is a weighted average of the value of each of its children according to its probability. We compute a value for each decision node by choosing the child with the maximal value, and select that action. With memoization, the number of nodes generated by the search is bounded by $|B|^2 \cdot |H|$, which is polynomial in the number of rounds in the game.

B. Computing a Policy for the Agent

The agent cannot observe the position of its goal on the board, and without interrupting the principal it receives no information relating to this position. We model its task as a No Observation Markov Decision Process (NOMDP), which is a special case of an MDP with no observations. The state space for this model includes the position l of the agent on the board, its belief $b \in \Delta B$ over its goal position, and current turn h. We denote the state for the agent as $S_A^h = \langle l, b \rangle$. As we are modeling the agent's individual task, rather than its interaction with the principal, we leave out the interruption action and use the set of actions A and reward function R is identical to the ones described for the principal agent. To compute the optimal policy we need to account for the fact that after each turn h, the agent updates its belief that the goal has moved to a new position in turn $h+1$. This is done by summing over each of the possible initial positions $c \in B$ for the goal. Formally, b is updated to $b^{'}$ according to the goal movement distribution P as follows:

$$\forall c' \in B, b'(c') = \sum_{c \in B} b(c) \cdot P(c', l, c) \tag{3}$$

The state at round $h+1$ for the agent is $S_A^{h+1} = \langle l', b' \rangle$, where l' is the position of the agent at $h+1$, and b' is computed using Equation 3. The value of an optimal policy for the agent Π_A^* at state $S_A^h = \langle l, b \rangle$ can thus be computed using Equation 2, substituting Π_A^* for Π_P^* and S_A for S_P.

Because the belief of the agent about its goal position is incorporated into the state space, there are an infinite number of states to consider, and using ExpecitMax in a straightforward fashion is not possible. However, applying the belief update function after each turn, only a small number of states turn out to be reachable. The deterministic belief update function maps each combination of states with full information (i.e., states in which the agent knows the correct position of its goal) and the number of turns since full information to a single belief state, thus to a single state. As a result, we can grow the search tree "on the fly", and only expand those states that are reachable after each turn. Memoization is not possible in this technique, and thus the complexity of the complete search is exponential in the length of the horizon.

C. Computing the Benefit of Interruption

To compute the benefit of an interruption, its effect on both the agent's and the principal's individual performance must be taken into account. It is the aggregate of these two effects that determines the utility of interruption. An interruption is initiated by the agent, but it is only established if the principal accepts it. The effect of an interruption on the individual game play of a player is the difference of the values of two states; one in which an interruption is established, and other in which it is not. Given the principal and its goal are located on squares p and g respectively in game round h, let $EU_P^{NI}(S_P^h = \langle p, g\rangle)$ denote the expected utility of the principal when it is not interrupted, and pursues its individual task. This is equal to the value for the principal of carrying out its individual task as shown in Equation 2. Thus we have

$$EU_P^{NI}(S_P^h) = V^{\Pi_P^*}(S_P^h) \tag{4}$$

For the agent that does not get to observe its own goal position, let $EU_A^{NI}(S_A^h = \langle l, b\rangle)$ denote the expected utility of the agent when it is not interrupted, and pursues its individual task. In a similar fashion to the principal's case, this is the value to the agent of carrying out its individual task. Thus we have,

$$EU_A^{NI}(S_A^h) = V^{\Pi_A^*}(S_A^h) \tag{5}$$

We now show how to compute the expected value of interruption for both participants. Let $EU_P^I(S_P^h = \langle p, g\rangle)$ denote the expected utility for the principal when it accepts an interruption. If the principal player is interrupted, it cannot move for one round, but its goal may move stochastically according to the probability distribution P. We denote the new goal position as g^{h+1}. Given our constraint that there can only be one interruption made in the game, the benefit of interruption for the principal is the expected value of its individual task in future rounds, for any possible position of its goal. Formally, the utility of interruption for state S_p^h, denoted $EU_P^I(S_p^h)$ is computed as follows:

$$EU_P^I(S_P^h) = \sum_{g^{h+1}\in B} P(g^{h+1}, p, g)\cdot V^{\Pi_P^*}(S_P^{h+1} = \langle p, g^{h+1}\rangle) \tag{6}$$

If the agent successfully interrupts the principal, the principal will reveal the position of the agent's goal. The agent will update its belief about its goal position in the following round, and use this belief to perform its individual task in future rounds. However, when it deliberates about whether to interrupt in the current round, it needs to to sum over every possible position of its unobserved goal, according to its belief about the goal location. Let $S_A^h = \langle l, b\rangle$ be the current state of the agent, including its position on the board and belief about its goal position. Let g denote the current position of its goal. The expected value for interruption for the agent is denoted EU_A^I and is computed as follows:

$$EU_A^I(S_A^h) = \sum_{g\in B} b(g)\cdot V^{\Pi_A^*}(S_A^{h+1} = \langle l, b'\rangle) \tag{7}$$

Here, b' refers to the belief state of the agent in which probability 1 is given to the true position of its goal as revealed by the principal, and updated once to reflect the stochastic movement of the goal in turn h.

D. Deciding Whether to Interrupt

By combining the expected values of the principal and agent players with and without interruption, it is now possible to compute the agent's estimate of the benefit of an interruption. We denote $EBI_P(S_P^h)$ to be the expected benefit of interruption for the principal given S_P^h, which is simply the difference in utility of the principal between accepting and interruption and carrying out its individual task.

$$EBI_P(S_P^h) = EU_P^I(S_P^h) - EU_P^{NI}(S_P^h) \tag{8}$$

The expected benefit of interruption for the agent is denoted $EBI_A(S_A^h)$ and is computed similarly:

$$EBI_A(S_A^h) = EU_A^I(S_A^h) - EU_A^{NI}(S_A^h) \tag{9}$$

The interruption game is collaborative in that the combined performance of both participants determines their individual scores. The agent can observe the state S_P^h of the principal, and for any combined state $S^h = (S_P^h, S_A^h)$, the agent will consider the joint expected benefit of interruption to both participants, EBI, and choose to interrupt if this joint benefit is positive.

$$EBI(S^h) = EBI_P(S_P^h) + EBI_A(S_A^h) \tag{10}$$

For any combined state S^h, the optimal strategy for the agent is to interrupt the principal if the expected benefit of the interruption (EBI) is larger than the expected value of acting individually. The agent cannot observe the correct position of its goal and estimates the benefit of interrupting under this uncertainty. Thus, not every interruption initiated by the agent is truly beneficial for the team. In contrast, the principal can observe the position of agent's goal and can capture the actual benefit of the interruption, denoted ABI with certainty. Any interruption with positive ABI offers a positive benefit to the team. The value of ABI is the sum of the individual benefits of interruption to both the principal and the agent. However, the principal knows the belief distribution of the agent's goal, as well as the correct position of its goal, and is thus able to compute an informed estimate about the value of revealing this information to the agent.

Let g_a be the agent's goal position, the actual benefit of interruption for both participants given states S_P^h and S_A^h is

$$ABI(S^h) = EBI_P(S_P^h) + EBI_{P,A}(S_A^h) \tag{11}$$

Here, the term $EBI_{P,A}(S_A^h)$ refers to the principal's perception of the agent's benefit from revealing the goal position g_a, where l is the current position of the agent, b' refers to the belief state of the agent in which probability 1 is given to g_a and updated once.

$$EBI_{P,A}(S_A^h) = V^{\Pi_A^*}(S_A^{h+1} = \langle l, b' \rangle) - EU_A^{NI}(S_A^h) \quad (12)$$

The advantage of the decision-making model introduced above is that it reduces the complexity of the multi-agent decision making process to that of two separate single agent decision making processes. Because the agent is allowed to interrupt only once during a game-play, the decoupling method is able to accurately capture the benefit of an interruption initiated by the agent. Our model assumes that principal players are fully rational and computationally unbounded. In the case of a computationally unbounded and rational principal player, we would expect any interruption with positive ABI to be accepted, and any interruption with negative ABI to be rejected. In the empirical investigations described in the next section, baseline values are compared with the subject responses to detect the possible mismatch between a computer's estimate of the benefit of an interruption and a person's perception of it, and to identify a subset of factors that affect the way that humans perceive the effectiveness of interruptions.

V. Empirical Methodology

This section evaluates the computational strategies we derived for playing the interruption game in an empirical setting. A total of 26 subjects participated in the study. The subjects were between ages of 19 and 46 and were given a 20 minute tutorial of the game. Subjects played between 25 to 35 games each, and were compensated in a manner that was proportional to their total performance.

During the empirical evaluation, all subjects were allocated to play the roles of principals, while the role of the agent was assigned to a computer that used the methodology described in the previous section to play the interruption game. Each game proceeded in the manner described in Section III. In particular, the agent could observe the board, the location of the participants, the goal of the principal agent, but not its own goal location. The principal had full visibility of the board, including the belief of the agent about its goal location, as shown in Figure 1. At each game, the agent is allowed to initiate an interruption once to acquire the correct location of its goal from the principal. As we mentioned, at each turn of the game, the policy of the agent is to interrupt only if the expected benefit of an interruption (EBI) is positive, given its uncertainty about the position of its goal.

Interruptions were generated by the computer agents at different points in the game with varying actual benefits, game levels and perceived partner types. We measured people's responses to these requests given the game conditions at the time of interruptions, which included the number of turns left to play, the positions of both players on the board, and the agent's belief about the location of its goal.

As principal players have full visibility through the game, we hypothesized that subjects are likely to perceive the benefit of an interruption using a similar computation to that of Equation 11, which computes the actual benefit of interruption (ABI) given the current game conditions. In this case, they would be likely to accept an interruption if it offered a higher benefit to both participants. The game scenarios were varied to have different ABI values to understand the way subject responses change with respect to benefit of interruptions. The strategy used by the computer playing the agent role in each game was to interrupt if the joint expected benefit of interruption was positive, by computing the value of EBI using Equation 10. Because the agent makes interruption decisions under the uncertainty about its goal position we did not expect all of its interruption requests to be accepted by the principal. However, we hypothesized that an agent using the EBI estimate to interact with people will choose to interrupt when they are likely to accept the offer. To investigate the effect of the timing of an interruption on the subjects' likelihood of acceptance, we varied the level of the game that an interruption is initiated. Lastly, we expected that the type of agent player (whether a computer or human) will affect the way people respond to interruption requests. For this reason, subjects were told they would be interacting with a human for some games, however they were always paired with an agent [2].

Subjects were given randomly generated game scenarios that vary the actual benefit of interruption to both participants (ABI) to cover four types of values: -1.5 (small loss), 1.0 (small gain), 3.5 (medium gain), 6.0 (large gain). These values represent the smallest and largest benefit values that can be generated from interruptions with positive expected benefit (EBI), which is a necessary condition to initiate interruption requests by the agent player. The levels in the game in which interruptions occurred varied to cover the beginning of a game (level 3), the middle of a game (level 5) and the end of a game (level 7). There were 540 game instances played when the perceived agent was a computer (PP:Computer) and 228 data points when the perceived agent was a person (PP:Person).

VI. Results and Discussion

The following results analyze a total of 768 game instances collected in our study. Table 1 shows interruption acceptance rates for different levels and ABI values for the same game instances when perceived partner type (PP) is person or agent. The optimal policy for the principal agent is to accept an interruption if its associated benefit (ABI) is positive, and reject otherwise.

As the results of Table 1 show, the utility of an interruption is the major factor influencing the probability that an interruption will be accepted by a person. The interruption acceptance rate increases significantly as the benefit of interruption rises from -1.5 to 1.0 ($p < e^{-20}$, α=0.001) and from 1.0 to 3.5

[2] Approval was obtained for the use of human subjects in research for this misinformation.

PP:Computer	Level 3	Level 5	Level 7
ABI -1.5	0.16	0.16	0.41
ABI 1.0	0.27	0.7	0.81
ABI 3.5	0.91	0.97	0.79
ABI 6.0	0.91	0.95	0.95
PP:Person	**Level 3**	**Level 5**	**Level 7**
ABI -1.5	0	0.11	0.44
ABI 1.0	0.72	0.94	1
ABI 3.5	0.91	0.94	0.88
ABI 6.0	1	0.88	1

TABLE I
ACCEPTANCE RATE OF INTERRUPTIONS

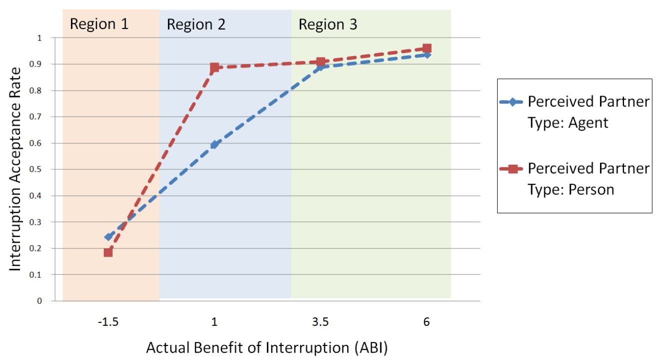

Fig. 2. Effect of interruption benefit and perceived partner type on interruption acceptance rate

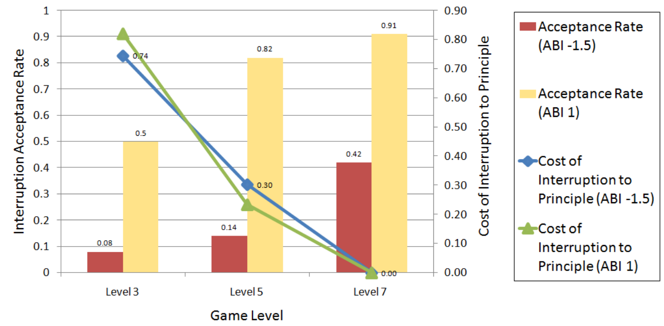

Fig. 3. Correlation of interruption acceptance rate with the cost of interruption to subjects (-ABI_P) for small gains and losses.

(p=0.0013, α=0.01). However the rise from 3.5 to 6.0 does not further improve the acceptance rate. These results confirm that people were successful at perceiving interruption benefits above a certain threshold. Similarly, when an interruption is costly for the collaboration, people are significantly more likely to reject the interruption. However, subjects varied in their responses to interruptions offering slightly positive gains, indicating the difficulty to estimate the benefit of interruption when its usefulness is ambiguous.

Figure 2 summarizes the acceptance rates of interruption as a function of the actual benefit of interruption and perceived partner type (person vs. computer). We divide the figure into three regions of interruption benefits: small losses (Region 1), small gains (Region 2), and large gains (Region 3). The analysis shows that for large gains (Region 3), changing the perceived partner type does not affect the likelihood that the interruption will be accepted. Similarly, for interruptions offering small losses (Region 1), the perceived partner type does not affect the interruption acceptance rate. In contrast, for interruptions offering small gains (Region 2), the acceptance rate is significantly larger if the perceived partner type is a person ($p = 3 \times 10^{-5}$, $\alpha = 0.001$). This result implies that when the benefit of interruption is straightforward, people do not care whether the initiator of the interruption is a person or a computer. However, for those cases in which the benefit of interruption is ambiguous, people prefer to accept those interruptions that originate from other people. This result suggests that agent designers need to consider the way they present interruptions to their users in cases where their perceived benefit is ambiguous. It also aligns with recent studies showing that mutual cooperation is more difficult to achieve in human-computer settings as compared to settings exclusively involving people [12].

Figure 3 shows the effect of interruption timing (the level of the game) on people's acceptance rates for interruptions of small losses and small gains (The interruption timing does not affect the acceptance rate for interruptions of large gains). We expected that interruptions occurring late in the game (i.e., with fewer number of turns left in the game) will be more likely to be accepted when they incur positive benefit, and rejected when incurring a loss. However, as shown by the Figure 3, as the game level increases, so does the acceptance rate for interruptions of both small losses (ABI -1) and small gains (ABI 1). There is a significant increase in the acceptance rate when game level increases from 3 to 5 (p=0.002, α=0.01) and from 5 to 7 (p$< 10^{-6}$, α=0.001).

One factor that may explain the correlation between the acceptance rate and the game level for interruptions of small gains and losses, is the cost of interruption to the subject. As shown in Figure 3, the cost of interruption to the subject (ABI_P) decreases as game level increases. Thus, for interruptions of small gains and losses, we found that the acceptance rate is negatively correlated with the cost of interruption to the principal. In addition, it was revealed that the benefit of the interruption to the principal (ABI_P) is a better predictor of the acceptance rate than ABI_A, the benefit of interruption to the agent (logistic regression $SE = 0.05, R^2 = 0.19, p < 0.001$). Thus, human subjects tend to overestimate their own benefit from interruptions as compared to the benefit for the agent. Consequently, the benefit of interruption to person may be weighted more in person decision making model than the benefit of the interruption, and people may be more likely to accept an interruption with low ABI_P among interruptions with identical benefit. Further study is required to determine whether these conjectures hold and better understand the correlation of acceptance rate with the cost of interruption to the person. This conjecture is supported by some subject responses to survey questions regarding their strategies for accepting interruptions. Answers include:

"If the agent was in the totally wrong direction and I had several moves left, I would allow the interruption. I always wanted the sure thing for myself."
"If the collaborator was way off in knowing and had enough moves to likely catch it after I told the location, I accepted. If it compromised my ability to get my goal, I declined."

Lastly, we emphasize that these results are a first step in understanding the human perception of interruptions in collaborative settings. Our goal is not to suggest that the computational strategies we derived for the interruption game are directly applicable for interruption management other domains. Rather, it is to suggest that effective interruption management needs to consider the collaborative benefit of interruption to both user and system, and to point system designers to the types of factors that people consider when they reason about interruptions. In future work, we plan to extend the study to better understand the effects of computational and cognitive complexity on people's interruption strategies, focusing on the possible role of trust and overestimation of costs.

VII. Conclusion and Future Work

We have presented a new computational model for interruption management based on computing the collaborative utility of an interruption to participants, and showed that this computation can be done efficiently and correctly. We evaluated the model in a specially designed human-computer collaborative setting. We showed that the actual benefit of interruptions to both computer agents and people is the major factor affecting the likelihood that people will accept interruption requests.

We also showed that our suggested methods for computing the perceived benefit of interruption align with people's behavior. In particular, we found that interruptions of high collaborative benefit are very likely to be accepted by people independent of the level of the game and the perceived partner type. For interruptions of small gains people were more likely to accept interruptions initiated by other people than by computers. This result implies that when the benefit of interruption is straightforward, people do not care whether the initiator of the interruption is a person or a computer. However, for those cases in which the benefit of interruption is ambiguous, people prefer to accept those interruptions that originate from other people.

Our empirical study yielded a number of insights about human decision making in the context of human-computer collaboration and fast-paced domains. Understanding the way that people make interruption decisions will enable the development of better mechanisms for initiating interruptions, focusing on the interruptions that are more likely to be accepted. These investigations will provide the foundation for building agents that collaborate with people efficiently without overburdening them.

References

[1] T. Babaian, B. Grosz, and S. Shieber, "A writer's collaborative assistant," in *Proceedings of the 7th international conference on Intelligent user interfaces*. ACM New York, NY, USA, 2002, pp. 7–14.

[2] P. Adamczyk and B. Bailey, "If not now, when?: the effects of interruption at different moments within task execution," *CHI*, 2004.

[3] B. Grosz, S. Kraus, S. Talman, B. Stossel, and M. Havlin, "The influence of social dependencies on decision-making: Initial investigations with a new game," *AAMAS*, 2004.

[4] P. Scerri, D. Pynadath, W. Johnson, P. Rosenbloom, M. Si, N. Schurr, and M. Tambe, "A prototype infrastructure for distributed robot-agent-person teams." in *AAMAS*, 2003.

[5] E. Horvitz and J. Apacible, "Learning and reasoning about interruption," *ICMI*, 2003.

[6] M. Fleming and R. Cohen, "A user modeling approach to determining system initiative in mixed-initiative ai systems," in *UM*, 2001, pp. 54–63.

[7] D. Avrahami, J. Fogarty, and S. E. Hudson, "Biases in human estimation of interruptibility: effects and implications for practice," in *CHI*, 2007.

[8] J. Gluck, A. Bunt, and J. McGrenere, "Matching attentional draw with utility in interruption," in *CHI*, 2007.

[9] A. Bunt, J. McGrenere, and C. Conati, "Understanding the utility of rationale in a mixed-initiative system for GUI customization," in *UM*, 2007.

[10] E. Kamar and B. J. Grosz, "Applying mdp approaches for estimating outcome of interaction in collaborative human-computer settings," *MSDM*, 2007.

[11] D. S. Bernstein, S. Zilberstein, and N. Immerman, "The complexity of decentralized control of markov decision processes," in *UAI*, 2000.

[12] J. Rilling, A. Sanfey, J. Aronson, L. Nystrom, and J. Cohen, "neural correlates of theory of mind within interpersonal interactions," pp. 1694–1703.