Crowdsourcing Human-Robot Interaction: New Methods and System Evaluation in a Public Environment

Cynthia Breazeal, Nick DePalma, Jeff Orkin
Massachusetts Institute of Technology
Sonia Chernova
Worcester Polytechnic Institute
and
Malte Jung
Stanford University

Supporting a wide variety of interaction styles across a diverse set of people is a significant challenge in human-robot interaction (HRI). In this work, we explore a data-driven approach that relies on crowdsourcing as a rich source of interactions that cover a wide repertoire of human behavior. We first develop an online game that requires two players to collaborate to solve a task. One player takes the role of a robot avatar and the other a human avatar, each with a different set of capabilities that must be coordinated to overcome challenges and complete the task. Leveraging the interaction data recorded in the online game, we present a novel technique for data-driven behavior generation using case-based planning for a real robot. We compare the resulting autonomous robot behavior against a Wizard of Oz base case condition in a real-world reproduction of the online game that was conducted at the Boston Museum of Science. Results of a post-study survey of participants indicate that the autonomous robot behavior matched the performance of the human-operated robot in several important measures. We examined video recordings of the real-world game to draw additional insights as to how the novice participants attempted to interact with the robot in a loosely structured collaborative task. We discovered that many of the collaborative interactions were generated in the moment and were driven by interpersonal dynamics, not necessarily by the task design. We explored using bids analysis as a meaningful construct to tap into affective qualities of HRI. An important lesson from this work is that in loosely structured collaborative tasks, robots need to be skillful in handling these in-the-moment interpersonal dynamics, as these dynamics have an important impact on the affective quality of the interaction for people. How such interactions dovetail with more task-oriented policies is an important area for future work, as we anticipate such interactions becoming commonplace in situations where personal robots perform loosely structured tasks in interaction with people in human living spaces.

Keywords: Human-robot teaming, crowdsourcing, bids analysis, humanoid dialog

Journal of Human-Robot Interaction, Vol. 2, No. 1, 2013, Pages 82-111. DOI 10.5898/JHRI.2.1.Breazeal

1. Introduction

Socially aware robots require a broad range of capabilities in order to work together with human beings as a team. Such robots must be able to detect and recognize action and intent (Gray, Breazeal, Berlin, Brooks, & Lieberman, 2005; Kelley et al., 2008), produce situationally appropriate actions (Breazeal, 1998; Mutlu, Shiwa, Kanda, Ishiguro, & Hagita, 2009), and engage with social back-channeling (Sidner & Lee, 2007) and dialogue (Kulyukin, 2004). Managing such a wide array of requirements is a major challenge to the machine learning and robotics community. We are inspired by efforts that focus on interactive data-driven approaches (Calinon & Billard, 2008; Knox & Stone, 2009; Thomaz & Breazeal, 2006), and we are interested in extending their methods to more loosely structured tasks. Crowdsourcing offers access to quick, easy examples and demonstrations that can be useful to the community. This work is an exploration into such a strategy.

Many robots find applications in museum guidance (Burgard et al., 1998), elder care (Graf, Hans, & Schraft, 2004), or, more recently, manufacturing (Wilcox, Nikolaidis, & Shah, 2012). Research into action and dialogue generation has also been conducted in the gaming community to develop new characters and non-playable characters (NPCs) for role-playing games (Kacmarcik, 2005; McNaughton, Schaeffer, Szafron, Parker, & Redford, 2004). Authoring behavior and language by hand is a labor-intensive process and limits the robots interactions to situations anticipated by an individual or a small team of designers. These systems are therefore typically restricted to engaging with the user under a small set of domains that are predetermined by the programmer at design time.

Interactive learning and data-driven models provide an alternative to hand-crafted techniques. These approaches usually involve a dataset and a set of machine learning techniques that utilize demonstrations or example traces. Successful techniques have been demonstrated in a number of applications, such as teleoperator dialogue (Gorin, Riccardi, & Wright, 1997; S. Singh, Litman, Kearns, & Walker, 2002), task learning (Chao, Cakmak, & Thomaz, 2010), and stunt helicopters (Abbeel, Coates, Quigley, & Ng, 2007).

Our work aims to tap into an incredibly broad population on the Internet to understand how large-scale, unscripted interactions recorded online can be leveraged to produce functionally valid behaviors that can be applied to physically embodied systems. We take a different perspective from more traditional machine learning methods in several important ways.

First, we chose to explore a loosely structured, high-level task that is highly complex in the number of states, state transitions, interaction styles, and teaming strategies that might be observed. For instance, in performing collaborative tasks people exhibit different team coordination behaviors, such as divide and conquer, mixed initiative, etc. Likewise, some people are more passive and let a robot take the lead more often, while others are more dominant and proactive in style. Traditional machine learning methods like Partially Observed Markov Decision Processes (POMDPs) and Reinforcement Learning (RL) do not readily scale to such high-dimensional state-action spaces. More importantly, it is unclear what the optimality or value criteria are for such a wide variety of social interactions. Second, we wanted to preserve the richness of the behavioral data, exploring methods that harness "the wisdom of the crowd," rather than abstracting what people tend to do in personal interactions into a trained policy of many interactions. This approach helps us to identify the important dimensions of human teaming behavior that might be very difficult to capture in a single trained policy. Third, we were curious if we could endow the robot with more natural and human-like behavior through generating functionally capable data-driven behaviors that are sourced from human behavior. Finally, we were also motivated by the on-demand nature of crowdsourcing as a means of reducing the overall development time for producing interesting and novel behaviors.

Hence, in this work, we were motivated and encouraged by the success of the Restaurant Game (Orkin & Roy, 2007), an interactive game driven by crowdsourced data that elegantly leverages a large number of interactions to produce functional agents. We make use of plan networks (Orkin &

Roy, 2009) as well as case-based planning (Kolodner, 1993; Spalzzi, 2001), a memory-based machine learning approach, to harness the wisdom of the crowd. We evaluate the resulting autonomous robot behaviors using a real-world reproduction of our online game environment.

A second aim of our work is to evaluate the affective quality of HRI that results from the use of a crowdsourced model in a real-world setting. We focus on the affective quality of the interactions because affect has been shown to be an important determinant of the quality of human-robot teamwork (Barsade & Gibson, 2007; Breazeal, Kidd, Thomaz, Hoffman, & Berlin, 2005; Jung, Lee, et al., 2012). Unfortunately, not much has been done to examine the role of affect in teamwork from a behavioral perspective; however, recent work by Jung, Chong, and Leifer (2012) showed that the application of a behavioral coding scheme to measure emotionally expressive behaviors in marital conflict could be used to identify behavior patterns that are predictive of performance in programming teams. We take inspiration from this work and provide qualitative affective analysis of the interactions we observed in the study.

2. Related Work

As mentioned above, our work is inspired by the Restaurant Game project (Orkin & Roy, 2007, 2009), in which data collected from thousands of players in an online game is used to create an automated data-driven behavior and dialogue authoring system. The Restaurant Game is a minimal investment multiplayer online (MIMO) game that enables users to log in to a virtual environment and take the role of one of two charactersa customer or a waiterat a restaurant. Players are randomly paired with another online player and can interact freely with each other and the objects in the environment. In addition to standard game controls, the users can maintain dialogue with each other and other simulated characters by typing freeform text. Logs of over 5,000 games were used by the authors to analyze this interactive human behavior and acquire contextualized models of language and behavior for collaborative activities. Following this approach, we have trained a model for human interaction that includes freeform, novice behavior.

The idea of data-driven dialogue systems has been heavily explored, and there are a number of successful applications outside of gaming and robotics. For example, Rieser and Lemon (2010) presented an automated system for information-seeking dialogue, such as a speech-driven navigation and track selection system for an MP3 player. This study utilized a dialogue corpus that was collected from 21 subjects for a total of over 17,000 dialogue turns.

The use of games as a data mining technique has also been shown to be highly successful in a number of applications. The term "Games With A Purpose," or GWAP, was coined by Luis von Ahn to describe games that address computational problems (vonAhn & Dabbish, 2008). The idea behind this approach is to make work fun by turning a scientific task, such as classification, into a multiplayer game, thereby harnessing the computational power of Internet users.

Another example of human-driven computation is the Open Mind Initiative, a world-wide research endeavor for developing intelligent software by leveraging human skills to train computers (Stork, 1999). This approach is based on a network of volunteers that provide answers to questions that computers cannot yet answer; for example, describing the content of an image.

Our work differs from previous approaches in that the data collection is performed in a different domain than the one in which it is applied, though there have been a few initiatives that have begun to leverage the crowd in the robotics community. Sorokin, Berenson, Srinivasa, and Hebert (2010) have made a novel interface for the crowd to annotate and inform a system to segment and build dense 3D reconstructions to assist a Rapidly-exploring Random Tree (RRT) grasp planner. We differ from that approach in this work in that we are interested less in the perceptual data and more in the strategies for interaction. Crick, Osentoski, Jay, and Jenkins (2011) also successfully enabled a robot to find its way through a maze using a crowd to handle direct control of its navigation. New

technologies have also been developed that enable crowds to have direct access to the robots, such as rosbridge (Crick, Jay, Osentoski, & Jenkins, 2012). We differ from that work as well in that we are interested in higher level interaction data that includes more goal-directed actions, rather than direct wheel control. We focus on behaviors that incorporate many different sources of information: Both dialogue from speech recognition systems and text-based dialogue, as well as state information used for task-oriented behaviors, were used to train the behavior of a real-world agent. This emphasis on crowdsourcing virtual interactions for real-world agents is directly applicable to robot behavior and learning models. We argue for a data-driven approach to HRI that can incorporate large corpora for new, interesting behaviors.

3. Technical Approach

In Section 3.1, we present an overview of the Mars Escape game that we developed for data collection. Section 3.2 presents the dataset we acquired using the game. Section 3.3 describes our data-driven system for generating the behavior of our robot. Section 4 describes the real-world variant of the Mars Escape task that we used to evaluate the performance of our data-driven method. Sections 4.2 and 4.3 present the two study conditions used in our evaluation, with results presented in Section 5. Sections 6.1 and 6.2 present a detailed analysis of the video-recorded real-world interactions in which we evaluate team cohesiveness and the use of nonverbal cues in the HRI.

3.1 Crowdsourcing Interaction Using an Online Game

As described in Chernova, DePalma, Morant, and Breazeal (2011), we began by designing an online multi-player game called Mars Escape, using the original code for the Restaurant Game (Orkin & Roy, 2007) in order to collect data about HRI in social, task-oriented settings. Mars Escape, shown in Figure 1, is a two-player game in which randomly paired players take on the roles of a human astronaut and a robot on Mars. We designed the game to model a collaborative search-and-retrieval task in which the players must locate and retrieve the following five items to successfully complete the mission: research journal, captured alien, canister, memory chip, and sample box. The object retrieval task is incorporated into the back-story of the game, in which players are told that the oxygen generator on their remote research station has failed, and that the pair must salvage the most critical items and return to the spaceship before oxygen supplies run out (10 min). Collaboration and communication between players are required to complete the entire task.

The attributes that the characters take on mimic real-world constraints (e.g., the robot cannot climb stairs due to its wheel base, and the human cannot walk on toxic sludge). Models for autonomous robot behavior are then trained from these data from the robots perspective, with a focus on modeling high-level collaborative behavior through high-level parameterized actions (e.g., `PICKUP`, `ANALYZE[Shelf]`). Our dataset does not attempt to exploit some of the other potential applications of crowdsourcing, such as low-level manipulation control or vision tasks. The role that the player takes on during this interaction is modeled using common machine learning techniques to build a behavior model that can be used in the real world.

In order to study a diverse set of collaborative behaviors, we designed the domain such that the retrieval of each of the five items required a different kind of interaction. Table 1 presents a list of the five inventory items in the game, how they can be accessed by each player, and how each object can be generalized to a broader class of problems. The item locations and character abilities were designed to balance the roles of both players and to ensure that collaboration and communication were necessary to perform the complete mission successfully.

At the end of the game, all participants were asked to complete a survey with the following eight questions. Their responses were recorded on a 5-point Likert scale from “strongly disagree” (1) to “strongly agree” (5):

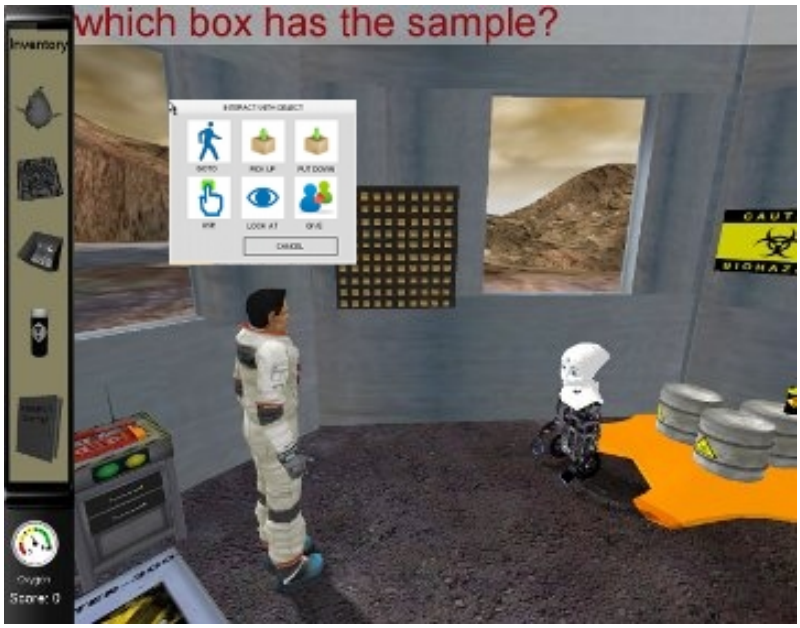


Figure 1. A screenshot of the Mars Escape game showing the action menu and dialogue between the players.

1. My overall game experience was enjoyable.
2. The other player's performance was an important contribution to the success of the team.
3. The human-robot team did well on the task.
4. The actions of the other player were rational.
5. The other player communicated in a clear manner.
6. The other player performed well as part of the team.
7. The other player's behavior was predictable.
8. The other player was controlled by a human.

3.2 Interaction Dataset

Within 3 months, we captured interaction data from 558 two-player games. Approximately 700 of the 1116 player logs were retained after excluding logs in which players exited the program unexpectedly or had not filled out the survey. Figure 2 shows an example interaction in which the astronaut (A) and the robot (R) retrieve the "alien" object from the elevator.

We observed that, as intended, the retrieval of different items provided different degrees of challenge to the players. The majority of players first picked up the items that were clearly visible and could be retrieved individuallythe canister and the journalleaving the retrieval of collaborative items until the end. We found that only 57% of the games contained successful retrieval of all five objects (86% collected three or more). When all five items were collected, in 75% of cases the sample box was the last item to be collected, in line with the design that made this the item that required the greatest degree of collaboration and communication (and therefore the greatest challenge) to retrieve. Of the games in which the participants only collected four items, 89% of them were missing the sample box.

These behavioral patterns directly affect the quantity and quality of the resulting interaction dataset. We observed that the number of examples of successful retrievals for some objects, partic-

Table 1: Description of the five objects players must obtain to successfully complete the game.

Item	Game Context	Generalization
Research Journal	Located on a stack of boxes. Reachable only by the astronaut by climbing the nearby objects.	can be performed only by one of the players
Captured Alien	Located in a cage on a raised platform. Reachable by either player after lowering the platform using wall mounted controls.	can be performed by either player
Canister	Located near a spill of toxic hazardous chemicals. Reachable by either player, but the astronaut loses 10 points for coming in contact with the chemicals.	one player is better suited than the other
Memory Chip	The appearance of this item is triggered by both players standing on a weight sensor at the same time. Once active, this item can be retrieved by either player.	requires action synchronization.
Sample Box	One of 100 identical boxes located on a high shelf. Astronaut must pick up and look at each box until the correct one is found. Robot can identify the exact box using its organic spectral scanner, but can not reach the box due to its height. Optimal solution is for the robot to scan the boxes and then tell the astronaut the sample's exact location.	requires coupled actions and dialog

ularly the sample box, was much smaller. Furthermore, there were significant temporal patterns in the data, such that some objects were commonly retrieved before others. In later sections we will discuss how these data patterns impact the behavior of the robot in real-world experiments.

R: "i'm too short, I cannot see what is on the elevator platform."
R: "can you bring it down?"
A: "i'll try to use the elevator buttons"
<astronaut *use* platform>
<robot *pick up* alien>
<astronaut *go to* crates>
A: "i think all the boxes are empty..."
R: "oh ok"
A: "do you know what this thing on the floor is???"

Figure 2. Sample retrieval transcript from online game.

3.3 Crowdsourced Case-Based Planning Algorithm

We introduce a novel algorithm for autonomous behavior generation that uses case-based planning (Spalzzi, 2001) and a plan network statistical model (Orkin & Roy, 2007) to identify patterns of behavior in noisy crowdsourced data. This algorithm is an extension of prior work demonstrating the use of case-based reasoning in this context (Chernova, DePalma, Morant, & Breazeal, 2011).

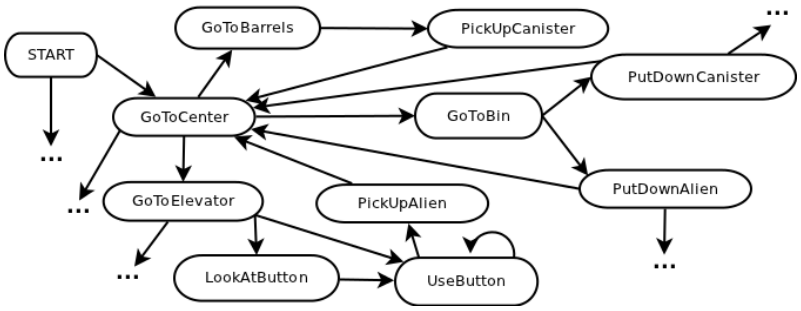


Figure 3. A plan network learned from the online human-human interaction corpus.

3.3.1 Plan Networks. In our analysis of player actions, we were interested in identifying statistically significant action transitions that represent typical player behavior. To model what a "typical" interaction might look like, and to ignore unusual or unexpected behaviors, we utilized the plan network representation (Orkin & Roy, 2007). A plan network is a statistical model that encodes context-sensitive expected patterns of behavior and language. Given a large corpus of data, a plan network provides a mechanism for analyzing action ordering, eliminating action sequences that are present in only a small number of training examples, and visualizing the graphical structure of the resulting action sequences. Figure 3 presents a subset of the plan network learned from the interaction data corpus. This model represents the most common behavior sequences observed in human players playing the role of the robot in the online Mars Escape game. While many players chose to deviate from the norm at some point in their gaming experience, these aberrant interactions were statistically eliminated in the analysis of the complete corpus.

Unlike other approaches, the goal of this work is to support a wide variety of interactions, exhibiting a wide range of behaviors but focusing on the goal directedness that plan networks offer. The goal of the plan networks is to encode action sequences and behaviors to support an incredibly large action space. Frequently observed action sequences allow the robot to be more goal directed (e.g., pick up toxic canister, take it to the bucket, and then put it in bucket). To help us achieve crowdsourcing with little annotation, novice behavior was not removed from the corpus, although the wide range of novice behavior can frequently misdirect models that rely on clear, consistent labels. Furthermore, we wanted to capture the expressivity of all of the data, not just the goal-directed behaviors, so that the robot and participant could reliably achieve the goal despite being "off track." Plan networks attempt to achieve these ends by modeling the transitions across the action space to model the diversity of the behavior. Our corpus (further discussed later in this section) captured over 70,000 unique, high-level states.

Plan networks, in short, are a representation of the action sequences executed in a given game. The graph begins with the START action and models every action from a set, $a \in A$. From the given online demonstrations, every action transition to a new action is modeled as a tuple (a_{prev}, a_{next}) for each agent. The corpus is meant to handle a large number of interaction scenarios rather than represent a concise representation of the task. The full graph is represented without generalization and does not throw out any "aberrant" interactions. So for each action in our log, n action tuples are generated (including the (A_{start}, a_{first}) tuple) for each agent. A frequency distribution is modeled for each action a_i for each agent to all of its consequents a_{i+1} in a given log. This concise representation represents all paths through the task from examples. Speech is a special, parameterized case, and when encountered, is modeled as a triplet with the utterance punctuating the actions.

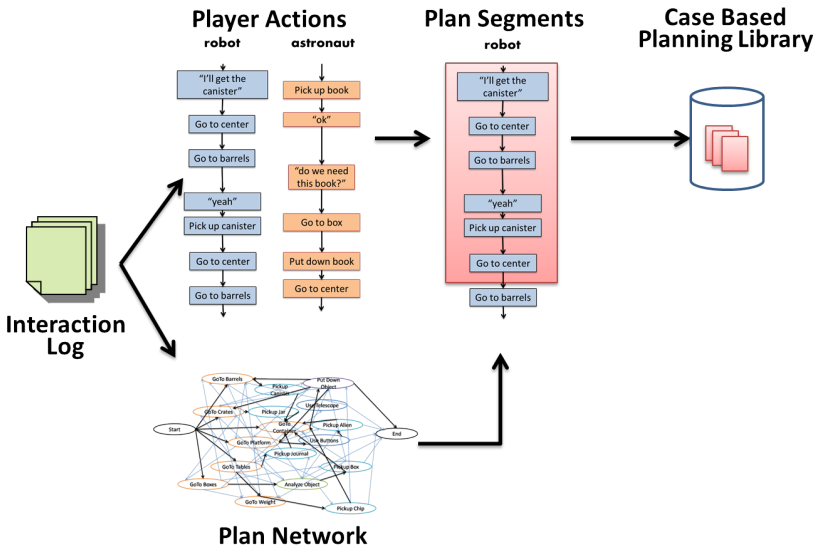


Figure 4. An overview of the generation of plan segments and the case-based planning library.

Once a plan network is trained, it is used to filter all robot actions observed when a human plays the role of a virtual robot in the online game. During this process, illustrated in Figure 4, each game log describing the temporal sequence of actions performed by a player is partitioned into sections called *plan segments*. A plan segment consists of a sequence of consecutive state-action pairs in which all action transitions are considered typical behaviors as determined by the plan network. Thus plan segments represent sequences of action that are commonly encountered in the dataset, and these segments form the case library of our case-based planning system. Note that in the current implementation, all dialogue between players is represented by a "Say" action, and thus spoken phrases can be directly encoded into a plan segment, as shown in Figure 4.

3.3.2 Case Based Planning. Case-based planning (CBP) is the reuse of past successful plans in order to solve new planning problems (Spalzzi, 2001). Once a case library is constructed, the robot utilizes its memory of example behaviors for autonomous decision making at runtime. As shown in Figure 5, during task execution the robot compares its current state to state-action pairs stored in the case library. The algorithm identifies the most closely matching state in the memory and retrieves the corresponding plan segment. The action sequence represented in the segment is then executed by the robot, with possible adaptations to the current environment, if necessary.

Within the planning library, the state of the robot is represented by 13 features: the previous robot action, the previous astronaut action, the previous robot spoken phrase, the previous astronaut spoken phrase, object held by astronaut, object held by robot, and the area location (e.g., center, near shelf, near barrels, etc.) of the astronaut, robot, journal, alien, chip, canister, and sample box.

In our case, case-based planning utilizes the plan networks to generate behavior from the state of the world. Using plan network graph G, each state query from the world requests a list of actions of size n from the database, $[a_1 \dots a_n]$. Using state s, a list of potential actions of some arbitrary size m is requested from the CBP database $[a_1 \dots a_m]$ to initialize the search on G. A tuple is generated (a_{prev}, a_{next}), where a_{prev} is the last action the robot took from the state feature and a_{next} is the

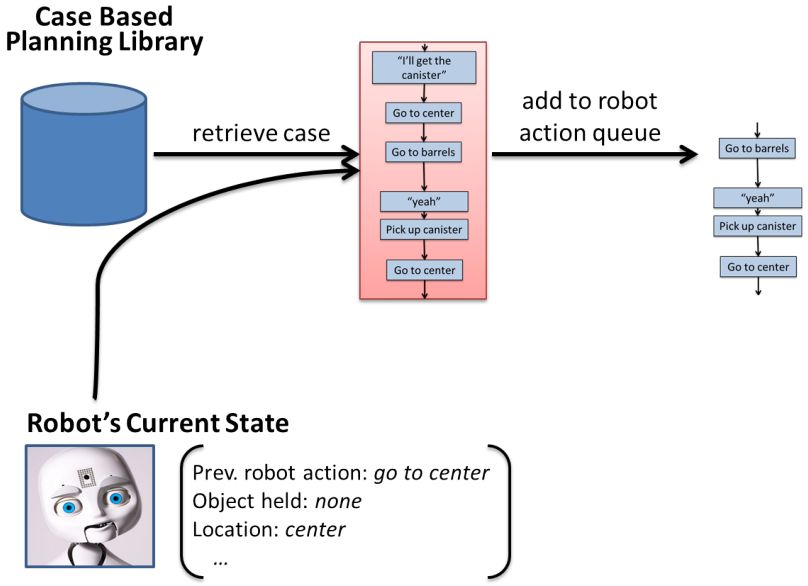


Figure 5. The autonomous action selection process.

first action the robot should take next according to some set of logs of size m. The search on G proceeds by popping each new action from the list of m actions from the database and verifying that it is a valid course of action according to the graph until a valid list of actions is generated. Speech was treated as it was in the plan network construction: by executing it and ignoring it as a major action the robot will take and querying the graph for the triplet once a third action is popped off the CBP query results.

The list of m actions from the database was selected by calculating distance for each feature. Distance is normalized for each feature; in other words, a distance of 0 is an exact match and a distance of 1 is a "no-match." For categorical data, the inverse Kronecker delta was used, and for speech, the Jaccard similarity coefficient was used. The distance is then calculated to be the weighted sum of squares: $state\_dist(s_1, s_2) = \sum_f w_f d_f^2$ for each feature weight w_f and feature distance d_f in s_1 and s_2. Weights are likewise interpreted, with smaller weights representing a more important match. The closest m matches are selected, and the next action from each log is returned for each of the closest m states.

We discretize the representation of continuous actions, such as movement across the space, to reduce the number of cases and to generalize across similar actions. Specifically, movement actions result in changes to the players state only when the player moves from one area to another and not when moving within the same area. When an object is picked up, its location is marked as being with the agent that picked it up (e.g., `ASTRONAUT`) or in a semantic location (e.g., `BUCKET`). Significant sensing errors (e.g., occlusions that made the object disappear from view for an extended period of time) were corrected by a human operator.

Based on the representation described above, our interaction dataset consisted of over 70,000 unique states. When comparing the robots current world state to the data corpus, similarity between the query and states in the corpus was calculated based on a weighted sum of differences between features. Feature weights were tuned for each feature based on the accuracy of the measure the

feature represented. For example, the weighting for all object locations was high because we were able to track this information with high accuracy, whereas a low weight value was used for the "previous astronaut spoken phrase" feature because speech recognition was noisy.

4. Real-World Mars Escape Experiment

To test our data-driven behavior models, we created a real-life version of the Mars Escape game at the Boston Museum of Science. Our goal was to evaluate whether data gathered in an online setting could be used to successfully produce robust, diverse, and natural interactions with participants of different technical backgrounds, ages, and experience.

4.1 The Task and Set

The experimental setup was designed to imitate the virtual game as closely as possible. Five objects were placed in similar placements to their in-game counterparts, including tall shelves that held the journal, a working elevator that participants had to lower by pressing a button to retrieve the alien, toxic barrels that held the canister, a working lockbox for the computer chip that required the human and robot to stand on a scale to open, and a chest of 100 drawers of which only one contained the sample box. The set also contained a number of additional props, such as empty crates and tools. Figure 6(a) shows the robot reaching for the button that activates the elevator while the human participant retrieves the journal from the shelf.

Museum visitors were recruited one at a time at random to perform the Mars Escape task with our MDS robot *Nexi* (Figure 6[a]). The MDS is a mobile upper-torso humanoid that combines a wheeled mobile base with a socially expressive face and two dexterous hands that allow it to grasp and lift objects. Auditory inputs support a microphone array for sound localization as well as a dedicated channel for speech recognition via a wearable microphone.

Due to the complexity of the search and retrieval task, an off-board Vicon MX motion tracking system was used to augment the robots onboard sensors. The Vicon system utilizes lightweight, low-impact passive reflective markers. This tracking system enabled the robot to monitor important events and provide real-time feedback with low latency. The Vicon system was used to track the position of the target objects, the human, the robot, and specific locations. Participants were asked to wear a hat and glove with Vicon markers. This enabled the robot to monitor the persons location and what object he or she was holding. Note that although no nonverbal cues could be captured in the online game, we decided to program the real robot to look at the person when he or she was near the robot. We expected that it would be important for the robot to demonstrate visual awareness of the person as they carried out the tasks. The robot also had Vicon markers on its body to help it track its own movement in the environment to aid in navigation. The robot also had Vicon markers on its hand to provide real-time feedback to assist in autonomous manipulation during picking up items and putting them in the bucket.

Finally, the environment was enclosed by a curtain from the rest of the exhibit hall, preventing participants from observing others performing the task. Therefore, participants had no familiarity with the robot, no knowledge of the floor layout, and no knowledge of the locations of the target objects before beginning the study.

Three experimental staff members were responsible for running the experiment. One staff person was in charge of greeting each participant and introducing him or her to the components of the study: the glove, the hat, the microphone, and the filling out of paperwork. Two other staff members were responsible for the hardware: making sure that that the Vicon system was tuned properly, that the batteries were sufficiently charged, that the microphone and video cameras were on, and that the study environment was reset once the participant was out of the curtained area.

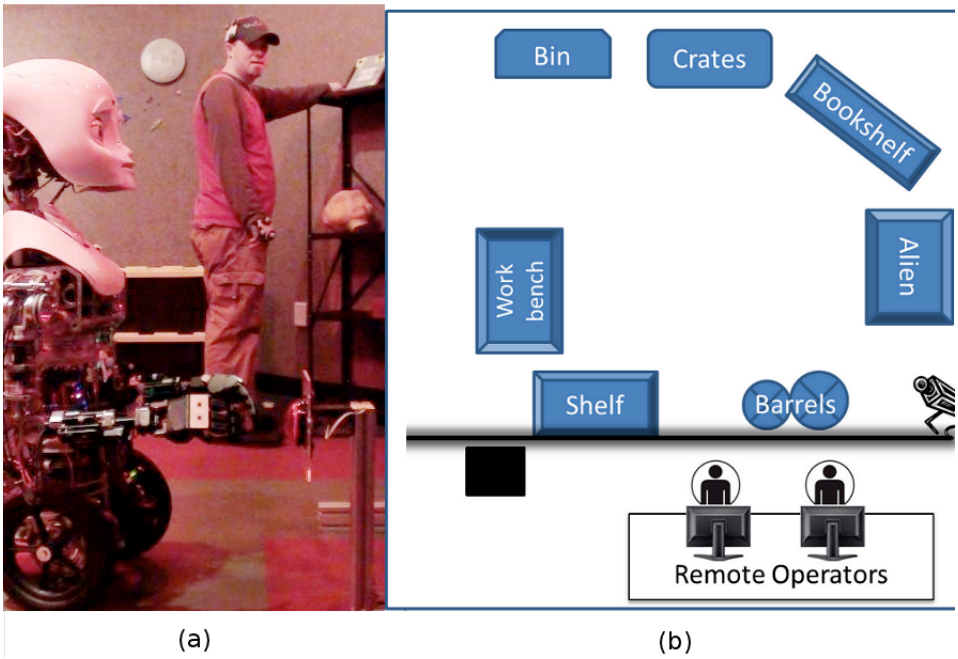


Figure 6. (a) The MDS robot platform. (b) The robot and a participant performing the study at the Boston Museum of Science.

One of the hardware staff members was responsible for teleoperating the robots arm to reliably grasp objects and to ensure that the Vicon-based "state" estimate was accurate. In general, the Vicon system estimated the state very well, but in the event that certain objects were not tracked or the participant picked up an object that was unseen, the operator could override the state and ensure a clean state signal. The second hardware staff member was responsible for ensuring that the robot was not injuring itself, that the study condition was proceeding without hardware failure, and that an override policy was followed if necessary (see below).

4.2 Wizard of Oz Condition

Many traditional HRI systems rely on a set of scripted rules, as well as a limited set of phrases, to determine the correct behavior in an interaction (Lee & Makatchev, 2009). In this study, we used a Wizard of Oz (WoZ) condition as a means of simulating such behavior and setting a baseline against which the CBP system could be compared. During the WoZ experiment, an operator controlled the robot from behind the scenes, unseen by the participant. The operator transitioned the robot through the sequence of actions shown in Figure 7.

Override Policy. In situations when an action or interaction failed, such as when a human participant did not understand the robots speech or when the robot dropped an object during a PICKUP action, the operator was allowed to repeat the previous action up to three times using an override interface. Overrides were used frequently to assist in manipulation or to repeat an utterance that the participant did not hear. The operator was instructed to use the override options only if the original script had failed to execute the appropriate action or the lower level control did not succeed.

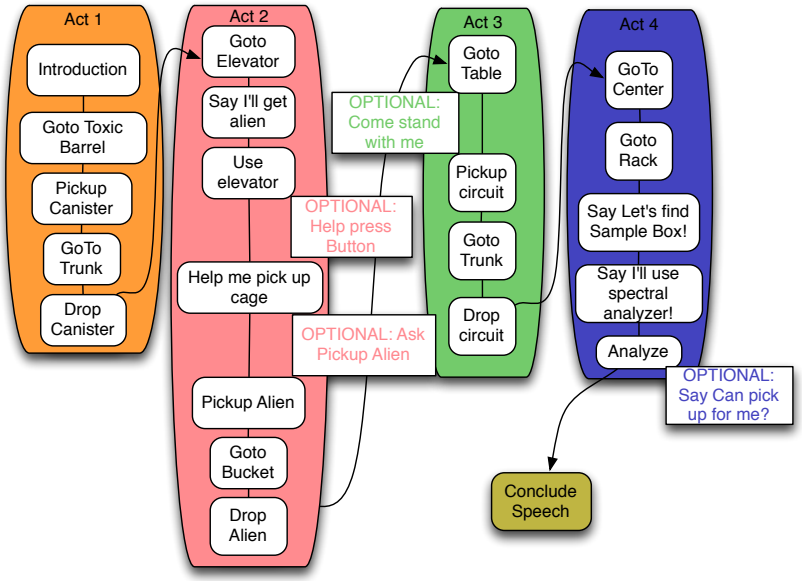


Figure 7. Wizard of Oz script

4.3 Crowdsourced CBP Condition

In the Crowdsourced CBP Condition, our case-based planning system was used to control the robots behavior. It is important to note that our goal was to evaluate the robots decision-making ability; namely, what it decided to say and do in response to the changing task state. Our goal was not to evaluate the robustness of the robots behavior to sensor and actuator uncertainty and failures. Hence, overrides were permissible if they addressed the robustness issue only. In fact, due to the size of the CBP model, the remote operator could not directly interfere with the generation of a new action–this was the responsibility of the robot.

Override Policy. Overrides were relatively unused and served only to facilitate the interaction in light of sensor or actuator uncertainty when necessitated. A strict override policy was adhered to in order to maintain two objectives: (1) the robot generated its own behavior in this condition, and (2) the selected action could be completed. The remote operator could issue an override if the robot could not successfully complete an action. For instance, this happened the most often when the robot was unable to successfully pick up an object or if it accidentally dropped an object. In these cases, the remote operator could intervene to help the robot complete the PICKUP action. The remote operator could also help to complete a failed action by triggering it again. For instance, if the person did not hear what the robot said, the operator could re-trigger the utterance. This often happened at the chest of drawers, where the robot would tell the human which box the chip was in, but the person did not always respond the first time. If the Vicon system hit a glitch and the robots environmental state was not properly updated, the operator could also override the environmental state so that the robot would know if it completed an action successfully.

In section 5.2.1, we discuss the robustness of the data-driven system in light of the above policy.

4.4 Protocol

A total of 54 participants were recruited to perform the museum study: 30 male and 24 female, with ages ranging from 11 to 36. One person had previously played the Mars Escape game online. Of the 54 participants, 23 were eliminated due to technical failures (i.e., the computers crashed, the network crashed, or the robot ran out of battery or physically broke in some way). The remaining 31 participants were assigned to one of two possible conditions: 16 participants were assigned to the WoZ condition, and 15 participants were assigned to the crowdsourced CBP condition[1].

The experiment began with the participant receiving basic instructions about the task and capabilities of the robot:

> *"Behind this curtain is a mockup of a research station on Mars. You will play the role of an astronaut. The researchers that used to work here have left the research station, but they forgot behind several important items. You and the robot were sent here to retrieve the items, which are shown on this list. Together, you must find as many items as you can and collect them all into the green bucket. The robot can help you accomplish your task – it can hear what you say, move around, pick up objects, and use its sensors to look for a sample box. Since no one has been here to maintain the research equipment, some of the barrels containing toxic chemicals have begun to leak. The robot has performed a preliminary sweep and marked the toxic area with red caution tape; you should avoid entering that area if possible."*

After receiving these instructions, participants entered the study area and were greeted by the robot using a standard greeting: "Hello. Let's find the five items on our list." After the greeting, the robot began the task. For each participant, the study continued until either the team retrieved all five objects, the robot failed (e.g., low battery), or the study participant requested to stop the session (none of the participants did). Data from incomplete trials were discarded. Following the study, participants were asked to fill out a short questionnaire consisting of the same eight questions as in the online Mars Escape game.

5. Results and Evaluation

In our system-level evaluation, we were particularly interested in exploring two questions. First, how well do online crowdsourced interactions transfer to the real world, where the robot must deal with sensing, navigation, and manipulation that are all very different from the virtual counterpart? Second, people are likely to have a very different experience when collaborating with a real robot compared to playing an online game with another human. How does the quality of the experience differ between the real-world task and the virtual counterpart?

This section presents an overall systems evaluation of the crowdsourced CBP system as it relates to the WoZ baseline, and when appropriate, to the online game. Due to the open and loosely structured nature of the task, participant behavior did indeed vary greatly from person to person, and even within the same task. In Section 5.1, we first characterize the macro-interaction patterns we observed from our video. This characterization strongly motivates the need for behavior generation algorithms that are robust to the diversity of human behavior and interaction styles, discussed in Section 5.2. Last, in Section 5.3, we examine peoples relative subjective experience in working with each of the three platforms: the online game (human with human), the WoZ baseline (human with human following script), and crowdsourced CBP autonomy (human with system).

[1] A sample video is available at http://www.vimeo.com/24546560

5.1 Macro Observations: Differences in Interaction Styles

To gain insight into the quality of the interactions that resulted from the crowdsourced autonomy model, we performed a preliminary macro-behavioral analysis to characterize the levels and modes of collaborative engagement employed by the people in the study. We coded 13 of the original 18 crowdsourced autonomy condition videos. (Five were lost due to technical failure of the video camera.) For each of the 13 videos, we isolated the segment between the introduction made by the robot that indicated the start of the game and the statement made by the robot that indicated the end of the game, "Great, were done. Lets get out of here." The total time to finish the task was defined as the time between beginning of the robots "start" utterance and the end of the "finished" utterance. We coded for the number of speaker turns and engagement level of the human participant (interacting with the robot, monitoring the robot, being inattentive to the robot). Due to the highly exploratory nature of our video coding, we omitted an analysis of coder reliability. Additionally, we extracted the number of utterances the robot made from the automatically generated log files. At the time of coding the videos, the coder was blind to the outcomes of the study.

Overall, the 13 analyzed human-robot dyads took between 3.9 and 10.8 min to solve the Mars Escape task ($\bar{x}$ = 7.2, s = 1.9). During that time the robot made between 3 and 50 utterances ($\bar{x}$ = 30.5, s = 15.6). The number of speaker turns varied between 1 and 73 ($\bar{x}$ = 37.4, s = 21.8).

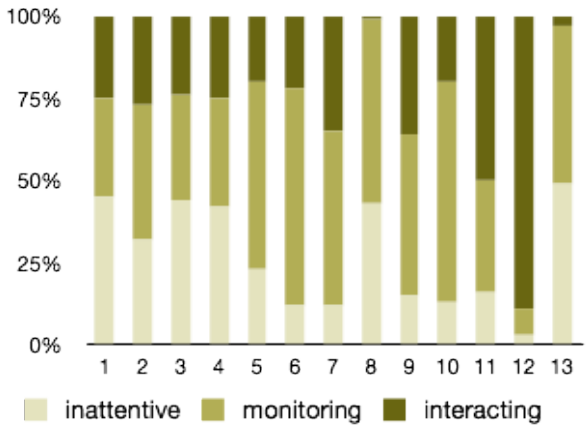


Figure 8. Distribution of engagement levels across participants. *Inattentive* often occurs when the human is following a divide-and-conquer approach. *Interactive* occurs when people are actively trying to coordinate with the robot, often in a mixed-initiative manner. *Monitoring* occurs when people watch the robot but do not try to advance the task.

On a general level, we noticed a high variability in peoples interactive engagement. Styles ranged dramatically, from a participant who solved the task by only interacting twice with the robot ($< 1\%$ of the time), to another participant who almost continuously engaged the robot in verbal interaction (almost 89% of the time). On average, participants spent 29% of the time interacting with the robot, 44% monitoring it, and 27% inattentive to it. The distribution of engagement levels is displayed in Figure 8.

We were surprised to find that people employed various ways of interacting with the robot that we had not anticipated (see Figure 9). For example, in an interaction style we called *cyber-gloving*, several participants mistook the glove they were wearing as a way to control the robots arm and

Style	Description
Cyber-gloving	Interaction style during which person mistakes positioning-glove as a data-glove and tries to teleoperate the robot through gesture. This often happens around the activity of the robot trying to pick up the canister.
Directing (Remote Control Style)	Interaction style during which person gives minute commands to the robot, directing every movement as if they were remote controlling the robot. Directions are not at the task level but at the movement level. Here is an example sequence of commands (CBP 21) "Raise your right hand" "Move forward!" "More!" "Extend your right hand!" "Extend your elbow!" "Let go with your right hand!" "Turn your right hand so your palm is down." "Turn your thumb down!" "Point your thumb down?"
Deferring Leadership	Interaction style during which person gives up leadership to the robot. For example in one interaction the person kept asking, "What should I do now?" throughout the entire duration of the game.

Figure 9. Some of the styles that people used in interacting with the robot.

hand. These participants tried to teleoperate the robot during various short periods, mainly when the robot was trying to pick up an item. Some people tried to command the robot at an action-by-action level as if to direct its every move, not at the task level but at the limb level. Others would defer leadership to the robot, asking it for advice on what to do next. Almost all participants spent a large amount of the time (44% on average) just watching the robot (monitoring), not contributing to the advancement of the game at all. Finally, some people actively coordinated with the robot in a mixed-initiative manner, making requests of the robot and responding to its requests (see Sections 6.1 and 6.2). Importantly, people often alternated between these strategies throughout the task.

5.2 Robustness of Crowdsourced CBP Autonomy

Given these observations, we think it is critical for system developers to consider that people are likely to employ a wide variety of interaction styles with robots, and it is likely that these strategies will also vary over time. This consideration also motivates the issue of robustness. We evaluate the ability of the crowdsourced CBP autonomy method with respect to overall robustness in two ways: (1) by examining the frequency of overrides by a human operator, and (2) by looking at the amount of variability in time to perform each object retrieval in comparison with the WoZ baseline.

5.2.1 Override Analysis. As part of our technical evaluation of the systems ability to generate robust behavior for a wide range of people and interactions, we conducted an evaluation of the interventions to the behavior system. Autonomous behavior selection by the robot occasionally required manual intervention via overrides by a human operator to keep the robot on track and

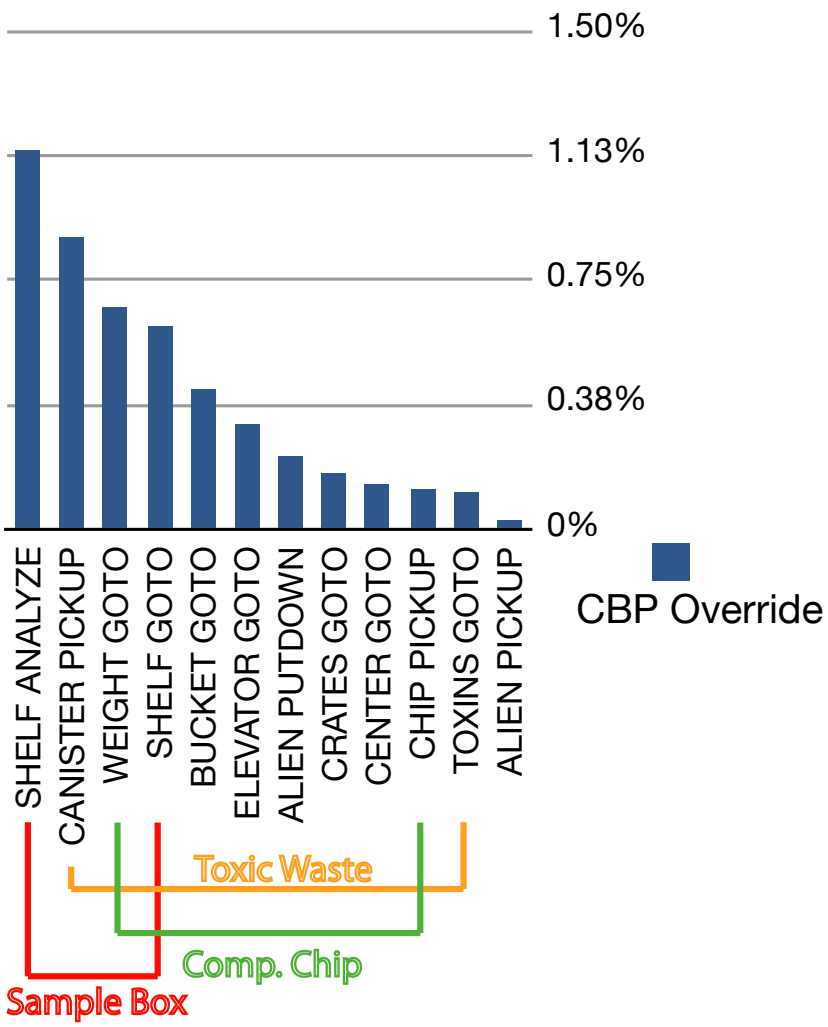


Figure 10. Distribution of overridden actions for our algorithm: expected frequency of override use per participant

maintain the expectations of the participant (as discussed in Section 4.3).

Figure 10 shows the distribution of all overrides performed. The top three actions that were overridden were `ANALYZE(SHELF)`, `PICKUP(CANISTER)`, `GOTO(WEIGHT)`, of which the most frequently used override was to analyze the shelf. By clustering the overridden actions by subtask (sample box, toxic waste, computer chip), we gain insight into why they were needed. Cross referencing this categorization with Table 1, it appears that approaching the shelf and using the robots "x-ray" to find the object in one of 100 bins (making it easier for the human to find the object) is critical to quickly solving this puzzle. The high-level `ANALYZE` was used to inform the person which bin contained the sample box. Interestingly, it frequently took many repetitions of this action before the participant followed the robots lead and opened the sample box. This unexpected behavior made the most challenging collaboration even more challenging. As mentioned before, only 57% of the participants in the online game collected all five objects, with recovery of the sample box being the least observed sequence of action. We were surprised that people were prone to not respond to the robot in this situation, given that the robot was actually giving very useful information to accomplish the task.

The toxic waste cluster override focused on manipulation, `PICKUP(CANISTER)`, which is emphasized by how rarely `GOTO(TOXINS)` was used. The robot was the only team member allowed to be present in the toxic sludge, which made the canister the object it most frequently picked up. Thus, overrides here were used to help the robot manipulate the canister.

Finally, the computer chip task was the second most collaborative challenge in Mars Escape, requiring both the human participant and the robot to stand in the same area before a mechanical box opened to reveal the computer chip. This was a challenging task in itself and required a timing element that was not well represented in the data. Future work on building a better joint-action model is further discussed in Section 7.

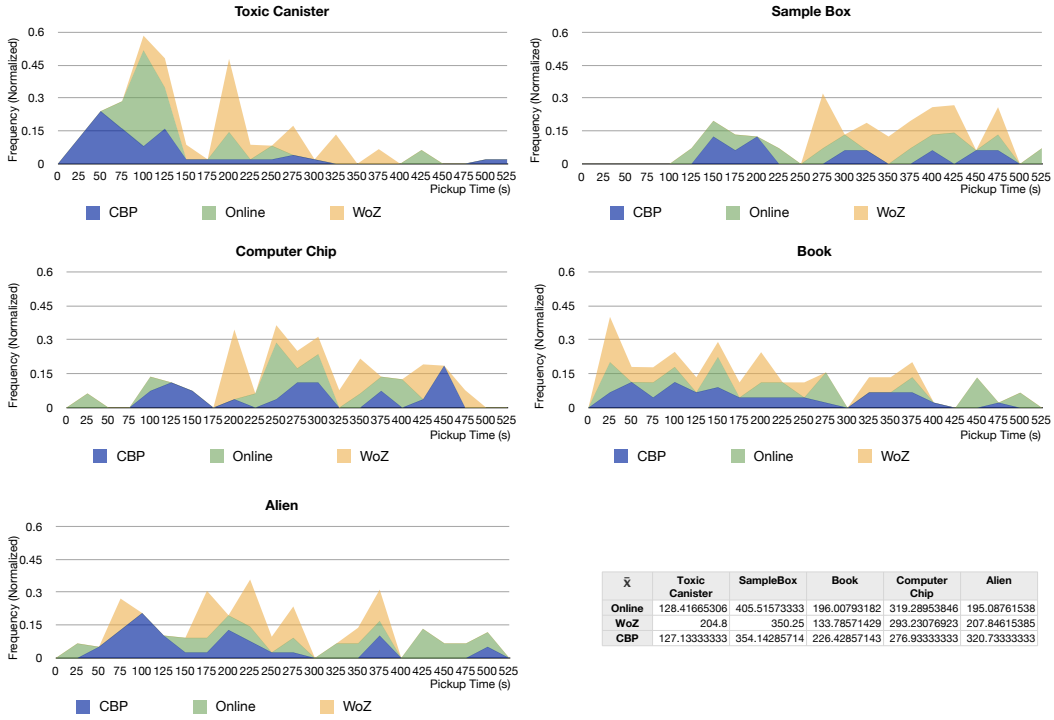
Finally, there were two unexpected entries into the override distribution: going to the center and going to the crates. The operators had discussed what to do when breaking the robot out of a looping behavior. The common strategy was to send the robot to the center of the room or to the crates to break it out of the loop. In theory, oscillating behavior should be minimized, and this issue should also be a focus of future work.

The relatively low frequency of the overrides (1% of the time at most for each participant) is a positive sign, but more often than not when the participant failed to jointly act on the collaborative task, an override was issued that was meant to encourage the joint action; for example, analyzing the shelf to inform the participant where the sample box was or moving towards the table to be jointly present. Hence, the overrides often happened when there was a breakdown in either the human or the robot responding to the bids of the other to engage in some way. This matter is something we will explore further in Section 6.2.

Overall, our new case-based planning approach to crowdsourced autonomy resulted in 84% of all robot actions being completed autonomously over the course of the study, a significant improvement over our previous result of 64% by our case-based reasoning system (Chernova, DePalma, Morant, & Breazeal, 2011). Given the difficulty of this task, we argue that this is a promising result, with the potential for generating robust robot behavior for a wide variety of people and interaction styles.

5.2.2 Timing Analysis The goal of this project was to support a wide variety of interaction styles. The WoZ script was rigid and more or less represented a finite-state machine. In contrast, the CBP system was meant to model the online interactions that were much more fluid in the moment of handling collaborative behavior. As in Chernova, DePalma, and Breazeal (2011), the object discovery times (Figure 11) of our crowdsourced CBP system closely resemble that of the game.

Looking strictly at the WoZ timing data, the clusters of pickup times for various items reflect



x̄	Toxic Canister	SampleBox	Book	Computer Chip	Alien
Online	128.41665306	405.51573333	196.00793182	319.28953846	195.08761538
WoZ	204.8	350.25	133.78571429	293.23076923	207.84615385
CBP	127.13333333	354.14285714	226.42857143	276.93333333	320.73333333

Figure 11. Frequency diagram of time to find and pick up representative objects in the game world and in the real-world CBP and WoZ conditions (measured in seconds).

the rigidity of the script. While the robot was responsible for the toxic canister and was needed for the computer chip, the human participant had to retrieve the book and could optionally pick up the alien. One of the interesting artifacts of the WoZ timing data was that soon after the toxic canister was retrieved, the computer chip was then frequently found and retrieved. The human participant could, in any order, retrieve the book and alien. The rather uniform timing distributions in the remaining subtasks show that the robot was more flexible to various interaction styles and also more flexible in finding objects regardless of order than the more rigid methods. In almost all cases, the sample box was retrieved last, since it could not be found without the help of the robot. The sample means reflect these observations as well. The latency in the pickup for the toxic canister was due mainly to the robot requiring a large amount of time to navigate to and pick up the canister.

In many ways, the expectation of the CBP behavior system is to track as closely to the online data as possible. Human participants were more likely to retrieve the easiest, most visible object first, which was the canister. This is reflected in the online data and subsequently the CBP algorithm. While our algorithm was at a general level equally likely to solve any of the subtasks first, it chose to retrieve the most visible object, reflecting how the online data informs its behavior. We argue that because of the relative uniformity of the timing distributions in the rest of the subtasks, the robot is indeed more flexible to various interaction styles and to finding objects in any order than more rigid methods. Finally, the computer chip and the sample box were generally not found and retrieved until after the toxic canister, reflecting the collaborative demands of these subtasks. The sample box took significantly longer to find and retrieve, reflecting both the subtasks collaborative demands and the

challenge of solving this task as a team.

The timing data also reflects the challenge of collaboration. In general, the more collaborative the subtask, the later the item tended to be found. The book and the canister frequently were found and retrieved within the first minute of the experiment, while the sample box, the computer chip, and the alien were generally found after about two minutes of interaction time. The variance in the distribution times of the retrieval tasks may reflect a more flexible behavior systemcertainly more flexible than the WoZ, which was highly rigid and structured. From a task-completion perspective, the task was modeled well using an instance-based learning method, but the subtleties of the interaction are interesting and will be discussed in Section 6.2.

5.3 Questionnaire Results

To assess the participants subjective assessment of each of the three platforms, we compared the post-study survey results of the online game, WoZ condition, and crowdsourced CBP condition. Note that this comparison is not intended to serve as a hypothesis-driven evaluation where specific, carefully controlled manipulations are performed to seek statistical significance. Rather, we are looking for broad trends in how the participants responded to the quality of interaction for each system. The WoZ condition serves as a baseline for the autonomy condition.

Note that each system has a different overall feel. The most flexible and responsive teaming is expected to occur in the online game, since two human players can engage in flexible and spontaneous interactions. One would expect this condition to receive the highest subjective ratings by participants.

The WoZ condition is more structured and rigid, given that the operators followed a script to remove any unwanted variability across performances. The operators could adjust the timing for when to trigger the next action in the script and had some flexibility to perform overrides to keep the task progressing. In general, however, the interaction resulted in the robot engaging in a divide and conquer style of coordination. For the two objects that required explicit coordination between robot and human, the operator initiated these interactions as part of the script. One would expect to see the WoZ system outperforming the crowdsourced CBP system on subjective interaction measures, given that the robot is controlled by a human, and therefore its behavior should be predictable and rational.

The CBP system can dynamically respond to the participants actions and hence has the potential to support more flexible teaming. The question here is how people subjectively rate the crowdsourced CBP system compared to the WoZ baseline. If there are ratings where CBP trends close to or is even more positive than WoZ, then that is interesting to note.

The survey contained the eight questions listed in Section 3.1. Figure 12 compares the survey responses, which have been collapsed to a 3-point Likert scale for ease of evaluation (agree, neutral, and disagree). The p-values were calculated using a Fishers exact test and pairwise t-tests, both giving similar results. The survey responses indicate that the vast majority of participants enjoyed the game experience (a), both online and in the real-world domain. Also, participants accurately predicted that the CBP condition was least likely to be under human control (h). We were interested to note that over 20% of online players were not certain that their human partner was indeed human.

Participants across all three conditions responded that the robot contributed to the success of the team (b), a critical positive result for our study where human-robot collaboration and teamwork are a central focus. In the real-world scenario, the WoZ script was rated more favorably than the autonomous CBP condition with respect to the rationality of robot actions (d), clear communication (e) and predictability (g). This result is partially due to the fact that behavior loops (performing the same action over and over, which can frequently happen when the state doesnt change after an action) and out-of-context speech occurred much more frequently through CBP autonomous

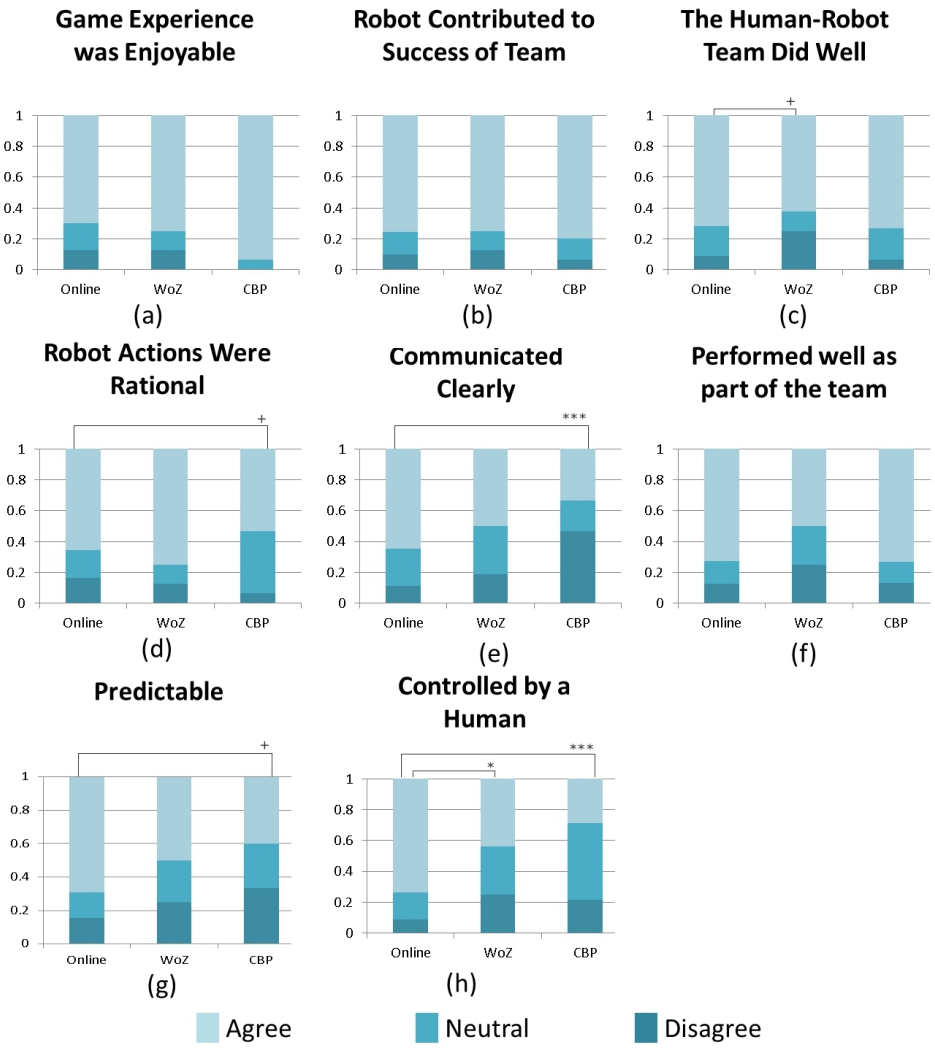


Figure 12. Comparison of survey responses for the online, Wizard of Oz and case-based planning conditions. **+** $p <= 0.1$, ***** $p <= 0.05$, ****** $p <= 0.01$, ******* $p <= 0.001$

behavior generation. However, there were additional in-the-moment, socially driven interactions that could be observed in the video that we also believe significantly contributed to these ratings. Many of these socially driven interactions were based on nonverbal social cues. We distinguish these from task-driven interactions. We discuss this in more detail in section 6.2.

Interestingly, the CBP condition outperformed WoZ with respect to the robot behaving as part of a team (c and f), suggesting that in the autonomous CBP condition, the robot was more reactive to the actions of the participant compared to the scripted scenario. This suggests that the CBP system was able to successfully address certain aspects of task-driven performance. Indeed, our overall system analysis of the time required to collect each item shows that our crowdsourced CBP approach performs favorably even compared to the online system. Hence, from an objective time-on-task measure, the CBP system was able to capture relevant task-based information from crowdsourced interaction data.

6. Discussion

Overall, these findings indicate that crowdsourced CBP from online interaction data holds a lot of potential for developing behaviors for human-robot interaction and collaboration. Along certain task-performance measures, the CBP system performed favorably, although there are many opportunities for improvement beyond this early exploration.

With CBP and WoZ, we employed two highly different approaches in driving a robots behavior. While all participants in both conditions successfully completed the task of finding the five items, we expected the approaches to lead to differences in interaction that people would notice. However, when we asked people about their perceptions of the real-world robot or the robot game avatar, we did not find many substantial differences (see Figure 12).

This result was both unexpected and interesting to us. We anticipated that people would favorably notice the richness of the crowdsourced robot behavior over the rigidly scripted actions of the WoZ-driven behavior. Why did people react so similarly across conditions? We observed that participants varied a great deal in (1) what prompted their interactions with the robot, and (2) how they reacted emotionally to the interactions with the robot. This high variability in reactions and interaction styles might have outweighed the perceptual differences between the systems.

In order to gain further insights that could inform future system development, we decided to examine the video recordings of the CBP interactions more closely. (Given that the WoZ robot was restricted to following a script, these videos would be less informative.) Our aim was to identify patterns that we had not considered in designing the crowdsourcing technique, but that could have a powerful impact on peoples reactive behavior and perceptions. In the following section, we share our insights from an exploratory investigation into the two areas introduced above.

6.1 Where Did Collaboration Really Happen?

We designed the task with explicit collaboration opportunities in mind, and we expected the majority of collaborative activity to happen there. For example, the task to retrieve the sample box and the task to retrieve the memory chip were designed so that they could only be solved in collaboration. We therefore expected collaborative interactions to happen around these two tasks, and we otherwise expected the human-robot pairs to take a divide-and-conquer approach in solving the rest of the Mars Escape scenario.

However, upon viewing the video record, we did not find interactive collaborations where we expected them to happen based on task demands. In fact, the task did not seem to be the defining criteria in structuring collaborative interactions. For example, the memory chip task (requiring both the robot and the human to stand on the scale simultaneously) was often solved when at some point

both the robot and the human accidentally stood on the scale. The person then usually just took the item without any interaction with the robot and put it in the bucket.

Moreover, we found the human-robot pairs to fluidly move in and out of collaborative segments. The need for collaboration seemed to emerge spontaneously out of the specific situation, rather than emerging out of pre-determinable task demands. Many, if not most, of the collaborative interactions therefore were organized around activities that were not intended by us to be collaborative. For example, several participants started to engage in collaborative interactions around placing a recently found object in the return bucket. This interaction could be as short as asking two questions, "Should I put it in the bucket?" ... "Shall I give it to you to put it in the bucket?" or it could be an extended interaction during which the participant seemingly guided the robot step by step in dropping the item into the bucket.

For instance, Figure 13 depicts a transcript of a 14-s interaction of a positive collaborative segment comprised of spoken turns and nonverbal cues during the retrieval of the canister. This item does not require explicit coordination to retrieve, but the person collaborated with the robot anyway. The robot is programmed to look toward the person when he or she is nearby. In this instance, the robot made a spontaneous glance toward the canister immediately after the person asked "Can I pick up the canister?" and just before the robot responded "Yes." This added a nice bit of in-the-moment coordination in the communication between them.

Figure 13. Example of a positive interaction with the robot around retrieving the canister.

6.2 How Did People React Emotionally to the Robot?

Finally, we noticed that people varied strongly in how they reacted emotionally to the robots behaviors. While some interactions were characterized by humor and high positive energy, others seemed more neutral, or even negative, with participants displaying signs of frustration. Seeing these emotional differences was particularly interesting for us since emotional dynamics were not considered in the automated development of the model.

For example, Figure 14 depicts a transcript of a 7-s interaction where mostly nonverbal cues

are coordinated around retrieving the alien from the elevator. The overall tone of the interaction is positive, with the robot responding affirmatively both verbally and nonverbally to the mans request to get the alien. The robot glances at the man as he steps aside for it to approach the elevator. He asks the robot to get the alien, and as he starts to walk away, the robot responds affirmatively to the humans request. The man then decides to rejoin the robot to finish the task as positive uptake on the robots response

Figure 14. Example of the human and robot positively responding each other.

The robots behavior certainly had hiccups during these exchanges. Sometimes within an interaction, the robot would make erroneous utterances, which added tension to the interaction even if the net result was positive. An example of this is shown in the transcript in Table 2. Although the erroneous utterance "I dont care" added a bit of tension to the interaction, the polite "Thank you" ended the overall interaction on a positive note.

Early in our investigation, one particular kind of interaction that seemed to be a source of frustration caught our attention. We found that sometimes a participant addressed the robot with a statement such as, "Would you like me to put this item into the bucket?" while looking at the robot and waiting for confirmation. The robot then sometimes just turned around without any sign of acknowledgment, seemingly ignoring the request. These interactions often created somewhat emotionally vulnerable moments, leaving the person standing with an object in hand waiting for a response from the robot, and then often showing signs of giving up with a frustrated shrug. In other words, the robots turning

Table 2: A bidding interaction with mixed affect, ending on a positive note.

Time [s]	Actor	Activity
0		Participant is standing next to the scale watching as robot approaches
1	R	*Please put the object in the bucket!*
7	R	*What next?*
8	P	*(Participant lifts up the item from the container, shows it to the robot) Shall I put this in the bucket?*
11	R	*What next?*
14	R	*Please put the object in the bucket!*
16	P	Participant starts walking towards bucket) *Ok*
17	R	(Looks after participant) *I don't care*
19	P	(Participants expresses tense laughter and drops the item into the bucket)
20	R	*Thank you.*
21	P	(Looks at robot and smiles) *You're welcome!*

Video: CBP 15
(R) = Robot
(P) = Participant

away in the middle of being asked a question was perceived as being rude. Researchers of marital interactions have found that interactions such as these can be referred to as bid-response interactions and can be distinguished by their emotional quality (Driver & Gottman, 2004). In the words of Driver and Gottman (2004), a bid "refers to the idea that one partner is attempting to interact with the other. They are bidding for attention or an emotional connection. These bids cover a wide range of interactions from a simple look to a funny story." Once a bid is made, the following response or non-response becomes emotionally relevant. Positive responses to bids can range from a simple glance, signaling "I noticed you," to an engaged reply or action. Negative responses to bids can range from passive non-responsiveness, to an active ignoring of the bid by turning away, to a highly negative hostile response such as a contemptuous remark. Table 4 is a simplified version of the original bids classification system, with examples for each bid or response type.

The concept of bids is relevant because it reveals that emotional expressivity does not rely on the robots capabilities of making specific facial expressions, body movements, or voice pitch modulations. Rather, the meaning of behavior is constructed in the moment of an interaction. For example, a robots behavior of turning away does not have any emotional meaning in itself, but if the behavior happens while someone is talking to the robot, that same movement suddenly expresses negative affect and ignoring of the person. Bid and response therefore form what Schegloff (2007) called an adjacency pair: The interpretation of the response is contingent upon the bid. Looking at the videos from this perspective, we found a multitude of instances that could be characterized as bids and responses, and they often gave us insight into what triggered participants expressions of frustration or delight.

Many of the interactions can also be viewed from the perspective of building common ground (Clark & Brennan, 1991), in which the bid could be seen as the *presentation phase* of a contribution to a conversation, and the response could be seen as the *acceptance phase* of a contribution to a

Table 3: A bidding interaction that begins positively but ends on a negative note when the robot ignores the person.

Time [s]	Actor	Activity
0		Participant places the last item into the bucket and walks towards the robot
3	R	*Thank you!*
5	P	(looks at robot) *Thank you!* (smiles)
7	R	*Great, we're done. Let's get out of here.*
11	P	(leaning towards robot, see figure below) *Wanna grab a beer?*
15		Participant walks away after receiving no reaction from robot.

Video: CBP 6
(R) = Robot
(P) = Participant

conversation. Viewing these interactions from a grounding perspective, however, does not allow us to see differences in their emotional meaning. Clark and Brennan do not give any insight into when and why the presentation or acceptance phases of a contribution carry additional emotional meaning. However, affective tone was certainly present in a number of our observed, in-the-moment human-robot exchanges.

The way that people issue and respond to bids is important in human relationships. For instance, it has been shown to be a key indicator of the health of a relationship (Driver & Gottman, 2004). Jung, Chong, and Leifer (2012) have applied a similar coding scheme to identify behavior patterns that are predictive of performance in programming teams. In recent work, such coding was even successfully applied to assess the quality of human-robot teamwork where a human collaborates with two autonomous humanoid robots (Jung, Lee, et al., 2012). Understanding what it means for a robot to recognize and respond to bids in a way that is perceived as positive could have important implications for the design of robots that engage with people for longer amounts of time and must establish and maintain positive rapport.

7. Implications and Future Research

The ability for a robot to engage in interactive parallel and sequential teamwork, as well as to perceive human coordination cues while also engaging in cuing behavior, are important skills to support high-quality human-robot teamwork. For highly trained tasks that follow rigid protocols, the ability to flexibly coordinate behavior in the moment and respond to bids is probably less critical. In this work, however, we intentionally explored the opposite case, where rich and diverse teaming behavior and interaction styles could arise from a loosely structured task. Whereas more traditional machine learning methods would serve to abstract human behavior into a generalized policy, we wanted to preserve this richness with the ultimate goal of respecting different peoples styles and approaches to human-robot teamwork. The large number of possible states for such domains presents a challenge to machine learning methods that do not scale well. We need to develop new methods that can succeed in supporting rich and flexible human-robot teaming over a diverse set of users for loosely structured activities. As robots begin to work alongside people as a part of daily life, loosely structured tasks and in-the-moment coordination will be the norm.

Table 4: Bids coding scheme.

	Code	Valence	Example Human	Example Robot
Bid	Comment Bid	Positive	"The item is over there!"	"The item is in box 56"
	Question Bid	Positive	"Can you tell me where box 56 is?"	"Can you help me?"
	Negative Bid	Negative	No Occurrence	No Occurrence
	Re-Bid	Depends on previous bid	Repeated bid: e.g. again: "The bucket is over there!"	Repeated bid: e.g. again: "The item is in box 56"
Response	Active Responding	Positive	"Yes, of course I can get the item"	"Yes!"
	Passive Responding	Positive	looks at speaker and nods silently	Seemingly turns to speaker
	No Responding	Negative	No reaction to bid, neither verbal nor non-verbal	No reaction to bid, neither verbal nor non-verbal
	Aversive Responding	NEgative	Active rejection, e.g. turning away while robot is speaking	Active rejection, e.g. turning away while person is speaking

Our novel approach to crowdsourcing human interaction data is an early attempt in exactly this direction. Although this attempt was promising, future methods clearly require better models and new approaches to facilitate the collection of better quality interaction data. Future work can take inspiration from interactive, crowdsourced agent literature as well as from the human-robot-agent teaming literature to improve on these task and dialogue models. We have a number of recommendations to make in this regard.

One of the challenges for purely data-driven approaches is dealing with data sparsity in demonstrations for complex tasks. Rarely observed examples of behavior and language may slip through the cracks of statistical learning algorithms; thus, future work must find ways to compensate for data sparsity. One possibility is to query knowledge from external sources when faced with unfamiliar words or action sequences; for example, by making use of common-sense databases and semantic networks like OpenMind and ConceptNet (P. Singh et al., 2002) or crowdsourced models of everyday activities (Li, Appling, Lee-Urban, & Riedl, 2011). Another option is to employ the crowd in the data interpretation process, leveraging humans to explain and preserve sparse but useful examples, an approach that is yielding promising results in video game artificial intelligence (Orkin & Roy, 2012). Other work in this direction includes managing interaction sparsity concerns at the cost of task modeling (Broz, Nourbakhsh, & Simmons, 2011; Clair & Matarić, 2011). Ideally, robots will one day be able to engage in online learning from experience to augment the corpus when previously unseen situations arise.

Moreover, the robustness analysis in Section 5.2 raises the issue of how to incentivize correct behavior from large crowds. In order to provide useful data for our approach, the behaviors that are

recorded in the online game must be appropriately directed. Incentives are now receiving focused attention in the crowdsourcing literature (Shaw, Horton, & Chen, 2011). The number of cases where people correctly completed the game were so few that frequency becomes a naive heuristic in building models from noisy data. The absence of incentives toward finishing the game had an obvious impact on our dataset. Our previous work (Chernova, DePalma, & Breazeal, 2011) also emphasized how closely our resulting behavior system matches the general populations data in terms of participant action traces through the task. By focusing on obtaining more coherent, cleaner data, our results will inevitably get better. The nature of the spoken dialogue between people and our robots as been the topic of prior publications (Chernova, DePalma, & Breazeal, 2011) and certainly has ample room for improvement. One issue we had to grapple with is that the online-sourced game dialogue was rich in distinctively Internet language that included acronyms like "LOL" and personal comments like "Hi Dad!". These types of utterances rarely show up in real-world conversation, and the speech synthesizer did not handle them gracefully. There is certainly an opportunity to look into the NLP literature to filter or translate such phrases into utterances that are useful in direct human-robot interaction. Nonetheless, the museum environment was so noisy that speech recognition was useless. As a result, most of the robot's context was derived from task and location information rather than on what people said. Given this, the system often did surprisingly well, although its utterances were often a bit quirky. We will save a detailed language analysis for later work where the acoustic environment is more amenable to speech input.

We focus our attention, instead, on a broader set of communication and coordination behaviors, including bids. Despite not having captured nonverbal communication cues in the online game, we found that under certain conditions random nonverbal behavior, such as a simple turning around, could gain social meaning based on how it was placed in the context of an interaction. Many of these nonverbal behaviors do not have any social or emotional meaning in themselves, but when a robot turns away while being addressed by a person, it breaks social politeness norms and appears rude or ignorant. Similarly, a gaze at the right moment and in the right context can signal acknowledgment of or even caring for a person. Participants naturally perceive these cues whether or not the intent to back-channel and communicate through nonverbal action is there. These unintentional acts often elicit emotive responses; human participants dont just perceive and interpret these acts, but expect them and react to them as if they matter. We argue that robot-teaming models require social awareness to sufficiently react to the humans back-channeling and bid initiation and response, and to support this type of communication and coordination.

When capturing datasets and building models from the Internet, a popular temptation is to just capture the task-related data and build human-robot teaming policies only around the task data. However, our video transcripts illustrate that the coordination demands of the task itself (e.g., those tasks designed to require synchronous action) did not determine when coordination attempts were initiated. Rather, people attempted to coordinate with the robot throughout the entire Mars Escape scenario, at different times, in different ways, and often serendipitously, rather than at specific times that might be captured by a trained policy. The notion that many collaborative interactions in human-robot teamwork are fluid and constructed in the moment brings up specific questions about how these interactions are initiated, maintained, and ended. How can robots more reliably notice bids and respond to them appropriately? How can robots select appropriate bids when a need for collaboration arises? What behavioral patterns end interactions in ways that build good rapport and trust? Future research and development efforts should address such questions in building high-performance human-robot teams that can simultaneously address both the task and the interpersonal demands of working alongside people.

Acknowledgements

This work was supported by Microsoft Research and the Office of Naval Research Award Numbers N000140910112 and N000140710749.

References

Abbeel, P., Coates, A., Quigley, M., & Ng, A. (2007). An application of reinforcement learning to aerobatic helicopter flight. *Advances in Neural Information Processing Systems*, *19*, 1.

Barsade, S. G., & Gibson, D. E. (2007). Why does affect matter in organizations? *The Academy of Management Perspectives*, *21*(1), 36–59, http://dx.doi.org/10.5465/AMP.2007.24286163.

Breazeal, C. (1998). A motivational system for regulating human-robot interaction. In *Aaai/iaai* (pp. 54–62). Menlo Park, CA, USA: American Association for Artificial Intelligence.

Breazeal, C., Kidd, C., Thomaz, A., Hoffman, G., & Berlin, M. (2005). Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In *Proceedings of international conference on intelligent robots and systems* (pp. 708–713). http://dx.doi.org/10.1109/IROS.2005.1545011: IEEE.

Broz, F., Nourbakhsh, I., & Simmons, R. (2011). Designing pomdp models of socially situated tasks. In *Proceedings of the 20th symposium on robot and human interactive communication* (pp. 39–46). IEEE. http://dx.doi.org/10.1109/ROMAN.2011.6005264.

Burgard, W., Cremers, A. B., Fox, D., Hähnel, D., Lakemeyer, G., Schulz, D., et al. (1998). The interactive museum tour-guide robot. In *Aaai/iaai* (pp. 11–18).

Calinon, S., & Billard, A. (2008, May). A framework integrating statistical and social cues to teach a humanoid robot new skills. In *Proceedings of international conference on robotics and automation, workshop on social interaction with intelligent indoor robots.* IEEE.

Chao, C., Cakmak, M., & Thomaz, A. (2010). Transparent active learning for robots. In *Proceedings of 5th international conference on human-robot interaction* (pp. 317–324). IEEE. http://dx.doi.org/10.1109/HRI.2010.5453178.

Chernova, S., DePalma, N., & Breazeal, C. (2011). Crowdsourcing real world human-robot dialog and teamwork through online multiplayer games. *AI Magazine*, *32*(4), 100–111.

Chernova, S., DePalma, N., Morant, E., & Breazeal, C. (2011). Crowdsourcing human-robot interaction : Application from virtual to physical worlds. In *Proceedings of the 20th symposium on robot and human interactive communication* (pp. 21–26). IEEE. http://dx.doi.org/10.1109/ROMAN.2011.6005284.

Clair, A., & Matarić, M. (2011). Modeling action and intention for the production of coordinating communication in human-robot task collaborations. In *In proceedings of the 21st international symposium on robot and human interactive communication: Workshop on robot feedback in HRI* (pp. 39–46). IEEE.

Clark, H., & Brennan, S. (1991). Grounding in communication. *Perspectives on socially shared cognition*, *13*(1991), 127–149, http://dx.doi.org/10.1037/10096-006.

Crick, C., Jay, G., Osentoski, S., & Jenkins, O. (2012). ROS and Rosbridge: Roboticists out of the loop. In *Proceedings of the seventh annual international conference on human-robot interaction* (pp. 493–494). ACM.

Crick, C., Osentoski, S., Jay, G., & Jenkins, O. C. (2011). Human and robot perception in large-scale learning from demonstration. In *Proceedings of the 6th international conference on human-robot interaction* (pp. 339–346). ACM. http://dx.doi.org/10.1145/1957656.1957788.

Driver, J. L., & Gottman, J. M. (2004). Daily marital interactions and positive affect during marital conflict among newlywed couples. *Family Process*, *43*(3), 301–314, http://dx.doi.org/10.1111/j.1545-5300.2004.00024.x.

Gorin, A. L., Riccardi, G., & Wright, J. H. (1997). How may I help you? *Speech Communication*, *23*(1-2), 113–127, http://dx.doi.org/10.1016/S0167-6393(97)00040-X.

Graf, B., Hans, M., & Schraft, R. D. (2004). Care-o-bot II—development of a next generation robotic home assistant. *Autonomous Robots*, *16*(2), 193–205, http://dx.doi.org/10.1023/B:AURO.0000016865.35796.e9.

Gray, J., Breazeal, C., Berlin, M., Brooks, A., & Lieberman, J. (2005). Action parsing and goal inference using self as simulator. In *Proceedings of the 14th symposium on robot and human interac-*

tive communication, workshop on robot and human interactive communication (pp. 202–209). IEEE. http://dx.doi.org/10.1109/ROMAN.2005.1513780.

Jung, M., Chong, J., & Leifer, L. J. (2012). Group hedonic balance and pair programming performance: Affective interaction dynamics as indicators of performance. In *Proceedings of the conference on human factors in computing systems.* ACM.

Jung, M., Lee, J., DePalma, N., Adalgeirsson, S., Hinds, P., & Breazeal, C. (2012). Engaging robots: Easing complex human-robot teamwork using backchanneling. In *Proceedings of the conference on computer supported cooperative work.* ACM.

Kacmarcik, G. (2005). Question-answering in role-playing games. In *Workshop on question answering in restricted domains.* American Association for Artificial Intelligence.

Kelley, R., Tavakkoli, A., King, C., Nicolescu, M., Nicolescu, M., & Bebis, G. (2008). Understanding human intentions via hidden markov models in autonomous mobile robots. In *Proceedings of the international conference on human-robot interaction* (pp. 367–374). ACM. http://dx.doi.org/10.1142/S0219843608001418.

Knox, W., & Stone, P. (2009). Interactively shaping agents via human reinforcement: The tamer framework. In *Proceedings of the fifth international conference on knowledge capture* (pp. 9–16). ACM.

Kolodner, J. (1993). *Case-based reasoning*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. http://dx.doi.org/10.1017/CBO9780511816833.015.

Kulyukin, V. (2004). Human-robot interaction through gesture-free spoken dialogue. *Autonomous Robots*, *16*(3), 239–257, http://dx.doi.org/10.1023/B:AURO.0000025789.33843.6d.

Lee, M. K., & Makatchev, M. (2009). How do people talk with a robot?: an analysis of human-robot dialogues in the real world. In *Proceedings of the 27th international conference extended abstracts on human factors in computing systems* (pp. 3769–3774). New York, NY, USA: ACM. http://dx.doi.org/10.1145/1520340.1520569.

Li, B., Appling, D., Lee-Urban, S., & Riedl, M. (2011). Learning sociocultural knowledge via crowdsourced examples. In *In proceedings of the 4th aaai workshop on human computation.* American Association for Artificial Intelligence.

McNaughton, M., Schaeffer, J., Szafron, D., Parker, D., & Redford, J. (2004). Code generation for ai scripting in computer role-playing games. In *Challenges in game ai workshop at aaai.* American Association for Artificial Intelligence. http://dx.doi.org/10.1109/ASE.2004.1342770.

Mutlu, B., Shiwa, T., Kanda, T., Ishiguro, H., & Hagita, N. (2009). Footing in human-robot conversations: how robots might shape participant roles using gaze cues. In *Proceedings of the international conference on human-robot interaction* (pp. 61–68). ACM.

Orkin, J., & Roy, D. (2007). The restaurant game: Learning social behavior and language from thousands of players online. *Journal of Game Development*.

Orkin, J., & Roy, D. (2009). Automatic learning and generation of social behavior from collective human gameplay. In *International conference on autonomous agents and multi-agent systems* (pp. 385–392).

Orkin, J., & Roy, D. (2012). Understanding speech in interactive narratives with crowdsourced data. In *In proceedings of the 8th conference on artificial intelligence and interactive digital entertainment.* American Association for Artificial Intelligence.

Rieser, V., & Lemon, O. (2010). Learning human multimodal dialogue strategies. *Natural Language Engineering*, *16*, 3–23, http://dx.doi.org/10.1017/S1351324909005099.

Schegloff, E. (2007). *Sequence organization in interaction: Volume 1: A primer in conversation analysis* (Vol. 1). Cambridge, UK: Cambridge University Press.

Shaw, A., Horton, J., & Chen, D. (2011). Designing incentives for inexpert human raters. In *Proceedings of the conference on computer supported cooperative work* (pp. 275–284). ACM.

Sidner, C., & Lee, C. (2007). Attentional gestures in dialogues between people and robots. In T. Nishida (Ed.), *Engineering approaches to conversational informatics.* Wiley and Sons. http://dx.doi.org/10.1002/9780470512470.ch6.

Singh, P., Lin, T., Mueller, E., Lim, G., Perkins, T., & Li Zhu, W. (2002). Open mind common sense: Knowledge acquisition from the general public. *On the Move to Meaningful Internet Systems 2002: CoopIS, DOA, and ODBASE*, 1223–1237, http://dx.doi.org/10.1007/3-540-36124-3_77.

Singh, S., Litman, D., Kearns, M., & Walker, M. (2002). Optimizing dialogue management with reinforcement learning: Experiments with the NJFun system. *Journal of Artificial Intelligence Research*, *16*(1), 105–133.

Sorokin, A., Berenson, D., Srinivasa, S., & Hebert, M. (2010). People helping robots helping people: Crowdsourcing for grasping novel objects. In *Proceedings of the international conference on intelligent robots and systems* (pp. 2117–2122). IEEE. http://dx.doi.org/10.1109/IROS.2010.5650464.

Spalzzi, L. (2001). A survey on case-based planning. *Artificial Intelligence*, *16*(1), 3–36, http://dx.doi.org/10.1023/A:1011081305027.

Stork, D. G. (1999). The open mind initiative. *Intelligent Systems and Their Applications*, *14*(3), 19-20, http://dx.doi.org/10.1109/ICDAR.1999.791712.

Thomaz, A., & Breazeal, C. (2006). Transparency and socially guided machine learning. In *Proceedings of 5th international conference on development and learning.* IEEE. http://dx.doi.org/10.1109/ROMAN.2006.314459.

vonAhn, L., & Dabbish, L. (2008). Designing games with a purpose. *Communications of the ACM*, *51*(8), 58–67.

Wilcox, R., Nikolaidis, S., & Shah, J. (2012). Optimization of temporal dynamics for adaptive human-robot interaction in assembly manufacturing. In *Proceedings of robotics science and systems.*

Authors' names and contact information: Cynthia Breazeal, Nick DePalma, Jeff Orkin, MIT Media Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA. Email: cynthiab@media.mit.edu, ndepalma@media.mit.edu, jorkin@media.mit.edu. Sonia Chernova, Department of Computer Science, Worcester Polytechnic Institute, Worcester, MA, USA. Email: soniac@cs.wpi.edu. Malte Jung, Department of Management Science and Engineering, Stanford University, Palo Alto, CA, USA. Email: mjung@stanford.edu.