Are You Messing with Me? Querying about the Sincerity of Interactions in the Open World

Sean Andrist[1,2,+], Dan Bohus[2], Zhou Yu[2,3,+], Eric Horvitz[2]
(1) University of Wisconsin–Madison, Madison, WI, USA
(2) Microsoft Research, Redmond, WA, USA
(3) Carnegie Mellon University, Pittsburgh, PA, USA
(+) work was conducted during internships at Microsoft Research
sandrist@cs.wisc.edu; dbohus@microsoft.com; zhouyu@cs.cmu.edu; horvitz@microsoft.com

ABSTRACT
When interacting with robots deployed in the open world, people may often attempt to engage with them in a playful manner or test their competencies. Such engagements are often associated with language and behaviors that fall outside of designed task capabilities and can lead to interaction failures. Detecting when users are driven by play and curiosity can help a robot to understand why some interactions are breaking down, respond more appropriately by conveying its capabilities to its users, and enhance perceptions of its situational awareness and social intelligence. We have been studying the intentions of everyday users in their engagement with a long-lived robot system that provides directions within an office building. We report on a pilot field-study exploring the use of direct queries to elicit the sincerity of user requests, in terms of their actual need for directions. We discuss early results from this initial study and frame research directions and design implications for robots deployed in the wild.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*evaluation/methodology, user-centered design*

Keywords
User intentions, off-task interaction, elicitation actions

1. INTRODUCTION
Interactive robots deployed in open, public spaces can encounter a variety of unforeseen situations that are not easily reproducible in the closed-world environment of controlled lab studies. An important class of such interactions stem from peoples' desires to explore and test boundaries when encountering novel technology, or to demonstrate a robot's capabilities to others. Such behaviors often lead to requests outside of the task domain, including utterances generated during playful natural discourse and chit-chat, further exacerbating what is typically an already challenging interaction task. In these instances, the robot is no longer being engaged on its primary competency, but instead as a novel appliance to be tested and experimented with.

In previous work, Makatchev and Simmons [4] found that 16.7% of users would reply to a robot receptionist's questions with a "display question" meant to test the robot's knowledge, revealing their playful intentions. Furthermore, only 41.5% of users were found to initiate interactions with the robot in order to seek information, with the rest engaging in social chatting, brief greetings while passing by, or "nonsense" [3]. Fischer [2] has explored strategies for using a robot's speech to guide users into behaviors that are more easily interpretable by the robot.

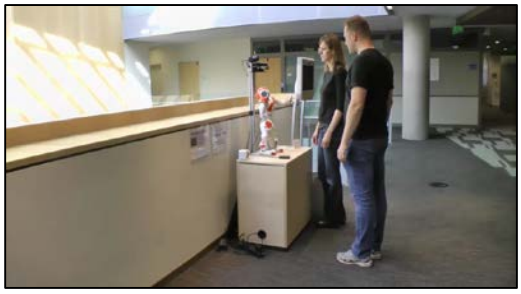

Figure 1. *Directions Robot* deployment setup.

Successful deployments of task-oriented interactive robots in the open world hinge in part on endowing them with the ability to better recognize intentions and detect off-task, playful interactions. This type of awareness would allow a robot to (1) better understand *why* it is misrecognizing or performing poorly with language, turns, and gestures when the interaction does not unfold smoothly, (2) convey the boundaries of its capabilities to the user, and (3) enhance its expression of social intelligence by recognizing and engaging appropriately per the assessed situation. In this paper, we report on our first phase of research in this area.

2. ROBOT SYSTEM AND SETTING
The experimental testbed for this work is *Directions Robot*, a Nao-based system that can engage in natural conversation (see http://sdrv.ms/15Yay8V) with one or multiple participants and give directions to locations inside a building [1]. Currently, three *Directions Robot* systems are deployed in our building, one on each floor. The robots are configured in a stationary, standing position on top of a desk in front of the bank of elevators (Figure 1). Interactions in this space involve people based within the building and visitors. Signs are mounted inside the elevators, as well as on the wall on approach to the robot's space. The signs advertise the robot's presence, provide an overview of the directions service, and disclose the privacy policy with regard to the video capture of interactions with the system. The deployment and collection of data for research purposes was reviewed by legal and ethics experts at our institution.

Some of the interactions in these spaces are driven by real needs for directions within the building, *e.g.*, people as individuals or in groups arriving for a meeting and looking for an office, restrooms, or conference room on the floor. However, we have observed that many interactions arise from curiosity about the robot. People that engage with the robot have a mix of goals beyond a true need for directions. They may simply be curious about how the robot works, try to test its limits, or try to engage in social chit-chat.

Intention	Action assessment	User response
Playful	Appropriate, [participants] have made several [directions] requests including elevator, timing good, comes after [participants] discuss what room to ask for.	Two [participants] laugh, another shows surprise then smiles. Interactor answers "yes" after a short pause, slightly guilty expression on face.
Sincere	[Participant] hasn't given any indication of playfulness. Dialog action disrupts interaction and leads to a misunderstood response - further disruption. [Participant] is visibly/audibly irritated.	Rolls eyes, sighs, says "No but you're wasting my time". [Participant] remains annoyed for rest of [interaction].

Table 1: Examples from annotator's assessment of elicitation actions based on ground truth labeling of users' intentions.

3. QUERYING INTENTIONS

In an initial study, an expert annotator with formal linguistics training reviewed videos of 71 interactions collected with two of the robots over a period of 5 days, and annotated users' intentions. The annotator assessed whether or not each engaged participant (as detected by the robot) had a real need for directions by considering the following question: "Is this an interaction that the person might have *sincerely* had with a *human* directions-giving receptionist, or is this a *playful* interaction meant to test or play with the *robot* system?" Participants had playful intentions 81% of the time (assessed at the robot's turns) in this dataset.

Our initial explorations using machine learning in conjunction with multimodal features indicate that automatic detection of these playful versus sincere intentions is a challenging problem. As such, and given the high amount of playfulness observed, we decided to explore the use of direct *elicitation actions* aimed at revealing this hidden state. The action was rendered as follows: *"Sometimes people like to play with me and don't really need directions. I'm curious, are you just testing my abilities right now?"* The action was designed to directly query the user's need for directions, while maintaining a curious and non-accusatory tone. The robot then waits for a "yes" or "no" response from the user, and then, after providing an acknowledgment of the response, proceeds to give the directions the user had requested.

We collected data again with these new elicitation actions in place. The actions were triggered in an exploratory fashion. Over a period of 10 days, the two robots engaged in 93 interactions and triggered 28 elicitation actions (at most once per interaction). The annotator assessed user intentions again, and this time also characterized the action and the user responses in context. Two examples of her observations are displayed in Table 1. Participants had playful intentions 65% of the time in this dataset.

In 15 out of 28 elicitation action instances, the participants were labeled as not having a real need for directions. For only four of the 15, participants admitted this verbally in response to the elicitation action, and their responses were accompanied by smiling or laughter. Interestingly, six playful participants denied being playful in response to the robot's elicitation action, insisting that they really needed directions. These responses were also sometimes accompanied with smiling and laughter, or with participants continuing to video the interaction with their phone at the time. In the five remaining instances assessed as playful, participants left after the robot asked the question.

For the 13 action instances where participants were labeled as having a sincere need for directions, seven gave a negative verbal response to the robot's query. In five cases, participants became visibly annoyed or impatient at the robot's question, or gave up on the interaction and left without responding. In nine cases, participants indicated amusement by smiling or laughing at the question. The annotator generally remarked that when participants had a real need, the action was disruptive to the interaction. We note that, in some instances, a system problem with one of the robots led to unnaturally long pauses before taking a turn. This may have influenced some of the interactions.

4. DISCUSSION

Robots can use elicitation actions to learn more about users' intentions in a targeted way. However, our analysis indicated that these actions are not as cost-free as we had originally assumed they would be and thus must be carefully designed and tested. The annotator (taking the mental stance of a "fly-on-the-wall" human observer) noted that the action often sounded suspicious in context, and that it was often disruptive when people had a real need for directions. Thus, it will be important to further refine the surface realization and rendering of this action. Beyond design, next steps include refining models to infer the likelihood that users are being playful or testing the system's capabilities, and employing a decision-theoretic policy that balances the expected benefits with the cost of such elicitations or other actions.

Given the observation that users do not always admit to being playful when directly asked, we cannot rely solely on verbal responses to these elicitations about sincerity. Future work should investigate the use of an extended memory of participants and interactions, as well as non-verbal features to help systems disambiguate truthful responses from denials. Promising features include detecting emotional responses, such as smiling or laughter, conflicting responses from multiple participants, and off-task behaviors, such as filming the robot.

Much of the playfulness we observed is likely due to the novelty of seeing an engaging humanoid robot in a public space. As robots become more pervasive, such novelty may diminish. However, robots in the open world will no doubt continue to encounter off-task interactions from users who seek to be playfully social or who are simply not aware of robots' limitations and domains. In these cases, it would be valuable for robots to identify and/or express awareness of off-task behavior in a socially competent way and to attempt to steer the interaction back to their domain-specific competencies. Endowing systems with the ability to make such inferences and to guide people to engage on designed competencies will raise the likelihood of successful experiences.

Acknowledgement. We thank Rebecca Hanson for her help with data annotation.

5. REFERENCES

[1] Bohus, D., Saw, C. W., & Horvitz, E. (2014). Directions robot: In-the-wild experiences and lessons learned. In *Proc. AAMAS '14.* (pp. 637-644).

[2] Fischer, K. (2011). How people talk with robots: Designing dialog to reduce user uncertainty. *AI Magazine*, 32(4), 31-38.

[3] Lee, M. K., & Makatchev, M. (2009). How do people talk with a robot?: An analysis of human-robot dialogues in the real world. In *CHI'09 Extended Abstracts*, ACM.

[4] Makatchev, M., & Simmons, R. G. (2010). Do you really want to know? Display questions in human-robot dialogues. In *AAAI Fall Symposium: Dialog with Robots.*