Towards Integrating Dialog, Planning, and Execution for Service Robots

Frank Broz, Alessandro Di Nuovo, Tony Belpaeme, and Angelo Cangelosi
Plymouth University, UK
{frank.broz, alessandro.dinuovo, tony.belpaeme, angelo.cangelosi}@plymouth.ac.uk

Abstract

This paper presents an experiment investigating what type of progress feedback users prefer in verbal updates by a robot about remotely performed tasks. Of primary concern is that users find the information presented useful. But as users in their home may be engaged in other activities while they wait for a service, it is also important that information is presented in a way and at a frequency that they do not find distracting or disruptive. We explore these issues through a human-robot interaction experiment involving a simulated food delivery service. We also discuss future research directions that involve giving naive users more input into the planning process.

Motivation

In order for robotic services to readily adopted, they must be usable by non-experts with varying levels of experience with technology. This raises new questions about how to inform users about the functionality of these complex systems. The ROBOT-ERA project seeks to provide everyday services for the elderly using a group of collaborating robots each designed for a particular environment (in the home, in shared indoor spaces, and outdoors) and coordinated by a central planner (Rocco et al. 2014). During the performance of these services, the robot that is with the user in the home is often not the robot that is actively carrying out the current stage of a task. However, the planner gives this robot access to the activities of the other robots as well as a potentially accurate estimate of when the service will be completed. In these situations, what information should a domestic robot provide to the user about task progress?

Related Work

Most research on speech-based interaction with service robots involves giving feedback about tasks that the robot is performing with the user (Marge et al. 2009; Rosenthal and Veloso 2010). To assist users that are non-experts and may have physical or cognitive impairments, a system should be capable of delivering services autonomously without making them a critical part of task performance. The user should be informed about a service's progress without needing to monitor the robot's actions in order to allow them to focus their attention on other activities while they wait. Therefore, the user needs information that provides important context about task progress with minimal distraction.

In this sense, the updates given by the robot are related to the kinds given by reminding agents, and the literature on task interruption is relevant to designing them. There is research on task interruption that focusses specifically on elderly users, but this work involves interactions with ambient intelligence or virtual agents rather than physical robots (McGee-Lennon, Wolters, and Brewster 2011). In interactions with a virtual agent, users preferred interruptions that contained social and empathetic content (Bickmore et al. 2008). A study on multimodal interruptions found that elderly and non-elderly users reacted to the interruption modalities similarly and that the modalities evaluated all had similar effects in terms of task interruption (Warnock, McGee-Lennon, and Brewster 2013). This suggests that the best modality to use may depend on the situation.

These interactions are also related to research in human-in-the-loop planning. However, most work on human-in-the-loop planning for robots is focussed on supporting experienced, expert users. The planning tasks often require the user to provide low-level control to the robot and/or assign tasks to robots themselves (Leeper et al. 2012; Miller and Parasuraman 2007). Research that involves humans interacting with autonomously generated plans highlights the difficulty of interpreting these plans for even experienced technical users (Roth et al. 2004). Therefore, how to involve non-expert users in re-planning in cases where plans fail is a difficult open issue.

Experiment

Design and Procedure

The experiment was conducted on the campus of Plymouth University in a classroom set up to simulate a home environment. Coro, one of the indoor mobile robots designed for the ROBOT-ERA project was used for the experiment. Users interact with the robot using a multimodal speech and tablet-based interface (Broz et al. 2012). Updates given via tablet may be missed by the user if they are engaged in another activity and not holding it. Therefore, we focus on verbal updates.

The experiment has a within-subjects design with 3 conditions: no feedback, event-based feedback, and time-based feedback. In all conditions, the user orders a meal using the tablet interface. Five minutes later, the robot goes to the door of the room as if to collect a meal delivery. While they waited, participants were able to engage in a pleasurable leisure activity (reading magazines provided for them). After experiencing each condition, the participant filled out a short questionnaire consisting of Likert scale responses to four statements. The robot's speech acts varied by condition:

- **No feedback**: the robot informs the user that the order has been placed and when their order has arrived.
- **Event feedback**: In addition to the statements in the no feedback condition, the robot informs the user after 3 minutes that their meal is being dispatched for delivery.
- **Time feedback**: In addition to the statements in the no feedback condition, the robot initially informs the user that their order will arrive in 7 minutes. 3 minutes later, the robot announces that the delivery time has changed and will be in 2 minutes.

Participants were non-elderly adults recruited from the Plymouth University campus. Twenty-one people participated in the study (M = 10, F=11). One user's data was excluded because they failed to complete the questionnaires.

Results

Table 1: Median Likert-scale responses for each condition with 95% confidence intervals on the median of the difference for the planned comparisons. Statistical significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

	No feedback	Event feedback	Time feedback
S1	2 [-4 -2] ***	2.5 [-3.5 -1.5] ***	5
S2	1	1	1
S3	2.5 [0.5 2.5] *	3 [0.5 2.5] **	1
S4	3 [-3 -1] ***	4 [-2 -1]***	4.5

We hypothesised that users would prefer the time-based condition, finding it the most informative and useful. Therefore, we planned comparisons between the time condition and the other two conditions. The measure used was the median response for each statement. A Wilcoxon signed rank test (exact) was used to test for statistical significance in the cases where the medians differ (see Table 1).

- **S1 (Informativeness)**: "The robot gave me enough information about when the delivery would arrive."
  Users strongly agreed that the time condition gave them adequate information about when the service would complete. They were in slight disagreement in the other conditions.
- **S2 (Annoyance)**: "The robot spoke too often."
  This statement investigates whether increasing the number of speech acts would negatively effect users' opinion of the interaction. In no condition was the amount of speech judged to be excessive.
- **S3 (Time Perception)**: "The delivery took longer than I thought it would."
  This statement measured the effect of the robot's speech on the users' subjective impression of the passage of time. Users strongly disagreed that the delivery took longer than expected in the time condition and were neutral in the other conditions.
- **S4 (Usefulness)**: "I found the robot's statements about the delivery useful."
  Users were neutral about the usefulness of the robot's speech in the no feedback condition. They somewhat agreed that event-based feedback and the strongly agreed that time-based feedback were useful.

Future Work

The experiment presented concerns a "best case" example of a robotic service in which no failures occurred and delays were minimal. But delays and failures are common occurrences in autonomous systems, and how to communicate with users about them is an important part of creating acceptable and interpretable robot behaviour. The constraint-based configuration planner used for ROBOT-ERA can deal with task deadlines, limited resources, and concurrent goals and can support giving accurate estimates of task completion times (Di Rocco, Pecora, and Saffiotti 2013). It also performs execution monitoring, allowing the system to detect if and when the current plan becomes unachievable.

Because the services provided effect the users directly, it is reasonable to expect that they would want to have feedback about why plans have failed and input on what constraints should be relaxed to find an achievable plan. But finding an explanation for why a plan has become unachievable is not straightforward. Coming up with a concise, reasonable, and interpretable set of possible relaxations to present to the user as options for re-planning is an even more complex problem.

The second level of complexity arises from the medium of communication. Users interact with the planner through a domestic robot in their home. In addition to what information should be provided, care should be taken in determining when and how often the user should be notified. Determining how long of a delay requires notification is likely to depend on a variety of factors including: the service underway, the user's current activity, and potential impact on the user's future plans.

Conclusions

This was an initial experiment on how to provide feedback about remote robotic services. Users were found to prefer time-based statements about task progress, though the results suggest that they also found event-based feedback to be somewhat useful. This study only investigated the case where the service was executed more or less "as planned". Users may want more or different types of information in cases where task execution fails or a service becomes severely delayed. How to involve the user in re-planning decisions when problems arise and when and in which cases users should be notified about changes to a planned service are also issues for further investigation.

Acknowledgments

This work was supported by European Commission FP7 project ROBOT-ERA (grant agreement ICT-288899).

References

Bickmore, T.; Mauer, D.; Crespo, F.; and Brown, T. 2008. Negotiating task interruptions with virtual agents for health behavior change. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 3*, AAMAS '08, 1241–1244. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.

Broz, F.; Nuovo, A. D.; Belpaeme, T.; and Cangelosi, A. 2012. Multimodal robot feedback for eldercare. In *Workshop on Robot Feedback in Human-Robot Interaction: How to make a Robot Readable for a Human Interaction Partner at Ro-Man*, Ro-Man '12.

Di Rocco, M.; Pecora, F.; and Saffiotti, A. 2013. When robots are late: Configuration planning for multiple robots with dynamic goals. In *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, 9515–5922.

Leeper, A.; Hsiao, K.; Ciocarlie, M.; Takayama, L.; and Gossow, D. 2012. Strategies for human-in-the-loop robotic grasping. In *Human-Robot Interaction (HRI), 2012 7th ACM/IEEE International Conference on*, 1–8.

Marge, M.; Pappu, A.; Frisch, B.; Harris, T. K.; and Rudnicky, A. I. 2009. Exploring spoken dialog interaction in human-robot teams. In *Robots, Games, and Research: Success stories in USARSim Workshop at IROS*, IROS '09.

McGee-Lennon, M. R.; Wolters, M. K.; and Brewster, S. 2011. User-centred multimodal reminders for assistive living. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, 2105–2114. New York, NY, USA: ACM.

Miller, C. A., and Parasuraman, R. 2007. Designing for flexible interaction between humans and automation: delegation interfaces for supervisory control. *Human Factors* 49(1):57–75.

Rocco, M. D.; Sathyakeerthy, S.; Grosinger, J.; Pecora, F.; Saffiotti, A.; Bonaccorsi, M.; Cavallo, F.; Limosani, R.; Manzi, A.; Teti, G.; and Dario, P. 2014. A planner for ambient assisted living: From high-level reasoning to low-level robot execution and back. In *AAAI Spring Symposium: Qualitative Representations for Robots*, volume SS-14-06 of *AAAI Technical Report*. AAAI.

Rosenthal, S., and Veloso, M. M. 2010. Mixed-initiative long-term interactions with an all-day-companion robot. In *AAAI Fall Symposium: Dialog with Robots*, volume FS-10-05 of *AAAI Technical Report*. AAAI.

Roth, E. M.; Hanson, M. L.; Hopkins, C.; Mancuso, V.; and Zacharias, G. L. 2004. Human in the loop evaluation of a mixed-initiative system for planning and control of multiple uav teams. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 48(3):280–284.

Warnock, D.; McGee-Lennon, M.; and Brewster, S. 2013. Multiple notification modalities and older users. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, 1091–1094. New York, NY, USA: ACM.