

The “Curious Robot” learns grasping in multi-modal interaction

Ingo Lütkebohle, Julia Peltason, Robert Haschke, Britta Wrede and Sven Wachsmuth

I. INTRODUCTION

Practical robots have come a long way, from being confined to strong cages in industrial manufacturing halls to open environments shared with humans. One consequence if robots are to share spaces with humans is that they must be able to learn from them – that much is well accepted. The reverse – that the robot becomes the “teacher” and the human the “student” – is less commonly seen, however. This is despite the fact that many applications tacitly assume that humans learn about the robot, e.g. from a manual or through instruction by an expert.

We surmise that there is a great deal of potential in an *explicit* reversal of the roles. Therefore, we have investigated how this reversal of the traditional roles can improve HRI. Concretely: *How could a robot structure the dialog such that a naive human partner is aware of her/his possible dialog actions?* The goal is to make humans able to act with confidence despite having absolutely no prior knowledge of either the robot’s goals or its capabilities.

To achieve this ambitious goal, several hard problems must be addressed. One important issue is the vocabulary problem [1], that describes the fact that humans do not know what the system understands, in particular at the beginning of an interaction [2]. Another well known problem is that user’s expectations about a system are strongly shaped by appearance [3], [4], which may lead to erroneous assumptions [5]. Last, but not least, it is not clear how to provide guidance in an easy to understand way and this requires an iterative, study-based approach towards system development [6].

To investigate how robot guidance can improve upon this, we have introduced the “Curious Robot” interactive scenario for learning about real-world objects [5]. In it, we have used a mixed-initiative [7] approach, that has the robot query the human for information at appropriate points during the interaction. For example, the robot queries a human about object labels and how to grasp an object. Initiative is guided by visual saliency information [8].

In this scenario, our results indicate that closed questions provide excellent guidance to the human, resulting in con-

fidant and very consistent (across subjects) answers. The reverse has also been found, with partially open questions leading to considerable more confusion and inconsistency.

II. MULTI-MODAL INTERACTION

A particular problem during speech based interaction is that many activities are hard to describe verbally. Instead, we found that participants prefer a mixture of demonstration and description [5]. Therefore, we have now added hand-posture sensing as an input device to describe grasping using a CyberGlove.

An issue with posture sensing through a glove is to determine when to use it, particularly when motions are only mimicked for demonstration. In the video, we demonstrate how verbal and haptic information are combined to overcome this issue.

III. PROGRESS INQUIRIES

Not surprisingly, we found subjects in the learning scenario to be interested in knowing what the robot has learned. This led them to interrupt the current activity through questions about the system’s knowledge and the current state.

We can accommodate this through our grounding-based dialog [9], which allows nesting of individual exchanges and demonstrate this capability in the video. To encourage subjects to ask questions, the system has also been equipped with voice activity detection, to slow down upon sensing speech. This provides users with feedback that their question is currently possible and the system is attending. The latter aspect is also supported through gaze feedback.

IV. CONCLUSION

We describe how to extend a scenario based on the idea of robot guidance with posture sensing for multi-modal descriptions and improved learning feedback. We also summarize user studies on a previous iteration of the scenario to motivate the chosen approach.

The video is available at <http://aiweb.techfak.uni-bielefeld.de/cr-icair-2010>.

REFERENCES

- [1] G. W. Furnas, T. K. Landauer, L. M. Gomez, and S. T. Dumais, “The vocabulary problem in human-system communication,” *Commun. ACM*, vol. 30, no. 11, pp. 964–971, November 1987. [Online]. Available: <http://dx.doi.org/10.1145/32206.32212>
- [2] M. Hanheide and G. Sagerer, “Active memory-based interaction strategies for learning-enabling behaviors,” in *International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Munich, 2008.
- [3] M. Lohse, F. Hegel, and B. Wrede, “Domestic applications for social robots - a user study on appearance and function,” *Journal of Physical Agents*, vol. 2, pp. 21–32, 2008.

This work was partially supported by the German Federal Ministry of Education and Research (BMBF) under grant no. 01IME01N and the German Research Society (DFG) within the Collaborative Research Centre 673 “Alignment in Communication”.

Ingo Lütkebohle, Julia Peltason and Britta Wrede are with the Applied Informatics Group, Bielefeld University, Germany. {iluetkeb,jpeltaso,bwrede}@techfak.uni-bielefeld.de

Robert Haschke is with the Neuroinformatics Group, Bielefeld University, Germany. {rhaschke}@techfak.uni-bielefeld.de

Sven Wachsmuth is with the Central Lab Facilities, Cognitive Interaction Technology Excellence Cluster, Bielefeld University, Germany. {swachsmu}@techfak.uni-bielefeld.de

- [4] M. Lohse, "The role of expectations in hri," in *New Frontiers in Human-Robot Interaction*, 2009.
- [5] I. Lütkebohle, J. Peltason, L. Schillingmann, C. Elbrechter, B. Wrede, S. Wachsmuth, and R. Haschke, "The Curious Robot - Structuring Interactive Robot Learning," in *International Conference on Robotics and Automation*, Robotics and Automation Society. Kobe, Japan: IEEE, May 2009.
- [6] M. Lohse, M. Hanheide, K. Rohlfing, and G. Sagerer, "Systemic interaction analysis (sina) in hri," in *Conference on Human-Robot Interaction (HRI)*, IEEE. San Diego, CA, USA: IEEE, 11/03/2009 2009, pp. 93–100.
- [7] J. F. Allen, "Mixed-initiative interaction," *IEEE Intelligent Systems*, vol. 14, no. 5, pp. 14–23, 1999.
- [8] Y. Nagai, K. Hosada, A. Morita, and M. Asada, "A constructive model for the development of joint attention," *Connection Science*, vol. 15, no. 4, pp. 211–229, December 2003. [Online]. Available: <http://dx.doi.org/10.1080/09540090310001655101>
- [9] J. Peltason, F. H. Siepmann, T. P. Spexard, B. Wrede, M. Hanheide, and E. A. Topp, "Mixed-initiative in human augmented mapping," in *International Conference on Robotics and Automation*, IEEE. Kobe, Japan: IEEE, 14/05/2009 2009.