Dialogue management for cooperative, symmetrical human-robot interaction

Mary Ellen Foster

Robotics and Embedded Systems Group
Department of Informatics, Technical University of Munich
Boltzmannstraße 3, 85748 Garching, Germany
foster@in.tum.de

Abstract

We describe the JAST human-robot dialogue system, which supports fully symmetrical collaboration between a human and a robot on a joint construction task. We concentrate on the dialogue manager, which is based on Blaylock and Allen's (2005) collaborative problem-solving model of dialogue and which supports joint action between the dialogue participants at both the planning and the execution levels.

1 Human-robot dialogue in JAST

The overall goal of the JAST project ("**J**oint **A**ction **S**cience and **T**echnology"; http://www.euprojects-jast.net/) is to investigate the cognitive and communicative aspects of jointly-acting agents, both human and artificial. The JAST human-robot dialogue system (Foster et al., 2006) is designed as a platform for integrating the project's empirical findings on cognition and dialogue with its work on autonomous robots, by supporting symmetrical human-robot collaboration on a joint construction task.

The robot (Figure 1) consists of a pair of mechanical arms, mounted to resemble human arms, and an animatronic talking head capable of producing facial expressions, rigid head motion, and lip-synchronised synthesised speech. The system input channels are speech recognition, object recognition, and face tracking; the outputs include synthesised speech, facial expressions and rigid head motion, and robot actions. The human user and the robot work jointly to assemble a Baufix wooden construction toy (Figure 2), coordinating their actions through speech, gestures, and facial

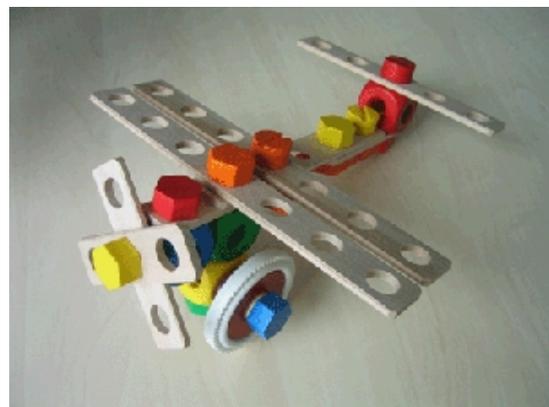
Figure 1: The JAST human-robot dialogue system

Figure 2: Assembled Baufix airplane

motions. Joint action may take several forms in the course of an interaction: for example, the robot may ask the user to provide assistance by holding one part of a larger assembly, or may delegate entire sub-tasks to be done independently. In the current version of the system, the robot is able to manipulate objects in the workspace (e.g., picking them up, putting them down, or giving them to the user) and to perform simple assembly tasks.

2 Dialogue management in JAST

The JAST human-robot dialogue system has several features that distinguish it from many existing dialogue systems. First, the roles of the user and the robot are, in principle, completely symmetrical at all levels: either agent may propose a goal or a strategy for addressing one, and either—or both—may perform any of the actions necessary to achieve it. Also, the interaction must deal with both the selection of the actions to take in the execution of those actions, and may switch between the two tasks at any point. Finally, *joint action* is central to the dialogue at all levels: the participants work together to create domain plans, and also jointly execute the selected plans.

The distinctive requirements of the JAST dialogue system are most similar to those addressed by Blaylock and Allen (2005) in their collaborative problem-solving (CPS) model of dialogue. In collaborative problem solving, multiple agents jointly select and pursue goals, in three interleaved phases: selecting the goals to address, choosing procedures for achieving the goals, and executing the selected procedures. The central process in the CPS model is the selection of values (or sets of values) to fill roles, such as the goal to pursue or the allocation of sub-tasks among the participants. Slot-filler negotiations of this sort make up a large part of collaborative communication.

Dialogue management in the JAST system is based on this CPS model. As in COLLAGEN (Rich et al., 2001), the JAST dialogue state consists of three parts: the active set of goals and procedures, a set of open issues, and the interaction history. An *open issue* corresponds to any request, proposal or action that has occurred during the course of the dialogue and that has not yet been fully addressed; these are essentially the same objects as Ginzburg's (1996) *questions under discussion* (QUD). As an interaction proceeds, two parallel processes are active: the participants must complete domain goals such as locating and assembling objects, and must also address open issues that arise during the conversation. These two processes are tightly linked; for example, if an agent proposes a procedure for a particular sub-goal and the other agrees (and closes the open issue), the next step in the interaction is likely to be executing the agreed-upon sequence of actions. Similarly, when an sub-goal is completed, the participants must address the open issue of how to proceed. The dialogue manager therefore maintains explicit links between the open issues and the current state of the domain plan to enable information to flow in both directions.

3 Current status and future work

At the moment, an initial dialogue-manager prototype based on the CPS model has been implemented in Java. This prototype supports a limited range of simple interactions with a cooperative user, using template expansion to create the domain plans. We are currently developing a more full-featured interaction manager, using a hierarchical planner to create the action sequences. As the system develops, we aim to expand its coverage to support phenomena such as failed actions and incorrect beliefs about the world, and to increase its robustness on incomplete or ill-formed messages from the input-processing modules.

Once a full working dialogue system has been developed, we intend to use it to implement and test the findings from the human-human joint-action dialogues that are currently being recorded and analysed by other participants in the JAST project; for example, we hope to derive strategies for confirmation, grounding, role assignment, and error handling. We will then perform a range of user studies to compare the success of the different strategies, as well as to measure the impact of factors such as feedback from the talking head, using both objective task-success measures and subjective measures of satisfaction and engagement.

References

N. Blaylock and J. Allen. 2005. A collaborative problem-solving model of dialogue. In L. Dybkjær and W. Minker, editors, *Proceedings, 6th SIGdial Workshop on Discourse and Dialogue*, pages 200–211.

M. E. Foster, M. Rickert, and A. Knoll. 2006. Human-robot dialogue for joint-action construction tasks. In *Proceedings, 8th International Conference on Multimodal Interfaces (ICMI 2006)*. To appear.

J. Ginzburg. 1996. Dynamics and the semantics of dialogue. In J. Seligman and D. Westerstahl, editors, *Language, Logic and Computation, Volume 1*, CSLI Lecture Notes.

C. Rich, C. L. Sidner, and N. Lesh. 2001. COLLAGEN: Applying collaborative discourse theory to human-computer interaction. *AI Magazine*, 22(4):15–25.