

Introduction to the Special Issue on Dialogue with Robots

Dan Bohus, Eric Horvitz, Takayuki Kanda, Bilge Mutlu, Antoine Raux

■ This special issue of *AI Magazine* on dialogue with robots brings together a collection of articles on situated dialogue. The contributing authors have been working in interrelated fields of human-robot interaction, dialogue systems, virtual agents, and other related areas and address core concepts in spoken dialogue with embodied robots or agents. Several of the contributors participated in the AAI Fall Symposium on Dialog with Robots, held in November 2010, and several articles in this issue are extensions of work presented there. Others include invited contributions. The articles in this collection address diverse aspects of dialogue with robots, but are unified in addressing opportunities with spoken language interaction, physical embodiment, and enriched representations of context.

For decades, people have reflected with whimsy and imagination upon the possibility of one day conversing with robotic systems in a natural manner as part of fluid, ongoing coordination with them. Research on computational models and mechanisms for supporting spoken dialogue dates back to the earliest days of AI research, including Alan Turing's reflection about how machine intelligence could be evaluated. Fueled by advances in enabling technologies, such as speech recognition, understanding, and synthesis, efforts on supporting fluid and natural conversations between people and computing systems have matured over the last 25 years. Fundamental theories, formalisms, and models for spoken dialogue have been developed and evaluated within the spoken dialogue community. Simple systems, operating within constrained domains and limited communication contexts, have been deployed widely and are in daily use by millions across the planet.

In parallel with these efforts, significant advances have also been made in robotics. Innovations in sensing, reasoning, and manipulation have allowed autonomous robots to move beyond the walls of computing labs into the workplace, home, and street. Bringing robots into real-world environments has made it clear to researchers that robots need not only accurately navigate and manipulate objects, but also to work alongside and, ultimately, *interact* and *collaborate* with humans.

Subsequently, efforts at the intersection of spoken dialogue and human-robot interaction (HRI) have sought to broaden studies of spoken dialogue to richer, more natural, physically situated settings, and have brought to the fore the rich research area of *situated dialogue*, focused on challenges and opportunities at the intersection of natural language, robotics, and commonsense reasoning. Projects in this realm have addressed challenges with the use of dialogue as enabling coordination among multiple actors, taking into consideration not only the details of the task at hand, but also the dynamic physical and social context in which the actors are immersed and the affordances that embodiment provides.

This special issue of *AI Magazine* on dialogue with robots brings together a collection of articles on situated dialogue. The contributing authors have been working in interrelated fields of human-robot interaction, dialogue systems, virtual agents, and

other related areas and address core concepts in spoken dialogue with embodied robots or agents. Several of the contributors participated in the AAAI Fall Symposium on Dialog with Robots, held in November 2010, and several articles in this issue are extensions of work presented there. Others include invited contributions. The articles in this collection address diverse aspects of dialogue with robots, but are unified in addressing opportunities with spoken language interaction, physical embodiment, and enriched representations of context.

The first three articles in this issue focus on research challenges at the intersection of embodiment and human-robot dialogue. Bilge Mutlu provides an overview of embodied communicative cues used in human-human dialogue and summarizes a body of work that showcases how such cues can be used by robots to influence rapport, motivation, persuasion, task performance, and learning in interactive settings. Next, Kerstin Fischer describes opportunities to reduce the mismatch between users' expectations and the actual affordances of robots and their interaction abilities and demonstrates how subtle linguistic cues can be used to shape these expectations. Reid Simmons and his colleagues summarize a body of work aimed at developing believable robot characters that afford everyday interactions in public contexts. The authors discuss the need for tight integration between a consistent storyline, verbal and nonverbal behaviors, and sociocultural context, and highlight some of the challenges and opportunities posed by in-the-wild, long-term deployments of robots that engage in dialogue.

The next three articles describe technical efforts on developing computational models for three core competencies for situated spoken language interaction: turn taking, grounding of spatial language, and spoken language understanding. Andrea Thomaz and Crystal Chao present a framework for turn taking in human-robot interaction, report empirical results that highlight the importance of information flow in coordinating turn taking, and discuss the opportunities around developing a general turn-taking model for HRI. Stefanie Tellex and her colleagues describe a probabilistic inference modeling framework that enables a robot to use data to learn how to ground spatial language, that is, how to map natural language descriptions with concrete objects, places, paths, and events in the real world. Matthias Scheutz, Rehj Cantrell, and Paul Schermerhorn review current research challenges in natural language processing for situated, human-robot interaction settings, and describe an architecture and body of work that begins to address some of these challenges.

The final two articles address important research questions on authoring of human-robot dialogue

systems. Julia Peltason and Britta Wrede conduct an extensive investigation in which they compare and contrast four existing dialogue management frameworks in the context of task-based human-robot dialogue and highlight research challenges in the interplay between dialogue and task modeling. Sonia Chernova, Nick DePalma, and Cynthia Breazeal present a novel approach that involves online crowd sourcing for collecting rich data sets for authoring human-robot dialogue and discuss empirical results and directions for future research on data-driven methods for dialogue authoring.

The articles in this issue frame a rich set of questions on dialogue and other interactions with robots that take into consideration physical embodiment and context. The diversity of questions raised and approaches taken reflect activities brewing within a nascent and vibrant community. We thank the authors for their contributions and hope that the collection will stimulate others to take on the challenges in communication and collaboration between people and intelligent systems in the open world.

Dan Bohus is a researcher at Microsoft Research. His work focuses on the challenges of developing situated interactive systems that can engage in collaboration and dialogue with people in open, dynamic, and relatively unconstrained environments. Prior to joining Microsoft Research, Bohus obtained his Ph.D. degree at Carnegie Mellon University in 2007.

Eric Horvitz has pursued principles of intelligence for successful decision making in the open world, including problems with inference and action under uncertainty arising in collaboration and dialogue. He is a distinguished scientist at Microsoft Research and is the immediate past president of AAAI.

Takayuki Kanda is a senior researcher at ATR Intelligent Robotics and Communication Laboratories, Kyoto, Japan. His current research interests include intelligent robotics, human-robot interaction, and vision-based mobile robots. He received his B.Eng, M.Eng, and Ph.D. degrees in computer science from Kyoto University in 1998, 2000, and 2003, respectively.

Bilge Mutlu is an assistant professor of computer science, psychology, and industrial engineering at the University of Wisconsin–Madison where he directs a research program on designing social robots and their applications in education, health and well-being, and collaborative work. He received his Ph.D. from Carnegie Mellon University's Human-Computer Interaction Institute in 2009.

Antoine Raux is a senior scientist at Honda Research Institute USA, which he joined in 2009. After completing his undergraduate studies at Ecole Polytechnique in France, Antoine Raux received his master's from Kyoto University in 2002 and his Ph.D. from Carnegie Mellon University in 2008, all in computer science. His research interests lie in all aspects of human-machine interaction through spoken and multimodal dialogue, including speech recognition, speech synthesis, turn taking, and language understanding.