

Preface

Human–robot interaction (HRI) is a growing area in several research communities, and its research topics cut across various fields in interaction technology, speech technology, computer science, cognitive science, neuroscience, engineering, and many other related disciplines. It is expected that robots will play a bigger role in future environments and societies, and in particular their capabilities for natural language communication will increase: besides being able to perform various physical tasks, the robot will understand human social signals and support extensive, flexible, adaptable, and reliable spoken language interaction. Such communicating robot agents are generally called *social robots*, and they will provide interactive interfaces to digital data, adapting to the needs and requirements of different users.

In spoken dialogue research, interaction with robots is an increasingly active research and development area in both academia and industry. This is understandable, because intelligent communicative agents have been an application area for dialogue management since the late 1970s, when the first Belief–Desire–Intention (BDI) agents were developed with plan-based dialogue modelling as the mainstream technique. Nowadays spoken dialogue technology is ripe enough to be integrated in the robot platforms, while statistical and rule-based dialogue modelling techniques allow versatile experimentation with natural multimodal communication strategies between humans and intelligent agents. On the other hand, robot applications also provide useful testing grounds for applying dialogue models in practical contexts, as well as evaluating usability of spoken dialogue systems in real contexts. It should be noticed that “intelligent agent” here refers both to robots (situated agents able to move in the physical environment) and to animated and virtual agents (embodied conversational agents that appear on the screen).

Although the title of the book refers to robots, the research area also concerns chat systems. These are popular text-based dialogue systems whose development draws on dialogue modelling, and which can provide tools for data collection and further studies on dialogue management.

As an emergent field of research, human–robot interaction has many open questions. Important topics concern various aspects in multimodal communication,

open-domain dialogue management, emotion detection, social acceptability, and natural interaction modelling. One of the major questions is how to evaluate and benchmark human–robot dialogues, and how to best model issues related to a social robot’s dialogue capabilities. Studies need data from natural interaction situations, and the data is needed for two purposes: to build appropriate models for dialogue interactions, and to evaluate the models and modules, and the interaction as a whole with an intelligent agent system.

Communication via novel interfaces, especially with situated social robots, does not only involve interaction with an information-providing tool but also includes aspects which are typical of human–human communication: timing of responses, feedback-giving strategies, error management, use of multimodal behaviour, and understanding of the partner’s contributions. The effect of a human-like appearance and the acceptability of the application in general are also crucial topics brought in by the development of novel autonomous interactive systems. It is important to study how dialogue strategies are carried over into new situations and new applications, i.e. to explore *affordance* of social robots. Moreover, advances in research and development of autonomous situated agents propel discussions on the ethics of social robots: their decision-making, application domains, conflicting goals between robots and their users, and the responsibilities of researchers. Such issues elevate research and development on social robotics to another level of social research: that of improving human well-being.

The book describes work on spoken dialogue systems and intelligent social agents. The chapters explore novel ways to model and use interaction strategies, and present improvements and new developments beyond previous work. The contributions are selected, adapted, and expanded from the papers presented at the Seventh International Workshop on Spoken Dialogue Systems (IWSDS 2016), and they offer different perspectives and solutions for the important questions, through papers by leading researchers covering key topics in the field. We hope the book will contribute to the design and study of interaction patterns that support successful human–robot communication, and will encourage ethical discussion and evaluation of the applications.

The IWSDS 2016 Conference

The series of IWSDS conferences aims to bring together researchers from both industry and academia working within the various disciplines that constitute the field of spoken dialogue systems. The Seventh IWSDS (IWSDS 2016) was held from January 13 to 16, 2016 at Hotel Riekonlinna in Saariselkä, Finland. This was the northernmost spoken dialogue workshop ever, and it allowed participants to discuss implementation and analytical work as well as to enjoy beautiful winter scenes and activities in Finnish Lapland. The *Kaamos*, the Arctic night when the sun stays below the horizon, had just ended, and at the time of the conference the daylight was less than three hours long; however, it contributed to wonderful sunrises and sunsets.

The special theme of IWSDS 2016 was *Dialogues with Social Robots: Enablements, Analyses, and Evaluation*. The full programme included presentations and demonstrations, a gala dinner and a trip to Inari to learn about the indigenous Sami culture at the Siida Museum for the Sami people and at Sajos, the Sami Parliament in Finland.

The conference also featured two invited speakers. Riitta Hari (Aalto University) spoke *About the Primacy of Social Interaction in Human Brain Function*, describing how current neuroscience, even “social neuroscience”, examines brain functions of isolated persons who observe other persons’ actions from a third-person view. Professor Hari advocated the idea that research on the brain basis of social cognition and interaction should move from this “spectator science” to studies of engaged participants. Recent advances in neuroimaging now allow study of the brain basis of *social interaction* by simultaneous recordings of brain activity of two participants engaged in natural communication. An important research question is whether social interaction emerges from lower-level perceptual, motor and cognitive functions, as is typically assumed, or whether social interaction is in fact the primary default mode governing perception and action, which would challenge many current ideas about human brain function.

David Traum (ICT, University of Southern California) spoke on *The Role of a Lifetime: Dialogue Models for Virtual Human Role-players*. Some dialogue systems are instruments to allow a user to efficiently solve a task, but other dialogue systems act as role-players in learning exercises, games, or other activities. For these cases, human-like dialogue is more important, especially if the purpose of the dialogue activity is to learn or practice interactive skills that should transfer to other humans. However, for role-play dialogue the most appropriate metric is not the familiar “Turing test” of indistinguishability from humans, but rather activation of the same sociocognitive skills that are appropriate for human interaction. Professor Traum presented examples of role-play dialogue systems from a wide variety of activities, genres, and roles, focussing on virtual humans created at the USC Institute for Creative Technologies.

Besides the normal sessions of papers and posters, the conference included three special sessions: *Dialogue State Tracking Challenge 4*, *Evaluation of Human–Robot Dialogue in Social Robotics*, and *Sociocognitive Language Processing*, which are briefly described below.

Dialogue State Tracking Challenge 4

Organisers: Luis F. D’Haro, Seokhwan Kim, Rafael E. Banchs, Matthew Henderson and Jason Williams.

Dialogue state tracking is one of the key sub-tasks of dialogue management, which defines the representation of dialogue states and updates them at each moment in a given ongoing conversation. In this challenge, participants used the TourSG corpus to develop the components. TourSG consists of dialogue sessions

about touristic information for Singapore collected from Skype calls between 3 tour guides and 35 tourists. The challenge included a main task on dialogue state tracking at sub-dialogue level; four optional pilot tasks on spoken language understanding, speech act prediction, spoken language generation, and the implementation of end-to-end systems; and an optional open track for participants to explore any task of their interest over the provided dataset.

Evaluation of Human–Robot Dialogue in Social Robotics

Organisers: Laurence Devillers, Kristiina Jokinen, Joseph Mariani, Haizhu Li, Alex Waibel and Wolfgang Minker.

The aim of this special theme was to make a bridge between the social robotics and spoken dialogue research communities. In the spoken dialogue community the focus is on evaluation of verbal interaction including spontaneous speech recognition and understanding, whereas in the social robotics community the focus is on evaluation of engagement measures in interaction, with non-verbal features such as acoustics and gestures but without spontaneous speech information. Several robotics competitions already incorporate human-robot interaction (such as RoboCup@Home, which has benchmarks to evaluate robots in realistic home environments, or the AAIL Grand Challenge in which the robot attends a conference and delivers a talk) but these competitions do not focus on interactive spoken dialogues. Similarly, several spoken dialogue competitions have been proposed by Darpa (such as ATIS or Communicator), but as yet there are no established protocols and no ongoing evaluation campaigns for human-robot interaction.

Sociocognitive Language Processing

Organisers: Björn Schuller and Michael McTear.

Sociocognitive Language Processing (SCLP) is the idea of coping with everyday language, including slang and multilingual phrases and cultural aspects, and in particular with irony/sarcasm/humour and paralinguistic information such as the physical and mental state and traits of the dialogue partner (affect, age group, personality dimensions) and social aspects. Multimodal aspects such as facial expression, gestures, or bodily behaviour should also be included in the analysis. At the same time, SCLP can render future dialogue systems more “chatty” by not only feeling natural but also being emotionally and socially competent, ideally leading to a more symmetrical dialogue. For that, the computer should itself have a “need for humour” and an “increase of familiarity”, enabling computers to experience or at least better understand emotions and personality.

We warmly thank the invited speakers, the steering committee, the reviewers in the scientific committee, the organisers of the special sessions, the members of the local arrangements committee, and all the participants for their contributions to the success of IWSDS 2016.

Helsinki, Finland
July 2016

Kristiina Jokinen
Conference Chair IWSDS 2016
Graham Wilcock
Local Arrangements Chair IWSDS 2016

Dialogues with Social Robots

Enablements, Analyses, and Evaluation

Jokinen, K.; Wilcock, G. (Eds.)

2017, XX, 489 p. 117 illus., 70 illus. in color., Hardcover

ISBN: 978-981-10-2584-6