WILEY SERIES IN AGENT TECHNOLOGY

WILEY

Constructive Dialogue Modelling

Speech Interaction and Rational Agents

Kristiina Jokinen

TP18
J74

CONSTRUCTIVE DIALOGUE MODELLING

SPEECH INTERACTION AND RATIONAL AGENTS

Kristiina Jokinen
University of Helsinki, Finland

WILEY
A John Wiley and Sons, Ltd., Publication

This edition first published 2009
© 2009 John Wiley & Sons, Ltd

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

The right of the author to be identified as the author of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data

Jokinen, Kristiina.
Constructive dialogue modelling: speech interaction and rational agents/Kristiina Jokinen.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-06026-1 (cloth)
1. Human-computer interaction. 2. Automatic speech recognition. 3. Intelligent agents (Computer software) 4. Dialogue–Computer simulation. I. Title.
QA76.9.H85J65 2009
004.01'9–dc22

2008046994

A catalogue record for this book is available from the British Library.

ISBN 978-0-470-06026-1(H/B)

Typeset in 11/13 Times by Laserwords Private Limited, Chennai, India
Printed and bound in Great Britain by CPI Antony Rowe, Chippenham, Wiltshire

CONSTRUCTIVE DIALOGUE MODELLING

SPEECH INTERACTION AND RATIONAL AGENTS

Wiley Series in Agent Technology

Series Editor: Michael Wooldridge, *University of Liverpool, UK*

The 'Wiley Series in Agent Technology' is a series of comprehensive practical guides and cutting-edge research titles on new developments in agent technologies. The series focuses on all aspects of developing agent-based applications, drawing from the Internet, Telecommunications, and Artificial Intelligence communities with a strong applications/technologies focus.

The books will provide timely, accurate and reliable information about the state of the art to researchers and developers in the Telecommunications and Computing sectors.

Titles in the series:

Padgham/Winikoff: Developing Intelligent Agent Systems 0-470-86120-7 (June 2004)

Bellifemine/Caire/Greenwood: Developing Multi-Agent Systems with JADE 0-470-05747-5 (February 2007)

Bordini/Hübner/Wooldrige: Programming Multi-Agent Systems in AgentSpeak using Jason 0-470-02900-5 (October 2007)

Nishida: Conversational Informatics: An Engineering Approach 0-470-02699-5 (November 2007)

此为试读,需要完整PDF请访问: www.ertongbook.com

Foreword

Although we might like to believe otherwise, the truth is that computers are strange and difficult beasts. The most beautiful user interface in the world can at times be baffling, leaving us bewildered and confused. The most carefully engineered program can crash, leaving us frustrated and angry. And of course, all these problems stem from the fact that computers are not like me and you: although we dress them up to appear friendly, colourful, and helpful, they ultimately understand the world only as strings of 1s and 0s. There is a fundamental barrier to communication between human and machine. To fully exploit the potential of computers, we need machines that can relate to us in our terms; that can model, understand, and make use of the methods of communication that we use as humans. The present text addresses exactly these issues. It presents a thorough overview of the area of interaction, focusing particularly on the idea of interaction as a rational process, and the notion of rational dialogue. As well as considering the underlying principles, the book makes a solid contribution to the engineering of dialogue and interaction systems. It represents a valuable step in the path to building systems that can overcome the fundamental barrier between human and machine.

Michael Wooldridge

Preface

State-of-the-art speech and language technology has reached a level that allows us to build interactive applications which the users can have short conversations with in order to search for information. We are already dealing with electronic banking facilities, information providing systems, restaurant guides, timetable services, assisting translation systems, emails, web browsers, etc., which can understand the users' speech and which can also reply using speech. Speech-based interfaces have also brought forward novel applications for situations where interactions using keyboard or mouse clicks are cumbersome or are not possible at all: car navigation, telephone services, home appliances, etc. Moreover, they also provide solutions for users for whom the ordinary mouse and keyboard would not be a possible means of interaction with computer services, thus contributing to the requirement of universal access to digital services and databases. However, the challenge that present-day speech and language research faces in an ever-expanding information society is not so much in producing tools and systems that would enable interaction with automatic services in the first place, but rather, to design and build systems that would allow interaction to take place in a natural way.

This book is about natural interaction in dialogue management. It introduces the Constructive Dialogue Modelling (CDM) approach to interaction modelling, based on a view of dialogue as a shared activity. CDM focuses especially on the aspects of rational and cooperative communication that allow humans to transmit, exchange, mediate, argue and ask for information in an efficient and flexible manner. The CDM approach provides an account of various issues concerning cooperative communication, in which the participants together construct dialogues by exchanging new information on a particular topic in a given context.

The book is not about communication and debate, rhetoric or persuasion, and it does not teach successful conversational styles or strategies. Rather, it focuses on the conversational features and preconceptions that would make the interaction between humans and computers more natural and intuitive. From the interaction design point of view, the goal can be expressed as increasing the usability of speech-based interactive systems: making interfaces more usable and intuitive to use, giving extra value to the system by enabling it to converse naturally with the user.

In order to achieve this goal, we need to equip systems with a component that takes care of interaction management and uses a special model to extract information that goes beyond the compositional meaning of the observed utterances and words. In other words, we need to abstract the intended meaning from the observed utterances and model the speaker's intentions and goals and roles in the activities the speaker is engaged in. Deliberations about the next dialogue action take place in this kind of intention space, and the observed utterances function as a kind of a trigger for the speakers to update the intention space appropriately. This means that usability and user-friendliness in speech-based interactive interfaces are related to the system's communicative capabilities. The users' perception of the system depends on the system's communicative capabilities, which affect the user's view of how satisfactorily the system assists the user in the task that it is meant to accomplish, and how much the system's operation can be trusted. Natural communication appears to support user satisfaction even if the interaction would contain such obviously undesirable features like long waiting times and minor errors: it provides a means to resolve misunderstandings by negotiation and talking. Thus it becomes possible to build systems that function in a more satisfactory manner from the user's point of view: interactions are perceived by the user as useful for the task at hand. It is the system's intelligent interaction that provides the basis for good interfaces and services.

The book provides a concise overview of the different interactional models as well as concepts that enable us to build practical interactive systems, and to test hypotheses for friendly and flexible human-computer interaction, with special attention paid to multimodality. It is intended for communication researchers and computer scientists aiming to design complex interactive systems and experimenting with various types of complex systems using natural language. Taking the view that computers are not only tools but agents that users need to interact with, the possibilities for making interaction more flexible and natural by taking some human language capabilities into account are examined. In particular, the focus is on the aspects of interaction that contribute to smoothness of communication, and on adaptation of the system to the user's level of expertise.

It has often been argued that theoretical approaches to dialogue management produce descriptive models which do not necessarily address concrete problems in interaction technology, while practical approaches provide heuristic and ad-hoc solutions which are not easy to extend to other domains or applications. The CDM model attempts to bring the two views together by insisting on a theoretical rather than add-hoc basis for modelling, and on proof-of-concept experimentation and open evaluation with different systems. The theoretical part of the CDM model is drawn from the existing approaches to dialogues, especially from dialogue planning and rational agency, and from empirical research on human-human communicative behaviour. Practical aspects are related

to various research projects where the CDM model has been applied and partially implemented.

Content of the Book

The content of the book can be divided into two parts. The first two chapters form the introductory part where the background and the starting point are introduced. The chapters provide an overview of dialogue management and the current state of the art in general. The second part of the book provides a more detailed account of the CDM approach and various topics related to its development, including discussion of some future views of cooperative, rational dialogue management. In detail, the book is organised as follows.

After the introduction to human-human and human-computer communication in Chapter 1, Chapter 2 introduces different dialogue management models and dialogue systems as well as system architectures and representations. It also highlights the problematic areas of current state-of-the-art interactive systems so as to lay a foundation for the following chapters. The work done in different projects on adaptive speech-based human-computer interaction is also reviewed, providing an overview of the various interactive systems. Chapter 3 then proceeds by presenting the Constructive Dialogue Model (CDM) and the basic concepts that deal with cooperation, rationality and adaptation. Chapter 4 discusses a few examples of the implementation of the basic concepts in dialogue systems. The book then continues with two specific topics: Chapter 5 discusses the management of dialogue and domain knowledge which is necessary for a dialogue system in order to exhibit intelligent interaction, and Chapter 6 focuses on the learning and adaptation necessary for the system to operate and manage interactions in dynamic contexts. Finally, Chapter 7 concludes with future challenges, and contains a roadmap of the interactive systems that we might expect to see in ten years' time.

Introduction

In this chapter, the main objectives of the book are introduced. The new metaphor for human-computer interaction, i.e. the computer as an agent as opposed to a tool, is presented and discussed. The emphasis is on the view of the dialogue as "dialogue in natural language", as opposed to "dialogue by icon clicking". The challenges and opportunities that this kind of interaction bring to speech-based application design are discussed, in particular the need for the modelling of rationality, agenthood, cooperation and natural language interaction. Speech adds much more than just a sound to interactive applications, and presupposes that the different aspects are addressed properly. The main claim of the book is also put forward in the chapter: that natural language communication is the genuine human factor necessary for building flexible and intelligent interactive systems Combined with the agent metaphor, this claim will be substantiated in the Constructive Dialogue Model.

Dialogue Models

This chapter provides the background for the development of CDM. It surveys various dialogue models and their implementation as dialogue management engines, and also gives a short history of (natural language) interactive systems and dialogue management development as opposed to point-and-click interface design. The systems are highlighted as examples of the ideas which have functioned as sources of inspiration for CDM rather than providing a comprehensive overview of the field. Dialogue modelling and practical system building are contrasted, so as to go back to the goals of Chapter 1 in order to argue that we need a deeper understanding of rationality and natural language conversation modelling so we can develop intelligent applications.

Constructive Dialogue Model

In this chapter, the framework of the Constructive Dialogue Model is presented. The basic principles related to the activity-based analysis of communication, rational agency, and to the concept of Ideal Cooperation are discussed. The construction of shared context and mutual knowledge are also studied, with reference made to the previous work on grounding. The possibilities for enabling conversations with computer agents are investigated, and special attention is paid to the socio-cultural context in which the dialogues take place. The focus is on dialogue obligations and trust as indications that the participants are rational agents.

Construction of Dialogue and Domain Information

This chapter digs deeper into one particular aspect that is essential for CDM: the process of planning and producing responses that would be perceived as natural and intuitive reactions in the ongoing dialogue. The starting point is the information that is exchanged in contributions to the dialogue, and the basic unit for this is the concept of NewInfo (new information), defined in terms of intonation phrases. The construction of a shared context takes place by evaluating and accommodating NewInfo with respect to one's own understanding of the dialogue situation, and providing implicit or explicit feedback of the success of this accommodation to the partner. While it is obvious that some higher-level expectations of the goal and the appropriate dialogue strategies are needed in order to guide reasoning, the basic approach in the accommodation is bottom-up: the identification of the NewInfo and its accommodation with the current dialogue situation is managed locally. The actual structure of the dialogue is recorded in the conceptual links between the pieces of NewInfo, and can be constructed by tracing the paths that show how the linking of NewInfo to the dialogue topic has manifested itself in the course of the interaction. In practical dialogue systems, the reasoning itself requires well-defined and rich semantic representation and an ontology-based reasoning engine.

Dialogue Systems

In this chapter, different dialogue system architectures, with special reference to agent-based architectures, and some basic dialogue management techniques are introduced. Suitable representations are also discussed briefly. The chapter clarifies what has been implemented and what is capable of implementation on one hand, and what is still required for further experimentation and research, on the other. The chapter attempts to establish where CDM can go beyond application development and then furnish interactive systems, taking a wider view about how spoken language interaction takes place and how it could be managed. Examples of CDM-based dialogue systems are also presented.

Constructive Information Technology

Some variations of the CDM-style dialogue management are discussed in this chapter. In particular, issues related to adaptivity and learning are investigated, as well as the notion of "full-blown communication" in the context of multimodal and non-verbal communication. This brings us back to the discussion of the general applicability of the aspects of human-human communication to human-computer interaction, and to the main claim that the genuine human factor in the design and development of intelligent interactive systems is the system's ability to use natural language communication, i.e. usable, flexible, and robust interactive systems should *afford* natural interaction.

Conclusions and Future Views

This chapter summarises the contributions of the book and presents some future challenges for CDM-based interaction management. It defines the book as a theoretically-based overview of interaction technology, which will hopefully be useful for researchers and students of interactive systems as well as for developers of commercial spoken dialogue systems.

此为试读,需要完整PDF请访问: www.ertongbook.com

Acknowledgements

Writing this book has been a lengthy and challenging project and would not have been possible without the help and support of many colleagues and friends. It is impossible to list all of those who have contributed to the ideas and their development and implementation in one way or another, either as my students, team members, and co-workers in research projects, or as dialogue partners and commentators in meetings and casual coffee conversations. A collective thank you must suffice to express my gratitude to all of you.

However, there are a few people I would especially like to mention since their influence has been crucial in various stages of the writing, with explicit comments and/or advice on the text.

First, I would like to mention Jens Allwood (University of Gothenburg, Sweden) whose work on Activity-based Communication Analysis has long been a source of inspiration for the CDM model. I have enjoyed our many discussions and arguments on the linguistic-philosophical aspects of communication, rational agents, communicative functions and language-based interaction, and his comments on the earlier versions of some of the chapters have been insightful.

I would also like to mention Michael McTear (University of Ulster, Northern Ireland) who has supplied me with another angle on spoken interaction. Through commenting on some of the chapters, and collaborating in various workshop organisation and writing activites, Mike has provided direct and indirect help with a more practical view of interaction, as well as useful guidance for developing advanced dialogue systems.

The late Karen Spärck-Jones (University of Cambridge, UK) should also be mentioned here. Besides providing a general model for scientific thinking, her encouragement and sharp observations have greatly shaped the structure and arguments of the book, especially in the early phase of its writing, when I was a visting Fellow at the Computer Laboratory and had the pleasure of using her "library" as my office.

I would also like to mention my husband Graham Wilcock, for his comments on many of the chapters, including his patient corrections of typos and grammar. Without his continuous support and love the whole book would have been impossible.

Finally, I would like to thank my publishers for their patience and kind understanding throughout the various stages of the project.

Kyoto

Contents

Foreword ix

Preface xi

Acknowledgements xvii

1 Introduction 1

1.1 Two Metaphors for Interaction Design 1
1.1.1 The Computer as a Tool 1
1.1.2 The Computer as an Agent 6
1.2 Design Models for Interactive Systems 11
1.2.1 Reactive Interaction 12
1.2.2 Interaction Modelling 13
1.2.3 Communication Modelling 14
1.3 Human Aspects in Dialogue System Design 17
1.3.1 User-centred Design 17
1.3.2 Ergonomics 18
1.3.3 User Modelling 19
1.3.4 Communication as a Human Factor 20

2 Dialogue Models 23

2.1 Brief History 24
2.1.1 Early Ideas 24
2.1.2 Experimental Prototypes 28
2.1.3 Large-scale Dialogue Projects 30
2.1.4 From Written to Spoken Dialogues 31
2.1.5 Dialogue Corpora 33
2.1.6 Dialogue Technology 34
2.2 Modelling Approaches 36
2.2.1 Grammar-based Modelling 37

2.2.2 *Intention-based Modelling* 39
2.2.3 *Corpus-Based Modelling* 42
2.2.4 *Conversational Principles* 45
2.2.5 *Collaboration and Teamwork* 46
2.3 Dialogue Management 47
2.3.1 *Dialogue Control* 49
2.3.2 *Representation* 50

3 Constructive Dialogue Model (CDM) 53

3.1 Basic Principles of Communication 54
3.1.1 *Levels of Communication* 55
3.1.2 *Rational Agency* 57
3.1.3 *Ideal Cooperation* 60
3.1.4 *Relativised Nature of Rationality* 62
3.2 Full-blown Communication 64
3.2.1 *Communicative Responsiveness* 65
3.2.2 *Roles and Activities of Participants* 67
3.3 Conversations with Computer Agents 69

4 Construction of Dialogue and Domain Information 73

4.1 Coherence and Context – Aboutness 73
4.2 Information Structure of Utterances – New and Old Information 76
4.3 Definitions of NewInfo and Topic 80
4.3.1 *Discourse Referents and Information Status* 82
4.3.2 *Other Two-dimensional Work* 84
4.4 Topic Shifting 87
4.5 Information Management as Feedback Giving Activity 89
4.6 Information Management and Rational Agents 95

5 Dialogue Systems 99

5.1 Desiderata for Dialogue Agents 99
5.1.1 *Interface* 99
5.1.2 *Efficiency* 100
5.1.3 *Natural Language Robustness* 101
5.1.4 *Conversational Adequacy* 101
5.2 Technical Aspects in CDM 102
5.2.1 *Distributed Dialogue Management* 102
5.2.2 *Generation from NewInfo* 105
5.2.3 *Adaptive Agent Selection* 107
5.2.4 *Multimodal Route Instructions* 109
5.3 Summary 111

6 Constructive Information Technology **113**

6.1 Learning and Adaptation 113
6.1.1 Adaptive User Modelling 116
6.1.2 Dialogue Strategies 118
6.2 Cognitive Systems and Group Intelligence 120
6.3 Interaction and Affordance 124

7 Conclusions and Future Views **127**

References **133**

Index **155**

1

Introduction

1.1 Two Metaphors for Interaction Design

1.1.1 The Computer as a Tool

The world around us is changing rapidly and we are increasingly surrounded by electronic devices which provide us with various types of information. Internet and mobile applications are commonplace and future predictions describe a world where computers are embedded increasingly in the environment in which we live and in the ways in which we work: in a few years' time, for instance intelligent home robotics will cater for our many everyday needs, and we will be required to interact with a complex environment which will consist not only of people, but also of computers embedded in our daily surroundings. Technological roadmaps have envisaged how smart environments will be populated by several context-aware devices which communicate with each other and with the users, and how the future information and communication systems will contain computers built into products such as clothes, books, beds and sporting gear (Sjöberg & Backlund, 2000; Plomp et al., 2002). The systems will identify their current context of use, adapt their behaviour and allow for natural interaction. Computers will also have senses and be able to interpret human expressions, will be able to smell, feel, hear, see and taste, and there will be intuitive human-computer interfaces that mimic human communication.

In such environments, our interactions with the computer will also become more complex. The internet provides an information source as well as an infrastructure for embedded computing and virtual interaction, while mobile communication and wireless appliances can be deployed for building various digital services. Many of our everyday tasks already require the use of a computer as a tool: text editing, image processing, communication with friends and colleagues by emails and web-phones, information seeking in the internet and in various digital databases. Computers also enter into human activities where they replace the other participant: interactive situations with automatic services are commonplace, and different types of online services (news, banking, shopping, hotel and flight

Constructive Dialogue Modelling Kristiina Jokinen
© 2009 John Wiley & Sons, Ltd

此为试读,需要完整PDF请访问: www.ertongbook.com