Theory and Applications of Natural Language Processing

Series Editors:
Graeme Hirst (Textbooks)
Eduard Hovy (Edited volumes)
Mark Johnson (Monographs)
Aims and Scope

The field of Natural Language Processing (NLP) has expanded explosively over the past decade: growing bodies of available data, novel fields of applications, emerging areas and new connections to neighboring fields have all led to increasing output and to diversification of research.

“Theory and Applications of Natural Language Processing” is a series of volumes dedicated to selected topics in NLP and Language Technology. It focuses on the most recent advances in all areas of the computational modeling and processing of speech and text across languages and domains. Due to the rapid pace of development, the diversity of approaches and application scenarios are scattered in an ever-growing mass of conference proceedings, making entry into the field difficult for both students and potential users. Volumes in the series facilitate this first step and can be used as a teaching aid, advanced-level information resource or a point of reference.

The series encourages the submission of research monographs, contributed volumes and surveys, lecture notes and textbooks covering research frontiers on all relevant topics, offering a platform for the rapid publication of cutting-edge research as well as for comprehensive monographs that cover the full range of research on specific problem areas.

The topics include applications of NLP techniques to gain insights into the use and functioning of language, as well as the use of language technology in applications that enable communication, knowledge management and discovery such as natural language generation, information retrieval, question-answering, machine translation, localization and related fields.

The books are available in printed and electronic (e-book) form:

* Downloadable on your PC, e-reader or iPad
* Enhanced by Electronic Supplementary Material, such as algorithms, demonstrations, software, images and videos
* Available online within an extensive network of academic and corporate R&D libraries worldwide
* Never out of print thanks to innovative print-on-demand services

For further volumes:
[http://www.springer.com/series/8899](http://www.springer.com/series/8899)
Reinforcement Learning for Adaptive Dialogue Systems

A Data-driven Methodology for Dialogue Management and Natural Language Generation
Preface

The past decade has seen a revolution in the field of spoken dialogue systems. As in other areas of Computer Science and Artificial Intelligence, data-driven methods are now being used to drive new methodologies for system development and evaluation. These methods are proving to be more robust, flexible, and adaptive than the largely rule-based approaches which preceded them.

We hope that this book is a contribution to that ongoing change. It describes, in detail, a new methodology for developing spoken dialogue systems – in particular the Dialogue Management and Natural Language Generation components – which starts with human data, and culminates in evaluation with real users. The journey therefore starts and ends with human behaviour in interaction, and explores methods for learning from the data, for building simulation environments for training and testing systems, and for evaluating the results.

The detailed material covers: Spoken and Multimodal dialogue systems, Wizard-of-Oz data collection, User Simulation methods, Reinforcement Learning, and Evaluation methodologies.

This book is therefore intended as research guide which navigates through a detailed case study in data-driven methods for development and evaluation of spoken dialogue systems. Common challenges associated with this approach are discussed and example solutions provided, for example, how to learn from limited amounts of data. As such, we hope it will provide insights, lessons, and inspiration for future research and development – not only for spoken dialogue systems in particular, but for data-driven approaches to human-machine interaction in general.

Edinburgh, September 2011

Verena Rieser
Oliver Lemon
Acknowledgements

The research presented here would not have been possible without interaction with our research community, and we owe special thanks to the participants of the TALK and CLASSiC projects for discussing the ideas presented in this book.

There is a lively international research community in Spoken Dialogue Systems whose interest, criticism, and encouragement have helped to shape this work. We are especially indebted to Steve Young and his colleagues at Cambridge University, who have had a deep influence on our ideas. Members of the Interaction Lab1 at Heriot-Watt University’s School of Mathematics and Computer Science have also helped to discuss and develop these ideas over several years.

In particular we would also like to thank Dr. Xingkun Liu, Dr. Helen Hastie, Dr. Ivana Kruijff-Korbayová, Dr. Tilman Becker, and other colleagues from Saarland and Edinburgh Universities for helping with the data collection and evaluation involved in this work. We especially thank Professor Manfred Pinkal for his guidance.

The research leading to these results has received funding from the European Community’s Seventh Framework Programme (FP7, 2007-2013) under grant agreement number 216594 (“Computational Learning in Adaptive Systems for Spoken Conversation”, CLASSiC project: www.classic-project.org), the EC FP6 project “TALK: Talk and Look, Tools for Ambient Linguistic Knowledge” (IST 507802, www.talk-project.org), from EPSRC grant numbers EP/E019501/1 and EP/G069840/1, and from the International Research Training Group “Language Technology and Cognitive Systems”, Saarland University, funded by the German Science Foundation (DFG).

Finally, the first author would like to thank her parents Franz and Tatjana Rieser for their support and encouragement. The second author thanks his family for the decades of rewarding learning experiences which have made this book possible.

1 http://www.macs.hw.ac.uk/InteractionLab/
# Contents

1 Introduction ................................................... 1
  1.1 The Design Problem for Spoken Dialogue Systems ............. 1
  1.2 Overview .................................................. 2
  1.3 Structure of the Book ........................................ 4

Part I Fundamental Concepts

2 Background .................................................... 9
  2.1 Human-Computer Interaction ................................. 10
  2.2 Dialogue Strategy Development .............................. 11
    2.2.1 Conventional Development Lifecycle .................... 12
    2.2.2 Evaluation and Strategy Quality Control ................ 13
    2.2.3 Strategy Implementation ................................ 17
    2.2.4 Challenges for Strategy Development .................... 19
  2.3 Literature review: Learning Dialogue Strategies ............ 21
    2.3.1 Machine Learning Paradigms ............................ 21
    2.3.2 Supervised Learning for Dialogue Strategies .......... 22
    2.3.3 Dialogue as Decision Making under Uncertainty ....... 23
    2.3.4 Reinforcement Learning for Dialogue Strategies ...... 24
  2.4 Summary .................................................... 26

3 Reinforcement Learning ........................................ 29
  3.1 The Nature of Dialogue Interaction .......................... 30
    3.1.1 Dialogue is Temporal .................................. 30
    3.1.2 Dialogue is Dynamic .................................. 31
  3.2 Reinforcement Learning-based Dialogue Strategy Learning ..... 32
    3.2.1 Dialogue as a Markov Decision Process ................. 32
    3.2.2 The Reinforcement Learning Problem .................... 36
    3.2.3 Model-based vs. Simulation-based Strategy Learning .... 42
  3.3 Dialogue Simulation .......................................... 45
    3.3.1 Wizard-of-Oz Studies .................................. 45
3.3.2 Computer-based Simulations ................................................. 46
3.3.3 Discussion .......................................................................... 47
3.4 Application Domains ................................................................. 48
  3.4.1 Information-Seeking Dialogue Systems ................................. 48
  3.4.2 Multimodal Output Planning and Information Presentation .... 49
  3.4.3 Multimodal Dialogue Systems for In-Car Digital Music Players ............................................................................ 52
3.5 Summary ..................................................................................... 52

4 Proof-of-Concept: Information Seeking Strategies .................................. 53
  4.1 Introduction ............................................................................. 53
    4.1.1 A Proof-of-Concept Study ...................................................... 54
  4.2 Simulated Learning Environments .............................................. 55
    4.2.1 Problem Representation .......................................................... 55
    4.2.2 Database Retrieval Simulations ............................................... 56
    4.2.3 Noise Model ........................................................................ 57
    4.2.4 User Simulations .................................................................. 58
    4.2.5 Objective and Reward Function ............................................. 59
    4.2.6 Application Scenarios ............................................................. 60
  4.3 Threshold-based Baseline ............................................................ 61
  4.4 Reinforcement Learning Method ................................................. 63
    4.4.1 Training the Policies ............................................................... 63
  4.5 Results ....................................................................................... 65
  4.6 Summary ..................................................................................... 69

Part II Policy Learning in Simulated Environments

5 The Bootstrapping Approach to Developing Reinforcement Learning-based Strategies ......................................................... 73
  5.1 Motivation ................................................................................. 74
    5.1.1 Term Definition ..................................................................... 75
    5.1.2 Related Work ....................................................................... 76
  5.2 Advantages for Learning from WOZ Data ..................................... 77
    5.2.1 Challenges for Learning from WOZ Data ................................... 78
  5.3 The Bootstrapping Method .......................................................... 79
    5.3.1 Step 1: Data Collection in a Wizard-of-Oz Experiment ............ 79
    5.3.2 Step 2: Build a Simulated Learning Environment ................... 81
    5.3.3 Step 3: Train and test a strategy in simulation ......................... 81
    5.3.4 Step 4: Test with Real Users .................................................. 82
    5.3.5 Step 5: Post-Evaluation ......................................................... 82
  5.4 Summary ..................................................................................... 82
6 Data Collection in a Wizard-of-Oz Experiment .............................................. 85
   6.1 Experimental Setup ......................................................................................... 86
      6.1.1 Recruited Subjects: Wizards and Users .................................................. 89
      6.1.2 Experimental Procedure and Task Design .............................................. 90
   6.2 Noise Simulation ............................................................................................ 90
      6.2.1 Related Work .......................................................................................... 90
      6.2.2 Method .................................................................................................. 91
      6.2.3 Results and Discussion .......................................................................... 91
   6.3 Corpus Description ........................................................................................ 92
   6.4 Analysis .......................................................................................................... 94
      6.4.1 Qualitative Measures .............................................................................. 94
      6.4.2 Subjective Ratings from the User Questionnaires .................................. 95
   6.5 Summary and Discussion .............................................................................. 98

7 Building Simulation Environments from Wizard-of-Oz Data ...................... 101
   7.1 Dialogue Strategy Learning with Simulated Environments ......................... 101
      7.1.1 Method and Related Work ...................................................................... 103
      7.1.2 Outline .................................................................................................. 106
   7.2 Database Description ...................................................................................... 107
   7.3 Action Set Selection ........................................................................................ 108
      7.3.1 Method and Related Work ...................................................................... 108
      7.3.2 Annotation Scheme ................................................................................ 108
      7.3.3 Manual Annotation ................................................................................ 110
      7.3.4 Action Set for Learning ......................................................................... 111
   7.4 State Space Selection ...................................................................................... 112
      7.4.1 Method and Related Work ...................................................................... 112
      7.4.2 Task-based State Space Features ............................................................ 113
      7.4.3 Feature Selection Techniques for Domain-specific State Space Features .................................................. 114
      7.4.4 Further System Behaviour ...................................................................... 118
   7.5 MDP and Strategy Design .............................................................................. 118
      7.5.1 Motivation ............................................................................................... 118
      7.5.2 Implementation ....................................................................................... 118
      7.5.3 Hierarchical Reinforcement Learning in the ISU Approach ................ 119
      7.5.4 Further System Behaviour ...................................................................... 120
   7.6 Wizard Behaviour ........................................................................................... 122
      7.6.1 Method and Related Work ...................................................................... 122
      7.6.2 Supervised Learning: Rule-based Classification ....................................... 124
   7.7 Noise Simulation: Modelling the Effects of Mis-Communication ................ 125
      7.7.1 Method and Related Work ...................................................................... 125
      7.7.2 Simulating the Effects of Non- and Mis-Understandings ....................... 127
   7.8 User Simulation .............................................................................................. 128
      7.8.1 Method and Related Work ...................................................................... 129
      7.8.2 User Actions ........................................................................................... 132
      7.8.3 A Simple Bi-gram Model ....................................................................... 133
      7.8.4 Cluster-based User Simulation ............................................................... 134
### Part III Evaluation and Application

#### 8 Comparing Reinforcement and Supervised Learning of Dialogue Policies with Real Users

8.1 Policy Integration into a Dialogue System ............................................. 168
  8.1.1 The DUDE Rapid Dialogue Development Tools .............................. 168
  8.1.2 Extensions to DUDE ................................................................. 170
8.2 Experimental Setup .............................................................................. 174
  8.2.1 Technical Setup ........................................................................... 174
  8.2.2 Primary Driving Task .................................................................... 174
  8.2.3 Subjects and Procedure .................................................................. 175
  8.2.4 Task Types .................................................................................. 176
  8.2.5 User Questionnaires ...................................................................... 176
8.3 Results .................................................................................................. 177
  8.3.1 Subjective User Ratings ............................................................... 178
  8.3.2 Objective Dialogue Performance .................................................. 181
8.4 Discussion of Real User Evaluation Results ......................................... 182
8.5 Meta-Evaluation .................................................................................... 183
  8.5.1 Transfer Between Simulated and Real Environments ....................... 183
  8.5.2 Evaluation of the Learned Reward Function ................................. 184
8.6 Summary ............................................................................................... 188

#### 9 Adaptive Natural Language Generation

9.1 Introduction ............................................................................................ 190
  9.1.1 Previous Work on Information Presentation in SDS ....................... 190
9.2 NLG as Planning Under Uncertainty ...................................................... 192
9.3 Wizard-of-Oz Data Collection ............................................................... 192
  9.3.1 Experimental Setup and Data Collection ...................................... 193
## Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASR</td>
<td>Automatic Speech Recognition</td>
</tr>
<tr>
<td>DA</td>
<td>Dialogue Act</td>
</tr>
<tr>
<td>DB</td>
<td>Database</td>
</tr>
<tr>
<td>DM</td>
<td>Dialogue Management</td>
</tr>
<tr>
<td>GUI</td>
<td>Graphical User Interface</td>
</tr>
<tr>
<td>HCI</td>
<td>Human Computer Interaction</td>
</tr>
<tr>
<td>IP</td>
<td>Information Presentation</td>
</tr>
<tr>
<td>ISU</td>
<td>Information State Update</td>
</tr>
<tr>
<td>MDP</td>
<td>Markov Decision Process</td>
</tr>
<tr>
<td>ML</td>
<td>Machine Learning</td>
</tr>
<tr>
<td>NLG</td>
<td>Natural Language Generation</td>
</tr>
<tr>
<td>NLP</td>
<td>Natural Language Processing</td>
</tr>
<tr>
<td>PARADISE</td>
<td>PARAdigm for DIalogue System Evaluation</td>
</tr>
<tr>
<td>POMDP</td>
<td>Partially Observable Markov Decision Process</td>
</tr>
<tr>
<td>RL</td>
<td>Reinforcement Learning</td>
</tr>
<tr>
<td>SA</td>
<td>Speech Act</td>
</tr>
<tr>
<td>SASSI</td>
<td>Subjective Assessment of Speech System Interfaces</td>
</tr>
<tr>
<td>SDS</td>
<td>Spoken Dialogue System</td>
</tr>
<tr>
<td>SL</td>
<td>Supervised Learning</td>
</tr>
<tr>
<td>SLU</td>
<td>Spoken Language Understanding</td>
</tr>
<tr>
<td>TTS</td>
<td>Text-to-Speech</td>
</tr>
<tr>
<td>VOIP</td>
<td>Voice-Over-Internet Protocol</td>
</tr>
<tr>
<td>WER</td>
<td>Word-Error Rate</td>
</tr>
<tr>
<td>WOZ</td>
<td>Wizard-of-Oz</td>
</tr>
</tbody>
</table>