

Contents

Foreword by T. V. Raman	v
Preface	vii
Introduction	1
Who Should Read This Book?	2
How This Book Is Organised	2
Part I. Background to Spoken Dialogue Technology	5
1. Talking with Computers: Fact or Fiction?	7
A Little Fantasy	7
Further Reading	17
Exercises	18
2. Spoken Dialogue Applications: Research Directions and Commercial Deployment	19
Spoken Dialogue Applications	19
Research in Spoken Dialogue Technology	24
The Commercial Deployment of Spoken Dialogue Technology ...	37
Summary	41
Further Reading	42
Exercises	42
3. Understanding Dialogue	45
Dialogue: A Definition	45
Key Characteristics of Dialogue	46
Modelling Dialogue: Structure and Process	59
Summary	75
Further Reading	76
Exercises	77
4. Components of a Spoken Dialogue System – Speech Input and Output	79
Overview of a Spoken Dialogue System	79
Speech Recognition	81
Language Understanding	91

Language Generation	99
Text-to-Speech Synthesis	102
Summary	104
Further Reading	105
Exercises	105
5. Dialogue Management	107
Dialogue Initiative	108
Dialogue Control	111
Grounding	117
Accessing Information from an External Knowledge Source	121
Knowledge Sources for Dialogue Management	123
Issues for Developers	124
Summary	126
Further Reading	126
Exercise	126
Part II. Developing Spoken Dialogue Applications	127
6. Dialogue Engineering: The Dialogue Systems	
Development Lifecycle	129
Requirements Analysis	129
Requirements Specification	135
Design	139
Implementation	146
Testing	150
Evaluation	154
Summary	160
Further Reading	160
Exercises	161
7. Developing a Spoken Dialogue System Using the CSLU Toolkit ...	163
Tutorial 1: Basic Functions in RAD: The Pizza Application	164
Tutorial 2: Developing Basic Functions for a Student Information System	171
Tutorial 3: The Default Repair Subdialogue	174
Tutorial 4: Subdialogues	176
Tutorial 5: Digit Recognition	179
Tutorial 6: DTMF Input	181
Tutorial 7: Alpha-digit Recognition	184
Tutorial 8: Recognition Grammars	187
Tutorial 9: Speech Output in RAD	191
Tutorial 10: Using Tcl in RAD	192
Tutorial 11: Adding New Pronunciations and Creating a Dynamic Recogniser	194
Tutorial 12: Linking a RAD Application to a Database Using TclODBC	197
Summary	201

Further Reading	201
Ideas for Further Projects	202
8. Developing Multimodal Dialogue Systems Using the CSLU Toolkit	205
Tutorial 1: Setting Up the Animated Character	206
Tutorial 2: Login Dialogue	208
Tutorial 3: Exploring Emotions	211
Tutorial 4: Identifying Two-dimensional Shapes	217
Tutorial 5: Improving the Interface	224
Some Additional Features	227
Current Developments	229
Further Reading	229
Ideas for Further Projects	230
9. Developing a Directed Dialogue System Using VoiceXML	231
VoiceXML: A Language for Voice Enabling the World Wide Web	231
VoiceXML: The Language	233
Tutorials	233
Tutorial 1: Prompts, Responses and Verification	234
Tutorial 2: Repair in VoiceXML	241
Tutorial 3: Subdialogues	243
Tutorial 4: Digit Recognition	247
Tutorial 5: DTMF	249
Tutorial 6: Alphadigit Recognition	253
Tutorial 7: Modifying Speech Output	254
Tutorial 8: Navigation and Global Commands	257
Notes on ECMAScript	260
Further Reading	261
10. More Advanced VoiceXML	263
Tutorial 1: Mixed-initiative Dialogue	263
Tutorial 2: The Form Interpretation Algorithm (FIA)	266
Tutorial 3: Recognition Grammars	270
Tutorial 4: Variable and Grammar Scope	274
Tutorial 5: Dynamic VoiceXML and Web Server Applications	278
Tutorial 6: Dynamic Grammars	284
Further Reading	288
Ideas for Further Projects	288
11. Multimodal Web-based Dialogue: XHTML + Voice and SALT	291
XHTML + Voice	291
SALT	304
Summary	318
Further Reading	319
Ideas for Further Projects	320

Part III. Advanced Applications	321
12. Advanced Dialogue Systems: Some Cases Studies	323
DARPA Communicator	323
TRAINS and TRIPS	332
Conversational Architectures Project	341
Summary	347
Further Reading	348
13. Research Topics in Spoken Dialogue Technology	349
Information State Theory	349
Error Handling in Spoken Dialogue Systems	353
Adaptive Dialogue Systems	362
Learning Optimal Dialogue Strategies	371
Summary	375
Further Reading	376
14. Future Directions: The Way Ahead	379
Advanced Multimodal Systems	379
Future Directions	385
Concluding Remarks	388
Appendixes	389
Appendix 1. Grammar Formats	391
The XML and ABNF Grammar Formats	391
Appendix 2. The CSLU Toolkit and Related Software for Chapters 7 and 8	395
CSLU Toolkit	395
TclODBC	395
Setting Up a Data Source (ODBC)	395
Appendix 3. Software for Chapters 9 and 10	397
Downloading the IBM WebSphere VoiceServer SDK	397
Downloading and Installing Apache Tomcat	399
Java Server Pages (JSP)	399
Appendix 4. Software for Chapter 11. IBM Multimodal Tools and Microsoft.NET Speech SDK	400
IBM Multimodal Tools	400
Microsoft .NET Speech SDK	400
Appendix 5. Spoken Dialogue Technology: Projects and Links	401
Appendix 6. Spoken Dialogue Technology: Toolkits and Resources ...	406
Toolkits and Resources for Spoken Dialogue Technology	406
Web Sites for Hosted VoiceXML Applications	407
References	409
Index	425