

Contents

1	Introducing the Conversational Interface	1
1.1	Introduction	1
1.2	Who Should Read the Book?	2
1.3	A Road Map for the Book	2
1.3.1	Part I: Conversational Interfaces: Preliminaries	4
1.3.2	Part II: Developing a Speech-Based Conversational Interface	4
1.3.3	Part III: Conversational Interfaces and Devices	5
1.3.4	Part IV: Evaluation and Future Prospects	6
Part I Conversational Interfaces: Preliminaries		
2	The Dawn of the Conversational Interface	11
2.1	Introduction	11
2.2	Interacting with a Conversational Interface	12
2.3	Conversational Interfaces for Smart Watches and Other Devices	15
2.4	Explaining the Rise of the Conversational Interface	15
2.4.1	Technological Developments	16
2.4.2	User Acceptance and Adoption	18
2.4.3	Enterprise and Specialized VPAs	19
2.4.4	The Cycle of Increasing Returns	20
2.5	The Technologies that Make up a Conversational Interface	20
2.6	Summary	22
	References	24
3	Toward a Technology of Conversation	25
3.1	Introduction	25
3.2	Conversation as Action	26
3.3	The Structure of Conversation	31
3.3.1	Dealing with Longer Sequences	34

3.4	Conversation as a Joint Activity.	35
3.4.1	Turn Taking in Conversation	36
3.4.2	Grounding.	37
3.4.3	Conversational Repair.	40
3.5	The Language of Conversation	42
3.5.1	Prosodic, Paralinguistic, and Nonverbal Behaviors . . .	43
3.5.2	Implications for the Conversational Interface	44
3.6	Summary	45
	References	47
4	Conversational Interfaces: Past and Present	51
4.1	Introduction.	51
4.2	Conversational Interfaces: A Brief History	52
4.2.1	A Typical Interaction with a Spoken Dialog System	52
4.2.2	An Interaction that Goes Wrong.	54
4.2.3	Spoken Dialog Systems.	55
4.2.4	Voice User Interfaces	56
4.2.5	Embodied Conversational Agent, Companions, and Social Robots	56
4.2.6	Chatbots	57
4.3	What Have We Learned so Far?	58
4.3.1	Making Systems More Intelligent.	58
4.3.2	Using Incremental Processing to Model Conversational Phenomena	60
4.3.3	Languages and Toolkits for Developers.	61
4.3.4	Large-Scale Experiments on System Design Using Techniques from Machine Learning.	62
4.4	Summary	65
	References	68
 Part II Developing a Speech-Based Conversational Interface		
5	Speech Input and Output.	75
5.1	Introduction.	75
5.2	Speech Recognition	75
5.2.1	ASR as a Probabilistic Process.	77
5.2.2	Acoustic Model	78
5.2.3	Language Model	81
5.2.4	Decoding	82
5.3	Text-to-Speech Synthesis	83
5.3.1	Text Analysis	83
5.3.2	Waveform Synthesis.	86
5.3.3	Using Prerecorded Speech.	87
5.3.4	Speech Synthesis Markup Language	87

- 5.4 Summary 89
- References 91
- 6 Implementing Speech Input and Output 93**
 - 6.1 Introduction. 93
 - 6.2 Web Speech API 95
 - 6.2.1 Text-to-Speech Synthesis. 95
 - 6.2.2 Speech Recognition 98
 - 6.3 The Android Speech APIs. 100
 - 6.3.1 Text-to-Speech Synthesis. 103
 - 6.3.2 Speech Recognition 110
 - 6.3.3 Using Speech for Input and Output. 117
 - 6.4 Summary 121
- 7 Creating a Conversational Interface Using Chatbot Technology 125**
 - 7.1 Introduction. 125
 - 7.2 Introducing the Pandorabots Platform 127
 - 7.3 Developing Your Own Bot Using AIML 131
 - 7.3.1 Creating Categories 131
 - 7.3.2 Wildcards 133
 - 7.3.3 Variables. 134
 - 7.3.4 Sets and Maps 135
 - 7.3.5 Context. 136
 - 7.4 Creating a Link to Pandorabots from Your Android App 139
 - 7.4.1 Creating a Bot in the Developer Portal 139
 - 7.4.2 Linking an Android App to a Bot. 143
 - 7.5 Introducing Mobile Functions 147
 - 7.5.1 Processing the <ooob> Tags. 147
 - 7.5.2 Battery Level. 149
 - 7.5.3 Search Queries. 149
 - 7.5.4 Location and Direction Queries 151
 - 7.6 Extending the App. 153
 - 7.7 Alternatives to AIML 154
 - 7.8 Some Ways in Which AIML Can Be Further Developed 154
 - 7.8.1 Learning a Chatbot Specification from Data. 154
 - 7.8.2 Making Use of Techniques from Natural Language Processing. 156
 - 7.9 Summary 157
 - References 157
- 8 Spoken Language Understanding 161**
 - 8.1 Introduction. 161
 - 8.2 Technologies for Spoken Language Understanding 163
 - 8.3 Dialog Act Recognition 164
 - 8.4 Identifying Intent 165

- 8.5 Analyzing the Content of the User’s Utterances 166
 - 8.5.1 Tokenization 166
 - 8.5.2 Bag of Words 166
 - 8.5.3 Latent Semantic Analysis 167
 - 8.5.4 Regular Expressions 168
 - 8.5.5 Part-of-Speech Tagging 168
 - 8.5.6 Information Extraction 169
 - 8.5.7 Semantic Role Labeling 169
- 8.6 Obtaining a Complete Semantic Interpretation of the Input 170
 - 8.6.1 Semantic Grammar 170
 - 8.6.2 Syntax-Driven Semantic Analysis 172
- 8.7 Statistical Approaches to Spoken Language Understanding 175
 - 8.7.1 Generative Models 176
 - 8.7.2 Discriminative Models 178
 - 8.7.3 Deep Learning for Natural and Spoken Language Understanding 179
- 8.8 Summary 180
- References 182
- 9 Implementing Spoken Language Understanding 187**
 - 9.1 Introduction. 187
 - 9.2 Getting Started with the Api.ai Platform 188
 - 9.2.1 Exercise 9.1 Creating an Agent in Api.ai 188
 - 9.2.2 Exercise 9.2 Testing the Agent. 189
 - 9.3 Creating an Android App for an Agent. 191
 - 9.3.1 Exercise 9.3 Producing a Semantic Parse. 191
 - 9.3.2 Exercise 9.4 Testing the App 194
 - 9.4 Specifying Your Own Entities and Intents. 195
 - 9.4.1 Exercise 9.5 Creating Entities 195
 - 9.4.2 Exercise 9.6 Creating an Intent 197
 - 9.4.3 Exercise 9.7 Testing the Custom Entities and Intents 199
 - 9.5 Using Aliases 199
 - 9.6 Using Context 199
 - 9.6.1 Exercise 9.8 Defining Contexts 200
 - 9.7 Creating a Slot Filling Dialog 200
 - 9.7.1 Exercise 9.9 Creating a Slot-Filling Dialog 201
 - 9.7.2 Exercise 9.10 Additional Exercises 202
 - 9.8 Overview of Some Other Spoken Language Understanding Tools 202
 - 9.8.1 Tools Using Intents and Entities. 203
 - 9.8.2 Toolkits for various other NLP Tasks 206
 - 9.9 Summary 207
 - References 208

- 10 Dialog Management** 209
 - 10.1 Introduction. 209
 - 10.2 Defining the Dialog Management Task 210
 - 10.2.1 Interaction Strategies. 211
 - 10.2.2 Error Handling and Confirmation Strategies. 214
 - 10.3 Handcrafted Approaches to Dialog Management 215
 - 10.4 Statistical Approaches to Dialog Management 217
 - 10.4.1 Reinforcement Learning 219
 - 10.4.2 Corpus-Based Approaches. 225
 - 10.5 Summary 228
 - References 229
- 11 Implementing Dialog Management** 235
 - 11.1 Introduction. 235
 - 11.2 Development of a Conversational Interface
 - Using a Rule-Based Dialog Management Technique. 236
 - 11.2.1 Practical Exercises Using VoiceXML 242
 - 11.3 Development of a Conversational Interface
 - Using a Statistical Dialog Management Technique 257
 - 11.4 Summary 262
 - References 263
- 12 Response Generation** 265
 - 12.1 Introduction. 265
 - 12.2 Using Canned Text and Templates. 265
 - 12.3 Using Natural Language Generation Technology 268
 - 12.3.1 Document Planning 269
 - 12.3.2 Microplanning 270
 - 12.3.3 Realization 271
 - 12.4 Statistical Approaches to Natural Language Generation. 272
 - 12.5 Response Generation for Conversational Queries 273
 - 12.5.1 Question Answering 274
 - 12.5.2 Structured Resources to Support Conversational
 - Question Answering 275
 - 12.5.3 Text Summarization 276
 - 12.6 Summary 276
 - References 278

Part III Conversational Interfaces and Devices

- 13 Conversational Interfaces: Devices, Wearables, Virtual Agents, and Robots** 283
 - 13.1 Introduction. 283
 - 13.2 Wearables 284
 - 13.2.1 Smartwatches and Wristbands 284
 - 13.2.2 Armbands and Gloves. 285

- 13.2.3 Smart Glasses 286
- 13.2.4 Smart Jewelry 287
- 13.2.5 Smart Clothing. 288
- 13.3 Multimodal Conversational Interfaces for Smart Devices
and Wearables 289
- 13.4 Virtual Agents 293
- 13.5 Multimodal Conversations with Virtual Agents 294
- 13.6 Examples of Tools for Creating Virtual Agents 295
- 13.7 Social Robots 296
- 13.8 Conversational Interfaces for Robots 297
- 13.9 Examples of Social Robots and Tools for Creating Robots 298
 - 13.9.1 Aldebaran Robots. 298
 - 13.9.2 Jibo 299
 - 13.9.3 FurHat 300
 - 13.9.4 Aisoy 300
 - 13.9.5 Amazon Echo 302
 - 13.9.6 Hello Robo 302
 - 13.9.7 The Open Robot Hardware Initiative. 302
 - 13.9.8 iCub.Org: Open-Source Cognitive Humanoid
Robotic Platform 302
 - 13.9.9 SPEAKY for Robots. 303
 - 13.9.10 The Robot Operating System (ROS). 303
- 13.10 Summary 304
- References 305
- 14 Emotion, Affect, and Personality. 309**
 - 14.1 Introduction. 309
 - 14.2 Computational Models of Emotion. 310
 - 14.2.1 The Dimensional Approach 310
 - 14.2.2 The Discrete Approach 312
 - 14.2.3 The Appraisal Approach 312
 - 14.3 Models of Personality. 314
 - 14.3.1 The Detection of Personality 315
 - 14.3.2 Simulating Personality 316
 - 14.4 Making Use of Affective Behaviors in the Conversational
Interface 319
 - 14.4.1 Acknowledging Awareness and Mirroring
Emotion 319
 - 14.4.2 Dealing with and Provoking the User’s Emotions. 320
 - 14.4.3 Building Empathy 321
 - 14.4.4 Fostering the User’s Engagement 322
 - 14.4.5 Emotion as Feedback on the System’s
Performance 322
 - 14.5 Summary 323
 - References 325

15 Affective Conversational Interfaces 329

15.1 Introduction. 329

15.2 Representing Emotion with EmotionML 330

15.3 Emotion Recognition 333

15.3.1 Emotion Recognition from Physiological Signals 335

15.3.2 Emotion Recognition from Speech 339

15.3.3 Emotion Recognition from Facial Expressions
and Gestures 344

15.4 Emotion Synthesis 349

15.4.1 Expressive Speech Synthesis 349

15.4.2 Generating Facial Expressions, Body Posture,
and Gestures 350

15.4.3 The Uncanny Valley. 352

15.5 Summary 353

References 354

**16 Implementing Multimodal Conversational Interfaces
Using Android Wear 359**

16.1 Introduction. 359

16.2 Visual Interfaces for Android Wear 361

16.3 Voice Interfaces for Android Wear. 363

16.3.1 System-Provided Voice Actions 364

16.3.2 Developer-Defined Voice Actions. 366

16.4 Summary 375

Part IV Evaluation and Future Directions

17 Evaluating the Conversational Interface 379

17.1 Introduction. 379

17.2 Objective Evaluation 380

17.2.1 Overall System Evaluation. 381

17.2.2 Component Evaluation 382

17.2.3 Metrics Used in Industry. 387

17.3 Subjective Evaluation. 388

17.3.1 Predicting User Satisfaction. 389

17.4 Evaluation Procedures 394

17.4.1 Evaluation Settings: Laboratory Versus Field. 394

17.4.2 Wizard of Oz. 395

17.4.3 Test Subjects. 396

17.5 Summary 398

References 399

- 18 Future Directions** 403
 - 18.1 Introduction. 403
 - 18.2 Advances in Technology. 403
 - 18.2.1 Cognitive Computing 404
 - 18.2.2 Deep Learning 404
 - 18.2.3 The Internet of Things 406
 - 18.2.4 Platforms, SDKs, and APIs for Developers 406
 - 18.3 Applications that Use Conversational Interfaces 408
 - 18.3.1 Enterprise Assistants 408
 - 18.3.2 Ambient Intelligence and Smart Environments 409
 - 18.3.3 Health care 410
 - 18.3.4 Companions for the Elderly 411
 - 18.3.5 Conversational Toys and Educational Assistants 412
 - 18.3.6 Bridging the Digital Divide for Under-Resourced Languages 414
 - 18.4 Summary 415
 - References 417
- Index** 419



<http://www.springer.com/978-3-319-32965-9>

The Conversational Interface

Talking to Smart Devices

McTear, M.; Callejas, Z.; Griol, D.

2016, XXII, 422 p. 76 illus., 29 illus. in color., Hardcover

ISBN: 978-3-319-32965-9