

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
7.5	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

7.8.5	Smoothed Bi-gram User Simulation	136
7.8.6	Evaluation of User Simulations	138
7.8.7	Speech Act Realisation Dependent on the User Goal	139
7.9	Reward and Objective Functions	142
7.9.1	Method and Related Work	142
7.9.2	Linear Regression for Information Acquisition	146
7.9.3	Non-linear Rewards for Information Presentation	148
7.9.4	Final Reward	150
7.10	State-Space Discretisation	151
7.11	Learning Experiments	152
7.11.1	Training with SHARSHA	152
7.11.2	Results for Testing in Simulation	154
7.11.3	Qualitative Strategy Description	155
7.11.4	Strategy Implementation	157
7.11.5	Discussion and Error Analysis	158
7.12	Summary	162

Part III Evaluation and Application

8	Comparing Reinforcement and Supervised Learning of Dialogue Policies with Real Users	167
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	189
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

9.3.2	Surface Realiser	193
9.3.3	Human “Wizard” Baseline Strategy	194
9.4	The Simulation / Learning Environment	195
9.4.1	User Simulations	195
9.4.2	Database Matches and “Focus of Attention”	197
9.4.3	Data-driven Reward Function	197
9.5	Reinforcement Learning Experiments	198
9.5.1	Experimental Set-up	199
9.5.2	Results	199
9.6	Evaluation with real users	202
9.7	Conclusion	203
10	Conclusion	205
10.1	Contributions	206
10.2	Discussion	207
10.2.1	Lessons Learned	208
10.2.2	RL for Commercial Dialogue Strategy Development	209
10.3	Outlook: challenges for future statistical dialogue systems	210
	Example Dialogues	213
A.1	Wizard-of-Oz Example Dialogues	213
A.2	Example Dialogues from Simulated Interaction	216
A.3	Example Dialogues from User Testing	218
	Learned State-Action Mappings	223
	References	229
	About the Authors	253