

Table of contents

Preface IX

CHAPTER 1

Preliminaries 1

1. Introduction 1
2. Basic notions for summarization 6
3. Abstract architecture for summarization 14
4. Summarization approaches 18
5. Current applications 19
6. Conclusion 21
7. Review 22

CHAPTER 2

Professional summarizing 27

1. Introduction 27
2. The stages of abstracting 29
3. Abstracting strategies 32
4. Reading for abstracting 35
5. Revision 37
6. Psychological experiments 39
7. Structure of empirical abstracts 40
8. Conclusion 41
9. Review 43

CHAPTER 3

Extraction 45

1. Introduction 45
2. The Edmundsonian paradigm 47

3. Corpus based sentence extraction 53
 - 3.1 General considerations 53
 - 3.2 Aspects of learning approaches 59
4. Coherence of extracts 70
5. Conclusion 71
6. Review 74

CHAPTER 4

Revision 77

1. Introduction 77
2. Shallow coherence smoothing 78
3. Full revision to improve informativeness 81
 - 3.1 Case study: Full revision 81
 - 3.2 Related work 84
 - 3.3 Implications 85
4. Text compaction 85
5. Conclusion 89
6. Review 90

CHAPTER 5

Discourse-level information 91

1. Introduction 91
2. Text cohesion 93
 - 2.1 Introduction 93
 - 2.2 Cohesion graph topology 94
 - 2.3 Topic characterization 102
3. Text coherence 106
 - 3.1 Introduction 106
 - 3.2 Coherence relations 108
 - 3.3 Rhetorical Structure Theory (RST) 110
 - 3.4 Rhetorical structure and cue phrases 113
 - 3.5 The document scheme, revisited 122
4. Conclusion 124
5. Review 126

CHAPTER 6

Abstraction 129

1. Introduction 129
2. Abstraction from templates 130
 - 2.1 Introduction 130
 - 2.2 Case study: Sketchy scripts 130
 - 2.3 Modern information extraction 134
3. Abstraction by term rewriting 139
4. Abstraction using event relations 144
5. Abstraction using a concept hierarchy 148
 - 5.1 Domain knowledge base activation 148
 - 5.2 Generic thesaurus activation 150
6. Synthesis for abstraction 152
 - 6.1 Pretty printing 152
 - 6.2 Graphical output 152
 - 6.3 Extraction 153
 - 6.4 Generation for synthesis 154
7. Conclusion 163
8. Review 165

CHAPTER 7

Multi-document summarization 169

1. Introduction 169
2. Types of relationships across documents 173
3. MDS methods 179
 - 3.1 Overview 179
 - 3.2 Specific approaches 181
4. Case study: Biographical summarization 194
 - 4.1 Introduction 194
 - 4.2 Example architecture 196
 - 4.3 Algorithm steps 197
 - 4.4 Bio summarizer components 201
 - 4.5 Assessment 202
5. Conclusion 202
6. Review 206

CHAPTER 8

Multimedia summarization 209

1. Introduction 209
2. Dialog summarization 209
3. Summarization of video 211
4. Summarization of diagrams 214
5. Automatic multimedia briefing generation 215
6. Conclusion 220
7. Review 220

CHAPTER 9

Evaluation 221

1. Introduction 221
2. Intrinsic methods 224
 - 2.1 Assessing agreement between subjects 225
 - 2.2 Quality 226
 - 2.3 Informativeness 227
 - 2.4 Component-level tests 238
3. Extrinsic methods 241
 - 3.1 Relevance assessment 241
 - 3.2 Reading comprehension 248
 - 3.3 Presentation strategies 251
 - 3.4 Mature system evaluation 252
4. Conclusion 253
5. Review 256

CHAPTER 10

Postscript 261

References 263

Index 279