Contents

Introduction ix

Chapter 1. The Sphere of Lexicons and Knowledge 1

1.1. Lexical semantics 1
1.1.1. Extension of lexical meaning 1
1.1.2. Paradigmatic relations of meaning 6
1.1.3. Theories of lexical meaning 16
1.2. Lexical databases 23
1.2.1. Standards for encoding and exchanging data 25
1.2.2. Standard character encoding 25
1.2.3. Content standards 32
1.2.4. Writing systems 40
1.2.5. A few lexical databases 45
1.3. Knowledge representation and ontologies 49
1.3.1. Knowledge representation 49
1.3.2. Ontologies 63

Chapter 2. The Sphere of Semantics 75

2.1. Combinatorial semantics 75
2.1.1. Interpretive semantics 75
2.1.2. Generative semantics 80
2.1.3. Case grammar 82
2.1.4. Rastier's interpretive semantics 84
2.1.5. Meaning–text theory 92

2.2. Formal semantics 95
2.2.1. Propositional logic 95
2.2.2. First-order logic 106
2.2.3. Lambda calculus 113
2.2.4. Other types of logic 121

Chapter 3. The Sphere of Discourse and Text 123

3.1. Discourse analysis and pragmatics 123
3.1.1. Fundamental concepts 123
3.1.2. Utterance production 125
3.1.3. Context, cotext and intertextuality 128
3.1.4. Information structure in discourse 130
3.1.5. Coherence 137
3.1.6. Cohesion 138
3.1.7. Ellipses 142
3.1.8. Textual sequences 143
3.1.9. Speech acts 144
3.2. Computational approaches to discourse 146
3.2.1. Linear segmentation of discourse 146
3.2.2. Rhetorical structure theory and automatic discourse analysis 148
3.2.3. Discourse interpretation: DRT 154
3.2.4. Processing anaphora 159

Chapter 4. The Sphere of Applications 169

4.1. Software engineering for NLP software 169
4.1.1. Lifecycle of an NLP software 169
4.1.2. Software architecture for NLP 170
4.1.3. Serial architectures 171
4.1.4. Data-centered architectures 173
4.1.5. Object-oriented architectures 177
4.1.6. Multi-agent architectures 178
4.1.7. Syntactic–semantic cooperation: from cognitive models to software architecture 180
4.1.8. Programming languages for NLP 184
4.1.9. Evaluation of NLP systems 186
4.2. Machine translation (MT) 191
4.2.1. Why is translation difficult? 192
4.2.2. History of MT systems 194
4.2.3. Typology of MT systems 196

4.2.4. The use of MT 198
4.2.5. MT techniques 199
4.2.6. Example of a translation system: Verbmobil 208
4.3. Information retrieval (IR) 211
4.3.1. IR and related domains 211
4.3.2. Lexical information and IR 213
4.3.3. Information retrieval approaches 219
4.4. Big Data (BD) and information extraction 234
4.4.1. Structured, semi-structured and unstructured data 234
4.4.2. Architectures of BD processing systems 235
4.4.3. Role of NLP in BD processing 237
4.4.4. Information extraction 238

Conclusion 259

Bibliography 263

Index 301