ONTOLOGY LEARNING FOR THE SEMANTIC WEB

Alexander Maedche

Kluwer Academic Publishers

2_005_498_1

2-005-498-1

ONTOLOGY LEARNING FOR THE SEMANTIC WEB

Alexander Maedche
University of Karlsruhe, Germany

KLUWER ACADEMIC PUBLISHERS
Boston / Dordrecht / London

Contents

List of Figures ix
List of Tables xiii
Preface xv
Acknowledgements xviii
Foreword by R. Studer xix

Part I Fundamentals

1. INTRODUCTION 3
   1 Motivation & Problem Description 3
   2 Research Questions 4
   3 Reader's Guide 6
2. ONTOLOGY — DEFINITION & OVERVIEW 11
   1 Ontologies for Communication – A Layered Approach 15
   2 Development & Application of Ontologies 21
   3 Conclusion 25
3. LAYERED ONTOLOGY ENGINEERING 29
   1 Ontology Engineering Framework 30
   2 Layered Representation 34
   3 Conclusion 49
      3.1 Further Topics in Ontology Engineering 50
      3.2 Ontology Learning for Ontology Engineering 51

Part II Ontology Learning for the Semantic Web

4. ONTOLOGY LEARNING FRAMEWORK 59
1 A Taxonomy of Relevant Data for Ontology Learning 60
2 An Architecture for Ontology Learning 66
2.1 Overview of the Architecture Components 66
2.2 Ontology Engineering Workbench OntoEdit 68
2.3 Data Import & Processing Component 70
2.4 Algorithm Library 71
2.5 Graphical User Interface & Management Component 72
3 Phases of Ontology Learning 73
3.1 Import & Reuse 74
3.2 Extract 75
3.3 Prune 76
3.4 Refine 77
4 Conclusion 78

5. DATA IMPORT & PROCESSING 81
1 Importing & Processing Existing Ontologies 83
1.1 Ontology Wrapper & Import 84
1.2 FCA-Merge — Bottom-Up Ontology Merging 85
2 Collecting, Importing & Processing Documents 95
2.1 Ontology-focused Document Crawling 95
2.2 Shallow Text Processing using SMES 97
2.3 Semi-Structured Document Wrapper 105
2.4 Transforming Data into Relational Structures 107
3 Conclusion 112
3.1 Language Processing for Ontology Learning 112
3.2 Ontology Learning from Web Documents 113
3.3 (Multi-)Relational Data 114

6. ONTOLOGY LEARNING ALGORITHMS 117
1 Algorithms for Ontology Extraction 118
1.1 Lexical Entry Extraction 118
1.2 Taxonomy Extraction 122
1.3 Non-Taxonomic Relation Extraction 130
2 Algorithms for Ontology Maintenance 140
2.1 Ontology Pruning 140
2.2 Ontology Refinement 142

3 Conclusion 144
3.1 Multi-Strategy Learning 145
3.2 Taxonomic vs. Non-Taxonomic Relations 145
3.3 A Note on Learning Axioms — $\mathcal{A}^{\mathcal{O}}$ 146

Part III Implementation & Evaluation

7. THE TEXT-TO-ONTO ENVIRONMENT 151
1 Component-based Architecture 153
2 Ontology Engineering Environment OntoEdit 154
3 Components for Ontology Learning 163
4 Conclusion 168

8. EVALUATION 171
1 The Evaluation Approach 172
2 Ontology Comparison Measures 173
2.1 Precision and Recall 174
2.2 Lexical Comparison Level Measures 175
2.3 Conceptual Comparison Level Measures 177
3 Human Performance Evaluation 183
3.1 Ontology Engineering Evaluation Study 184
3.2 Human Evaluation – Precision and Recall 185
3.3 Human Evaluation – Lexical Comparison Level 187
3.4 Human Evaluation – Conceptual Comparison Level 188
4 Ontology Learning Performance Evaluation 190
4.1 The Evaluation Setting 191
4.2 Evaluation of Lexical Entry Extraction 191
4.3 Evaluation of Concept Hierarchy Extraction 193
4.4 Evaluation of Non-Taxonomic Relation Extraction 194
5 Conclusion 196
5.1 Application-oriented Evaluation 197
5.2 Standard Datasets for Evaluation 198

Part IV Related Work & Outlook

9. RELATED WORK 203
- 1 Related Work on Ontology Engineering 204
- 2 Related Work on Frameworks of KA & ML 209
- 3 Related Work on Data Import & Processing 212
- 4 Related Work on Algorithms 214
- 5 Related Work on Evaluation 219

10. CONCLUSION & OUTLOOK 223
- 1 Contributions 223
- 2 Insights into Ontology Learning 224
- 3 Unanswered Questions 225
- 4 Future Research 226

References 228

Index 242