

Contents

1. Introduction	1
1.1. Linked Data on the WWW	1
1.2. Approaches to Query Linked Data	4
1.3. Problem Statement	6
1.4. Contributions	7
1.5. Thesis Outline	10
 I. Foundations of Queries over a Web of Linked Data	 13
2. Models for Linked Data Queries	15
2.1. Data Model	15
2.1.1. Structural Elements	15
2.1.2. Queries	22
2.1.3. Related Work	24
2.2. Computation Model	25
2.2.1. Existing Models of Computation on the WWW	26
2.2.2. The LD Machine	27
2.2.3. LD Machine Computability	29
2.2.4. LD Machine Decidability	30
 3. Full-Web Query Semantics	 33
3.1. Related Work	34
3.1.1. Web Query Languages	34
3.1.2. Navigational Query Languages for Linked Data	34
3.1.3. SPARQL as a Query Language for Linked Data	35
3.2. Definition	38
3.2.1. SPARQL	38
3.2.2. SPARQL _{LD}	42
3.3. Theoretical Properties	42
3.3.1. Satisfiability, (Un)bounded Satisfiability, and Monotonicity	43
3.3.2. LD Machine Decidability of Termination	47
3.3.3. LD Machine Computability	53
3.3.4. Finiteness of Expected Query Results	57
3.4. Summary	59

4. Reachability-Based Query Semantics	61
4.1. Definition	61
4.1.1. Reachability	62
4.1.2. SPARQL _{LD(R)}	63
4.2. Result Containment and Infiniteness	64
4.3. Reachability Criteria	69
4.3.1. Comparing Reachability Criteria	69
4.3.2. Combining Reachability Criteria	71
4.3.3. Reachability Criteria that Ensure Finiteness	73
4.3.4. Constant Reachability Criteria	74
4.4. Theoretical Properties	76
4.4.1. Satisfiability, (Un)bounded Satisfiability, and Monotonicity	77
4.4.2. LD Machine Decidability of Termination	86
4.4.3. LD Machine Computability	90
4.5. Differences between SPARQL _{LD} and SPARQL _{LD(R)}	95
 II. Execution of Queries over a Web of Linked Data	 97
5. Overview of Query Execution Techniques	99
5.1. Data Source Selection	100
5.1.1. Live Exploration Approaches	100
5.1.2. Index-Based Approaches	101
5.1.3. Hybrid Approaches	104
5.2. Data Source Ranking	104
5.3. Integration of Data Retrieval and Result Construction	106
5.3.1. Separated Execution Approaches	106
5.3.2. Integrated Execution Approaches	107
5.4. Traversal-Based Query Execution	107
5.5. Summary	108
 6. A Traversal-Based Strategy	 111
6.1. Conjunctive Linked Data Queries	111
6.2. Informal Description	114
6.3. Query Execution Model	115
6.3.1. Overview	116
6.3.2. Partial Solutions	116
6.3.3. Constructing (Partial) Solutions	117
6.3.4. Discovered Subwebs of the Queried Web	118
6.3.5. Traversing Data Links	119
6.3.6. Combining Construction and Traversal	121
6.3.7. Abstract Query Execution Procedure	125
6.4. Soundness and Completeness	126
6.5. Summary	128

7. An Iterator-Based Implementation	129
7.1. Definition	130
7.1.1. Iterators for Query Execution over RDF Data	130
7.1.2. Iterators for Traversal-Based Query Execution	134
7.2. Formal Analysis	137
7.2.1. Examples for Incompleteness	137
7.2.2. Alignment with the Execution Model	140
7.3. Experimental Analysis	151
7.3.1. Our Query Execution System	152
7.3.2. WWW-Based Experiment	153
7.3.3. Simulation Based Experiments	158
7.4. Summary	176
 III. Conclusions	 177
8. Conclusions	179
8.1. Main Results	179
8.2. Directions for Future Work	182
8.2.1. Schema Heterogeneity	182
8.2.2. Coreferences	183
8.2.3. Trustworthiness and Data Quality	183
8.2.4. Dynamic Environment	184
8.2.5. Query Expressiveness	184
 IV. Appendix	 187
A. Commonly Used Symbols	189
B. Encoding of Structures Related to Query Computation	193
B.1. Encoding Basic Elements	193
B.2. Encoding RDF Triples	194
B.3. Encoding Webs of Linked Data	194
B.4. Encoding Valuations	194
C. Basic Properties of SPARQL Queries	195
C.1. Satisfiability	195
C.2. Monotonicity	201
C.3. Bounded Satisfiability and Unbounded Satisfiability	204
D. Supplementary Information about the Experiments	211
D.1. Queries for the WWW-Based Experiment	211
D.2. Measurements of the WWW-Based Experiment	215
D.3. Queries for the Simulation-Based Experiments	216

E. Proofs of Auxiliary Results	219
E.1. Proof of Lemma 3.1	219
E.2. Proof of Lemma 3.2	219
E.3. Proof of Lemma 4.1	221
E.4. Proof of Lemma 4.2	222
E.5. Proof of Lemma 4.3	223
E.6. Proof of Lemma 4.4	224
E.7. Proof of Lemma 4.5	225
E.8. Proof of Lemma 4.6	226
E.9. Proof of Lemma 4.7	228
E.10. Proof of Lemma 6.1	229
E.11. Proof of Lemma 6.2	230
E.12. Proof of Lemma 6.3	231