

## Table of Contents

Chapter 1.	Introduction.....	1
1.1.	Research Motivation .....	1
1.2.	Research Problem .....	2
1.3.	Research Objective .....	3
1.4.	Research Flow.....	3
1.5.	Organization of Thesis .....	6
Chapter 2.	Literature Review.....	7
2.1.	XML Query Capability .....	7
2.2.	XML Benchmarks.....	10
2.2.1.	XMark .....	11
2.2.2.	XMach-1 .....	13
2.2.3.	XOO7.....	15
2.3.	XML Benchmarks Comparison .....	18
2.4.	Ontology .....	20
2.4.1.	Ontology and Information Integration .....	20
2.4.2.	Ontology and Reasoning.....	21
2.4.3.	Ontology and Benchmark .....	24
Chapter 3.	Research Method .....	26
3.1.	Research Structure .....	26
3.2.	Research Model .....	28
3.3.	XML Data Model.....	28
3.4.	XML Query Model .....	31
3.4.1.	Exact Match .....	33
3.4.2.	Joins .....	34
3.4.3.	Regular Path Expressions .....	36
3.4.4.	Document Construction .....	37
3.4.5.	Ordered Access .....	38
3.4.6.	Sorting.....	39
3.4.7.	Missing Elements.....	40
3.4.8.	Text Search.....	41
3.4.9.	Data-type Cast.....	42
3.4.10.	Function Application.....	42
3.5.	Ontology Data Model .....	44
3.6.	Ontology Query Model.....	46
3.7.	Test Database Generation.....	49
3.8.	Control Model.....	53

3.9.	Performance Metrics .....	54
Chapter 4.	Prototype Implementation.....	57
4.1.	Prototype Platform and Architecture .....	57
4.2.	Experimental Design.....	59
4.3.	Prototype Functions .....	64
4.3.1.	Data Loader.....	65
4.3.2.	Query Generator.....	67
4.3.3.	Scheduler.....	71
4.3.4.	Result Collector .....	72
Chapter 5.	Research Discussions and Limitations.....	74
5.1.	Research Implications .....	74
5.2.	Research Limitations .....	76
Chapter 6.	Conclusions and Future Research Directions .....	78
6.1.	Summary .....	78
6.2.	Future Research Directions.....	79
References	.....	81
Appendix A:	Global Schema .....	85
Appendix B:	Relations in Microsoft SQL Server .....	88
Appendix C:	Schemas in Tamino XML Server .....	90

## List of Tables

Table 2.1: Desired Functionalities of XML Query Languages .....	8
Table 2.2: Queries Specified in XMark Benchmark .....	12
Table 2.3: Queries Specified in XMach-1 Benchmark .....	14
Table 2.4: Queries Specified in XOO7 Benchmark .....	16
Table 2.5: Comparison of Benchmarks over Workload Characteristics .....	18
Table 2.6: Comparison of Benchmarks over Query Functionalities .....	19
Table 2.7: Query Services .....	22
Table 2.8: Reasoning Tasks .....	23
Table 3.1: List of Complexity Factors .....	43
Table 3.2: Scaling the Test Database .....	51
Table 3.3: Retrieved and Relevant Data .....	56
Table 4.1: The Description of Prototype System Platform Specification .....	57
Table 4.2: Query Response Time at Different Logical Database Sizes .....	59
Table 5.1: This Workload Model Category .....	76

## List of Figures

Figure 1.1: Research Flow .....	5
Figure 2.1: DL System Architecture (Horrocks, 2002).....	25
Figure 3.1: Research Structure.....	27
Figure 3.2: Graphic Representation of XML Data Model .....	31
Figure 3.3: XML Query Model.....	33
Figure 3.4: Ontology Query Model .....	47
Figure 3.5: Billions of Textual Documents Indexed (Sullivan, 2003).....	52
Figure 3.6: Retrieved and Relevant Data .....	55
Figure 4.1: Prototype Architecture.....	58
Figure 4.2: ER Diagram of SQL Server .....	61
Figure 4.3: Tree Structure of Department Schema .....	61
Figure 4.4: Tree Structure of Faculty Schema .....	62
Figure 4.5: Tree Structure of Student Schema .....	62
Figure 4.6: Tree Structure of GradStudent Schema .....	63
Figure 4.7: Tree Structure of Course Schema .....	63
Figure 4.8: Tree Structure of Publication Schema .....	63
Figure 4.9: Prototype Function .....	64
Figure 4.10: The Main Menu of the Prototype .....	65
Figure 4.11: Data Loader – Set Logical Database Size .....	66
Figure 4.12: Data Loader – Modify Logical Database Size of Each Data Source .....	66
Figure 4.13: Data Loader – Modify Number of Tuples of Each Relation .....	67
Figure 4.14: Query Generator – Select Test Query Type.....	68
Figure 4.15: Query Generator – Input Open Query .....	69
Figure 4.16: Query Generator – Delete Open Query .....	69
Figure 4.17: Query Generator – Select Standard Query Type .....	70
Figure 4.18: Query Generator – Set Standard Query Complexity.....	70
Figure 4.19: Scheduler – Test Sequence .....	71
Figure 4.20: Scheduler – Test Repetition.....	72
Figure 4.21: Result Collector .....	73