# TABLE OF CONTENTS

1. INTRODUCTION 01

2. BACKGROUND AND LITERATURE REVIEW 06

2.1. Semantic Web VS Current Web ---------------------------------------------------06
2.2. The Web Ontology---------------------------------------------------------------09
2.3. Engineering Ontologies for Semantic Web----------------------------------------13
2.4. Populating Web Ontologies -----------------------------------------------------15
2.5. Integrating Web Ontologies ------------------------------------------------------17

3. PROPOSED TECHNIQUE FOR SIMILARITY IDENTIFICATION AND MEASUREMENT 28

3.1 The Basis of Proposed Technique---------------------------------------------------28
3.1.1. Comparison of Concepts Rather Than Terms--------------------------------------28
3.1.2. Use of Domain-specific Semantics Rather Than Linguistic Semantics-----------30
3.1.3. Super-Concepts Based Contextual Similarity Measurement------------------------30
3.1.4. Explicit Semantics Based Measurement of Degree of Similarity (DoS) and Semantic Relation (SR) -----------------------------------------------30
3.1.5. The Add-on of Both the DoS and SR with Each Pair of Similar Concepts --------31
3.1.6. Matching Strategy -------------------------------------------------------------31
3.2 Definitions and Structure Diagram--------------------------------------------------36
3.3 IPS - Identifying Primary Similarity ---------------------------------------------41
3.4 ICS - Identifying Contextual Similarity ------------------------------------------44
3.5 IRS - Identifying Role-based Similarity-------------------------------------------47
4. IMPLEMENTATION AND TESTING

4.1 Implementation

4.1.1. Main Module

4.1.2. Ontology Loading Module

4.1.3. Concepts Acquisition Module

4.1.4. Primary Similarity Measurement Module

4.1.5. Super-concepts Acquisition Module

4.1.6. Contextual Similarity Measurement Module

4.1.7. Roles Acquisition Module

4.1.8. Role-based Similarity Measurement Module

4.2. Testing

4.2.1. Case Study 1: Software Development Organization

4.2.1.1. List of Roles

4.2.1.2. List of Concepts of First Input Ontology

4.2.1.3. List of Concepts of Second Input Ontology

4.2.1.4. List of Pairs of Similar Concepts

4.3. Case Study 2: University – The Education Organization

4.3.1. List of Sample Concepts of Input Ontologies

4.3.2. Primary Similarity Identification and Measurement

4.3.3. Contextual Similarity Identification and Measurement

4.3.4. Role-based Similarity Identification and Measurement

4.4. Observations about Results of Different Test-scenarios

4.5. Results: Analysis and Discussion

4.5.1. Completeness
4.5.2. Correctness

4.5.3. Overall Quality of Result

4.5.4. Test Cases for Evaluating Performance

  4.5.4.1. Evaluating Performance with First Test Case

  4.5.4.2. Evaluating Performance with Second Test Case

  4.5.4.3. Evaluating Performance with Third Test Case

  4.5.4.4. Evaluating Performance with Fourth Test Case

4.6. Discussion about Overall Performance

5. CONCLUSION AND FUTURE WORK

5.1 Conclusion

5.2 Recommendations for Further Research

REFERENCES

APPENDICES

I. Ontology of a person’s family

II. A Subset of Roles from Software Development Organization

III. Domain Vocabulary of a Software Development Organization

IV. datasoft.owl

V. ridos.owl
LIST OF FIGURES

2.1 A layered model of semantic web ------------------------------------------- 11
2.2 A sample slice of person’s family ontology --------------------------------- 12
2.3 A slice of taxonomy of furniture ontology-A -------------------------------- 25
2.4 A slice of taxonomy of furniture ontology-B -------------------------------- 26
2.5 A slices of ontologies A & B of a university domain------------------------- 26
3.1 Concept vs. terms ---------------------------------------------------------- 29
3.2: The use of concepts rather than terms in measurement of DoS -------------- 29
3.3 The DoS through edit-distance based formula and through proposed formula ------ 29
3.4 Similarity measurement strategy used in proposed technique----------------- 32
3.5 Degree of similarity measurement through existing techniques --------------- 33
3.6 Degree of similarity measurement through proposed technique
along with new criterion (role-based) ------------------------------------------ 34
3.7 Similarity measurement strategy used in current techniques------------------ 34
3.8 Structure diagram of proposed technique ------------------------------------ 40
3.9 A slice of pseudo code for acquisition of concepts from input ontologies ------- 42
3.10 A slice of pseudo code for identifying primary similarity------------------- 43
3.11 Acquisition of parents of primarily similar concepts------------------------ 45
3.12 A slice of pseudo code for identifying contextual similarity----------------- 46
3.13 Acquisition of roles of contextually similar concepts----------------------- 49
3.14 A slice of pseudo code for identifying role-based similarity----------------- 51
4.1 A prototype for similarity identification and measurement------------------- 54
4.2 A subset of roles in software development organization---------------------- 59
4.3 Completeness wise comparison of results with respect to first test case ------------------- 71
4.4 Correctness wise comparison of results with respect to first test case ------------------- 71
4.5 Overall quality wise comparison of results with respect to first test case ---------------- 72
4.6 Completeness wise comparison of results with respect to second test case --------------- 73
4.7 Correctness wise comparison of results with respect to second test case----------------- 74
4.8 Overall quality wise comparison of results with respect to second test case ----------- 74
4.9 Completeness wise comparison of results with respect to third test case--------------- 76
4.10 Correctness wise comparison of results with respect to third test case-------------- 76
4.11 Overall quality wise comparison of results with respect to third test case--------- 77
4.12 Completeness wise comparison of results with respect to fourth test case----------- 78
4.13 Correctness wise comparison of results with respect to fourth test case----------- 79
4.14 Overall quality wise comparison of results with respect to fourth test case------- 79
4.15 Role-based similarity: expected results vs. produced results--------------------- 80
# LIST OF TABLES

2.1 Relevant results (from top 20 results) of search some engines .......................... 07
2.2 Semantic web vs. current web ................................................................................. 08
2.3 Examples of semantic web systems ....................................................................... 10
2.4 A comparison of some techniques for similarity measurement between ontologies .... 23
3.1 Lexical & linguistic similarity matchers vs. 1st level matcher ............................... 35
3.2 Contextual similarity matcher vs. 2nd level matcher ............................................. 35
3.3 Third level or explicit semantic matcher ............................................................... 35
3.4 Existing techniques vs. proposed technique .......................................................... 52
4.1 A sample slice of intellectual concepts form A ontology ..................................... 60
4.2 A sample slice of intellectual concepts form B ontology ..................................... 61
4.3 A sample list of similar concepts, including pairs having minimum degree of similarity ...................................................................................................................................... 61
4.4 A slice of role-based similar concepts with a threshold-value .............................. 62
4.5 Different scenarios for a pair of concepts to similarity identification .................. 63
4.6 Results of first, second and third levels of similarities ....................................... 65
4.7 Observations about results of different test-scenarios ......................................... 66
4.8 Test-cases for evaluating performance of different techniques ............................ 69
4.9 Results of different techniques with first test case ............................................... 70
4.10 Results of different techniques with second test case ....................................... 73
4.11 Results of different techniques with third test case ........................................... 75
4.12 Results of different techniques with fourth test case ........................................ 77
4.13 Test cases along with status of WordNet and Domain Vocabulary ..................... 81
4.14 Comparison of results related to Heterogeneous Pairs of Similar Concepts (HPoSCs) 81