

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**An Ontology-based
Knowledge Support System
for Requirements Analysis**

A thesis presented
in partial fulfilment of the requirements
for the degree of
Doctor of Philosophy
in Computer Science
at Massey University, Manawatū,
New Zealand

Gang (Jason) Liu

2013

Abstract

An Ontology-based Knowledge Support System for Requirements Analysis (OKSSRA) is proposed and developed, in order to help requirements analysts obtain the preliminary business knowledge.

Requirements Engineering (RE) is a sub-discipline of Software Engineering (SE). It is involved in the whole software life cycle from the very first step throughout the process of software development. Thus, the performances of requirements analysts are crucial to RE outcomes since requirements analysts bridge the communication gap between business stakeholders and development teams.

However, normally, there is a knowledge gap between requirements analysts and business stakeholders, especially when analysts work for out-sourcing contractors. The existence of this knowledge gap may seriously lower analysts' efficiency of communicating with business stakeholders and hinder their performances on preparing requirements documentation. Obviously, preliminary business knowledge related to the project will help analysts to narrow down the knowledge gap and improve their performances. However, based on our survey, there is no existing RE tools providing such knowledge support to analysts. Therefore, we proposed and developed OKSSRA to help analysts obtain the preliminary business knowledge for narrowing down the knowledge gap.

There are three key modules in OKSSRA: (i) a semantic similarity measure module, (ii) an ontology mapping module, and (iii) an automatic use case generating module.

In the semantic similarity measure module of OKSSRA, we proposed and developed a new semantic similarity measure utilising WordNet and

Normalised Google Distance (NGD). In the new measure, NGD will be used to calculate a unique length for each edge in the shortest path between two candidate concepts in the WordNet graph. The semantic similarity measure enables our system (i) to assign related concepts for a user's queries to extend the queries; and (ii) to identify the related business processes from the business knowledge repository.

The ontology mapping module of OKSSRA employs a newly developed ontology mapping method based on MIMapper (Kaza and Chen 2008). In this new ontology mapping method, our newly proposed semantic similarity measure will be used to matching class names and to locate the most informative instances of their class. The ontology mapping module enables our system (i) to update the ontology-based repositories in the system, and (ii) to integrate the ontology-based repositories with other repositories.

In the OKSSRA module for automatically generating Use Cases, we propose a set of mapping rules for the system to automatically generating Use Cases based on the information retrieved from business processes. The set of mapping rules specified how the components of a business process are transformed into use case elements, e.g., actors, goals, and steps of scenarios. With this module, our system is able to generate essential Use Cases automatically using business processes retrieved from MIT Process Handbook.

A set of three test use scenarios and a questionnaire has been carefully designed to evaluate the efficiency and effectiveness of OKSSRA. The experimental results show that (i) our system is useful for obtaining business knowledge, (ii) our system is more effective than existing system developed for similar purpose, and (iii) our system is able to provide a pleasant user experiences.

Dedicated to my parents
for their everlasting love, encouragement and support

Acknowledgement

I would like to take this opportunity to express my appreciation and gratitude to those people who have supported me to achieve this qualification.

My first sincere thanks go to my supervisor, Dr. Ruili Wang, for his invaluable guidance and tremendous support throughout this research. Without his tireless directions and continuing encouragement, it would have been unfeasible for me to achieve my PhD degree. His enthusiasm, support and encouragement have influenced me far beyond this research.

I also wish to express my deep gratitude to my co-supervisors Dr. Jeremy Buckley, Dr. Dave Parry, Dr. Helen Zhou, and A/Prof. Elizabeth Kemp, for the time and effort they have spent with me, during my PhD study. I appreciate their valuable suggestions and constructive comments.

I thank all the participants for their help in the evaluations. Thanks to Jingli, Frank, and Ali, and other friends at Massey University, for their valuable support and friendship.

I gratefully acknowledge the funding from the Foundation for Research, Science and Technology towards my PhD study.

Last but not least, my special thanks go to my parents for their everlasting support, understanding and encouragement.

Contents

Chapter 1. Introduction and Scope.....	1
1.1 Introduction	1
1.2 Scope of this thesis	4
Chapter 2. Motivation and Research Objectives.....	6
2.1 Motivations.....	6
2.1.1 Importance of business knowledge for requirements analysts ..	7
2.1.2 Review of current Requirements Engineering tools	8
2.2 Ontology-based Knowledge Support System for Requirements Analysis (OKSSRA).....	11
2.2.1 High level system structure.....	12
2.2.2 Research issues and proposed solutions	15
2.2.3 Application scope	18
2.3 Summary	20
Chapter 3. Semantic Similarity Measure Module.....	21
3.1 Previous research on semantic similarity measure	22
3.1.1 WordNet.....	22
3.1.2 WordNet-based semantic similarity measures.....	23
3.2 The semantic similarity measure in OKSSRA	28
3.2.1 The Internet as a corpus	29
3.2.2 Use context words to tag word senses	30
3.2.3 Normalised Google distance	31
3.2.4 Calculation of semantic similarity	31
3.3 Evaluation and discussion	32
3.3.1 Comparing with human judgement.....	32
3.3.2 Evaluating with a given NLP application	36
3.4 Summary	40
Chapter 4. Ontology Mapping Module	41
4.1 Introduction	41
4.2 Previous research on ontology mapping	43
4.2.1 Ontology in Computer Science	44
4.2.2 Current research in ontology mapping	47
4.3 The ontology mapping method in OKSSRA.....	48
4.3.1 Linguistic-based matching	50
4.3.2 Instance-based matching	53
4.4 Evaluation and discussion	55
4.4.1 Task description	55
4.4.2 Implementation.....	56
4.4.3 Results and discussion.....	57
4.5 Summary	60

Chapter 5. Automatic Use Case Generating Module.....	62
5.1 Introduction	62
5.2 Previous research on generating Use Cases	65
5.2.1 Use Cases	65
5.2.2 Current research in generating Use Cases from business processes.....	68
5.3 Automatic Use Cases generating in OKSSRA.....	74
5.4 Evaluation and Discussion	79
5.4.1 Research on evaluating the use case quality.....	79
5.4.2 Experiment design	80
5.4.3 Experimental results and discussion.....	82
5.5 Summary	85
 Chapter 6. System Evaluation	 86
6.1 Experiment design.....	86
6.1.1 Scenarios for test use	86
6.1.2 Questionnaires	88
6.1.3 Participants.....	89
6.1.4 Procedures.....	90
6.2 Experimental results and discussion.....	91
6.3 Summary	95
 Chapter 7. Conclusions and Future Work.....	 97
7.1 Summary of main findings and contributions	97
7.1.1 Semantic similarity measure combining knowledge from the Internet and WordNet.....	98
7.1.2 Ontology mapping based on advanced semantic similarity measure and Internet knowledge.....	100
7.1.3 Automatically generating Use Cases using retrieved business processes	101
7.1.4 An efficient and effective knowledge support system for requirements analysts	102
7.2 Opportunities for further research	102
 Bibliography	 104
 Appendix A. Scenarios for Test Use of OKSSRA.....	 116
Appendix B. Questionnaire for OKSSRA Evaluation	118
Appendix C. Questionnaire for Investigating the Quality of Use Cases Generated by AUCG Module.....	120
Appendix D. Review of Current Requirements Engineering Tools.....	121
Appendix E. Previous Research on Ontology Mapping.....	130
Appendix F. Data Obtained from Evaluations	137

List of Figures

FIGURE 2. 1 THE ARCHITECTURE OF PROPOSED SYSTEM	12
FIGURE 2. 2 THE FLOW CHART OF ONE QUERY THREAD	14
FIGURE 2. 3 THE REQUIREMENTS ENGINEERING TOOL TAXONOMY.....	19
FIGURE 3. 1 THE SCATTER PLOT WITH LINEAR MODEL FITS BETWEEN HUMAN BENCHMARK AND JIANG & CONRATH'S ESTIMATIONS.....	34
FIGURE 3. 2 THE SCATTER PLOT WITH LINEAR MODEL FITS BETWEEN HUMAN BENCHMARK AND OUR ESTIMATIONS.....	35
FIGURE 5. 1 DISTRIBUTION OF PATICIPANTS' OPINIONS	83
FIGURE 6. 1 THE STATISTICS OF HOW SUBJECTS RETRIEVED RELEVANT PROCESSES.....	91
FIGURE 6. 2 DISTRIBUTIONS OF SUBJECTS' ANSWERS FOR QUESTIONS IN PART 1 OF THE QUESTIONNAIRE.....	92
FIGURE 6. 3 DISTRIBUTIONS OF SUBJECTS' ANSWERS FOR QUESTIONS IN PART 2 OF THE QUESTIONNAIRE.....	93

List of Tables

TABLE 2. 1 THE FUNCTIONALITIES OF REIEWED RE TOOLS	11
TABLE 3. 1 THE CORRELATION COEFFICIENTS ON MC-SET	34
TABLE 3. 2 THE CORRELATION COEFFICIENTS ON RG-SET	34
TABLE 3. 3 THE CORRELATION COEFFICIENTS OBTAINED BY OUR METHOD ON MC-SET WHEN DIFFERENT VALUES OF α AND β ARE APPLIED.....	36
TABLE 3. 4 THE RECALL RATES OF DETECTING MALAPROPISM ACHIEVED BY JIANG AND CONRATH’S METHOD AND OUR METHOD.....	39
TABLE 4. 1 THE BEST PERFORMANCES OF OUR METHOD AND MIMAPPER...	58
TABLE 4. 2 THE PERFORMANCES OF OUR METHOD AND MIMAPPER.....	59
TABLE 5. 1 EXAMPLES OF THE CONVENTIONAL USE CASE AND THE ESSENTIAL USE CASE FOR “GETTING CASH FROM ATM” SCENARIO	67
TABLE 5. 2 THE MAPPING PROPOSED BY DIJKM AND JOOSTEN (2002)	73
TABLE 5. 3 THE AVERAGE SCORES OF SUBJECTS’ OPINIONS IN AUCG MODULE EVALUATION.....	84
TABLE 6. 1 THE MEAN ABSOLUTE DEVIATION OF SUBJECTS' OPINIONS ON OKSSRA	95