การค้นคืนยูสเคสโดยใช้การคำนวณความคล้ายกันของพจน์และโครงสร้างของยูสเคส

นายอัคเดช อุดมชัยพร

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมซอฟต์แวร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2548

ISBN 974-17-3756-4

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

USE CASE RETRIEVAL USING TERMS AND USE CASE STRUCTURE SIMILARITY COMPUTATION

Akadej Udomchaiporn

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science Program in Software Engineering

Department of Computer Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 2005

ISBN 974-17-3756-4

Thesis Title	Use Case Retrieval Using Terms and Use Case Structure Similarity Computation
By	Akadej Udomchaiporn
Field of study	Software Engineering
Thesis Advisor	Nakornthip Prompoon
Thesis Co-advisor	Pizzanu Kanongchaiyos, Ph.D.

Accepted by the Faculty of Engineering, Chulalongkorn University in Partial Fulfillment of Requirements for the Master's Degree

..... Dean of the Faculty of Engineering
(Professor Direk Lavansiri, Ph.D.)

THESIS COMMITTEE

..... Chairman
(Associate Professor Wanchai Rivepaiboon, Ph.D.)

..... Thesis Advisor
(Nakornthip Prompoon)

..... Thesis Co-advisor
(Pizzanu Kanongchaiyos, Ph.D.)

..... Member
(Associate Professor Veera Boonjing, Ph.D.)

..... Member
(Daricha Sutivong, Ph.D.)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

นายอัคเดช อุดมชัยพร : การค้นคืนยูสเคสโดยใช้การคำนวณความคล้ายกันของพจน์และโครงสร้างของยูสเคส. (USE CASE RETRIEVAL USING TERMS AND USE CASE STRUCTURE SIMILARITY COMPUTATION) อ. ที่ปรึกษา : อาจารย์นครทิพย์ พร้อมพูล. อ.ที่ปรึกษาร่วม : อาจารย์ ดร.พิษณุ คนองชัยยศ จำนวน 161 หน้า. ISBN 974-17-3756-4.

การนำซอฟต์แวร์กลับมาใช้ใหม่ถือเป็นหัวใจสำคัญของวิธีการทางด้านวิศวกรรมซอฟต์แวร์ เพราะการนำซอฟต์แวร์กลับมาใช้ใหม่สามารถช่วยผู้พัฒนาซอฟต์แวร์ลดต้นทุนของการพัฒนาซอฟต์แวร์ลงได้ เนื่องจากในปัจจุบันอุตสาหกรรมซอฟต์แวร์มีการแข่งขันกันสูง ผู้พัฒนาซอฟต์แวร์จึงพยายามหาวิธีการที่จะลดต้นทุนการพัฒนาซอฟต์แวร์ให้ได้มากที่สุด ซึ่งการนำซอฟต์แวร์ในส่วนของข้อกำหนดความต้องการซอฟต์แวร์ที่จัดเก็บอยู่ในรูปของเอกสารการบรรยายยูสเคสกลับมาใช้ใหม่นั้นจะสามารถช่วยให้ผู้พัฒนาซอฟต์แวร์ลดการทำงานที่ซ้ำซ้อนลงไปได้มาก จึงจะนำไปสู่การลดต้นทุนการพัฒนาซอฟต์แวร์ลงได้มากตามไปด้วย

วิทยานิพนธ์ฉบับนี้จึงมีวัตถุประสงค์เพื่อวิเคราะห์และออกแบบวิธีการค้นคืนข้อกำหนดความต้องการซอฟต์แวร์ในรูปแบบของเอกสารการบรรยายยูสเคส โดยพิจารณาความคล้ายกันของพจน์และโครงสร้างของยูสเคส ซึ่งได้นำทฤษฎีการจัดเก็บและค้นคืนข้อมูล เช่น การสร้างดรรชนีอัตโนมัติ การกำหนดน้ำหนักของพจน์ การคำนวณความคล้ายกันของเอกสาร และการประเมินผลการค้นคืนเอกสาร มาประยุกต์ใช้งานกับวิธีการค้นคืนเอกสารการบรรยายยูสเคสนี้ งานวิจัยนี้ได้พัฒนาเครื่องมือขึ้นมาเพื่อทดสอบแนวคิดที่นำเสนอ และยังได้มีการออกแบบการทดลองเพื่อเปรียบเทียบผลการค้นคืนเอกสารการบรรยายยูสเคสโดยพิจารณาความคล้ายกันของพจน์และโครงสร้างของยูสเคสซึ่งเป็นวิธีการที่ได้นำเสนอ กับการค้นคืนเอกสารการบรรยายยูสเคสโดยพิจารณาจากความคล้ายกันของพจน์เพียงอย่างเดียว ซึ่งเหมือนกับการค้นคืนเอกสารโดยทั่วไป มาตรวัดของการจัดเก็บและค้นคืนข้อมูลคือ รีคอล พรีซิชัน และฮาร์โมนิกมีน ได้นำมาใช้ในการประเมินผลการค้นคืน ผลการทดลองของงานวิจัยนี้แสดงให้เห็นว่าการค้นคืนเอกสารการบรรยายยูสเคสโดยการพิจารณาโครงสร้างของยูสเคสให้ผลลัพธ์ที่ดีกว่าการค้นคืนเอกสารการบรรยายยูสเคสโดยใช้วิธีการค้นคืนที่เหมือนกับการค้นคืนเอกสารโดยทั่วไป

ภาควิชา วิศวกรรมคอมพิวเตอร์	ลายมือชื่อนิสิต.....อัคเดช อุดมชัยพร.....
สาขาวิชา วิศวกรรมซอฟต์แวร์	ลายมือชื่ออาจารย์ที่ปรึกษา.....นครทิพย์ พร้อมพูล.....
ปีการศึกษา 2548	ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

4770536321 : MAJOR SOFTWARE ENGINEERING

KEY WORD: USE CASE / USE CASE DESCRIPTION / INFORMATION STORAGE AND RETRIEVAL / SIMILARITY COMPUTATION / SOFTWARE REUSE

AKADEJ UDOMCHAIPORN : USE CASE RETRIEVAL USING TERMS AND USE CASE STRUCTURE SIMILARITY COMPUTATION. THESIS ADVISOR : NAKORNTHIP PROMPOON, THESIS COADVISOR : PIZZANU KANONGCHAIYOS, Ph.D., 161 pp. ISBN 974-17-3756-4.

Software reuse is a key component of software engineering methodology as it helps software developers to reduce cost of software development process. At present, the software industry is highly competitive among organizations. Therefore, software developers have to find ways to reduce cost of software production as much as they can. Software reuse in a part of software requirements specification collected in a form of use case description can reduce much redundant works. Thus, this leads to the reduction of cost and time of software production.

The objective of this thesis is to analyze and design an approach for retrieving use case considering similarity of terms and structure of use cases. Information storage and retrieval theories such as automatic indexing, term weighting system, document similarity computation, and retrieval evaluation are applied in this approach. As a result, this research develops a tool to test the proposed approach. The experiment is also designed to compare this proposed approach to the existing approach which considers only terms similarity. The 3 metrics named recall, precision, and harmonic mean of information storage and retrieval theory are used to evaluate use case retrieval results. The results of this experiment indicate that this proposed approach is more effective than the existing approach.

Department Computer Engineering
Field of study Software Engineering
Academic year 2005

Student's signature..Akadej Udomchaiporn
Advisor's signature..Nakornthip Prompoon
Co-advisor's signature..

ACKNOWLEDGEMENTS

Many people have greatly contributed to the success of this thesis. I would like to especially thank my thesis advisor, Nakornthip Prompoon, for her insightful and intimate suggestions, comments, and feedback throughout this thesis. Also, I would like to thank my thesis co-advisor, Pizzanu Kanongchaiyos, Ph.D. for many helpful comments and discussions on issues related to this thesis.

I would like to sincerely thank my other thesis committee members, Associate Professor Wanchai Rivepaiboon, Ph.D., Associate Professor Veera Boonjing, Ph.D., and Daricha Sutiwong, Ph.D. who always provided valuable advice.

I would like to thank my friends, software engineering students, the volunteers, who participated in my experiments. Your patience in enduring the long query sessions is highly appreciated.

I would like to thank all faculties in the Department of Computer Engineering, Chulalongkorn University, and my all friends, both in software engineering lab and in my class, for their kindness and helpfulness all along.

Most of all, I would like to extremely thank my father, my mother, and everyone in my family, who best take care of me all along since I was born.

Lastly, I would like to thank King Mongkut's Institute of Technology Ladkrabang and Office of the Higher Education Commission for their financial supporting.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

CONTENTS

PAGE

ABSTRACT (THAI) iv

ABSTRACT (ENGLISH) v

ACKNOWLEDGEMENT vi

CONTENTS vii

LIST OF TABLE xi

LIST OF FIGURES xvi

CHAPTER

1 INTRODUCTION 1

1.1 Background and importance of research 1

1.2 Research objective 3

1.3 Research scope 3

1.4 Research benefit 3

1.5 Research procedure 4

1.6 Research organization 5

2 CONCEPT, THEORIES, AND RELATED WORKS 7

2.1 Concept and theories 7

2.1.1 Software reuse 7

2.1.2 A use case diagram 8

2.1.2.1 Components of a use case diagram 9

2.1.2.2 A use case description 11

2.1.3 Information storage and retrieval 15

2.1.3.1 Information storage process 15

2.1.3.1.1 Automatic indexing 15

2.1.3.1.2 Term weighting system 17

CHAPTER PAGE

2.1.3.2 Information retrieval process 18
2.1.3.3 The evaluation process 18
2.2 Related works 20
2.2.1 Reusing use case descriptions for requirements specifications: Towards use case patterns 20
2.2.2 Reuse of scenario specifications using an automated relational learner: A lightweight approach 20
2.2.3 Reusing UML specifications in a constrained application domain ... 21

3 THE APPROACH 22
3.1 Use case collection and index creation 24
3.2 Query generation and retrieval process 25
3.3 The system evaluation 27

4 TOOL DEVELOPMENT 28
4.1 Supporting tools 28
4.1.1 Apache Tomcat web server 28
4.1.2 MySQL database server 29
4.1.3 Java Server Page (JSP) 29
4.2 Tool architecture 29
4.3 Function of the tool 30
4.3.1 Use case storage 31
4.3.2 Use case retrieval by use case keywords 33
4.3.3 Use case retrieval by use case structure 34
4.4 Data model 36

5 EXPERIMENTS 41

CHAPTER PAGE

5.1 Experimental objective 41
5.2 Experimental method 41
5.2.1 Use cases 41
5.2.2 Testers 41
5.2.3 Topics 42
5.2.4 Queries 43
5.3 Experimental procedure 45
5.4 Experimental environment 46
5.4.1 Similarity threshold 46
5.4.2 High frequency threshold 47
5.5 Experimental results 47
5.6 Experimental result analysis 50
5.7 Experimental result summary 55
5.8 Discussion 56

6 RESEARCH SUMMARY 61
6.1 Research summary 61
6.2 Future works 62
6.3 Publication 62

REFERENCES 63

APPENDICES 65
A Selected use case domain 66
B Tester's use case query 83
C Similarity score and recall precision, and harmonic mean of the experiments 113

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

CHAPTER PAGE

D Publication 153

BIOGRAPHY 161

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

LIST OF TABLES

TABLE		PAGE
4.1	Use case details	37
4.2	Use case domain	37
4.3	Use case component	37
4.4	All terms in the collection	38
4.5	All terms in each use case	38
4.6	Stop list	38
4.7	Tester	38
4.8	Logged data of tester's query by use case keywords	39
4.9	Logged data of tester's query by use case structure	39
4.10	Topic	39
4.11	Result of query by use case query	40
5.1	The number of all queries for each tester	43
5.2	Average and maximum similarity score of query by use case structure and keywords (1) for each subject	47
5.3	Average recall, precision, and harmonic mean of query by use case structure and keywords (1) for each subject	48
5.4	Average and maximum similarity score of query by use case structure and keywords (2) for each subject	48
5.5	Average recall, precision, and harmonic mean of query by use case structure and keywords (2) for each subject	49
5.6	z-test (1) of recall, precision, and harmonic mean for subject A	51
5.7	z-test (1) of recall, precision, and harmonic mean for subject B	51
5.8	z-test (1) of recall, precision, and harmonic mean for subject C	51
5.9	z-test (1) of recall, precision, and harmonic mean for subject D	51
5.10	z-test (1) of recall, precision, and harmonic mean for subject E	52
5.11	z-test (1) of recall, precision, and harmonic mean for all subjects	52

TABLE PAGE

5.12 z-test (1) summary of the experiment 52
5.13 z-test (2) of recall, precision, and harmonic mean for subject A 53
5.14 z-test (2) of recall, precision, and harmonic mean for subject B 53
5.15 z-test (2) of recall, precision, and harmonic mean for subject C 53
5.16 z-test (2) of recall, precision, and harmonic mean for subject D 53
5.17 z-test (2) of recall, precision, and harmonic mean for subject E 54
5.18 z-test (2) of recall, precision, and harmonic mean for all subjects 54
5.19 z-test (2) summary of the experiment 54
5.20 Summary of weighted value of query by use case structure 57
5.21 Average recall, precision, and harmonic mean of query by use case structure using 3 components and by normal use case structure for each subject 58
5.22 Average recall, precision, and harmonic mean of query by use case structure using 7 components and by normal use case structure for each subject 60

A.1 All selected use case domains 66
A.2 ICT blocking and filtering system 67
A.3 Departmental information system for curriculum and course offerings Management 68
A.4 Resource management system 69
A.5 Automatic restaurant lookup system 70
A.6 Library management system 71
A.7 Mobile phone selling on the internet system 72
A.8 Payroll system 72
A.9 Personal stock manager on PDA system 74
A.10 Basic product trading system 75
A.11 Mobile phone service paying system 75

TABLE PAGE

A.12 Basic MLM computerize system 76
A.13 Content management and platform upgrade for the official website of Nation Channel 77
A.14 Purchase, producing, and export system 78
A.15 Management information system for insurance company 79
A.16 E-learning system 80
A.17 Management information system for an academic 81

B.1 Tester's query set of query by use case keywords for subject A 83
B.2 Tester's query set of query by use case keywords for subject B 85
B.3 Tester's query set of query by use case keywords for subject C 87
B.4 Tester's query set of query by use case keywords for subject D 89
B.5 Tester's query set of query by use case keywords for subject E 91
B.6 Tester's query set of query by use case structure for subject A 93
B.7 Tester's query set of query by use case structure for subject B 97
B.8 Tester's query set of query by use case structure for subject C 101
B.9 Tester's query set of query by use case structure for subject D 105
B.10 Tester's query set of query by use case structure for subject E 109
B.11 Components of the use case description 112

C.1 Average and maximum similarity score of query by use case structure and keywords (1) for subject A 113
C.2 Average and maximum similarity score of query by use case structure and keywords (1) for subject B 115
C.3 Average and maximum similarity score of query by use case structure and keywords (1) for subject C 117

TABLE PAGE

C.4 Average and maximum similarity score of query by use case structure and keywords (1) for subject D 119
C.5 Average and maximum similarity score of query by use case structure and keywords (1) for subject E 121
C.6 Average and maximum similarity score of query by use case structure and keywords (2) for subject A 123
C.7 Average and maximum similarity score of query by use case structure and keywords (2) for subject B 125
C.8 Average and maximum similarity score of query by use case structure and keywords (2) for subject C 127
C.9 Average and maximum similarity score of query by use case structure and keywords (2) for subject D 129
C.10 Average and maximum similarity score of query by use case structure and keywords (2) for subject E 131
C.11 Recall, precision, and harmonic mean of query by use case structure and keywords (1) for subject A 133
C.12 Recall, precision, and harmonic mean of query by use case structure and keywords (1) for subject B 135
C.13 Recall, precision, and harmonic mean of query by use case structure and keywords (1) for subject C 137
C.14 Recall, precision, and harmonic mean of query by use case structure and keywords (1) for subject D 139
C.15 Recall, precision, and harmonic mean of query by use case structure and keywords (1) for subject E 141
C.16 Recall, precision, and harmonic mean of query by use case structure and keywords (2) for subject A 143
C.17 Recall, precision, and harmonic mean of query by use case structure and keywords (2) for subject B 145

TABLE PAGE

C.18 Recall, precision, and harmonic mean of query by use case structure and keywords (2) for subject C 147

C.19 Recall, precision, and harmonic mean of query by use case structure and keywords (2) for subject D 149

C.20 Recall, precision, and harmonic mean of query by use case structure and keywords (2) for subject E 151

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

LIST OF FIGURES

FIGURE PAGE

1.1 The activity diagram of research procedure 6

2.1 An example of the use case diagram for credit card validation system 8
2.2 An example of components of a use case diagram 9
2.3 An example of relationships of a use case diagram 10
2.4 An example of a use case description for withdrawing funds 13
2.5 An example of a use case description for transferring funds 14
2.6 The activity diagram of indexing process 16

3.1 The activity diagram of the overview approach 23
3.2 The activity diagram of index creation 24
3.3 The activity diagram of retrieval process 25
3.4 The activity diagram of the evaluation process 27

4.1 The component diagram of developed tool architecture 30
4.2 The menu screen of use case retrieval system 31
4.3 The use case storage screen of use case retrieval system 32
4.4 The use case retrieval by use case keywords screen of use case retrieval system 33
4.5 The use case retrieval by use case structure screen of use case retrieval system 34
4.6 The use case retrieval result screen of use case retrieval system 35
4.7 The entity relationship diagram of use case retrieval system 36

5.1 The activity diagram of the experimental method 44
5.2 The activity diagram of the experimental procedure of the tester 45
5.3 A use case description template for use case retrieval using 3 components 58
5.4 A use case description template for use case retrieval using 7 components 59

CHAPTER I

INTRODUCTION

In this chapter, the overall concept of this thesis is introduced. It includes background and importance, objective, scope, benefit, and procedure of our research. Details of them are described as follows.

1.1 Background and importance of research

Nowadays, there is a high competition among software developers in software development industry. Therefore, software developers have to find ways to reduce cost of software development process as much as they can in order to decrease software production cost to gain competition advantage in the software market. Software reuse [1] is one of efficient approaches to support this idea. There are many components of software which can be reused. Each of them may be a consequent product from early phases of software development process, and reusing software components is more efficient when they are reused in early phases of software development process such as requirements analysis phase. When those software components are reused in the early phases, their consequent products in the following phases can be identified and reused too. This finally leads to reduction in the time and cost of software production and to increase software quality.

In requirements analysis and design phase, UML (Unified Modeling Language) [2] is always selected among software developers to represent user's requirements and software characteristics from many points of view and it helps developers to plan activities in the next phases efficiently. The UML is composed of many kinds of diagrams, but one of them which is very important for requirements analysis phase is a use case diagram. This diagram is always used for capturing user's requirements and for indicating software functional requirements. Moreover, it is also used as an agreement between users and developers. Each use case has a use case description associated with it. It is used for describing details of each use case. A use case description counts as a software component in requirements analysis phase which can

be efficiently reused [3]. Due to the fact that the details of use case descriptions are collected in a natural language or textual format, information storage and retrieval theory can be applied for use case retrieval. Therefore, some theories of information storage and retrieval such as automatic indexing, weighting system, similarity computation, and retrieval evaluation can be applied in use case reuse process.

According to several research conducted on this issue at the moment, there are many related works [4-7] about use case retrieval but they still have some disadvantages such as high complexity for users, low performance process, manual process, and lack of use case structure consideration. For example, in Woo's research [7], low usability is emerged because the users have to generate normal flow of events as query data. In addition to low usability, his approach has low performance because those query data is transformed into graph format. Using similarity computation between query data and data in the collection by using graph format is time consuming. For another disadvantage, Blok's research [5] is manual processing in both storage and retrieval process. Some activities in his approach are performed by the specialists. Those activities depend on experience of the specialists, and the retrieved results may be biased. Another disadvantage which is very important is lack of use case structure consideration. All current related research [5-7] do not consider the whole structures of use cases. Some research [5, 7] consider only normal flow of events of use cases while the other [6] does not consider use case structure. Thus, the retrieved results may not be efficient because of missing use case structure consideration.

From those limitations, this thesis proposes an approach to reduce disadvantages in previous works. It presents an approach for retrieving use cases by using terms and use case structure similarity computation. Some information storage and retrieval theories are applied to some processes of the approach. Weighting technique is also applied to consider all structures of a use case by weighting each component of them. Retrieving use cases by considering their structure helps developers to retrieve them more efficient than retrieving them by considering only their terms.

1.2 Research objective

1. To design an approach for retrieving use cases by considering terms and use case structure similarity between use case query and use cases in the collection.
2. To develop a tool for testing concept of this approach.

1.3 Research scope

1. The developed tool supports only the storage and retrieval process.
2. Input of the storage process is use case descriptions, and output is a set of indices and its weighted values stored in the library.
3. Input of the retrieval process is a set of keywords in a use case query and weighted values generated by the users, and output is a list of use case descriptions.
4. Input of the evaluation process is a list of relevant use cases defined by the users, and output is the value of recall, precision, and harmonic mean.
5. Recall, precision, and harmonic mean are 3 metrics for comparing the results of this approach.
6. Output of the retrieval process displayed by the tool is only first 10 use cases which their similarity value are more than a predefined threshold value.
7. Output of the retrieval process may not be displayed by the tool if no use case description has similarity value more than a predefined threshold value.
8. The approach supports use case descriptions in only English language.
9. The number of test data is 16 systems specified with use case description, and each system has the number of use case descriptions at least 10.
10. The approach does not handle ambiguous term, so terms which are the same word are considered as the same meaning.

1.4 Research benefit

1. A developed tool from this approach can be used for use case retrieval.
2. The proposed approach helps users to reduce cost and time in software development process.

3. The proposed approach is a model for researchers who are interested in software reuse and information storage and retrieval.

1.5 Research procedure

The research procedure is composed of two main parts, proof of concept and tool development.

Part 1 : Proof of concept

1. Select system domain and study use case diagram of each domain including its use case description in details.
2. Analyze and design the overview approach for retrieving use cases.
3. Design processes for storing use cases in a proper form which is convenient for retrieval.
   - 3.1 Design automatic indexing process.
   - 3.2 Design term weighting process.
4. Design processes for retrieving use cases which are relevant to user's use case query.
   - 4.1 Design an algorithm for computing terms similarity between user's use case query and each use case in database.
   - 4.2 Define weighted values for each element of use case structure for computing structure similarity between user's use case query and each use case in database.
   - 4.3 Design an algorithm for computing terms and structure similarity between user's use case query and each use case in database.
5. Select some metrics for evaluating the system results.
6. Test the proposed approach by the developed tool with test data which is a set of example use cases.
7. Evaluate system results with selected metrics.
8. Summarize the results and the proposed approach and document thesis.

Part 2 : Tool development

9. Design overview architecture of the tool for supporting storage and retrieval process.
10. Design function of the tool.
11. Design user's interface of the tool.
12. Design relational database structure.
13. Develop the designed tool.

1.6 Research organization

In this thesis, the overview concept of the proposed approach is introduced in chapter 1. After that, background theories such as software reuse, use case diagram, information storage and retrieval, and some related research are mentioned in chapter 2. In chapter 3, the approach for retrieving use cases is described in detail. It includes storage, retrieval, and evaluation process. In chapter 4, design and development of supporting tool are described. The objective, method, and procedure of our experiment including their results are described and discussed in chapter 5. Finally, research summary and future works are mentioned in chapter 6. The overall procedure of our research is shown as the activity diagram in figure 1.1.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

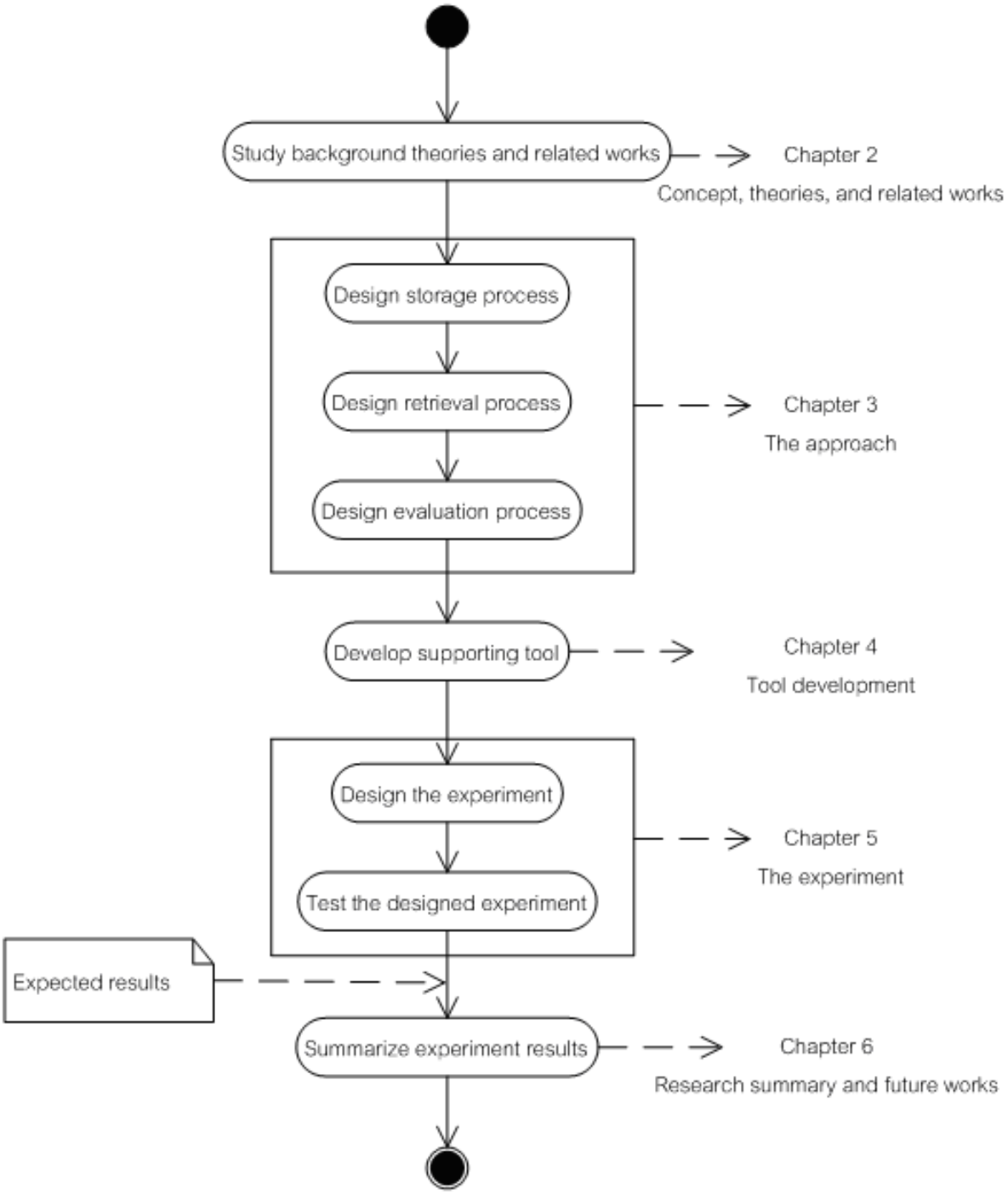


Figure 1.1 The activity diagram of research procedure

CHAPTER II

CONCEPT, THEORIES, AND RELATED WORKS

In this chapter, concept, theories, and related works which are related to our research are introduced. They are both approach part and tool development part. Their details are described as follows.

2.1 Concept and theories

2.1.1 Software reuse

Software reuse [1, 3, 8] is the process of creating software systems from existing software rather than building software systems from scratch. Software systems are generally composed of many parts or components. The components of each new software system can be assembled from predefined reusable components. Reusable components may be program sourcecode, but the bigger benefits from software reuse often come from a broader and higher-level view of reusable components. Software specifications, designs, test scripts, project plans, documentation, object frameworks, and subroutines are all examples of reusable components [3]. In general, any software components could be produced from reusable components. Increasing opportunities to reuse enables significant software productivity, quality and cost improvements. The major benefits of software reuse [8] are to

- increase software productivity
- shorten software development time
- develop software with fewer people
- move personnel, tools and methods more easily from project to project
- reduce software costs
- produce better quality software
- improve software system interoperability
- provide a competitive advantage

In the ideal case, reuse in the software development life cycle is performed at a higher, broader level such as planning or requirements analysis phase. Examples of reusable software components in this phase are prose, use cases, state diagrams, sequence diagrams, data flow diagrams, and sketch user interfaces.

2.1.2 A use case diagram

A use case diagram [2] is one of the UML diagrams always used for capturing functional requirements from requirements specification in requirements analysis phase. It can indicate the capabilities of a software system, so it can be used as an agreement between users and developers. Moreover, it is used for planning activities in the following phases of software development process. A use case diagram consists of many parts; there are a set of actors, use cases, and their relationships. An example of a use case diagram of credit card validation system is depicted in figure 2.1.

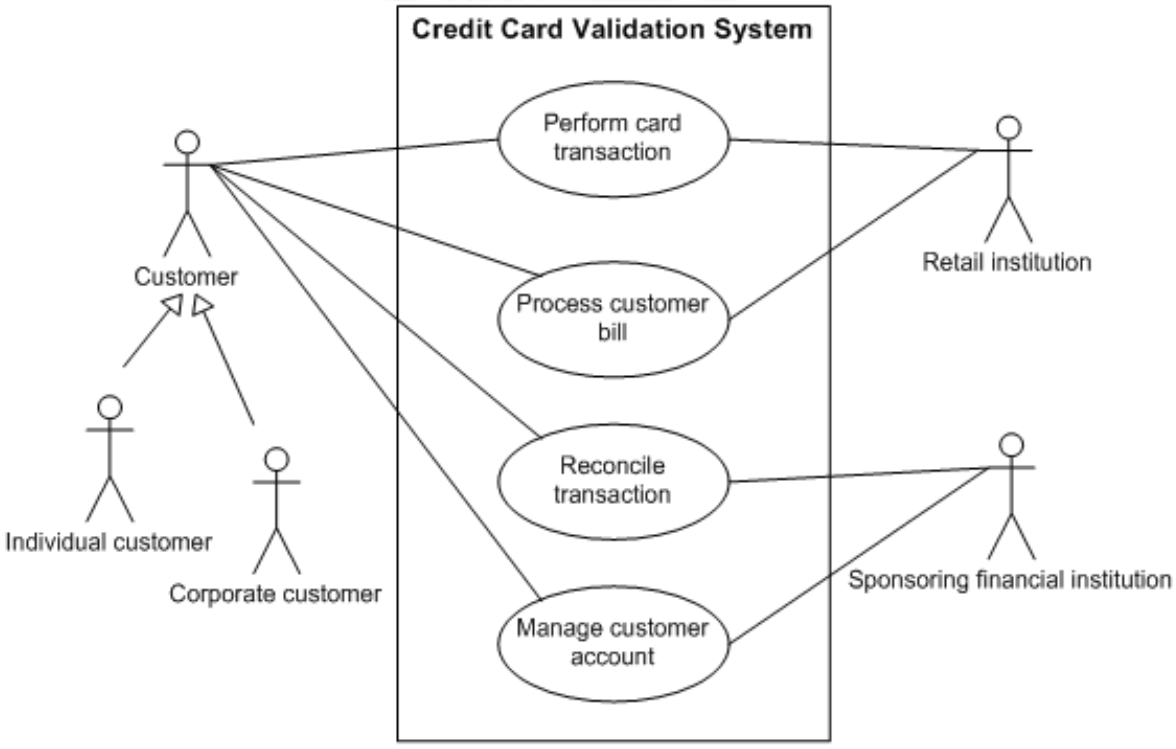


Figure 2.1 An example of the use case diagram for credit card validation system

2.1.2.1 Components of a use case diagram

There are many components of a use case diagram. An illustration of them is shown as an example in figure 2.2, and details of them are described below the figure.

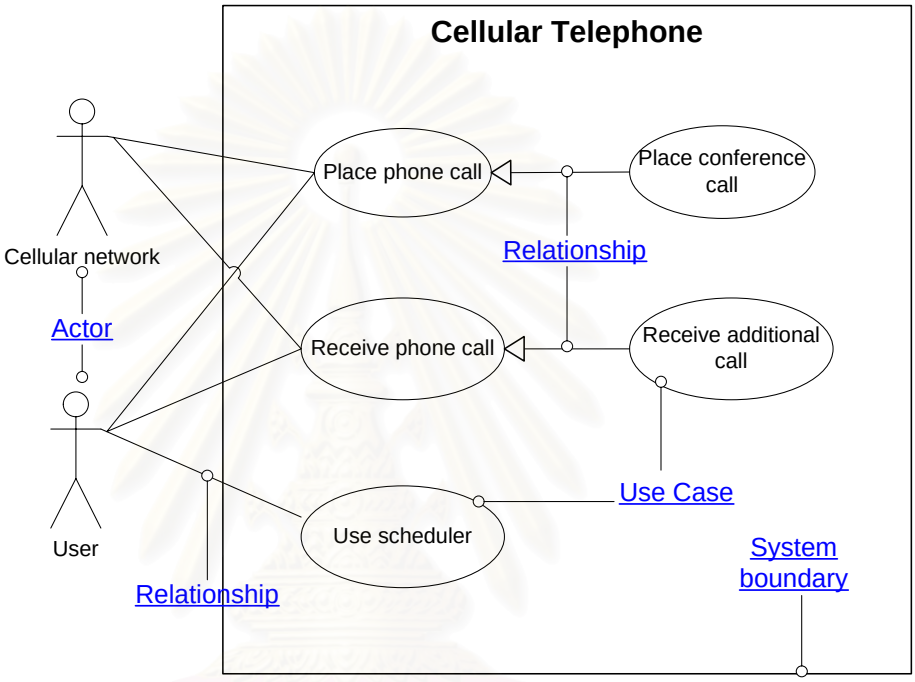


Figure 2.2 An example of components of a use case diagram

1) ***Use case***: In UML, a use case is a complete task of a system that provides a measurable result of value for an actor. More formally, a use case defines a set of use case instances or scenarios. Graphically, a use case is rendered as an ellipse.
2) ***Actor***: An actor is someone or something outside the system that interacts with the system. It can be connected to only use cases by association relationship.
3) ***Relationship***: Relationship is a semantic connection between model elements. In a use case diagram, relationships are consisted of associations, dependencies, and generalizations. Graphically, a relationship rendered as a path, with different kinds

of lines used distinguish the kinds of relationships. An example of relationships is illustrated in figure 2.3.

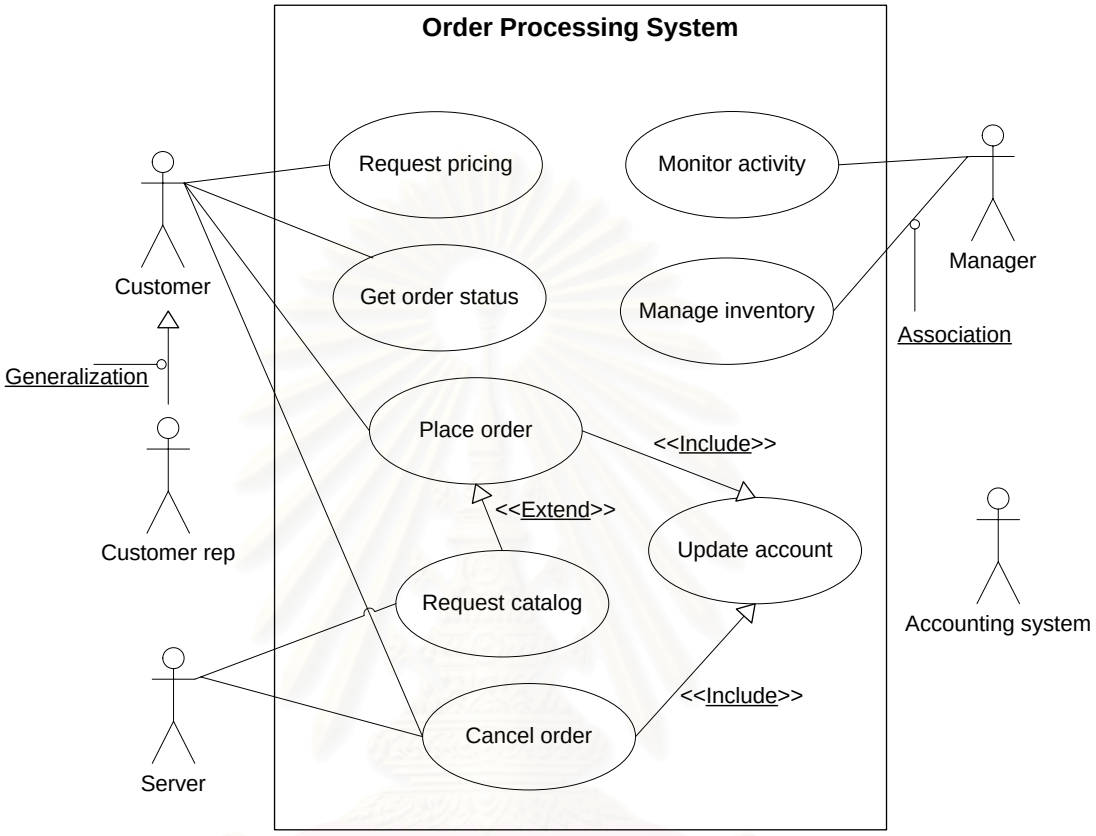


Figure 2.3 An example of relationships of a use case diagram

- ***Dependency*** is a relationship between use cases working together. The dependency relationship can be divided in two sub-relationships which are “include” and “extend”.

  Include is a relationship from a base use case to an included use case specifying how the behavior defined for the included use case can be inserted into the behavior defined for the base use case.

  Extend is a relationship from an extending use case to a base use case specifying how the behavior defined for the extending use case can be

operationally inserted into the behavior defined for the base use case.

- ***Generalization*** is a relationship between a general thing (called the superclass or parent) and a more specific kind of that thing (called the subclass or child). Generalization is sometimes called an "is-a-kind-of" relationship. Generalization can be a relationship between both use cases and both actors.
- ***Association*** is a structural relationship that specifies that objects of one thing are connected to objects of another. In a use case diagram, the objects are use cases and actors, so an association in use case diagram is a relationship between use cases and actors.

2.1.2.2 A use case description

A use case description [9, 10] is a document describing details of a use case. The elements of a use case description are not standardized but the fundamental elements of it which cover important content and which are generally used consist of 12 elements. The examples [11] of a use case description for withdrawing funds and transferring funds are depicted in figure 2.4 and 2.5, and details of them are described as follows.

- ***Use Case Name:*** Use case name is a name of use cases. Every use case must have a name that distinguishes it from other use cases.
- ***Objective***: An objective is a part describing about an objective of use cases.
- ***Actor***: An actor is someone or something outside the system that interacts with the system.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

- ***Relationship***: Relationship is a semantic connection between model elements. In a use case description, relationships are composed of 4 sub-elements; there are
    - *Association*
    - *Include*
    - *Extend*
    - *Generalization*
- ***Precondition***: Precondition is a constraint that must be true when a use case is invoked.
- ***Postcondition***: Postcondition is a constraint that must be true when a use case has ended.
- ***Normal Flow of Events***: Normal flow of events is the part of a use case that describes its most common implementation. The basic flow is written assuming that no errors or alternatives exist. Also called basic path or normal path.
- ***Subflow***: In some cases, normal flow of events can be decomposed into a set of subflows to keep the normal flow of events as simple as possible.
- ***Alternative or Exceptional Flow of Events***: Alternative or exceptional flow of events is the part of a use case that describes its alternative implementations. It is also used to describe error conditions, since errors can be considered a kind of alternative. It is also called alternative path.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

<table>
<tr><td colspan="3"><u>Use Case Description for Withdrawing Funds</u></td></tr>
<tr><td>Use Case Name :</td><td colspan="2">Withdraw Funds</td></tr>
<tr><td>Objective :</td><td colspan="2">Customer withdraws a specific amount of funds from a valid bank account</td></tr>
<tr><td>Actor :</td><td colspan="2">ATM Customer</td></tr>
<tr><td rowspan="4">Relationship :</td><td>Association :</td><td>ATM Customer</td></tr>
<tr><td>Include :</td><td>-</td></tr>
<tr><td>Extend :</td><td>-</td></tr>
<tr><td>Generalization :</td><td>-</td></tr>
<tr><td>Precondition :</td><td colspan="2">ATM is idle, displaying a welcome message</td></tr>
<tr><td>Postcondition :</td><td colspan="2">Customer funds have been withdrawn</td></tr>
<tr><td>Normal Flow of Events :</td><td colspan="2">1. Customer inserts the ATM card into the card reader.
2. The system reads the card number.
3. System prompts customer for PIN number.
4. Customer enters PIN.
5. System checks the expiration date and whether the card is lost or stolen.
6. System checks whether the user-entered PIN matches the card PIN maintained by the system.
7. System checks what account are accessible with the ATM card.
8. System displays customer accounts and prompts customer for transaction type.
9. Customer selects Withdrawal, enters the amount, and selects the account number.
10. System checks whether customer has enough funds in the account.
11. System authorizes dispensing of cash.
12. System dispenses the cash amount.
13. System prints a receipt.
14. System ejects card.</td></tr>
<tr><td>Subflow :</td><td colspan="2">-</td></tr>
<tr><td>Alternative or Exceptional Flow of Events :</td><td colspan="2">2-a. If the system does not recognize the card, the card is ejected.
5-a. If the system determines that the card has expired, the card is confiscated.
5-b. If the system determines that the card has been reported lost or stolen, the card is confiscated.
6-a. If the customer entered PIN does not match the PIN number for this card, the system re-prompts for the PIN.
6-b. If the customer enters the incorrect PIN three times, the system confiscates the card.
7-a. If the system determines the account number is invalid, it ejects the card.
10-a. If the system determines there are insufficient funds in the customer's account, it ejects the card.
11-a. If the ATM out of funds, the system displays an apology, ejects the card, and shuts down the ATM.
If the customer enters Cancel, the system cancels the transaction and ejects the card.</td></tr>
</table>

Figure 2.4 An example of a use case description for withdrawing funds

A Use Case Description for Transferring Funds		
Use Case Name :	Transfer Funds	
Objective :	Customer transfers funds from one valid bank account to another.	
Actor :	ATM Customer	
Relationship :	Association :	ATM Customer
	Include :	-
	Extend :	-
	Generalization :	-
Precondition :	ATM is idle, displaying a welcome message	
Postcondition :	Customer funds have been transfered	
Normal Flow of Events :	1. Customer inserts the ATM card into the card reader. 2. The system reads the card number. 3. System prompts customer for PIN number. 4. Customer enters PIN. 5. System checks the expiration date and whether the card is lost or stolen. 6. System checks whether the user-entered PIN matches the card PIN maintained by the system. 7. System checks what account are accessible with the ATM card. 8. System displays Transfer, enters amount, from account, and to account. 9. Customer selects withdrawal, enters the amount, and selects the account number. 10. System performs the transfer. 11. System prints a receipt. 12. System ejects card.	
Subflow :	-	
Alternative or Exceptional Flow of Events :	2-a. If the system does not recognize the card, the card is ejected. 5-a. If the system determines that the card has expired, the card is confiscated. 5-b. If the system determines that the card has been reported lost or stolen, the card is confiscated. 6-a. If the customer entered PIN does not match the PIN number for this card, the system re-prompts for the PIN. 6-b. If the customer enters the incorrect PIN three times, the system confiscates the card. 8-a. If the system determines the from account number is invalid, it ejects the card. 8-b. If the system determines the to account number is invalid, it ejects the card. 10-a. If the system determines there are insufficient funds in the customer's from account, it ejects the card. If the customer enters Cancel, the system cancels the transaction and ejects the card.	

Figure 2.5 An example of a use case description for transferring funds

2.1.3 Information storage and retrieval

The information storage and retrieval theory relating to this thesis has 3 main processes. They are storage, retrieval, and evaluation process. Details of them applied to this thesis are described in this section.

2.1.3.1 Information storage process

In storage process, the processes included in this thesis are automatic indexing and term weighting system [12, 13]. Both of them are described in details as follows.

2.1.3.1.1 Automatic indexing

The indexing task applied to a use case description consists of assigning to each stored item terms, or concepts, capable of representing use case content. Therefore, automatic indexing is a process which usually considered worthwhile to preprocess the text of the documents in the collection to determine the term to be used as index terms because the computer storage of the full text of documents is expensive and is rarely possible except as a by-product of automatic typesetting operations. During automatic indexing process, some text operations can be performed such as elimination of stopwords and stemming (reduction of a word to its grammatical root). The activity diagram of indexing process is depicted in figure 2.6, and details of each step are introduced as follows.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

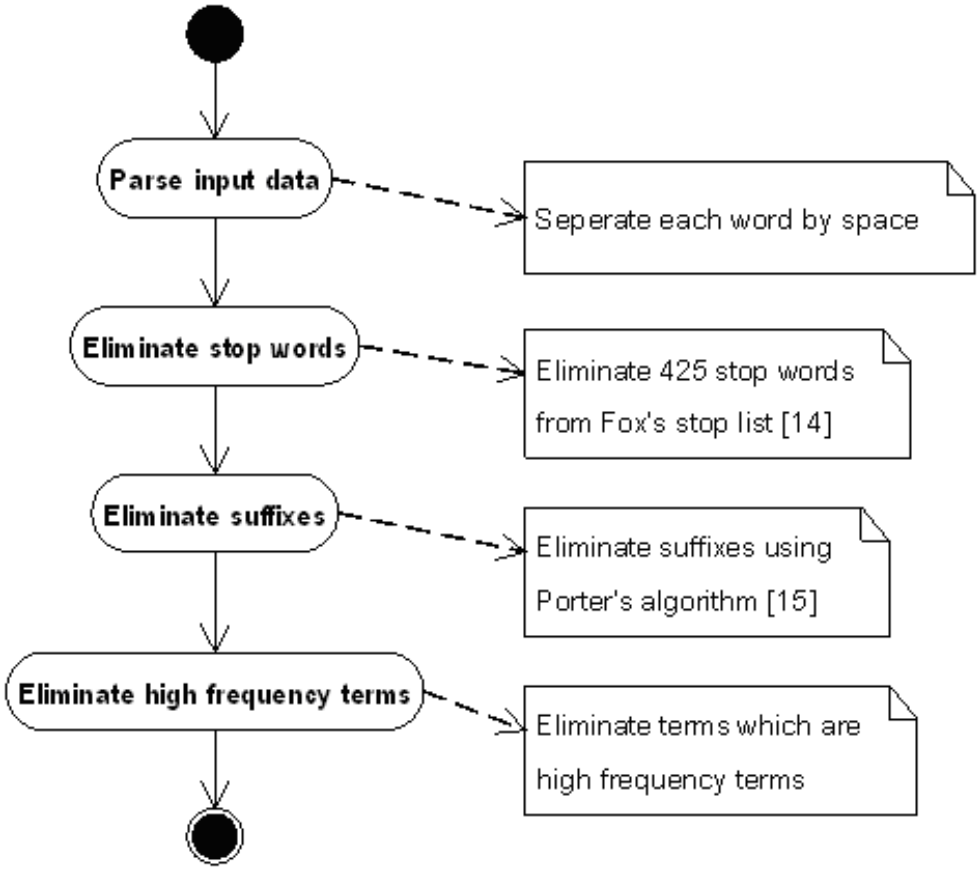


Figure 2.6 The activity diagram of indexing process

1) Parsing information from the use cases is to separate each word from the use cases for preparing them to the next step.
2) Eliminating words from the stop list is removal of stopwords from a stop list because these words are poor discriminators and cannot be possibly used by themselves to identify use case content. In English, about 425 common words are involved, and it is easy to include them in a dictionary or the stop list [14].
3) Stemming words into its grammatical root is to reduce the original words to word stem for reducing a variety of different forms. The generation of word stems, and subsequent identification of common stems, is relatively easy to do in English and serves as a recall enhancing device. One of the models for stemming words which is popular is Porter's algorithm [15].

4) Eliminating high frequency words is the following step from eliminating words from the stop list. The purpose of this step is to eliminate high frequency function words because they are considered as poor discriminators.

2.1.3.1.2 Term weighting system

Term weighting system is the process of assigning each term a weight, or value, reflecting its presumed importance for purposes of content identification. The first and most obvious place where appropriate content identifiers might be found is the text of the use cases themselves. Currently, there are many techniques for term weighting system but a simple technique used in this thesis for the extraction of content terms from documents and a document excerpts and with the assignment of term weights in order of term importance is the inverse document frequency (IDF) weighting system. It is computed as equation (1).

$$W_{ik} = \frac{Freq_{ik}}{TotFreq_k} \qquad (1)$$

Where W_{ik} is a weighted value of term k in document i.

$Freq_{ik}$ is a frequency of term k appearing in document i.

$TotFreq_k$ is a total frequency of term k appearing in the collection.

In IDF weighting system, a composite expression measuring the importance, or weight, of term k in a given document i would increase as the frequency of the term in the use case, $Freq_{ik}$, increases but decrease as the document frequency $TotFreq_k$ increases.

2.1.3.2 Information retrieval process

In retrieval process, the important process is similarity computation. This computation is to find similarity between a set of user's query terms and a set of terms represented documents in the library. The similarity measurement often used is Dice's coefficient [12] defined as equation (2).

$$Similarity(Doc_i, Query_j) = \frac{2\left[\sum_{k=1}^{t} (Term_{ik} \bullet QTerm_{jk})\right]}{\sum_{k=1}^{t} Term_{ik} + \sum_{k=1}^{t} QTerm_{jk}} \quad (2)$$

Where Similarity(Doc_i ,$Query_j$) is a similarity value between document i and query j.

$Term_{ik}$ is the weighted value of term k assigned to the document i.

$QTerm_{jk}$ is the weighted value of term k assigned to the query j.

2.1.3.3 The evaluation process

The purpose of evaluation is to measure the performance of the system. This measure can reflect performance of an information storage and retrieval system. For evaluation process [12, 13], there are 3 metrics applied to this thesis; that is recall, precision, and harmonic mean. Their details are introduced as follows.

1) Recall

Recall is defined as the proportion of relevant documents retrieved. It is given by the equation (3).

$$Recall = \frac{RetRel}{RetRel + NRetRel} \quad (3)$$

Where RetRel is the number of retrieved and relevant documents.

NRetRel is the number of relevant but not retrieved documents.

2) Precision

Precision is the proportion of retrieved documents that is relevant. It is given by the equation (4)

$$Precision = \frac{RetRel}{RetRel + RetNRel} \quad (4)$$

The variable extending from those of recall is RetNRel which is the number retrieved but not relevant documents.

3) Harmonic mean

Harmonic mean is a single measure which combines recall and precision. It assumes values in the interval [0, 1]. It is 0 when no relevant documents have been retrieved and is 1 when all ranked documents are relevant. Further, the harmonic mean F assumes a high value when both recall and precision are high. Therefore, determination of the maximum value for F can be interpreted as an attempt to find the best possible compromise between recall and precision. Harmonic mean is computed as the equation (5).

$$F(j) = \frac{2}{\frac{1}{r(j)} + \frac{1}{p(j)}} \quad (5)$$

Where F(j) is the harmonic mean F.

r(j) is the recall for j-th document in the ranking.

p(j) is the precision for j-th document in the ranking.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2.2 Related works

2.2.1 Reusing use case descriptions for requirements specifications: Towards use case patterns [6]

The purpose of this research is to present a pattern for reusing requirements specification. This requirements specification received from its pattern can be transformed into use case descriptions collecting the details of use cases. This research also presents a template of a use case description for collecting those requirements specification. However, this research can reuse only the pattern for getting requirements specification, but it cannot reuse information of that requirements specification.

Therefore, in this thesis, the approach for reusing requirements specification collected in a new use case description template is proposed. It covers both storage and retrieval process.

2.2.2 Reuse of scenario specifications using an automated relational learner: A lightweight approach [7]

The purpose of this research is to present a technique for reusing use case diagrams using a lightweight approach. The authors of this research develop the tool named “ScenAsst” for reusing use case diagram conveniently. The main processes of ScenAsst are storage and retrieval process. For their storage process, ScenAsst transforms information of use cases into graph format, clusters them, and then storing them into the library. In their retrieval process, user’s query data is transformed into graphs and is compared similarity to graphs collected in the library. Thus, this technique can reuse use cases by considering some structures of them but it is high complexity in comparing similarity process causing time consumption in consequence.

For improving usability and performance in their process, this thesis proposes the approach using textual format as input data instead of graph format.

2.2.3 Reusing UML specifications in a constrained application domain [5]

The purpose of this research is to present a method for reusing UML specifications focusing on use case diagrams. The proposed method applied the information storage and retrieval theory to its process. In storage process, a set of indices from use case descriptions is identified and clustered by the experts while retrieval process, a user has to generate query data in English text and those data is transformed into a set of indices. A set of indices from the library and from the user's data is computed to similarity scores. The highest similarity score of use cases can indicate to the most relevant use cases to the user's query data. This method has the experts generate indices for both storage and retrieval process. Thus, the results are rather precise. On the other hand, the retrieved results from work supported by the experts are depended on their experience, so they may be bias and did not develop a tool for supporting this method automatically.

To reduce those limitations, this thesis proposes automatic processes for reusing use cases in both storage and retrieval process, so it can be supported by a developed tool and the results of them do not depend on the experience of the experts.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

CHAPTER III

THE APPROACH

In this chapter, the approach for retrieving use cases using terms and use case structure similarity computation is described. The details of them are described in this section.

The activities of our approach consist of 3 main processes. They are storage, retrieval and evaluation process. They are depicted as the activity diagram in figure 3.1. In our approach, a tool is developed for supporting the storage and retrieval process, but not in the evaluation process. This tool can help users to store and retrieve use cases automatically. The steps of the overview approach are briefly introduced as follows.

Step 1 : Use case collection and index creation

1) Use cases developed from some example domains are collected.
2) Use cases are transformed into a set of indices and their weighted values. (See details in use case collection and index creation, section 3.1)

Step 2 : Query generation and retrieval process

3) A user's use case query and its weighted values are generated in a predefined use case format.
4) A user's use case query is transformed into a set of indices.
5) A set of use cases from the collection is retrieved according to their similarity with the user's use case query.
6) A set of retrieved use cases is selected and presented to a user. (See details in query generation and retrieval process, section 3.2)

Step 3 : System evaluation

7) The retrieved use cases are evaluated which one is relevant or irrelevant to the user's use case query in order to compute recall, precision, and harmonic mean. (See details in system evaluation, section 3.3)

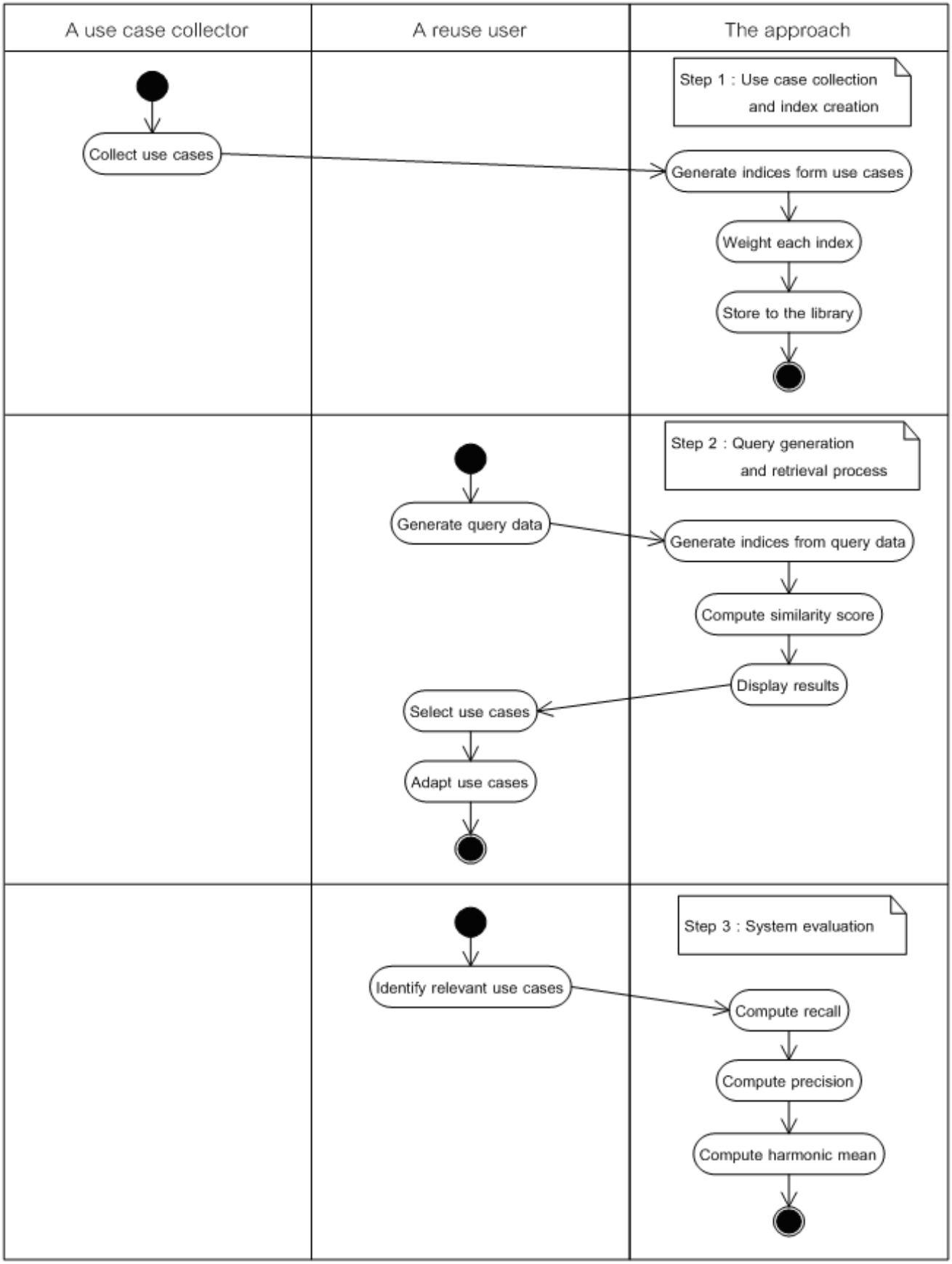


Figure 3.1 The activity diagram of the overview approach

3.1 Use case collection and index creation

The activities of this step are shown as the activity diagram in figure 3.2 and details of them are described in the below figure.

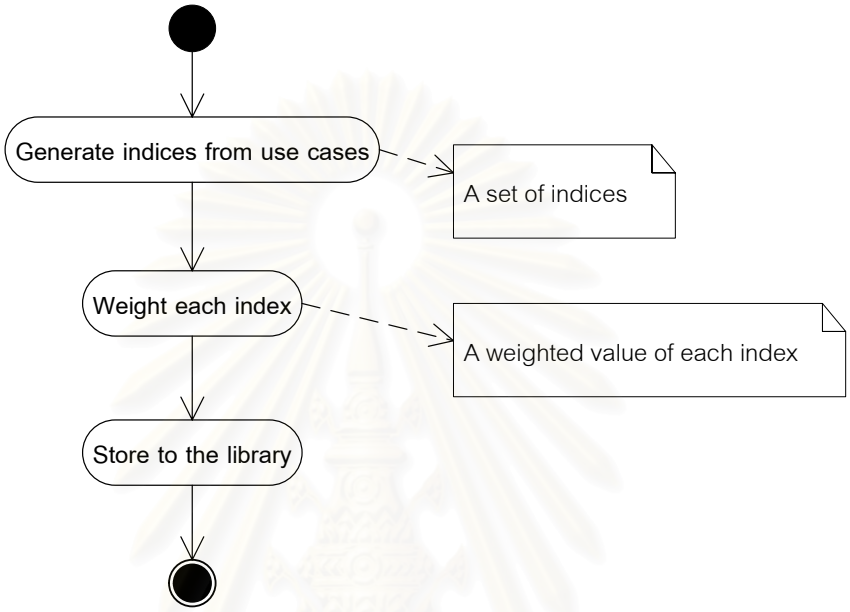


Figure 3.2 The activity diagram of index creation

- After a user collects use cases, natural English language collected in use cases are parsed and transformed to be a set of indices by using automatic indexing process mentioned in section 2.1.3.1.1. The output of this step is a set of indices.
- Each index is weighted by using inverse document frequency (IDF) technique introduced in section 2.1.3.1.2. The weighted value of each index is the output of this step.
- All indices and their weighted values are stored into the database as an inverted file.

3.2 Query generation and retrieval process

The activities of this step are shown as the activity diagram in figure 3.3 and the details of them are described in the below figure.

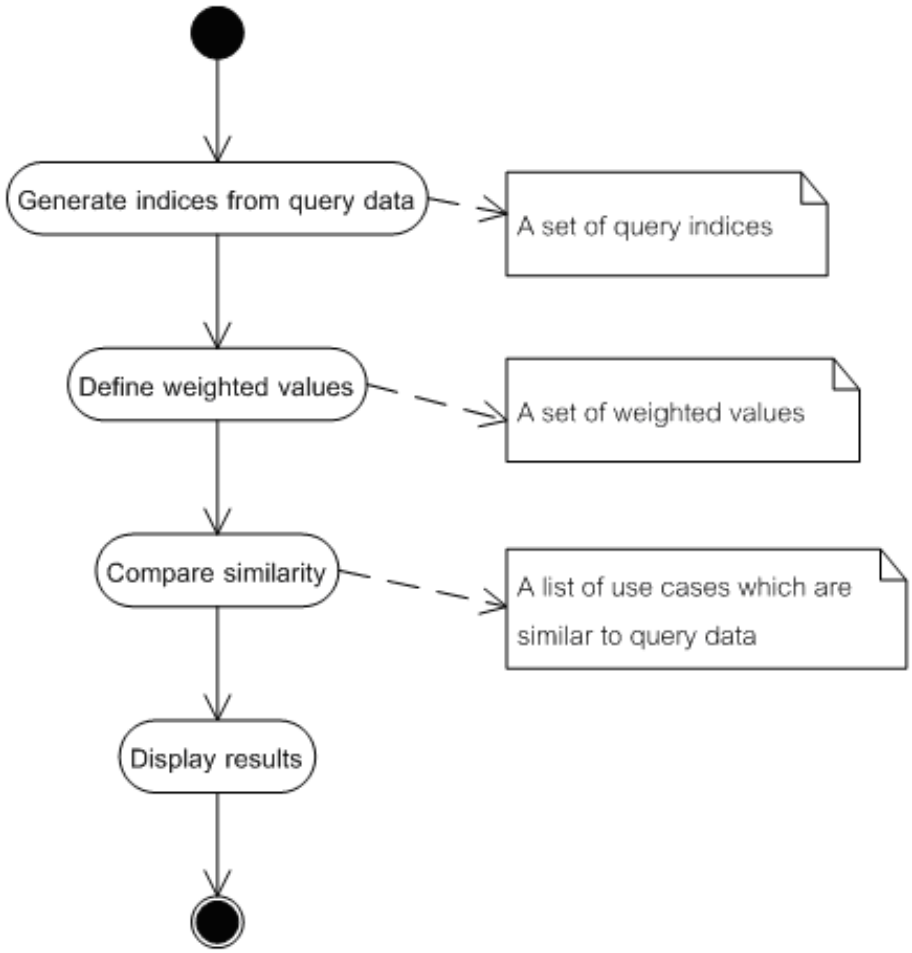


Figure 3.3 The activity diagram of retrieval process

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

- Firstly, a user has to generate a simple use case as query data. The simple use case is composed of a set of keywords. The user's use case query is transformed into a set of indices by automatic indexing process which is the same as those in the index creation step.
- A user defines weighted values of each element of use case query. These weighted values are used as a factor in similarity computation.
- Similarity scores between each element of user's use case query and each element of each use case collected in database is computed by

equation (6). Output of this computation is a set of similarity scores between each element of the user's use case query and each element of each use case in the collection.

$$Similarity(E_{mi}, E_{mj}) = \frac{2\left[\sum_{k=1}^{t}(Term_{mik} \bullet W_{ik}) \bullet Term_{mjk}\right]}{\sum_{k=1}^{t} Term_{mik} + \sum_{k=1}^{t} Term_{mjk}} \qquad (6)$$

Where

E_{mi} is an element m of use case i.

E_{mj} is an element m of use case j.

$Term_{mik}$ is 1 when term k appears in element m of use case i and is 0 when term k does not appear in element m of use case i.

$Term_{mjk}$ is 1 when term k appears in element m of use case j and is 0 when term k does not appear in element m of use case i.

W_{ik} is the weighted value of term k in use case i introduced in section 2.1.3.1.2 in the equation (1).

- Similarity scores between user's use case query and each use case in the collection are computed by equation (7). Output of this computation is a set of similarity scores between user's use case query and each use case in the collection.

$$Similarity(UC_i, Query_j) = \frac{\sum_{m=1}^{e}\left[Similarity(E_{mi}, E_{mj}) \bullet WE_m\right]}{TotalWeight} \qquad (7)$$

Where

$Similarity(UC_i, Query_j)$ is the similarity score between user's use case query and each use case in the library.

$Similarity(E_{mi}, E_{mj})$ is the similarity between each element of use case i and query use case j computed by equation (6).

WE_m is the weighted value of the element m defined by the user.

TotalWeight is the summary values of weighted values (WE_m) of all components.

- The results which are the use cases in the collection have topmost T similarity scores computed from equation (7) are displayed. Where T is the predefined threshold number.

3.3 The system evaluation

The activities of this step are shown as the activity diagram in figure 3.4 and their steps are described as follows.

- In the first step of a system evaluation, a user has to identify all use cases with relevant to his/her use case query in the library. This information is used for computing the values of metrics such as recall, precision, and harmonic mean.
- The values of the metrics are computed by the equation (3), (4), and (5). Details of them were introduced in section 2.1.3.3.

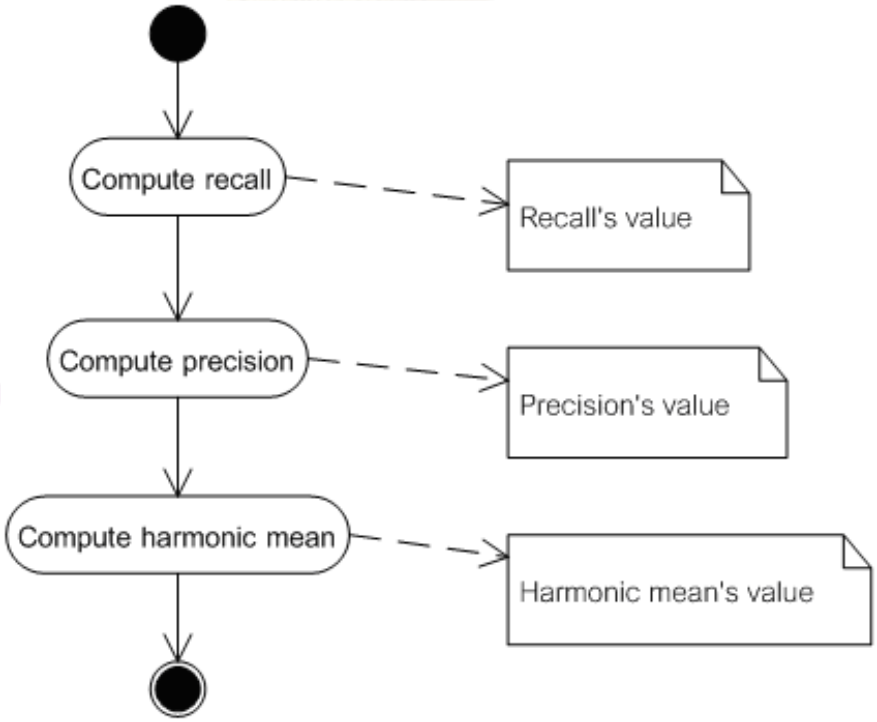


Figure 3.4 The activity diagram of the evaluation process

CHAPTER IV

TOOL DEVELOPMENT

In this chapter, details of tool development for supporting use case retrieval in our approach are described. These details are composed of supporting tools, tool architecture, function of the tool, and data model.

4.1 Supporting tools

4.1.1 Apache Tomcat web server

Web server is a program to operate information or source code of a domain system contained in its server. A client can request information or operation via web browser, so web server can retrieve data collected in database server in order to operate client's requested information or operation. Finally, web server sends client's requested information or operation back to the client.

Apache Tomcat web server is a web server program which is a freeware. It can support operation on many platforms such as UNIX, Linux, or Windows platform, and can support work on hypertext transfer protocol (HTTP). Therefore, this program is selected to be our supporting tool.

4.1.2 MySQL database server

Database server is a program to maintain data of a domain system. Data collected in this server is designed to reduce data redundancy, and to keep data consistency. Thus, some relationships among those data collected in a form of entities or tables are identified. Database server can maintain these relationships of data in a domain.

MySQL database server is a database server program which can manage or maintain data collected in its server. This program is a freeware program, so it is also selected to be our supporting tool.

4.1.3 Java Server Page (JSP)

Java Server Page is a script language which can support web application for a system domain. JSP normally works together with hypertext markup language (HTML). An application domain developed from JSP is collected in a web application server or a web server to be processed when a client requests information or operation.

JSP is one of java technology, so it can work in many platforms. In addition, JSP is also freeware program; therefore, it is selected to be our supporting tool too.

4.2 Tool architecture

The developed tool is a web based application using 3 tiers architecture. These tiers are web browser, web server, and database server. The architecture of developed tool is shown as component diagram in figure 4.1.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

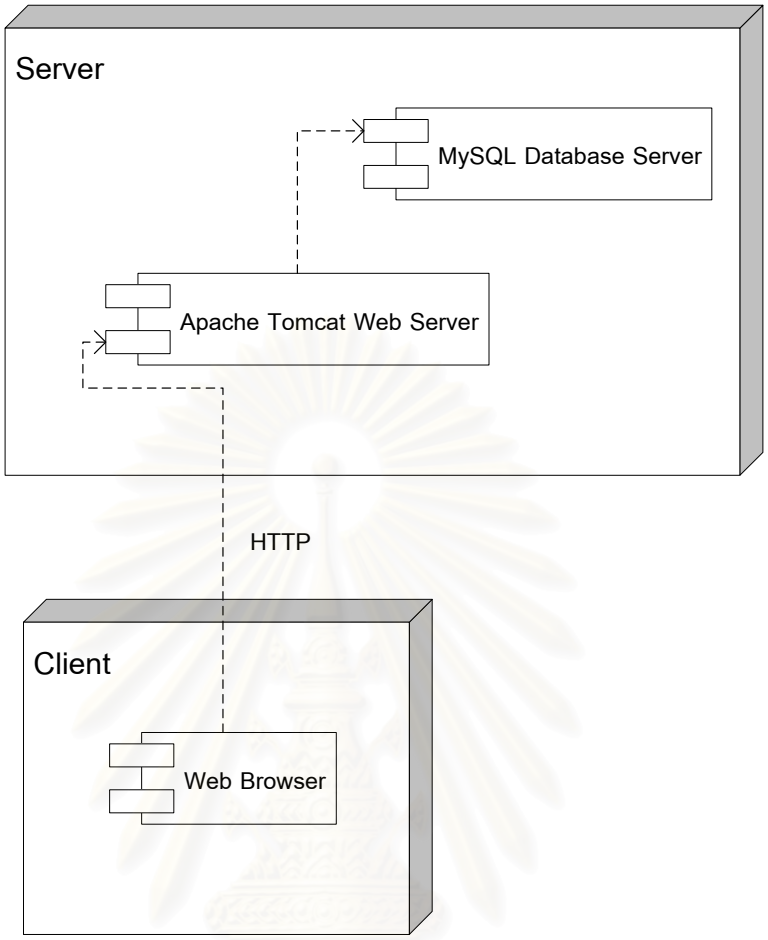


Figure 4.1 The component diagram of developed tool architecture

4.3 Function of the tool

The developed tool has 2 main functions which are use case storage and retrieval. In use case retrieval, this function can be divided to 2 subfunction; there are query use cases by use case keywords, and query use cases by use case structure. The menu screen of use case retrieval system is illustrated in figure 4.2.

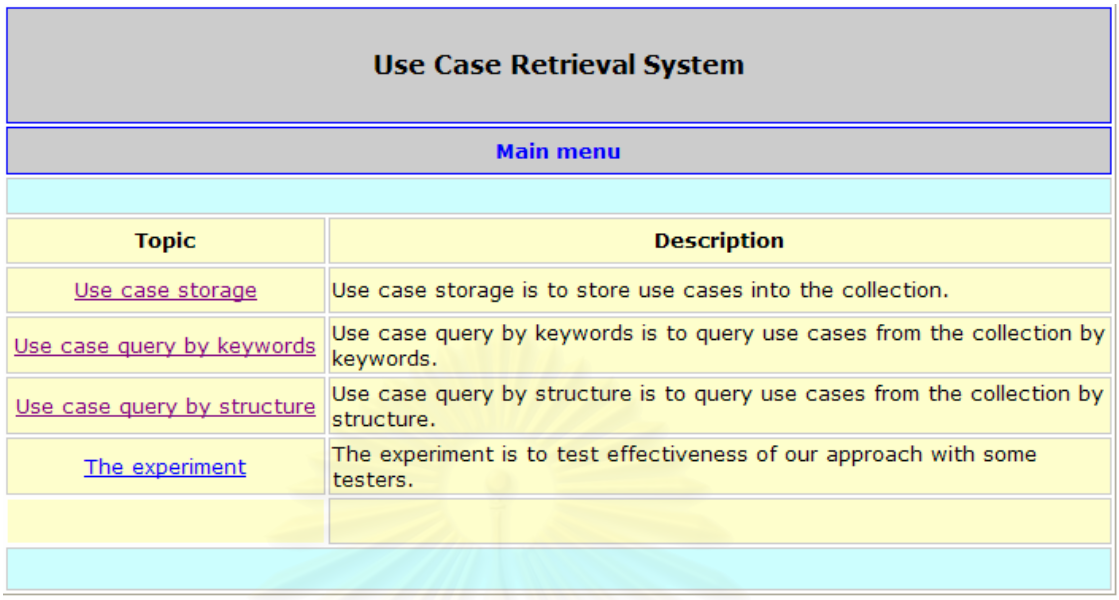


Figure 4.2 The menu screen of use case retrieval system

The menu screen has 4 topics to be selected. The description of each topic is briefly introduced as follows.

- **Use case storage** is for a use case collector to store use cases into the collection. Details of this function are described in section 4.3.1.
- **Use case query by keywords** is for a user who wants to retrieve use cases from the collection to retrieve them by keywords. Details of this function are described in section 4.3.2.
- **Use case query by use case structure** is for a user who wants to retrieve use cases from the collection to retrieve them by use case structure. Details of this function are described in section 4.3.3.
- **The experiment** is for a tester to test effectiveness of our use case retrieval system. Details of our experiment are described in detail in chapter 5.

4.3.1 Use case storage

Use case storage is a screen for a use case collector to collect use cases into the collection. A use case collector can insert use case information in the "Store Use Case" screen. This screen is shown in figure 4.3

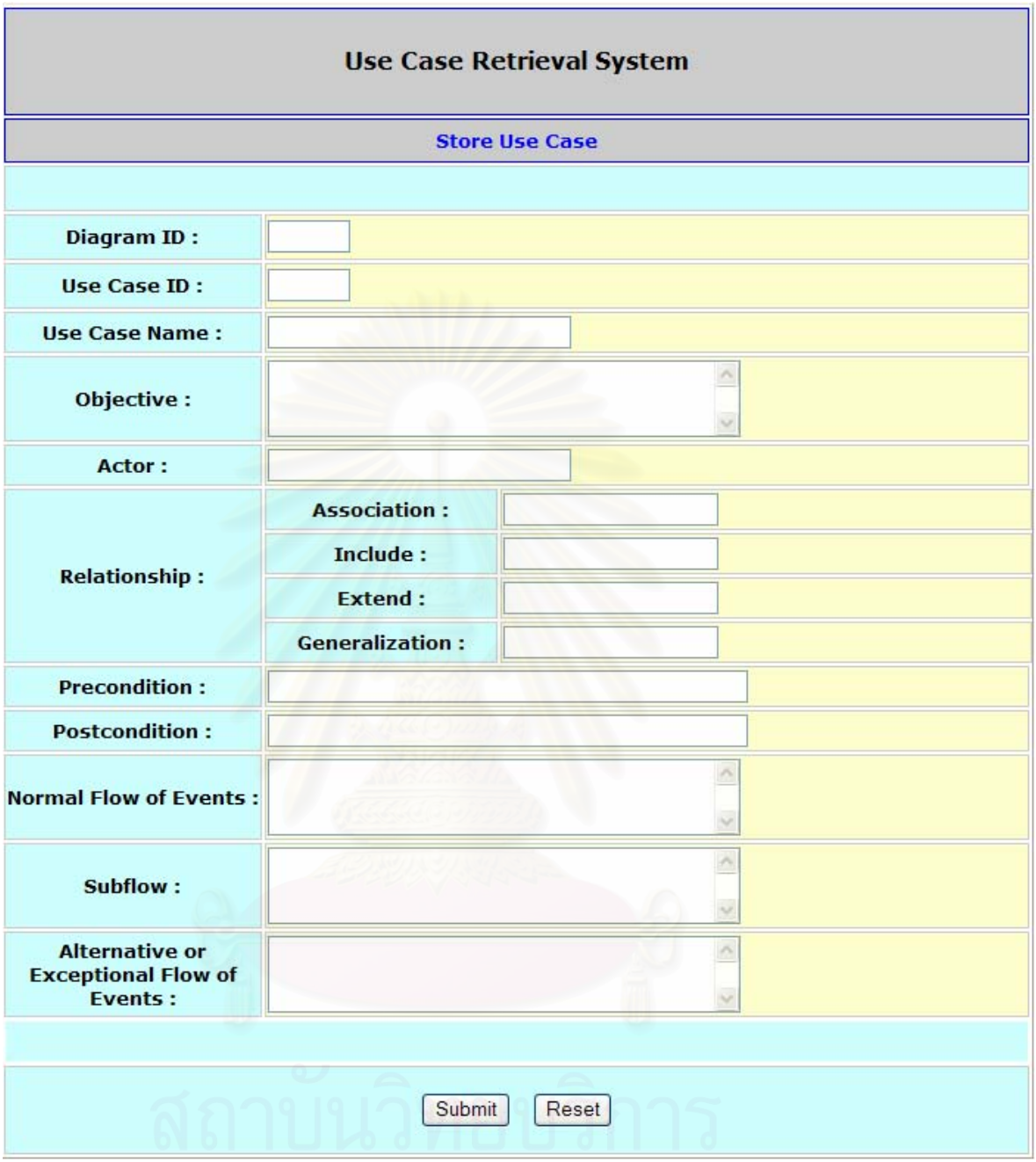


Figure 4.3 The use case storage screen of use case retrieval system

The input text area behind “Diagram ID” is an area for inserting an ID of a use case diagram or a use case domain, and the input text area behind “Use Case ID” is an area for inserting ID of a use case description in the use case domain. For other input text areas, they are areas for a use case collector to insert use case information in each component of use case. When the use case collector selects “Submit” button, the inserted use case information is transformed into a set of terms or indices by automatic indexing process described in section 2.1.3.1.1,

and then collected in the database. On the other hand, the use case collector can clear all fields in the screen by selecting “Reset” button.

4.3.2 Use case retrieval by use case keywords

Use case retrieval by use case keywords is a screen for a user who wants to query or retrieve use cases by keywords. The user has to generate keywords which he/she thinks they are relevant to his/her wanted use case description or use case domain. The use case retrieval by use case keywords screen is shown in figure 4.4.

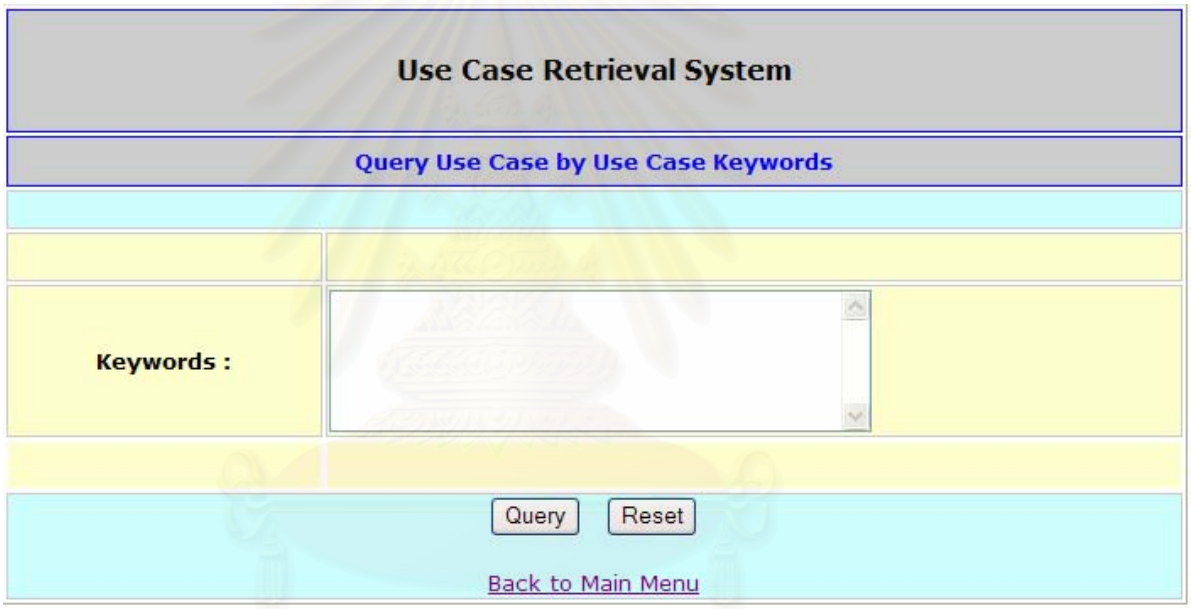


Figure 4.4 The use case retrieval by use case keywords screen of use case retrieval system

The user can generate just one or more keywords. When the user selects “Query” button, the user’s generated keywords are transformed into a set of terms or indices, and then compared similarity with each use case in the collection by Dice’s coefficient in equation (6), described in section 3.2. The results of query are shown in the next section. On the other hand, the user can clear the keyword field in the screen by selecting “Reset” button.

4.3.3 Use case retrieval by use case structure

Use case retrieval by use case structure is a screen for a user who wants to query or retrieve use cases by use case structure. The user has to generate keywords in each use case component and identify weighted value of each use case component. The range of weighted value is the integers from 1 to 5, where 1 is the minimum weighted value, and 5 is the maximum weighted value. The use case retrieval by use case structure screen is shown in figure 4.5.

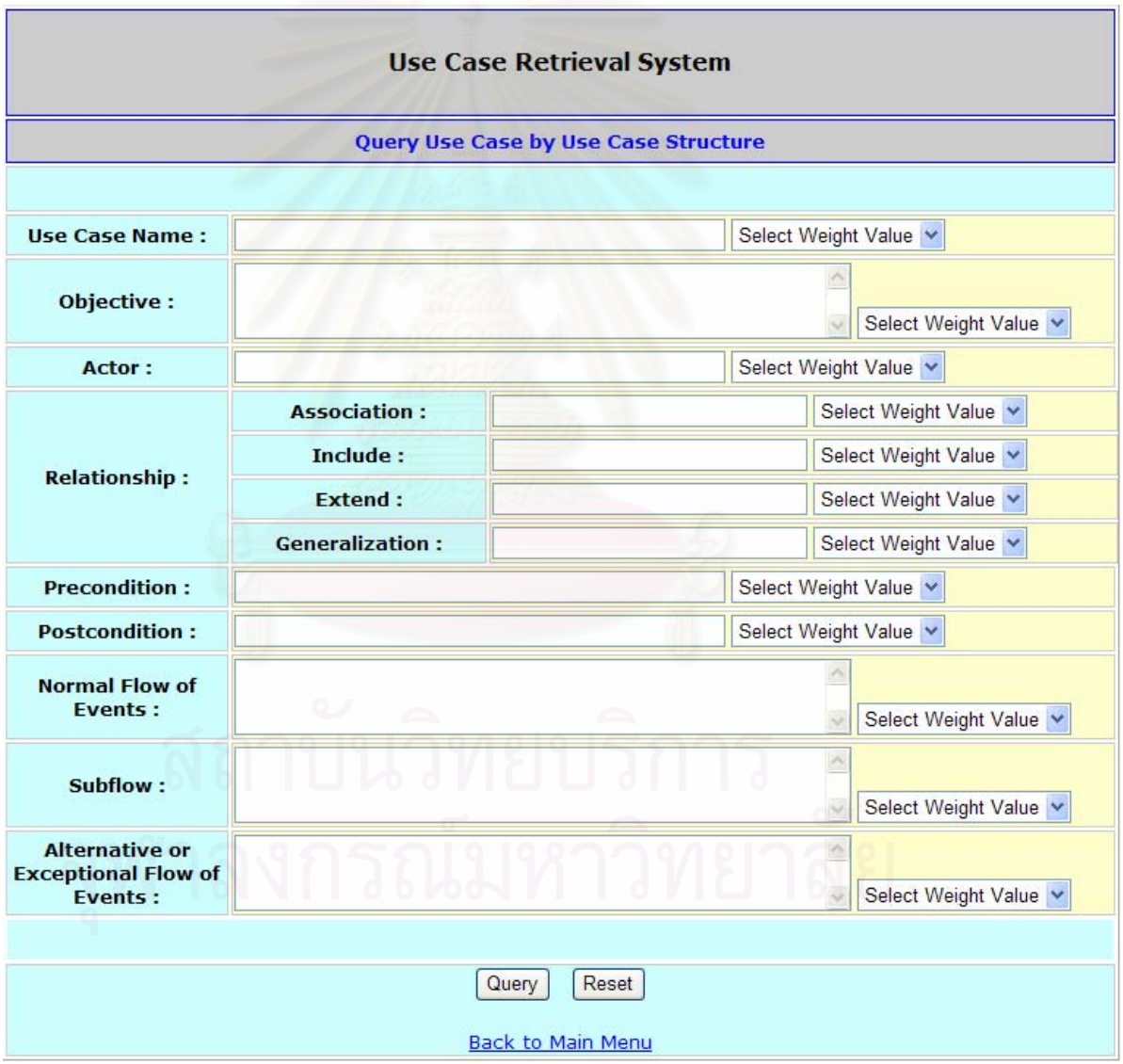


Figure 4.5 The use case retrieval by use case structure screen of use case retrieval system

The user can generate no, just one, or many keywords in each component, and select weighted value in the component he/she generated keywords. When the user selects “Query” button, the user’s generated keywords are transformed into a set of terms or indices, and then compared similarity with each use case in the collection by the equation (7), described in section 3.2. The results of query are shown in the next section. On the other hand, the user can clear the keyword field in the screen by selecting “Reset” button.

The result of use case retrieval screen is shown in figure 4.6. The system displays only first 10 use cases which their similarity value are more than a predefined threshold value.

Use Case Retrieval System

Display Results

Display 9 of 10 Results

No	ID	Use Case Name	Use Case Domain
1	0209	Save subject details	Departmental Information System for Curriculum and Course Offeings Management
2	0211	Store subject report	Departmental Information System for Curriculum and Course Offeings Management
3	0216	View subject operation report	Departmental Information System for Curriculum and Course Offeings Management
4	0213	View subject operation details	Departmental Information System for Curriculum and Course Offeings Management
5	0214	View subject operation status	Departmental Information System for Curriculum and Course Offeings Management
6	0210	Save subject operation status	Departmental Information System for Curriculum and Course Offeings Management
7	0221	Save subject information	Departmental Information System for Curriculum and Course Offeings Management
8	0223	View subject information	Departmental Information System for Curriculum and Course Offeings Management
9	0212	Save subject operation result	Departmental Information System for Curriculum and Course Offeings Management
10			

Back to Main Menu

Figure 4.6 The use case retrieval result screen of use case retrieval system

The “No” area in the screen is the order of use cases which was retrieved. The “ID” is an ID of use cases. “Use Case Name” is the name of retrieved use cases, and “Use Case Domain” is the name of use case domain of use cases which was retrieved.

4.4 Data model

The use cases collected in use case retrieval system is transformed and collected in relational database management system (RDBMS). The information collected in this database is not only use case information, but also experiment information such as tester information, query information, and result information. All information collected in this relational database management system is designed as the entity relationship diagram depicted in figure 4.7.

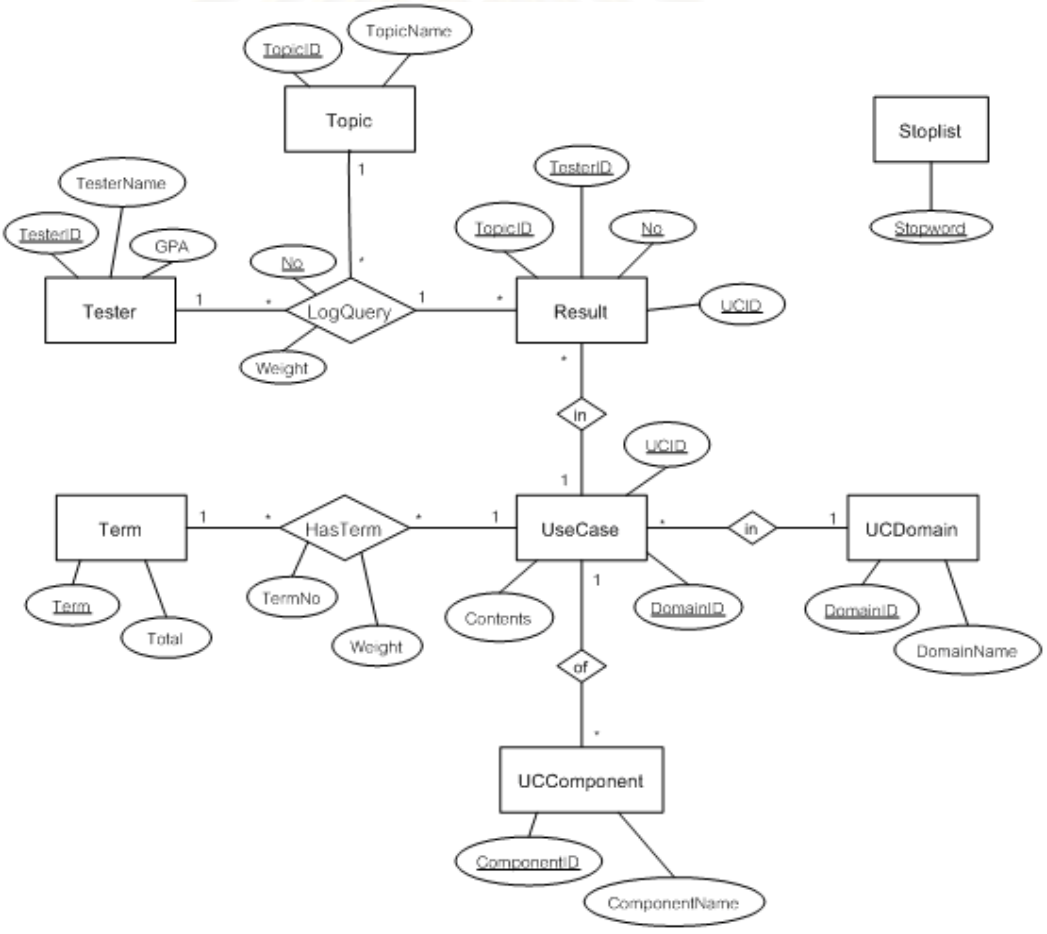


Figure 4.7 The entity relationship diagram of use case retrieval system

All information collected in this relational database management system can be separated into 12 tables. Details of each table are briefly introduced in table 4.1 to 4.11 as follows.

Table 4.1 Use case details

Table name : UseCase		
Table description : This table is to collect use case details		
Name	**Type**	**Description**
UCID	Char (4)	ID of a use case
UCDomainID	Char (2)	ID of a use case domain
Name	Text	Name of use case
Objective	Text	Objective of use case
Actor	Text	Actor of use case
Association	Text	Association relationship of use case
Include	Text	Include relationship of use case
Extend	Text	Extend relationship of use case
Generalization	Text	Generalization relationship of use case
Normalflow	Text	Normal flow of use case
Subflow	Text	Subflow of use case
Alternativeflow	Text	Alternative flow of use case

Table 4.2 Use case domain

Table name : UCDomain		
Table description : This table is to collect use case domain		
Name	**Type**	**Description**
DomainID	Char (2)	ID of a use case domain
DomainName	Char (50)	Name of a use case domain

Table 4.3 Use case component

Table name : UCComponent		
Table description : This table is to collect each element of use case components		
Name	**Type**	**Description**
ComponentID	Char (2)	ID of a use case component
ComponentName	Char (50)	Name of a use case component

Table 4.4 All terms in the collection

Table name : Term		
Table description : This table is to collect all terms and amount of them in the collection		
Name	**Type**	**Description**
Term	Char (20)	Term
Total	Integer	The number of terms in the collection

Table 4.5 All terms in each use case

Table name : HasTerm		
Table description : This table is to collect all terms and amount of them in each use case		
Name	**Type**	**Description**
Term	Char (20)	Term
DomainID	Char (2)	ID of a use case domain
UCID	Char (4)	ID of a use case
ComponentID	Char (2)	ID of a use case component
TermNo	Integer	The number of term appearing in a use case
Weight	Float	Weighted value of term

Table 4.6 Stop list

Table name : Stoplist		
Table description : This table is to collect all stop words from a stop list		
Name	**Type**	**Description**
Stopword	Char (20)	Stop words in a stop list

Table 4.7 Tester

Table name : Tester		
Table description : This table is to collect all testers in the experiments		
Name	**Type**	**Description**
TesterID	Char (2)	ID of tester
TesterName	Char (50)	Name of tester
GPA	Char (4)	Average grade of tester

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table 4.8 Logged data of tester's query by use case keywords

Table name : LogQuery1		
Table description : This table is to collect all terms in query by use case keywords of a tester in the experiments		
Name	**Type**	**Description**
TesterID	Char (2)	ID of tester
Topic	Char (1)	Topic number
No	Char (1)	Number of query
Term	Char (20)	Query term

Table 4.9 Logged data of tester's query by use case structure

Table name : LogQuery2		
Table description : This table is to collect all terms in query by use case structure of a tester in the experiments.		
Name	**Type**	**Description**
TesterID	Char (2)	ID of tester
Topic	Char (1)	Topic number
No	Char (1)	Number of query
ComponentID	Char (2)	ID of use case component
Weight	Char (1)	Weighted value of each component
Term	Char (20)	Query term

Table 4.10 Topic

Table name : Topic		
Table description : This table is to collect all topics in the experiments		
Name	**Type**	**Description**
TopicID	Char (1)	ID of topic
TopicName	Char (100)	Name of topic

Table 4.11 Result of use case query

Table name : Result		
Table description : This table is to collect all results of use case query in the experiments.		
Name	**Type**	**Description**
TesterID	Char (2)	ID of tester
Topic	Char (1)	Topic number
No	Char (1)	Number of query
QueryType	Char(1)	Type of query (keyword or structure)
UCID	Char (4)	ID of use case which was retrieved

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

CHAPTER V

EXPERIMENTS

In this chapter, details of our experiments are described. These experiments are to test our approach with the developed tool. Details of the experiments including objective, method, procedure, result, result analysis, and discussion of the experiments are described as follows.

5.1 Experimental objective

Objective of our experiments is to test our research assumption that whether use case retrieval by use case structure is more effective than use case retrieval by use case keywords. Therefore, the experiments are provided in order to compare the effectiveness of our approach, use case retrieval by use case structure, and the effectiveness of a general approach, use case retrieval by use case keywords. 3 selected metrics which are recall, precision, and harmonic mean, mentioned in section 2.1.3.3 are used in effectiveness measurement in the experiments.

5.2 Experimental method

This experimental method is designed for eliminating bias in our experiments, so it has many controlled factors. These factors are separated into 4 groups, use cases, testers, subjects, and queries. Their details are explained as follows.

5.2.1 Use cases

315 use cases from 16 use case domains are selected from practicable system for our experiments, and they were written in English language. Their contents are introduced in the appendix A.

5.2.2 Testers

Our experiments has 10 testers for testing our use case retrieval system. For all testers, everyone has studied in a Master's degree program in software

engineering field of study. Therefore, they have had experience and capability about use case modeling, and their knowledge about English language is quite fair because English score of all testers is about 500 for TOEFL equivalent score. As a result, all testers have ability to test these experiments.

5.2.3 Subjects

These experiments have 5 sample subjects for testers to test the use case retrieval system. The testers have to query use cases from our 16 use case domains by generating keywords which are relevant to 5 given sample subjects. All 5 sample subjects used in the experiments are

A) Teaching-studying system
B) Product trading system
C) Customer's information management
D) Financial calculation
E) Report generation

The answer set of use cases for each subject is identified and collected in the database in order to compute recall, precision, and harmonic mean later. The objective of our 5 sample subjects is to compare effectiveness of both retrieval systems, query use case by use case structure and by use case keywords in 2 points of view. There are view of functional requirements and a system domain. The reason of defining sample subjects in 2 points of view is because we want to observe broad and narrow subjects whether they have an effect on effectiveness of use case retrieval in our approach. Subject A and B are in view of system domain because they are not identified functional requirements or main functions, but they just indicate their wanted system domains. Meanwhile, subject C and D are in views of both functional requirements and system domain because they are identified both functional requirements and their system domain. Therefore, scope of subject C and D are narrower than subject A and B. The last subject, E is in view of only functional requirements because it is identified functional requirements or main function of the system, but the system domain is not

identified. The reason of defining 5 sample subjects in 2 introduced main points of view is variety of query which may eliminate tester's bias about their points of view.

5.2.4 Queries

The testers have to query both by use case keywords and by use case structure for all 5 subjects, A, B, C, D, and E. For each subject, the tester can query 5 times. Therefore, each tester has to generate 50 set of queries, 25 sets for query by use case keywords, and another 25 sets for query by use case structure. Finally, the experiments get 500 sets of queries for all testers, 250 sets for query by use case keywords, and another 250 sets for query by use case structure. In other words, the experiments get 100 sets of queries for each subject, 50 sets from query by use case keywords, and another 50 sets from query by use case structure. These sets of queries are used to calculate recall, precision, and harmonic mean later.

For easier understanding, the number of all queries for each tester is depicted in table 5.1. The values in the table are the number of queries which each tester has to generate for each subject, A, B, C, D, and E.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table 5.1 The number of all queries for each tester

Tester	The number of all queries for each tester											
	Query by use case keywords. Subject						Query by use case structure. Subject					
	A	B	C	D	E	Summary	A	B	C	D	E	Summary
1	5	5	5	5	5	25	5	5	5	5	5	25
2	5	5	5	5	5	25	5	5	5	5	5	25
3	5	5	5	5	5	25	5	5	5	5	5	25
4	5	5	5	5	5	25	5	5	5	5	5	25
5	5	5	5	5	5	25	5	5	5	5	5	25
6	5	5	5	5	5	25	5	5	5	5	5	25
7	5	5	5	5	5	25	5	5	5	5	5	25
8	5	5	5	5	5	25	5	5	5	5	5	25
9	5	5	5	5	5	25	5	5	5	5	5	25
10	5	5	5	5	5	25	5	5	5	5	5	25
Summary	50	50	50	50	50	250	50	50	50	50	50	250

The activity diagram of the experimental method is shown in figure 5.1. After use cases were stored, a tester can query use cases by the controlled factors mentioned previously. For each subject, all 50 sets of queries from query by use case keywords and all 50 sets of queries from query by use case structure are used to compare effectiveness of retrieval. The retrieved results of those queries are used to calculate recall, precision, and harmonic mean for comparing effectiveness of both retrievals in the experiments.

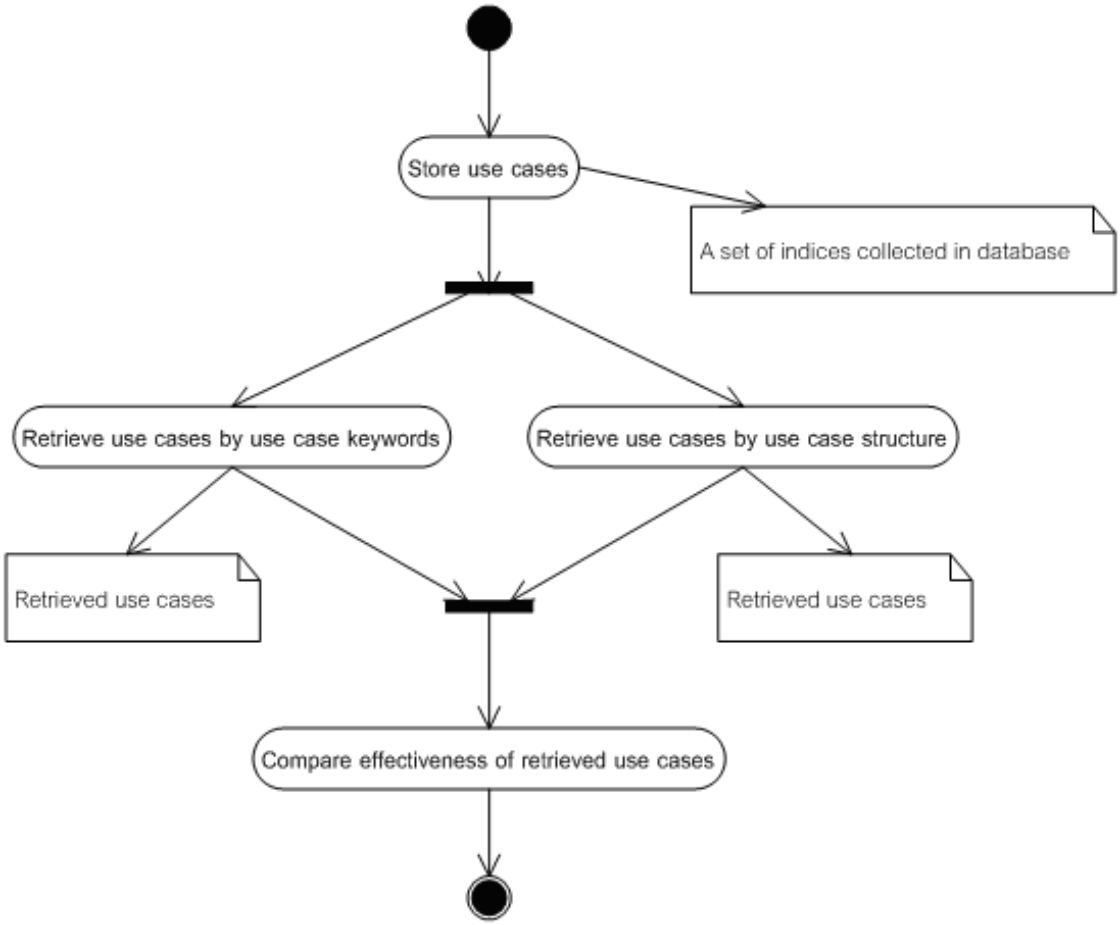


Figure 5.1 The activity diagram of the experimental method

5.3 Experimental procedure

The experimental procedure of the tester is shown as the activity diagram in figure 5.2. The experimental procedure is separated into 4 steps as follows.

1) The tester has to insert his/her personal profile which is tester's ID, first name, last name, and GPA to the developed tool.
2) The tester learns about 16 use case domains collected in our collection. The brief description of each domain is shown to the tester by graphical user interface of the developed tool.
3) The tester queries use cases by use case keywords for all 5 subjects, A, B, C, D, and E, and queries 5 times for each subject using the developed tool.
4) The tester queries use cases by use case structure for the same 5 subjects, A, B, C, D, E, and queries 5 times for each subject using the developed tool too.

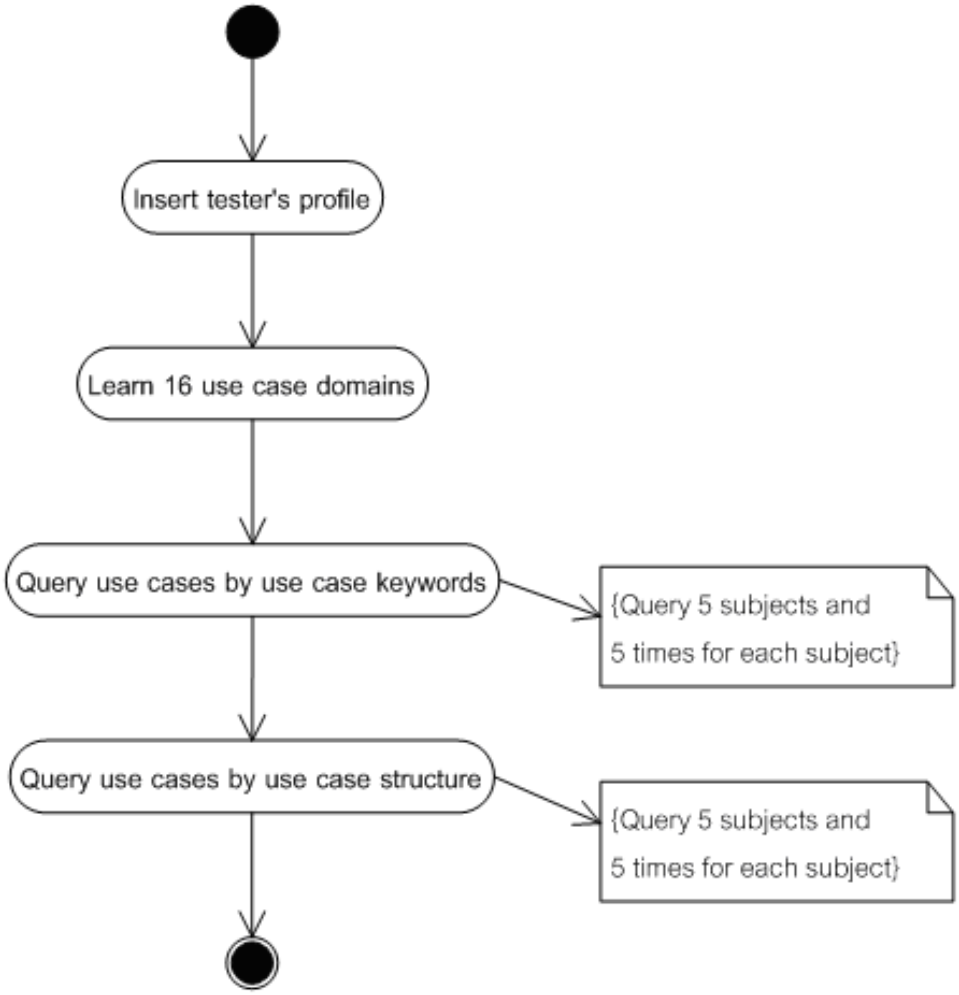


Figure 5.2 The activity diagram of the experimental procedure of the tester

5.4 Experimental environment

5.4.1 Similarity threshold

The similarity threshold of our approach for our experiments was defined as 0.00039. This threshold value is computed from

$$\text{mean} - \alpha$$

Where mean is average value of similarity score for all queries in the experiments (0.00077).

α is standard deviation value of all queries in the experiments (0.00038).

The reason of mean – α is because size of use case collection used in the experiments is small. The use case collection is composed of 315 use cases from

16 use case domains. Therefore, some use cases may fall out if similarity threshold was defined higher such as mean or mean + α. However, if the use case collection is larger, the similarity threshold can be adjusted higher.

5.4.2 High frequency threshold

In our experiments, high frequency terms are not eliminated in our experiments. The reason is small size of use case collection used in the experiments. The number of all terms in the collection is 651. Thus, some significant terms may be lost if high frequency threshold was defined higher. However, if the use case collection is larger, the high frequency threshold can be adjusted higher.

5.5 Experimental results

The summary results of our experiments are shown in table 5.2 to 5.5. Results in the tables are summarized and compared between use case query by use case structure and use case query by keywords. For use case query by keywords, 2 characteristics of query by use case keywords are tested in the experiments; there are

(1) Use case query using keywords generated by testers.

(2) Use case query using keywords generated from all keywords appeared in all components of query by use case structure.

Table 5.2 shows average and maximum similarity score of query by use case structure and keywords (1) for each subject, A, B, C, D, and E, and table 5.3 shows average recall, precision, and harmonic mean of query by use case structure and keywords (1) for each subject like table 5.2. Meanwhile table 5.4, average and maximum similarity score of query by use case structure and keywords (2) for each subject, A, B, C, D, and E, are shown, and table 5.5 shows average recall, precision, and harmonic mean of query by use case structure and keywords (2) for each subject like table 5.4.

The abbreviations in table 5.3 and 5.5 are introduced as follows.

- R is the recall.
- P is the precision.

- H is the harmonic mean.

Results in table 5.2, 5.3, 5.4, and 5.5 are summary results. Details of them are shown in the appendix D.

Table 5.2 Average and maximum similarity score of query by use case structure and keywords (1) for each subject

Average and maximum similarity score of query for each subject						
Query by use case keywords (1)				Query by use case structure		
Subject	Average	Maximum		Subject	Average	Maximum
A	0.00041	0.02381		A	0.00123	0.05556
B	0.00061	0.02893		B	0.00114	0.06586
C	0.00058	0.02364		C	0.00086	0.04001
D	0.00046	0.03587		D	0.00098	0.06432
E	0.00062	0.03140		E	0.00112	0.04724
Average	**0.00054**	**0.02873**		**Average**	**0.00107**	**0.05460**

The results indicate that average and maximum similarity scores of use case query by use case structure is more than use case query by use case keywords (1) in average.

Table 5.3 Average recall, precision, and harmonic mean of query by use case structure and keywords (1) for each subject

Average recall, precision, and harmonic mean for each subject									
Query by use case keywords (1)					Query by use case structure				Precision
Subject	R	P	H		Subject	R	P	H	Improvement
A	0.2290	0.4791	0.2725		A	0.4075	0.6945	0.4781	**44.96%**
B	0.2789	0.1863	0.2032		B	0.3180	0.2559	0.2554	**37.36%**
C	0.6054	0.1162	0.1887		C	0.7429	0.2785	0.3519	**139.67%**
D	0.3490	0.3073	0.2863		D	0.4615	0.4956	0.4344	**61.28%**
E	0.6125	0.1831	0.2717		E	0.5438	0.2572	0.3237	**35.01%**
Average	**0.4149**	**0.2544**	**0.2445**		**Average**	**0.4947**	**0.3963**	**0.3687**	**55.78%**

From table 5.3, recall, precision, and harmonic mean indicate that recall, precision and harmonic mean of use case query by use case structure is higher than they of use case query by keywords (1) for almost all sample subjects. The average recall, precision, and harmonic mean of all subjects also indicate that recall, precision and harmonic mean of use case query by use case structure is higher than they of use case query by keywords (1). According to precision value, the precision improvement is in a high rate for all sample subjects. Finally, the precision improvement of this experiment is concluded to be 55.78%

Table 5.4 Average and maximum similarity score of query by use case structure and keywords (2) for each subject

Average similarity score of query for each subject						
Query by use case keywords (2)				Query by use case structure		
Subject	Average	Maximum		Subject	Average	Maximum
A	0.00069	0.02605		A	0.00123	0.05556
B	0.00075	0.02634		B	0.00114	0.06586
C	0.00069	0.02084		C	0.00086	0.04001
D	0.00066	0.03341		D	0.00098	0.06432
E	0.00077	0.03134		E	0.00112	0.04724
Average	**0.00071**	**0.02760**		**Average**	**0.00107**	**0.05460**

The results indicate that average and maximum similarity scores of use case query by use case structure is more than use case query by use case keywords (2) in average.

Table 5.5 Average recall, precision, and harmonic mean of query by use case structure and keywords (2) for each subject

Average recall, precision, and harmonic mean for each subject									
Query by use case keywords (2)					Query by use case structure				Precision
Subject	R	P	H		Subject	R	P	H	Improvement
A	0.5147	0.5565	0.4915		A	0.4075	0.6945	0.4781	**24.80%**
B	0.4922	0.2168	0.2895		B	0.3180	0.2559	0.2554	**18.04%**
C	0.7232	0.1094	0.1859		C	0.7429	0.2785	0.3519	**154.57%**
D	0.5865	0.3322	0.3937		D	0.4615	0.4956	0.4344	**49.19%**
E	0.7125	0.1590	0.2525		E	0.5438	0.2572	0.3237	**61.76%**
Average	**0.6058**	**0.2748**	**0.3226**		**Average**	**0.4947**	**0.3963**	**0.3687**	**44.21%**

From table 5.5, recall, precision, and harmonic mean indicate that precision of use case query by use case structure is higher than they of use case query by keywords (2) for all sample subjects, while recall and harmonic mean are fluctuated. The average recall, precision, and harmonic mean of all subjects also indicate that precision and harmonic mean of use case query by use case structure is higher than they of use case query by keywords (2), but recall is not. According to precision value, the precision improvement is in a high rate for almost all sample subjects. Finally, the precision improvement of this experiment is concluded to be 44.21%

5.6 Experimental result analysis

According to the experimental results, statistical analysis is used for supporting reliability of it. z-test for hypothesis testing [16] is applied to our experimental results. The reason of z-test is because query sets in our experiments are more than 30 queries.

The hypothesis of the experiments is that use case retrieval by use case structure is more effective than use case retrieval by use case keywords. Three selected metrics which are recall, precision, and harmonic mean are used to measure effectiveness of all use case retrieval systems. Therefore, H_0 and H_1 is defined as follows

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_1 : \mu_1 - \mu_2 > 0$$

Where, μ_1 is the average value of recall, precision, and harmonic mean of use case retrieval by use case structure.

μ_2 is the average value of recall, precision, and harmonic mean of use case retrieval by keywords (1) and (2).

Next, the significance value is defined to be 0.05, so confidence value is 0.95. Therefore, the hypothesis H_0 will be rejected when $z > z_{.95}$, and $z_{.95}$ is 1.645 referred from z table [16]. Our hypothesis H_1 will be accepted when z values of recall, precision, and harmonic mean are more than 1.645. According to table 5.3 and 5.5, summary results of z-test for the experiments are shown in table 5.6 through 5.19. z-test (1) in table 5.6 through 5.12 are z-test of average recall, precision, and harmonic mean of use case query by use case structure and use case query by keywords (1), and z-test (2) in table 5.3 through 5.19 are z-test of average recall, precision, and harmonic mean of use case query by use case structure and use case query by keywords (2).

Table 5.6 z-test (1) of recall, precision, and harmonic mean for subject A

z-test of recall, precision, and harmonic mean of subject A			
Metrics	**z**	**$z > z_{.95}$**	**H_1**
Recall	3.6046	3.6046 > 1.645	**Accept**
Precision	3.6075	3.6075 > 1.645	**Accept**
Harmonic mean	4.2498	4.2498 > 1.645	**Accept**

Table 5.7 z-test (1) of recall, precision, and harmonic mean for subject B

z-test of recall, precision, and harmonic mean of subject B			
Metrics	**z**	**$z > z_{.95}$**	**H_1**
Recall	1.1512	1.1512 < 1.645	**Reject**
Precision	2.3498	2.3498 > 1.645	**Accept**
Harmonic mean	2.3926	2.3926 > 1.645	**Accept**

Table 5.8 z-test (1) of recall, precision, and harmonic mean for subject C

z-test of recall, precision, and harmonic mean of subject C			
Metrics	z	$z > z_{.95}$	H_1
Recall	2.2765	2.2765 > 1.645	Accept
Precision	3.6218	3.6218 > 1.645	Accept
Harmonic mean	4.2025	4.2025 > 1.645	Accept

Table 5.9 z-test (1) of recall, precision, and harmonic mean for subject D

z-test of recall, precision, and harmonic mean of subject D			
Metrics	z	$z > z_{.95}$	H_1
Recall	1.9698	1.9698 > 1.645	Accept
Precision	3.2639	3.2639 > 1.645	Accept
Harmonic mean	2.9908	2.9908 > 1.645	Accept

Table 5.10 z-test (1) of recall, precision, and harmonic mean for subject E

z-test of recall, precision, and harmonic mean of subject E			
Metrics	z	$z > z_{.95}$	H_1
Recall	-1.7343	-1.7343 < 1.645	Reject
Precision	3.1068	3.1068 > 1.645	Accept
Harmonic mean	2.0910	2.0910 > 1.645	Accept

Table 5.11 z-test (1) of recall, precision, and harmonic mean for all subjects

z-test of recall, precision, and harmonic mean of all subjects			
Metrics	z	$z > z_{.95}$	H_1
Recall	1.3834	1.3834 < 1.645	Reject
Precision	2.5076	2.5076 > 1.645	Accept
Harmonic mean	3.0594	3.0594 > 1.645	Accept

According to z-test in table 5.6 through 5.11, z-test results can be summarized in table 5.12, z-test (1) summary of the experiment.

Table 5.12 z-test (1) summary of the experiment

Z-test summary of the experiment						
Metrics	Topic A	Topic B	Topic C	Topic D	Topic E	% Accept H_1
Recall	Accept	Reject	Accept	Accept	Reject	**60%**
Precision	Accept	Accept	Accept	Accept	Accept	**100%**
Harmonic mean	Accept	Accept	Accept	Accept	Accept	**100%**

Table 5.13 z-test (2) of recall, precision, and harmonic mean for subject A

z-test of recall, precision, and harmonic mean of subject A			
Metrics	z	$z > z_{.95}$	H_1
Recall	-2.0865	-2.0865 > 1.645	Reject
Precision	3.1744	3.1744 > 1.645	**Accept**
Harmonic mean	-0.3034	-0.3034 > 1.645	Reject

Table 5.14 z-test (2) of recall, precision, and harmonic mean for subject B

z-test of recall, precision, and harmonic mean of subject B			
Metrics	z	$z > z_{.95}$	H_1
Recall	-4.3595	1.1512 < 1.645	Reject
Precision	1.2833	1.2833 < 1.645	Reject
Harmonic mean	-1.4176	-1.4176 < 1.645	Reject

Table 5.15 z-test (2) of recall, precision, and harmonic mean for subject C

z-test of recall, precision, and harmonic mean of subject C			
Metrics	z	$z > z_{.95}$	H_1
Recall	0.4507	0.4507 < 1.645	Reject
Precision	3.8058	3.8058 > 1.645	**Accept**
Harmonic mean	4.3551	4.3551 > 1.645	**Accept**

Table 5.16 z-test (2) of recall, precision, and harmonic mean for subject D

z-test of recall, precision, and harmonic mean of subject D			
Metrics	z	$z > z_{.95}$	H_1
Recall	-2.6430	-2.6430 < 1.645	Reject
Precision	2.9082	2.9082 > 1.645	**Accept**
Harmonic mean	0.8767	0.8767 < 1.645	Reject

Table 5.17 z-test (2) of recall, precision, and harmonic mean for subject E

z-test of recall, precision, and harmonic mean of subject E			
Metrics	z	$z > z_{.95}$	H_1
Recall	-6.7714	-6.7714 < 1.645	Reject
Precision	4.5565	4.5565 > 1.645	**Accept**
Harmonic mean	3.2927	3.2927 > 1.645	**Accept**

Table 5.18 z-test (2) of recall, precision, and harmonic mean for all subjects

z-test of recall, precision, and harmonic mean of all subjects			
Metrics	z	$z > z_{.95}$	H_1
Recall	-2.2034	-2.2034 < 1.645	Reject
Precision	2.2007	2.2007 > 1.645	**Accept**
Harmonic mean	1.1035	1.1035 > 1.645	Reject

According to z-test in table 5.13 through 5.18, the z-test results can be summarized in table 5.19, z-test (2) summary of the experiment.

Table 5.19 z-test (2) summary of the experiment

z-test summary of the experiment						
Metrics	Topic A	Topic B	Topic C	Topic D	Topic E	% Accept H_1
Recall	Reject	Reject	Reject	Reject	Reject	**0%**
Precision	Accept	Reject	Accept	Accept	Accept	**80%**
Harmonic mean	Reject	Reject	Accept	Reject	Accept	**40%**

5.7 Experimental result summary

According to the experimental results in section 5.5, and experimental result analysis in section 5.6, results of the experiments can be summarized as follows.

1) Recall of all sample subjects mentioned in section 5.2.3, subject A, B, C, D, and E, recall of them in z-test (1) and z-test (2) cannot indicate that use case retrieval by use case structure is more effective than use case retrieval by use case keywords.
2) Precision of all sample subjects, A, B, C, D, and E in z-test (1) indicates that results of use case retrieval by use case structure is more precise than those of use case retrieval by keywords (1), and precision of almost all subjects, in z-test (2) indicates that results of use case retrieval by use case structure is more precise than those of use case retrieval by keywords (2).
3) Harmonic mean of all sample subjects, A, B, C, D, and E in z-test (1) indicates that use case retrieval by use case structure is more effective than use case retrieval by keywords, and harmonic mean of some subjects in z-test (2) indicates that use case retrieval by use case structure is more effective than use case retrieval by keywords (2).
4) The results from average recall, precision, and harmonic mean of query by use case structure and keywords (1) for each subject, shown in table 5.3, and from average recall, precision, and harmonic mean of query by use case structure and keywords (2) for each subject, shown in table 5.5 can be summarized that improving rate of precision called precision improvement of use case retrieval by use case structure increases for all sample subjects.
5) Although recall is not assured that use case retrieval by use case structure is more effective than use case retrieval by keywords, precision and harmonic mean from table 5.12 and 5.19 may indicate that results of use case retrieval by use case structure is more effective than that of use case retrieval by keywords.

5.8 Discussion

According to the experiments and its results, some issues can be obviously seen. They can be discussed as follows.

1) Query form of use case retrieval by use case structure may help users generate more suitable keywords because the users have a guideline to generate appropriate keywords in each component. For example, the users always generate keywords which are verb in “Use Case Name” component, or generate keywords which are noun about system stakeholders in “Actor” component. This suitable keyword generation helps users retrieve use cases effectively. According to the experimental results, this issue can be concluded that use case structure has an effect on use case retrieval.
2) User’s defined weighted value of each component in use case query by use case structure has an effect on the results of use case retrieval. If the users defined high weighted value in a use case component which its query keywords are highly similar to its keywords of the same use case component in the collection, the retrieval results would be more precise. This can concluded that defining suitable weighted value of the components helps users retrieve use case more effective.
3) The characteristic of the sample subjects has an effect on use case retrieval in the experiments. According to the experiments, the precision improvement is in a high rate, when the scope of the topic is defined narrowly. From table 5.3 and 5.5, precision improvement rate of topic C and D is higher than those of any other topics. This may result from topic C and D indicating both system domain and functional requirements, while the others indicate just a system domain or functional requirements. This can be concluded that the results of use case retrieval by use case structure are more precise when the scope of a subject is narrowly defined.

4) According to table 5.20, almost all users do not generate keywords in the components such as association, include, extend, generalization, precondition, postcondition, normal flow, subflow, and alternative or exceptional flow of events, but they always generate in any other components such as use case name, objective, and actor. As a result, our suggestion is that some components can be cut or reduced in use case query by use case structure. This reduction may help users are not bored and pay more attention in keyword generation. These may help users concentrate on keyword generation in the remaining components.

Table 5.20 Summary of weighted value of query by use case structure

Summary of weighted value of query by use case structure				
Component	Q (0-200)	Ratio (0-1)	Average weighted value (1-5)	Ratio (0-1)
Use case name	197	0.420	3.89	0.126
Objective	143	0.300	3.54	0.112
Actor	101	0.210	3.63	0.114
Association	7	0.010	3.57	0.113
Include	0	0.000	1.00	0.032
Extend	2	0.004	2.50	0.079
Generalization	0	0.000	1.00	0.032
Precondition	9	0.020	2.89	0.091
Postcondition	5	0.010	4.00	0.126
Normal flow of events	7	0.015	2.71	0.085
Subflow of events	1	0.002	2.00	0.063
Alternative flow of events	0	0.000	1.00	0.032
Average		**0.0833**		**0.0833**

Note : Q is the number of queries in which keyword was generated.

5) From summary weighted value of query by use case structure in table 5.20, the value of Q can be normalized to be ratio value. This ratio is computed from "Q divided by summary of Q in all 12 components of use case description". The components whose Q ratio is more than the average Q ratio are regarded as important components. It can be observed that use case name, objective, and actor are 3 components whose ratio is more than average ratio. Therefore, use case retrieval considering those 3 components shown in figure 5.3 is tested, and results of them are shown in table 5.21.

A Use Case Description Template for Use Case Retrieval	
Use Case Name :	
Objective :	
Actor :	

Figure 5.3 A use case description template for use case retrieval using 3 components

Table 5.21 Average recall, precision, and harmonic mean of query by use case structure using 3 components and by normal use case structure for each subject

Average recall, precision, and harmonic mean for each subject									
Query by use case structure					Query using 3 components				Precision
Subject	R	P	H		Subject	R	P	H	Improvement
A	0.4075	0.6945	0.4781		A	0.4083	0.6991	0.4799	**0.66%**
B	0.3180	0.2559	0.2554		B	0.3125	0.2537	0.2529	**- 0.86%**
C	0.7429	0.2785	0.3519		C	0.7375	0.2796	0.3520	**0.40%**
D	0.4615	0.4956	0.4344		D	0.4635	0.4993	0.4371	**0.75%**
E	0.5438	0.2572	0.3237		E	0.5438	0.2626	0.3263	**2.10%**
Average	**0.4947**	**0.3963**	**0.3687**		**Average**	**0.4931**	**0.3989**	**0.3696**	**0.66%**

From table 5.21, the results show that precision of use case query considering 3 mentioned components a bit increases in average.

6) From table 5.20, the average weighted value can be normalized to be ratio value like Q ratio value. The components whose average weighted value ratio is more than the average ratio are regarded as important components. It can be observed that use case name, objective, actor, association, precondition, postcondition, and normal flow of events are 7 components whose average weighted value ratio is more than average ratio. Therefore, use case retrieval considering those 7 components shown in figure 5.4 is tested, and results of them are shown in table 5.22.

A Use Case Description Template for Use Case Retrieval	
Use Case Name :	
Objective :	
Actor :	
Association :	
Precondition :	
Postcondition :	
Normal Flow of Events :	

Figure 5.4 A use case description template for use case retrieval using 7 components

Results from table 5.22 show that precision of use case query considering 7 mentioned components a bit increases in average.

Table 5.22 Average recall, precision, and harmonic mean of query by use case structure using 7 components and by normal use case structure for each subject

Average recall, precision, and harmonic mean for each subject									
Query by use case structure					Query using 7 components				Precision
Subject	R	P	H		Subject	R	P	H	Improvement
A	0.4075	0.6945	0.4781		A	0.3980	0.7113	0.4785	**2.42%**
B	0.3180	0.2559	0.2554		B	0.2938	0.2570	0.2511	**0.43%**
C	0.7429	0.2785	0.3519		C	0.6875	0.2939	0.3637	**5.53%**
D	0.4615	0.4956	0.4344		D	0.4604	0.5165	0.4486	**4.22%**
E	0.5438	0.2572	0.3237		E	0.5188	0.2714	0.3344	**5.52%**
Average	**0.4947**	**0.3963**	**0.3687**		**Average**	**0.4717**	**0.4100**	**0.3753**	**3.46%**

7) From results of table 5.21 and 5.22, they can be concluded that the components of use case description have an effect on use case retrieval results. Components that the experts always generated and defined high weighted value are regarded as important components of use case description. According to the precision of use case retrieval in table 5.21 and 5.22, we can conclude that the results of use case retrieval considering only important components are more precise than those of use case retrieval considering all components of a use case description. Therefore, a use case description template which is proper to use case retrieval is composed of 3 components, use case name, objective, and actor, and of 7 components, use case name, objective, actor, association, precondition, postcondition, and normal flow of events. These templates are shown in figure 5.3 and 5.4.

CHAPTER VI

RESEARCH SUMMARY

In this chapter, conclusion of this research is mentioned. This conclusion is composed of research summary, future works, and publication.

6.1 Research summary

This research proposes an approach for retrieving use cases using terms and use case structure similarity computation. Information storage and retrieval theories such as automatic indexing, term weighting system, and similarity computation are applied to this approach. For use case storage, contents of a use case are transformed to a set of indices and their weighted values. A user has to generate use case query for use case retrieval, and use case query is also transformed to a set of indices. Sets of indices from both user's use case query and each use case in the collection are computed to be similarity score. Use cases which their similarity score is more than predefined threshold are regarded as use cases which are similar to use case query, and they are retrieved.

In order to test effectiveness of the proposed approach, the tool for use case retrieval is developed, and the experiment is designed. This developed tool can support retrieving use case by both use case keywords and use case structure. The designed experiment is to compare effectiveness between use case retrieval by use case keywords and use case retrieval by use case structure which is our proposed approach. The experiment has 10 testers to test use case retrieval system. Many factors such as testers, use cases, subjects, and queries are controlled for unbiased results. Three selected metrics which are recall, precision, and harmonic mean are used to evaluate the effectiveness of both use case retrieval systems.

According to the results of the experiments, although recall is not assured that use case retrieval by use case structure is more effective than use case retrieval by use case keywords, but precision and harmonic mean which is the metric combining recall and precision, can conclude that use case retrieval by use case structure is more

effective than use case retrieval by use case keywords. Therefore, from the experimental results, and from the experimental statistic result analysis, they can be concluded that use case retrieval by use case structure which is our approach is more effective than use case retrieval by use case keywords.

6.2 Future works

6.2.1 For improving recall, or correctness rate of retrieval system, this research can extends thesaurus in the retrieval system. While terms similarity computation in this research considered only similarity of word but not similarity of meaning, thesaurus helps the retrieval system consider similarity of meaning of terms. This thesaurus may improve recall of the retrieval system.

6.2.2 Other information storage and retrieval theories such as user relevance feedback, probabilistic model, or document clustering can be included in this research for improving effectiveness and efficiency of use case retrieval.

6.3 Publication

This research is selected to present and publish in "The 1st National Conference on Computing and Information Technology" on May 24-25, 2005 at King Mongkut's Institute of Technology North Bangkok. The publication is "Enhancing Use Case Reusability Using Term Similarity Computation". This publication appears in the appendix C.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

REFERENCES

1. Charles, W. Krueger. Software reuse. Journal of ACM computing surveys 24 (1992): 131-183.
2. Booch, G., Rambaugh, J., and Jarcobson, I. The unified modeling language: The user guide. (n.p.): Addison Wesley, 1998.
3. Broad, E. J. Software engineering: An object-oriented perspective. (n.p.): John Wiley & Sons, 2001.
4. Saeki, M. Patterns and aspects for use cases: reuse techniques for use case descriptions. Requirements engineering conference., 2000.
5. Blok, M. C., and Cybulski, J. L. Reusing UML specifications in a constrained application domain. Proceeding of Asia-pacific software engineering conference., 1998.
6. Saeki, M. Reusing use case descriptions for requirements specification: towards use case patterns. Proceeding of Asia-pacific software engineering conference., 1999.
7. Woo, H. G., and Robinson, W. N. Reuse of scenario specifications using an automated relational learner: a lightweight approach. Proceedings of IEEE Joint requirements engineering international conference., 2002.
8. Mcclure, C. Software reuse techniques. 1st ed. (n.p.): Prentice Hall, 1997.
9. Winters, G. Use case terminology. Journal of IEEE software 22 (2005): 67.
10. Dennis, A., Tegarden, D., and Wixom, B. H. Systems analysis and design with UML version 2.0, An object-oriented approach. 2nd ed. New Jersey: John Wiley & Sons, 2005.
11. Gomaa, H. Designing concurrent, distributed, and real-time applications with UML. 1st ed. (n.p.): Addison Wesley, 2000.
12. McGill, M. J., and Salton, G. Introduction to modern information retrieval. (n.p.): McGraw Hill, 1983.
13. Neto, B. R., and Yates, R. B. Modern information retrieval. (n.p.): Addison Wesley, 1999.

14. Yates, R. B., and William, B. Information retrieval: Data structures and algorithms. (n.p.): Prentice Hall, 1992.

15. Porter, M. F. An algorithm for suffix stripping: Program. Automated library and information systems 14 (1980): 130-137.

16. กัลยา วานิชย์บัญชา. การวิเคราะห์สถิติ: สถิติสำหรับการบริหารและวิจัย. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, 2546.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

APPENDICES

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

APPENDIX A

SELECTED USE CASE DOMAIN

This chapter describes details of all use case domain used in this thesis. They are composed of 16 use case domains from various systems. All of them are practical use case domain, and they were written in English language. All 16 selected use case domains are introduced in Table A.1.

Table A.1 All selected use case domains

Use Case Domain	Use case name
1	ICT blocking and filtering system
2	Departmental information system for curriculum and course offerings management
3	Resource management system
4	Automatic restaurant lookup system
5	Library management system
6	Mobile phone selling on the internet system
7	Payroll system
8	Personal stock manager on PDA system
9	Basic product trading system
10	Mobile phone service paying system
11	MLM computerize system
12	Content management and platform upgrade for the official website of Nation Channel
13	Purchase, producing, and export system
14	Management information system for insurance company
15	E-learning system
16	Management information system for an academic

Table A.2 ICT blocking and filtering system

Domain name : ICT blocking and filtering system	
Domain description : This system is to detect and prevent bad or unsuitable things on the internet, and to monitor a game player whose age is under 18 years old. The system helps parents, teachers, or cyber inspectors to reduce risk, danger, or unsuitable information which may be obtained with their children.	
Use case ID	**Use case name**
0101	Register for a member
0102	Edit members information
0104	Cancel member status
0105	Request report about members
0106	Manage information for users
0107	Provide other information
0108	Display information
0109	Create token
0110	Create log file
0111	Verify log file
0112	Add game online users information
0113	Manage rules
0114	Receive log file from a game online system
0115	View monitoring reports
0116	Pre-Screening
0117	Manage a bad website
0118	Manage a good website
0119	Manage suspended keywords
0120	View checking blocked information reports

Table A.2 ICT blocking and filtering system (Continued)

Use case ID	Use case name
0121	Notify a trace
0122	Initialize sending SMS system
0123	Notify SMS for emergency case
0124	Check notifying a trace
0125	View a notified trace report
0126	Define a system authority for a member
0127	View a report for defining authority for a member

Table A.3 Departmental information system for curriculum and course offerings management

Domain name : Departmental information system for curriculum and course offerings management **system**	
Domain description : This system is to manage course and curriculum information in an academic.	
Use case ID	**Use case name**
0201	Save course details
0202	Save course operation status
0203	Save course operation report
0204	Save course operation result
0205	View course operation details
0206	View course operation status
0207	View course operation report
0208	View course operation result
0209	Save subject details
0210	Save subject operation status
0211	Store subject report
0212	Save subject operation result
0213	View subject operation details

Table A.3 Departmental information system for curriculum and course offerings management **(Continued)**

Use case ID	Use case name
0214	View subject operation status
0215	View subject operation result
0216	View subject operation report
0217	Save course information
0218	Edit course information
0219	View course information
0220	View course statistic information
0221	Save subject information
0222	Edit subject information
0223	View subject information
0224	View teaching schedule

Table A.4 Resource management system

Domain name : Resource management system	
Domain description : This system is to manage operation in an organization. The system is devided into 6 subsystems. There are customer management, sales management, employee management, supplier management, inventory/product management, and supplier order management subsystem.	
Use case ID	Use case name
0301	Process new application
0302	Process add new customer
0303	Process customer verification
0304	Process Credit Customer Payment
0305	Process Cash Sales transactions
0306	Process Credit Sales
0307	Process Lay-Bye Sales
0308	Process Sales Returns

Table A.4 Resource management system (Continued)

Use case ID	Use case name
0309	Capture employees details
0310	Process employee records
0311	Process existing supplier
0312	Process New Supplier
0313	Process Incoming Inventory
0314	Process Self-Manufactured Products
0315	Process Check Inventory Available
0316	Process Inventory Order
0317	Access control
0318	Process Generate Password
0319	Process Order Details
0320	Process Receive Order Details
0321	Process order payment
0322	Process Account Payments
0323	Generate Sales Figures
0324	Process Petty Cash Withdrawals
0325	Process Supplier Payments
0326	User Trail
0327	Supplier Order Trail
0328	Customer Payment Trail
0329	Customer Order Trail
0330	Inventory Trail

Table A.5 Automatic restaurant lookup system

Domain name : Automatic restaurant lookup system
Domain description : This system is to search restaurants and their favorite food. The restaurants can be searched depending on their areas, kinds of food, or favorite. In addition, this system can work on both PC and PDA.

Table A.5 Automatic restaurant lookup system (Continued)

Use case ID	Use case name
0401	New customer
0402	Manage restaurant
0403	Manage customer
0404	Find a restaurant
0405	View restaurants information
0406	Vote restaurant
0407	Manage map
0408	Post webboard
0409	Manage webboard
0410	View FAQ
0411	Manage FAQ
0412	Log in
0413	Find restaurant on PDA
0414	View restaurant's information on PDA
0415	Vote restaurant on PDA

Table A.6 Library management system

Domain name : Library management system	
Domain description : This system is to manage works on a library using RFID technology. The system provides resource management and services management for a librarian, a student, or other users.	
Use case ID	**Use case name**
0501	Manage users information
0502	Summarize users report
0503	Manage library media information
0504	Define library media status
0505	Summarize library media
0506	Add members information

Table A.6 Library management system (Continued)

Use case ID	Use case name
0507	Manage members information
0508	Summarize members
0509	Borrow library media
0510	Return library media
0511	Pay a fine

Table A.7 Mobile phone selling on the internet system

Domain name : Mobile phone selling on the internet system	
Domain description : This system is to sell mobile phone on the internet. The system can serve searching mobile phone's information, selling system, payment system, and making suggestion about mobile phone selection for a user.	
Use case ID	**Use case name**
0601	Search product's information
0602	View product's information
0603	Comment
0604	Compare products
0605	Purchase products
0606	Manage a product list
0607	Check purchase order
0608	Pay money
0609	Use credit card
0610	Transfer money
0611	Confirm transferring money
0612	Verify user's payment

Table A.8 Payroll system

Domain name : Payroll system
Domain description : This system is to schedule works of popcorn seller in many places and to calculate popcorn seller's salary depending on his/her real work schedule.

Table A.8 Payroll system (Continued)

Use case ID	Use case name
0701	Add new selling place
0702	Edit selling places information
0703	Cancel selling place
0704	Edit sale staff information
0705	Delete sale staff information
0706	View sale staff information
0707	Display work schedule
0708	Save working time
0709	Calculate staff salary
0710	Calculate social assurance fee
0711	Calculate over time fee
0712	Fine for working late
0713	Calculate special fee
0714	Fine for absent work.
0715	Decrease tax
0716	Collect additional money from selling place
0717	Calculate net salary
0718	Display net salary
0719	Create report for executive
0720	Create report for sale staff or seller
0721	View selling place information
0722	Add new sale staff or seller
0723	Select a sale staff or seller
0724	Calculate salary
0725	Manage salary
0727	Save taking money in advance
0728	View taking money in advance information

Table A.9 Personal stock manager on PDA system

Domain name : Personal stock manager on PDA system	
Domain description : This system is to manage information about stock on PDA. A user can view stock's information on realtime, and can trade his/her share or stock by using PDA.	
Use case ID	**Use case name**
0801	Log-In
0802	Verify user's account
0803	Search by Abbreviate
0804	Search stock
0805	View top ten
0806	View most gain
0807	View most lost
0808	View most active
0809	View Portfolio
0810	View transaction
0811	View index
0812	View search results
0813	View other currencies
0814	View stock details
0815	View graph
0816	Buy stock
0817	Sell stock
0818	Verify PIN
0819	Check regulation
0820	Calculate price and commission
0821	Cancel transaction
0822	Check user regulation
0823	Maintain user account
0824	Maintain regulation

Table A.10 Basic product trading system

Domain name : Basic product trading system	
Domain description : This system is to manage product selling of a company on the internet. The system includes customer management, product management, selling management, and billing management subsystem.	
Use case ID	**Use case name**
0901	Download purchase order
0902	Maintain purchase order
0903	Maintain sale order
0904	Transfer purchase order to sale order
0905	Approve purchase order
0906	Approve price
0907	Approve credit
0908	Approve stock balance
0909	Maintain invoice
0910	Transfer sale order to invoice
0911	Confirm invoice
0912	Update stock balance
0913	Update customer credit
0914	Maintain customer credit
0915	Maintain product price
0916	Maintain product balance

Table A.11 Mobile phone service paying system

Domain name : Mobile phone service paying system
Domain description : This system is to manage mobile phone's paying service of a company. The system can manage customer's information, compute service's fee depending on customer's package, and manage reports for information system of an executive.

Table A.11 Mobile phone service paying system (Continued)

Use case ID	Use case name
1001	Register telephone number
1002	Manage customer and telephone number information
1003	Manage customer information
1004	Manage telephone number information
1005	Manage service information
1006	Cancel service
1007	Manage service package
1008	Save service information
1009	Ask for information
1010	Check service bill
1011	Check remaining money
1012	Create reports
1013	Print paying report
1014	Print paying report by terminal network
1015	Print income summary report

Table A.12 Basic MLM computerize system

Domain name : Basic MLM computerize system	
Domain description : This system is to manage information of a MLM system. The system can manage customer's information, manage product warehouse, manage product selling, and calculate staff's benefit.	
Use case ID	**Use case name**
1101	Sign in member
1102	Maintain member information
1103	Show child member
1104	Cancel member
1105	Adjust position
1106	Purchase order

Table A.12 Basic MLM computerize system (Continued)

Use case ID	Use case name
1107	Receive product
1108	Adjust product quantity
1109	Maintain information
1110	New product information
1111	Change product information
1112	New position information
1113	Change position information
1114	Order product
1115	Create invoice
1116	Save point value
1117	Print invoice
1118	Cancel invoice
1119	Print commission
1120	Calculate commission
1121	Promote position

Table A.13 Content management and platform upgrade for the official website of Nation Channel

Domain name : Content management and platform upgrade for the official website of Nation Channel	
Domain description : This system is to manage news content on website automatically. This website provides news information, file service, photo service, activate live broadcast. The system mainly supports news information management and other service management.	
Use case ID	**Use case name**
1201	Login
1202	Manage Photo
1203	Load Photo Dialogue

Table A.13 Content management and platform upgrade for the official website of Nation Channel (Continued)

Use case ID	Use case name
1204	Edit Schedule Table
1205	Reply Messages
1206	Manage Program
1207	Manage Today News
1208	Link Video Clip
1209	View Today News
1210	Post Photo Captions
1211	Leave message to anchorman
1212	Vote Poll
1213	Read Breaking News
1214	Activate Live

Table A.14 Purchase, producing, and export system

Domain name : Purchase, producing, and export system	
Domain description : This system is to manage material purchase order, production, product import, and product export of a company.	
Use case ID	**Use case name**
1301	Purchase system
1302	Product system
1303	Warehouse and export system
1304	Check raw material
1305	Make Contract and Order Raw Material
1306	Track inventory and generate report
1307	Configure flow, hire rate and raw material
1308	Create barcode and preparation
1309	Produce
1310	Calculate hire and create report

Table A.14 Purchase, producing, and export system (Continued)

Use case ID	Use case name
1311	Check inventory
1312	Contact port
1313	Contact duty
1314	Export

Table A.15 Management information system for insurance company

Domain name : Management information system for insurance company	
Domain description : This system is to support insurance selling of an insurance company. The system mainly provides information about insurance details for both staff and customer. In addition, this system supports information on both normal PC and wireless technology.	
Use case ID	**Use case name**
1401	Place a claim
1402	View claim status
1403	External approve claim
1404	Maintain company information
1405	View company information
1406	View product information
1407	Calculate an insurance application
1408	Verify underwriting rule
1409	Calculate premium
1410	Validate insurance application
1411	View agent information
1412	View agent holder information
1413	View policy holder information
1414	Authentication

Table A.16 E-learning system

Domain name : E-learning system	
Domain description : This system is to support E-learning system on website. A student and other users can register and study via website. The system can manage course, student's information, teacher's information, examination, and studying-teaching evaluation.	
Use case ID	**Use case name**
1501	Register a new user for E-Learning system
1502	Generate password
1503	Log in to E-Learning system
1504	Change password
1505	Announce course information
1506	Upload materials
1507	Create examinations and examination solutions
1508	Post answer
1509	View course detail
1510	Register course
1511	Verify condition
1512	Study registered course
1513	Take examination and get result of the examination
1514	Record time
1515	View history's study information
1516	Post question
1517	View study result
1518	Follow study result
1519	View study result
1520	Number of learner or student
1521	Maintain employee information

Table A.17 Management information system for an academic

Domain name : Management information system for an academic	
Domain description : This system is to manage course in some academics. The system is divided in 3 main parts; there are students, faculties, and administrators. The system can manage student's information, faculty's information, studying schedule, and subject's information.	
Use case ID	**Use case name**
1601	Log-In
1602	Student Adjust Schedule
1603	Student View All Courses
1604	Student View Record
1605	Student View Schedule
1606	Faculty View Courses
1607	Faculty View Enrollment
1608	Faculty View Student Records
1609	Add another admin
1610	Add another building
1611	Add another class
1612	Add another course
1613	Add another department
1614	Add another faculty
1615	Add student
1616	Modify existing admin
1617	Add existing building
1618	Modify existing classroom
1619	Add existing course
1620	Modify existing department
1621	Admin Modify Enrollment List
1622	Modify existing faculty member
1623	Modify existing student

Table A.17 Management information system for an academic (Continued)

1624	Remove administrator
1625	Remove building
1626	Remove classroom
1627	Remove course
1628	Remove department
1629	Remove faculty member
1630	Remove student

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

APPENDIX B

TESTER'S USE CASE QUERY

Table B.1 Tester's query set of query by use case keywords for topic A

Tester	Query	Tester's query set
1	1	manage schedule
	2	payment
	3	student
	4	teacher
	5	course material
2	1	course lecturer grade subject student instructor gpa credit effective course core teach register
	2	training score instructor name exam room
	3	course material description prerequisute thesis instructor topic biography schedule
	4	assignment course name curriculum research resources library class announcement scholarship thesis defense
	5	activities bachelor master doctoral degree publications thesis exam schedule
3	1	course
	2	search research
	3	statistic
	4	search senior project
	5	course description
4	1	learn on web online
	2	instructor management
	3	search curriculum
	4	manage course
	5	examination

Table B.1 Tester's query set of query by use case keywords for topic A (Continued)

Tester	Query	Tester's query set
5	1	grade
	2	teacher
	3	student
	4	course
	5	register
6	1	take examination
	2	teacher section class grade
	3	withdraw examination
	4	learn study
	5	schedule academic education
7	1	master thesis conference
	2	search university
	3	subject credit
	4	register subject
	5	grade report
8	1	department faculty university
	2	classroom building
	3	teaching studying
	4	query grade student course
	5	view subject classroom

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table B.2 Tester's query set of query by use case keywords for topic B

Tester	Query	Tester's query set
1	1	credit card verification
	2	customer payment
	3	add shoping cart
	4	search product
	5	select item
2	1	trade study product name description price
	2	cost vendor acquisition advertisement promotion discount newest recommended new release
	3	negotiation exchange barter trade
	4	trade off trader merchandise sell buy low price
	5	special worth money shipping time span damage defect second hand used old excellent firsthand similar product categories types
3	1	add item
	2	delete item
	3	show list of product
	4	add product to cart
	5	delete product from cart
4	1	transfer money
	2	buy with credit card
	3	buy product supplier sell
	4	customer sell buy product on hand
	5	customer query product item supplier
5	1	customer contact product list
	2	manage inventory
	3	sale system
	4	receivable billing
	5	shipment delivery product

Table B.2 Tester's query set of query by use case keywords for topic B (Continued)

Tester	Query	Tester's query set
6	1	trade product
	2	money paying finance
	3	selling product
	4	buying product
	5	product price description
7	1	selling bill
	2	credit card payment
	3	trading system
	4	movie ticket payment online
	5	invoice management product
8	1	sell buy customer
	2	purchase order management
	3	product management
	4	product price list
	5	paying billing invoice receipt

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table B.3 Tester's query set of query by use case keywords for topic C

Tester	Query	Tester's query set
1	1	add customer profile
	2	search customer
	3	manage customer
	4	customer level
	5	sell order list
2	1	life span money purchase valid expire regular customer favorite
	2	customer name address age birthday road province career occupation
	3	location postal zip code district amphur
	4	events activities participation details
	5	top excellent bad good list customer
3	1	analyze user needs
	2	add customer
	3	store customer order history
	4	search by customer name
	5	view customer profile
4	1	add customer profile
	2	remove customer profile
	3	query customer
	4	maintain customer contact
	5	exchange customer information between company
5	1	update customer profile
	2	contact management
	3	customer relationship management
	4	manage contact information
	5	manage selling buying history

Table B.3 Tester's query set of query by use case keywords for topic C (Continued)

Tester	Query	Tester's query set
6	1	name
	2	address telephone number
	3	balance tax credit
	4	customer information management
	5	new customer
7	1	customer management
	2	information
	3	customer
	4	user information
	5	phone update information e-mail customer
8	1	customer name address credit
	2	customer detail firstname lastname title
	3	add insert update delete customer
	4	query display customer report
	5	customer payment management

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table B.4 Tester's query set of query by use case keywords for topic D

Tester	Query	Tester's query set
1	1	value per unit
	2	number item
	3	tax including
	4	total value
	5	payment
2	1	price money purchase monetary years months
	2	yearly annual income expense expenditure fee charge
	3	net profit benefit cost risk
	4	gain advantage interests lost turnover business
	5	tread increase decrease
3	1	calculate interest
	2	income summary report
	3	current balance rate
	4	set interest rate
	5	calculate return on investment
4	1	calculate tax invoice
	2	calculate salary staff
	3	calculate bonus dividen
	4	query capital asset company
	5	account product credit debit
5	1	account receivable
	2	calculate income
	3	calculate depreciation
	4	calculate revenue
	5	general ledger

Table B.4 Tester's query set of query by use case keywords for topic D (Continued)

Tester	Query	Tester's query set
6	1	money
	2	tax
	3	rate
	4	debit credit
	5	amount summary
7	1	financial calculation
	2	financial information
	3	financial management
	4	money calculation
	5	finance summary report
8	1	rate vat calculation
	2	price percent calculation
	3	summary discount total system
	4	rate per unit total system
	5	calculate vat input customer

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table B.5 Tester's query set of query by use case keywords for topic E

Tester	Query	Tester's query set
1	1	item summary
	2	customer name
	3	create receipt
	4	create item
	5	summary report
2	1	template field require option summary
	2	tailoring items row column print annual report monthly daily
	3	summary report recommendation
	4	page logo standard id
	5	number of row column report
3	1	create report
	2	search annual report
	3	generate report template
	4	add update delete report
	5	print report
4	1	create report all product
	2	create summary report profit year
	3	create sale income summary report
	4	summary report staff
	5	query report inventory product
5	1	monthly report
	2	information data filter
	3	analytic company report
	4	report evaluation process
	5	education report

Table B.5 Tester's query set of query by use case keywords for topic E (Continued)

Tester	Query	Tester's query set
6	1	cost sell order report
	2	generate report
	3	report system
	4	make new report
	5	report information
7	1	report invoice summary
	2	total type category column
	3	sort information group report
	4	montyly weekly daily annual report summary
	5	display show summary report
8	1	create transaction summary report
	2	course subject summary report
	3	head detail description report
	4	curriculum student report
	5	total money calculation report

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table B.6 Tester's query set of query by use case structure for topic A

Tester	Query	Tester's query set	Component	Weight
1	1	register course	1	5
		student	3	4
	2	course scheduling	1	5
		teacher	3	4
	3	teaching scheduling	1	3
		schedule for teaching	2	5
		teacher instructor	3	5
	4	study in subject	1	3
		student learns study	2	4
		student	3	5
	5	take examination	1	5
		student	3	5
2	1	course	1	5
		course description	2	5
		instructor teacher	3	5
		credit	4	3
		prerequisite	8	3
	2	schedule	1	5
	3	register	1	5
		student	3	5
	4	search course	1	5
		student teacher instructor	3	5
	5	select course	1	5
		course description	2	5
		student teacher instructor	3	5
3	1	course scheduling	1	5
		teacher officer	3	5

Table B.6 Tester's query set of query by use case structure for topic A (Continued)

Tester	Query	Tester's query set	Component	Weight
3	2	search course	1	5
		search all opened course	2	4
		student	3	5
	3	view student profile	1	5
		views details of student profile	2	4
		teacher	3	5
	4	view research progress	1	5
		view progression of student research	2	3
		teacher advisor	3	5
	5	select subject	1	4
		student	3	4
4	1	query course	1	4
		search course description	2	4
		student teacher staff	3	3
		query	4	5
		print course	6	2
		search course description	10	3
	2	register course	1	5
		register course student	2	3
		student	3	3
		apply for course	6	3
		query course register course	10	3
	3	learn course	1	3
		learn course website	2	3
		student	3	4
		register course	8	5
		open website learn course	10	4

Table B.6 Tester's query set of query by use case structure for topic A (Continued)

Tester	Query	Tester's query set	Component	Weight
4	4	query student	1	4
		to query student information	2	3
		student teacher	3	4
	5	withdraw course	1	3
		to withdraw registered course	2	3
		student	3	4
5	1	course	1	5
		add update course	2	4
	2	teacher lecturer	3	4
	3	register course	1	4
		register course view status	2	4
	4	test	1	3
		student test examination	2	4
		student	3	4
	5	room allocation	1	3
		room classroom scheduling allocation	2	3
6	1	student	1	4
		register course gpa	2	3
		student teacher course	3	4
		register	10	2
		take course	11	2
	2	teacher	1	4
		teach grading	2	3
		teacher student grade course	3	4
	3	course	1	4
		contain min max name	2	2
		teacher student	3	3

Table B.6 Tester's query set of query by use case structure for topic A (Continued)

Tester	Query	Tester's query set	Component	Weight
6	4	classroom	1	4
		classroom serve reservation	2	3
	5	internet course learning	1	4
		course description learn	2	2
		student teacher course	3	2
		register	4	4
7	1	educational curriculum	1	3
		search curriculum	2	3
		student	3	5
	2	course	1	3
		find course	2	3
	3	program	1	3
		find program study	2	3
	4	manage course	1	3
		manage course description	2	3
		instructor	3	3
	5	educational information	1	4
8	1	manage instructor	1	3
		manage instructor name information	2	3
	2	register	1	3
		student register course	2	3
		student	3	3
	3	grade report	1	3
		student	3	3
	4	teacher student professor	3	3
	5	curriculum	1	3
		add insert update course curriculum	2	3

Table B.7 Tester's query set of query by use case structure for topic B

Tester	Query	Tester's query set	Component	Weight
1	1	sell product	1	5
		customer seller	3	5
	2	customer information	1	3
		customer staff	3	5
	3	search product	1	4
		search product in stock	2	3
		customer staff	3	5
	4	buy product	1	4
		staff seller	3	3
	5	add product	1	3
		add product into stock	2	3
2	1	purchase	1	5
		vendor customer	3	5
	2	trade study	1	5
		decision making	2	4
		vendor customer	3	4
	3	sell	1	5
		vendor customer	3	5
		determined price	8	5
		bill	9	5
	4	search product	1	5
		product description	2	5
		customer	3	5
	5	product comparison	1	5
		product	4	5
3	1	search product	1	5
		customer	3	5

Table B.7 Tester's query set of query by use case structure for topic B (Continued)

Tester	Query	Tester's query set	Component	Weight
3	2	buy product	1	5
		customer	3	4
	3	sell product	1	4
		customer sales	3	4
	4	view customer order	1	4
		view list of products ordered by customer	2	4
		sale	3	4
	5	purchase order	1	5
		vendor producer	3	4
4	1	sell product	1	5
		sell product stock	2	4
		seller sale	3	4
	2	buy product	1	5
		supplier	3	4
	3	query product	1	5
		query product	2	4
		customer sale	3	4
	4	purchase order	1	5
		purchase product	2	4
		seller	3	4
	5	order product	1	4
		order product	2	4
		customer	3	3
5	1	purchase	1	5
		purchase product	2	5
		customer	3	5
	2	billing system	1	5

Table B.7 Tester's query set of query by use case structure for topic B (Continued)

Tester	Query	Tester's query set	Component	Weight
5	3	inventory	1	4
		manage inventory stock	2	4
	4	sale seller	3	5
	5	sale tracking	1	4
		track sale progress status	2	3
6	1	trade	1	3
		trade product	2	4
	2	buy stock	1	4
		buy stock to company	2	3
		supplier company	3	3
	3	product selling	1	4
		saleman manager	3	4
	4	purchase product	1	4
		buy product	2	4
	5	trade	1	4
		trade product	2	4
7	1	add product	1	4
		add product category	2	2
	2	order product	1	4
		customer order product	2	2
	3	manage product category	1	4
		manage product category	2	2
		manage product category	10	2
	4	billing system	1	4
		payment	2	2
	5	sell product	1	4
		sale	3	2

Table B.7 Tester's query set of query by use case structure for topic B (Continued)

Tester	Query	Tester's query set	Component	Weight
8	1	buy product	1	4
		buy product list	2	4
	2	product order	1	4
		get product order from customer	8	2
	3	product trading	1	4
		customer	3	3
		search product	10	2
	4	product report	1	3
		summary report product	2	3
	5	deliver product	1	3
		send product to customer	2	3
		saleman seller	3	4

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table B.8 Tester's query set of query by use case structure for topic C

Tester	Query	Tester's query set	Component	Weight
1	1	add customer information	1	4
		add customer information	2	4
		staff	3	3
	2	delete customer	1	4
		staff	3	3
	3	search customer	1	4
	4	promote customer	1	4
	5	view customer	1	4
		view customer profile	2	5
		staff	3	4
2	1	search customer	1	5
		customer details	2	4
		product	4	2
	2	add customer	1	5
	3	delete customer	1	5
	4	update customer	1	5
	5	customer summary	1	5
3	1	create customer	1	4
		create new customer profile	2	3
		staff	3	3
	2	analyze customer needs	1	4
		analyze customer needs by customer order history	2	3
		staff	3	3
	3	delete customer	1	4
		delete customer profile	2	3
		staff	3	3

Table B.8 Tester's query set of query by use case structure for topic C (Continued)

Tester	Query	Tester's query set	Component	Weight
3	4	store customer order	1	4
		store customer order for further customer needs analysis	2	3
		customer	3	3
	5	update customer profile	1	4
		update customer profile	2	3
		staff	3	3
4	1	add customer	1	5
		add new customer	2	4
		staff manager	3	3
	2	remove customer	1	5
		remove customer profile	2	4
		staff manager	3	3
	3	update customer	1	5
		update customer profile	2	4
		manager	3	3
		available customer	8	2
	4	query customer	1	5
		query customer information	2	4
		staff manager	3	3
		query	4	4
	5	check credit	1	5
		check customer credit	2	4
		employee staff	3	3
5	1	schedule	1	1
		calendar schedule events	2	4
	2	address	1	4

Table B.8 Tester's query set of query by use case structure for topic C (Continued)

Tester	Query	Tester's query set	Component	Weight
5	2	manage customer address	2	3
		customer	3	5
	3	purchase history	1	3
		purchase history track	2	5
	4	contact management	1	3
		manage contact management address	2	5
	5	analyze customer data	1	3
		analyze customer purchase data mining	2	5
6	1	customer	1	5
		normal activities of customer	2	5
		customer	3	5
	2	retrieve data	1	4
		get information of customer	2	4
		searcher	3	3
		get information	9	4
	3	customer buying detail	1	4
		get or set customer buyer detail	2	3
		customer searcher	3	4
		set data	8	1
		get data	9	3
	4	set promotion	1	4
		set appropriate data for each customer	2	4
		customer definer	3	3
	5	advertising for customer	1	2
		create promotion for advertise customer	2	3
		customer creator	3	5
		set advertising for customer	8	3

Table B.8 Tester's query set of query by use case structure for topic C (Continued)

Tester	Query	Tester's query set	Component	Weight
7	1	customer	1	5
	2	customer	1	5
		customer information	2	3
	3	new customer	1	3
		customer administrator	3	3
	4	customer management	1	3
		customer administrator	3	3
	5	information	1	3
8	1	add customer information	1	2
		add customer information	2	2
	2	update customer information	1	2
		update customer information	2	2
	3	delete customer information	1	2
		delete customer information	2	2
	4	change status	1	2
		change customer status active inactive	2	2
	5	report customer information	1	2
		report customer information summary daily monthly	2	2

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table B.9 Tester's query set of query by use case structure for topic D

Tester	Query	Tester's query set	Component	Weight
1	1	calculate price	1	3
	2	calculate tax	1	4
	3	salary calculation	1	5
	4	net benefit	1	2
	5	price information	1	3
2	1	net profit	1	5
	2	annual profit	1	5
	3	calculate interest	1	5
		calculate annual interest	2	4
		user	3	4
	4	set interest rate	1	5
		set interest rate for interest calculation	2	3
	5	calculate investment	1	3
		calculate investment summary	2	3
3	1	create net profit report	1	4
		create net profit report	2	3
		user	3	3
	2	calculate tax	1	4
		calculate tax product	2	4
		employee manager	3	4
	3	calculate amount payment	1	5
		calculate total payment	2	4
		get total payment	9	4
	4	calculate money	1	4
		money summary calculation	2	4
	5	financial information	1	4
		income outcome money	2	4

Table B.9 Tester's query set of query by use case structure for topic D (Continued)

Tester	Query	Tester's query set	Component	Weight
4	1	calculate tax	1	4
		calculate tax	2	4
		officer	3	4
	2	manage dividen	1	4
		calculate dividen share	2	3
		manager	3	2
	3	calculate salary	1	4
		calculate salary employee	2	4
	4	price summary	1	4
		calculate price summary	2	4
		staff	3	3
	5	payment summary	1	4
		total amount payment	2	4
5	1	calculate net profit	1	2
		calculate net profit general ledger	2	4
	2	calculate depriciation	1	2
		calculate depriciation assets	2	4
	3	calculate net income	1	3
		calculate net income taxes	2	4
	4	calculate balance account	1	3
		calculate balance account	2	5
	5	cash flow prediction	1	3
		predict cash flow money	2	3
6	1	get finance information	1	4
		view finance money information	2	3
		searcher	3	4
	2	set criteria	1	2

Table B.9 Tester's query set of query by use case structure for topic D (Continued)

Tester	Query	Tester's query set	Component	Weight
6	2	set criteria for financial calculation	2	4
		calculator	3	3
	3	calculate process	1	2
		calculation information for each record	2	2
		calculator	3	2
	4	calculate process	1	2
		price money calculation	2	2
		calculator	3	2
	5	financial report	1	3
		summary money price report	2	3
7	1	financial calculation	1	5
	2	financial report	1	3
		summary money report	2	3
	3	manage calculation type	1	3
		manage calculation type group category	2	3
	4	insurance rate calculation	1	3
		insurance rate calculation precent	2	3
	5	stock index calculation	1	3
		stock index calculation price amount	2	3
8	1	interest	1	3
		calculate display	2	3
	2	calculate	1	4
		tax interest	2	4
	3	generate calculate	1	3
		debit credit finance	2	3
	4	interest deposit calculation	1	3
		interest deposit calculation	2	3

Table B.9 Tester's query set of query by use case structure for topic D (Continued)

Tester	Query	Tester's query set	Component	Weight
8	5	price calculation	1	4
		price summary billing calculation	2	4

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table B.10 Tester's query set of query by use case structure for topic E

Tester	Query	Tester's query set	Component	Weight
1	1	print report	1	4
	2	update report	1	4
	3	delete report	1	4
	4	summary information report	1	4
	5	create generate receipt	1	4
2	1	create summary report	1	5
	2	formatting report	1	5
	3	adjust or tailor report	1	5
	4	delete report	1	5
		delete unnecessary information report	2	5
	5	search report	1	5
		search summary document report	2	5
3	1	generate report template	1	5
		generate report template	2	5
		user	3	2
	2	search report	1	5
		search report	2	5
		user	3	2
	3	print report	1	5
		print report	2	5
		user	3	2
	4	view report	1	5
		view report	2	5
		user	3	2
	5	generate receipt	1	4
		generate receipt summary	2	4
		buyer	3	2

Table B.10 Tester's query set of query by use case structure for topic E (Continued)

Tester	Query	Tester's query set	Component	Weight
4	1	create sale report	1	4
		create manager report sale product	2	4
		manager	3	3
	2	create employee report	1	4
		employee list	2	3
		employee	3	3
	3	create product report	1	4
		summary product report	2	2
		producer	3	2
	4	create profit report	1	3
		create profit report investment	2	3
		company	3	3
	5	create customer report	1	4
		customer list	2	3
		company	3	3
		generate query	4	2
		available customer	8	2
5	1	create summary report	1	4
		create summary report	2	5
	2	create monthly report	1	4
		create monthly report	2	5
	3	detailed analysis report	1	3
		detailed analysis report analytical purpose	2	4
	4	special report	1	3
		create special report	2	4
	5	prediction report	1	2
		data mining analysis report	2	3

Table B.10 Tester's query set of query by use case structure for topic E (Continued)

Tester	Query	Tester's query set	Component	Weight
6	1	get information for create report	1	4
		get information for create report	2	3
		reportor	3	3
	2	set criteria	1	3
		set criteria for retrieve report	2	3
		reportor	3	3
	3	create product report	1	4
		summary product report	2	2
		producer	3	2
	4	print report	1	3
		print report by approve information	2	3
		reportor manager	3	4
		set criteria	8	3
		get report	9	4
	5	make new report	1	4
		make new report	2	4
7	1	manage report	1	4
		add edit delete report	2	4
	2	report	1	4
		report system	2	4
	3	create report	1	4
		create report	2	4
	4	total order report	1	4
		total order report daily weekly	2	4
	5	cost production report	1	4
		cost production report group supplier	2	4
		customer	3	3

Table B.10 Tester's query set of query by use case structure for topic E (Continued)

Tester	Query	Tester's query set	Component	Weight
8	1	customer information report	1	2
		customer information report	2	2
	2	production report	1	3
		production report line port lot	2	3
	3	staff salary report	1	4
		staff salary report department	2	4
	4	summary report	1	3
		display view show	2	3
	5	report	1	4
		header detail	10	3

Table B.11 Components of the use case description

Component ID	**Component name**
1	Use case name
2	Objective
3	Actor
4	Association
5	Include
6	Extend
7	Gneralization
8	Precondition
9	Postcondition
10	Normal flow of events
11	Subflow of events
12	Alternative or exceptional flow of events

APPENDIX C

SIMILARITY SCORE AND RECALL, PRECISION, AND HARMONIC MEAN OF THE EXPERIMENTS

Table C.1 Average and maximum similarity score of query by use case structure and keywords (1) for subject A

Average and maximum similarity score of all queries for topic A					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
1	0.00053	0.01761		0.00177	0.07010
2	0.00022	0.01282		0.00171	0.05405
3	0.00019	0.00840		0.00155	0.09795
4	0.00023	0.00593		0.00140	0.05313
5	0.00051	0.01760		0.00122	0.10714
6	0.00126	0.02747		0.00121	0.03457
7	0.00053	0.02778		0.00133	0.12162
8	0.00110	0.03050		0.00140	0.09091
9	0.00099	0.05311		0.00198	0.06049
10	0.00039	0.03367		0.00157	0.04033
11	0.00022	0.00937		0.00211	0.04865
12	0.00021	0.00889		0.00116	0.04210
13	0.00032	0.05556		0.00135	0.02542
14	0.00062	0.10000		0.00090	0.01881
15	0.00038	0.00904		0.00117	0.05442
16	0.00065	0.03623		0.00083	0.01748
17	0.00049	0.01562		0.00128	0.05333
18	0.00021	0.00889		0.00085	0.02623
19	0.00050	0.00904		0.00131	0.05584
20	0.00021	0.02381		0.00147	0.10000

Table C.1 Average and maximum similarity score of query by use case structure and keywords (1) for subject A (Continued)

Average and maximum similarity score of all queries for topic A					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
21	0.00000	0.00000		0.00116	0.03748
22	0.00023	0.00593		0.00206	0.06636
23	0.00019	0.00840		0.00193	0.08523
24	0.00022	0.00937		0.00086	0.04806
25	0.00020	0.01136		0.00071	0.04103
26	0.00021	0.02381		0.00107	0.03483
27	0.00052	0.03922		0.00063	0.02622
28	0.00041	0.06452		0.00125	0.03203
29	0.00049	0.02589		0.00144	0.23834
30	0.00024	0.01678		0.00080	0.01909
31	0.00000	0.00000		0.00064	0.04947
32	0.00021	0.00889		0.00095	0.03795
33	0.00044	0.01324		0.00056	0.04375
34	0.00042	0.01324		0.00142	0.04815
35	0.00028	0.00828		0.00073	0.01749
36	0.00041	0.01111		0.00091	0.02889
37	0.00034	0.03602		0.00142	0.06667
38	0.00050	0.04274		0.00111	0.05442
39	0.00059	0.06897		0.00165	0.06531
40	0.00063	0.03333		0.00043	0.00904
Average	**0.00041**	**0.02381**		**0.00123**	**0.05556**

Table C.2 Average and maximum similarity score of query by use case structure and keywords (1) for subject B

Average and maximum similarity score of all queries for topic B					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
1	0.00066	0.04823		0.00185	0.05385
2	0.00043	0.01313		0.00154	0.03191
3	0.00019	0.00256		0.00143	0.02560
4	0.00052	0.00939		0.00138	0.13333
5	0.00038	0.02098		0.00145	0.02233
6	0.00125	0.02899		0.00109	0.05769
7	0.00039	0.06061		0.00073	0.03730
8	0.00024	0.03175		0.00081	0.03590
9	0.00097	0.05405		0.00122	0.02355
10	0.00102	0.04493		0.00056	0.02368
11	0.00034	0.02098		0.00124	0.02992
12	0.00035	0.02098		0.00145	0.12963
13	0.00054	0.00844		0.00182	0.05385
14	0.00049	0.00797		0.00116	0.01852
15	0.00050	0.00797		0.00060	0.06410
16	0.00054	0.03195		0.00168	0.05615
17	0.00060	0.03111		0.00080	0.12963
18	0.00091	0.03011		0.00099	0.02210
19	0.00103	0.05714		0.00098	0.05325
20	0.00104	0.06667		0.00101	0.02705
21	0.00094	0.01871		0.00122	0.04808
22	0.00053	0.02665		0.00079	0.24000
23	0.00049	0.01042		0.00138	0.13125
24	0.00052	0.05714		0.00115	0.11571
25	0.00052	0.04286		0.00103	0.04762

Table C.2 Average and maximum similarity score of query by use case structure and keywords (1) for subject B (Continued)

Average and maximum similarity score of all queries for topic B					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
26	0.00058	0.03509		0.00069	0.05442
27	0.00052	0.02987		0.00107	0.15657
28	0.00053	0.01588		0.00152	0.05385
29	0.00052	0.03218		0.00147	0.09333
30	0.00072	0.01818		0.00061	0.04762
31	0.00052	0.05714		0.00119	0.02446
32	0.00060	0.03037		0.00094	0.02990
33	0.00049	0.03529		0.00094	0.01920
34	0.00039	0.01556		0.00072	0.16000
35	0.00080	0.01027		0.00156	0.07180
36	0.00062	0.03111		0.00143	0.19445
37	0.00057	0.01357		0.00068	0.02392
38	0.00060	0.00844		0.00097	0.02244
39	0.00079	0.01818		0.00146	0.03526
40	0.00093	0.05217		0.00077	0.03500
Average	**0.00061**	**0.02893**		**0.00114**	**0.06586**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.3 Average and maximum similarity score of query by use case structure and keywords (1) for subject C

Average and maximum similarity score of all queries for topic C					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
1	0.00058	0.01563		0.00116	0.01685
2	0.00043	0.00990		0.00066	0.03114
3	0.00050	0.00617		0.00068	0.03556
4	0.00041	0.02885		0.00084	0.15000
5	0.00045	0.01588		0.00100	0.01313
6	0.00075	0.02597		0.00063	0.01746
7	0.00077	0.05366		0.00094	0.02881
8	0.00035	0.03158		0.00049	0.04115
9	0.00044	0.04132		0.00046	0.02881
10	0.00086	0.05455		0.00068	0.10000
11	0.00040	0.03333		0.00084	0.02299
12	0.00042	0.00469		0.00058	0.02400
13	0.00065	0.03889		0.00068	0.02853
14	0.00059	0.01026		0.00113	0.13222
15	0.00063	0.01653		0.00054	0.01293
16	0.00058	0.01563		0.00111	0.01643
17	0.00056	0.01563		0.00153	0.04327
18	0.00046	0.07407		0.00065	0.01345
19	0.00074	0.01975		0.00063	0.01257
20	0.00064	0.02604		0.00095	0.02527
21	0.00061	0.01563		0.00066	0.07297
22	0.00054	0.01961		0.00077	0.01904
23	0.00073	0.07692		0.00112	0.10000
24	0.00073	0.01852		0.00072	0.05440
25	0.00087	0.03011		0.00048	0.01500

Table C.3 Average and maximum similarity score of query by use case structure and keywords (1) for subject C (Continued)

Average and maximum similarity score of all queries for topic C					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
26	0.00019	0.00444		0.00087	0.03788
27	0.00034	0.01667		0.00028	0.00475
28	0.00088	0.06838		0.00077	0.04978
29	0.00069	0.00717		0.00058	0.05455
30	0.00023	0.00436		0.00073	0.01789
31	0.00050	0.00617		0.00046	0.03601
32	0.00023	0.00183		0.00059	0.02830
33	0.00023	0.00436		0.00148	0.03241
34	0.00042	0.00414		0.00168	0.03537
35	0.00107	0.01767		0.00091	0.02332
36	0.00071	0.01324		0.00127	0.02317
37	0.00040	0.00884		0.00090	0.02258
38	0.00097	0.00748		0.00108	0.03295
39	0.00087	0.06924		0.00100	0.04545
40	0.00068	0.01257		0.00177	0.10000
Average	**0.00058**	**0.02364**		**0.00086**	**0.04001**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.4 Average and maximum similarity score of query by use case structure and keywords (1) for subject D

Average and maximum similarity score of all queries for topic D					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
1	0.00018	0.00941		0.00132	0.09356
2	0.00014	0.02139		0.00164	0.25000
3	0.00039	0.05000		0.00169	0.08260
4	0.00037	0.01399		0.00051	0.10667
5	0.00022	0.01282		0.00120	0.08000
6	0.00089	0.02548		0.00051	0.10667
7	0.00058	0.05000		0.00000	0.00000
8	0.00059	0.06897		0.00083	0.03028
9	0.00072	0.08696		0.00042	0.06058
10	0.00041	0.03448		0.00075	0.03263
11	0.00024	0.00691		0.00113	0.04648
12	0.00084	0.10247		0.00140	0.11657
13	0.00088	0.06838		0.00104	0.04779
14	0.00050	0.04348		0.00152	0.05442
15	0.00040	0.01649		0.00081	0.05000
16	0.00068	0.03682		0.00128	0.11893
17	0.00066	0.01335		0.00069	0.02361
18	0.00037	0.02381		0.00145	0.05966
19	0.00045	0.06897		0.00091	0.05818
20	0.00068	0.01160		0.00092	0.05185
21	0.00043	0.00766		0.00081	0.07272
22	0.00056	0.06667		0.00059	0.02797
23	0.00023	0.00666		0.00144	0.11317
24	0.00023	0.00666		0.00146	0.12500
25	0.00000	0.00000		0.00113	0.11111

Table C.4 Average and maximum similarity score of query by use case structure and keywords (1) for subject D (Continued)

Average and maximum similarity score of all queries for topic D					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
26	0.00026	0.01441		0.00062	0.01012
27	0.00024	0.03704		0.00025	0.02307
28	0.00030	0.03333		0.00100	0.03729
29	0.00021	0.01253		0.00115	0.03729
30	0.00045	0.04000		0.00123	0.03236
31	0.00047	0.07596		0.00085	0.04661
32	0.00045	0.07419		0.00110	0.03526
33	0.00052	0.07407		0.00090	0.02712
34	0.00048	0.01361		0.00104	0.06089
35	0.00053	0.03997		0.00148	0.12182
36	0.00050	0.03000		0.00029	0.01036
37	0.00063	0.04482		0.00072	0.05556
38	0.00064	0.03529		0.00088	0.03200
39	0.00065	0.02857		0.00085	0.05455
40	0.00057	0.02778		0.00133	0.06804
Average	**0.00046**	**0.03587**		**0.00098**	**0.06432**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.5 Average and maximum similarity score of query by use case structure and keywords (1) for subject E

Average and maximum similarity score of all queries for topic E					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
1	0.00041	0.04000		0.00155	0.06452
2	0.00041	0.00489		0.00101	0.03728
3	0.00042	0.01361		0.00105	0.04115
4	0.00039	0.02098		0.00165	0.09135
5	0.00055	0.04263		0.00095	0.05333
6	0.00077	0.03571		0.00155	0.09135
7	0.00085	0.08318		0.00092	0.03728
8	0.00053	0.03997		0.00154	0.11429
9	0.00069	0.06061		0.00113	0.03232
10	0.00027	0.00799		0.00101	0.03236
11	0.00053	0.00828		0.00093	0.03111
12	0.00047	0.00855		0.00096	0.02407
13	0.00050	0.01231		0.00117	0.04032
14	0.00085	0.00772		0.00140	0.03278
15	0.00055	0.01894		0.00036	0.03467
16	0.00081	0.00819		0.00172	0.01916
17	0.00074	0.03762		0.00088	0.02859
18	0.00122	0.09108		0.00125	0.02299
19	0.00074	0.03997		0.00072	0.01562
20	0.00098	0.06897		0.00066	0.01314
21	0.00045	0.03377		0.00125	0.07613
22	0.00051	0.03704		0.00124	0.04630
23	0.00050	0.02564		0.00091	0.02616
24	0.00047	0.00799		0.00177	0.28571
25	0.00028	0.00828		0.00065	0.02386

Table C.5 Average and maximum similarity score of query by use case structure and keywords (1) for subject E (Continued)

Average and maximum similarity score of all queries for topic E					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
26	0.00071	0.07407		0.00112	0.01876
27	0.00052	0.01333		0.00016	0.00497
28	0.00050	0.00894		0.00125	0.02299
29	0.00029	0.00860		0.00088	0.01673
30	0.00050	0.00828		0.00088	0.03169
31	0.00077	0.03997		0.00154	0.02768
32	0.00050	0.01299		0.00090	0.03097
33	0.00048	0.00799		0.00122	0.02874
34	0.00084	0.11803		0.00118	0.15385
35	0.00073	0.03997		0.00165	0.06061
36	0.00097	0.03762		0.00147	0.02397
37	0.00092	0.03762		0.00140	0.04132
38	0.00085	0.06274		0.00135	0.03244
39	0.00044	0.00799		0.00091	0.05329
40	0.00089	0.01406		0.00078	0.02556
Average	**0.00062**	**0.03140**		**0.00112**	**0.04724**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.6 Average and maximum similarity score of query by use case structure and keywords (2) for subject A

Average and maximum similarity score of all queries for topic A					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
1	0.00057	0.01355		0.00177	0.07010
2	0.00067	0.01678		0.00171	0.05405
3	0.00086	0.05198		0.00155	0.09795
4	0.00084	0.02453		0.00140	0.05313
5	0.00039	0.02317		0.00122	0.10714
6	0.00090	0.01815		0.00121	0.03457
7	0.00026	0.01802		0.00133	0.12162
8	0.00037	0.01277		0.00140	0.09091
9	0.00094	0.01860		0.00198	0.06049
10	0.00108	0.01859		0.00157	0.04033
11	0.00085	0.01622		0.00211	0.04865
12	0.00058	0.00872		0.00116	0.04210
13	0.00096	0.02096		0.00135	0.02542
14	0.00085	0.08696		0.00090	0.01881
15	0.00062	0.01430		0.00117	0.05442
16	0.00142	0.06397		0.00083	0.01748
17	0.00094	0.06667		0.00128	0.05333
18	0.00096	0.02429		0.00085	0.02623
19	0.00080	0.06897		0.00131	0.05584
20	0.00074	0.06061		0.00147	0.10000
21	0.00062	0.00872		0.00116	0.03748
22	0.00046	0.01531		0.00206	0.06636
23	0.00081	0.01943		0.00193	0.08523
24	0.00054	0.07519		0.00086	0.04806
25	0.00061	0.03571		0.00071	0.04103

Table C.6 Average and maximum similarity score of query by use case structure and keywords (2) for subject A (Continued)

Average and maximum similarity score of all queries for topic A					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
26	0.00074	0.01294		0.00107	0.03483
27	0.00078	0.03663		0.00063	0.02622
28	0.00081	0.02099		0.00125	0.03203
29	0.00032	0.03333		0.00144	0.23834
30	0.00125	0.01700		0.00080	0.01909
31	0.00037	0.00823		0.00064	0.04947
32	0.00022	0.00937		0.00095	0.03795
33	0.00050	0.03241		0.00056	0.04375
34	0.00081	0.01528		0.00142	0.04815
35	0.00022	0.00178		0.00073	0.01749
36	0.00083	0.01481		0.00091	0.02889
37	0.00057	0.01355		0.00142	0.06667
38	0.00044	0.00799		0.00111	0.05442
39	0.00040	0.00752		0.00165	0.06531
40	0.00077	0.00816		0.00043	0.00904
Average	**0.00069**	**0.02605**		**0.00123**	**0.05556**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.7 Average and maximum similarity score of query by use case structure and keywords (2) for subject B

Average and maximum similarity score of all queries for topic B					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
1	0.00090	0.01492		0.00185	0.05385
2	0.00064	0.00530		0.00154	0.03191
3	0.00110	0.01023		0.00143	0.02560
4	0.00089	0.03011		0.00138	0.13333
5	0.00075	0.00797		0.00145	0.02233
6	0.00049	0.01282		0.00109	0.05769
7	0.00067	0.03030		0.00073	0.03730
8	0.00102	0.04955		0.00081	0.03590
9	0.00084	0.01064		0.00122	0.02355
10	0.00031	0.00844		0.00056	0.02368
11	0.00071	0.01108		0.00124	0.02992
12	0.00071	0.03111		0.00145	0.12963
13	0.00094	0.01492		0.00182	0.05385
14	0.00114	0.00892		0.00116	0.01852
15	0.00060	0.07500		0.00060	0.06410
16	0.00113	0.01741		0.00168	0.05615
17	0.00073	0.03111		0.00080	0.12963
18	0.00095	0.06897		0.00099	0.02210
19	0.00079	0.02149		0.00098	0.05325
20	0.00054	0.00910		0.00101	0.02705
21	0.00079	0.01388		0.00122	0.04808
22	0.00052	0.05733		0.00079	0.24000
23	0.00076	0.02785		0.00138	0.13125
24	0.00048	0.01808		0.00115	0.11571
25	0.00101	0.09524		0.00103	0.04762

Table C.7 Average and maximum similarity score of query by use case structure and keywords (2) for subject B (Continued)

Average and maximum similarity score of all queries for topic B					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
26	0.00058	0.03509		0.00069	0.05442
27	0.00090	0.03333		0.00107	0.15657
28	0.00074	0.01492		0.00152	0.05385
29	0.00077	0.03111		0.00147	0.09333
30	0.00058	0.03509		0.00061	0.04762
31	0.00061	0.00797		0.00119	0.02446
32	0.00054	0.00910		0.00094	0.02990
33	0.00069	0.00797		0.00094	0.01920
34	0.00070	0.05472		0.00072	0.16000
35	0.00076	0.01538		0.00156	0.07180
36	0.00072	0.03111		0.00143	0.19445
37	0.00054	0.00910		0.00068	0.02392
38	0.00093	0.03175		0.00097	0.02244
39	0.00083	0.03997		0.00146	0.03526
40	0.00088	0.01522		0.00077	0.03500
Average	**0.00075**	**0.02634**		**0.00114**	**0.06586**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.8 Average and maximum similarity score of query by use case structure and keywords (2) for subject C

Average and maximum similarity score of all queries for topic C					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
1	0.00080	0.00558		0.00116	0.01685
2	0.00064	0.01026		0.00066	0.03114
3	0.00043	0.00990		0.00068	0.03556
4	0.00048	0.02308		0.00084	0.15000
5	0.00081	0.01603		0.00100	0.01313
6	0.00089	0.01064		0.00063	0.01746
7	0.00042	0.00469		0.00094	0.02881
8	0.00044	0.00823		0.00049	0.04115
9	0.00046	0.00562		0.00046	0.02881
10	0.00050	0.04000		0.00068	0.10000
11	0.00081	0.01515		0.00084	0.02299
12	0.00082	0.03125		0.00058	0.02400
13	0.00078	0.01515		0.00068	0.02853
14	0.00042	0.03889		0.00113	0.13222
15	0.00080	0.01515		0.00054	0.01293
16	0.00085	0.00625		0.00111	0.01643
17	0.00097	0.01471		0.00153	0.04327
18	0.00096	0.01506		0.00065	0.01345
19	0.00105	0.06667		0.00063	0.01257
20	0.00101	0.01438		0.00095	0.02527
21	0.00024	0.01678		0.00066	0.07297
22	0.00062	0.00794		0.00077	0.01904
23	0.00079	0.02972		0.00112	0.10000
24	0.00066	0.01852		0.00072	0.05440
25	0.00081	0.03509		0.00048	0.01500

Table C.8 Average and maximum similarity score of query by use case structure and keywords (2) for subject C (Continued)

Average and maximum similarity score of all queries for topic C					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
26	0.00046	0.04329		0.00087	0.03788
27	0.00082	0.03509		0.00028	0.00475
28	0.00090	0.03571		0.00077	0.04978
29	0.00118	0.03704		0.00058	0.05455
30	0.00081	0.03704		0.00073	0.01789
31	0.00023	0.00436		0.00046	0.03601
32	0.00044	0.00545		0.00059	0.02830
33	0.00048	0.00464		0.00148	0.03241
34	0.00073	0.00599		0.00168	0.03537
35	0.00023	0.00183		0.00091	0.02332
36	0.00062	0.00574		0.00127	0.02317
37	0.00065	0.00543		0.00090	0.02258
38	0.00063	0.00909		0.00108	0.03295
39	0.00080	0.03953		0.00100	0.04545
40	0.00119	0.08879		0.00177	0.10000
Average	**0.00069**	**0.02084**		**0.00086**	**0.04001**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.9 Average and maximum similarity score of query by use case structure and keywords (2) for subject D

Average and maximum similarity score of all queries for topic D					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
1	0.00051	0.02228		0.00132	0.09356
2	0.00046	0.03814		0.00164	0.25000
3	0.00046	0.01334		0.00169	0.08260
4	0.00040	0.03243		0.00051	0.10667
5	0.00050	0.01905		0.00120	0.08000
6	0.00022	0.01778		0.00051	0.10667
7	0.00000	0.00000		0.00000	0.00000
8	0.00042	0.00643		0.00083	0.03028
9	0.00070	0.04167		0.00042	0.06058
10	0.00048	0.03750		0.00075	0.03263
11	0.00085	0.01826		0.00113	0.04648
12	0.00113	0.03487		0.00140	0.11657
13	0.00076	0.02198		0.00104	0.04779
14	0.00073	0.03750		0.00152	0.05442
15	0.00092	0.06677		0.00081	0.05000
16	0.00066	0.03682		0.00128	0.11893
17	0.00047	0.00621		0.00069	0.02361
18	0.00066	0.01541		0.00145	0.05966
19	0.00093	0.03529		0.00091	0.05818
20	0.00080	0.03529		0.00092	0.05185
21	0.00042	0.02090		0.00081	0.07272
22	0.00022	0.00643		0.00059	0.02797
23	0.00093	0.05882		0.00144	0.11317
24	0.00088	0.06838		0.00146	0.12500
25	0.00069	0.09524		0.00113	0.11111

Table C.9 Average and maximum similarity score of query by use case structure and keywords (2) for subject D (Continued)

Average and maximum similarity score of all queries for topic D					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
26	0.00065	0.01295		0.00062	0.01012
27	0.00063	0.07072		0.00025	0.02307
28	0.00082	0.00871		0.00100	0.03729
29	0.00090	0.02034		0.00115	0.03729
30	0.00116	0.06667		0.00123	0.03236
31	0.00047	0.07596		0.00085	0.04661
32	0.00096	0.06897		0.00110	0.03526
33	0.00079	0.00688		0.00090	0.02712
34	0.00069	0.02955		0.00104	0.06089
35	0.00110	0.05167		0.00148	0.12182
36	0.00044	0.00694		0.00029	0.01036
37	0.00046	0.03814		0.00072	0.05556
38	0.00061	0.01119		0.00088	0.03200
39	0.00037	0.02870		0.00085	0.05455
40	0.00100	0.05217		0.00133	0.06804
Average	**0.00066**	**0.03341**		**0.00098**	**0.06432**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.10 Average and maximum similarity score of query by use case structure and keywords (2) for subject E

Average and maximum similarity score of all queries for topic E					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
1	0.00055	0.01894		0.00155	0.06452
2	0.00051	0.00828		0.00101	0.03728
3	0.00049	0.00828		0.00105	0.04115
4	0.00074	0.03997		0.00165	0.09135
5	0.00064	0.01515		0.00095	0.05333
6	0.00077	0.03997		0.00155	0.09135
7	0.00047	0.03226		0.00092	0.03728
8	0.00054	0.03077		0.00154	0.11429
9	0.00066	0.00881		0.00113	0.03232
10	0.00087	0.05556		0.00101	0.03236
11	0.00067	0.01143		0.00093	0.03111
12	0.00066	0.01018		0.00096	0.02407
13	0.00072	0.01824		0.00117	0.04032
14	0.00071	0.01049		0.00140	0.03278
15	0.00063	0.03529		0.00036	0.03467
16	0.00122	0.01119		0.00172	0.01916
17	0.00091	0.01487		0.00088	0.02859
18	0.00127	0.07358		0.00125	0.02299
19	0.00069	0.02367		0.00072	0.01562
20	0.00150	0.06199		0.00066	0.01314
21	0.00077	0.03997		0.00125	0.07613
22	0.00068	0.03268		0.00124	0.04630
23	0.00063	0.05882		0.00091	0.02616
24	0.00070	0.06061		0.00177	0.28571
25	0.00043	0.03509		0.00065	0.02386

Table C.10 Average and maximum similarity score of query by use case structure and keywords (2) for subject E (Continued)

Average and maximum similarity score of all queries for topic E					
	Query by use case keywords			Query by use case structure	
Query	Average	Maximum		Average	Maximum
26	0.00069	0.00771		0.00112	0.01876
27	0.00067	0.03846		0.00016	0.00497
28	0.00127	0.07358		0.00125	0.02299
29	0.00126	0.03448		0.00088	0.01673
30	0.00029	0.00860		0.00088	0.03169
31	0.00106	0.01161		0.00154	0.02768
32	0.00050	0.00894		0.00090	0.03097
33	0.00053	0.00828		0.00122	0.02874
34	0.00083	0.12541		0.00118	0.15385
35	0.00116	0.06897		0.00165	0.06061
36	0.00069	0.00799		0.00147	0.02397
37	0.00096	0.04725		0.00140	0.04132
38	0.00093	0.00923		0.00135	0.03244
39	0.00093	0.03762		0.00091	0.05329
40	0.00048	0.00950		0.00078	0.02556
Average	**0.00077**	**0.03134**		**0.00112**	**0.04724**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.11 Recall, precision, and harmonic mean of query by use case structure and keywords (1) for subject A

Recall, precision, and harmonic mean of all queries for topic A							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
1	0.0476	0.0435	0.0455		0.4444	0.8485	0.5833
2	0.0000	0.0000	0.0000		0.4921	0.9394	0.6458
3	0.2698	0.8095	0.4048		0.2540	0.8889	0.3951
4	0.1905	0.8000	0.3077		0.4286	0.7941	0.5567
5	0.3810	0.7059	0.4948		0.1746	0.7857	0.2857
6	0.7937	0.6173	0.6944		0.7460	0.8393	0.7899
7	0.3810	0.4068	0.3934		0.0476	0.6000	0.0882
8	0.7302	0.6216	0.6715		0.1587	0.6250	0.2532
9	0.6349	0.4301	0.5128		0.5556	0.8537	0.6731
10	0.0476	0.1765	0.0750		0.7937	0.8772	0.8333
11	0.3810	0.8889	0.5333		0.5873	0.5692	0.5781
12	0.0476	0.1034	0.0652		0.4603	0.8286	0.5918
13	0.0635	0.8000	0.1176		0.5079	0.4706	0.4885
14	0.0476	0.0968	0.0638		0.5079	0.5079	0.5079
15	0.6825	0.8431	0.7544		0.3333	0.8400	0.4773
16	0.2063	0.3611	0.2626		0.8889	0.6829	0.7724
17	0.0794	0.0704	0.0746		0.5079	0.7805	0.6154
18	0.0476	0.1034	0.0652		0.4603	0.6905	0.5524
19	0.3810	0.2667	0.3137		0.5079	0.5000	0.5039
20	0.0476	0.7500	0.0896		0.4603	0.7838	0.5800
21	0.0000	0.0000	0.0000		0.4286	0.5400	0.4779
22	0.1905	0.8000	0.3077		0.1905	1.0000	0.3200
23	0.2698	0.8095	0.4048		0.6032	0.5758	0.5891
24	0.3810	0.8889	0.5333		0.2540	0.8000	0.3855
25	0.0159	0.0833	0.0267		0.0794	0.5556	0.1389

Table C.11 Recall, precision, and harmonic mean of query by use case structure and keywords (1) for subject A (Continued)

Recall, precision, and harmonic mean of all queries for topic A							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
26	0.0476	0.7500	0.0896		0.6032	0.8261	0.6972
27	0.2540	0.7273	0.3765		0.3810	0.8889	0.5333
28	0.0476	0.6000	0.0882		0.5556	0.8974	0.6863
29	0.2063	0.5200	0.2955		0.0317	0.6667	0.0606
30	0.0476	0.2727	0.0811		0.7937	0.8475	0.8197
31	0.0000	0.0000	0.0000		0.1587	0.5882	0.2500
32	0.0476	0.1034	0.0652		0.3492	0.9565	0.5116
33	0.2222	0.4667	0.3011		0.0476	0.4286	0.0857
34	0.2381	0.5769	0.3371		0.6825	0.5181	0.5890
35	0.0635	0.1026	0.0784		0.1429	0.1875	0.1622
36	0.2698	0.3617	0.3091		0.2063	0.1566	0.1781
37	0.0952	1.0000	0.1739		0.5079	0.8000	0.6214
38	0.1429	0.8182	0.2432		0.2222	0.4516	0.2979
39	0.6032	0.8636	0.7103		0.3492	0.8800	0.5000
40	0.5556	0.5224	0.5385		0.3968	0.5102	0.4464
Average	**0.2290**	**0.4791**	**0.2725**		**0.4075**	**0.6945**	**0.4781**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.12 Recall, precision, and harmonic mean of query by use case structure and keywords (1) for subject B

Recall, precision, and harmonic mean of all queries for topic B							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
1	0.1250	0.2353	0.1633		0.4688	0.3000	0.3659
2	0.3125	0.2083	0.2500		0.5313	0.1560	0.2411
3	0.0625	0.0426	0.0506		0.6250	0.1754	0.2740
4	0.2500	0.1290	0.1702		0.4375	0.2414	0.3111
5	0.1563	0.0806	0.1064		0.2813	0.1200	0.1682
6	0.3438	0.0991	0.1538		0.3750	0.3158	0.3429
7	0.0000	0.0000	0.0000		0.1563	0.1389	0.1471
8	0.0625	0.4000	0.1081		0.2813	0.2368	0.2571
9	0.3438	0.3793	0.3607		0.3125	0.1190	0.1724
10	0.2500	0.0611	0.0982		0.1563	0.2500	0.1923
11	0.2188	0.1321	0.1647		0.2813	0.2093	0.2400
12	0.2188	0.2414	0.2295		0.3125	0.2439	0.2740
13	0.2813	0.1098	0.1579		0.5000	0.3137	0.3855
14	0.2500	0.1053	0.1481		0.5313	0.1504	0.2345
15	0.2813	0.1552	0.2000		0.2500	1.0000	0.4000
16	0.0938	0.1500	0.1154		0.4375	0.2373	0.3077
17	0.2188	0.3500	0.2692		0.1875	0.3750	0.2500
18	0.4375	0.2500	0.3182		0.4063	0.2241	0.2889
19	0.6875	0.2683	0.3860		0.5313	0.4146	0.4658
20	0.5000	0.2078	0.2936		0.3125	0.1887	0.2353
21	0.5313	0.1491	0.2329		0.5313	0.2833	0.3696
22	0.1563	0.0769	0.1031		0.0625	0.2500	0.1000
23	0.4063	0.1625	0.2321		0.1563	0.0962	0.1190
24	0.4375	0.3182	0.3684		0.0938	0.5000	0.1579
25	0.2500	0.2162	0.2319		0.4375	0.4516	0.4444

Table C.12 Recall, precision, and harmonic mean of query by use case structure and keywords (1) for subject B (Continued)

Recall, precision, and harmonic mean of all queries for topic B							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
26	0.2813	0.2250	0.2500		0.1875	0.1935	0.1905
27	0.0625	0.0800	0.0702		0.1250	0.1481	0.1356
28	0.3750	0.2553	0.3038		0.3438	0.1746	0.2316
29	0.3125	0.2564	0.2817		0.4375	0.3684	0.4000
30	0.3438	0.1467	0.2056		0.1875	0.1935	0.1905
31	0.1563	0.3333	0.2128		0.2188	0.1321	0.1647
32	0.1563	0.2500	0.1923		0.3750	0.2353	0.2892
33	0.0938	0.0536	0.0682		0.2813	0.1125	0.1607
34	0.1250	0.2222	0.1600		0.0313	0.0833	0.0455
35	0.4063	0.1477	0.2167		0.3750	0.4615	0.4138
36	0.5313	0.2931	0.3778		0.2813	0.2195	0.2466
37	0.5625	0.2250	0.3214		0.2813	0.3913	0.3273
38	0.2813	0.1125	0.1607		0.2813	0.1500	0.1957
39	0.4063	0.1413	0.2097		0.1875	0.1250	0.1500
40	0.1875	0.1818	0.1846		0.4688	0.2542	0.3297
Average	**0.2789**	**0.1863**	**0.2032**		**0.3180**	**0.2559**	**0.2554**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.13 Recall, precision, and harmonic mean of query by use case structure and keywords (1) for subject C

Recall, precision, and harmonic mean of all queries for topic C							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
1	0.7857	0.1222	0.2115		0.8571	0.1034	0.1846
2	0.7143	0.1515	0.2500		0.7857	0.2075	0.3284
3	0.7143	0.0990	0.1739		0.7857	0.7333	0.7586
4	0.7143	0.2083	0.3226		0.7857	0.9167	0.8462
5	0.1429	0.0308	0.0506		0.8571	0.1043	0.1860
6	0.7143	0.1408	0.2353		0.7857	0.2444	0.3729
7	0.7857	0.1279	0.2200		0.7857	0.4583	0.5789
8	0.0000	0.0000	0.0000		0.7857	0.9167	0.8462
9	0.1429	0.0426	0.0656		0.7857	0.9167	0.8462
10	0.7143	0.1087	0.1887		0.7857	0.9167	0.8462
11	0.2857	0.0656	0.1067		0.7857	0.1447	0.2444
12	0.7857	0.1310	0.2245		0.7857	0.1667	0.2750
13	0.7143	0.2000	0.3125		0.7857	0.1528	0.2558
14	0.7857	0.1111	0.1947		0.7857	0.2444	0.3729
15	0.7857	0.1111	0.1947		0.7857	0.1642	0.2716
16	0.7857	0.1222	0.2115		0.7857	0.0940	0.1679
17	0.7143	0.1754	0.2817		0.7857	0.1058	0.1864
18	0.7143	0.2222	0.3390		0.7857	0.1746	0.2857
19	0.7857	0.1774	0.2895		0.8571	0.1101	0.1951
20	0.8571	0.1200	0.2105		0.8571	0.1622	0.2727
21	0.7143	0.0971	0.1709		0.0000	0.0000	0.0000
22	0.2143	0.0448	0.0741		0.7857	0.1833	0.2973
23	0.7143	0.0990	0.1739		0.1429	0.1333	0.1379
24	0.3571	0.0417	0.0746		0.3571	0.1220	0.1818
25	0.2143	0.0385	0.0652		0.7857	0.3235	0.4583

Table C.13 Recall, precision, and harmonic mean of query by use case structure and keywords (1) for subject C (Continued)

Recall, precision, and harmonic mean of all queries for topic C							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
26	0.2857	0.0816	0.1270		0.7857	0.2444	0.3729
27	0.3571	0.2381	0.2857		0.8571	0.1905	0.3117
28	0.3571	0.2174	0.2703		0.7857	0.1930	0.3099
29	0.8571	0.0805	0.1472		0.7857	0.2444	0.3729
30	0.7143	0.2273	0.3448		0.7857	0.1864	0.3014
31	0.7143	0.0990	0.1739		0.7857	1.0000	0.8800
32	0.2143	0.0469	0.0769		0.8571	0.1739	0.2892
33	0.7143	0.2273	0.3448		0.7857	0.1122	0.1964
34	0.5000	0.0588	0.1053		0.7857	0.1068	0.1880
35	0.8571	0.0736	0.1356		0.2143	0.0625	0.0968
36	0.7857	0.1250	0.2157		0.8571	0.1319	0.2286
37	0.7143	0.1282	0.2174		0.8571	0.1538	0.2609
38	0.7857	0.0719	0.1317		0.8571	0.1500	0.2553
39	0.7857	0.0853	0.1538		0.7857	0.2619	0.3929
40	0.7143	0.0990	0.1739		0.8571	0.1277	0.2222
Average	**0.6054**	**0.1162**	**0.1887**		**0.7429**	**0.2785**	**0.3519**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.14 Recall, precision, and harmonic mean of query by use case structure and keywords (1) for subject D

Recall, precision, and harmonic mean of all queries for topic D							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
1	0.0417	0.0526	0.0465		0.4583	0.7857	0.5789
2	0.0417	0.1667	0.0667		0.5000	0.9231	0.6486
3	0.1250	0.6000	0.2069		0.4583	0.7857	0.5789
4	0.1250	0.1111	0.1176		0.0417	0.5000	0.0769
5	0.0833	0.2000	0.1176		0.0417	0.0196	0.0267
6	0.3333	0.1667	0.2222		0.0417	0.5000	0.0769
7	0.3333	0.5333	0.4103		0.0000	0.0000	0.0000
8	0.3333	0.6667	0.4444		0.5833	0.1944	0.2917
9	0.0417	0.1667	0.0667		0.5417	0.8125	0.6500
10	0.1250	0.6000	0.2069		0.5833	0.8235	0.6829
11	0.6667	0.5333	0.5926		0.0833	0.0294	0.0435
12	0.0833	0.0488	0.0615		0.6250	0.2083	0.3125
13	0.0833	0.0952	0.0889		0.6667	0.6957	0.6809
14	0.0833	0.2000	0.1176		0.7917	0.7600	0.7755
15	0.7083	0.4595	0.5574		0.2083	0.0893	0.1250
16	0.7083	0.3542	0.4722		0.5833	0.2593	0.3590
17	0.6667	0.2078	0.3168		0.5833	0.1918	0.2887
18	0.6667	0.5000	0.5714		0.5833	0.6667	0.6222
19	0.0833	0.2222	0.1212		0.5833	0.2295	0.3294
20	0.6250	0.2113	0.3158		0.0833	0.2500	0.1250
21	0.6250	0.2308	0.3371		0.5833	0.8235	0.6829
22	0.6667	0.4848	0.5614		0.5833	0.8750	0.7000
23	0.6667	0.5333	0.5926		0.5833	0.7368	0.6512
24	0.6667	0.5333	0.5926		0.5833	0.5385	0.5600
25	0.0000	0.0000	0.0000		0.2083	0.4545	0.2857

Table C.14 Recall, precision, and harmonic mean of query by use case structure and keywords (1) for subject D (Continued)

Recall, precision, and harmonic mean of all queries for topic D							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
26	0.2500	0.4286	0.3158		0.2083	0.0610	0.0943
27	0.1250	0.7500	0.2143		0.5417	0.8125	0.6500
28	0.0833	0.2500	0.1250		0.6667	0.1928	0.2991
29	0.1250	0.1875	0.1500		0.8750	0.4200	0.5676
30	0.2083	0.2381	0.2222		0.2917	0.2188	0.2500
31	0.6667	0.5333	0.5926		0.4583	0.9167	0.6111
32	0.0833	0.0317	0.0460		0.2500	0.2143	0.2308
33	0.1667	0.0625	0.0909		0.5833	0.2692	0.3684
34	0.8333	0.5000	0.6250		0.5833	0.7368	0.6512
35	0.0833	0.0500	0.0625		0.5833	0.3500	0.4375
36	0.6667	0.4571	0.5424		0.5417	0.4483	0.4906
37	0.6667	0.4000	0.5000		0.5000	0.9231	0.6486
38	0.2500	0.1111	0.1538		0.5417	0.5417	0.5417
39	0.2917	0.1207	0.1707		0.6250	0.8824	0.7317
40	0.8750	0.2917	0.4375		0.6250	0.6818	0.6522
Average	**0.3490**	**0.3073**	**0.2863**		**0.4615**	**0.4956**	**0.4344**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.15 Recall, precision, and harmonic mean of query by use case structure and keywords (1) for subject E

Recall, precision, and harmonic mean of all queries for topic E							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
1	0.3750	0.4615	0.4138		0.5625	0.4500	0.5000
2	0.1875	0.0349	0.0588		0.4375	0.3500	0.3889
3	0.3125	0.1351	0.1887		0.4375	0.3684	0.4000
4	0.3125	0.1250	0.1786		0.4375	0.1077	0.1728
5	0.6875	0.2750	0.3929		0.3750	0.4615	0.4138
6	0.3750	0.2609	0.3077		0.5625	0.3913	0.4615
7	0.6875	0.2444	0.3607		0.4375	0.3889	0.4118
8	0.6875	0.2750	0.3929		0.4375	0.3333	0.3784
9	0.1250	0.0392	0.0597		0.5000	0.1096	0.1798
10	0.6250	0.2632	0.3704		0.5000	0.3333	0.4000
11	0.8125	0.2167	0.3421		0.6875	0.1964	0.3056
12	0.6250	0.1515	0.2439		0.6250	0.1754	0.2740
13	0.6875	0.2245	0.3385		0.6875	0.2037	0.3143
14	0.6875	0.0859	0.1528		0.6250	0.1087	0.1852
15	0.8125	0.3023	0.4407		0.1250	0.4000	0.1905
16	0.8750	0.1728	0.2887		0.8125	0.1140	0.2000
17	0.8125	0.2131	0.3377		0.5625	0.2250	0.3214
18	0.8750	0.1591	0.2692		0.6250	0.1961	0.2985
19	0.7500	0.1290	0.2202		0.6250	0.3448	0.4444
20	0.7500	0.1765	0.2857		0.5625	0.1525	0.2400
21	0.6875	0.2683	0.3860		0.6250	0.3448	0.4444
22	0.0000	0.0000	0.0000		0.6875	0.3548	0.4681
23	0.6250	0.2174	0.3226		0.5000	0.2222	0.3077
24	0.6250	0.1471	0.2381		0.6250	0.3448	0.4444
25	0.6875	0.2821	0.4000		0.5000	0.4000	0.4444

Table C.15 Recall, precision, and harmonic mean of query by use case structure and keywords (1) for subject E (Continued)

Recall, precision, and harmonic mean of all queries for topic E							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
26	0.6250	0.1923	0.2941		0.6250	0.1266	0.2105
27	0.7500	0.2400	0.3636		0.5000	0.3810	0.4324
28	0.7500	0.1446	0.2424		0.6250	0.1961	0.2985
29	0.6875	0.2750	0.3929		0.6250	0.1266	0.2105
30	0.6875	0.1134	0.1947		0.5000	0.4000	0.4444
31	0.9375	0.2679	0.4167		0.5000	0.1013	0.1684
32	0.0625	0.0172	0.0270		0.5000	0.2105	0.2963
33	0.6875	0.1134	0.1947		0.6250	0.3448	0.4444
34	0.6875	0.2750	0.3929		0.5625	0.4286	0.4865
35	0.6875	0.1264	0.2136		0.5625	0.1139	0.1895
36	0.8125	0.1757	0.2889		0.5625	0.0947	0.1622
37	0.6875	0.1486	0.2444		0.5625	0.1837	0.2769
38	0.5625	0.0857	0.1488		0.5625	0.2250	0.3214
39	0.6250	0.1695	0.2667		0.4375	0.1014	0.1647
40	0.5625	0.1184	0.1957		0.4375	0.1750	0.2500
Average	**0.6125**	**0.1831**	**0.2717**		**0.5438**	**0.2572**	**0.3237**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.16 Recall, precision, and harmonic mean of query by use case structure and keywords (2) for subject A

Recall, precision, and harmonic mean of all queries for topic A							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
1	0.6032	0.7451	0.6667		0.4444	0.8485	0.5833
2	0.5079	0.6957	0.5872		0.4921	0.9394	0.6458
3	0.2857	0.5625	0.3789		0.2540	0.8889	0.3951
4	0.5238	0.7333	0.6111		0.4286	0.7941	0.5567
5	0.2857	0.8182	0.4235		0.1746	0.7857	0.2857
6	0.7619	0.6154	0.6809		0.7460	0.8393	0.7899
7	0.0476	0.2727	0.0811		0.0476	0.6000	0.0882
8	0.2698	0.5862	0.3696		0.1587	0.6250	0.2532
9	0.6984	0.5500	0.6154		0.5556	0.8537	0.6731
10	0.9206	0.5524	0.6905		0.7937	0.8772	0.8333
11	0.5873	0.4744	0.5248		0.5873	0.5692	0.5781
12	0.6190	0.5652	0.5909		0.4603	0.8286	0.5918
13	0.6190	0.3980	0.4845		0.5079	0.4706	0.4885
14	0.5556	0.4667	0.5072		0.5079	0.5079	0.5079
15	0.6190	0.5132	0.5612		0.3333	0.8400	0.4773
16	0.9524	0.4380	0.6000		0.8889	0.6829	0.7724
17	0.6032	0.7308	0.6609		0.5079	0.7805	0.6154
18	0.6349	0.6250	0.6299		0.4603	0.6905	0.5524
19	0.6825	0.4624	0.5513		0.5079	0.5000	0.5039
20	0.6032	0.7308	0.6609		0.4603	0.7838	0.5800
21	0.5714	0.3186	0.4091		0.4286	0.5400	0.4779
22	0.1905	0.8000	0.3077		0.1905	1.0000	0.3200
23	0.6349	0.4040	0.4938		0.6032	0.5758	0.5891
24	0.2857	0.8182	0.4235		0.2540	0.8000	0.3855
25	0.0794	0.3333	0.1282		0.0794	0.5556	0.1389

Table C.16 Recall, precision, and harmonic mean of query by use case structure and keywords (2) for subject A (Continued)

Recall, precision, and harmonic mean of all queries for topic A							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
26	0.6825	0.7544	0.7167		0.6032	0.8261	0.6972
27	0.6984	0.8627	0.7719		0.3810	0.8889	0.5333
28	0.8095	0.5795	0.6755		0.5556	0.8974	0.6863
29	0.0317	0.5000	0.0597		0.0317	0.6667	0.0606
30	0.8889	0.6154	0.7273		0.7937	0.8475	0.8197
31	0.2857	0.3750	0.3243		0.1587	0.5882	0.2500
32	0.3810	0.8889	0.5333		0.3492	0.9565	0.5116
33	0.0952	0.4286	0.1558		0.0476	0.4286	0.0857
34	0.6984	0.3793	0.4916		0.6825	0.5181	0.5890
35	0.2540	0.2581	0.2560		0.1429	0.1875	0.1622
36	0.5873	0.2327	0.3333		0.2063	0.1566	0.1781
37	0.6032	0.7451	0.6667		0.5079	0.8000	0.6214
38	0.3333	0.3559	0.3443		0.2222	0.4516	0.2979
39	0.4603	0.8056	0.5859		0.3492	0.8800	0.5000
40	0.6349	0.2703	0.3791		0.3968	0.5102	0.4464
Average	**0.5147**	**0.5565**	**0.4915**		**0.4075**	**0.6945**	**0.4781**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.17 Recall, precision, and harmonic mean of query by use case structure and keywords (2) for subject B

Recall, precision, and harmonic mean of all queries for topic B							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
1	0.6875	0.2558	0.3729		0.4688	0.3000	0.3659
2	0.6563	0.1429	0.2346		0.5313	0.1560	0.2411
3	0.8125	0.1595	0.2667		0.6250	0.1754	0.2740
4	0.5313	0.1789	0.2677		0.4375	0.2414	0.3111
5	0.3750	0.1200	0.1818		0.2813	0.1200	0.1682
6	0.6563	0.3443	0.4516		0.3750	0.3158	0.3429
7	0.3750	0.2105	0.2697		0.1563	0.1389	0.1471
8	0.5625	0.2727	0.3673		0.2813	0.2368	0.2571
9	0.5000	0.1356	0.2133		0.3125	0.1190	0.1724
10	0.2188	0.2000	0.2090		0.1563	0.2500	0.1923
11	0.5000	0.1758	0.2602		0.2813	0.2093	0.2400
12	0.5625	0.2500	0.3462		0.3125	0.2439	0.2740
13	0.9063	0.2816	0.4296		0.5000	0.3137	0.3855
14	0.8125	0.1529	0.2574		0.5313	0.1504	0.2345
15	0.5313	0.6538	0.5862		0.2500	1.0000	0.4000
16	0.7813	0.2577	0.3876		0.4375	0.2373	0.3077
17	0.3125	0.2174	0.2564		0.1875	0.3750	0.2500
18	0.7813	0.2551	0.3846		0.4063	0.2241	0.2889
19	0.6563	0.3500	0.4565		0.5313	0.4146	0.4658
20	0.5000	0.2319	0.3168		0.3125	0.1887	0.2353
21	0.7188	0.2911	0.4144		0.5313	0.2833	0.3696
22	0.0313	0.0189	0.0235		0.0625	0.2500	0.1000
23	0.3125	0.1149	0.1681		0.1563	0.0962	0.1190
24	0.3750	0.3243	0.3478		0.0938	0.5000	0.1579
25	0.4375	0.2029	0.2772		0.4375	0.4516	0.4444

Table C.17 Recall, precision, and harmonic mean of query by use case structure and keywords (2) for subject B (Continued)

Recall, precision, and harmonic mean of all queries for topic B							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
26	0.2813	0.2250	0.2500		0.1875	0.1935	0.1905
27	0.2500	0.1818	0.2105		0.1250	0.1481	0.1356
28	0.4375	0.1556	0.2295		0.3438	0.1746	0.2316
29	0.6563	0.3818	0.4828		0.4375	0.3684	0.4000
30	0.2813	0.2250	0.2500		0.1875	0.1935	0.1905
31	0.2500	0.0941	0.1368		0.2188	0.1321	0.1647
32	0.5000	0.2319	0.3168		0.3750	0.2353	0.2892
33	0.2813	0.0968	0.1440		0.2813	0.1125	0.1607
34	0.1563	0.0893	0.1136		0.0313	0.0833	0.0455
35	0.7188	0.3067	0.4299		0.3750	0.4615	0.4138
36	0.3750	0.1395	0.2034		0.2813	0.2195	0.2466
37	0.5000	0.2319	0.3168		0.2813	0.3913	0.3273
38	0.5625	0.1895	0.2835		0.2813	0.1500	0.1957
39	0.2500	0.1231	0.1649		0.1875	0.1250	0.1500
40	0.5938	0.2000	0.2992		0.4688	0.2542	0.3297
Average	**0.4922**	**0.2168**	**0.2895**		**0.3180**	**0.2559**	**0.2554**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.18 Recall, precision, and harmonic mean of query by use case structure and keywords (2) for subject C

Recall, precision, and harmonic mean of all queries for topic C							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
1	0.9286	0.0751	0.1390		0.8571	0.1034	0.1846
2	0.7857	0.0880	0.1583		0.7857	0.2075	0.3284
3	0.7143	0.1515	0.2500		0.7857	0.7333	0.7586
4	0.7143	0.2041	0.3175		0.7857	0.9167	0.8462
5	0.7857	0.0743	0.1358		0.8571	0.1043	0.1860
6	0.7143	0.0813	0.1460		0.7857	0.2444	0.3729
7	0.7857	0.1310	0.2245		0.7857	0.4583	0.5789
8	0.7857	0.1719	0.2821		0.7857	0.9167	0.8462
9	0.7143	0.1020	0.1786		0.7857	0.9167	0.8462
10	0.7143	0.1961	0.3077		0.7857	0.9167	0.8462
11	0.7143	0.0763	0.1379		0.7857	0.1447	0.2444
12	0.7143	0.0917	0.1626		0.7857	0.1667	0.2750
13	0.7857	0.0866	0.1560		0.7857	0.1528	0.2558
14	0.7143	0.2041	0.3175		0.7857	0.2444	0.3729
15	0.7143	0.0617	0.1136		0.7857	0.1642	0.2716
16	0.7857	0.0625	0.1158		0.7857	0.0940	0.1679
17	0.7143	0.0610	0.1124		0.7857	0.1058	0.1864
18	0.7143	0.0719	0.1307		0.7857	0.1746	0.2857
19	0.8571	0.0642	0.1194		0.8571	0.1101	0.1951
20	0.8571	0.0811	0.1481		0.8571	0.1622	0.2727
21	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000
22	0.7857	0.1019	0.1803		0.7857	0.1833	0.2973
23	0.1429	0.0645	0.0889		0.1429	0.1333	0.1379
24	0.5000	0.0875	0.1489		0.3571	0.1220	0.1818
25	0.7143	0.1538	0.2532		0.7857	0.3235	0.4583

Table C.18 Recall, precision, and harmonic mean of query by use case structure and keywords (2) for subject C (Continued)

Recall, precision, and harmonic mean of all queries for topic C							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
26	0.7143	0.2000	0.3125		0.7857	0.2444	0.3729
27	0.8571	0.1008	0.1805		0.8571	0.1905	0.3117
28	0.7143	0.1220	0.2083		0.7857	0.1930	0.3099
29	0.7857	0.1571	0.2619		0.7857	0.2444	0.3729
30	0.7143	0.1266	0.2151		0.7857	0.1864	0.3014
31	0.7143	0.2273	0.3448		0.7857	1.0000	0.8800
32	0.8571	0.1224	0.2143		0.8571	0.1739	0.2892
33	0.8571	0.1304	0.2264		0.7857	0.1122	0.1964
34	0.8571	0.0923	0.1667		0.7857	0.1068	0.1880
35	0.2143	0.0469	0.0769		0.2143	0.0625	0.0968
36	0.9286	0.1008	0.1818		0.8571	0.1319	0.2286
37	0.8571	0.0816	0.1491		0.8571	0.1538	0.2609
38	0.9286	0.1193	0.2114		0.8571	0.1500	0.2553
39	0.7857	0.1196	0.2075		0.7857	0.2619	0.3929
40	0.7857	0.0853	0.1538		0.8571	0.1277	0.2222
Average	**0.7232**	**0.1094**	**0.1859**		**0.7429**	**0.2785**	**0.3519**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.19 Recall, precision, and harmonic mean of query by use case structure and keywords (2) for subject D

Recall, precision, and harmonic mean of all queries for topic D							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
1	0.6667	0.4103	0.5079		0.4583	0.7857	0.5789
2	0.6667	0.5161	0.5818		0.5000	0.9231	0.6486
3	0.6667	0.5333	0.5926		0.4583	0.7857	0.5789
4	0.3333	0.7273	0.4571		0.0417	0.5000	0.0769
5	0.1250	0.0405	0.0612		0.0417	0.0196	0.0267
6	0.2917	0.7778	0.4242		0.0417	0.5000	0.0769
7	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000
8	0.7083	0.1977	0.3091		0.5833	0.1944	0.2917
9	0.6667	0.4324	0.5246		0.5417	0.8125	0.6500
10	0.6667	0.4444	0.5333		0.5833	0.8235	0.6829
11	0.4167	0.0885	0.1460		0.0833	0.0294	0.0435
12	0.7083	0.1504	0.2482		0.6250	0.2083	0.3125
13	0.7500	0.4286	0.5455		0.6667	0.6957	0.6809
14	0.8333	0.4348	0.5714		0.7917	0.7600	0.7755
15	0.2917	0.0946	0.1429		0.2083	0.0893	0.1250
16	0.6667	0.2388	0.3516		0.5833	0.2593	0.3590
17	0.7083	0.2000	0.3119		0.5833	0.1918	0.2887
18	0.6667	0.3478	0.4571		0.5833	0.6667	0.6222
19	0.6667	0.1839	0.2883		0.5833	0.2295	0.3294
20	0.3333	0.2667	0.2963		0.0833	0.2500	0.1250
21	0.6667	0.5333	0.5926		0.5833	0.8235	0.6829
22	0.6667	0.5333	0.5926		0.5833	0.8750	0.7000
23	0.6667	0.4706	0.5517		0.5833	0.7368	0.6512
24	0.7083	0.3469	0.4658		0.5833	0.5385	0.5600
25	0.2500	0.3333	0.2857		0.2083	0.4545	0.2857

Table C.19 Recall, precision, and harmonic mean of query by use case structure and keywords (2) for subject D (Continued)

Recall, precision, and harmonic mean of all queries for topic D							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
26	0.2917	0.0619	0.1022		0.2083	0.0610	0.0943
27	0.6667	0.5000	0.5714		0.5417	0.8125	0.6500
28	0.7917	0.1508	0.2533		0.6667	0.1928	0.2991
29	0.9167	0.2933	0.4444		0.8750	0.4200	0.5676
30	0.3750	0.1343	0.1978		0.2917	0.2188	0.2500
31	0.6667	0.5333	0.5926		0.4583	0.9167	0.6111
32	0.3333	0.1455	0.2025		0.2500	0.2143	0.2308
33	0.7083	0.1288	0.2179		0.5833	0.2692	0.3684
34	0.7083	0.4250	0.5313		0.5833	0.7368	0.6512
35	0.7083	0.2537	0.3736		0.5833	0.3500	0.4375
36	0.6667	0.1860	0.2909		0.5417	0.4483	0.4906
37	0.6667	0.5161	0.5818		0.5000	0.9231	0.6486
38	0.7917	0.3654	0.5000		0.5417	0.5417	0.5417
39	0.6667	0.5000	0.5714		0.6250	0.8824	0.7317
40	0.7083	0.3617	0.4789		0.6250	0.6818	0.6522
Average	**0.5865**	**0.3322**	**0.3937**		**0.4615**	**0.4956**	**0.4344**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table C.20 Recall, precision, and harmonic mean of query by use case structure and keywords (2) for subject E

Recall, precision, and harmonic mean of all queries for topic E							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
1	0.8125	0.3023	0.4407		0.5625	0.4500	0.5000
2	0.7500	0.1176	0.2034		0.4375	0.3500	0.3889
3	0.6875	0.1774	0.2821		0.4375	0.3684	0.4000
4	0.6875	0.1134	0.1947		0.4375	0.1077	0.1728
5	0.3750	0.1304	0.1935		0.3750	0.4615	0.4138
6	0.8125	0.2131	0.3377		0.5625	0.3913	0.4615
7	0.6875	0.2683	0.3860		0.4375	0.3889	0.4118
8	0.6875	0.2619	0.3793		0.4375	0.3333	0.3784
9	0.6250	0.0943	0.1639		0.5000	0.1096	0.1798
10	0.6875	0.1594	0.2588		0.5000	0.3333	0.4000
11	0.7500	0.1200	0.2069		0.6875	0.1964	0.3056
12	0.6875	0.0991	0.1732		0.6250	0.1754	0.2740
13	0.8125	0.1340	0.2301		0.6875	0.2037	0.3143
14	0.6875	0.0846	0.1507		0.6250	0.1087	0.1852
15	0.5000	0.3200	0.3902		0.1250	0.4000	0.1905
16	0.8125	0.0977	0.1745		0.8125	0.1140	0.2000
17	0.6875	0.1000	0.1746		0.5625	0.2250	0.3214
18	0.9375	0.1807	0.3030		0.6250	0.1961	0.2985
19	0.6875	0.1667	0.2683		0.6250	0.3448	0.4444
20	0.7500	0.0800	0.1446		0.5625	0.1525	0.2400
21	0.8125	0.2131	0.3377		0.6250	0.3448	0.4444
22	0.7500	0.1967	0.3117		0.6875	0.3548	0.4681
23	0.5625	0.1154	0.1915		0.5000	0.2222	0.3077
24	0.7500	0.2000	0.3158		0.6250	0.3448	0.4444
25	0.5625	0.2308	0.3273		0.5000	0.4000	0.4444

Table C.20 Recall, precision, and harmonic mean of query by use case structure and keywords (2) for subject E (Continued)

Recall, precision, and harmonic mean of all queries for topic E							
	Query by use case keywords				Query by use case structure		
Query	R	P	H		R	P	H
26	0.7500	0.1062	0.1860		0.6250	0.1266	0.2105
27	0.5625	0.1385	0.2222		0.5000	0.3810	0.4324
28	0.9375	0.1807	0.3030		0.6250	0.1961	0.2985
29	0.8125	0.0942	0.1688		0.6250	0.1266	0.2105
30	0.6875	0.2750	0.3929		0.5000	0.4000	0.4444
31	0.6875	0.0797	0.1429		0.5000	0.1013	0.1684
32	0.7500	0.1446	0.2424		0.5000	0.2105	0.2963
33	0.8125	0.2167	0.3421		0.6250	0.3448	0.4444
34	0.5625	0.2000	0.2951		0.5625	0.4286	0.4865
35	0.7500	0.1176	0.2034		0.5625	0.1139	0.1895
36	0.7500	0.0909	0.1622		0.5625	0.0947	0.1622
37	0.7500	0.1905	0.3038		0.5625	0.1837	0.2769
38	0.8125	0.1262	0.2185		0.5625	0.2250	0.3214
39	0.6875	0.0932	0.1642		0.4375	0.1014	0.1647
40	0.6250	0.1282	0.2128		0.4375	0.1750	0.2500
Average	**0.7125**	**0.1590**	**0.2525**		**0.5438**	**0.2572**	**0.3237**

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

APPENDIX D

PUBLICATION

This research is selected to present and publish in “The 1st National Conference on Computing and Information Technology” on May 24-25, 2005 at King Mongkut’s Institute of Technology North Bangkok. The publication is “Enhancing Use Case Reusability Using Term Similarity Computation”.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การเพิ่มความสามารถในการนำยูสเคสกลับมาใช้ใหม่โดยใช้การคำนวณพจน์ความคล้ายกัน
Enhancing Use Case Reusability Using Term Similarity Computation

อัคเดช อุดมชัยพร พิษณุ คนองชัยยศ และ นครทิพย์ พร้อมพูล

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

อีเมล *: akadej@akadej.com pizzanu@cp.eng.chula.ac.th* ***และ*** *nakornthip.s@chula.ac.th*

บทคัดย่อ

งานวิจัยนี้นำเสนอวิธีการเพิ่มความสามารถในการนำยูสเคสกลับมาใช้ใหม่ ยูสเคสเป็นส่วนประกอบของซอฟต์แวร์ที่ถูกพัฒนาในขั้นตอนการวิเคราะห์และออกแบบระบบและมักจะถูกนำมาใช้เพื่อเป็นข้อตกลงความต้องการระหว่างผู้ใช้งานกับผู้พัฒนาระบบ แต่ละยูสเคสมีเอกสารการบรรยายยูสเคสซึ่งถูกเขียนไว้ด้วยภาษาธรรมชาติซึ่งเอกสารนี้สามารถนำกลับมาใช้ใหม่ได้อีก ดังนั้นการพัฒนาซอฟต์แวร์ของระบบที่คล้ายกันอาจมีการนำยูสเคสที่ได้เคยออกแบบมาแล้วกลับมาใช้งานเพื่อช่วยลดงบประมาณและระยะเวลาในการพัฒนาซอฟต์แวร์ งานวิจัยนี้เสนอระบบการจัดเก็บและค้นคืนยูสเคสโดยพิจารณาจากความคล้ายกันของเอกสารการบรรยายยูสเคสกับชุดข้อมูลนำเข้าจากผู้ใช้งาน รีคอลและพรีซิชันจะถูกนำมาประเมินคุณภาพของผลลัพธ์ของวิธีการที่ได้นำเสนอ

คำสำคัญ: แผนภาพยูสเคส ยูสเคส เอกสารการบรรยายยูสเคส การนำกลับมาใช้ใหม่ การค้นคืนข้อมูล

Abstract

This research presents a method for enhancing reusability software components, developed in an analysis and design phase. Use case is always used during those phases as a contract between user and developer for capturing user requirements. Each use case has use case description, written in natural language. The software development of the similar systems may reuse use case description from previous works which results in reduce software development cost and time. This research purposes a use case storage and retrieval system using the similarity of the weighted indexing terms between use case descriptions and user input data. The recall and precision are used to evaluate the quality of the results from the proposed model.

Keywords: Use Case Diagram; Use Case; Use Case Description; Reuse; Information Retrieval.

1. บทนำ

ในปัจจุบันแผนภาพยูเอ็มแอล (UML Diagram : Unified Modeling Language Diagram) [1] มีบทบาทสำคัญในกระบวนการพัฒนาซอฟต์แวร์เชิงวัตถุ โดยเฉพาะอย่างยิ่งในขั้นตอนการวิเคราะห์และออกแบบระบบซอฟต์แวร์ แผนภาพยูเอ็มแอลมักจะถูกนำมาใช้เพื่อนำเสนอหน้าที่และขั้นตอนการทำงานในแต่ละมุมมองของระบบ ผลลัพธ์ที่ได้จากขั้นตอนนี้สามารถนำมาใช้เป็นข้อตกลงความต้องการระหว่างผู้ใช้งานกับผู้พัฒนาซอฟต์แวร์ แผนภาพหนึ่งที่มักจะถูกเลือกนำมาใช้งานคือแผนภาพยูสเคส (Use Case Diagram) [1, 2] เนื่องจากแผนภาพยูสเคสสามารถสื่อถึงความต้องการของผู้ใช้งานและช่วยให้ผู้ใช้งานที่ไม่เชี่ยวชาญสามารถทำความเข้าใจการทำงานของระบบจากแผนภาพยูสเคสได้ง่ายขึ้น แผนภาพยูสเคสจะประกอบด้วยส่วนประกอบต่างๆหลายส่วนซึ่งส่วนประกอบหนึ่งของแผนภาพยูสเคสที่สามารถนำกลับมาใช้ได้คือเอกสาร

การบรรยายยูสเคส (Use Case Description) [2, 3] เอกสารการบรรยายยูสเคสจะถูกเขียนด้วยภาษาธรรมชาติ ภาษาที่ได้รับการยอมรับว่ามีประสิทธิภาพในการสื่อสารมากกว่าภาษาใดๆในโลกอีกทั้งยังมีความเป็นสากลและได้รับการยอมรับใช้เป็นภาษาสื่อกลางในการติดต่อสื่อสารกันทั่วโลกก็คือภาษาอังกฤษ ด้วยเหตุนี้งานวิจัยนี้จึงมุ่งเน้นไปที่การทำงานกับเอกสารการบรรยายยูสเคสที่เป็นภาษาอังกฤษ

ในองค์กรหรือหน่วยงานต่างๆที่มีการพัฒนาซอฟต์แวร์มาเป็นระยะเวลาหนึ่ง การวิเคราะห์ความต้องการของผู้ใช้งานเพื่อนำมาเขียนเป็นข้อกำหนดความต้องการรวมไปถึงการออกแบบยูสเคสอาจมีความซ้ำซ้อนหรือใกล้เคียงกันกับงานที่ได้เคยทำมาแล้วในอดีต การทำงานที่ซ้ำซ้อนเหล่านี้จะก่อให้เกิดความสิ้นเปลืองทั้งระยะเวลาการทำงานและงบประมาณที่ใช้ในกระบวนการพัฒนาซอฟต์แวร์ งานวิจัยนี้จึงนำเสนอวิธีการเพิ่มความสามารถในการนำยูสเคสกลับมาใช้ใหม่ เพื่อช่วยประหยัดระยะเวลาและงบประมาณในกระบวนการพัฒนาซอฟต์แวร์ซึ่งการนำกลับมาใช้ใหม่ถือเป็นหัวใจสำคัญของวิธีการทางด้านวิศวกรรมซอฟต์แวร์ วิธีการนำยูสเคสกลับมาใช้ใหม่จะใช้วิธีการจัดเก็บและค้นคืนยูสเคสซึ่งจะได้กล่าวถึงรายละเอียดของวิธีการในส่วนต่อๆไป

2. แผนภาพยูสเคส

แผนภาพยูสเคสเป็นแผนภาพหนึ่งในส่วนของแผนภาพยูเอ็มแอล จุดประสงค์ของแผนภาพยูสเคสมักจะถูกนำมาใช้แสดงถึงข้อกำหนดหลักของระบบและความสัมพันธ์ที่เกี่ยวกับบทบาทหน้าที่ของระบบกับผู้ที่เกี่ยวข้องในระบบ โดยแผนภาพยูสเคสจะแสดงถึงส่วนประกอบสำคัญ 3 ส่วน คือ 1) ชุดของยูสเคส 2) ชุดของผู้ที่เกี่ยวข้องกับระบบ และ 3) ความสัมพันธ์ต่างๆระหว่างชุดของยูสเคสกับชุดของผู้ที่เกี่ยวข้องกับระบบ ในชุดของยูสเคสจะประกอบไปด้วยยูสเคสต่างๆที่กล่าวถึงรายละเอียดการทำงานของระบบในแต่ละหน้าที่หลัก เอกสารที่ทำหน้าที่เก็บรายละเอียดของแต่ละยูสเคสคือเอกสารการบรรยายยูสเคส ในที่นี้จะขอกล่าวถึงเพียงรายละเอียดของเอกสารการบรรยายยูสเคส [2, 3, 7] ประกอบไปด้วยโครงสร้างหลัก 7 ส่วนคือ

1) ชื่อของยูสเคส คือ ชื่อที่ใช้เรียกยูสเคส
2) จุดประสงค์ คือ จุดประสงค์การทำงานหลักของยูสเคส
3) ผู้ใช้งานยูสเคส คือ ผู้ที่เกี่ยวข้องในระบบที่มีบทบาทเกี่ยวข้องกับการทำงานของยูสเคส
4) เงื่อนไขก่อนการทำงาน คือ เงื่อนไขหรือสถานะของระบบที่จะทำให้ยูสเคสเริ่มทำงาน
5) ลำดับเหตุการณ์การทำงาน คือ ลำดับการทำงานของยูสเคส จะบอกถึงการทำงานที่เป็นขั้นตอนตั้งแต่เริ่มจนจบการทำงานโดยจะต้องมีการบอกจุดเริ่มต้นและจุดสิ้นสุดของการทำงาน
6) ลำดับทางเลือกการทำงาน คือ ลำดับการทำงานที่มีขั้นตอนการทำงานที่ต่างจากลำดับเหตุการณ์การทำงานปกติ โดยจะทำงานเมื่อเกิดกรณีที่เป็นไปตามเงื่อนไขบางอย่างของลำดับเหตุการณ์ปกติ
7) เงื่อนไขหลังการทำงาน คือ เงื่อนไขหรือสถานะของระบบที่ต้องเป็นจริงเสมอหลังจบการทำงานยูสเคส

3. งานวิจัยที่เกี่ยวข้อง

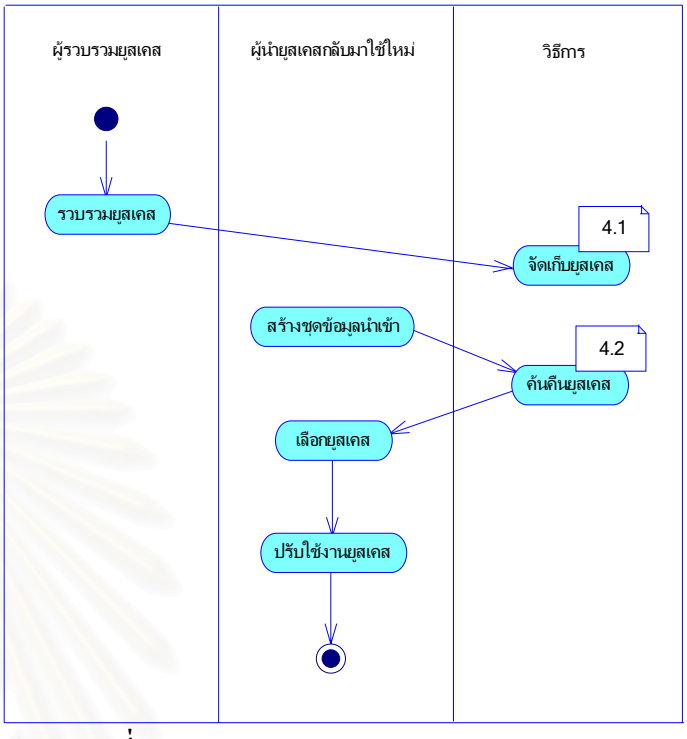
Krueger [4] ได้ทำการสำรวจงานวิจัยที่เกี่ยวกับการนำซอฟต์แวร์กลับมาใช้ใหม่และได้ทำการแบ่งหมวดหมู่ของงานวิจัยที่เกี่ยวกับการนำซอฟต์แวร์กลับมาใช้ใหม่ออกเป็น 4 ส่วนด้วยกัน สำหรับงานวิจัยนี้จะอยู่ในส่วนที่ 3 คือส่วนของการเลือกชิ้นส่วนของซอฟต์แวร์ที่กลับมาใช้ใหม่ให้เหมาะสมซึ่งนักวิจัยหลายท่านได้นำเสนองานวิจัยที่อยู่ในส่วนเดียวกันนี้และมีความเกี่ยวข้องกับการวิจัยเรื่องนี้ เช่น งานวิจัยของ Woo [5] ได้นำเสนอเทคนิควิธีการนำยูสเคสกลับมาใช้ใหม่โดยอัตโนมัติโดยใช้การแปลงลำดับเหตุการณ์การทำงานให้อยู่ในรูปแบบของกราฟและใช้ขั้นตอนวิธี SUBDUE ในการเปรียบเทียบความคล้ายกันของกราฟเพื่อค้นคืนยูสเคส แต่การแปลงข้อมูลเป็นกราฟนั้นจะทำให้การค้นคืนใช้เวลาค่อนข้างนานโดยเฉพาะอย่างยิ่งเมื่อข้อมูลในฐานข้อมูลมีจำนวนมากๆ ในขณะที่งานวิจัย

ของ Blok [6] ได้เสนอวิธีการจัดเก็บและค้นคืนแผนภาพยูเอ็มแอลที่เน้นในส่วนของการค้นคืนยูสเคส กระบวนการของการจัดเก็บและค้นคืนยูสเคสนั้นจะถูกพิจารณาจากลำดับเหตุการณ์ (Event Flow) ของแต่ละยูสเคส โดยมีการจัดกลุ่มความคล้ายกันของดรรชนีที่ได้มาจากลำดับเหตุการณ์ การกำหนดดรรชนีและการจัดกลุ่มความคล้ายกันให้กับยูสเคสจะถูกตัดสินใจด้วยผู้เชี่ยวชาญก่อนที่จะถูกจัดเก็บลงฐานข้อมูล การค้นคืนยูสเคสจะมีการเปรียบเทียบความคล้ายกันของดรรชนีที่ถูกกำหนดมาจากชุดข้อมูลนำเข้ากับลำดับเหตุการณ์ทั้งหมดที่เก็บอยู่ในฐานข้อมูลเพื่อให้ผู้ใช้งานสามารถเลือกยูสเคสที่คล้ายกันไปปรับใช้งานต่อไป แต่การนำผู้เชี่ยวชาญมาช่วยจัดทำดรรชนีและจัดกลุ่มความคล้ายกันอาจจะเกิดความคลาดเคลื่อนเอนเอียงอีกทั้งยังไม่สะดวกต่อการนำไปสร้างเครื่องมือเพื่อช่วยในการทำงานด้วยคอมพิวเตอร์ และงานวิจัยของ Saeki [7] เสนอวิธีการนำเอกสารการบรรยายยูสเคสกลับมาใช้ใหม่ โดยได้เสนอรูปแบบการวิเคราะห์และเก็บรวบรวมความต้องการของผู้ใช้งานให้เป็นแผนภาพยูสเคส รวมทั้งสร้างแบบฟอร์มเพื่อจัดเก็บรายละเอียดของเอกสารการบรรยายยูสเคส งานวิจัยของ Saeki นี้จะเน้นไปที่การวิเคราะห์ความต้องการเพื่อจัดเก็บแผนภาพยูสเคสให้สะดวกต่อการค้นคืนแผนภาพยูสเคส แต่ไม่ได้กล่าวถึงรายละเอียดของวิธีการค้นคืนเอกสารการบรรยายยูสเคสเพื่อนำกลับมาใช้ใหม่

เพื่อช่วยลดข้อจำกัดของงานวิจัยที่กล่าวมา งานวิจัยนี้จึงนำเสนอทั้งวิธีการจัดเก็บและค้นคืนยูสเคสเพื่อนำยูสเคสกลับมาใช้ใหม่ โดยการจัดเก็บยูสเคสจะจัดเก็บข้อมูลให้อยู่ในรูปแบบที่เหมาะสมเพื่อความรวดเร็วในการค้นคืนยูสเคส และขั้นตอนต่างๆของการจัดเก็บและค้นคืนยูสเคสจะมีการทำงานที่เป็นอัตโนมัติทำให้สามารถนำไปสร้างเครื่องมือเพื่อช่วยในการค้นคืนด้วยคอมพิวเตอร์ได้

4. วิธีการนำยูสเคสกลับมาใช้ใหม่

วิธีการนำยูสเคสกลับมาใช้ใหม่จะแสดงได้ดังภาพที่ 1 ซึ่งจะเห็นได้ว่าในส่วนของวิธีการจะมีวิธีการหลักอยู่ 2 ส่วนคือ ส่วนของการจัดเก็บยูสเคสและส่วนของการค้นคืนยูสเคส รายละเอียดการทำงานของทั้ง 2 ส่วนมีดังนี้



ภาพที่ 1 แผนภาพกิจกรรมของการนำยูสเคสกลับมาใช้ใหม่

4.1 การจัดเก็บยูสเคส

วิธีการจัดเก็บยูสเคสจะทำการจัดเก็บยูสเคสลงฐานข้อมูลในรูปแบบที่สะดวกต่อการค้นคืนยูสเคสซึ่งวิธีการจัดเก็บยูสเคสจะมี 3 ส่วนหลักๆคือ การสร้างดรรชนี การสร้างพจนานุกรมคำพ้อง (Thesaurus) และการกำหนดน้ำหนักให้กับดรรชนี ซึ่งมีรายละเอียดดังต่อไปนี้

4.1.1 การสร้างดรรชนี

การสร้างดรรชนีจากยูสเคสจะมีจุดประสงค์เพื่อสร้างชุดของคำหรือพจน์ที่สามารถระบุถึงเอกสารการบรรยายยูสเคสได้ การสร้างดรรชนี [8] มีขั้นตอนดังนี้ คือ

1) นำข้อมูลทั้งหมดที่ปรากฏอยู่ในเอกสารมาแยกเป็นพจน์โดยพจน์ที่ซ้ำกันให้นำมาเพียงพจน์เดียวเท่านั้น
2) ตัดพจน์ที่ไม่สื่อความหมายทิ้งไปโดยอ้างอิงรายการของพจน์ที่ไม่สื่อความหมาย (Stop List) จากรายการพจน์ที่มีอยู่แล้วหรือจะทำการสร้างขึ้นมาใหม่ก็ได้
3) ตัดส่วนของหน่วยพจน์ตัวหน้า (Prefix) และหน่วย

พจน์ตัวหลัง (Suffix) ทิ้งไปเพื่อเป็นการลดรูปแบบความแตกต่างกันของพจน์

4) ตัดพจน์ที่มีความถี่สูงหรือต่ำกว่าค่าขีดแบ่ง (Threshold) ที่กำหนดไว้ทิ้งไป เพราะพจน์ที่มีความถี่สูงหรือต่ำเกินไปมักจะเป็นพจน์ที่ไม่สื่อความหมาย

หลังจากผ่านขั้นตอนการทำดรรชนีแล้วจะได้ชุดของพจน์ที่เป็นดรรชนีของแต่ละเอกสารการบรรยายยูสเคสที่พร้อมจะนำไปสร้างพจนานุกรมคำพ้องในขั้นตอนต่อไป

4.1.2 การสร้างพจนานุกรมคำพ้อง

การสร้างพจนานุกรมคำพ้องจะทำหลังจากที่สร้างดรรชนีเสร็จสิ้น ชุดของดรรชนีจะถูกนำมาสร้างพจนานุกรมคำพ้องเพื่อเก็บรวบรวมและจัดแบ่งหมวดหมู่ความคล้ายกันของชุดของดรรชนีที่ได้มาจากเอกสารการบรรยายยูสเคส การสร้างพจนานุกรมคำพ้อง [8] มีขั้นตอนดังนี้

1) สร้างตารางที่สัมพันธ์กันของดรรชนีและเอกสารการบรรยายยูสเคสดังแสดงในตารางที่ 1 โดยให้

- T_1, T_2 จนถึง T_t หมายถึงดรรชนี T ตั้งแต่ดรรชนีตัวที่ 1 จนถึงดรรชนีตัวที่ t
- D_1, D_2 จนถึง D_n หมายถึงเอกสารการบรรยายยูสเคส D ตั้งแต่เอกสารที่ 1 จนถึงเอกสารที่ n
- d_{ij} คือ จำนวนความถี่ของดรรชนี T_i ที่ปรากฏอยู่ในเอกสารการบรรยายยูสเคส D_j

ตารางที่ 1 ตารางดรรชนีของเอกสารการบรรยายยูสเคส

	$\mathbf{T_1}$	$\mathbf{T_2}$	…	$\mathbf{T_t}$
$\mathbf{D_1}$	d_{11}	d_{12}	…	d_{1t}
$\mathbf{D_2}$	d_{21}	d_{22}	…	d_{2t}
⋮	⋮	⋮	⋮	⋮
$\mathbf{D_n}$	d_{n1}	d_{n2}	…	d_{nt}

2) หาค่าความคล้ายกันของแต่ละคู่ของดรรชนีโดยใช้สูตรการหาค่าความคล้ายกันจากสมการที่ 1

$$S(T_i, T_j) = \frac{\sum_{d=1}^{n} t_{di}\, t_{dj}}{\sum_{d=1}^{n}(t_{di}) + \sum_{d=1}^{n}(t_{dj}) - \sum_{d=1}^{n} t_{di}\, t_{dj}} \qquad \textbf{(1)}$$

- *d* คือเอกสารการบรรยายยูสเคสที่ *1* จนถึง *n*
- *S* (T_i, T_j) คือค่าความคล้ายกันของคู่ดรรชนี T_i และ T_j

3) นำค่าที่ได้มาสร้างตารางความคล้ายกันของดรรชนีดังตารางที่ 2

ตารางที่ 2 ตารางความคล้ายกันของดรรชนี

	$\mathbf{T_1}$	$\mathbf{T_2}$	…	$\mathbf{T_t}$
$\mathbf{T_1}$	$s(T_1,T_1)$	$s(T_1,T_2)$	…	$s(T_1,T_t)$
$\mathbf{T_2}$	$s(T_2,T_1)$	$s(T_2,T_2)$	…	$s(T_2,T_t)$
⋮	⋮	⋮	⋮	⋮
$\mathbf{T_t}$	$s(T_t,T_1)$	$s(T_t,T_2)$	…	$s(T_t,T_t)$

เมื่อค่า T_1, T_2 จนถึง T_t หมายถึงดรรชนี T ตั้งแต่ดรรชนีตัวที่ 1 จนถึงดรรชนีตัวที่ t และค่า $s(T_i,T_j)$ ที่อยู่ในตารางก็จะเป็นค่าความคล้ายกันของดรรชนีแต่ละคู่ i และ j โดยที่ค่า i และ j จะมีค่าตั้งแต่ 1 จนถึง t

ตารางความคล้ายกันของดรรชนีจะแสดงให้เห็นถึงค่าความคล้ายกันของดรรชนีซึ่งสามารถนำมาสร้างพจนานุกรมคำพ้องได้โดยนำคู่ของดรรชนีที่มีค่าความคล้ายกันที่มากกว่าค่าขีดแบ่งที่กำหนดไว้มาจัดอยู่ในหมวดหมู่เดียวกัน ค่าขีดแบ่งนั้นจะขึ้นอยู่กับผู้ใช้งานที่จะกำหนดค่าขีดแบ่งตามความเหมาะสม กล่าวคือ ถ้าค่าขีดแบ่งนั้นยิ่งมีค่ามากก็เท่ากับว่าพจนานุกรมคำพ้องที่จะถูกสร้างขึ้นก็จะยิ่งมีความละเอียดของหมวดหมู่ข้อมูลมากขึ้นซึ่งก็จะส่งผลโดยตรงกับความความถูกต้องแม่นยำของการค้นคืนยูสเคส

4.1.3 การกำหนดน้ำหนักให้ดรรชนี

การกำหนดน้ำหนักของดรรชนีคือการกำหนดค่าความสำคัญให้กับแต่ละดรรชนี ทั้งนี้เพราะแต่ละดรรชนีในแต่ละเอกสารการบรรยายยูสเคสอาจมีความสำคัญไม่เท่ากัน ดังนั้นการกำหนดน้ำหนักให้แต่ละดรรชนีก่อนที่จะนำชุดของดรรชนีนั้นไปเปรียบเทียบเพื่อค้นคืนยูสเคสจะช่วยทำให้ผลลัพธ์ของการค้นคืนแม่นยำมากขึ้น การกำหนดค่าน้ำหนักให้กับดรรชนี [8] มีวิธีการกำหนดค่าน้ำหนักดังสมการที่ 2

$$W_{ik} = \frac{Freq_{ik}}{TotFreq_k} \qquad \textbf{(2)}$$

เมื่อ W_{ik} คือค่าน้ำหนักของดรรชนี k ในเอกสาร i

$Freq_{ik}$ คือ ความถี่ของดรรชนี k ที่ปรากฏทั้งหมดในเอกสาร i

$TotFreq_k$ คือ ความถี่ของดรรชนี k ที่ปรากฏในเอกสารทั้งหมดในฐานข้อมูล

จากวิธีการดังกล่าวค่าน้ำหนักที่ถูกกำหนดให้กับแต่ละดรรชนีจะขึ้นอยู่กับจำนวนของดรรชนีที่เกิดขึ้นในเอกสารนั้นเมื่อเทียบกับจำนวนดรรชนีที่มีในเอกสารทั้งหมด นั่นคือดรรชนีตัวใดที่ปรากฏอยู่ในเอกสารเป็นจำนวนมาก ดรรชนีนั้นก็จะยิ่งถือว่ามีความสำคัญสำหรับเอกสารนั้นมากขึ้นตามไปด้วย

4.2 การค้นคืนยูสเคส

ในการค้นคืนยูสเคส ผู้ใช้งานที่ต้องการจะนำยูสเคสกลับมาใช้ใหม่จะต้องสร้างข้อมูลนำเข้าเบื้องต้นเกี่ยวกับยูสเคสที่ต้องการ ข้อมูลที่ถูกสร้างขึ้นมาควรจะสามารถสื่อถึงยูสเคสที่ต้องการได้ เช่น จุดประสงค์การทำงานของยูสเคส ลำดับการทำงานของยูสเคส เป็นต้น จากนั้นก็จะทำการสร้างดรรชนีจากข้อมูลดังกล่าวซึ่งก็จะมีขั้นตอนเหมือนกับการสร้างดรรชนีของการจัดเก็บยูสเคส หลังจากนั้นก็จะนำชุดของดรรชนีเหล่านั้นไปเปรียบเทียบกับข้อมูลในฐานข้อมูลเพื่อทำการค้นคืนยูสเคสที่มีความคล้ายกัน วิธีการที่สำคัญของการค้นคืนยูสเคสคือการเปรียบเทียบความคล้ายกันซึ่งมีรายละเอียดดังนี้

4.2.1 การเปรียบเทียบความคล้ายกัน

การเปรียบเทียบความคล้ายกันของยูสเคสจะกระทำเพื่อค้นคืนเอกสารการบรรยายยูสเคสที่มีความคล้ายกันกับชุดข้อมูลนำเข้ามากที่สุด การเปรียบเทียบความคล้ายกันจะนำดรรชนีจากชุดข้อมูลนำเข้ามาเปรียบเทียบกับดรรชนีของแต่ละเอกสารการบรรยายยูสเคสในฐานข้อมูล ดรรชนีที่ถูกจัดอยู่ในหมวดหมู่เดียวกันในพจนานุกรมข้อมูลจะถือว่าเป็นดรรชนีที่มีความคล้ายกัน ความคล้ายกันของดรรชนีจะถูกคำนวณเป็นคะแนนความคล้ายกันซึ่งการคำนวณคะแนนความคล้ายกันจะคำนวณได้จากสมการที่ 3

$$Sim(X,UC_j) = \sum_{i=1}^{n} t_i w_i \qquad \textbf{(3)}$$

- $Sim(X,UC_j)$ คือค่าคะแนนความคล้ายกันของชุดข้อมูลนำเข้า X และเอกสารการบรรยายยูสเคส UC_j
- w_i คือค่าน้ำหนักของดรรชนี t_i
- t_i คือดรรชนีที่ i จากชุดข้อมูลนำเข้า X โดยกำหนดให้

$t_i = 0$ เมื่อดรรชนีนั้นไม่ปรากฏในเอกสารบรรยายยูสเคส UC_j

$t_i = 1$ เมื่อดรรชนีนั้นปรากฏอยู่ในเอกสารการบรรยายยูสเคส UC_j

ในกรณีที่ดรรชนีใดๆจากชุดข้อมูลนำเข้าไม่ได้ปรากฏอยู่ในเอกสารการบรรยายยูสเคสที่ต้องการจะเปรียบเทียบในฐานข้อมูล แต่มีดรรชนีที่อยู่ในหมวดหมู่เดียวกันกับดรรชนีพจน์นั้นในพจนานุกรมคำพ้องหรือกล่าวได้ว่ามีดรรชนีที่มีความหมายคล้ายกันปรากฏอยู่ในเอกสารการบรรยายยูสเคสนั้น กรณีนี้จะให้มีการกำหนดค่าน้ำหนักให้กับดรรชนีใดๆโดยจะคำนวณได้จากค่าเฉลี่ยของดรรชนีที่มีความหมายคล้ายกันจากพจนานุกรมคำพ้องซึ่งปรากฏอยู่ในเอกสารการบรรยายยูสเคสนั้นในฐานข้อมูล

เมื่อนำเอกสารการบรรยายยูสเคสจากฐานข้อมูลทุกเอกสารที่ต้องการจะเปรียบเทียบมาคำนวณคะแนนความคล้ายกันแล้ว เอกสารการบรรยายยูสเคสที่มีคะแนนความคล้ายกันมากที่สุดเมื่อเทียบกับชุดข้อมูลนำเข้านั้นจะถือว่าเป็นเอกสารการบรรยายยูสเคสที่มีความคล้ายกันมากที่สุด ผู้ใช้งานสามารถเลือกนำเอายูสเคสนั้นไปปรับใช้งานให้เข้ากับระบบของตัวเองต่อไป

5. การประเมินผลลัพธ์ที่ได้

การค้นคืนยูสเคสจะมีการประเมินความถูกต้องแม่นยำของผลลัพธ์ที่ได้เพื่อใช้เป็นมาตรวัดวิธีการที่กล่าวมาทั้งหมด การประเมินผลลัพธ์ที่ได้จะมีการใช้ค่ารีคอล (Recall) และค่าพรีซิชัน (Precision) [8] เป็นมาตรวัดความถูกต้องแม่นยำของวิธีการที่ได้นำเสนอมา ค่ารีคอลและพรีซิชันมีรายละเอียดดังนี้

รีคอล

รีคอลเป็นค่าที่บอกถึงอัตราส่วนความเกี่ยวข้องของ

ข้อมูลที่ค้นคืนมาได้ต่อข้อมูลที่มีความเกี่ยวข้องทั้งหมดโดยค่ารีคอลที่จะนำมาประเมินวิธีการค้นคืนยูสเคสคำนวณได้จาก

$$Recall = \frac{RelUC}{TotRelUC} \quad \textbf{(4)}$$

- *RelUC* คือจำนวนของเอกสารการบรรยายยูสเคสที่เกี่ยวข้องกับชุดข้อมูลนำเข้าที่ถูกค้นคืน

- *TotRelUC* คือจำนวนของเอกสารการบรรยายยูสเคสที่เกี่ยวข้องกับชุดข้อมูลนำเข้าทั้งหมด

พรีซิชัน

พรีซิชันเป็นค่าที่บอกถึงอัตราส่วนความเกี่ยวข้องของข้อมูลที่ค้นคืนมาได้ต่อข้อมูลที่สามารถค้นคืนมาได้ทั้งหมด โดยค่าพรีซิชันที่นำมาประเมินวิธีการค้นคืนยูสเคสคำนวณได้จาก

$$Precision = \frac{RelUC}{TotRetUC} \quad \textbf{(5)}$$

- *RelUC* คือจำนวนของเอกสารการบรรยายยูสเคสที่เกี่ยวข้องกับชุดข้อมูลนำเข้าที่ถูกค้นคืน

- *TotRetUC* คือจำนวนของเอกสารการบรรยายยูสเคสที่ถูกค้นคืนทั้งหมด

จากการประเมินผลลัพธ์ที่ได้ จะได้ค่ารีคอลและค่าพรีซิชันซึ่งเป็นค่าที่สื่อให้ทราบถึงความถูกต้องและความแม่นยำของการค้นคืนเอกสารการบรรยายยูสเคส ค่ารีคอลและค่าพรีซิชันจะมีค่าอยู่ระหว่าง 0 ถึง 1 โดยถ้าค่ารีคอลและค่าพรีซิชันมีค่าเข้าใกล้ 1 มากเท่าไรก็จะสามารถประเมินได้ว่าวิธีการนำยูสเคสกลับมาใช้ใหม่มีประสิทธิภาพมากขึ้นตามไปด้วย

6. กรณีตัวอย่าง

เพื่อเป็นการนำเสนอวิธีการที่ได้กล่าวมาแล้วให้มีความเข้าใจมากยิ่งขึ้น ในหัวข้อนี้จะนำเสนอกรณีตัวอย่างในการค้นคืนเอกสารการบรรยายยูสเคสจากฐานข้อมูลโดยจะทำการเปรียบเทียบความคล้ายกันระหว่างชุดข้อมูลนำเข้าและข้อมูลที่เก็บอยู่ในฐานข้อมูลเพื่อค้นคืนยูสเคสที่มีความคล้ายกันมากที่สุด ข้อมูลสมมติของกรณีตัวอย่างมีดังนี้

- สมมติให้ความถี่ของจำนวนดรรชนีที่ปรากฏในแต่ละเอกสารการบรรยายยูสเคสในฐานข้อมูลมีค่าดังตารางที่ 3 เมื่อ T คือดรรชนีและ UC คือยูสเคส
- สมมติให้ดรรชนี T_1, T_2, T_3 อยู่ในพจนานุกรมคำพ้องหมวดหมู่เดียวกัน
- และสมมติให้ชุดข้อมูลนำเข้ามีดรรชนีคือ T_1, T_2, T_4

ตารางที่ 3 ความถี่ของดรรชนีที่ปรากฏในเอกสารการบรรยายยูสเคส

	UC_1	UC_2	UC_3	UC_4	**Total**
T_1	5	2	0	1	8
T_2	4	1	1	2	8
T_3	2	2	7	0	11
T_4	1	5	5	5	16

จากข้อมูลในตารางที่ 3 จะสามารถนำมากำหนดค่าน้ำหนักให้กับแต่ละดรรชนีของแต่ละเอกสารโดยใช้สมการที่ 2 จะได้ผลลัพธ์ดังตัวอย่างที่แสดงในตารางที่ 4

ตารางที่ 4 ตัวอย่างการกำหนดค่าน้ำหนักให้กับดรรชนี

	UC_1	UC_2	UC_3	UC_4
T_1	5/8	2/8	0	1/8
T_2	4/8	1/8	1/8	2/8
T_3	2/11	2/11	7/11	0
T_4	1/16	5/16	5/16	5/16

และจากชุดข้อมูลนำเข้าที่มีดรรชนีคือ T_1, T_2, T_4 เมื่อนำมาเปรียบเทียบความคล้ายกันกับข้อมูลในฐานข้อมูลสามารถนำมาคำนวณหาคะแนนความคล้ายกันจากสมการที่ 3 ได้ดังตารางที่ 5

ตารางที่ 5 ตัวอย่างการคำนวณคะแนนความคล้ายกัน

	คะแนนความคล้ายกัน
UC_1	(5/8) + (4/8) + 0 + (1/16) = **1.19**
UC_2	(2/8) + (1/8) + 0 + (5/16) = **0.69**
UC_3	(((1/8) + (7/11)) / 2) + (1/7) + 0 + (5/16) = **0.84**
UC_4	(1/8) + (2/8) + 0 + (5/16) = **0.69**

เมื่อพิจารณาการคำนวณคะแนนความคล้ายกันของยูสเคสที่

3 จะเห็นได้ว่าดรรชนี T_1 ไม่ได้ปรากฏอยู่ในยูสเคสที่ 3 แต่จะมีดรรชนี T_2 และ T_3 ซึ่งอยู่ในหมวดหมู่เดียวกันกับดรรชนี T_1 ในพจนานุกรมคำพ้องปรากฏอยู่ จึงนำค่าน้ำหนักของดรรชนี T_2 และ T_3 มาคิดเฉลี่ยเป็นค่าน้ำหนักแทนค่าน้ำหนักของดรรชนี T_1 และในท้ายที่สุดเมื่อพิจารณาคะแนนความคล้ายกันในตารางที่ 5 จะทราบได้ว่าชุดข้อมูลนำเข้ามีความคล้ายกันกับยูสเคสที่ 1 มากที่สุด

7. สรุปผลการวิจัย

งานวิจัยนี้นำเสนอวิธีการเพิ่มความสามารถในการนำยูสเคสกลับมาใช้ใหม่โดยได้นำวิธีการทางด้านการจัดเก็บและค้นคืนข้อมูลเข้ามาใช้งานกับยูสเคสในส่วนของเอกสารการบรรยายยูสเคสซึ่งเป็นเอกสารสำคัญที่เก็บรายละเอียดการทำงานของแต่ละยูสเคสในแผนภาพยูสเคส การค้นคืนเอกสารการบรรยายยูสเคสจะอ้างอิงถึงยูสเคสในแผนภาพยูสเคสซึ่งจะสามารถนำไปสู่การค้นคืนแผนภาพยูสเคสได้ ข้อดีของงานวิจัยนี้คือวิธีการนำยูสเคสกลับมาใช้ใหม่ที่มีขั้นตอนการทำงานแบบอัตโนมัติ รวมทั้งการสร้างข้อมูลนำเข้าจากผู้ใช้งานสามารถใช้งานได้โดยง่ายซึ่งให้ความสะดวกกับผู้ใช้งานในการนำยูสเคสกลับมาใช้ใหม่ แต่อย่างไรก็ตามงานวิจัยนี้จะมีการพัฒนางานวิจัยต่อไปในส่วนของการทดลองวิธีการกับชุดของยูสเคสที่มีจำนวนมากๆและจัดสร้างเครื่องมือเพื่อทดสอบวิธีการที่ได้นำเสนอมาซึ่งก็ถือเป็นส่วนหนึ่งของงานวิจัยที่จะได้ทำต่อไปในอนาคต และยังรวมไปถึงการพัฒนาวิธีการกำหนดน้ำหนักให้กับดรรชนีโดยวิเคราะห์จากความสัมพันธ์ของยูสเคสและโครงสร้างของเอกสารการบรรยายยูสเคสเพื่อให้วิธีการนำยูสเคสกลับมาใช้ใหม่โดยอัตโนมัติมีประสิทธิภาพมากยิ่งขึ้น

8. บรรณานุกรม

[1] Grady Booch, James Rambaugh, and Ivar Jacobson, “The Unified Modeling Language, User Guide,” Addison-Wesley, 1998.

[2] Geri Schneider, and Jason P. Winters, “Applying Use Cases, A Pratical Guide,” Addison-Wesley, 2001.

[3] Alan Dennis, Barbara Haley Wixom, and David Tegarden, “Systems Analysis and Design, An Object-Oriented Approach with UML,” Wiley, 2002.

[4] Charles W. Krueger, “Software Reuse,” ACM computing surveys, vol. 24, no. 2, 1992, pp. 131-183.

[5] Han G. Woo, and William N. Robinson, “Reuse of Scenario Specifications Using an Automated Relational Learner: A Lightweight Approach,” International Conference on Requirements Engineering, 2002.

[6] Maurits C. Blok, and Jacob L. Cybulski, “Reusing UML Specifications in a Constrained Application Domain,” Asia-Pacific Software Engineering Conference, 1998.

[7] Motoshi Saeki, “Reusing Use Case Descriptions for Requirements Specification: Towards Use Case Patterns,” Asia-Pacific Software Engineering Conference, 1999.

[8] Gerard Salton, and Michael J. McGill. “Introduction to Modern Information Retrieval,” McGraw-Hill, 1983.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

BIOGRAPHY

Akadej Udomchaiporn was born on September 23, 1981, received a Bachelor's degree of Computer Science from Faculty of Science, King Mongkut's Institute of Technology Ladkrabang in March, 2004. After that, he has studied a Master's degree of Software Engineering, Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย