

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

Improving Effectiveness of Dialogue in Learning Communities

A thesis presented in partial fulfilment of the requirements for the degree of

PhD in Information Systems

at Massey University, Palmerston North, New Zealand.

Jingyu Yang

2009

ABSTRACT

In a learning community, conventional discussion forums are integral to web-based interventions in traditional classrooms as well as on-line learning environments. Despite popular belief that they are a great success in fostering deep and meaningful discussions and support active learning; research has found that there are millions of messages posted by users to express such an opinion, but it is hard to be directly delivered to all users. Finally there are millions of postings in databases across the country stored away and never reused. This thesis introduces a PhD student's current research work. It proposes a distributed intelligent discussion forum system dedicated to supporting both students and teachers. The system is developed with the primary goal of reducing the number of problems associated with conventional discussion forum systems in web-based environments and improving the effectiveness of dialogue between students with each other and with teachers so that it can enhance each individual student's ability to share and learn knowledge.

ACKNOWLEDGEMENTS

In the last four years, the writing process of the thesis included both joyful and painful moment. Even there were times when I thought the work will never finish. Therefore I must be sincerely grateful to my supervisor, Dr. Kinshuk, Professor for his guidance, advice and inspiring wisdom throughout this project and for his lots of valuable support and supervision from the beginning of the research project, without which this thesis would not have been completed. He gave me such patient, effective support and encourages in such long time,

In addition, I would like to thank my whole family including my parents, my brother and my partner for their support and encouragement. And also thank my second supervisor, Dr. Lynn Jeffrey and all my friends who gave me such help and encourages during last four years.

Financial support towards this study from E-Learning Collaborative Development Fund (eCDF) (Ministry of Education, New Zealand), Massey University and IBM New Zealand is highly appreciated.

Your thought, help and kindness will encourage me to keep going!

Thanks You All!

Yours truly,
Jingyu Yang

CONTENTS

ABSTRACT	II
ACKNOWLEDGEMENTS	III
LIST OF TABLES.....	IX
LIST OF FIGURES.....	X
LIST OF ABBREVIATIONS	XII
CHAPTER 1 GENERAL INTRODUCTION	1
1.1 INTRODUCTION	1
1.2 DIALOGUE IN THE LEARNING COMMUNITIES.....	2
1.3 RESEARCH QUESTIONS, TREATMENT, METHODOLOGY AND SCOPE.....	3
1.3.1 Limitations of Discussion Forums in Educational Context	3
1.3.2 Research Questions.....	4
1.3.3 Treatment of Research Question.....	5
1.3.4 Methodology.....	6
1.3.5 Scope - eQuake Project.....	7
1.4 DISSERTATION STRUCTURE	8
CHAPTER 2 LITERATURE REVIEW.....	10
2.1 INTRODUCTION	10
2.2 ONLINE LEARNING COMMUNITIES.....	10
2.2.1 Community-based Learning.....	10
2.2.2 Online Learning Communities.....	13
2.2.3 C4P Model.....	14
2.2.4 Dialogue via Technologies	18
2.3 DISCUSSION FORUM SYSTEMS	21
2.3.1 Overview	21
2.3.2 The Use of Discussion Forums.....	22
2.3.3 General Introduction of Major Systems.....	24

2.3.4	Limitations of Current Discussion Forums	27
2.4	INFORMATION RETRIEVAL.....	29
2.4.1	Overview	29
2.4.2	Text Similarity.....	35
2.4.3	Relevance Feedback	37
2.4.4	Latent Semantic Indexing	42
2.4.5	Search Engine	48
2.5	RECOMMENDER SYSTEMS	52
2.6	CONCLUSION	56
CHAPTER 3 REQUIREMENTS ANALYSIS.....		57
3.1	INTRODUCTION	57
3.2	REQUIREMENTS GATHERING	57
3.2.1	Overview	57
3.2.2	Challenges in Discussion Forum	58
3.2.3	Requirement Analysis.....	61
3.2.4	Requirements Summary.....	63
CHAPTER 4 SOLUTION DESIGN		65
4.1	INTRODUCTION	65
4.2	EXISTING SOLUTIONS FOR TRADITIONAL DISCUSSION FORUMS.....	65
4.2.1	Overview	65
4.2.2	EQuake	66
4.2.3	Intelligent Discussion-Bot (ID-B).....	68
4.2.4	Limitations of Current Solutions	71
4.3	PROPOSED SOLUTION	72
4.3.1	Introduction.....	72
4.3.2	Requirements Recap	72
4.3.3	Solution Overview	73
4.3.4	High-Level System Architecture and Components Design	74
4.4	SUMMARY	78

CHAPTER 5 DIALOGUE MECHANISM DESIGN	79
5.1 INTRODUCTION	79
5.2 RELATED APPROACHES	79
5.2.1 Syntactically Enhanced Latent Semantic Analysis (LSA).....	81
5.2.2 Assessment Model to Analyzing Student’s Performance	83
5.2.3 Ontology-based Information Retrieval System.....	88
5.2.4 Dialogue Manager	90
5.2.5 Summary.....	91
5.3 DIALOGUE MECHANISM DESIGN	93
5.3.1 Introduction.....	93
5.3.2 Cosine Methodology Based on LSA.....	93
5.3.3 Framework Design.....	96
5.3.4 Concept Classification & Information Retrieval.....	97
5.3.5 Relevance Feedback and Concept Reclassification	101
5.3.6 Summary	104
CHAPTER 6 DEVELOPMENT PLAN.....	106
6.1 INTRODUCTION	106
6.2 DEVELOPMENT OVERVIEW	106
6.3 FORUM SELECTION	109
6.3.1 Requirement Analysis.....	109
6.3.2 Survey of Existing Open Source Forum Systems	110
6.3.3 Base forum selection.....	114
6.4 DEVELOPMENT APPROACHES AND TECHNOLOGIES USED	117
6.4.1 Web Services	117
6.4.2 Agents.....	118
6.4.3 Object Relational Mapping (ORM) Tools	118
6.4.4 Development Approaches.....	119
6.4.5 Development Environment and Tools	121
CHAPTER 7 ARCHITECTURE AND IMPLEMENTATION	123

7.1 INTRODUCTION	123
7.2 ARCHITECTURE OVERVIEW	123
7.2.1 Plug-in Component	124
7.2.2 Dialogue Mechanism Component.....	126
7.2.3 Dispatch Component	131
7.2.4 Forum Display Component.....	132
7.2.5 Data Architecture.....	133
7.2.6 Deployment Architecture.....	136
7.3 IMPLEMENTATION	139
7.3.1 Overview	139
7.3.2 General Features	139
7.3.3 New Features	141
7.4 SUMMARY	145
CHAPTER 8 RELATED EVALUATION APPROACHES	146
8.1 INTRODUCTION	146
8.2 EVALUATION METHODOLOGIES	146
8.2.1 Evaluation for IR Performance	147
8.2.2 Evaluation for User Performance and Satisfaction	149
8.3 RELATED EVALUATION APPROACHES	150
8.3.1 Overview	150
8.3.2 User Performance versus Precision Measures for Simple Search Tasks	151
8.3.3 User-centred Evaluation of an E-learning Repository	151
8.3.4 Evaluating Effectiveness of a Context-based IR System	152
8.3.5 Evaluating an Intelligent Discussion-Bot	153
8.3.6 Assessing Distance Learning Student Performance.....	153
8.4 SUMMARY	154
CHAPTER 9 EVALUATION APPROACH	155
9.1 INTRODUCTION	155
9.2 EVALUATION GOALS	155

9.2.1 Research Objectives.....	156
9.2.2 Evaluation Goals.....	156
9.3 EVALUATION APPROACHES.....	158
9.3.1 Baseline Systems.....	158
9.3.2 Evaluation Approach.....	161
9.4 EVALUATION EXPERIMENTS.....	163
9.4.1 The First Experiment.....	164
9.4.2 The Second Experiment.....	166
9.5 SUMMARY.....	169
CHAPTER 10 EVALUATION RESULTS AND ANALYSIS.....	170
10.1 INTRODUCTION.....	170
10.2 THE FIRST EXPERIMENT.....	170
10.2.1 Experiment Procedure.....	170
10.2.2 Results and Analysis.....	173
10.3 THE SECOND EXPERIMENT.....	178
10.3.1 Experiment Procedure.....	178
10.3.2 Experiment Results and Analysis.....	182
10.3.3 User Satisfaction.....	187
10.4 DISCUSSION AND CONCLUSION.....	192
10.4.1 Further Discussion.....	192
10.4.2 Conclusion.....	195
CHAPTER 11 CONCLUSIONS AND FUTURE WORK.....	197
11.1 DISSERTATION SUMMARY.....	197
11.2 CONCLUSION: CONTRIBUTIONS, LIMITATIONS, AND FUTURE RESEARCH.....	200
REFERENCES.....	205
APPENDIX.....	228
APPENDIX I – STOP-WORD LIST.....	228
APPENDIX II – QUESTIONNAIRE.....	230

LIST OF TABLES

TABLE 2-1: FIVE COMPONENTS OF A COMMUNITY-BASED LEARNING ENVIRONMENT (KANG AND BYUN, 2001).....	12
TABLE 2-2: A QUERY STRING “COMPUTER-BASED INFORMATION LOOK-UP”	43
TABLE 3-1: ADDRESSING THE FIVE CHALLENGES IN DISCUSSION FORUMS.....	62
TABLE 4-1: CHARACTERISTICS OF A TRADITIONAL FORUM, EQUAKE & ID-B	71
TABLE 5-1: THE SUMMARY OF RELATED APPROACHES	92
TABLE 6-1: NEW SYSTEM DEVELOPMENT PLAN.....	108
TABLE 6-2: A LIST OF OPEN SOURCE FORUMS.....	110
TABLE 10-1: THE GENERAL INFORMATION OF PARTICIPANTS	172
TABLE 10-2: RELEVANCE RESULTS IN THE FIRST EVALUATION EXPERIMENT.....	175
TABLE 10-3: RATIO OF RELEVANCE RESULTS IN THE FIRST EVALUATION EXPERIMENT	175
TABLE 10-4: AVERAGE NUMBER OF RELEVANT POSTS FOUND FOR EACH SUB-TASK.....	183
TABLE 10-5: AVERAGE NUMBER OF RELEVANT POSTS FOUND FOR EACH SUB-TASK.....	184
TABLE 10-6: AVERAGE NUMBER OF RELEVANT POSTS FOUND FOR EACH SUB-TASK.....	185
TABLE 10-7: SUMMARY OF EVALUATORS.....	187
TABLE 10-8: STATISTICAL DATA FOR THE SECTION 1	188
TABLE 10-9: GENERAL INFORMATION OF EVALUATORS	189

LIST OF FIGURES

FIGURE 1-1: OUTLINE OF THE THESIS.....	9
FIGURE 2-1: THE PROCESS OF COMMUNITY-BASED LEARNING (PALLOFF & PRATT, 1999)	11
FIGURE 2-2: MODEL OF AN ELECTRONIC LEARNING COMMUNITY (SHEARD, 2004)	21
FIGURE 2-3: A WEBSITE USED FOR A COURSE (HTTP:IS157730.MASSEY.AC.NZ)	24
FIGURE 2-4: DIFFERENT TOPIC FOLDERS IN THE DISCUSSION FORUM.....	25
FIGURE 2-5: DISCUSSING AREA OF THE DISCUSSION FORUM	25
FIGURE 2-6: SCREENSHOT OF EQUAKE WEB FORUM (FROM HTTP://WWW.EQUAKE.ORG.NZ/M1).....	27
FIGURE 2-7: GENERAL MODEL OF INFORMATION RETRIEVAL (SALTON, 1983)	30
FIGURE 2-8: SEARCH ENGINE ARCHITECTURE IN GENERAL (ARASU, ET AL., 2001).....	50
FIGURE 4-1: USE OF TRADITIONAL FORUMS VS. EQUAKE FORUMS (SMESTAD, 2006)	66
FIGURE 4-2: EQUAKE SYSTEM ARCHITECTURE (SMESTAD, 2006)	67
FIGURE 4-3: SCREENSHOT OF EQUAKE WEB FORUM (SMESTAD, 2006)	68
FIGURE 4-4: DISCUSSION-BOT ARCHITECTURE (FENG ET AL., 2006)	69
FIGURE 4-5: USE OF PROPOSED SYSTEM FOR DIFFERENT INSTITUTIONS	74
FIGURE 4-6: SYSTEM LAYER DIAGRAM	75
FIGURE 4-7: PLUG-IN COMPONENT	76
FIGURE 4-8: DIALOGUE MECHANISM COMPONENT	77
FIGURE 5-1: VISUALIZATION OF FILTERED CONCEPT SPACE (ZHANG ET AL., 2002).....	85
FIGURE 5-2: AN EXAMPLE OF KEYWORD DENSITY	86
FIGURE 5-3: AN ONTOLOGY-BASED DOCUMENT RETRIEVAL (PARALIC AND KOSTIAL, 1999).....	89
FIGURE 5-5: AN EXAMPLE OF CONCEPT CLASSIFICATION IN A DISCUSSION FORUM.....	98
FIGURE 5-6: AN EXAMPLE OF RELATIONSHIPS BETWEEN SUB-CONCEPTS AND CONCEPTS	99
FIGURE 5-7: AN EXAMPLE OF CONCEPT RECLASSIFICATION – BEFORE USER EVALUATION	103
FIGURE 5-8: AN EXAMPLE OF CONCEPT RECLASSIFICATION – AFTER USER FEEDBACK.....	104
FIGURE 6-1: HIGH LEVEL SYSTEM OVERVIEW	107
FIGURE 6-2: A SAMPLE SCREENSHOT OF A BEEHIVE FORUM WEBSITE.....	112
FIGURE 6-3: A SAMPLE SCREENSHOT OF PHPBB FORUM	113
FIGURE 6-4: A SAMPLE SCREENSHOT OF PUNBB FORUM	113

FIGURE 6-5: MODEL-VIEW-CONTROLLER FRAMEWORK (GAMMA ET AL., 1995)	116
FIGURE 7-1: HIGH-LEVEL SYSTEM ARCHITECTURE	124
FIGURE 7-2: SYSTEM ARCHITECTURE FOR PLUG-IN COMPONENT	125
FIGURE 7-3: HIGH-LEVEL SYSTEM ARCHITECTURE FOR DIALOGUE MECHANISM	127
FIGURE 7-4: ARCHITECTURE FOR CONCEPT FORMULATION AGENT	128
FIGURE 7-5: ARCHITECTURE FOR QUERY MONITORING AGENT	129
FIGURE 7-6: WORK FLOW CHART IN DIALOGUE MECHANISM	130
FIGURE 7-7: ARCHITECTURE FOR DISPATCH COMPONENT	132
FIGURE 7-8: SYSTEM ARCHITECTURE FOR FORUM DISPLAY COMPONENT	133
FIGURE 7-9: INFORMATION STRUCTURE AND STORAGE STRUCTURE	134
FIGURE 7-10: ENTITY RELATIONSHIP DIAGRAM	136
FIGURE 7-11: DEPLOYMENT ARCHITECTURE FOR PROTOTYPE SYSTEM	137
FIGURE 7-12: LAYERS OF PROTOTYPE SYSTEM	138
FIGURE 7-13: TOPIC CREATION IN PROTOTYPE DISCUSSION FORUM	140
FIGURE 7-14: VIEWING ALL TOPICS IN PROTOTYPE DISCUSSION FORUM	140
FIGURE 7-15: VIEW MESSAGES IN PROTOTYPE DISCUSSION FORUM	141
FIGURE 7-16: POST A REPLY IN PROTOTYPE DISCUSSION FORUM	141
FIGURE 7-17: ARCHITECTURE FOR INTEGRATION WITH PREVIEW FUNCTION	142
FIGURE 7-18: MESSAGE REPLY PAGE	143
FIGURE 7-19: MESSAGE PREVIEW PAGE	143
FIGURE 7-20: USER EVALUATION FOR RETRIEVED RELEVANT MESSAGES	144
FIGURE 7-21: USER FEEDBACK SUBMITTING	144
FIGURE 10-1: STATISTICS OF THREAD LENGTH	172
FIGURE 10-2: DISTRIBUTION OF TESTING QUESTION LENGTH	173
FIGURE 10-3: THE RELEVANCE RESULTS IN THE FIRST EVALUATION EXPERIMENT	176
FIGURE 10-4: AVERAGE LENGTH OF SUBMITTED POST FOR EACH TASK	182
FIGURE 10-5: AVERAGE RELEVANT POSTS FOUND IN THREE FORUMS FOR THREE TASKS	186
FIGURE 10-6: RELATIONSHIP BETWEEN EFFECTIVENESS RATIO AND AVERAGE LENGTH	193

LIST OF ABBREVIATIONS

ACM	-	Association of Computing Machinery
AI	-	Artificial Intelligence
AWBES	-	Adaptive Web-based Educational Systems
BSD	-	Berkeley Software Distribution License
C4P	-	Content, Conversation, Connections, (information) Context, and Purpose
Equake	-	Electronic Question and Answer Knowledge Environment
ER	-	Entity Relationship
GPL	-	General Public License
HTTP	-	Hypertext Transfer Protocol Overview
IDE	-	Integrated Development Environment
IDF	-	Inverse Document Frequency
ID-B	-	Intelligent Discussion-Bot
IEEE	-	Institute of Electrical & Electronics Engineers
IIS	-	Information Interactive Systems
IR	-	Information Retrieval
IRS	-	Information Retrieval Systems
LSA	-	Latent Semantic Analysis
LSI	-	Latent Semantic Indexing
MVC	-	Model-View-Controller
NPL	-	Natural Language Process
SOAP	-	Simple Object Access Protocol
SVD	-	Single-Value Decomposition

TF	-	Term Frequency
URL	-	Uniform Resource Locator
W3C	-	World Wide Web Consortium
WSDL	-	Web Services Description Language
XML	-	Extensible Markup Language

CHAPTER 1

GENERAL INTRODUCTION

1.1 Introduction

Computers have been used in education for more than 40 years. Over the last 25 years tertiary educational institutions have increasingly employed computers as tools for learning in parallel with the evolution of personal computers (PCs) (Freiberger & Swaine, 2000). Currently computer based educational systems are used in distance learning as well as for class-room learning environments and since around 1999, the use of Intranet or Internet based e-learning systems has also been markedly on the rise within institutions (Kanejiya et al., 2003)..

An important consideration for Intranet or Internet based learning is the provision of effective communication within the electronic environment. The Internet has enabled many online forms of conversation and communication, such as emails, chat groups, news-groups, message boards, and more recently, weblogs (Glance et al., 2005). Asynchronous communication, typically in the form of discussion forums, is widely used to enable online communication between learners, and between learners and teachers. An additional benefit of discussion forums is that they can enable the development of online learning communities by providing a social structure that supports learners working in the remoteness of the electronic environment (Sheard, 2004).

1.2 Dialogue in the Learning Communities

According to many educational theories, learning is a process of dialogue between teachers and learners and between learners (Cutts et al., 2005). Meanwhile educational research literature has long identified interaction among learners as a key variable in learning (Brookfield, 1986; Slavin, 1983; Whitsed, 2004). Technology can play an important role in enabling people to interact with each other. The Web is one such technology which provides the capacity to share information and for connecting people to people (Giurgensohn & Lee, 2002). Various studies have also shown the importance of providing facilities to enable learners working online to interact with other learners and with their teachers (Ho, 2002; Lowder, 1999; McKenzie, 2000).

Learners working in an online environment are often isolated from their teachers and other learners. Asynchronous discussions are able to provide additional forums outside the classroom for strengthening interaction and enriching the learning experience for learners (Bhagyavati et al., 2005). The online discussions can potentially make learners' ideas visible and support collaboration or competition. The Web is one of many active technologies that can be used to develop critical thinking skills and teamwork (Lawson, 1997).

The existence of online communities is both dependent upon, and a result of, online technologies (Sheard, 2004). The availability of technologies has seen the development of online asynchronous discussion boards for use in teaching and learning (Barker, 2002). The establishment of an online learning community by means of a discussion forum is a strategy that may be used to support the effective dialogue between learners and between learners and teachers (Sheard, 2004). The

online interaction within the discussion forums in learning communities does however have some inherent problems. More discussion on this issue is presented in the next section.

1.3 Research Questions, Treatment, Methodology and Scope

1.3.1 Limitations of Discussion Forums in Educational Context

As one of the web communication tools, a discussion forum (known also by various other names such as discussion group, discussion board, message board, and online forum) is an asynchronous communication module. It works much like a bulletin board - users post messages that can then be read and responded to by other users (Glossary Website, 2008). Although online discussion forums are used in a range of undergraduate and graduate courses in many universities, there are various limitations to how well discussion forums can be used.

One drawback of all discussion forums in use today is that there is no way for the system to understand the meaning of a user's message. Fully understanding natural language is still an open problem in artificial intelligence.

Secondly, in discussion forums, many messages are repetitions of older messages. This commonly happens when a forum has grown to contain a sizable number of messages.

Thirdly, duplication or repeated questions can make the moderation of the forum more time consuming than necessary for the teachers.

Fourthly, courses are very much like separate islands in tertiary education. Learners in different courses do not have access to each other's forum, even if the course content overlaps and similar questions might come up in different courses.

Finally, reusing parts of the knowledge base accumulated from previous courses that have been run in past years is often desirable in education. This is not supported in forum systems, as they normally only give the options to keep everything or remove everything.

1.3.2 Research Questions

Reflecting the problems with existing online discussion forums described above, the key research question is: **How can the effectiveness of dialogue between learners and between teachers and learners in online discussion forum systems be improved?**

In order to effectively answer this question, several detailed research questions are explored in this study, as listed below:

1-1: How can knowledge be easily shared and reused for various courses in different institutions.

1-2: How can a forum provide immediate responses for users' messages.?

1-3: How can a forum accurately search and provide relevant information when users try to post a message.

1-4: How can the learners' feedback be effectively utilised to help users of the discussion forum easily identify relevant content from previous postings in the

discussion forum? This question is hard to understand. I tried to re-write it and realised I didn't really understand what it was saying.

1-5: How can the teachers' time consuming burden of managing the discussion forum and answering duplicated questions be reduced?

1.3.3 Treatment of Research Question

The research in this study aims at improving the effectiveness of communication between learners and between teachers and learners in online learning communities. This section describes the proposed approach to address the research questions. The methodological part of this study is discussed in section 1.3.4. The approach of this study has the following steps:

Step 1: Study and understand the utilisation of the discussion forums in online learning communities; explore and analyse the existing advanced technologies that could contribute to the solutions.

Step 2: Propose model system architecture on the basis of issues and gaps identified through literature review analysis, and by incorporating novel design methodologies.

Step 3: Implement a prototype of the proposed system.

Step 4: Evaluate the implemented prototype using evaluation experiments and quantitative analysis of collected data.

Step 5: Summarise the conclusion from the evaluation results and discuss future outlook for further research directions.

1.3.4 Methodology

The key task of this research is to develop a methodology which overcomes existing limitations of online discussion forum thereby improving communication between learners and between teachers and learners. The following technological solutions are reviewed in this research to identify their suitability:

Natural Language Processing: This area is very important for this research because ‘understanding’ the meaning of users’ message text is a core objective of this research.

Information Retrieval: Traditional and advanced information retrieval methodologies are reviewed and adopted in this research.

Semantic-based Text Similarity: Measures of text similarity have been used for a long time in applications in natural language process and related areas. Currently semantic-based text similarity is a major approach for text extraction and classification.

Relevant Feedback: Relevant feedback is another important area for automatic text classification in the text similarity area.

User Model: User modelling is concerned with representing the user's knowledge and interaction within the systems and adapting those systems to the needs of the users.

1.3.5 Scope - eQuake Project

This research study extends the research in the eQuake project which was funded by the Tertiary Education Commission of New Zealand through their e-Learning Collaborative Development Fund. The official title of the project was “agent based intelligent help system for the New Zealand student community”. The project's primary objectives were identified as:

- a) providing better support for discourse-based learning styles, as identified in priority groups such as Maori;
- b) providing wider student interaction across various institutions hence increasing exposure to multiple perspectives;
- c) reducing workload on teachers by reducing repetitive explanations and creating long-term archives of core student queries and solutions; and
- d) alerting teachers to potential problem areas in students' understanding process.

The resulting system was developed on a model similar to asynchronous web-based discussion forums, since that was seen as the technology most easily extended to fulfil the project objectives.

1.4 Dissertation Structure

Based on the five steps of the research approach (Section 1.3.3) and the technological solutions (Section 1.3.4), this dissertation is organised as follows (Figure 1-1). The first (this) chapter is the general introduction to this study. The second chapter reviews literature in the context of this study. Important areas reviewed include current technologies and existing systems.

Following the background discussion, the third and fourth chapters then describe the requirements analysis and propose the solution. This is followed by the introduction of the core component of the proposed solution in the chapter five and the development plan in the chapter six. The chapter seven illustrates the architecture and the implementation of the developed system.

The eighth, ninth and tenth chapters present an evaluation of this research. The eighth chapter describes related work and the evaluation experiment design process. The ninth chapter introduces the step by step process of the evaluation experiment and the tenth chapter presents the analysis of data from the evaluation experiments, along with interpretation of findings and recommendations.

Finally, the eleventh chapter presents a summary of this dissertation and highlights major contributions and limitations of this study. Future research directions are also pointed out in this chapter.

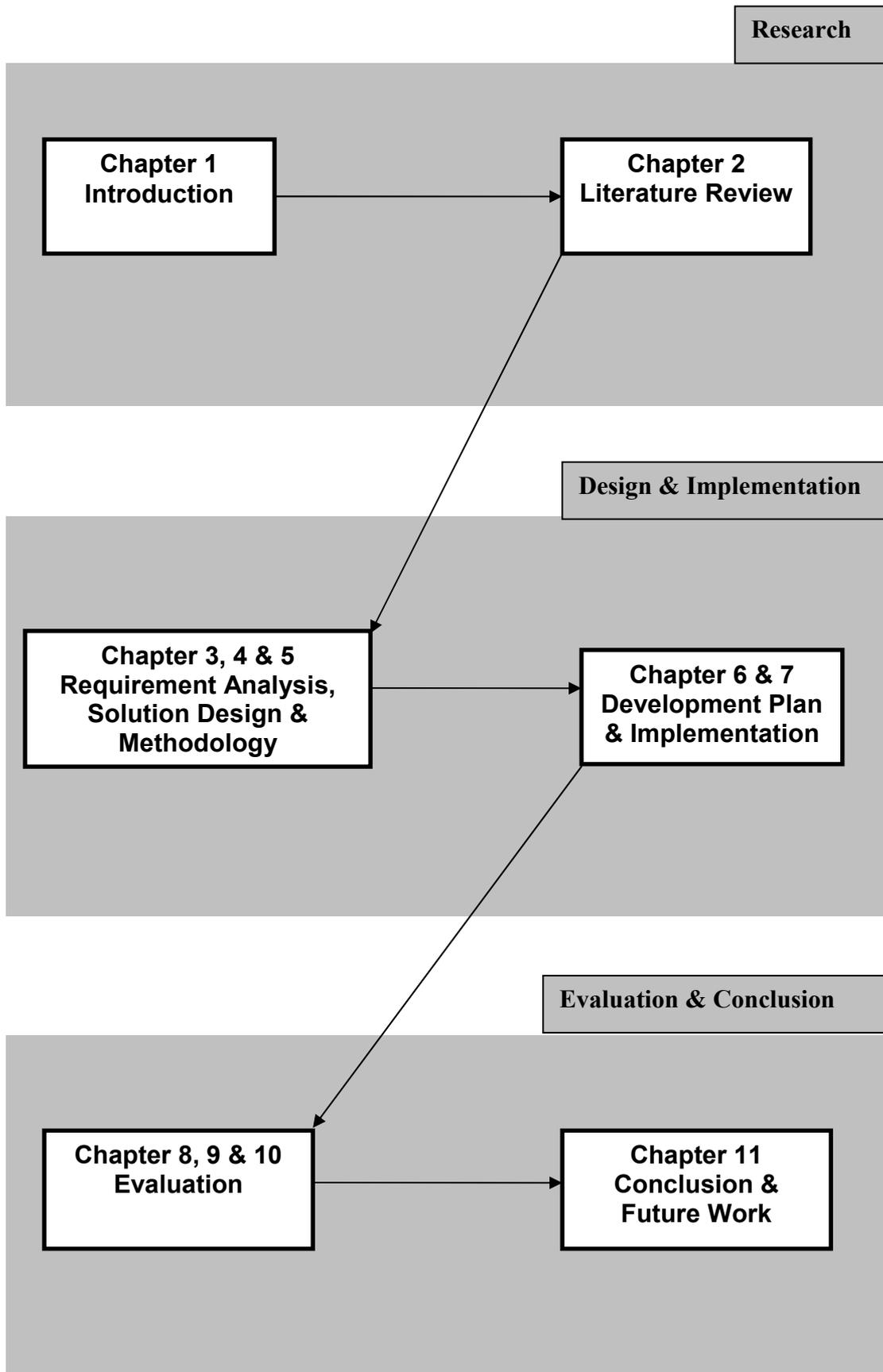


Figure 1-1: Outline of the Thesis

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter reviews the previous literature. Based on the literature review, the researchers were able to analyze the major limitations of discussion forums used in learning communities and then propose a solution which combined information retrieval, text similarity and user model concept to overcome those limitations. This chapter firstly provides an overview of online learning communities, various discussion forum systems, and other online learning environments. This is followed by an overview of some technological solutions which are used in this research, such as information retrieval and text similarity. Finally the chapter discusses the recommended systems.

2.2 Online Learning Communities

2.2.1 Community-based Learning

With the rapid development of a knowledge-based society, knowledge creation through collaboration is growing in importance. Recently, with the expansion of online communities, community-based learning has evolved new learning methods where learners form communities around common objectives and create knowledge through communication with other members (Rovai, 2003). According to Sue et al.'s (2005) definition, community-based learning is a learner centred approach and is based on the principles of collaborative knowledge construction and the learning

community. Students within the learning community set up common goals, generate ideas collaboratively, and share their production (Wilson and Ryder, 1998).

The process of community-based learning consists of five phases which are described in Figure 2-1. These five phases are: (1) Common goals assignment phase, (2) Group rules forming phase, (3) Assignment recognition phase, (4) Collaborative accomplishment phase, and (5) Production generation phase (Palloff & Pratt, 1999).

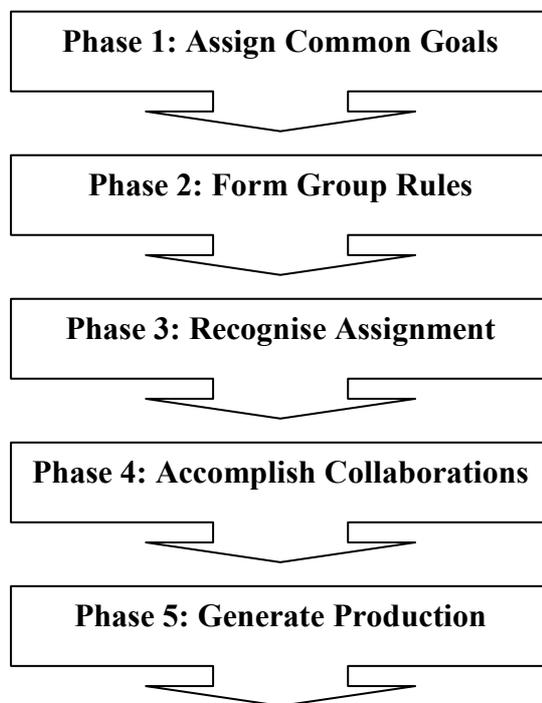


Figure 2-1: The Process of Community-based Learning (Palloff & Pratt, 1999)

These five phases are integrated as a whole process. In the first phase, common goals can be defined and community members' demands and interests can be distinguished. Then students set up group rules and all group members are made aware of the rules in the second phase. The third phase follows where students can find out the problem or task. In the next phase, students can work together in group collaboration and

communicate with external instructors. In the final phase, students can create and share the final outcome and the outcome can be evaluated by peers (Palloff & Pratt, 1999).

Table 2-1 indicates the five components of a community-based learning environment including collaborative learning, social interaction, planning & reflection, knowledge base, and evaluation & compensation (Kang & Byun, 2001).

Table 2-1: Five Components of a Community-based Learning Environment (Kang and Byun, 2001)

Dimension	Description	Function
Collaborative Learning	Exchange opinions and interact with members during group work or discussion	Discussion forum, whiteboard or blackboard, group workplace, communication tools(e.g. MSN), etc.
Social Interaction	Informal communication and social discussion	Free-board, announcement, Q/A, etc.
Planning & Reflection	Set up the common goal for person and group, make the plan schedule for the project, confirm member roles, and reflect learning process	Reflection note, schedule, goal/role specification, etc.
Knowledge Base	Share and store learning resources and products	e-library, mybase, teambase
Evaluation & Compensation	Peer evaluation, instructor evaluation, internal/external compensation	Evaluation rubrics, comment form, avatar-point, etc.

2.2.2 Online Learning Communities

Technology has an important role to play in supporting knowledge-building communities and communities of practice in general. In the past decade since widespread adoption of the World-Wide Web, there has been a trend towards using the Internet to support traditional classroom teaching and substitute traditional teaching with online learning (Hoadley & Kilner, 2005). Normally in higher education, “curricular learning communities are classes that are linked or clustered during an academic term, often around an interdisciplinary theme, and enrol a common cohort of students” (Wikipedia Contributors, 2008).

While many of these phenomena can take place on an individual level, in other cases, the “killer app” has been the used by technology to support groups of people in in e-business, e-government, and e-learning domains (Wenger, 2000), and most notably, online communities (Rheingold, 1994). Although online communities have been used in a variety of applicable areas, ranging from medical support groups to online dating, one of the more promising reasons for online communities is as a support for knowledge-building communities. Online communities can provide a forum where information may be exchanged and questions may be asked. This is often a significant component of online learning environments (Harasim, 1995; Hiltz, 1993).

Online learning communities currently play an important role in traditional and distance education (Sheard, 2004). A common place on the Internet that addresses the learning needs of its members through proactive and collaborative partnerships is called an online learning community. In the learning community, students work together to capture knowledge and exchange information in order to achieve a

common learning objective through social networking and online communication. When online learning communities are created, students do not miss out on existing social and extracurricular cognitive opportunities, and financial costs are easily afforded by institutions of higher education (Bauman, 1997).

In general, there are two types of online learning communities.

1. E-learning communities – groups for interaction and discussion which are connected solely via technology.
2. Blended learning communities – groups that utilise face-to-face meetings as well as online meetings.

Online learning communities may also be categorized as knowledge-based, practice-based, or task-based and online learning communities may focus on personal aspects, process, or technology (Riel & Pollin, 2004).

A wide variety of technologies and tools have been used in online learning communities such as synchronous (e.g. instant) messaging, asynchronous message boards, content management (e.g. Moodle and Lectureshare), collaborative information pooling (e.g. wikis), social networking (e.g. Facebook and Flickr), learning object repositories, and blogs.

2.2.3 C4P Model

This section presents a traditional model – the C4P model. It is abbreviation for content, conversation, connections, (information) context, and purpose of learning

takes place in knowledge-building communities (Hoadley & Enyedy, 1999). The C4P model is based on the assumption that knowledge is created and shared when there is purposeful conversation around the content in context. These elements comprise a non-linear system that occurs in a community of practice (Hoadley & Kilner, 2005). If one of the elements of these five has increased, it will result in increases in all of them, although the relationships among them are not one-to-one. The C4P model defines each of these terms in a specific way. Content refers to explicit, static knowledge objects such as learning materials or documents, whereas conversation refers to face to face communication or online discussions. The key difference between content and conversation is that content is about a one-way communication of information (monologue), whereas conversation necessarily contains at least a two-way exchange of information (dialogue) (Hoadley & Enyedy, 1999). Connections, as used in C4P, refer to interpersonal communication between community members who are connected in some level of relationship, e.g., when one member sends a message or email to another member, a connection has occurred. Information context is the: who, what, where, when, why, and how that is able to let community members assess whether and how information is relevant to them. This context provides the many details that make information meaningful and memorable. Finally, purpose is the reason why the members come together in the community.

The five elements of C4P can influence and improve each other. None of these five elements should be ignored when building an effective knowledge-based online community. If content is missing, it is difficult to get the conversation started and keep conversation focused on the community's purpose. If we do not have enough conversation, it is difficult to create the new knowledge even though the knowledge

may transfer (Constant & Sproull, 1996). If there is no connection, contributions of content and conversation will be unlikely to occur, and the contributions will have less context. If information context is absent, it is easy for the community to misunderstand content or apply knowledge incorrectly to new scenarios. Finally, there will be no knowledge building without purpose. A clear communal purpose makes meaning of content, provides direction to conversation, fosters connections, and is the unifying context for all activities in the community (Hoadley & Kilner, 2005). The following paragraphs present an explanation of each of the elements, why they are important to a knowledge-building community, and how they integrate as a system.

In a knowledge-building community, quality content is recognised as a foundational element. Content is used for four important purposes: community members can capture immediate value from the content; new members can be socialised when they understand the topic of the content; it is used as a basis for conversation; and it encourages members as they consider themselves jointly building their domain of knowledge.

One of the great challenges of enhancing a knowledge-building community is to have more quality content because, for a variety of reasons, people are hesitant to contribute content (Goman, 2002). The other elements of C4P, such as conversation, help to solve this challenge and improve content generation. In this study, conversation is the core research foundation in that it is the most effective model of knowledge transfer and generation, because the personal connection and back-and-forth nature of conversation provides the best context for information (Allee, 2003).

The challenge within a knowledge-building community is to generate conversations that describe meaningful knowledge, not useless talk. Normally conversation is based on content. Meaningful conversation is enabled by quality content, clear purpose, and personal connections. Connections are an essential element of a knowledge-building community. Without connections, there is no active community and an online space is merely a document repository (content) or chat room (conversation). Connections can foster relationships and subsequently trust enabling people, wherever they live, to learn and work together on the common goal of building their knowledge domain (Saint-Onge & Wallace, 2003). Connections are facilitated by the other elements described by C4P and they do not spontaneously occur in an online community.

Information context enables learners to learn more efficiently and effectively (Brown et al., 1996). Community members are able to clearly identify the location of a knowledge object and its history according to information context. Information context enables members of a community to know what a contributor is communicating, to judge whether the information is related to them, and to apply the knowledge to their current situation.

Another challenge in a learning community is to enable people who are not physically co-located to learn (Hoadley & Kilner, 2005). The elements of the C4P framework help members understand the context of information. A community's purpose inheres in every piece of content, every conversation, every connection. Clarity of purpose creates energy and produces results and shared purpose is a defining element in collaboration and community (Woodruff, 1999). Purpose alone has the ability to relate everything that occurs within the community, and shared purpose is a big step

to generating trust and connections. Even if a community has a stated purpose, its actual purpose will be evident in its content, conversations, connections, and context. Ideally, the actual and stated purposes are identical. It is most likely that in a learning community, a meaningful purpose will be supported by the 4Cs due to the elements of C4P building off each other. Conversely, every piece of content, every conversation, every additional element of context serves either to improve or to impair the community's stated purpose.

In summary, the C4P model provides an explanation of learning related processes that have occurred in successful knowledge-building communities of practice (Dixon et al., 2005). Communication is the most important factor which supports every other component of this model for building a successful learning community. How people can achieve effective communication to improve learning has been discussed for a long time. The next section introduces how technologies are able to enhance the capacity of communication with others.

2.2.4 Dialogue via Technologies

The provision of effective communication (dialogue) in online learning communities is an important consideration for effective learning (Sheard, 2004). Students in the online learning community engage with each other and develop strong bonds of friendship so that they are able to provide one another with support. They also create intensive working relationships that involve frequent communication (Haythornthwaite et al., 2000). People use technologies to build their shared learning world, taking time to learn how to use available technologies and how to implement them effectively in both traditional and distance education.

When building a learning community in which they interact primarily online, students develop norms for sharing communication, activity, and technology, and for creating space (Kazmer, et al., 2001). Students communicate mostly using online mechanisms such as email, text chat and discussion forums. They must co-create acceptable and shared procedures governing many facets of interaction. Those facets include (Kazmer, 2005):

- Appearance of messages: e.g., length and grammar;
- Content of messages: social, task-related, or both;
- Temporal issues: e.g., when to communicate, how often, acceptable response time, synchronous such as chat versus asynchronous such as email or discussion forum;
- Audience and medium: one to one such as email, one-to-many via email or discussion forum.

Communication via technology provides the platform for many of the activities students share, including attending lectures, completing group projects, and participating in Web-board discussions. Thus the norms and procedures for communicating and using technology control much of students' activities in general. In building an online learning community, students learn to use and depend on technologies that support shared interaction. Using an online threaded discussion board enables participation in rewarding intellectual exchange within the learning environment, not just reading a sequence of postings on the computer screen (Kazmer, 2005).

Online learning communities rely heavily on technology to support their members' activities and communication with one another. It is important to provide online mechanisms that enable learners working online to interact with other learners and with their teachers. A mechanism widely used for this purpose is the asynchronous discussion forums. The provision of a discussion forum offers the learner access to sources of assistance with the removal of the restrictions of time and space. In addition, online discussion forums can encourage learners to share and exchange their information and perspectives during the learning process.

Figure 2-2 (Sheard, 2004) introduces a model of electronic learning communities which depicts learner and educator interactions within an online discussion forum. These interactions (dialogues) are between learners and between teachers (lecturers and tutors) and learners.

In the diagram, the closed arrows represent postings by students (A and C) and the open arrows represent postings by teaching staff (B and D), with the solid lines representing lecturers and the dotted lines representing tutors. The interactions may be further explained by the fact whether they are proactive or reactive. Proactive interactions (left of diagram) represent postings that are questions or comments initiated by students or staff. Reactive interactions (right of diagram) are responses to comments or questions posted by others. The numbers of arrows give some indication of the relative volumes of postings.

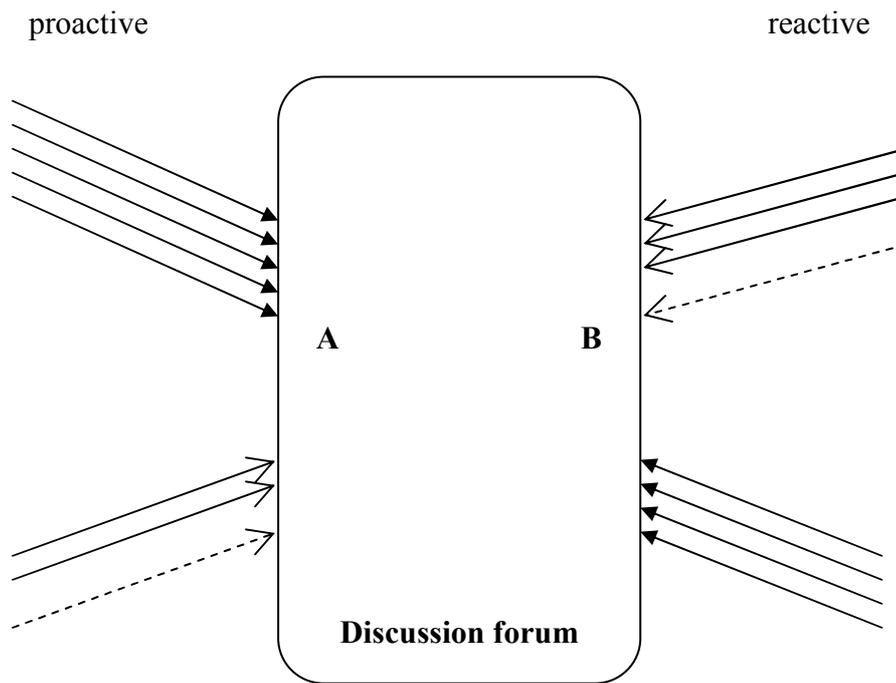


Figure 2-2: Model of an Electronic Learning Community (Sheard, 2004)

The provision of effective communication (dialogue) via facilities such as online discussion forums is fundamental to the process of establishing electronic learning communities. How to design a knowledge-building community, and what role technology can play to improve the effectiveness of communication in the learning communities are major research questions of this study. The next section introduces the communication facilities of online learning communities – discussion forum systems.

2.3 Discussion Forum Systems

2.3.1 Overview

The widespread use of the World Wide Web in the education area has enabled the development of on-line learning environments. Many educators use Web technology

to broaden the learning experience of their students beyond what is possible in a traditional teaching environment. Courses may now be enhanced by the addition of varied and interesting Web-based resources containing interactive elements that encourage student engagement in learning (Sheard, et al., 2001). The adoption of the Web technology in tertiary education has encouraged a shift in pedagogical thinking by educators and educational courseware designers (Lefoe, 1998). Meanwhile Web technologies provide an opportunity for students to learn and communicate with teachers on the Internet. Such online communication technology is already being used in many areas.

Nowadays, discussion forum is a facility commonly provided as an integral component of typical online learning communities. This enables synchronous or asynchronous communication between learners and instructors who are physically separated (Liaw & Huang, 2000). Synchronous communication, for example, chat sessions or conferencing, requires all participants to be present at the same time. Asynchronous communication, for example, discussion forums, do not require participants to be present at the same time and has the advantage of allowing the learner and instructor the opportunity to analyse and reflect before they express their ideas.

2.3.2 The Use of Discussion Forums

In tertiary education, web-based discussion forums are used by students in a range of courses across institutions to discuss course content or structure, or offer comments on any issues arising during the study semester. In addition, the facility is seen as "a place to encourage further learning", particularly as many questions not asked in

lectures or tutorials could be asked in the forum (Sheard, et al., 2003). Users can ask queries and browse the forum to capture, share and retrieve knowledge. In many cases, anonymity of postings is provided to encourage students to use it freely without revealing their identity (Lowder & Hagan, 1999). The impetus for this development came partly from a workshop titled "Computer-supported cooperative work" (Brugge & Houghton, 1996) which led some staff to consider ways to introduce student interaction in the computer-supported course environment and turn it into a cooperative learning environment.

A particular type of previous collaboration systems, called on-line discussion databases, facilitates formulation and sharing of metaphors and analogies to facilitate capturing tacit knowledge. In this type of discussion forums, features such as access control enable discussions to be restricted to team members and this has been proven to encourage frankness and engender feelings of trust amongst the members (Bradner et al., 1999).

Not only in the area of education but also in other areas, newsgroups and similar forums offer comparable opportunities but they generally differ in the fact that they are open to all. Many people have been found to offer advice and guidance in newsgroups due to a mixture of motivations such as altruism, seeking recognition and offering acknowledgement. This is despite the fact that newsgroups also present the potential threat of flaming, personal abuse and irrelevant postings. The archives of the forums serve as useful repositories of knowledge and the number of questions a person has answered and the speed of response could be utilised as a means to index the person as an expert in the domain (Marwick, 2001).

2.3.3 General Introduction of Major Systems

Discuss

The course ‘Web-based Multimedia Systems’ at Massey University (New Zealand) for postgraduate students incorporates “Discus” discussion forum on the course website. This website’s contents (Figure 2-3) include Course News, Information and the Discussion Forum. The forum is for communication between teachers and students. It is available through the campus local area network and the Internet. All discussion topics and replies can be posted from any computer connected through the Internet.

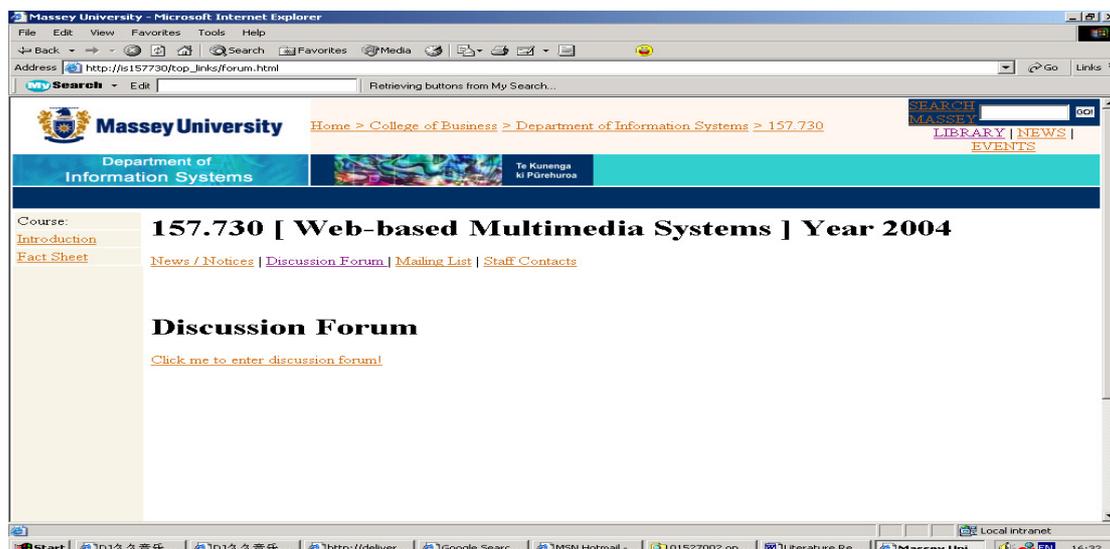


Figure 2-3: A website Used for a Course (<http://is157730.massey.ac.nz>)

After students register, they become authorised users in the discussion forum system. The discussion forum has five topic folders (Figure 2-4) for students to post their messages, such as Introduction, Discussion on Specific Topic, Projects, Discussion on Assignment and Other Issues.

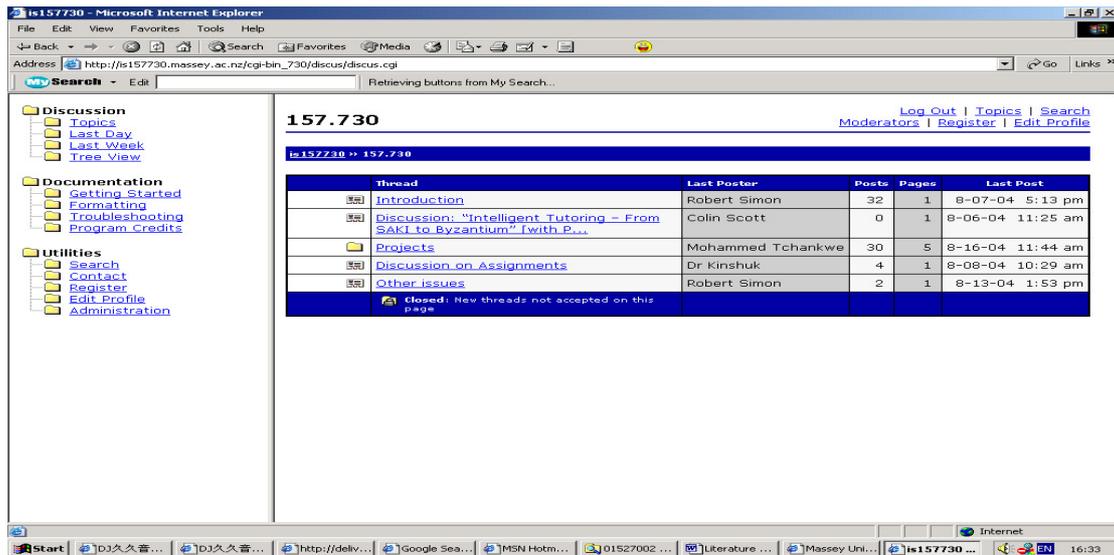


Figure 2-4: Different Topic Folders in the Discussion Forum

In each topic folder (Figure 2-5), students can post their messages and instructors also respond to students' queries or post important information related to the course.

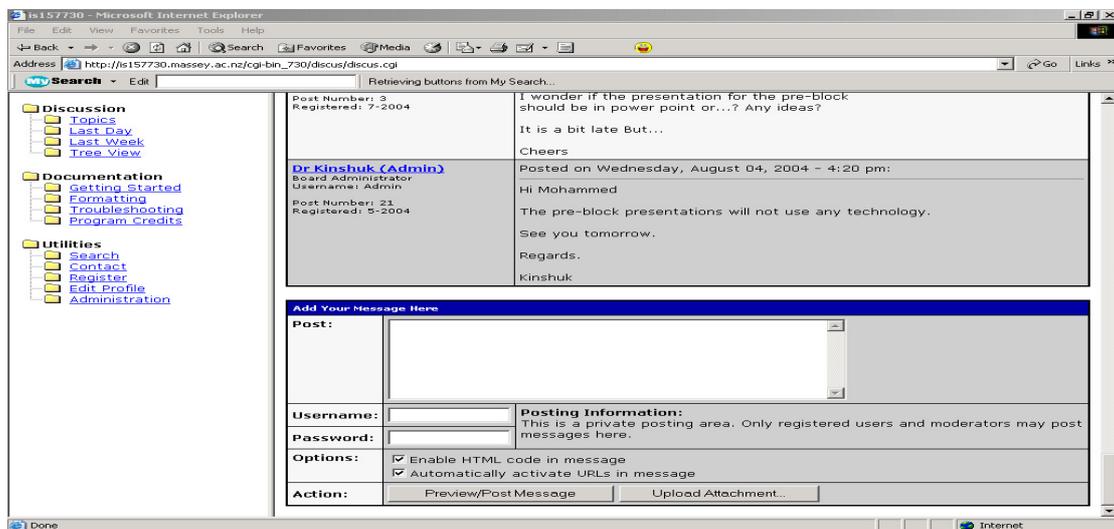


Figure 2-5: Discussing Area of the Discussion Forum

CRPP's web portal

Another type of discussion forum system, which is called a web portal, was developed by CRPP (Centre for Research in Practice & Pedagogy, National Institute of Education, Singapore). It has a number of discussion forums tailored to different contexts of use in the CRPP's web portal. Protecting the confidentiality of knowledge

content residing in the knowledge portal is of prime concern, and therefore different users have different access rights to log into the online discussion forum. A general forum is designed for the public who do not have login rights to post and discuss issues of relevance and importance. All members who are assigned rights of log-in into the portal are able to access project specific discussion forums. The discussion forums have database-driven search engines which have a search that is limited to information hosted in the discussion forum only. The ability of the search is pegged to the access rights of the individual participating in the forum (Muthukumar & Hedberg, 2004). The CRPP's web portal is developed for the public and all members to have a different right to post, discuss and search relevant information. It offers a better solution than other traditional discussion forums.

EQuake

EQuake, a web-based learning management system, includes a new type of discussion forum, enhanced with software agents which are able to: automatically find when a question has been answered before; link together several institution's forums; notify users of answers to questions they are interested in; and notify moderators when several similar questions are unanswered in the system so they can create an FAQ entry. Figure 2-6 depicts the discussing area of this new type of discussion forum in the eQuake system.

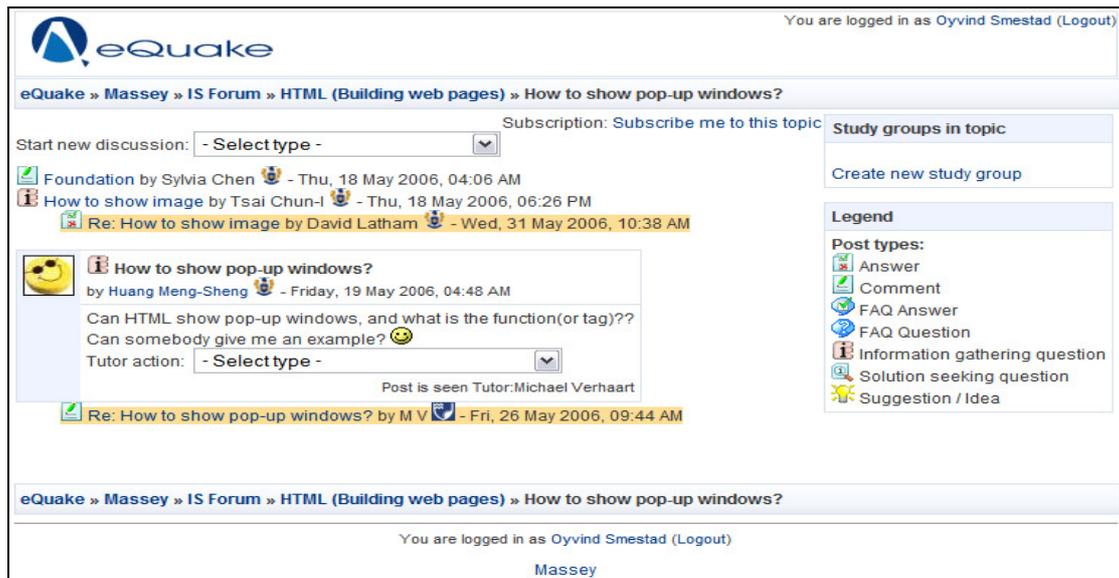


Figure 2-6: Screenshot of EQuake Web Forum (from <http://www.equake.org.nz/m1>)

2.3.4 Limitations of Current Discussion Forums

Although web-based discussion forums are used in a range of undergraduate and graduate courses in many universities, the enthusiastic adoption of these facilities by students has impacted upon teachers' workloads, and concerns have been raised about the time and effort required by teachers to manage them. Furthermore, it has been difficult to determine the pedagogical value of these facilities and the level of educational support they provide to the teaching and learning community (Sheard et al., 2003). Another issue is that not all students are comfortable with the discussion forum medium. They might be facing a number of difficulties, for instance, not feeling at ease with the technology, finding it hard to use, and waiting a long time for a response (Mazza & Dimitrova, 2004).

Additionally, duplication or repeated questions can make the moderation of the forum more time-consuming than necessary for the teachers. Assuming a teacher-led style of

discussion forum, this would cause the teachers to have to spend time answering questions that have been answered before. This in turn limits how the teachers use the forum as they might not find time to encourage student discussions by asking challenging questions. If a system can minimise the number of duplicated questions that have to be answered by a human, more time can be spent encouraging new discussions.

Courses are very much like separate islands in tertiary education. Students in different courses do not have access to each other's forums, even if the course content overlaps and similar questions might come up in different courses. This is true both inside an institution and between different institutions offering similar degrees or courses. Allowing forums to be shared between several courses would increase the available knowledge base and expose students to multiple perspectives. Students can then capture and share knowledge on a common platform.

Reusing certain parts of the knowledge bases from the previous instances of a course is often desirable in education. This is not supported in existing forum systems, as they normally only give the options to keep everything or remove everything. It is desirable, for instance, to be able to keep some high quality discussions or messages that answer important questions. These could be identified based on the number of similar questions being posted in the course forums and the feedback from students regarding a specific topic. It would then also be helpful to be able to hide the identities of those users who originally contributed these messages. A partial reuse of existing messages in this manner might be very useful in some settings, though the mechanisms for message selection would be a key factor.

Some existing systems, such as the eQuake system, have been developed as an attempt to solve the above problems. However they do have serious limitations in their capabilities such as less effective communication and poor performance in information seeking.

Several issues for improving discussion forums in education have been identified. Better tracking solutions (based on systems with more effective information retrieval) and ways to gather opinions from users would save time and effort of teachers and students. Having forums cross-used by multiple courses would allow for more knowledge resources to be connected.

2.4 Information Retrieval

2.4.1 Overview

Usually, an information retrieval system is considered to have the function of “leading the user to those documents that will best enable him/her to satisfy his/her need for information”. Somewhat more generally, “the goal of an information retrieval system is for the user to obtain information from the knowledge resource which helps him/her in problem management” (Robertson, 1981). Such functions, or goals, of information retrieval have been described in models of the type shown in Figure 2-7 (Salton, 1983). This model indicates basic entities and processes in the information retrieval situation.

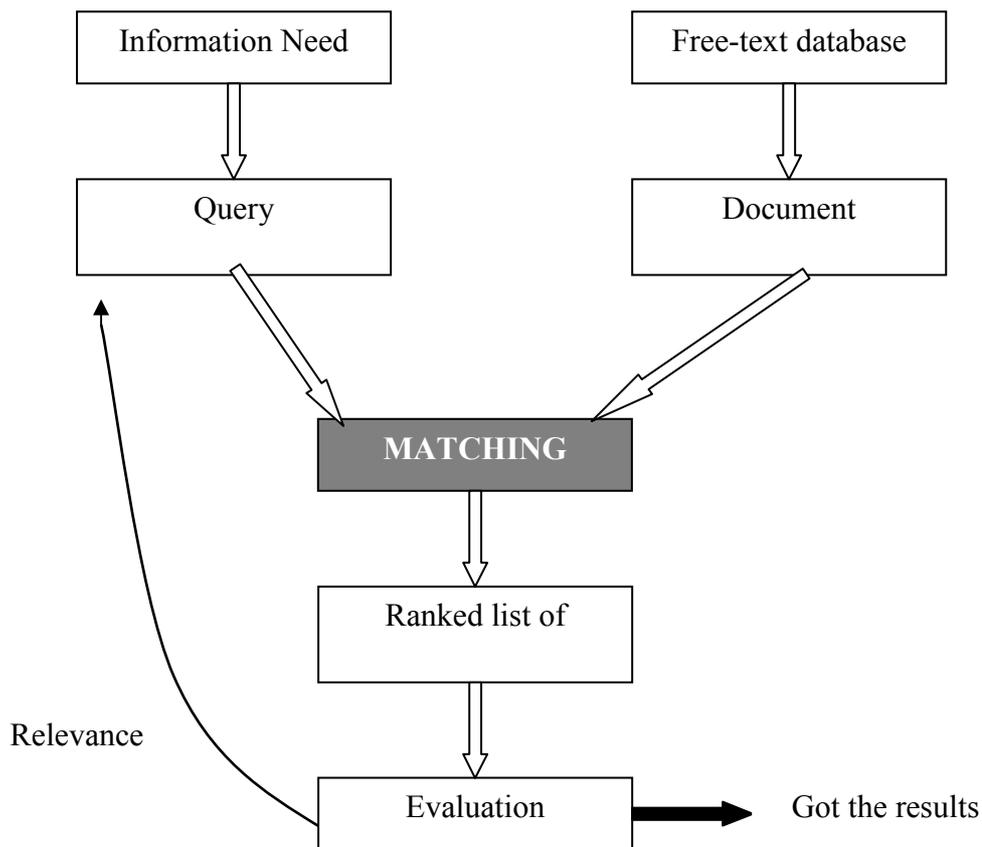


Figure 2-7: General Model of Information Retrieval (Salton, 1983)

In this model, a person with goals and intentions related to, for instance, a learning task, finds that these goals cannot be attained because the person’s resources or knowledge are in some way inadequate. A characteristic of such a “problematic situation” is an anomalous state of knowledge or information need, which prompts the person to engage in active information-seeking behaviour, such as submitting a query to an information retrieval system. The query, which must be expressed in a language understood by the system, is a representation of the information need. This is shown on the left-hand side of Figure 2-7. Due to the inherent difficulty of representation, the query in an information retrieval system is always regarded as approximate and imperfect (Belkin et al., 1987),

On the other side of Figure 2-7, the focus is on the information resources that the user of the information retrieval system will eventually access. Here, the model considers the producers or authors of texts, the groupings of texts into collections (e.g., databases), the representation of text, and the organization of these representations into databases of text surrogates. The process of representing the meaning of texts in a form more amenable to processing by computers (sometimes called indexing) is of central importance in information retrieval. A typical surrogate would consist of a set of index terms or keywords.

The comparison of a query and surrogates leads to the selection of possibly relevant retrieved texts. These retrieved texts are then evaluated or used. The user would then leave the information retrieval system, or the evaluation would lead to some modification of the query, the information need or, in rare cases, the surrogate. The process of query modification through user evaluation is known as relevance feedback in information retrieval (Salton, 1983).

Some information retrieval models focus primarily on the comparison process. The three major models are the boolean, vector space and probabilistic retrieval models. The first of these is based on what is called the “extract match” principle, whereas the other two are based on the concept of “best match” (Belkin et al., 1987).

Boolean retrieval is based on the concept of an extract match of query specification with one or more text surrogates. The term “Boolean” is used because the query specifications are expressed as words or phrases, combined with the standard operators of Boolean logic. In this retrieval model, all texts containing the

combination of words or phrases specified in the query are retrieved and there is no distinction made between any of the retrieved documents. Thus the result of the comparison operation in Boolean retrieval is partition of the database into a set of retrieved documents and a set of not-retrieved documents.

Boolean retrieval is the simplest retrieval method and relies on the use of Boolean operators. The terms in a query are linked together with AND, OR and NOT (Heaps, 1978). This method is often used in search engines on the Internet because it is fast and can therefore be used online. This method has its problems. The user has to have some knowledge in the search topic for the search to be efficient, e.g., a wrong word in a query could rank a relevant document non relevant. The retrieved documents are all equally ranked with respect to relevance. And the number of retrieved documents can only be changed by reformulating the query (Agosti et al., 1996).

Best-match retrieval models have been proposed in response to the problems of extract-match retrieval. The most widely known of these is the Vector Space model (Salton, 1983). This model treats texts and queries as vectors in a multidimensional space, the dimensions of which are the words used to represent the texts. Queries and texts are compared by comparing the vectors using, for example, the cosine correlation similarity measure. The assumption is that the more similar a vector representing a text is to a query vector, the more likely it is that the text is relevant to that query. In this model, an important refinement is that the terms (or dimensions) of a query, or text representation, can be weighted to take account of their importance. These weights are computed on the basis of the statistical distributions of the terms in the database, and in the texts.

The vector space model process can be divided into three major stages. The first stage is document indexing where content bearing terms are extracted from the document text. The second stage is weighting of indexed terms to enhance retrieval of documents relevant to the user. The last stage ranks the document with respect to the query according to a similarity measure. The vector space model has been criticised for being ad hoc.

Stage 1: Document Indexing

It is obvious that many of the words in a document do not describe the content, for example, words like “the” or “is”. By using automatic document indexing those non-significant words (function words) are removed from the document vector, so the document is only represented by content bearing words (Salton, 1983). This indexing can be based on term frequency, where terms that have both high and low frequencies within a document are considered to be function words (Salton, 1983; Van Rijsbergen, 1979). In practice, term frequency has been difficult to implement in automatic indexing. Instead a stop list of common words to be removed as high frequency words (stop words) is used (Salton, 1983; Van Rijsbergen, 1979). This makes the indexing method language dependent. Research shows that 40%-50% of the total number of words in a document is removed with help of a stop (Salton, 1983).

Non-linguistic methods for indexing have also been implemented. Probabilistic indexing is based on the assumption that there is some statistical difference in the distribution of content bearing words and function words. Probabilistic indexing ranks the terms in the collection with respect to the term frequency in the whole collection. The function words are modelled with a Poisson distribution over all documents. The

use of the Poisson model has been expanded to the Bernoulli model (Charkrabarti et al., 1998).

Stage 2: Term Weighting

Term weighting has many effects related to recall and precision of retrieval. The term weighting for the vector space model is entirely based on single term statistics. There are three main factors in term weighting: the term frequency factor, the collection frequency factor and the length normalisation factor. These three factors are multiplied together to make the resulting term weight.

A common weighting scheme for terms within a document is to use the frequency of occurrence. The term frequency is somewhat content descriptive for the documents and is generally used as the basis of a weighted document vector (Salton & Buckley, 1996). It is also possible to use a binary document vector, but the results have not been as good as term frequency when using the vector space model (Salton & Buckley, 1996).

Various weighting schemes are used to discriminate one document from the other. In general this factor is called the collection frequency document. Most of them, e.g. the inverse document frequency, assume that the importance of the term is proportional to the number of documents the term appears in. Experimentally it has been shown that these document discrimination factors lead to a more effective retrieval, i.e. improvement in precision and recall (Salton & Buckley, 1996).

The third possible weighting factor is the document length normalisation factor. Long documents usually have a much larger term set than short documents, which makes long documents more likely to be retrieved than short documents (Salton & Buckley, 1996).

Different weight schemes have been investigated and the best results, with respect to recall and precision, are obtained by using term frequency with inverse document frequency and length normalisation (Lee et al., 1997).

Stage 3 Similarity Coefficients

The similarity in vector space models is determined by using associative coefficients based on the inner product of the document vector and the query vector, where word overlap indicates similarity. The inner product is usually normalised. The most popular similarity measure is the cosine coefficient, which measures the angle between the document vector and the query vector.

2.4.2 Text Similarity

Measures of text similarity have been used for a long time in applications in natural language processing and related areas. One of the easiest applications of text similarity is perhaps the vector model in information retrieval, introduced in the above section, where the document most relevant to the input query is determined by ranking documents in a collection in reverse order of their similarity to the given query (Salton & Lesk, 1971). Text similarity has also been used for relevance feedback and text classification (Rocchio, 1971), word sense disambiguation (Schutze, 1998) and more recently for extractive summarisation (Salton et al., 1997), automatic

evaluation of matching translation or text summarisation (Lin & Hovy, 2003). Measures of text similarity were also found useful for evaluation of text coherence (Lapata & Barzilay, 2005).

With few exceptions, the typical approach used in finding the similarity between two text segments is a simple lexical matching method which produces a similarity score based on the number of lexical units that occur in both input segments. Proposed improvements to this simple method include stemming, stop-word removal, part-of-speech tagging, longest subsequence matching and various weighting and normalisation factors (Salton & Buckley 1997). While successful to a certain degree, these lexical similarity methods cannot always identify the semantic similarity of texts. For instance, there is an obvious similarity between the text segments ‘I own a dog’ and ‘I have an animal’, but most of the current text similarity metrics will fail to identify any kind of connection between these texts.

There are a large number of word-to-word semantic similarity measures using approaches that are either knowledge-based (Leacock & Chodorow 1998) or corpus-based (Turney 2001). Such measures have been successfully applied to language processing tasks such as malapropism detection (Budanitsky & Hirst 2001), word sense disambiguation and synonym identification (Turney 2001). For text-based semantic similarity, perhaps the most widely used approaches are the approximations obtained through query expansion, as performed in information retrieval (Voorhees 1993), or the latent semantic analysis method (Landauer et al., 1998) that measures the similarity of texts by exploiting second-order word relations automatically acquired from large text collections. A related line of work consists of methods for

paraphrase recognition, which typically seek to align sentences in comparable corpora (Lapata & Barzilay, 2005), generate paraphrases using distributional similarity applied on paths of dependency trees or use bilingual parallel corpora (Lin & Pantel 2001). These methods identify paraphrases in large documents, or generate paraphrases from an input text, without necessarily providing a measure of their similarity. The more recently introduced textual entailment task (Lapata & Barzilay, 2005) is also related to some extent, however textual entailment identifies a directional inferential relationship between texts, which is different from textual similarity, and hence entailment systems are not overviewed here.

2.4.3 Relevance Feedback

Relevance feedback is a classical information retrieval (IR) technique where users relay their agreement with the system's evaluation of document relevance back to the system, which then uses this information to provide a revised list of search results (Vinay et al., 2005). In a system that involves user relevance feedback the user is given an opportunity to inspect the ranked list and indicate which documents are relevant to the user's query and which are not. This information is then used by the relevance feedback algorithm to induce a new ranking of documents. The new ranking, possibly including new documents, is displayed to the user and the process repeats. The feedback process comprises several phases. The display phase is the presentation of ten documents from the ranked list. The user feedback phase is a single action where the user nominates some subset of the displayed documents as being relevant to his or her information need. The document ranking phase applies one of the relevance feedback algorithms to create a new ranking of the document collection.

In relevance feedback, relevance information is gathered from documents retrieved in a ranked list generated from an initial request. This relevance information is used to modify the search query and perform a further retrieval pass. The two main factors in relevance feedback are the source from which expansion terms are determined and the method of ranking expansion terms. The quality of a query fed to an IR system has a direct impact on the success of the search outcome. In fact, one of the most important but frustrating tasks in IR is query formulation (French et al., 1997). Relevance feedback is a popular and widely accepted query reformulation strategy. The main idea consists of selecting important terms, or expressions, attached to the documents that have been identified as relevant by the user, and of enhancing the importance of these terms in a new query formulation. The expected effect is that the new query will be moved towards the relevant documents and away from the non-relevant ones.

Three standard relevance feedback algorithms are described below: The Rocchio algorithm, the Robertson/Sparck-Jones algorithm and the Bayesian feedback algorithm (Vinay et al., 2005).

The Rocchio Algorithm

The Rocchio relevance feedback scheme (Rocchio, 1971) is used in conjunction with the term frequency inverse-document-frequency (tf-idf) representation where documents and queries are represented as vectors of term weights and similarity is measured by the cosine distance between these vectors.

A document is a vector $d_i = (d_{i,1}, d_{i,2}, \dots, d_{i,T})$ where T is the number of words across the collection, excluding a predefined set of stopwords, and $d_{i,j} = t(i,j) \cdot s_j$. Here $t(i,j)$ corresponds to the number of occurrences of term j in document i and s_j is the inverse

document frequency of term j across the whole collection. A query $q = (q_1, q_2, \dots, q_T)$ is defined similarly, though their values are typically 0 or 1. Both documents and queries are normalised for length by setting

$$d' = \frac{d}{\|d\|} \text{ and } q' = \frac{q}{\|q\|} \text{ where } \|x\| = \sqrt{\sum_{j=1}^T x_j^2}$$

The similarity score between document d and query q is then given by the dot product of the normalized vectors, i.e., $\text{score}_{\text{rocchio}}(d_i, q) = d_i' \cdot q'$. The Rocchio algorithm takes a set R , of relevant documents and a set N , of non-relevant documents (as selected in the user feedback phase) and updates the query weights according to the following equation:

$$w_j' = \alpha w_j + \beta \frac{\sum_{i \in R} d_{i,j}}{n_R} + \gamma \frac{\sum_{i \in N} d_{i,j}}{n_N}$$

where n_R and n_N are the number of relevant and non-relevant documents respectively.

The parameters α , β , and γ control the relative effect of the original weights, the relevant documents, and the non-relevant documents. If there are no non-relevant documents, $\alpha = \beta = 1$ is used.

The Robertson/Sparck-Jones Algorithm

In the Robertson/Sparck Jones model of information retrieval (Robertson, 1995), the terms in a corpus are all assigned relevance weights which are updated for a particular query whenever relevant documents are identified. Initially the relevance weights are

given idf-based values. Documents are given ranking scores against a query based on the relevance weights of the query terms occurring in each document. The initial relevance weight for term j is given by

$$w_j = f(x) = \log\left(\frac{C}{n_j}\right)$$

where C is the total number of documents in the corpus and n_j is the number of documents containing term j.

A document d_i is assigned a score against query q as follows:

$$score_{r,s_j}(d_i, q) = \sum_{j \in Q} \frac{(K + 1) * t(i, j)}{K(1 - b) + \frac{b * |d_i|}{l} + t(i, j)}$$

where

$t(i, j)$ is the number of occurrences of term j in document d_i

K and b are parameters typically set to 2.0 and 0.75 respectively

$|d_i|$ is the length of document d_i

l is the average length of all documents in the corpus

Documents are then ranked in descending score order. If certain documents are flagged as relevant, the relevance weights are updated as follows:

$$w_j = \log\left(\left(\frac{(r_j + 0.5)}{(n_j - r_j) + 0.5}\right)\left(\frac{(C - n_j - n_r + r_j + 0.5)}{n_r - r_j + 0.5}\right)\right)$$

where

R is the number of relevant documents

r_j is the number of relevant documents containing term j

C and n_j are defined as before

In addition to updating the relevance weights, the relevant documents are used to select new (or additional) query terms according to the offer weights, o_j , where

$$o_j = r C w_j$$

Terms are ranked in decreasing order of offer weight, and the top terms are used as part of the subsequent query. How many such terms are to be chosen per iteration is another parameter of the system.

The Bayesian Algorithm

The Bayesian relevance feedback algorithm (Cox et al., 2000), first proposed for a Content-Based Image Retrieval System – PicHunter – is a recursive probabilistic formulation in which, at each iteration, k , the probability, P_k , of document, d_i , being the target document, d_T , is computed. This probability is conditioned on all current and past user actions and the history of displayed documents, which collectively is denoted by H_k . The concept of a current query, q , is not explicitly present in this formulation. Thus, at each iteration, the document rankings are given by

$$\text{score}_{\text{bayesian}}(d_i) = P_k(d_i = d_T | H_k) = P_{k-1}(d_i = d_T | H_{k-1}) * G(d_i, R)$$

where

P_{k-1} is the document's probability in the previous iteration

R is the set of documents marked relevant in this iteration

$G(d_i, R)$ is given by

$$G(d_i, R) = \prod_{j \in R} \frac{\exp\left(\frac{\text{sim}(d_i, d_j)}{\sigma}\right)}{\left(\sum_{((k \in D) \text{ and } (k \in R))} \exp\left(\frac{\text{sim}(d_i, d_k)}{\sigma}\right)\right) + \exp\left(\frac{\text{sim}(d_i, d_j)}{\sigma}\right)}$$

The term $\text{sim}(x,y)$ computes the similarity of document x with document y , which for textual documents can be taken as the cosine dot product of tf-idf vectors normalised for length. σ is a tuning noise parameter which is set according to the specific dataset.

2.4.4 Latent Semantic Indexing

Latent Semantic Indexing (LSI) is an information retrieval method that is based on the assumption that there is an underlying semantic structure in textual data, and that the relationship between terms and documents can be re-described in this semantic structure form (Dumais, 1995). The resulting semantic structure reflects the major associative patterns in the data while ignoring some of the smaller variations that may be due to different word usage of individual documents. This permits retrieval based on the ‘latent’ semantic content of the documents rather than just on keyword matches. Even with a clear idea of what factors are important for predicting interest, there is no guarantee that those factors can be identified easily. One of the simplest methods of determining whether information matches a certain user model is through keyword matching. If a user’s interests are clearly indicated by certain words, then information containing those words should be relevant. Nevertheless, keyword matching can still fail due to polysemy (more than one meaning for a single word) and synonymy (many ways of referring to the same concept). Research has shown that across people, the same word is seldom used by two people to describe an object (Furnas, et al 1987). This suggests that keyword matching in texts may fail. Thus any user model for

automatic information retrieval will be based on its ability to determine what part of the text actually allows the prediction of the user’s interest.

Below is an example to illustrate some of the problems with term-based information retrieval systems by means of a fictional matrix of terms by documents (Table 2-2). A fictional query “computer-based information look-up” is given here, which might have been passed against this database.

Table 2-2: A query string “computer-based information look-up”

	access	document	retrieval	information	theory	database	indexing	computer	RELATED	MATCH
Doc1	√	√	√			√	√		Yes	
Doc2				√*	√			√*		Yes
Doc3			√	√*				√*	Yes	Yes

A “Yes” in the column labelled RELATED indicates that the user would have judged the document relevant to the query (here documents 1 and 3 are relevant). Terms occurring in both the query and a document (computer and information) are indicated by an asterisk in the appropriate cell. A “Yes” in the MATCH column indicates that the document matches the query and would have been returned to the user by automatic system. Documents 1 and 2 illustrate common classes of problems with which the proposed method deals. Document 1 is a relevant document, which contains none of the words in the query. It would, therefore, not be returned by a straightforward term overlap retrieval scheme. Document 2 is a non-relevant document that does contain terms in the query, and therefore would be returned despite the fact that the query context makes it clear enough to a human observer that a different sense of at least one of the words is intended. Note that in this example

none of the meaningful conditioning terms in the query is found in the index. Thus intersecting them with the query terms would not have been a plausible strategy for omitting document 2.

Latent Semantic Indexing takes advantage of the implicit higher-order structure of the association of terms with documents to create a multi-dimensional semantic structure of information. Through the pattern of co-occurrences of words, LSI is able to infer the structure of relationships between documents and words (Zelikovitz & Hirsh, 2005). A single-value decomposition (SVD) of the term by document association matrix is computed to produce a reduced dimensionality matrix containing the best k orthogonal factors to approximate the original matrix as the model of “semantic” space for the collection (Dumais, 1995). This semantic space reflects the major associative patterns in the data while ignoring some of smaller variations that may be due to idiosyncrasies in the word usage of individual documents. In this way, LSI produces a representation of the underlying “latent” semantic structure of the information. Retrieving information in LSI overcomes some of the problems of keyword matching by retrieval based on the higher-level semantic structure rather than just the surface level word choice.

Latent Semantic Indexing is based on the assumption that there is an underlying semantic structure in textual data, and that the relationship between terms and documents can be re-described in this semantic structure form. Textual documents are represented as vectors in a vector space. Each position in a vector represents a term/word, with the value of a position i equal to 0 if the term does not appear in the document, and having a positive value otherwise. The positive values can be

represented as the log of the total frequency in that document weighted by the entropy of the term (log-entropy weighting scheme). As a result, the corpus can be looked at as a large term-by-document ($t \times d$) matrix X , with each position x_{ij} corresponding to the presence or absence of a term (row i) in a document (column j). This matrix is typically very sparse, as most documents contain only a small percentage of the total number of terms seen in the full collection of documents. Because of the way it represents terms and documents in a term-document space, it is considered a kind of vector-space information retrieval model.

Unfortunately, in this very large space, many documents that are related to each other semantically might not share any words and thus appear very distant. Occasionally documents that are not related to each other might share common words and thus appear to be closer. This is due to the nature of the text, where the same concept can be represented by many different words, and words can have ambiguous meanings. LSI reduces this large space to one that tends to capture the true relationship between documents.

Once the $t \times d$ matrix X has been created and properly weighted, a rank- k approximation ($k < \min(t, d)$) to X , X_k , is computed using an orthogonal decomposition known as the singular value decomposition (SVD) (Golub & Loan, 1989). The SVD of the matrix X is defined as the product of three matrices,

$$X = TSD^t,$$

The columns of T and D are the left and right singular vectors respectively, corresponding to the monotonically decreasing (in value) diagonal elements of S

document vectors, allowing clusters of terms and documents to be more easily identified. Within the reduced space, semantically related terms and documents presumably lie near each other since the SVD attempts to derive the underlying, semantic structure of the term-document space (Berry, et al., 1995).

In the LSI model, queries are formed into pseudo-documents that specify the location of the query in the reduced term-document space. Given q , a vector whose non-zero elements contain the weighted (using the same local and global weighting schemes applied to the document collection being searched) term-frequency counts of the terms that appear in the query, the pseudo-document, q , can be represented by

$$q = q^T T_k S_k^{-1}$$

Thus, the pseudo-document consists of the sum of the term vectors ($q^T T_k$) corresponding to the terms specified in the query scaled by the inverse of the singular values (S_k^{-1}). The singular values are used to individually weight each dimension of the term-document space (Berry, et al., 1995).

Once the query is projected into the term-document space, one of several similarity measures can be applied to compare the position of the pseudo-document to the positions of the terms or documents in the reduced term-document space. One popular similarity measure, the cosine similarity measure, is often used because one only needs to find the angle between the pseudo-document and the terms or documents in the reduced space. In this measure, the lengths of the documents can affect the distance between the pseudo-document and all the terms and documents in the space that have been computed. The terms or documents are ranked according to the results

of the similarity measure, and the highest-ranking terms or documents, or all the terms and documents exceeding some threshold value, are returned to the user (Dumais, 1995).

Relevance feedback is an effective method for iteratively improving a query without increasing the computational requirements to perform the query (Vinay et al., 2005). Relevance feedback uses the terms contained in relevant documents to supplement and enrich the user's initial keyword query, allowing greater retrieval performance (in terms of precision and recall). Since LSI explicitly represents both terms and documents in the same space, a relevance feedback query is constructed in essentially the same way as a regular query. Here two vectors are given:

q: terms in query;

d: a vector whose elements specify the documents in the query;

the pseudo-document representing the relevance feedback query is given by

$$Q = q^T T_k S_k^{-1} + d^T D_k$$

By matching the pseudo-document against each term or document vector and sorting the results, the highest-ranking terms or documents (or all those exceeding a threshold value) can be returned to the user.

2.4.5 Search Engine

This section describes how a Web search engine is typically put together. The approach of search engine helps the solution proposal in this research project. Figure 2-8 shows general search engine architecture. Search engines perform three basic tasks (Sonnenreich & Macinta, 1998): Firstly, search engines find and fetch web

pages using crawlers or spiders, and build lists of words and phrases found on each web page. Crawlers are small programs that browse the Web on the search engine's behalf, similar to how human users follow links to reach different pages (Arasu, et al., 2001). A crawler module is the most important core part for all engines because it provides the direction for its operation (shown on the left in Figure 2-8). Secondly, search engines keep an index (or database) of the words and phrases found on each web page the spiders are able to crawl. The part of the search engine that places the web pages into the database is called an indexer; and search engines then enable end users to search for keywords and keyword phrases found in their indices. Finally, search engines try to match the words typed in a search query with the web page that is most likely to have the information which end users are searching for. This part of the search engine is called the query processor.

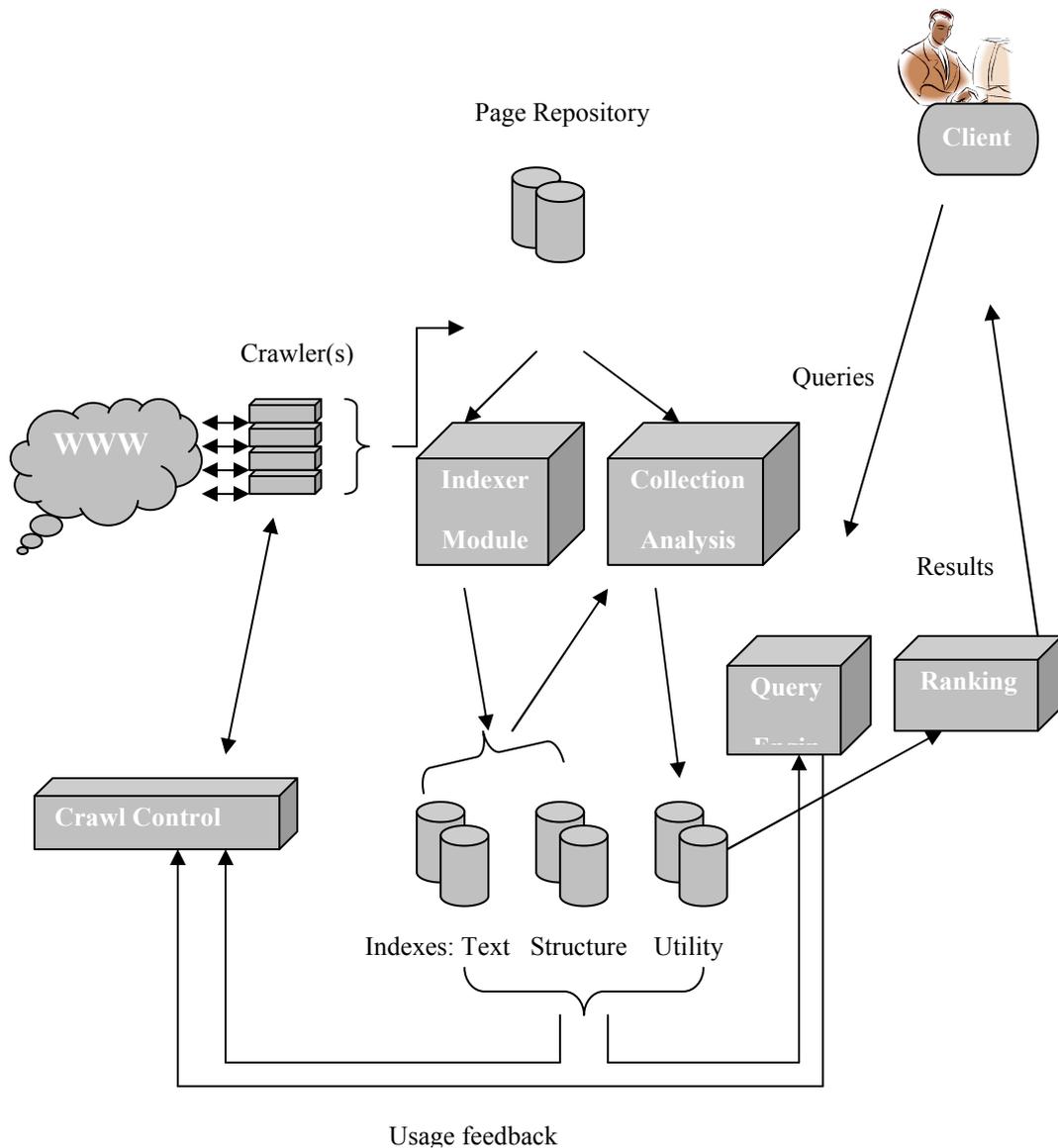


Figure 2-8: Search Engine Architecture in General (Arasu, et al., 2001)

Indices of search engines are continually updated by spiders crawling the web. The search engine index contains full-text indices of web pages. Thus, when people execute a query search on a search engine, they are actually searching this full-text index of retrieved web pages, not the web itself. The crawlers grab URLs appearing in the web pages and send this information to the crawler control module. The crawlers

also store the retrieved pages into a page repository. Then crawlers continue retrieving from the Web until storage is full.

The crawl control module may, for example, use a previous crawl's link graph (the structure index in Figure 2-8) to decide which links the crawlers should explore and which links they should ignore. Crawl control may also use feedback from usage patterns to guide the crawling process (collection between the query engine and the crawl control module in Figure 2-8). The indexer module extracts all the words from each page and records the URL where each word occurred. The collection analysis module is responsible for creating a variety of other indexes. The utility index is created by the collection analysis module. During a crawling and indexing run, search engines must store the pages they retrieve from the Web. The page repository holds this – possibly temporary – collection. Search engines sometimes maintain a cache of the pages they have retrieved beyond the time required to build the index. This cache allows them to serve out result pages very quickly, in addition to providing basic search facilities.

The query engine module is used for receiving and filling search requests from users. The engine relies heavily on the indexes, and sometimes on the page repository (Arasu, et al., 2001). Due to the Web's size and the fact that users typically only enter one or two keywords, result sets are usually very large. Hence the ranking module has the task of sorting the results so that results near the top are the most likely to be what the user is looking for. The query module is of special interest because traditional information retrieval techniques have run into selectivity problems when applied without modification to Web searching. Most traditional techniques rely on measuring

the similarity of query texts to texts in a collection's documents. The small queries over a vast collection will return a high proportion of irrelevant pages. There are some search algorithms that take advantage of the Web's interlinked nature. When deployed in conjunction with the traditional information retrieval techniques, these algorithms significantly improve retrieval precision in Web search scenarios.

Along with search engines, the next section introduces the technology that has been developed based on human interaction. The user model is used to predict human behaviours. Based on this, researchers have developed a different type of information system to meet various needs.

2.5 Recommender Systems

Recommenders are computer systems designed to help people find preferred items within a very large set of available alternatives such as movies, books or news (Popescul, et al., 2001). In order to provide effective, tailored advice (i.e. to predict whether the user will like/dislike a particular item) recommenders can make use of three different sources of information about their users (Carenini et al., 2003): (i) user demographics (e.g. user's age, gender), (ii) user preferences about features of the items (e.g. the movie director) and (iii) user ratings of experienced items (e.g. previously seen movies). All these sources of information can be effectively combined.

Recommender systems collect user feedback in the form of ratings for items in a particular domain and calculate similarities and differences from profiles of several

users in determining how to recommend an item. There are two prevalent approaches for constructing recommender systems (Sugiyama et al., 2004): collaborative filtering-based recommendation and content-based recommendation.

2.5.1 Collaborative Filtering-Based Recommendation

Collaborative filtering-based recommendation is the most successful recommendation technique to date. The term collaborative filtering was coined by Goldberg et al. (1992). Collaborative filtering means that people collaborate to help each other perform filtering by recording their reactions to documents they read. Based on this concept, Goldberg et al. (1992) developed a system called Tapestry that is one of the earliest implementations of collaborative filtering-based recommendation. This system is used to filter emails and it allows users to annotate messages. The collaborative filtering provided by complex queries in a special query language was designed for the task.

In addition, this system relied on clear opinions of people from a close-knit community, such as a group of office workers. However, recommender systems for large communities generally cannot depend on everyone knowing each other. Therefore, the framework in Tapestry is not appropriate to systems for large communities.

Rating-based automated collaborative filtering is quickly becoming a popular approach in reducing information overload by providing personalised recommendations for information, products or services. The k-nearest-neighbour collaborative filtering-based system is achieving widespread success on the Web. For

instance, the GroupLens research system (Konstan et al., 1997) which filters Usenet news, first introduced an automated collaborative filtering system using the k-nearest neighbour based algorithm. In this algorithm, a subset of appropriate k users is chosen based on their similarity to the active user, and a weighted aggregate of their rating is used to generate predictions for the active user. GroupLens then recommends Usenet news articles to these active users.

While Tapestry and GroupLens mentioned above rely on explicit ratings, some systems rely on implicit ratings. For example, Morita and Shinoda (1994) exploit “time-spent-reading” as a measure of implicit ratings. PHOAKS recommender system (People Helping One Another Know Stuff) (Terveen et al., 1997) also uses implicit ratings by examining Usenet news postings to find “endorsements” of Web sites. It then creates a listing of the top Web sites endorsed in each newsgroup. Some recommender systems also explore user preferences transparently without any extra effort from the users like the recommender systems relying on implicit ratings described above (Kelly et al., 2003). For example, Letizia (Lieberman, 1995 & Lieberman, 1997) and WebWatcher (Joachims et al., 1997) infer user preferences by observing users’ browsing behaviour. However, the main shortcomings in Letizia and WebWatcher are that they maintain persistent and slowly-changing user models and overlook the fact that different browsing sessions by the same user or even a single session may involve different user interests and goals.

In addition, at the E-commerce sites such as Amazon.com, CDnow.com and MovieFinder.com, automated collaborative filtering systems have been used with considerable success. Moreover, in the field of audio, Ringo (Shardanand & Maes,

1995) uses collaborative filtering techniques to provide users with recommendations for music albums and artists.

2.5.2 Content-Based Recommendation

A content-based approach provides recommendations by comparing representations of content contained in an item with representations of content that the user is interested in (Carenini et al., 2003). In this approach, a model of user ratings is first developed. Algorithms in this category use probabilities and envision the collaborative filtering process by computing the expected value of a user prediction given the user's ratings on other items. The model building process is performed by three different machine learning algorithms (Sugiyama et al., 2004): (1) Bayesian network (Breese et al., 1998), (2) clustering (Basu et al., 1998; Breese et al., 1998), and (3) rule-based models (Sarwar et al., 2000).

The systems described above only provide recommendations based on collaborative filtering. However, some systems provide better recommendations by combining collaborative filtering with content information. For instance, Fab (Balabanovic & Shoham, 1997) uses relevance feedback to simultaneously construct a personal filter along with a communal "topic" filter. Web pages are initially ranked by the topic filter and then sent to the user's personal filters. The user then provides relevance feedback for that Web page, and this feedback is used to modify both the personal filter and the originating topic filter. Basu et al. (1998) also integrate content and collaboration in a framework where they treat recommendation as a classification task. Melville et al. (2002) overcome drawbacks of collaborative filtering systems in their recommender system by exploiting content information of items already rated.

Schafer et al. (2002) introduce a new class of recommender system that provides users with personalised control over the generation of a single recommendation list formed from a combination of rich data using multiple information resources and recommendation techniques.

2.6 Conclusion

This chapter has reviewed the previous research background on online learning communities and discussion forum systems. In addition, it has introduced technological solutions which are used in this research, such as information retrieval, text similarity and the recommender systems. Based on the research background, the concerns of existing discussion forum systems used in learning environments and requirements of proposed system will be identified in the next chapter.

CHAPTER 3

REQUIREMENTS ANALYSIS

3.1 Introduction

The previous chapter presented the research background and described current problems with traditional discussion forums in learning environments. This chapter identifies the challenges of existing discussion forum systems used in learning environments. It then analyses and gathers the proposed discussion forum system requirements in order to solve current problems by overcoming communication obstacles in discussion forums. A requirement summary is finally presented according to requirements analysis and previous research contributions.

3.2 Requirements Gathering

3.2.1 Overview

This research focuses on discussion forum systems used in web-based distance learning environments and classrooms as a tool that can be enhanced to support the construction of knowledge and improve learning. The aim is to identify how to improve the effectiveness of communication in discussion forum systems. This chapter begins with a description of challenges in using discussion forum systems in learning environments and presents the requirements analysis for the proposed solution based on previous research contributions introduced in Chapter 2.

3.2.2 Challenges in Discussion Forum

Asynchronous communication tools, such as discussion forum and bulletin boards, are frequently used by instructors in web-based learning environments as a supplement to traditional classroom instruction. These tools allow instructors to communicate with students by posting questions, requesting information, generating a debate, surveying students, or conducting a discussion. Students can respond to the postings by the instructor or other students after reading and reflecting on the postings. The expectation is that responses to a posting will not be immediate since the nature of asynchronous communication suggests an undefined time lag between postings and responses.

Even though discussion forums are now used by majority of on-line courses and learning management systems, there are some challenges in incorporating them into learning environments. Five areas of concern are discussed below regarding the current discussion forum tools based on the prior research contributions.

First Area of Concern - discussion forums not being a catalyst for new ideas: The propagation of communication tools such as discussion forums has been recognised as a promoter of “active” learning. These discussion forums are part of almost all web course management systems such as WebCT, Blackboard, and Moodle. Previous research has suggested: that including a discussion on a topic and/or posting questions makes the course “active”; that students actively engage in a high level discussion on any given topic and in this process, knowledge building is a common occurrence. Sheard et al. (2003) on the other hand suggest that discussion forums are neither active nor do they consistently encourage thoughtful synthesis and postings. Students

frequently paraphrase each other instead of contributing something new to the discussion (Cunningham-Atkin et al., 2004; Weller et al., 2005).

Challenge #1: Scaffolding the learners into contributing new ideas and managing their learning from the discussion forums is important.

Second Area of Concern – discussion forums having too much ‘junk’: Previous research has described that in traditional discussion forums, instructors typically end up managing too many unnecessary topics and duplicate postings. Sometimes they have to answer similar queries more than once because students who asked at later stage would be expecting an answer (Yang, 2005).

Challenge #2: Automatic responses by the system for duplicate queries can reduce the number of responses to be made by the instructor or other students.

Third Area of Concern – information rarely being reused in discussion forums: Discussion forums in learning environments contain a wealth of knowledge that is frequently stored in databases and rarely used beyond the initial posting and response situation. Most colleges and universities have stored millions of posting by students and responses from peers and instructors and have rarely gone back and reused the information in those postings. A better way of organising and distributing the information may provide students with tools to access learning in discussion forums more effectively.

Challenge #3: Reusing data to support improved learning is a worthy goal that can be extended to assessment of learning.

Fourth Area of Concern – discussion forums have overwhelming amounts of information: The sheer number of postings can actually cause students and instructors to be overwhelmed (Reeves, 2003), forcing them to skim through vast number of postings trying to construct shallow disjointed knowledge from the discussion (Weller et al., 2005; Wilhelm et al., 2003). A proper learning experience from current discussion forums requires students to read multiple postings and analyse whether their own query has already been answered. One likely result is that they would post another message or query in the forum, further increasing the duplication. It is quite possible that a similar query or message had already been submitted by someone else. In such a scenario, it is very difficult for students to capture and retrieve relevant knowledge and useful information from the thousands of messages in the discussion forums. Discussion forums can be disjointed due to high volumes of irrelevant postings and duplicated queries.

Challenge #4: Creating a clear structured discussion and automatically identifying similar postings in the forums to help users contribute more meaningful content (Wijekumar & Spielvogel, 2006).

Fifth Area of Concern – discussion forum having delayed responses: The undefined delay in response time in asynchronous discussion forums can cause breakdowns in communication. The resulting discussion can be disjointed and students may lose the motivation for the discussion while waiting for a response. When misinformation is

(often inadvertently) circulated, the delay before corrections are posted can exacerbate the confusion resulting from misinformation (Anderson & Jackson, 2000).

Challenge #5: Immediate feedback can eliminate misconceptions, hold the interest of students and prompt them to think about their postings (Wijekumar & Spielvogel, 2006).

3.2.3 Requirement Analysis

New forum system requirements developed in an attempt to address the above challenges are identified in this section. Table 3-1 presents the challenges identified earlier and denotes how traditional and the new discussion forum systems handle them.

As shown in Table 3-1, in traditional forums, effective communication relies extensively on the instructor's ability to: foster a good learning environment, draw out students' thinking, limit the paraphrasing of student comments, scaffold the learners to concentrate on knowledge building and think deeply about their postings, provide immediate feedback, prompt the learners, reduce the time lag for feedback, and encourage all students to actively participate with less duplicated postings. However this research is proposing a new type of forums aiming to solve the current challenges of the discussion forums. This new type of forums should overcome the problems experienced in the traditional discussion forums in order to help students effectively communicate with instructors and each other. Such forums should be able to identify that the submitted posting is a duplicate or similar to an existing message in the forum and then inform the user about the existing information. Moreover when users post

messages in the forum, it should be able to provide immediate useful and relevant response. All existing information stored in the discussion forum should be easily reused. The new type of forums should be able to reduce instructors' workload of managing the forums as well as encouraging students to actively think and participate.

Table 3-1: Addressing the five challenges in discussion forums

Challenge	Traditional Discussion Forums	New type of Discussion Forums
1. Students frequently paraphrase each other	Instructor intervention needed to manage the postings	They should be able to automatically check postings for similarity and inform the students about existing information.
2. Instructors have to spend too much time on managing the forums	Instructor intervention required in deleting or responding to duplicate postings.	They should provide relevant postings to students during new posting submission in order to reduce the number of similar postings to which the instructor needs to respond.
3. Stored millions of postings but never used	No solution. Some instructors do review postings for research purpose but there have been no examples of assessment techniques for the postings.	They should provide all relevant information from the different course forums across different institutions.
4. Large numbers of postings that add little or no new information.	The instructor has limited ability to move threads and postings to organise the discussion. However, this is a	They should inform students about other relevant postings, and ask students to focus on new information in their postings in order to effectively

	complex task complicated by the large number of postings.	contribute to the discussion.
Delayed response time.	The instructor has to frequently read and respond to all postings.	They should provide immediate feedback on the quality of postings and refer users to any useful responses already available in the discussion forum.

3.2.4 Requirements Summary

Requirements for the proposed new type of discussion forum systems have been gathered and analysed in the last section. This section summarises the requirement details for the development phase of the new type of forums:

- Based on the assumption that being presented with multiple perspectives and sharing knowledge across larger groups is beneficial for students, a forum should be able to be shared by different courses. This should be possible both inside one institution and across different institutions (using different learning management systems and discussion forum systems). This is in tune with other efforts to define learning objectives (Friesen et al., 2004) for sharing among institutions and building repositories of reusable learning content (Baraniuk, 2006) for easing access to knowledge.
- Based on the first requirement that a forum should be shared by various courses in different institutions and since different institutions use various learning management systems and discussion forum, a key requirement for the proposed system is to be able to be integrated properly with the existing

discussion forum systems that are used in the different courses and different institutions.

- Queries submitted by users should be intercepted by the system, the existing knowledge base should be searched, and a list of possible answers and relevant information should be displayed to the users, before the queries are added to the system. This would allow the users to quickly find relevant information to their queries rather than having to browse or search through the entire forum to see if the information already exists. This feature should help to minimise the creation of duplicate messages in the system.
- Students should be able to give feedback to a searched message to indicate whether the displayed message is related to their queries. The system should then use this information to more accurately link messages with certain types of queries so that more accurate search results could be extracted for future queries.
- Teachers should be able to reduce the time-consuming burden of managing the discussion forum and answering duplicated questions when students submit duplicated information in the forums. Hence the system should be able to find all relevant information in the pre-existing postings across all forums.

A system implementing the above discussed features would be a substantial improvement to the forum systems currently used in academia. The next chapter will introduce existing discussion forum systems developed by other researchers to meet the above requirements but with major limitations in system and user performance. Based on comparative analysis of the existing solutions, this research aims to create a new type of discussion forum systems as a better solution to current problems.

CHAPTER 4

SOLUTION DESIGN

4.1 Introduction

The previous chapter proposed requirements for a new type of discussion forum systems to solve a host of current problems with such forums. This chapter first introduces existing innovative solutions proposed elsewhere, such as the eQuake system and intelligent discussion-bot system. This is followed by an analysis of strengths and limitations of those solutions. According to the proposed new system requirements and summary of existing solutions, this research investigates a new type of forum systems as a complement to existing solutions. The chapter concludes with the introduction to the high-level design of the new type of systems.

4.2 Existing Solutions for Traditional Discussion Forums

4.2.1 Overview

This research focuses on the use of discussion forum systems in learning settings such as web-based distance learning environments and classrooms. The core research objective is to improve the effectiveness of communication in discussion forum systems. The last two chapters introduced and summarised limitations of traditional discussion forums. This chapter begins with a description of typical and recent existing solutions such as eQuake and intelligent discussion-bot, which have proven to be able to overcome current limitations of traditional discussion forums used in

academia. A solution is then proposed on the basis of the identified requirements for solving major problems inherent in existing traditional discussion forums (see Chapter 3).

4.2.2 EQuake

The electronic Question and Answer Knowledge Environment (eQuake) system (Smestad, 2006) was designed and implemented for tertiary and postgraduate students within New Zealand. It includes a new type of discussion forum enhanced with software agents that are able to be used for various courses in different institutions. Figure 4-1 depicts one example of the distributed architecture of the eQuake system which was intended to support interaction between students in different courses and institutions.

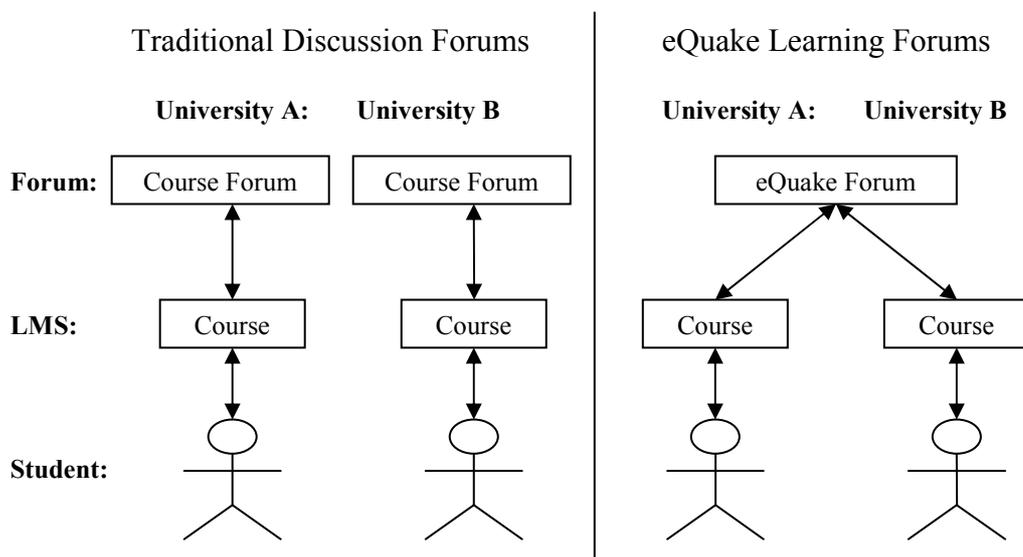


Figure 4-1: Use of Traditional Forums vs. eQuake Forums (Smestad, 2006)

The components of eQuake are shown in Figure 4-2. The system consists of three layers: Presentation Layer, Application Layer and Data Layer. Boxes signify components, cylinders represent databases and lines show interactions between

components. Lines marked with WS indicate Web Services being used for communication between components. Grey parts represent the existing Learning Management System (LMS) into which the particular instance of eQuake is integrated.

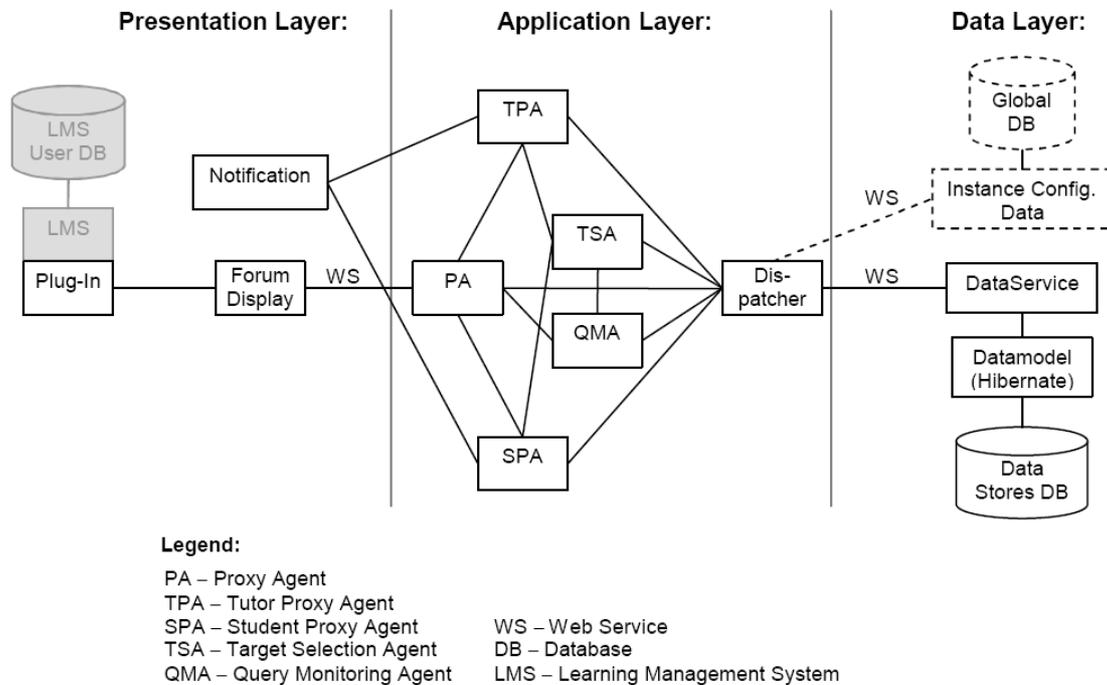


Figure 4-2: eQuake System Architecture (Smestad, 2006)

As an innovative solution attempting to solve the problems of traditional discussion forums, the main functions in eQuake included: automatic identification of previously answered questions, linking together several institutions' forums, and notifying users of answers to questions they are interested in. In the automatic answer search process, eQuake's simple Boolean keyword search performs the information retrieval task. As reported by Smestad (2006), the search approach in eQuake requires improvements most likely from further research in the advanced information retrieval methodology field. Figure 4-3 is a screenshot of the discussion area of the eQuake discussion forum.

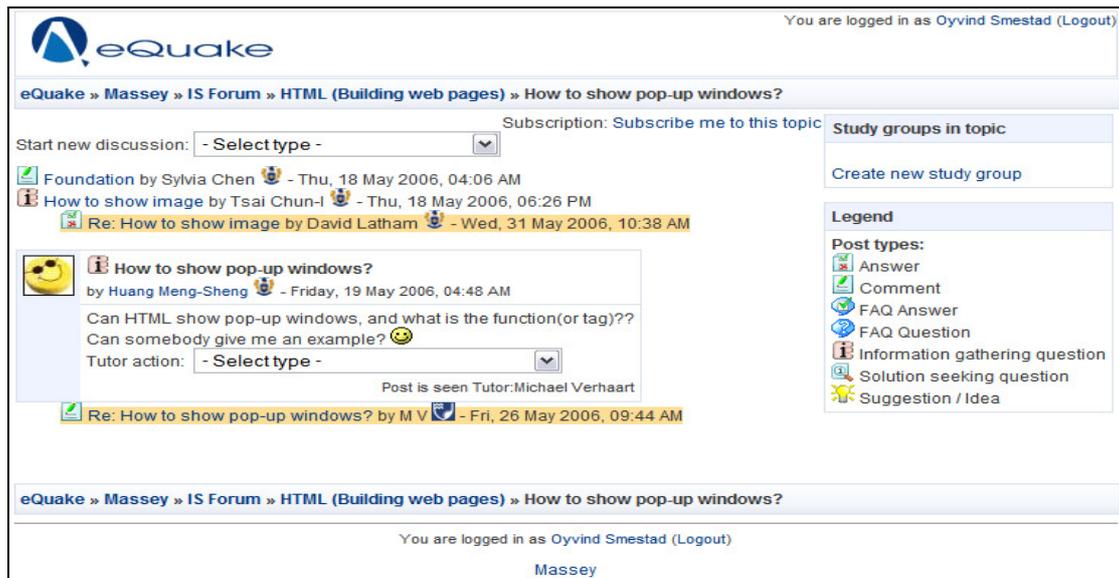


Figure 4-3: Screenshot of eQuake Web Forum (Smestad, 2006)

4.2.3 Intelligent Discussion-Bot (ID-B)

Feng et al. (2006) describe an intelligent discussion-bot system that aims to provide answers to students' discussion board questions in an unobtrusive and human-like manner. It uses an advanced search function that enhances communication between users in order to solve major problems of traditional discussion forums. This discussion-bot adopts information retrieval and natural language processing techniques to identify the questioner's preferences, mines appropriate answers from an annotated corpus of 1236 archived threaded discussion and 279 course documents and chooses an appropriate response. Feng et al. (2006) report a new modelling approach designed to analyse archived threaded discussions and facilitate answer extraction by the discussion-bot.

Figure 4-4 depicts the architecture of the discussion-bot system which was implemented within the ISI discussion board. When a query was posted to the discussion board, the discussion-bot system first extracted certain information from

the post, e.g., the words and word frequencies. Following that, the system tried to match the student's interest with archived data, including course documents and past discussions. The ensuing list of document tiles or messages related to the question was ranked based on predefined metrics. The answer extraction module selected the leading candidate from the list based on whether it was a segment of a document or a post. A variety of strategies were then applied to generate the answer which was automatically presented on the discussion board.

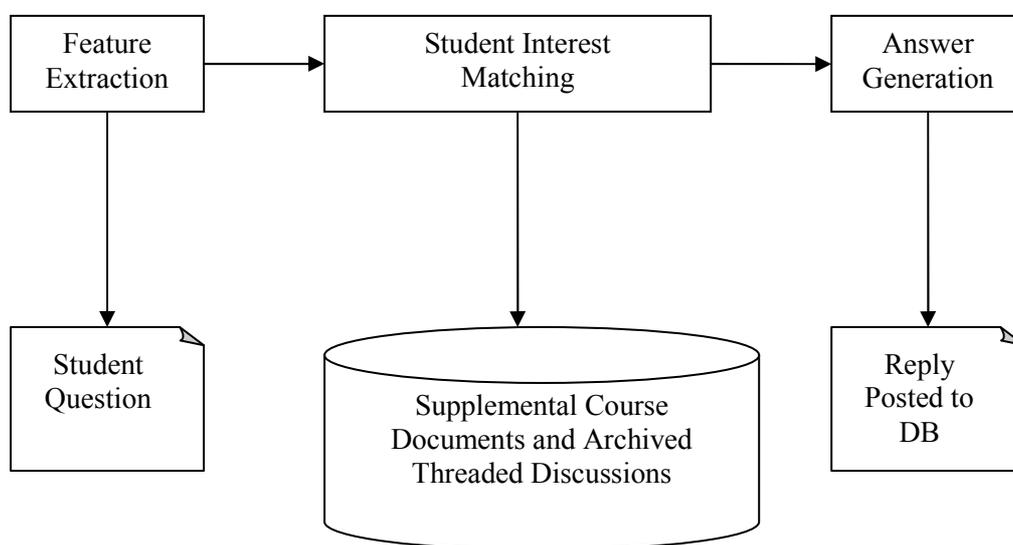


Figure 4-4: Discussion-Bot Architecture (Feng et al., 2006)

The discussion-bot did not apply the question-processing module of traditional (natural language) question answering systems to determine an answer type and extract query terms for the search engines. This was because it was difficult to recognise an extract question and thus identify a single answer type with submitted query terms. Instead, a set of semantically-related passages that match with a student's interest were retrieved using the TF*IDF technique which computes the cosine similarity between a question post and archived data.

To understand the approach adopted in the discussion-bot, let us assume that it had a total of n passages in the corpus, a student query q , and corpus passages p_1, p_2, \dots, p_n . Let us also assume that there were a total of m different unique words, w_1, w_2, \dots, w_m , found in all documents and posts. Let the number of occurrences of word w_j in passage p_i be tf_{ij} and the number of passages in which word w_j is found be c_j . The student's post and each passage in the corpus could be represented with a vector in the following format:

$$q = \langle w_{q1}, w_{q2}, \dots, w_{qm} \rangle$$

$$p_i = \langle w_{p_i1}, w_{p_i2}, \dots, w_{p_im} \rangle$$

where m is the total number of words in this domain, and $w_{ij} = 0$ if a word is missing in that passage. Each element then, after normalising the vectors, can be computed by

$$w_{ij} = \frac{tf_{ij} \log(n/c_j)}{\sqrt{\sum_{j=1}^m (tf_{ij})^2 [\log(n/c_j)]^2}}$$

When question q is posted, could the discussion-bot can retrieve a list of semantically-related passages (posts or document tiles) using the cosine similarities between the query post and the passages where:

$$\cos_sim(q, p_i) = \sum_{j=1}^m w_q * w_{p_i j}$$

According to the evaluation of this research, although the matching, modelling, extraction, and evaluation processes show promise and merit iteration, the understanding of complex questions remains very much an unsolved problem. Therefore the evaluation identifies the need to improve system performance and search accuracy (Feng et al., 2006).

4.2.4 Limitations of Current Solutions

The last two sections introduced two recent solutions, eQuake and an intelligent discussion-bot, which were designed and developed to solve major problems with traditional discussion forum systems. Both eQuake and discussion-bot aimed to improve communication in and usability of discussion forums in learning environments. Even though both solutions improve forum usability and meet some of the proposed system requirements for solving problems of traditional forums, they have obvious limitations that cannot be ignored.

Table 4-1 compares characteristics of traditional discussion forums with the above mentioned systems. The table shows that the eQuake forum has impressive system architecture that enables knowledge to be reused in different institutions, but it could be much enhanced by replacing the Boolean keyword search function with a more advanced information retrieval process. On the other hand, the intelligent discussion-bot uses advanced natural language processing methodology for information retrieval. However, the performance of its search function could be improved by information retrieval measure modification and other popular methods such as user feedback. Also, the limited architecture of the intelligent discussion-bot forum does not support the reuse of knowledge from different institutions.

Table 4-1: Characteristics of a Traditional Forum, eQuake & ID-B

Characteristics of discussion forums	Traditional Discussion Forum	eQuake Forum	Intelligent Discussion-Bot Forum
General functions (view, post	Yes	Yes	Yes

and reply)			
Reuse of knowledge from different institutions	No	Yes	No
Automatic search	No	Yes	Yes
Immediate response	No	Yes	Yes
Advanced natural language processing approach	No	No	Yes
Search function limited by search algorithm	—	Yes	Yes
Enables rating of search results by users	No	No	No

4.3 Proposed Solution

4.3.1 Introduction

Recently proposed solutions to the problems of traditional discussion forums have several limitations, as discussed in previous section. This section describes the design of a new type of discussion forums to overcome those limitations, enhance system usability and improve communication between forum users.

4.3.2 Requirements Recap

In Chapter 3, requirements of the proposed system were identified for solving the problems of traditional discussion forums. Before proposing a solution, those requirements are recapped below as system design criteria:

- The forum should include all general functions of a forum such as viewing messages, submitting a posting and replying in a thread.

- The forum should enable sharing and reuse of knowledge for various courses in different institutions.
- The forum should provide immediate responses for users' messages.
- The forum should automatically search and provide relevant information when users try to post a message.
- Users should be able to give feedback on search results.

All these system requirements are identified as part of the design criteria to solve the problems of traditional discussion forums. A new type of forum system is designed and developed in this research as an innovative solution based on these requirements. This newly designed system exploits the strengths of the current solutions while seeking to overcome their limitations. The new system architecture is presented below.

4.3.3 Solution Overview

In response to the system requirements list above, the proposed system's main functions are:

- that it should have general discussion forum functions (view, post, and reply);
- that knowledge in this forum can be shared between courses and institutions;
- that it should have synchronised function to respond to users' queries;
- that it should have a search function for existing relevant information within the forum.
- that users can rate the information retrieval results to enhance the forum's accuracy.

According to the main design functions, the system should combine the strengths of current solutions such as eQuake's system architecture and the information retrieval

approach of intelligent discussion-bot systems. The next section introduces the high-level design for the new system architecture and each main component of it.

4.3.4 High-Level System Architecture and Components Design

The architecture of the proposed system aims to support knowledge reuse and the interactions between users in separate courses and institutions. Different institutions use different learning management systems (e.g., Moodle, WebCT and BlackBoard) and discussion forums systems (e.g., Phpbb, Moodle forum and Jforum) installed and running on different platforms. Therefore the proposed system needs to have a platform independent interface that can be plugged into different existing systems in order to handle information exchange between different institutions. Figure 4-5 shows an example of how the proposed system can be utilised by different institutions.

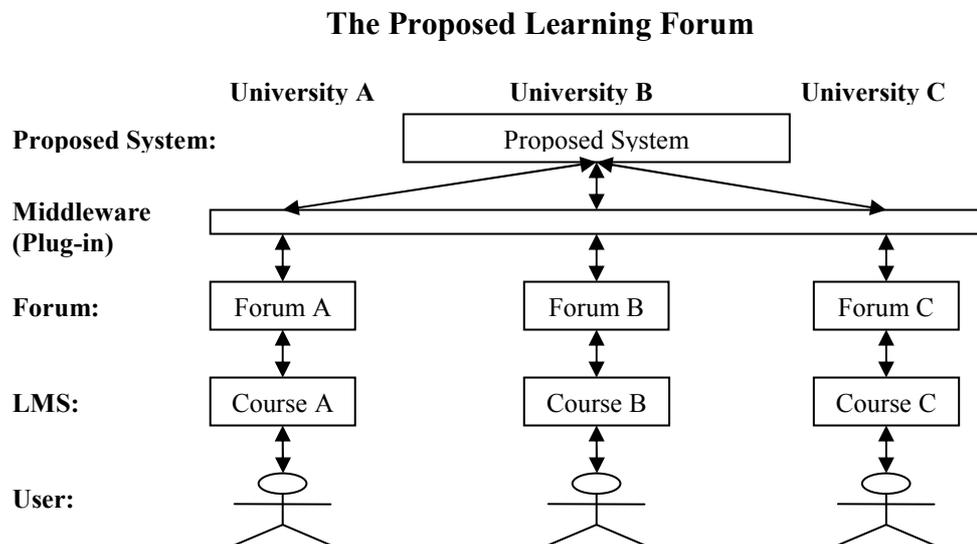


Figure 4-5: Use of Proposed System for Different Institutions

Various components of the proposed system are shown in Figure 4-6. The system consists of three layers: Presentation Layer, Application Layer and Data Layer. In

Figure 4-6, the proposed system is integrated with an existing discussion forum via a plug-in component. The lines marked with *WS* are where Web Services are used to communicate between components. In the Application Layer, a core component, the Dialogue Mechanism, is designed to receive data from existing forums, perform information retrieval, handle user interaction and return data to existing forums. With the exception of the forum display component, all system components are platform independent. The forum display component is developed with similar look and feel to the existing forum. The following sections describe each component.

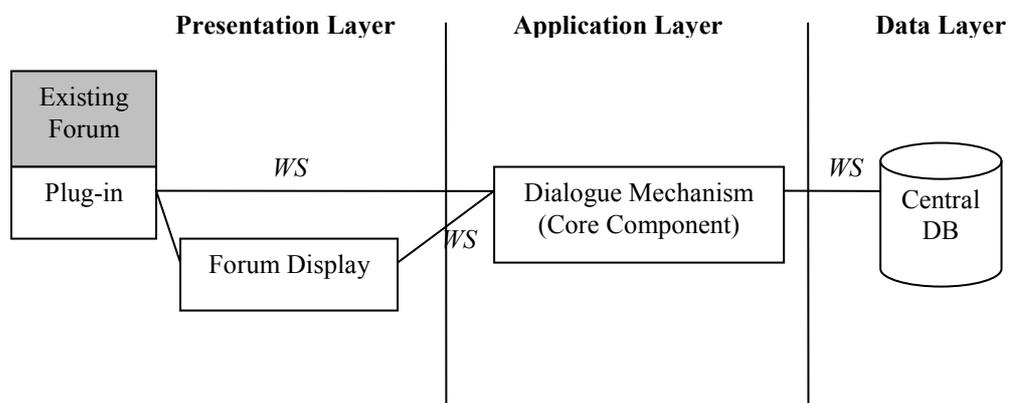


Figure 4-6: System Layer Diagram

Plug-in Component

A plug-in component is developed to integrate the new forum system with existing discussion forum systems. The main task of the plug-in component is to receive the data from the existing forums, encapsulate it and then send it to the Dialogue Mechanism (via Web Services as the communication method). Web Services is used so that, whatever system development languages and server platforms are used for existing forums, the new system can be easily integrated. After data is returned from the Dialogue Mechanism component, the plug-in component will send the data back

to existing forums and display it to users. How the plug-in component works with an existing forum is shown in Figure 4-7.

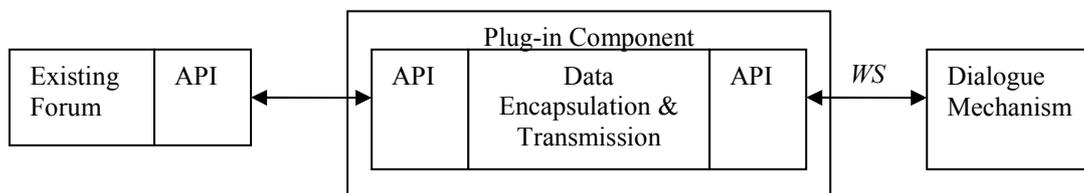


Figure 4-7: Plug-in Component

Dialogue Mechanism Component

The Dialogue Mechanism is a core component which performs text semantic analysis and computing tasks in the new system. As shown in Figure 4-8, the main tasks of the Dialogue Mechanism are to semantically analyse the input text and to search existing relevant information and knowledge in the database when a new message is submitted. Figure 4-8 shows the data flow between the student and the system when a message is submitted. When the Dialogue Mechanism receives a message it first calculates the semantic similarity between it and existing messages stored in the database and then searches for and returns the most relevant messages to users. It efficiently extracts keywords from all new messages and stores them with the message when the messages are added to the system. It also handles user interaction tasks. When a user gives feedback to the system, it again performs the data computation tasks. The design details and methodologies adopted for the Dialogue Mechanism component are introduced in Chapter 5.

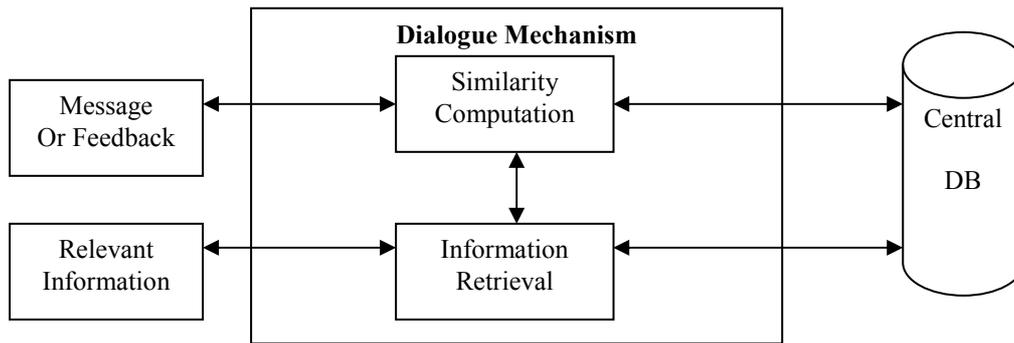


Figure 4-8: Dialogue Mechanism Component

Forum Display Component

The Forum Display component provides the user interface of the system. In other words, it generates the HTML that is sent to a user's web browser. Different discussion forum systems have different methods for their front-end development. For instance, the forum display component in eQuake is written in PHP. However, in Jforum, the display component was developed in Java. Therefore the new system's Forum Display component needs to be made as identical as possible to the existing discussion forum front-ends so that information can be easily exchanged with users.

Central DB Component

In the Data Layer of the proposed system, the central database component is used to store information from different institutions. However each integrated forum also has its own local databases to store data. The central database stores information from remote databases needed to perform the text semantic similarity calculation and information retrieval tasks. Whenever information is stored in a local database, a request is sent to the proposed system and then the required information is stored in the central database.

4.4 Summary

This chapter introduced two recent solutions which aim to solve major problems in traditional discussion forum systems. Both solutions have been implemented to meet the requirements summarised in Chapter 3. They improved the usability of the discussion forums and solved some traditional problems. However some limitations of those solutions were identified. Section 4.3 proposed outline of a novel system to overcome these limitations and enhance the usability of forum systems to improve the effectiveness of communication between students and instructors. In this proposed system, the core component is the Dialogue Mechanism that performs the text semantic similarity calculation and information retrieval tasks. This core component is designed in detail in the next chapter.

CHAPTER 5

DIALOGUE MECHANISM DESIGN

5.1 Introduction

The previous chapter proposed a new type of system to meet the identified requirements and overcome the limitations of existing systems in terms of solving major problems in traditional discussion forums. The high-level design of the new system architecture was then proposed. In the new system architecture, a core component, namely Dialogue Mechanism, was depicted that performed the natural language processing tasks in the new forum systems. This chapter first presents and summarizes some of the related approaches of natural language processing for text similarity calculation, information retrieval and user feedback that have been utilized in other research works. Based on the analysis of those approaches, an innovative approach is then proposed to design and develop the Dialogue Mechanism component. This proposed approach aims at improving the communication between the user and the system in the discussion forum systems.

5.2 Related Approaches

To improve the effectiveness of communication in the discussion forums, the two-way communication method between the user and system needs to be taken into account because the traditional discussion forums do not have the ability to

understand the meaning of users' messages. This two-way communication method is able to improve the understanding between the users and the system as follows:

- The system should be able to understand the user's text input more accurately in order to perform relevant information retrieval tasks more effectively for the users.
- The users should be able to rate the accuracy of system's understanding of the meaning of their text input in order for the system to improve the performance of relevant information retrieval tasks.

Besides the information retrieval approach used in the intelligent discussion-bot system (Feng et al., 2006) that was introduced in the previous chapter, a number of other innovative approaches are available in the literature that have been adopted in various system development processes in an attempt to improve the effective communication between the user and the system, particularly in the areas of information retrieval and user feedback. After introducing various related solutions in the chapter 4, this chapter focuses on the framework for designing the Dialogue mechanism component of the proposed system. The approaches used in the framework are as follows:

- Syntactically Enhanced Latent Semantic Analysis (LSA);
- Assessment Model to Analyzing Student's Performance;
- Ontology-based Information Retrieval System; and
- Dialogue Manager.

Following sections briefly describe these approaches.

5.2.1 Syntactically Enhanced Latent Semantic Analysis (LSA)

In 2003, Kanejiya et al. (2003) designed a new approach called Syntactically Enhanced Latent Semantic Analysis (SELSA) to automatically evaluate students' answers. This approach generalized LSA by considering a word along with its syntactic neighbourhood given by the part-of-speech (POS) tag of its preceding word, as a unit of knowledge representation.

This approach consisted of first identifying a sufficiently large corpus representing the domain of tutoring. Then a POS tagger was used to convert it to a POS tagged corpus. The next step was to construct a matrix whose rows correspond to word-prevtag pairs and columns correspond to documents in the corpus. Again, a document could be a sentence, a paragraph or a larger unit of text. If the vocabulary size was I , POS tag vocabulary size was J and number of documents in corpus was K , then the matrix would be $IJ \times K$. Let $c_{i,j,k}$ denote the frequency of word w_i with prevtag p_j i.e., $(i-1)J + j$ th row of the matrix. Bellegarda (2000) found in LSA the entropy $\varepsilon_{i,j}$ of each word-prevtag pair and scale the corresponding row of the matrix by $\varepsilon_{i,j}$. The document length normalization to each column of the matrix was also applied by dividing the entries of k th document by n_k , the number of words in document d_k . Let $t_{i,j}$ be the frequency of i,j th word-prevtag pair in the whole corpus, i.e., $t_{i,j} = \sum_{k=1}^K c_{i,j,k}$. Then $\varepsilon_{i,j}$ and the matrix element were given as:

$$\varepsilon_{i,j} = -\frac{1}{\log K} \sum_{k=1}^K \frac{c_{i,j,k}}{t_{i,j}} \log \frac{c_{i,j,k}}{t_{i,j}}$$

$$\chi_{i,j,k} = (1 - \varepsilon_{i,j}) \frac{c_{i,j,k}}{n_k}$$

Once the matrix X was obtained, performing its singular value decomposition (SVD) and approximate it by keeping the largest R singular values and setting the rest to zero.

Thus,

$$X \approx \hat{X} = USV^T$$

where, $U(IJ \times R)$ and $V(K \times R)$ were orthonormal matrices and $S(R \times R)$ was a diagonal matrix. It was this dimensionality reduction step through SVD that captures major structural associations between words-prevtags and documents (albeit, in different bases). This R -dimensional space could be called either syntactically enhanced latent semantic space or latent syntactic-semantic space.

Then projecting any new document as a R dimensional vector \vec{d}_L in this space, let d be $IJ \times 1$ the vector representing this document whose elements d_{i_j} were frequency counts i.e. number of times word w_i occurs with prevtag p_i , weighted by its corresponding entropy measure $(1 - \varepsilon_{i_j})$. It could be thought of as an additional column in the matrix X , and therefore could be thought of as having its corresponding vector v in the matrix V . then, $d = USV^T$ and

$$\vec{d}_L = Sv^T = U^T d$$

which is a $R \times 1$ dimensional vector representation of the document in the latent space.

A syntactic-semantic similarity measure could also be defined between any two text documents as the cosine of the angle between their projection vectors in the latent syntactic-semantic space. Although with this measure the problems that LSA has been applied on namely natural language understanding, cognitive modelling, statistical

language modelling, can be addressed, this presented approach has limited improvement over LSA.

5.2.2 Assessment Model to Analyzing Student's Performance

Wu and Chen (2004) proposed another approach which was used to assess the performance of students in the discussion forums. This is an approach to enhance the ability of the discussion forum systems in understanding the message semantic meaning of students' input. The core contribution was the algorithm to measure the quality of student's online messages. Aiming at distance learning improvements, Wu and Chen (2004) proposed this approach which applied natural language processing techniques to analyze the course messages to assess students' performance in the discussion forums. They classified the evaluation criteria into three categories:

- 1) quality assessment;
- 2) quantity assessment; and
- 3) participation assessment.

They considered the quality of a student's work to be a representation of the student's knowledge learnt directly or indirectly from the course, the quantity to be the indication of the amount of work the student had done in the class, and participation frequency to be a reflection of the activeness of the student in the class. Three measurements - Keyword Density (KD), Message Length (ML), and Message Count (MD), were accordingly derived from the online messages. The Performance Indicator (PI) was then computed from the above three measurements. The experimental results from their work suggested that the computer grader highly agreed with the human graders (Wu and Chen, 2004).

Quality Assessment

This approach assumed that quality of learning was revealed by the quality of discussion messages in some ways. The number of key concepts in a student's message indicated how much the student understood about the course topic and content. Therefore, there was an assumption made that the key concepts covered in the class message set would be those that a student could learn. By comparing the concepts from a student with those from the class message set, the instructors were able to estimate how well the student could learn the existing key concepts.

There were two stages before counting the measurement – Keyword Density. Firstly a keyword was defined as a simple, non-recursive noun phrase. This approach used a noun phrase extractor to identify basic noun phrases from free text. The extractor was implemented based on the lexical probability approach and the rule-based approach. The lexical probability was obtained from the WordNet lexical database, and the rules were defined manually. The second stage was multi-tag disambiguation to identify the noun phrase if this phrase were in more than one category. For example, “hit” could be either a noun or a verb. If the preceding word was a determiner (she, a, there, etc.), it would be tagged as a noun rather than a verb.

The quality of a student's message was measured by Keyword Density (KD), which was the proportion of the distinct noun phrase that was in the message generated by the student. In addition, the duplicated noun phrase posted by the same student was counted only once. It was denoted as

$$KD_i = \frac{NNP_i}{\sum NNP_j} = \frac{NNP_i}{NNP}$$

where KD_i was the keyword density for student i , NNP_i was the number of distinct noun phrases in student i 's messages, NNP_j was the number of distinct noun phrases in student j 's messages, and NNP was the total number of distinct noun phrases in all class messages.

The above approach proposed by Wu and Chen (2004) has used the concept space approach which is a method for users to retrieve documents in a digital library. It has been defined by the research (Zhang et al., 2002) that automatically generated keywords (terms) in document collections could be represented as the "concepts" and the "concepts" with their semantic relationships could be represented as the "concept space" as well. The similarity between two concepts (terms) was measured by the TF*TFD cosine formula in their definition. In the concept space, similar concepts were close to each other in the physical space to cover a topic. The figure 5-1 shows an example relationship among concepts (terms) and topics which are a cluster of concepts (same colour) in the visualization of concept space after a filtering process with a minimum similarity value.

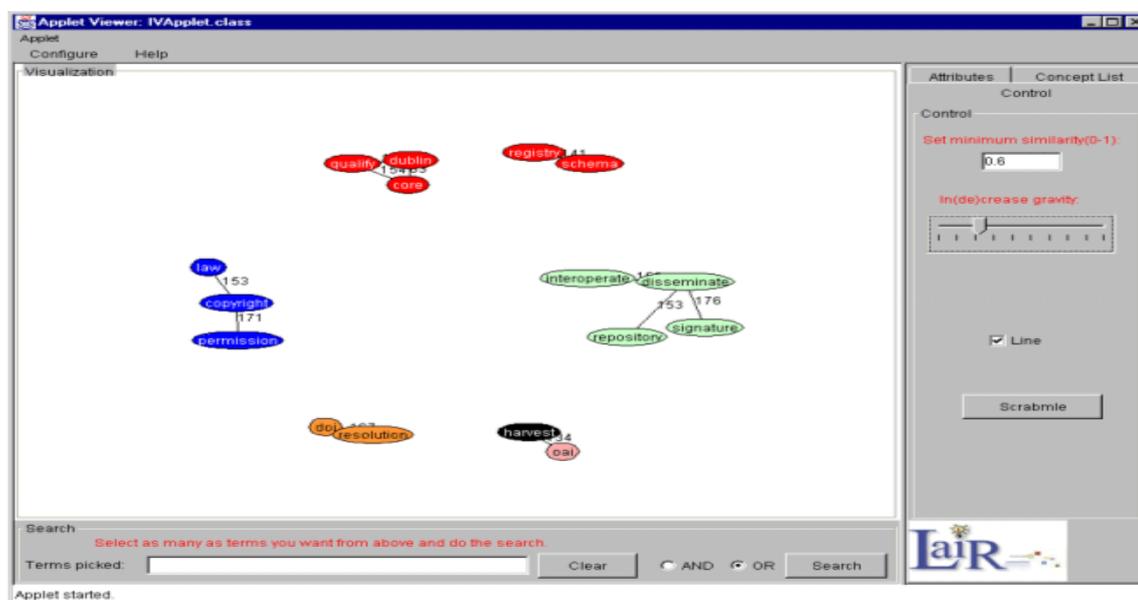


Figure 5-1: Visualization of filtered concept space (Zhang et al., 2002)

Therefore in the Quality Assessment model, all online messages in the discussion forum with their semantic relationships were assumed to be a concept space. Each message was represented as a sub-concept and a cluster of sub-concepts was constructing a big concept (topic). The model algorithm was calculating the semantic similarity of a sub-concept (an online message) with a big concept (a class messages set) to assess the quality of the students' messages. The figure 5-2 depicts an example of how the online messages are visualized as concepts in the discussion forum. In this example, there are two online messages and each of them is calculated for the semantic similarity with the class message set to assess the quality of those two messages in the discussion forum.

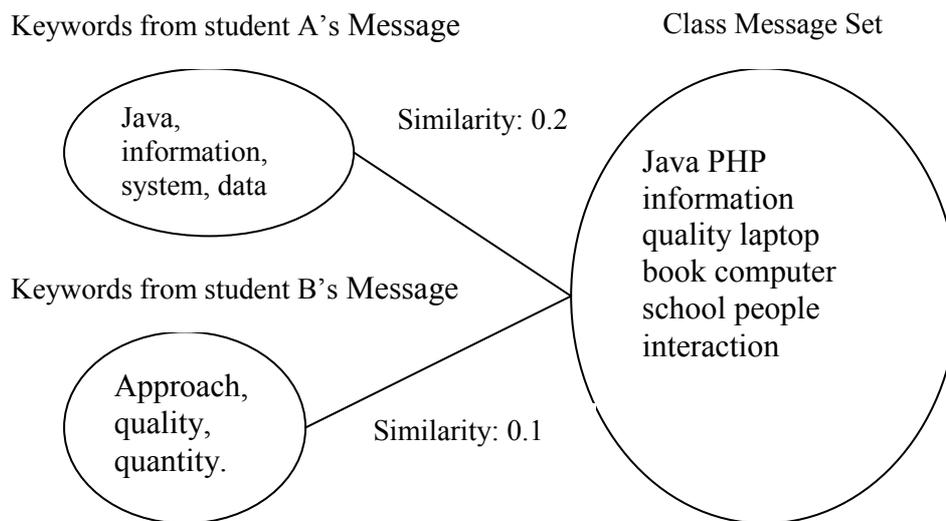


Figure 5-2: An Example of Keyword Density

Quantity Assessment

The quantity of a student's effort was measured by the length of the messages (ML) student generates. ML was calculated by counting all the words in the student's messages. The absolute number was proportioned by the local size, i.e. the number of

words in the entire class messages. Let ML_i be the number of words in message j of student i , and NW_{kj} be the number of words in message j of student k , the formula was denoted:

$$ML_i = \frac{\sum_j NW_{ij}}{\sum_k \sum_j NW_{kj}}$$

Participation Assessment

The participation frequency could be measured by the login times, but such information was not available inside the text messages. If posting a message is considered as a valid participation, participation could be measured by Message Count (MC), which was the proportion of messages that were generated by a student. MC_i was used to represent the message count of student i .

Performance Indicator

Taken together, the measures were proposed by Wu and Chen (2004) to compute a Performance Indicator (PI) score, which was defined as

$$PI_i = \alpha KD_i + \beta ML_i + \gamma MC_i$$

Where PI_i is the performance indicator assigned to student i , and the coefficients α , β , and γ were the weight of each of the three measures respectively. The coefficients were adjustable. Instructors could define the values by specifying the importance of each of the three evaluation aspects.

5.2.3 Ontology-based Information Retrieval System

Ontology is considered as any formalism with a well-defined mathematical interpretation which is capable of, at least, representing a sub-concept taxonomy, concept instances and user-defined relations between concepts (Nagypál, 2005). The use of ontologies to overcome the limitations of keyword-based search has been put forward as one of the motivations of the Semantic Web since its emergence in the late 90's (Vallet et al., 2004). Paralic and Kostial (1999) adopted ontology-based information retrieval approach in the Webocrate system. This proposed system was based on a domain knowledge representation schema in the form of ontology. Use of ontology enabled the definition of concepts and relations representing knowledge about a particular document in domain specific terms. In order to express the contents of a document explicitly, this system created links (associations) between the document and relevant parts of a domain model, i.e. links to those elements of the domain model, which were relevant to the contents of the document. Model elements could also be used for search and retrieval of relevant documents. In case all documents were linked to the same domain model, it was possible to calculate a similarity between documents using the proposed conceptual structure of this domain model. Such approach also supported 'soft' techniques, where a search engine could utilize the domain model to find concepts related to those specified by the user. The search engine could thus return every document linked to the concepts, which were close enough to the concepts mentioned in the user's query.

This approach did not consider type of relation in ontology for calculation of similarity between concepts. Moreover it assumed that the set of relevant concepts to the query was known. This condition could be achieved with any technique for

assigning concepts from ontology to a query, e.g. based on manual assignment or based on synonyms to query terms, making use of Wordnet or other similar techniques. The figure 5-3 depicts the framework of the Webocrate system process which was adopted with ontology concepts.

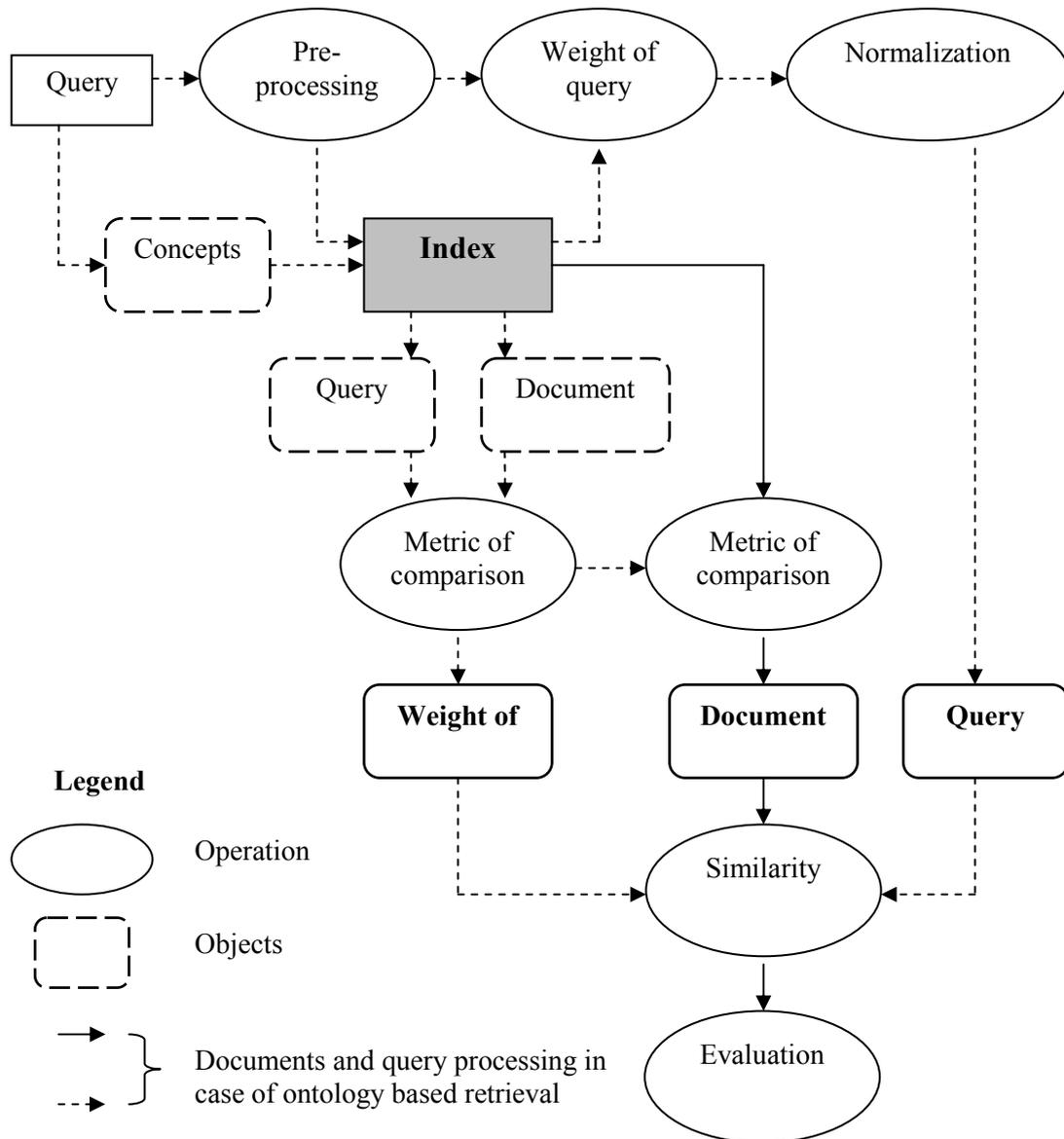


Figure 5-3: An Ontology-based Document Retrieval (Paralic and Kostial, 1999)

The way in which a query was processed by this approach is shown on the figure 5-3. In this approach, for a given query first appropriate concepts were retrieved – in this case manually from the user. Then the set of concepts associated with each document

was retrieved from the database. As next step, these two sets were compared using simple metric, which expressed the similarity between a document \bar{D}_i and given query \bar{Q} .

$$sim_{onto}(\bar{Q}, \bar{D}) = \frac{|Q_{con} \cup D_{i,con}|}{k} \neq 0$$

where Q_{con} was a set of concepts assigned to query Q and D_{con} was a set of concepts assigned to document \bar{D}_i , and k is small constant, e.g. 0.1. Resulted number represented ontology-based similarity measure. Better results had been achieved when this number had been combined with some of the previous two retrieval approaches described above (i.e. LSI approach or vector model). The final similarity was then computed as multiplication, e.g.

$$sim(Q, D_i) = sim_{onto}(Q, D_i) * sim_{TF-IDF}(Q, D_i)$$

Although the evaluation experiment had shown that ontology-based approach employed in the Webocrate system was very promising and might yield better precision-recall characteristics, the major problem of this approach was how to transform a user-defined query into a set of concepts from an actual ontology.

5.2.4 Dialogue Manager

User feedback is an important way to improve the system capability of understanding the meaning and preference of users. Goren-Bar and Kuflik (2004) introduced a Dialogue Manager component in their recommending restaurant system. The Dialogue Manager was able to provide the users with the three “best match” restaurants and perform the adaptation of the user model. The adaptation was based

on users' implicit feedback. Learning was performed by reinforcing the weight of the selected restaurants and weakening the weight of the non selected ones. If the user selected one of the three first recommended restaurants, the initial weight of the selected restaurant was increased and the weights of other two were decreased. If the user did not like the recommended restaurants and requested more restaurants, then if the user selects one of the next three restaurants, the selected restaurant weight was increased, and the weights of the other five were decreased. This approach could reinforce a specific positive or negative attitude towards the displayed items and update the user model (preference) every time the user provided relevance feedback regarding a recommended restaurant. This user feedback approach was recognized as a simple and effective method to constantly reinforce the accuracy of the system in understanding user preferences (Goren-Bar & Kuflik, 2004). For instance, if the restaurant R1 is selected and the restaurant R2 and R3 are not selected by the user, the weight of those three restaurants will be recalculated as:

$$W(R1)_{New} = W(R1)_{Original} + \beta F$$

$$W(R2)_{New} = W(R2)_{Original} - \beta F$$

$$W(R3)_{New} = W(R3)_{Original} - \beta F$$

where W represents Weight, F represents Feedback and β represents a parameter of how the Feedback influences the original Weight.

5.2.5 Summary

In other related research works, many novel approaches have been proposed and different methods have been adopted to enhance the system ability of understanding human natural language and preferences, particularly in the information retrieval and user feedback areas. The cosine similarity measure based on LSA was adopted as a

major measure for most information retrieval approaches because of its popularity and good text retrieval performance, but some research have modified the LSA in order to achieve a better performance. The Concept Relationship method and User Feedback approach were also adopted as important assistant methods to improve the communication between the user and system. The table 5-1 summarizes adopted method details, strengths and limitations of these approaches.

Table 5-1: The summary of related approaches

System or Approach Name	Cosine based on LSA	Concept Relationship	User Feedback	Strength	Limitation
Intelligent Discussion-bot	Yes	No	No	Good text retrieval performance.	Performance can be improved & no user feedback.
SELSA	Modified	No	No	Good text retrieval performance.	Performance can be improved & no user feedback.
Assessment Model	No	Yes	No	Improve the ability that the system understands user's input text.	Performance can be improved & no user feedback.
Webocrate	Modified	Yes	No	Good text retrieval performance.	Difficult to transform a user-defined query into a set of concepts from actual ontology.
Dialogue Manager	No	No	Yes	Improve the ability that the system understands user's preference	No natural language processing approach adopted.

In text similarity domain, LSA has been recognized to have an excellent performance for recall and precision of information retrieval (Milne, 2007). However, from the summarized related approaches in table 5-1, it can be seen that (modified) LSA with

cosine similarity measure, concept relationship and user feedback has various strengths and limitations to perform natural language processing tasks. The next section introduces the design and development in detail for the Dialogue Mechanism component, which integrates the modified cosine similarity measure, concept space and user feedback weighting to perform effective communication tasks in the proposed discussion forum systems.

5.3 Dialogue Mechanism Design

5.3.1 Introduction

Dialogue Mechanism is the core component in the proposed system for performing the communication tasks between the user and the system. As a result of the analysis of previous research contributions, this research proposes a novel information retrieval approach for this component to improve the communication between the user and the system in the proposed discussion forum systems. In this proposed approach, the popular cosine similarity measure based on LSA is used as the basic algorithm to calculate the text similarity of different postings in the proposed discussion forums. This measure assumes that each posting of the discussion forum is represented as a concept and all postings with their semantic relationships are represented as a concept space. LSA and cosine similarity approach are introduced next as the basic algorithm for this approach.

5.3.2 Cosine Methodology Based on LSA

LSA is a machine-learning model that induces representations of the meaning of words by analyzing the relation between words and passages in large bodies of

representative text. LSA is both a method (tool) used in industry to develop technology to improve educational applications, and a theory of knowledge representation used to model well known experimental effects in text comprehension and priming, among others (Landauer, 1997). It was originally developed in the context of information retrieval as a way of overcoming problems with polysemy and synonymy. Some words appear in the same contexts (synonyms) and an important part of word usage patterns is blurred by accidental and inessential information (Deerwester et al., 1990). LSA offers a way of assessing semantic similarity between any two samples of text in an automatic, unsupervised way and it has been used in applied settings with a surprising degree of success in areas like automatic essay grading (Foltz et al., 1999) and automatic tutoring to improve summarization skills in children (Kintsch et al., 2000).

The LSA methodology begins with the term-by-document matrix, an $n \times m$ matrix where each value in the matrix is the frequency of the n th word in the m th document. A weighting procedure is applied that weights each of the term frequencies (TF) by the inverse document frequency (IDF) (Salton et al., 1968). A very powerful mathematical procedure, known as Singular Value Decomposition (SVD), is then performed against the transformed matrix. SVD permits the reduction of any $n \times m$ matrix to a set of three matrices, such that $M = U\Sigma V^T$, where

$U = (m \times m \text{ matrix of left singular vectors})$

$\Sigma = (n \times m \text{ diagonal matrix containing the singular values of } M)$

$V^T = (n \times n \text{ transposed row matrix of the right singular vectors}).$

While the SVD solution of any given matrix can re-create the original matrix exactly, its primary value lies in its capacity to infer what the pattern of relationships and associations is for the words in the documents (Landauer et al., 1998).

Mathematically, let us assume a total of n postings in the discussion forum database, a student query posting q , and previous forum postings p_1, p_2, \dots, p_n . Let us also assume there were a total of m different unique words, w_1, w_2, \dots, w_m , found in all postings. LSA transforms all postings into a term-frequency matrix M , where m_{ij} is the weighted number of times word i appears in post j . It then decomposes M by Singular Value Decomposition (SVD) into three matrices such that $M = TSD$. The M matrix can be very large. For instance, it is using over 4500 postings which yield over 1100 terms (Landauer et al., 1998). LSA reduces S to k dimensions (where k is of the order of 300) resulting in the matrix S' . The matrices T , S' , and D can be multiplied, resulting in M' , the least squares best fit of M , where $M' = TS'D$ and $M \approx M'$. This finished semantic matrix is what we use to actually search the forum. Given terms in a posting query, the search system in the forum looks up the values for each search term/posts combination, calculates a cumulative score for every posting, and ranks the postings by that score, which is a measure of their similarity to the posting query. In practice, it will probably assign an empirically-determined threshold value to serve as a cut-off between relevant and irrelevant postings, so that the query does not return every posting in the discussion forum. Therefore the system finds all relevant postings from the discussion forum that are closest using cosine similarity measure of vector space.

$$sim(q, p_i) = cos_sim(q, p_i) = \sum_{j=1}^m w_q * w_{p_j}$$

The limitations of cosine similarity measure is in a large space many documents that are related to each other semantically might not share any words and thus appear very distant, and occasionally documents that are not related to each other might share common words and thus appear to be closer. Therefore this research is adopting the concept space and relevance feedback to reinforce the accuracy of relevant retrieval results based on the basic cosine measure to weaken its limitation. This proposed method assumes that in the discussion forum, all relevant postings are classified into different concepts across whole distributed discussion forum databases. The next section introduces how a posting is classified into a concept in the discussion forum.

5.3.3 Framework Design

The Dialogue Mechanism component is designed to perform information retrieval and user interaction tasks in the proposed discussion forum system. Dialogue Mechanism consists of two agents which handle concept classification process and query monitoring process. The figure 5-4 depicts the framework of Dialogue Mechanism. After the component receives the text of submitted postings, Concept Classification Agent performs a concept classification process. After concept classification processing is completed, Query Monitoring Agent then starts a relevant information search process. Relevant postings are found through the information retrieval processing and then returned to the users. Moreover, users are able to evaluate the retrieved postings and give feedback to the system. When Dialogue Mechanism receives the feedback from the user, Concept Classification Agent performs the concept reclassification process. The details of approaches adopted for Dialogue Mechanism development are introduced in the next two sections.

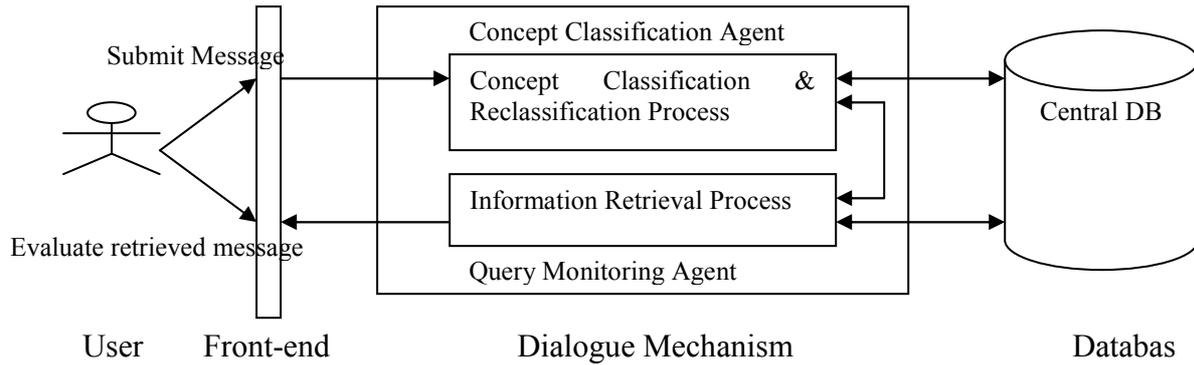


Figure 5-4: Framework of Dialogue Mechanism

5.3.4 Concept Classification & Information Retrieval

In the proposed system, the first task for Dialogue Mechanism is to receive the text of submitted postings from the front-end side and then start a concept classification process. Concept Classification Agent is designed for performing this process. In the discussion forum, there are semantic relationships among all postings. Some posting may be related in semantic meanings, but the relevant information can be posted anywhere in the discussion forum and relevant postings are specifically related to one topic (a big concept) such as J2EE or LSA. Thousands of postings may be submitted anywhere in the discussion forum and they might be related to hundreds of different topics, making it impossible for students to capture all relevant postings that they might need or be interested in. The proposed approach for concept classification processing in Dialogue Mechanism is to set up a concept space for the discussion forum where every posting belongs to a certain semantic concept. It assumes that each posting representing sub-concept belongs to a semantic concept in the discussion forum. The introduced cosine similarity algorithm based on LSA is used to calculate the similarity between a posting (p) and a semantic concept (C):

$$sim(q, C) = cos_sim(q, C)$$

The Concept Classification Agent calculates the similarity weight between each submitted posting with every existing semantic concept. The result could then be in two scenarios: 1) if there is no similarity weight greater than a default threshold, the submitted posting itself will become a concept in the concept space; or 2) if there is a group of concepts whose similarity weights are greater than the threshold, the submitted posting will be classified as a sub-concept of the semantic concept with maximum similarity value in the concept space. Therefore all submitted postings in the discussion forum should be classified into different semantic concepts in the concept space. The figure 5-5 describes a concept classification example for a submitted posting in a discussion forum. In this example, there are two existing concepts (Concept 1 and Concept 2) in the concept space and the newly submitted posting is classified into the semantic concept 1 (Java) because the similarity between them is greater than threshold (0.25) and the highest weight (0.4).

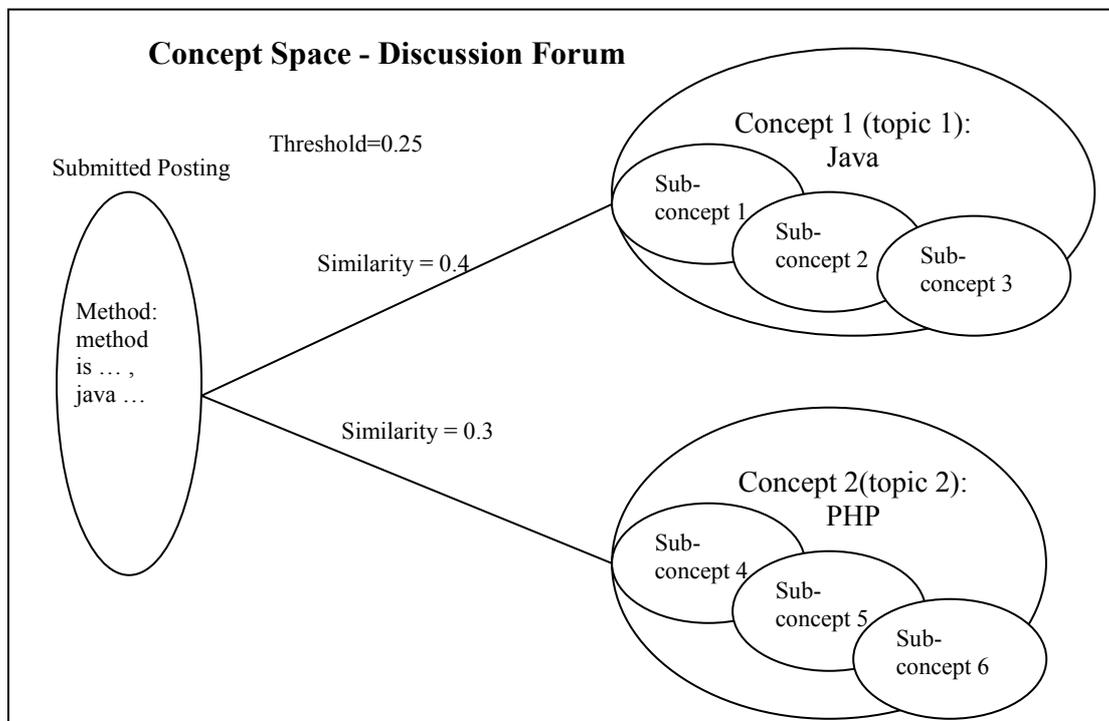


Figure 5-5: An Example of Concept Classification in a Discussion Forum

When a message is submitted into the discussion forum and sent to the Dialogue Mechanism component, the Concept Classification Agent performs the concept classification task to identify which concept this message belongs to. After the submitted posting is classified into a concept, the relationship between it with other postings (sub-concepts) can be identified and information retrieval process is performed by Query Monitoring Agent. This relationship is represented by the similarity among them. Such relationship is clearly shown in the concept space view. The figure 5-6 depicts the relationships between sub-concepts and between concepts in the concept space view, such as sub-concept 1 with sub-concept 2, sub-concept 2 with sub-concept 3, and concept 1 with concept 2.

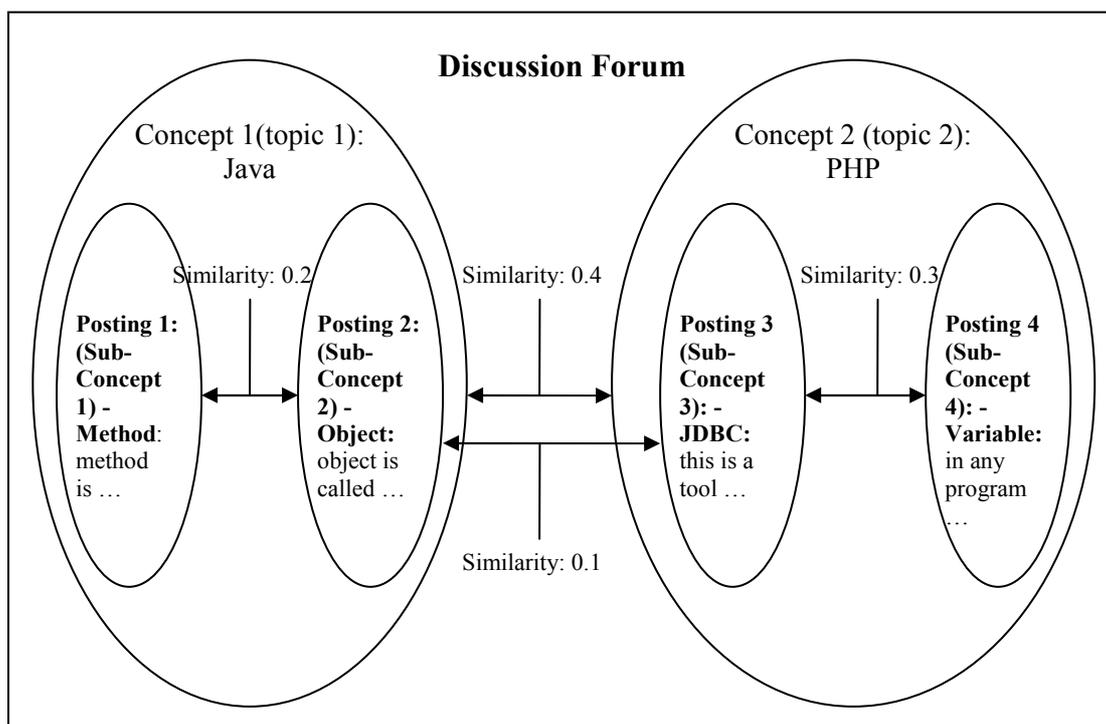


Figure 5-6: An Example of Relationships between Sub-Concepts and Concepts

Based on the concept classification measure, the basic similarity calculation measure is modified to calculate the relationship - similarity (*sim*) between a submitted posting (*p*) with each existing sub-concept s_1, s_2, \dots, s_n belonging to both same concept C_s and

other concepts C_o for the information retrieval process. Based on a set-up threshold of the similarity calculation, the Query Monitoring Agent finds out the most relevant sub-concepts and returns them to the users. Two assumptions made for this proposed similarity calculation algorithms based on the cosine similarity measure are denoted as follows:

- Assumption 1: the existing sub-concept s_i belongs to the same concept with submitted sub-concept p :

$$sim(p, s_i) = cos_sim(p, s_i)$$

- Assumption 2: the existing sub-concept s_i belongs to a different concept with submitted sub-concept p :

$$sim(p, s_i) = \alpha cos_sim(C_s, C_o) \times cos_sim(p, s_i)$$

where α is a parameter that controls the influence of the relationship between two semantic concepts.

In the concept space, the relationship between two semantic concepts is represented by the similarity between them. Therefore the relationship between two sub-concepts can be influenced if those two sub-concepts belong to different semantic concepts. This proposed approach for Dialogue Mechanism component design and development has combined a modified cosine similarity measure based on LSA and concept space to perform the text similarity calculation and information retrieval tasks. As previous research contributions have described that relevance feedback is another important assistant approach to reinforce the performance of how the system understands users meaning and preference, and that it is also able to weaken the word/phrase semantic confusion limitation of cosine similarity measure, such as different meaning for one word and different words for one meaning, relevance feedback is integrated with

above proposed approach to re-enhance the capability of Dialogue Mechanism component to perform natural language processing tasks. The next section introduces how relevance feedback can be integrated with above approach to reinforce the performance of Dialogue Mechanism component.

5.3.5 Relevance Feedback and Concept Reclassification

This section presents how the relevance feedback function is integrated with the proposed text similarity calculation approach for Dialogue Mechanism component to perform natural language processing tasks. In the last section, a synchronised response function was introduced in the proposed discussion forum system that, when a user submits a message, the Query Monitoring Agent retrieves most relevant existing postings in the discussion forum and then returns them to the user. Therefore, for the design of relevance feedback function in the proposed system, users should be able to evaluate each retrieved posting (s_i) after they submit a posting (p) into the proposed discussion forum system. Users can give a positive feedback (f_i) (relevant) against each retrieved posting within this function. After users evaluate retrieved postings, the Concept Classification Agent performs the concept reclassification process with users' feedback. The proposed relevance feedback approach is designed to recalculate the relationship (similarity) between the evaluated retrieved postings with the semantic concept corresponding to the submitted message. In this process, the relevance feedback takes into account the calculation of similarity weight.

After a posting (p) is submitted and sent to Dialogue Mechanism component to perform the concept classification process, the Concept Classification Agent classifies it into a semantic concept (C_p) and then the Query Monitoring Agent performs the

information retrieval task to search the existing relevant information in the discussion forum and returns a list of relevant postings to the user. In the list of returned relevant postings, each returned posting (s) which was previously classified into a semantic concept (C_s) can be evaluated. The user can submit feedback by identifying which returned posting is related to the submitted message (the feedback is f). After the user's feedback is sent back to Dialogue Mechanism, the Concept Classification Agent uses a modified similarity calculation algorithm to recalculate the similarity (relationship) between the evaluated posting (s) with the semantic concept (C_p) which the submitted posting corresponds to. This measure is denoted as the following algorithm:

$$sim(s_i, C_p) = sim(s_i, C_s) + \beta \times \sum f_i / N$$

where N is the total number of feedback against the posting (s_i) and β is the parameter of influence of user's feedback against the concept classification.

After the similarity recalculation between the evaluated posting (s) with the concept (C_p) corresponding to the submitted posting, if the new similarity weight is higher than the similarity weight with its current corresponding concept (C_s), the system reclassifies this evaluated posting from the current concept (C_s) into the other concept (C_p). The figure 5-7 and figure 5-8 present an example of the concept reclassification process. The figure 5-7 describes the concept classification and information retrieval processes and the figure 5-8 shows the concept reclassification process with a user feedback.

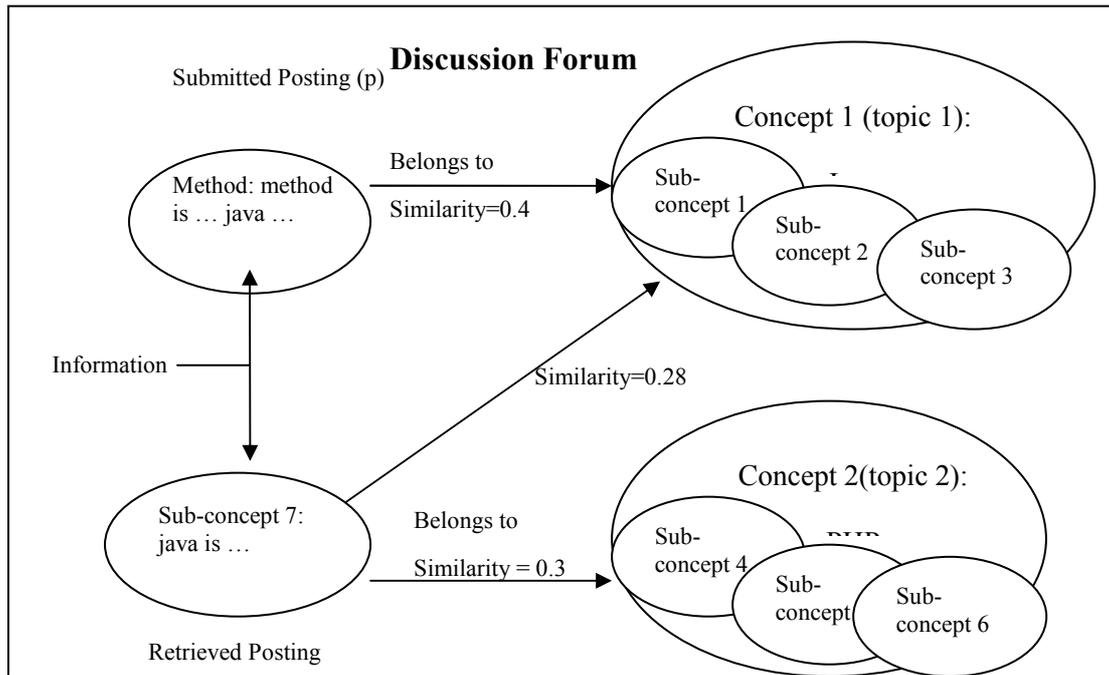


Figure 5-7: An Example of Concept Reclassification – Before User Evaluation

In the example (Figure 5-7), a posting p is submitted into a discussion forum and then sub-concept 7 is retrieved as a relevant posting and it belongs to concept 2. After concept classification process, p is calculated as a sub-concept of concept 1. The similarity between p with concept 1 is 0.4, between sub-concept 7 with concept 2 is 0.3 and between sub-concept 7 with concept 1 is 0.28. After a user evaluates that sub-concept 7 (the retrieved posting) is a relevant posting with the submitted posting p , the user's feedback is sent back to Dialogue Mechanism, and the concept reclassification process is preformed (Figure 5-8). The similarity between sub-concept 7 with concept 1 is recalculated from previous 0.28 to current 0.32 which is greater than the similarity weight 0.3 between sub-concept 7 with concept 2. Therefore sub-concept 7 is reclassified from concept 2 to concept 1. Finally with user's feedback submission, sub-concept 7 no longer belongs to concept 2 but concept 1.

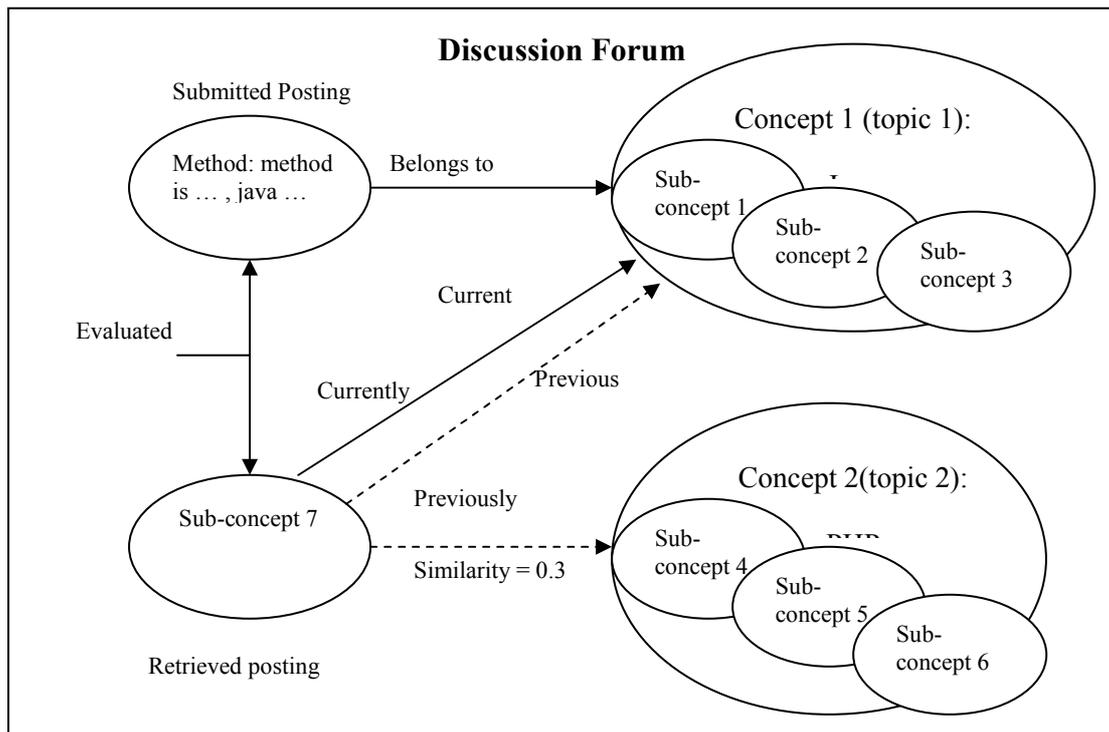


Figure 5-8: An Example of Concept Reclassification – After User Feedback

This concept reclassification approach is based on relevance feedback from users. The more relevance feedback is given from the users the more postings are accurately classified into semantic concepts. Moreover, the influence parameter β is also highly dependent on the number of relevance feedback. If a posting is classified into a correct concept, the proposed system can perform more effective natural language processing tasks so that it can retrieve more accurately relevant information for the users.

5.3.6 Summary

In order to achieve the core objective of this research, which is to improve the effectiveness of communication in the discussion forum, the core component of the proposed system, Dialogue Mechanism, was designed and developed to perform natural language processing tasks. Two agents were designed for the Dialogue

Mechanism component to perform the natural language processing tasks. For both agents to perform text semantic similarity calculation and information retrieval tasks, the adopted approach integrated the cosine similarity measure based on LSA and concept space theories to enhance the system's intelligence capability of natural language understanding in the learning communities. However, due to the known obstacles, it is impossible to understand 100% correctly the real meaning of users' input in the discussion forums. Because user feedback is recognized as an effective assistant approach to reinforce the performance for information retrieval, it has been integrated with the Dialogue Mechanism component in the proposed system. After the system's high-level architecture and components design, a prototype of the proposed system is implemented in this research. In the next chapter, the development plan of the prototype will be introduced.

CHAPTER 6

DEVELOPMENT PLAN

6.1 Introduction

The last two chapters described a new type of discussion forum systems that aims to overcome some limitations of traditional discussion forums and improve the effectiveness of communication in learning communities. A prototype of this type of systems was developed. This chapter first introduces an overview of the development plan and then presents various technologies selected for use in the development process. It then analyses the benefits and limitations of available open source discussion forum systems in order to select an appropriate forum system to be customised in the development process. This is followed by a description of the proposed development approach, and the development environment and tools.

6.2 Development Overview

A high level design overview of system components, main flows and controls for the proposed system are described in Chapter 4 and can be seen in Figure 6-1. This diagram represents the new system as split into three layers and shows how it integrates with existing forum systems. Web Services (WS) are used as the communication method between components. A key requirement of this new forum system is that it should integrate with existing discussion forum systems so it can be used for various courses in different institutions and for reuse of the knowledge within

and among different institutions. This new system should also have all the general features of traditional discussion forums, such as enabling students to view, post and reply to messages, and new features such as being able to view and evaluate information retrieval results.

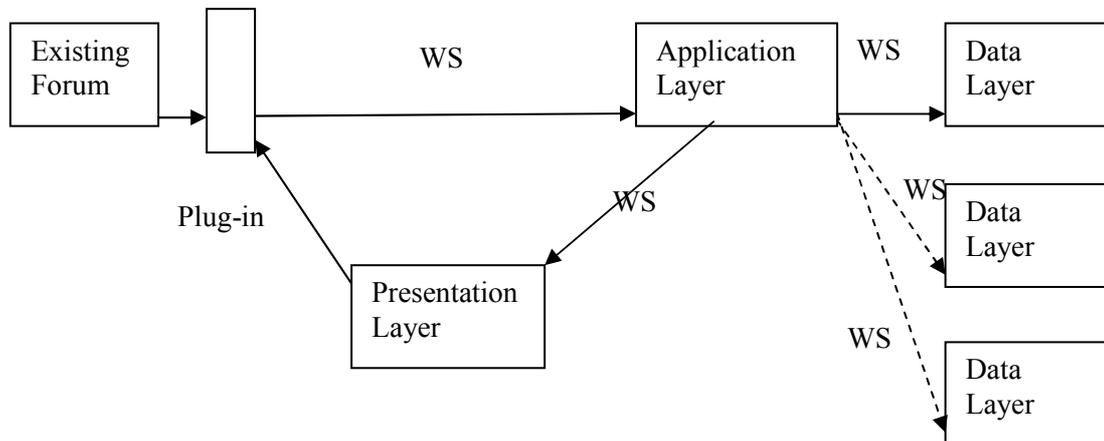


Figure 6-1: High level system overview

The new system was therefore not implemented from scratch but rather “built” upon existing systems. This is because traditional discussion forums have functions for general features which can be reused in the new system. The presentation layer of the new system was integrated with existing discussion forum systems via a plug-in component which shows relevant information on search results for users to view and evaluate.

Table 6-1 illustrates the components of the system that were implemented. Since the new system needs to communicate with existing forums, the plug-in component was developed as a bridge for sending and receiving data between two systems. The interface development of the presentation layer was intended to align with the interface development of the existing discussion forum systems to address the

integration issue. The presentation layer includes the user interface of the new system. It generates information in HTML format, which is then sent to a user's web browser. In the development plan for the presentation layer, the functions for general features were not developed because they could be reused from the existing discussion forum systems. Rather the functions for new features (the interface for information retrieval results and the user feedback function) were developed from scratch and integrated with the general functions of existing discussion forums. The dialogue mechanism component of the application layer was developed to perform the text similarity calculation and information retrieval tasks. The data layer was developed to store all data from different institutions into databases.

Table 6-1: New system development plan

Component Name		Development	Component Details
Plug-in		Yes	Integrates the new system with the existing discussion forum systems.
Presentation Layer	General Features	No	Reuses functions for general features from existing forum systems.
	New Features	Yes	Can be integrated with the presentation layer of existing forum systems.
Application Layer		Yes	The dialogue mechanism is the core component which performs the text similarity calculation and information retrieval tasks.
Data Layer		Yes	The data structure of the new system in the database.

According to the prototype system development plan, it was important to select an existing discussion forum system for customisation, which could then be integrated with the new system to achieve the research objectives. The next section describes the process for selection of an existing discussion forum system for the proposed system development.

6.3 Forum selection

6.3.1 Requirement Analysis

To be an appropriate existing discussion forum for this research, the candidate discussion forum system needs be customised and integrated with the new system. Based on the system requirements in Chapter 3, the forum should meet the following criteria:

- The candidate discussion forum should have the standard features such as the functionality for viewing, posting and replying messages.
- The candidate discussion forum should be available through open source licence.
- The candidate discussion forum should have a friendly user interface.
- The framework of the candidate discussion forum should be easily customisable.
- The candidate discussion forum should be able to be installed and run on different operating systems, such as Windows and Linux platforms.

To meet the above criteria, only those discussion forum systems which have an open source license were considered for selection. The next section outlines a number of open source discussion forum systems and describes the selection process.

6.3.2 Survey of Existing Open Source Forum Systems

A number of features proposed in the new system aim to enhance existing discussion forums. Existing open source discussion forums are therefore examined in order to select one of them as the basis, instead of starting development from scratch. Table 6-2 shows an overview of the open source forums evaluated in this study.

Table 6-2: A list of open source forums

Name	URL	Licence	Implementation Language	Message Structure
Beehive	http://beehiveforum.net	GPL	PHP	Threaded
phpBB	http://www.phpbb.com	GPL	PHP	Flat
PunBB	http://www.punbb.org	GPL	PHP	Flat
JavaBB	http://www.javabb.org	Apache License	Java	Flat
JForum	http://www.jforum.net	BSD License	Java	Flat
Yazd	http://www.forumsoftware.ca	Apache License	Java	Flat
Moodle Forum	http://moodle.org	GPL	PHP	Threaded / Flat

The three types of different open source licenses are:

- GPL – GNU General Public License (Free Software Foundation, 1991)
- BSD – (Modified) Berkeley Software Distribution License (XFree86, 1996)
- Apache License (Apache, 2004)

The main difference between these licenses is that with the GPL, any modifications or extensions made to the source code also have to be open source (under a compatible license). The other licenses are less strict in this sense and can easily be combined with closed source code. With the above candidate discussion forums, only those with an open source licence can be selected as a basis system to be customised.

Beehive Forum, phpBB and PunBB use the same programming language – PHP. Beehive is an open-source project for creating a highly configurable frame-based discussion forum (Figure 6-2). Beehive Forum has many features. The main ones are listed below (<http://beehiveforum.sourceforge.net>):

- Frame-based layout, for easy navigation.
- Reply-to-user posting and e-mail notification.
- Powerful HTML posting, without security risks.
- Extendable user profiles.
- Flexible relationships system.
- Private Messaging system.
- Multiple style options, including dyslexia support.
- Advanced word-filtering facility.
- Comprehensive user permissions.
- Full, easy to use admin tools.

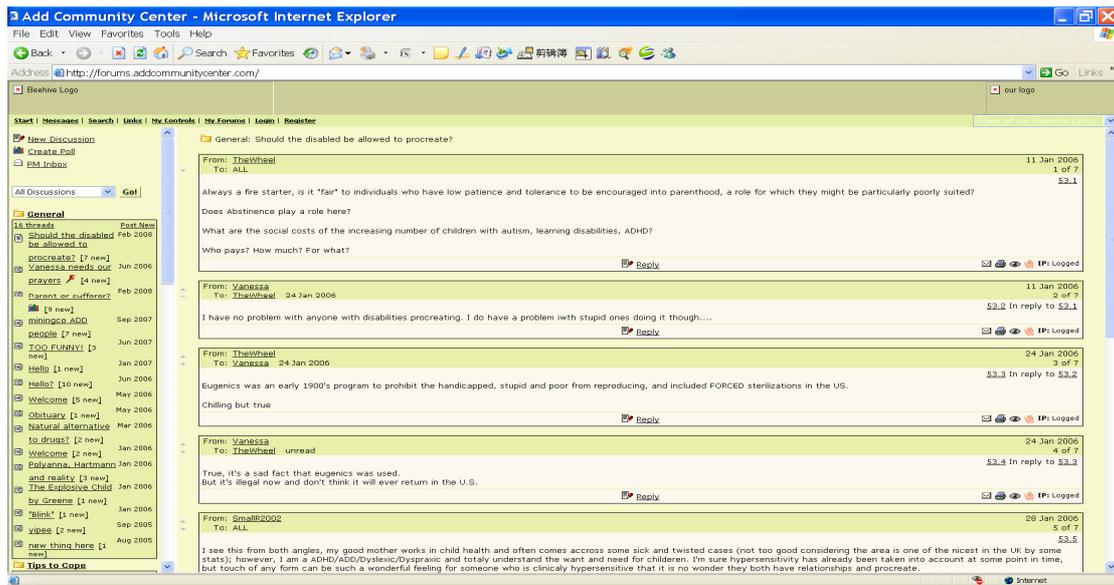


Figure 6-2: A Sample Screenshot of a Beehive Forum Website

PhpBB is a popular Internet forum package written in the PHP scripting language (Figure 6-3). Some of phpBB's major features are (<http://www.phpbb.com>)

- Modular design for the Admin Control Panel, Moderator Control Panel, and User Control Panel
- Support for multiple database management systems, including MySQL, Microsoft SQL Server, Oracle, PostgreSQL, SQLite, Firebird, OpenLink Virtuoso, and other ODBC-accessible DBMS
- Support for unlimited levels of sub-forums
- Ability to create custom-defined BB-Code
- Ability to create custom profile fields
- Permissions system

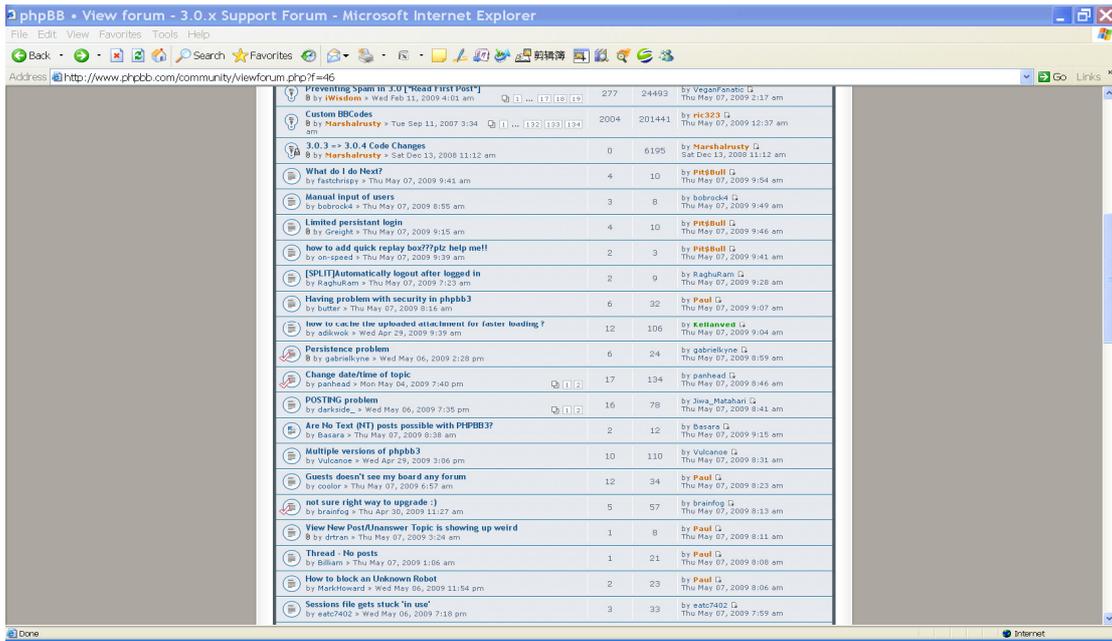


Figure 6-3: A Sample Screenshot of phpBB Forum

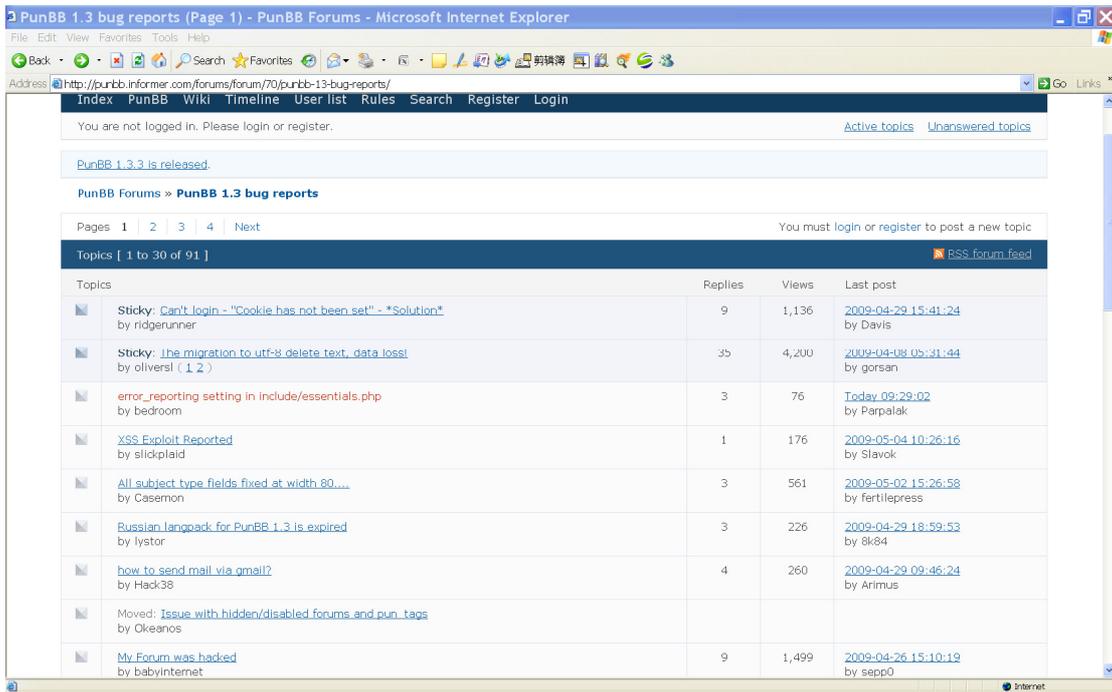


Figure 6-4: A Sample Screenshot of PunBB Forum

PunBB is a fast and lightweight PHP-powered discussion board (Figure 6-4) which has the following features (<http://www.punbb.org>):

- runs Fast
- simple layout and design
- easy to administrate and moderate
- non-output, huge, non-valid, (X)HTML pages
- source code easy to read and understand

Moodle forum is a component of Moodle system which is a Course Management System (CMS), also known as a Learning Management System (LMS) or a Virtual Learning Environment (VLE). The Moodle system is a free web application that educators can use to create effective online learning sites. Since the proposed system is integrated with an independent discussion forum system, Moodle forum was not seen as a suitable candidate in the research project.

6.3.3 Base forum selection

Evaluations of the features of all candidate discussion forums were undertaken. These included evaluations as to how well and in what language (some were not English) the source code was documented. Additionally, assessments of how well their data models would match with the proposed system were also carried out.

It was decided to base the system on the existing Jforum system. This was because the JAVA programming language had already been selected as the dialogue mechanism development language. As described in the previous chapter, this was because:

- JAVA is a powerful web application development language for complex applications that require huge data computation. The dialogue mechanism of

the proposed system does indeed require a complex algorithm to perform huge data computations.

- JAVA is platform independent which will make the proposed system easier to integrate with various discussion forum systems on different types of servers.

Since Web Services are used for communicating between the system's layers, the fact that this requires the Presentation Layer to be written in the same language is not seen as a major problem. Using Web Services hides the implementation languages of the various layers from each other anyway.

Therefore as Java is selected for the core Dialogue Mechanism component, the candidate discussion forum systems developed in Java are more easily customised and integrated with the proposed system. Jforum has a BSD license and is implemented in Java. Moreover it is a powerful and robust discussion forum system which provides an efficient forum engine, an easy to use administrative panel, and an advanced permission control system. The interface of Jforum is very attractive. It can be installed and run on a Servlet container or Application Server, such as Tomcat, Resin or JBoss (<http://www.jforum.net>). The other important advantage of Jforum is its system framework. Jforum has a clean design and implementation architecture, the Model-View-Controller (MVC) framework, which can be easily customised and extended.

MVC is a classic design pattern widely used in GUI and Web applications. Generally speaking, Model defines and maintains the data, View renders the interactions based

on the data, and Controller coordinates actions and events that affect the Model and View (Gamma et al., 1995). Figure 6-5 presents the MVC framework.

- Input --> Processing --> Output
- Controller --> Model --> View

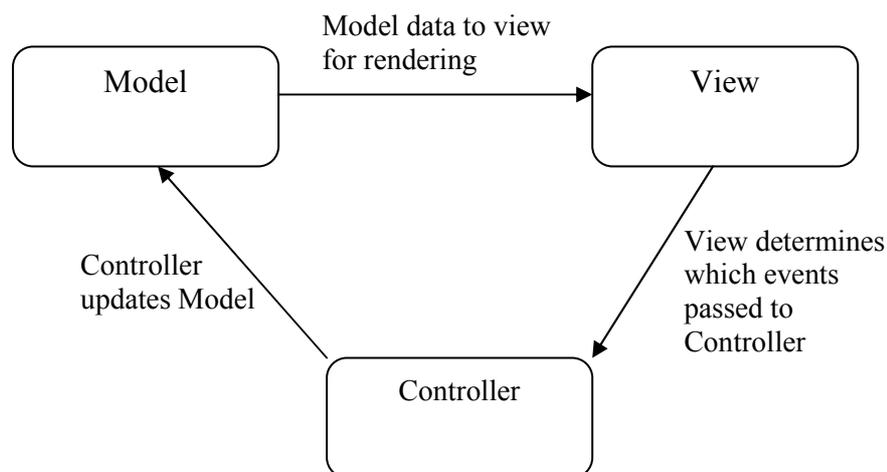


Figure 6-5: Model-View-Controller Framework (Gamma et al., 1995)

Based on the above analysis and comparison, Jforum was selected as the best match for the research objectives and system characteristics. Specifically, its BSD license, Java-based development language and simple framework (which can be easily customised and extended) make Jforum the best candidate for the base discussion forum system in this research project. The next section introduces the technologies used for developing the proposed system.

6.4 Development Approaches and Technologies Used

As outlined in previous chapters, the dialogue mechanism is the core component that performs information retrieval and user interaction tasks. It consists of two intelligent agents to handle text semantic similarity calculations and information retrieval processes. Meanwhile, Web Services are used as the communication method between components. This section first introduces concepts for Agent, Web Services, and a Java-based methodology tool. Then the development approaches, environments and tools for implementing the components of the proposed system are presented.

6.4.1 Web Services

A Web Service is a piece of software whose interface and binding can be defined, described, and discovered as XML artefacts (Alonso et al., 2003). The Web Services concept has recently gained popularity as part of the Service-Oriented Architecture (SOA) approach. SOA is often defined as nodes making resources available to other nodes in the form of independent services. The important properties of such services are usually defined as: being loosely coupled, having well specified interfaces and concealing their underlying implementation details (Kobielus & Manes, 2004). Web Services, as defined by the W3C (W3C, 2008), are based on Extensible Markup Language (XML) and are normally implemented using Web Services Description Language (WSDL) definitions to describe each service and by using SOAP envelopes to send requests and responses between the nodes.

6.4.2 Agents

There is no consensus in the literature on the definition of the term ‘agent’ because it has been used to signify numerous things within the field of computer science (Wooldridge, 1999). The agent paradigm has expanded from mere communication tools to embrace autonomous ‘intelligent’ software entities used in business and educational areas to represent their users or owners (Balke & Eymann, 2008). In this research, the reason for using agents for various components of the proposed system is that they are components acting and interacting on behalf of (representing) a user to classify input and gather and return relevant information for that user.

6.4.3 Object Relational Mapping (ORM) Tools

Object/Relational Mapping (ORM) provides a methodology and a mechanism for object-oriented systems to store their long-term data safely in a database while maintaining transactional control over it, yet having it expressed in program objects when needed (O’Neil, 2008). Two commonly used ORM tools for development of web applications are Hibernate and the Active Record part of the Ruby on Rails framework (Peak, 2005). Hibernate is an object/relational persistence and query service for Java (Hibernate, 2006) while Rails is an integrated web application framework written in the Ruby language. Basically, these tools automatically map data structures between a relational model stored in a database and an object model as used for object oriented programming. They are then in charge of updating the database when changes occur to data in the object model. The prototype system was implemented using these technologies and it successfully met the specified requirements.

6.4.4 Development Approaches

In the above sections, Java was selected as the programming language for the development of dialogue mechanism in the application layer of the prototype system (to be implemented from scratch). The integrated discussion forum system, Jforum, is also developed in Java. The advantages for Java are that as an object-oriented, platform independent and powerful programming language it is strict and statically typed (encouraging good design and coding standards). For other components of the new system such as the presentation layer, plug-in component, and data layer, the development approaches are presented as follows:

- To develop the forum display component of Presentation Layer, or user interface part, we can use FreeMarker to create a Jforum layout. FreeMarker is a "template engine"; a generic tool to generate text output (anything from HTML to auto-generated source code) based on templates and it is a Java package and a class library for Java programmers (<http://www.jforum.net>). For best integration with the presentation layer of Jforum, the forum display component of the proposed system is developed in FreeMarker.
- To integrate the proposed system with Jforum, the Plug-in component is implemented in Java and the communication approach for transferring data from Jforum to the proposed system uses Web Services which are language neutral in that they do not expose the language used in the implementation of the services they provide. Therefore the proposed Plug-in component can be integrated with any type of forum system regardless of which development language is used for the Presentation Layer development. Web Services are used to communicate between the Presentation Layer of Jforum and the Application Layer of the proposed system via the Plug-in component, and

between the Application Layer and Data Layer within the proposed system. This occurs in two ways. Firstly a request from the user interface implemented in the Presentation Layer is wrapped in a SOAP envelope and then a Web Service call is made to send it via the Plug-in component to the main Application Layer (Dialogue Mechanism). These two parts of the system can be located on the same physical server or on different servers. The second use of Web Services in the system is for making requests from the Application Layer to the Data Layer. In this case, one request from the Application Layer will normally call both the Data Layer Web Service within the system and the Data Layers of all other institutions participating in this discussion forum. How many Data Layers on different servers will be called depends on how many institutions are participating in the forum.

- A modelling tool was used to design an entity relationship (ER) diagram for the Data Layer. Originally, the design was based on the data structure of the Jforum module. It was then changed and extended considerably to support the specific requirements of the prototype system. The ER diagram was used to generate SQL script to create the actual database. The database is accessed to generate Java persistency classes for use with Hibernate in the Data Layer. Choosing a modelling approach for the Data Layer ensures that the documenting diagram remains up to date as the code is generated from the diagram.

A code driven approach is used for the Web Services implementation, in that the Java classes implementing the functionality are written first. The XML based Web Service Definition Language (WSDL) descriptions and supporting files are then generated

from the Java source code. This appears in the classes as Web Services. This is done using the Eclipse Web Tools Platform (WTP) (Eclipse Web Tools Platform, 2008) package and the files are tweaked manually to correct minor errors in generation. Apache Axis (Apache Axis, 2008) is used in combination with Apache Tomcat (Apache Tomcat, 2008) to operate Web Services.

For this study a range of development approaches were used for the implementations of various component. Each approach was considered in order to give the best advantage to each individual component implemented in the prototype system. A single unified approach would not have been an effective method given the diverse challenges of implementing each of the components. The next section will introduce the development environment (such as the server environment) and software tools (such as the development environment and database management software) chosen for the development process.

6.4.5 Development Environment and Tools

For the development environment, free software options are chosen when available. With the exceptions of the Sun JDK (Java Development Kit) (Sun JDK, 2008), Microsoft Windows (Microsoft Windows, 2008) and Microsoft Office (Microsoft Office, 2007), all software used is open source. Eclipse (Eclipse, 2008) is used as the Integrated Development Environment (IDE) with plug-ins for web development: Eclipse Web Tools Platform – WTP (Eclipse Web Tools Platform, 2008), Hibernate (Hibernate, 2006), object-persistence with MiddlegenIDE (MiddlegenIDE, 2006) and Subversion connection (Subclipse) (Subclipse, 2008). Both Putty for secure shell

access (Putty, 2008) and WinSCP for file transfer (WinSCP, 2008) are valuable tools used to manage the Window or Linux servers on PCs.

Various server-based softwares are used in the development process. Subversion (Subversion, 2008) is used as the version control system for source code and documents. This is mainly because it is considered superior to the other common open source choice of CVS (CVS, 2008) by the author of this thesis. One convincing argument for this is that it keeps track of a whole tree of files, and not just the separate files. This is a major advantage when refactoring the code. For the web server, Apache (Apache, 2008) is used in combination with Apache Tomcat for Java (Apache Tomcat, 2008) and Apache Axis for Web Services (Apache Axis, 2008). MySQL (MySQL, 2008) is chosen for the database server in the system, as it is able to provide the required features. It is currently a very common choice for web based applications.

The choice of development software tools shows that free software tools are more than adequate to handle the requirements. After the design and development plan for the proposed system were completed, a prototype system was implemented to meet the proposed system requirements and achieve the objectives of this research. The next chapter introduces how the proposed system was developed, presented and integrated with the Jforum discussion forum system.

CHAPTER 7

ARCHITECTURE AND IMPLEMENTATION

7.1 Introduction

Chapter 6 outlined a new type of discussion forum systems designed to solve the research problems introduced in earlier chapters. A prototype system was implemented based on the new system design. The development plan for this prototype system is described in Chapter 6 which also discusses the selection of Jforum as the base discussion forum system which was customised and integrated with the prototype system. In this chapter, the architecture of the prototype system is described first, followed by the details of the prototype system implementation.

7.2 Architecture Overview

The high level design of system components, main flows of control of the proposed system and development plan for the prototype system implementation were presented in chapters 4 and 6. The prototype system comprises three layers and it integrates with Jforum. Figure 7-1 shows the high-level architecture for each component of the prototype system. The prototype was implemented in this research to represent one instance of the proposed new type of discussion forum systems meeting the design requirements and research objectives. This was because the method for connecting the prototype system to the remote databases (other forum systems in different institutions) and retrieving data is same as for connecting to the

local database. It was therefore not necessary to implement remote database connection (the dotted line and box) in this research. Only the global database running on a central server and the local database running on the local server were implemented. Figure 7-1 shows how Jforum integrates with the core components of the prototype system via the plug-in component. The global database box represents a central database that stores global information from local and remote systems. The local database box represents the database for Jforum running on a local server. The following sub-sections introduce the system architecture for each component and then present the data architecture and deployment architecture for the prototype system.

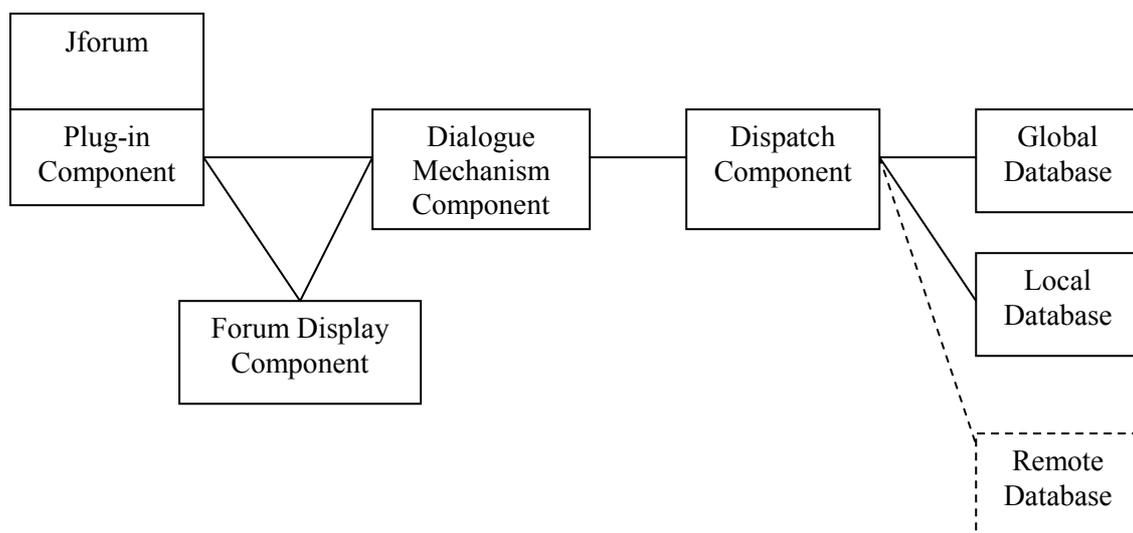


Figure 7-1: High-level System Architecture

7.2.1 Plug-in Component

The plug-in component was designed and developed to integrate the core part of the prototype system with Jforum (customised as the new type of discussion forum system). Jforum is an open source discussion forum system which provides an API for customisation and integration. The plug-in component works with the API to send and receive the data between Jforum and the core part of the prototype system. It has four

classes: the data receiving class, the data encapsulation class, the data sending class and the forum display class.

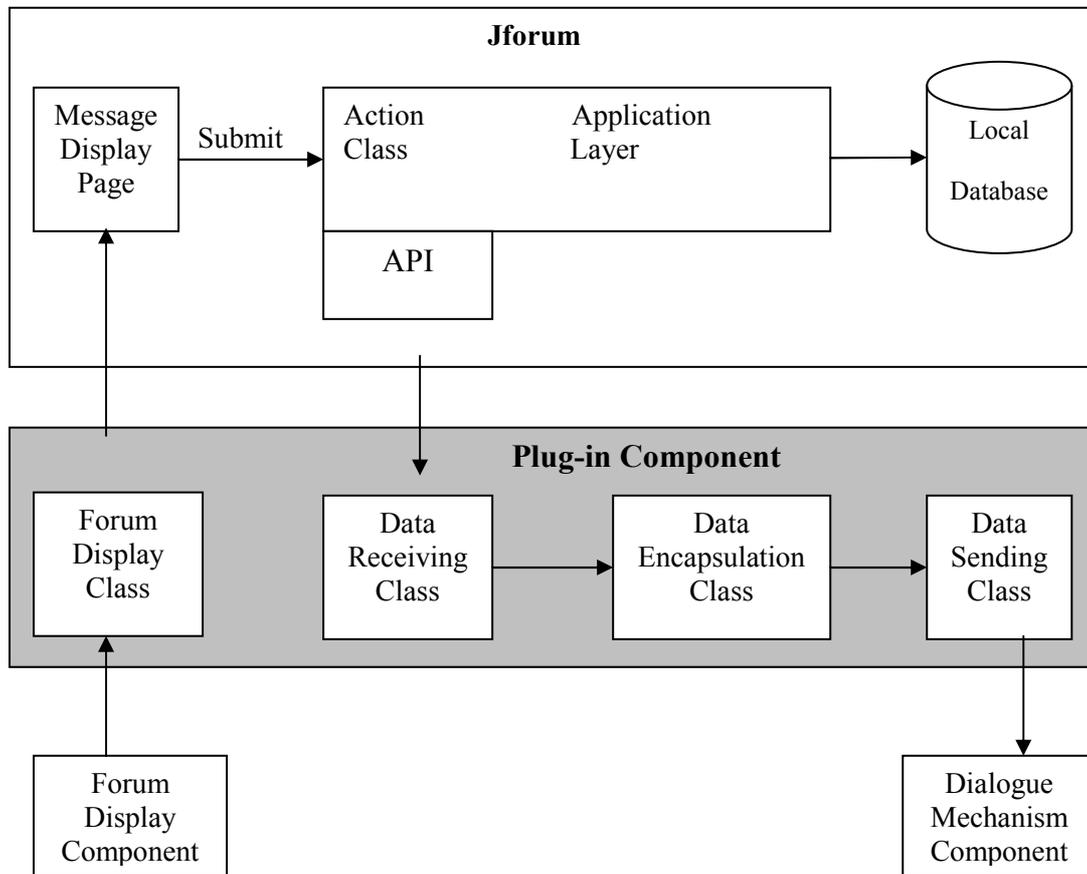


Figure 7-2: System Architecture for Plug-in Component

Figure 7-2 shows the system architecture for the plug-in component. This architecture diagram shows that the API exists where the action occurs in Jforum. Therefore when a message is submitted from the Jforum message display page, the action class of Jforum system sends the message to the data receiving class of the plug-in component via the Jforum API class. After the sent message is encapsulated by the data encapsulation class, the data is transferred to the Dialogue Mechanism component via the data sending class. When relevant information is retrieved from the Dialogue

Mechanism, it is returned via the forum display class to the Jforum message display page.

7.2.2 Dialogue Mechanism Component

The Dialogue Mechanism is the core component which is particularly designed to execute communication tasks between the users and system. This component consists of two agents: the Concept Formulation Agent and the Query Monitoring Agent. After data is transferred from the plug-in component to the Dialogue Mechanism, the text of the submitted messages is filtered and then sent to the Concept Formulation Agent. Both the Concept Formulation Agent and the Query Monitoring Agent need to communicate with each other and with the global, local, and remote databases via the dispatch component. The Concept Formulation Agent performs the semantic concept similarity calculation and recalculation processes to compare submitted messages with existing concepts in the database. The Query Monitoring Agent executes information retrieval tasks to search for relevant information in response to each submitted message and then returns it to the users. Figure 7-3 shows the high-level system architecture for the Dialogue Mechanism.

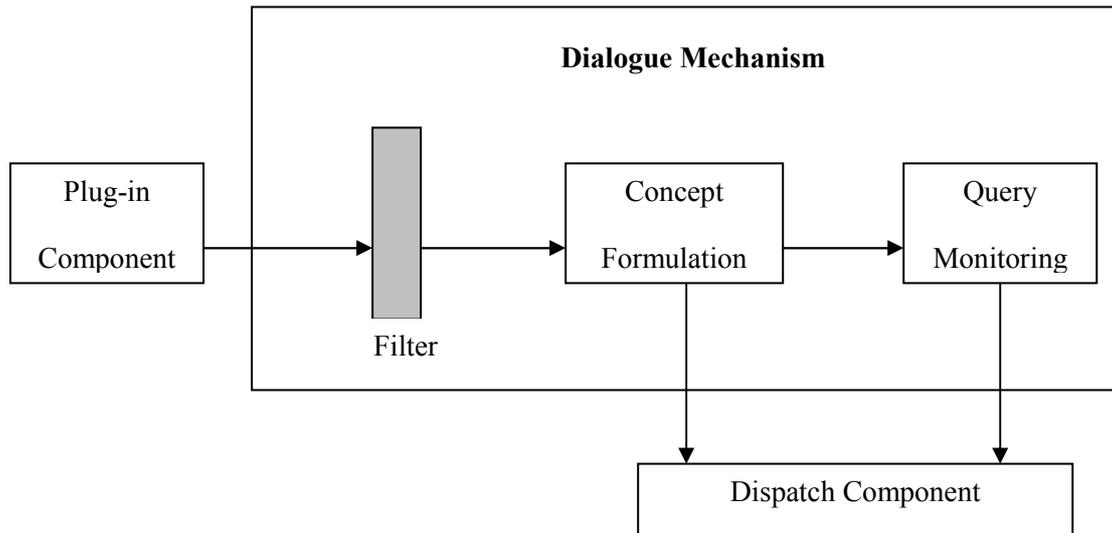


Figure 7-3: High-level System Architecture for Dialogue Mechanism

The system architecture details for the Concept Formulation Agent and the Query Monitoring Agent are introduced below.

Concept Formulation Agent

The Concept Formulation Agent is intended to act as an intelligent agent that understands the meanings of messages submitted by users and then classifies those messages into corresponding semantic concepts. The architecture for the Concept Formulation Agent is presented in Figure 7-4. This agent is used to handle two types of input. If the input is a submitted message, then the LSA Engine processes the message using Latent Semantic Analysis and the processed message is classified into an existing semantic concept according to calculations performed by the Concept Classification Class. If the feedback is submitted from the user, then it is sent to the Concept Reclassification Class where it undergoes the concept reclassification process. Both Concept Classification Class and Concept Reclassification Class are

connected to the Global Database via the dispatch component. The Global Database is used to store classified concepts and user feedback.

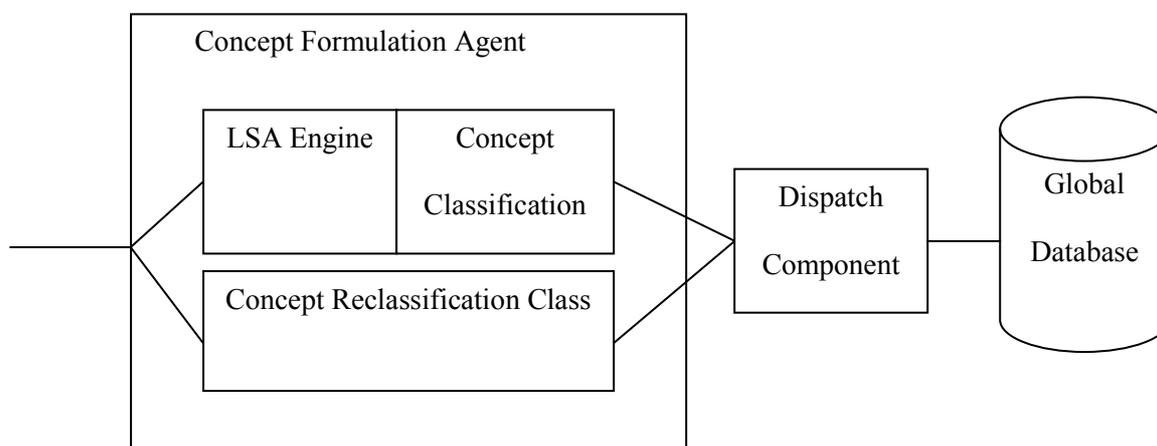


Figure 7-4: Architecture for Concept Formulation Agent

Query Monitoring Agent

The Query Monitoring Agent was developed to act as an intelligent agent that monitors each submitted query and performs the information retrieval process in the Dialogue Mechanism. Normally, after a submitted message is classified into an existing semantic concept by the Concept Classification Agent, the Query Monitoring Agent performs a relevant information retrieval task. The architecture for the Query Monitoring Agent is shown in Figure 7-5. When a message is sent to the Concept Formulation Agent, it is processed by the LSA Engine. Therefore, after a submitted message is classified into an existing semantic concept, the LSA processed message is sent to the Query Monitoring Agent and the Relevant Information Retrieval Class performs an information search. The Relevant Information Retrieval Class is connected to global, local and remote databases via the dispatch component. The global database stores concept classification and keyword information for each

existing posting and the local and remote databases store the original posting details for various courses in different institutions. However the remote part (the dotted line and box) were not implemented as discussed at the start of this chapter.

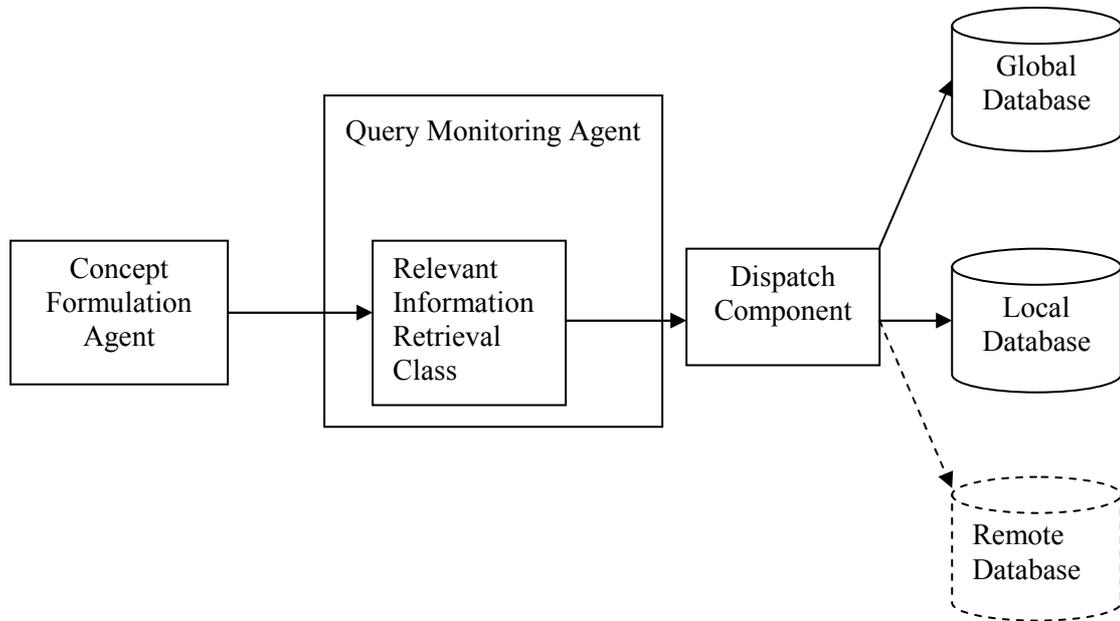


Figure 7-5: Architecture for Query Monitoring Agent

Text Similarity Calculation & Information Retrieval Process

The Dialogue Mechanism was designed and developed to perform the natural language process tasks of text similarity calculation and information retrieval. The methodologies for natural language processing in this component are outlined in Chapter 5.

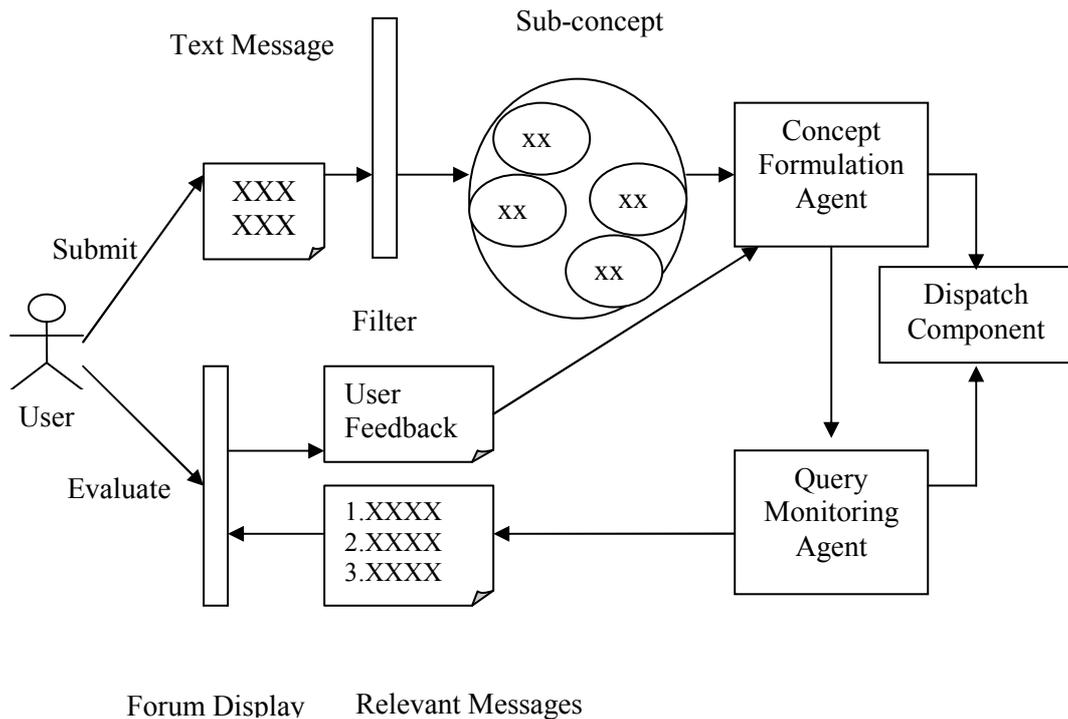


Figure 7-6: Work Flow Chart in Dialogue Mechanism

Figure 7-6 shows the work flow for the text similarity calculation (recalculation) and information retrieval processes in the Dialogue Mechanism. The text of a message is sent to the Dialogue Mechanism and this starts the process. The work flow for text similarity calculation and information retrieval processes is denoted in Figure 7-6. When a submitted message is sent to the Dialogue Mechanism, it is filtered first to remove all non-semantic words (the full list is presented in Appendix A). After the filtering process, only semantic keywords are left and they are used to identify a sub-concept which is sent to the Concept Formulation Agent. Here, the sub-concept is classified into an existing concept and then sent to the Query Monitoring Agent which performs a relevant information retrieval from local and remote databases via the Dispatch Component. A relevant information list is then sent back and displayed for users to view and evaluate through the Forum Display Component. If users submit

feedback, it is sent back to the Concept Formulation Agent which performs the concept reclassification process.

7.2.3 Dispatch Component

The Dispatch Component was developed to receive all calls from the Concept Classification and Query Monitoring Agents and to manipulate or retrieve data from global, local and remote databases. It uses Web Services to call the global, local, and any remote databases and, after gathering and sorting the information retrieval results, it returns one result set to the calling agent. The Data Service was designed and developed to expose the Data Layers of the local and remote systems as a Web Service that can be accessed by the Dispatch Component to retrieve data. It uses a programming method called “reflection” to translate incoming requests into objects that can be used with Hibernate as specified in the Datamodel package. The Datamodel package contains the Java Persistency Classes and XML descriptions used by Hibernate to map database tables to Java Classes. The Dispatch Component then uses Hibernate to access local and remote databases. Meanwhile the global database is available for direct connection to the prototype system via the Dispatch Component. The architecture for the Dispatch Component is shown in the figure 7-7.

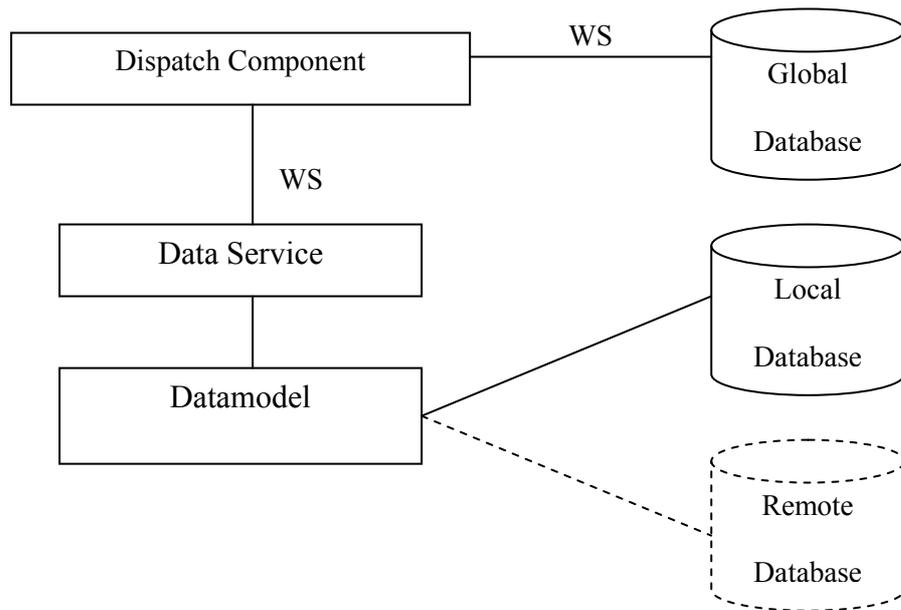


Figure 7-7: Architecture for Dispatch Component

7.2.4 Forum Display Component

The Forum Display Component is the interface that shows information to the users of the prototype system. This component requires functions to provide the general features of traditional discussion forums such as view, post and reply. According to the development plan used in this research, these functions were not developed from scratch but were reused from the customised and integrated Jforum system. However, the functions for news features were developed to display relevant retrieved results and provide an interface for users to evaluate those results. These new functions were then integrated with the general functions of Jforum. The system architecture for the Forum Display Component is shown in Figure 7-8.

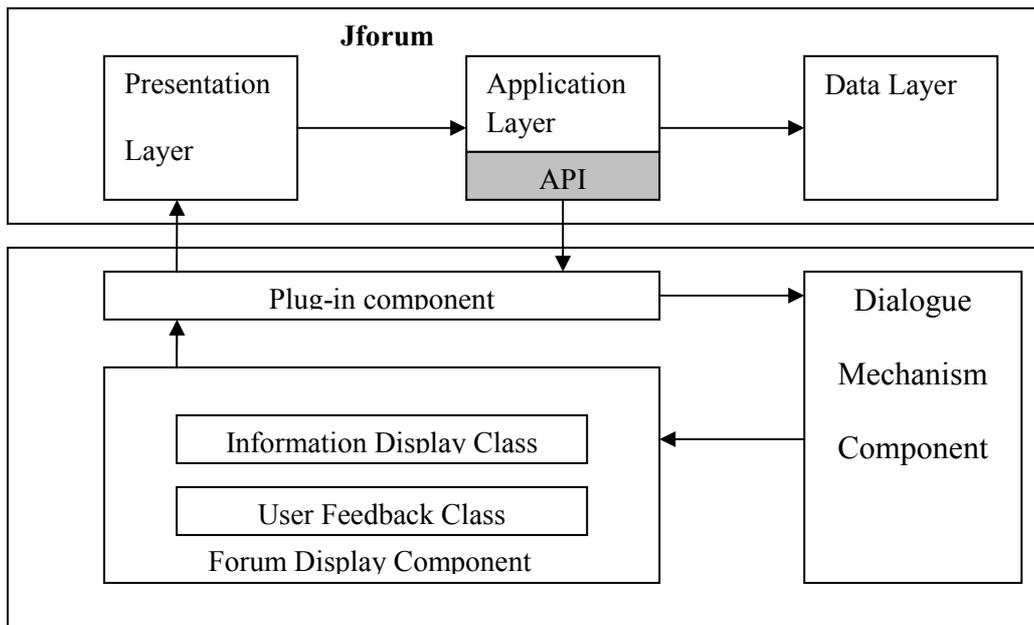


Figure 7-8: System Architecture for Forum Display Component

The Forum Display Component includes two function classes. One is the information display class which is used to generate the interface for displaying retrieved messages and the other is the user feedback class used to provide the interface framework for users to submit feedback on the relevance of the messages. Both classes were developed in the same way as the message display was developed for presentation layer of Jforum. This is because the Forum Display Component needs to be integrated with the existing functions of the presentation layer of Jforum via the plug-in component.

7.2.5 Data Architecture

The structure of information in the prototype system can be seen in Figure 7-9. This structure is quite similar to those commonly used in other web forums. A forum defines a wide subject area that can contain several topics. Each topic is a more narrowly defined area that can contain many messages or posts. Students may add

messages, while topics are set up by the teachers or tutors. A range of forums would ideally be set up centrally using an agreed ontology of subjects. This would encourage more institutions to participate in the discussion and the knowledge could be reused across different institutions.

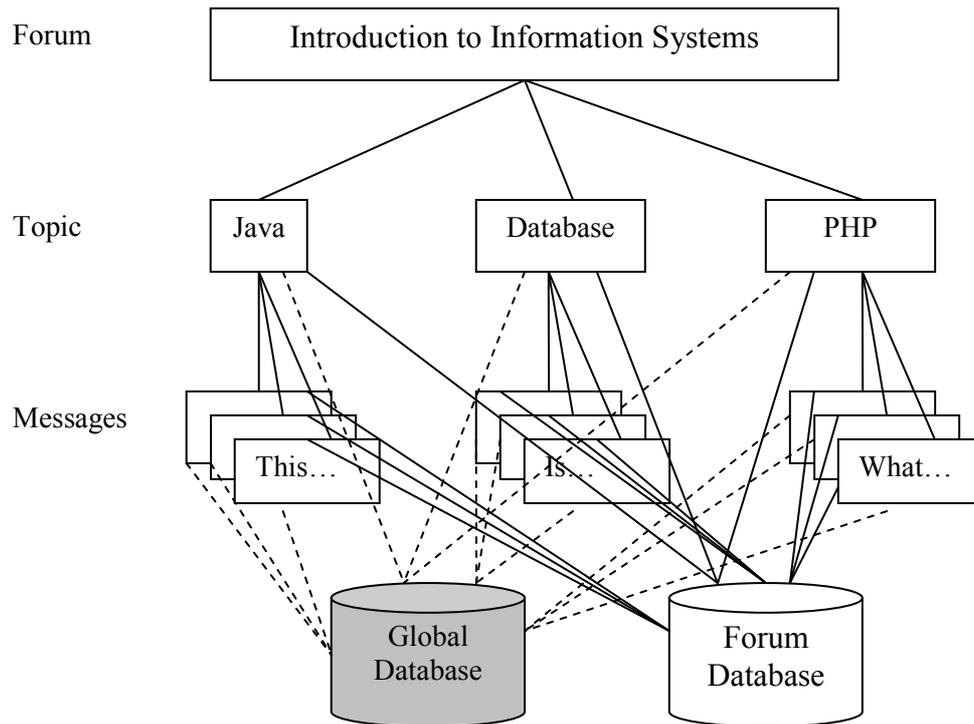


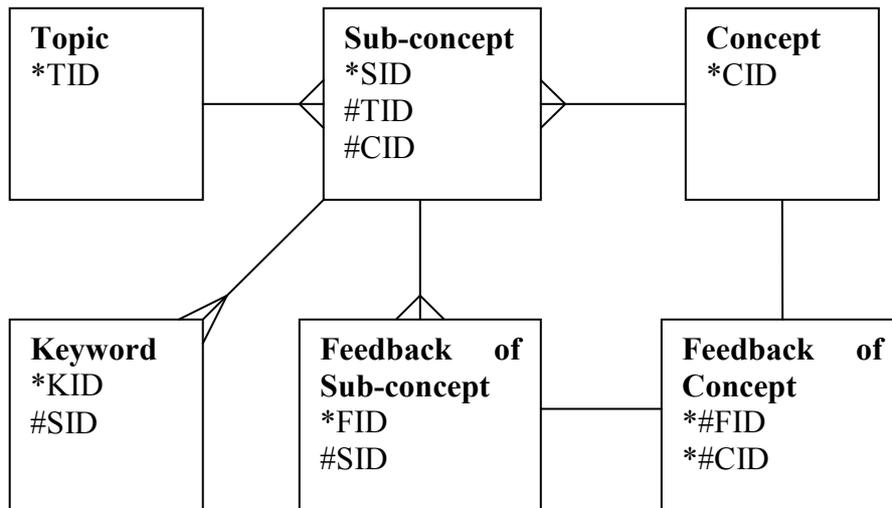
Figure 7-9: Information Structure and Storage Structure

The dotted lines and grey cylinders in Figure 7-9 show an example of one possible structure for data storage. The prototype system is significant in that all the data for natural language processing is stored in one central location. Other data is stored in the local system to which the user submits a message. For example, a message submitted by a student is stored in the local forum database or other remote databases, while the concept classification information for the message is stored in the global database. To generate the view of a forum the system therefore needs to access all connected servers and databases to retrieve all related information. The actual forum

definitions are here considered global in that they are replicated in all the databases. There is, in other words, currently no physically separate global database. The information structure of the prototype system is quite conventional for web forums, but storage of information in a distributed data structure is not common. This was implemented to allow each institution to be physically in control of their data.

Data Structure

The entity relationship (ER) model of the databases used in the prototype system is shown in Figure 7-10. This includes the data structure for the global database but not for the local database. The global database stores the duplicate information of topic and message text from the local database and the information for concept classification and user feedback. Keeping the local data protected from changes by users to different instances of the system works in most cases. One example of a task that would not be supported in this design is for a teacher in one course to perform actions on a message posted by a student in another course (on a different server). Teachers can choose to look after only their own students and then all work is done on the local server. This approach works well in that it does not generate any data consistency problems, since all data is stored in one location. It makes no difference to the teacher whether tasks are being performed on a locally stored message or a remotely stored one. This model also allows an information retrieval process to be performed across local and remote databases and even those located on servers of different institutions.



* Primary Key

Foreign Key

Figure 7-10: Entity Relationship Diagram

Data Access Layer

The fact that the data is distributed is hidden from the core system. This is achieved by all data access calls going through one component, called the Dispatcher. This component handles requests to retrieve data and to change data in two different ways. Requests to store, update or delete data are sent to the local instance (global database) of the Data Layer using its Web Service (which sits on the local server). Requests to retrieve data, on the other hand, are forwarded to the Web Services of all Data Layers that the prototype system is connected to (i.e. all institutions' databases) through Data Service and Datamodel and then the results received are combined before being returned to the calling component.

7.2.6 Deployment Architecture

Two approaches were identified and evaluated for deployment of the prototype system. The first architecture uses a central database to store global data in the

prototype system. The core components of the prototype system would run on a central server and the integrated discussion forum systems would run on each of the collaborating institution’s servers, communicating with the central server using Web Services, as seen in Figure 7-11. The components represented by the dotted line and boxes were not implemented in this research. Only the prototype system integrated with Jforum was implemented as one instance of the proposed new type of discussion forums intended to meet the research objectives.

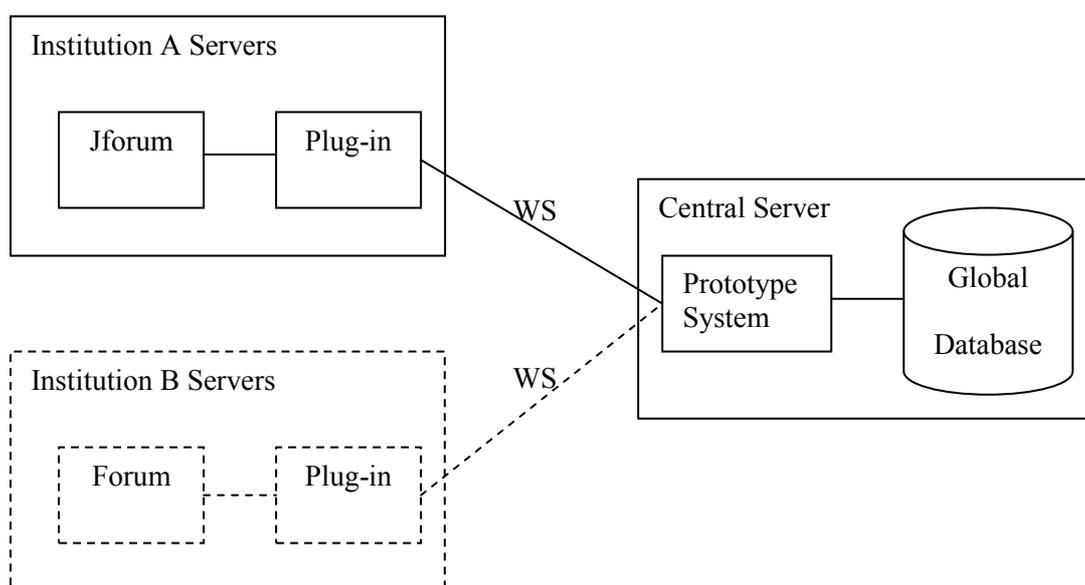


Figure 7-11: Deployment Architecture for Prototype System

The service-oriented approach where everything except the plug-in component runs centrally has the advantages that all data is stored in one place, and that most functionality is also located there. Maintenance of the system would therefore be easier, and scaling the system could be done by adding resources at one point. The negative sides to this architecture are that it can require large central resources to scale well and institutions may not be willing to store “their” data in a central location, especially if the central server is hosted at another institution. The second architecture

places everything within each institution. In either case, for performance and reliability, the global data structures are duplicated in each running database of the system, as they rarely change but are often queried.

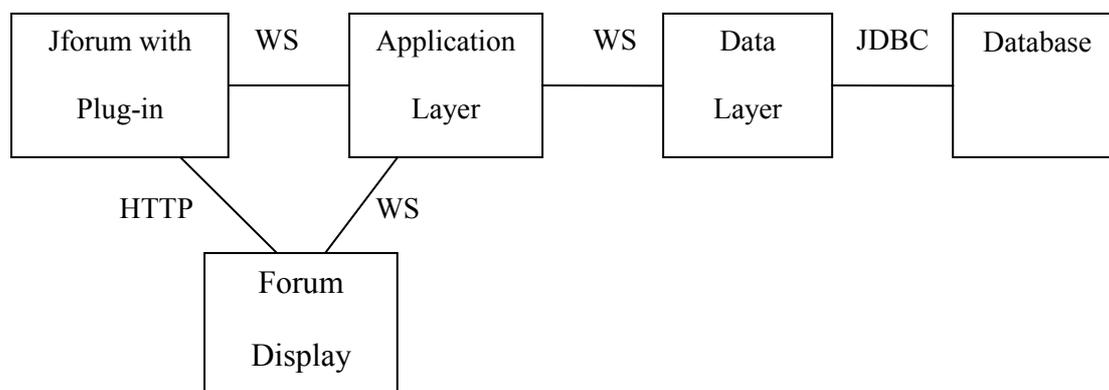


Figure 7-12: Layers of Prototype System

The developed prototype system is very flexible in how it can be installed on different servers. As seen in Figure 7-12, the system can be seen as a five layer architecture, where each layer, represented by a box, can be installed on a separate server if required. The lines between the boxes signify the communication channel used between the layers. In the case of Jforum, an HTTP GET request is sent from the forum display component to Jforum via the plug-in component to present the information to users. In this case, Jforum and the plug-in component have to be on the same server. Between the forum display component and the Application Layer, there is a Web Service connection and this is also how the Application Layer communicates with the Data Layer. The Application Layer communicates not only with the Data Layer within the local system where it resides, but also with the Data Layers of any other instances of the system that participate in the forum. The Data Layer then uses a JDBC connection to communicate with the actual Database. This gives a high degree

of freedom in the number of possible ways of configuring and installing the system on servers. This has the benefit that the system can be tailored to take advantage of available server resources. The next section outlines some of the specific function implementations for the prototype system.

7.3 Implementation

7.3.1 Overview

According to the research development plan, the prototype system implementation process includes core component development and integration with a Jforum system. The system functions for general features such as view, post and reply were not implemented from scratch because these general features of Jforum could be reused in the prototype system. The system functions for newly designed features of the prototype system such as information retrieval and user feedback were developed from scratch and integrated with the general functions of the Jforum system. The interfaces for general and new features of the prototype system are presented below.

7.3.2 General Features

All functions for general features were reused from the Jforum system. Jforum is a traditional discussion forum system which provides general discussion features including that:

- users can create a new topic in the forum;
- users can view existing topics and messages in the forum;
- users can post a reply to a topic in the forum;

- administrators can manage the forum, such as by deleting an existing message.

The prototype system uses all functions for the above general features. The screenshots for the interfaces of the prototype system show how a user creates a new topic (Figure 7-13), views all existing topics and messages (Figure 7-14 & 7-15), and posts a reply (Figure 7-16) in the prototype system.

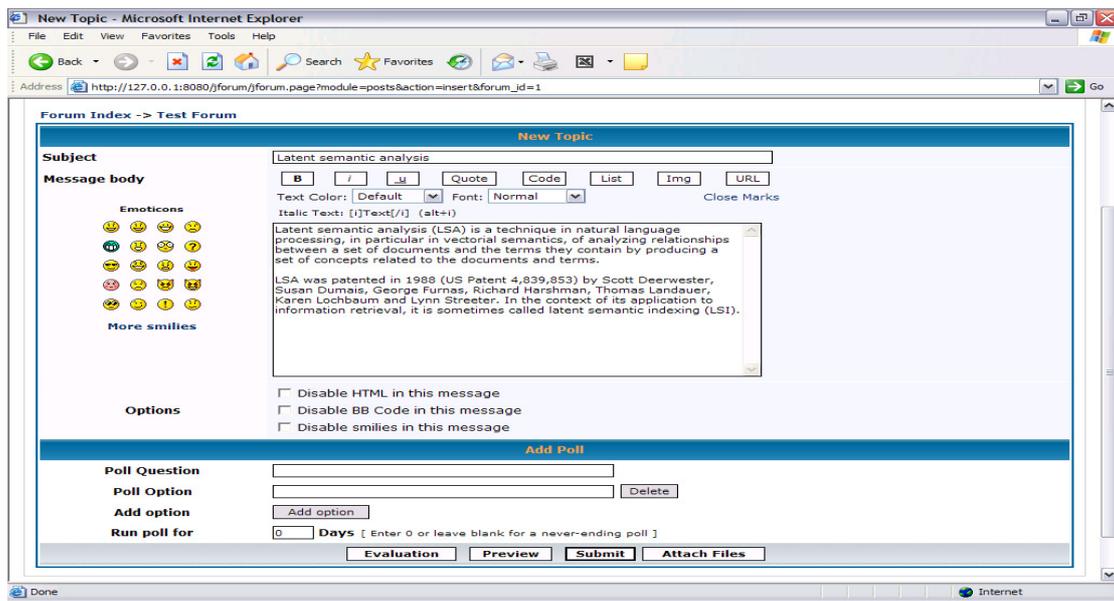


Figure 7-13: Topic Creation in Prototype Discussion Forum

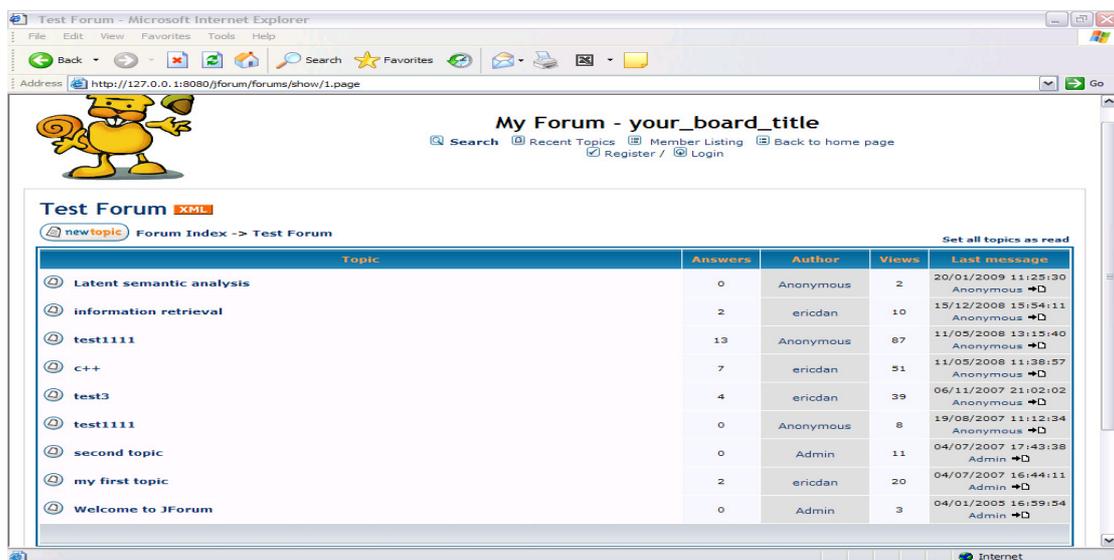


Figure 7-14: Viewing All Topics in Prototype Discussion Forum

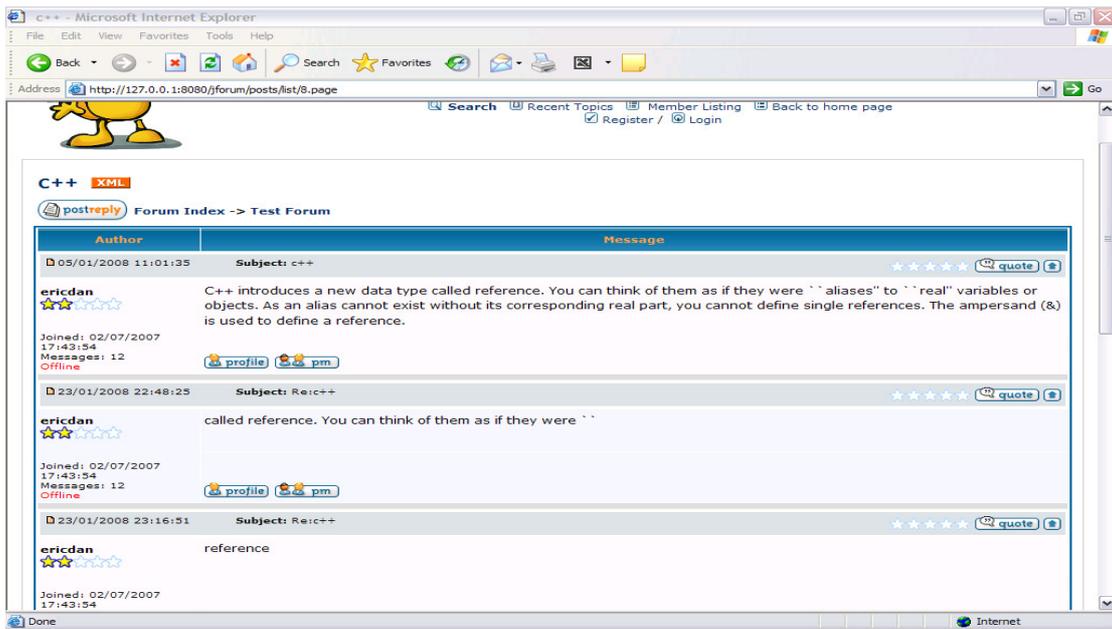


Figure 7-15: View Messages in Prototype Discussion Forum

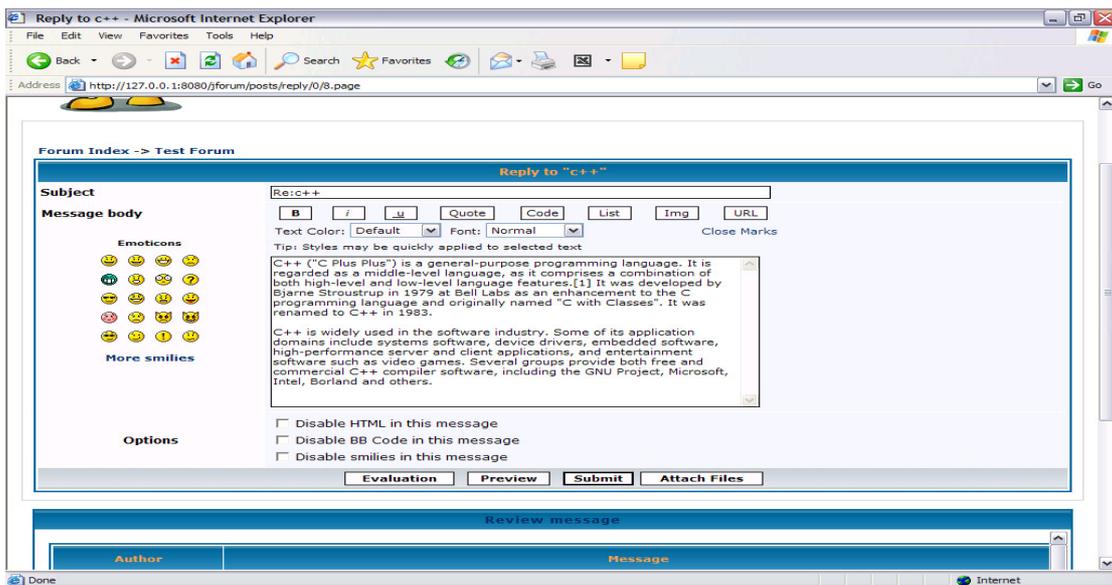


Figure 7-16: Post a Reply in Prototype Discussion Forum

7.3.3 New Features

Functions for new features in the prototype system were implemented from scratch and integrated with general functions of Jforum. These new features include the relevant information retrieval function and the user feedback function. In the

prototype discussion forum system, when a user submits a new message, the system automatically finds the most relevant existing messages from databases and then returns them to the user. Meanwhile, when users browse those relevant messages, they can evaluate which ones relate to their submitted message.

In the Jforum system, there is a message previewing function which allows users to review their new messages before the messages are submitted into the discussion forum. This is an appropriate function to integrate the new functions with because it allows users to submit their messages directly without a relevant information search or they can view relevant information via the preview function. The architecture for the preview function in Jforum is shown in Figure 7-17.

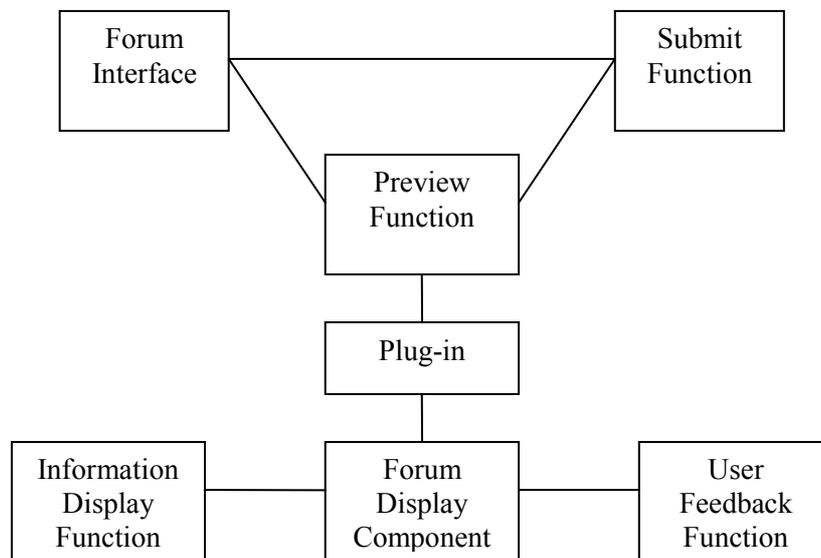


Figure 7-17: Architecture for Integration with Preview Function

New features of the prototype system are integrated with the Jforum preview function and the interfaces for the relevant information display and user feedback functions are presented in Figures 7-18 and 7-19. After a user completes a new message on the

‘reply’ page (Figure 7-18), he/she can click the ‘preview’ button and previously relevant messages are then shown on the ‘preview’ page (Figure 7-19).

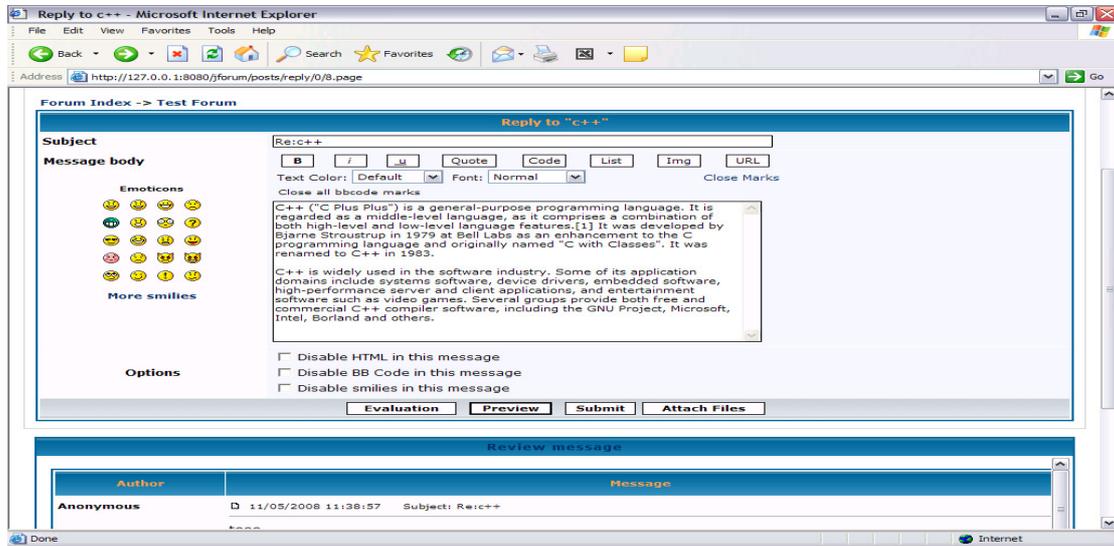


Figure 7-18: Message Reply Page

On the ‘preview’ page, relevant messages are retrieved from different topics and sorted by similarity weight with the submitted message.

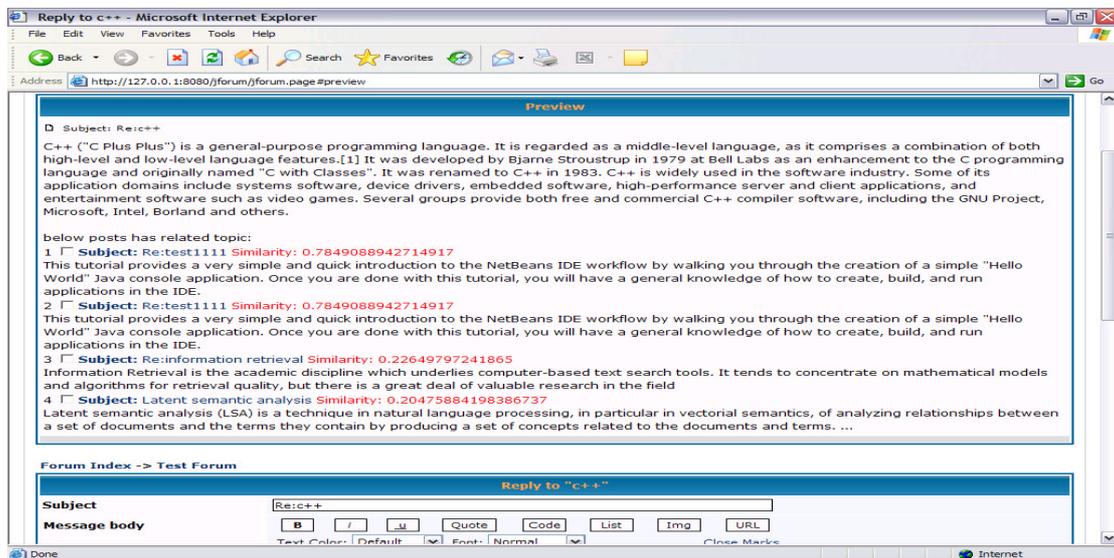


Figure 7-19: Message Preview Page

When relevant messages are shown on the ‘preview’ page, there is a checkbox for users to evaluate each returned message. Therefore users can select all messages which they think are related to the submitted message and then submit their feedback by clicking the ‘evaluation’ button (Figures 7-20 and 7-21). After users view and evaluate relevant messages, they can choose to submit their new message, or cancel it in case their query has already been answered.

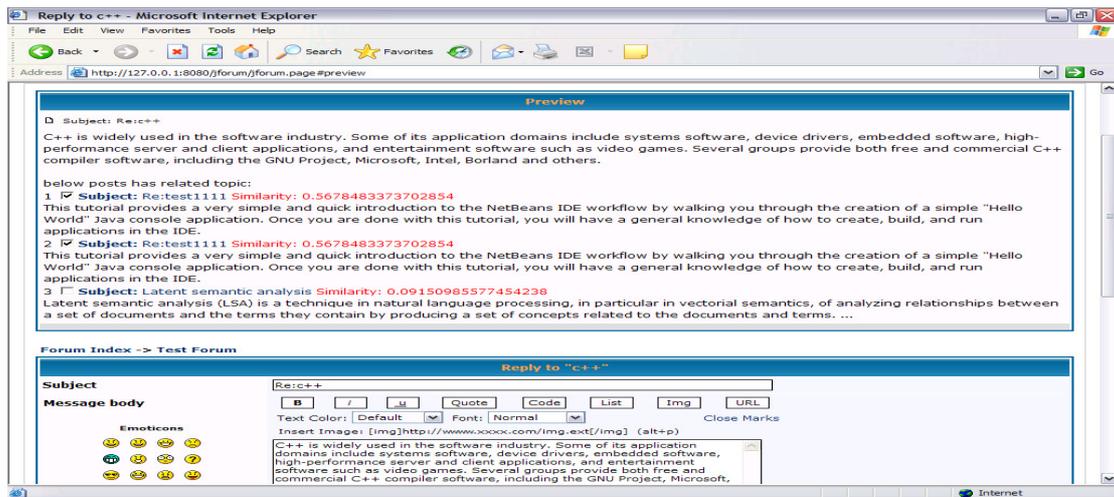


Figure 7-20: User Evaluation for Retrieved Relevant Messages

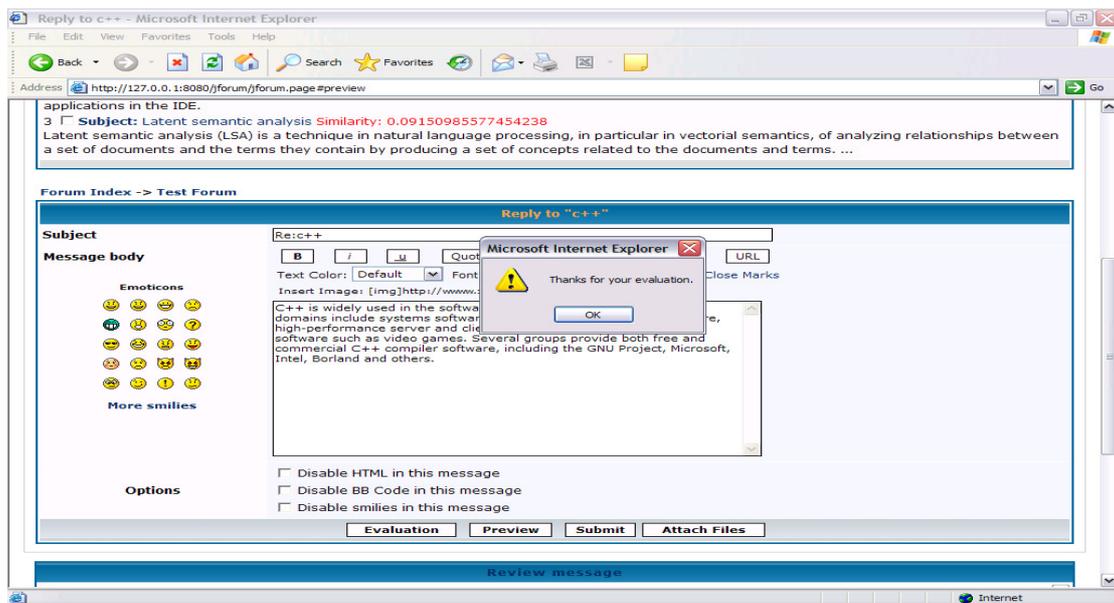


Figure 7-21: User Feedback Submitting

7.4 Summary

This chapter first introduced the architecture for the newly designed discussion forum system. In the system development process, a Jforum system is customised and integrated with the core components of the prototype system to meet the proposed system design requirements and achieve the research goals. The interfaces for general and new features of the prototype system were presented in the later part of this chapter. The general features include the traditional functions of discussion forum such as view, post and reply. The new features developed from scratch include relevant information search and user evaluation functions. After the prototype system was successfully implemented, it was evaluated from the perspectives of system performance and users to establish whether the research problems were solved and research goals achieved. The next three chapters outline the research background, experiment design and results analysis for the evaluation.

CHAPTER 8

RELATED EVALUATION APPROACHES

8.1 Introduction

A new type of discussion forum system has been proposed and designed in this research to overcome the main limitations of traditional discussion forum systems. A prototype of such system was then successfully implemented. How well the prototype system met the system requirements and achieved the research goals was then evaluated. Before presenting the evaluation of the prototype system, this chapter first provides the research background to traditional evaluation methods and approaches. The following sections introduce traditional evaluation methodologies for Information Retrieval (IR) and present evaluation approaches used by other related research.

8.2 Evaluation Methodologies

Evaluation is defined as a method for assessing performance or value of a system, process, product, or policy. As such, evaluation is a critical necessity in science and technology (Saracevic, 1995). The prototype system was designed and implemented to provide relevant information retrieval and user feedback functions which improve the effectiveness of communication between users in the discussion forum systems. Therefore to evaluate the prototype system, evaluation methodologies for both information retrieval performance and user performance and satisfaction are taken into account in this research.

8.2.1 Evaluation for IR Performance

Traditionally the effectiveness of information retrieval systems is measured by comparing information retrieval performance on a common set of queries and documents (Sanderson & Zobel, 2005). The evaluation criteria for IR systems include retrieval accuracy, cost-effectiveness, processing time, and effort required to learn the system. Evaluation methods are commonly based on classical recall and precision measures which are system-oriented. These methods evaluate average system accuracy but do not identify the influences of other factors such as retrieval task and system user (Croft, 1995).

An ideal IR system allows all relevant documents in a collection to pass through and no others. A real system is never as effective as an ideal one, but it meets the evaluation goals to a large degree. Retrieval accuracy can be defined as the effectiveness with respect to the goals (Harter, 1986). Kent et al. (1955) defined the measures of precision and relevance (later renamed recall) for evaluation of IR systems. Precision is the ratio of the number of relevant documents retrieved to all retrieved documents. Recall is the ratio of the number of relevant documents retrieved to the number of all relevant documents in the search space. Precision and recall are the most commonly used measures of effectiveness for the evaluation of IR systems (Billhardt, 2002). The best known example is the Text Retrieval Conference (TREC) evaluation methodology, which includes several metrics to evaluate different aspects of an IR system (Robertson, 2002).

TREC metrics are based on the measures of precision and recall and are regarded as representing the state-of-the-art of ad hoc evaluation. The TREC metrics include:

- Interpolated Recall-Precision Averages: This is based on the calculated precision at recalls of 11 equal intervals from 0.00 to 1.00. The calculated values are averaged over all queries (for each of the 11 recall levels).
- Average Precision over All Relevant Documents: This is based on the calculated precision after each relevant document is retrieved. All precision values are then averaged together to get a single number for the performance of a query.
- Precision at k Documents: This is based on the calculated precision after k documents have been retrieved. The calculated values are averaged over all queries.
- R-Precision: This is based on the calculated precision after R (number of relevant documents for a query) documents have been retrieved. The calculated values are also averaged over all queries.

Another well known metric is “Effectiveness” (E) proposed by Van Rijsbergen (1979):

$$E = 1 - \frac{(1 + \beta^2)PR}{\beta^2P + R}$$

where P and R are precision and recall, and β is a parameter reflecting the relative importance of recall to precision defined by the user. Effectiveness not only takes into account the users’ preferences but also integrates precision and recall in a single measure. Precision-oriented users are interested in how many irrelevant documents they have to view for a given retrieval set size. In contrast, recall-oriented users are interested in how large a retrieval set size they must view in order to obtain a certain number of relevant documents (Frei & Wyle, 1991).

Other metrics include Normalised Recall, Coverage Ratio, Novelty Ratio, Relative Recall and Recall Effort (Korfhage, 1997). These criteria, like many others, however, are not widely accepted and used.

Different approaches can be used to evaluate how good a system is or whether one is better than another (Beaulieu, 1997). These traditional metrics measure an IR system from different perspectives (Greisdorf, 2000). The choice of approach depends on the intent of evaluation (Saracevic, 1995). In the practical evaluation of IR systems, Interpolated Recall-Precision Averages are used ubiquitously (Frei & Wyle, 1991) and TREC is probably the greatest single source of information about IR evaluation methods and metrics (Jackson, 2002).

8.2.2 Evaluation for User Performance and Satisfaction

Improvements in IR effectiveness measures (recall and precision) do not necessarily reflect an improvement in user performance (Al-Maskari, 2007). Many researchers have recommended considering perceptions of the users to be as important as IR effectiveness measures, and both should be interpreted as measures of effectiveness. Therefore IR evaluation should not focus solely on maximising retrieval performance by refining IR techniques. It should also include understanding user satisfaction, behaviours and information needs.

Measures of user performance and satisfaction are considered as other aspects of the criteria of a system (Turpin & Scholer, 2006; Belkin et al., 2004). User performance and satisfaction are important to indicate how well a system meets users' information

needs. User performance normally includes measures of user effort and time spent. Normally the user effort is measured by the number of documents visited by the user to complete a search task. The time spent is the time taken by a user on a session for meeting an information need. User satisfaction is measured after completing tasks using an information retrieval system by asking the user to indicate the degree of satisfaction to which the information need is met (Chanana et al., 2004). Since the 1990's researchers have developed various methods for capturing user satisfaction with the retrieved results of IR systems and have examined how users assess their own performance in task completion. Furthermore some interactive evaluation frameworks have been customised from different approaches to bring together aspects of system effectiveness and user performance and satisfaction (AI-Maskari, 2007). The next sections introduce some related approaches to evaluating the effectiveness of IR or similar systems.

8.3 Related Evaluation Approaches

8.3.1 Overview

Different approaches can be used to evaluate or compare systems. The choice of approach depends on the research goals for the evaluation (Greisdorf, 2000). In the practical evaluation of IR or similar systems, single or combined traditional measures may be adopted. These metrics measure an IR or similar system from different perspectives. They evaluate the general performance of a system with average measures and can provide detailed information about the effects of different factors that determine retrieval results. In this way they contribute information to help make

improvements to IR or similar systems and processes. The next sections introduce and present some recent evaluation examples.

8.3.2 User Performance versus Precision Measures for Simple Search Tasks

Turpin and Scholer (2006) concluded that traditional evaluation approaches for IR cannot prove the benefits for users. They used two evaluation approaches to evaluate two different information retrieval systems for user performance. The approaches were based on two simple tasks. The first was a precision-oriented task, requiring users to find one document that was relevant to the supplied query. The second was a recall-based task, measured by the number of relevant documents that users could identify in a five minute time period. Data for the time taken to find the first relevant document and the number of relevant documents that could be identified in five minutes were collected and analysed after the evaluation process was completed. The evaluation results showed no significant relationship between system effectiveness measured by mean average precision and the precision-based task.

8.3.3 User-centred Evaluation of an E-learning Repository

Venturi and Bessis (2006) presented a user-centred approach to evaluate the DELTA system which is a distributed repository for learning resources. To identify the evaluation goals, they employed scenarios and claims analysis techniques. For instance, according to designed scenarios, they made a number of claims including:

- Simple and advanced search functionalities will be easy to use for people new to DELTA; and

- Search results will be easy to interpret and filter by advanced search for novices.

In their research, they set up an “evaluation toolkit” and six methods were selected comprising quantitative and qualitative research techniques as follows:

- Web questionnaire (quantitative and qualitative);
- Usability tests (quantitative and qualitative);
- Logs analysis (quantitative);
- Interviews (qualitative);
- User diaries (qualitative);
- Pedagogy workshop (qualitative).

The results showed that DELTA fulfilled four out of the nine evaluation goals.

8.3.4 Evaluating Effectiveness of a Context-based IR System

Chanana et al. (2005) proposed a new context-based information retrieval system to compete with the traditional keyword-based systems. To evaluate the proposed system, they designed an evaluation to assess whether the proposed system had better retrieval effectiveness than the baseline system. The evaluation measures used included:

- difference in performance between the two systems in terms of precision ratio, search effort and time spent; and
- difference in user satisfaction.

Ten students participated in the setup evaluation experiment to use both systems to complete a set of ten predetermined information search tasks. All data concerning the experiment was logged on a local server. The information recorded included documents read, the sequence in which documents were visited, the relevance score for each document, time spent for each session and the user's score for how well the information need was met by the documents retrieved. Statistical analysis showed that compared with baseline system, the precision ratio and user satisfaction of the proposed system were significantly increased, but user effort and time spent were decreased. It concluded that for IR effectiveness, the proposed system performed better than the baseline system.

8.3.5 Evaluating an Intelligent Discussion-Bot

In Feng et al.'s (2006) research, an intelligent discussion-bot was developed to automatically answer student queries. To evaluate this intelligent discussion-bot, the researchers set up an experiment in which 279 documents and 3093 posts were placed in the system. A total of 66 questions were written into a script to simulate a user posting new questions on the discussion forum and the related information and answers were returned. They classified the 66 replies as exact answer, good answer, related answer or unrelated answer. The results showed that good system performance depended highly on the system threshold value of the cosine similarity score for the intelligent discussion-bot system.

8.3.6 Assessing Distance Learning Student Performance

Wu and Chen (2004) proposed a model which combined three measures: keyword density, message length and message count in the discussion forum, to automatically

assess student performance. To validate this model, they set up an experiment which included two online courses. In the first course, the instructors gave out eight discussion topics which covered major issues discussed in the class and for each topic. Students were required to respond with an original post and reply to others before the due date. The second course had only one discussion conference, and there were no specific discussion topics. Students were encouraged to find interesting resources to share with the class. The experiment showed that automated performance assessment worked surprisingly well compared with human judges. The correlations between the score based on the proposed measures and the actual grades were generally high.

8.4 Summary

This chapter introduced traditional evaluation methodologies for IR and similar systems including system performance, and user performance and satisfaction. The latter sections presented different evaluation approaches adopted in recent research works. Normally an evaluation experiment is set up to test system usability, information retrieval performance, user performance, etc. Researchers analyse experimental results and compare systems to prove research goals such as whether one system is more effective than another.

In this research, the evaluation uses some of the above methods to evaluate how well the prototype system meets the proposed system requirements and achieves the research goals. This evaluation is introduced in the next chapter.

CHAPTER 9

EVALUATION APPROACH

9.1 Introduction

The previous chapter reviewed traditional and recent evaluation literature for information retrieval related systems. This chapter outlines the evaluation approach chosen in this project and the consequent evaluation experiments designed to determine how well the prototype system met the system design requirements and achieved research objectives. The first section summarises the research objectives and evaluation goals. The next section introduces the evaluation approaches proposed to assess both the objective and subjective performance of the prototype system. The details of the evaluation experiments are presented in the last section.

9.2 Evaluation Goals

The evaluation process for this research project started after the prototype system development phase was completed. The implemented prototype system has been assessed as to whether it achieved the research objectives and this is the core component of the entire evaluation process. The evaluation goals are based on the research objectives and the proposed system design requirements. This section recaps those research objectives and then establishes the evaluation goals used in the subsequent sections.

9.2.1 Research Objectives

The research objectives provided the foundation for the initial phase of the project. This research aims to overcome major communication limitations of traditional discussion forum systems by improving the effectiveness of dialogue between users. One important requirement of such a system is that the prototype forum should enable sharing and reuse of knowledge for various courses in different institutions. Such a feature of the prototype forum has been implemented and evaluated in the previous eQuake project (Smestad, 2006). Hence this research objective has not been considered in the evaluation section of the current research study.

The research objectives are listed below.

- The prototype forum should provide immediate responses for users' messages.
- The prototype forum should automatically search and provide relevant information when users try to post a message.
- Users should be able to give feedback on the relevancy of the search results in the prototype forum.
- The prototype forum should be able to reduce the teachers' time consuming burden of managing the discussion forum and answering duplicated questions.

9.2.2 Evaluation Goals

Building on previous solutions, the proposed discussion forum system was designed to meet the above research objectives. The prototype system was developed as an instance of the proposed system to meet the proposed system design requirements and achieve better system performance than those systems proposed in other research works. Therefore the overall evaluation goal is that **the developed system should**

enhance effective communication among learners and between learners and instructors in the discussion forum.

To establish whether the prototype discussion forum system has achieved the evaluation goal, the usability and performance of the system need to be assessed from two perspectives:

- For achieving effective communication among learners, learners should be able to easily find relevant and useful previously posted information in a timely manner;
- For achieving effective communication between learners and instructors, instructors should be able to reduce their efforts on discussion forum management (such as answering duplicated questions). Meanwhile learners should be able to find the related answers or knowledge they need by themselves in the prototype discussion forum so that they do not need to submit similar questions multiple times.

The evaluation approach was designed to assess the usability and performance of the prototype system developed in the research project. The evaluation of the prototype system focused on system use, such as how well the prototype system enables a user to find useful information thereby meeting an information need. If learners can easily find relevant needed information, they do not need to submit duplicate questions into the discussion forum so that instructors do not need to waste time and effort on answering similar questions and fewer similar questions reduces the number of postings which reinforces the ability of learners to easily find useful information.

The next section introduces the evaluation approach used to meet the evaluation goals summarised above.

9.3 Evaluation Approaches

9.3.1 Baseline Systems

According to the evaluation literature review, popular approaches for evaluating information retrieval related systems include objective system performance which is measured in terms of information retrieval precision ratio, time spent, user effort, and search effectiveness, and subjective system performance (Robertson, 2002; Chanana et al., 2004; Belkin et al., 2004; Sanderson & Zobel, 2005; Turpin & Scholer, 2006). Precision ratio is calculated by the number of documents judged relevant by a user, divided by the number of documents retrieved. User effort is measured as the number of documents visited by the user to complete a search session. Time spent is the time taken by a user on a session to meet an information need. Search effectiveness is calculated as the number of related information sources found within a given time. Subjective system performance is based on users' stated satisfaction with information retrieval related systems. For instance, the user can be asked to indicate the degree of satisfaction to which the information need is met.

An objective evaluation of the communication aspects of the system's performance requires comparison with baseline systems (Chanana et al., 2004). The criteria for selecting the baseline system for this evaluation are as follows.

- The baseline system should have similar discussion forum functions to those of the prototype system (Chanana et al., 2004).

- The baseline system should have the same auto-reply feature as the prototype system because this feature is the core research contribution to improve the effective communication in discussion forums.
- The IR function of the baseline discussion forum system should adopt Information Retrieval (IR) and Natural Language Processing (NLP) techniques – such as the Boolean keyword method and the TF*IDF Method and also the IR technique should be already evaluated as an effective way to improve effectiveness of communication in a discussion forum (Feng et al., 2006);

This research is an extension of the eQuake research project (Smestad, 2006). The Boolean keyword search method was the major information retrieval approach used in the eQuake prototype discussion forum. This research seeks to improve information retrieval functionality in order to provide more effective communication performance than that achieved by the eQuake system. Therefore the eQuake forum was the first baseline system used in the evaluation experiments.

The TF*IDF approach has been adopted in some recently designed systems as a new and popular approach to improve information retrieval performance in discussion forums such as intelligent discussion-bot system (Feng et al., 2006) and in student performance assessment systems (Wu & Chen, 2004). The TF*IDF method has already been evaluated as an effective method to improve the effectiveness of communication in discussion forums. Therefore as developed by using the TF*IDF information retrieval approach, the intelligent discussion-bot system was selected as

the second baseline discussion forum system and its mock-up system was installed for use in the evaluation process.

In the evaluation process, the developed prototype forum system was compared with the eQuake forum system and the intelligent-bot discussion forum system.

The measures selected to assess objective system performance for this evaluation were: the information retrieval precision ratio, user effort, time spent and search performance. The precision ratio measure was used to compare the prototype system with the baseline systems, and another measure, search performance (effectiveness), was used to assess system performance. The measures time spent and search effort were not selected for use in this study because it would have been difficult to record the time spent and user effort (how many links visited) in the evaluation process for as many as 30 participants who would be involved. In many previous studies, it has been shown that even without those two measures, the evaluation measures precision ratio and search performance were still able to provide a valid evaluation result (Feng et al.'s, 2006; Turpin & Scholer, 2006). In Feng et al.'s (2006) research, only the evaluation measure precision ratio was adopted to evaluate that their proposed system. Also Turpin and Scholer (2006) assessed their developed IR system using two evaluation approaches to that were based on two simple tasks. The first was a precision-oriented task, requiring users to find one document that was relevant to the supplied query. The second was a recall-based task measured by the number of relevant documents that users could identify in a five minute time period. Therefore, the selected evaluation measures in this study, precision ratio and search performance, were able to provide a valid evaluation result despite not using the two measures, time

spent and user effort. The details of the adopted evaluation approach are introduced in the next section.

9.3.2 Evaluation Approach

The innovative features of the prototype system are the relevant message retrieval function and the user feedback function. The user feedback function was also designed to enhance the accuracy of information retrieval. Therefore, the core evaluation aspect was to assess the system's performance for the information retrieval process. Furthermore, to evaluate the effectiveness of the prototype system, one needs to evaluate the system performance of the IR process in the prototype discussion forum.

The two measures for evaluating objective system performance for the prototype were finalized with information retrieval precision ratio and search performance. For subjective system evaluation, the measure was user satisfaction. The evaluation primarily involved learner users, but one evaluation in the user satisfaction with just instructors was performed.

Based on the evaluation goal in the previous sections the final evaluation approach was selected to combine measures from information retrieval evaluation methodology and dialogue (user interactive) systems evaluation methodology to evaluate how the proposed solution can effectively solve the research problems and answer the research questions. It was also considering with the convenience to set up experiments, recruit participants and record experiment data for the evaluation process. Two evaluation experiments were designed to compare information retrieval performance and

usability between the prototype system and the baseline systems with respect to the evaluation goals.

The first evaluation was set up to evaluate the quality of the auto-reply function of the prototype system which can synchronously find most relevant posts when users post a message. In order to evaluate the performance of the auto-reply function, the first experiment was designed to assess the information retrieval precision ratio against both the prototype system and baseline systems. Therefore the retrieval information relatedness was measured and the two systems compared. This was designed as an automated experiment to test the system's IR performance. A test script automatically triggered the designed questions to be submitted to both the prototype system and baseline systems and then all automatically retrieved information was recorded in a log.

The second evaluation experiment was set up as a task completion process for users to interact with the prototype and baseline systems. During this experiment, the objective system evaluation criterion for search effectiveness was assessed. The main focus of the second experiment was to assess how well the prototype system enabled a user to effectively find useful information that he/she needs in learning communities. After the task was completed, learners and instructors completed a questionnaire designed to assess their satisfaction. The user opinions identified from the questionnaire were then used to evaluate the subjective performance of the prototype system. Due to the time and resource limitations, the prototype system was not able to be used for a real course to complete the user performance evaluation experiment.

Therefore based on an assumed learning scenario, the second experiment was designed as a simulated task completion experiment.

Data from the above experiments were then analysed to determine whether evaluation goals were successfully achieved the following objectives:

- Compared with baseline systems, was the performance of information retrieval in the prototype discussion forum improved?
- Compared with baseline systems, was user capture of relevant information and useful knowledge improved?
- Compared with baseline systems, were users more or less satisfied with the new features and usability of the prototype discussion forum?

The evaluation experiment designs and details are presented in the next section.

9.4 Evaluation Experiments

Two evaluation experiments were designed to assess the objective and subjective performance of the prototype system. The first was an automated testing experiment used to collect the precision ratio data for the auto-reply function in the prototype system and baseline systems. The second was an interactive experiment where learners used all the prototype and baseline systems. A simulated task completion approach and a designed questionnaire were adopted in the second experiment.

9.4.1 The First Experiment

In the first experiment, the aim was to simulate a real course-used discussion forum in which all users were able to seek useful information or submit a post so that for both the prototype forum and baseline forums, their information retrieval performance could be assessed. To evaluate IR performance, a document collection or test collection was required as the information retrieval base in the objective system performance evaluation process. Additionally, in order to record if retrieved information was related or not with a submitted query in a system, the relevancy of all information in the test collection with the submitted query was marked before the experiment was started. This evaluation experiment used standard test collections that assumed that relevance could be approximated by topical similarity of the query to the document so the user was not required to make a relevance assessment about retrieved documents. The relevance is binary – relevant or non-relevant – and recall for a query post was always known. A time constraint and selection of documents from specialised databases restricted the domain to 950 articles published in ACM (Association of Computing Machinery) and IEEE (Institute of Electrical & Electronics Engineers) from 1980 to 2008.

All articles were collected from six topic groups and the content of each article was uploaded as a post into the corresponding thread folders of the prototype discussion forum by the researchers. This is because the post was the basic information representation in the discussion forum systems. The data for distribution of the length of each topic folder was recorded and grouped for further statistical analysis on the relationship between IR performance and the length of posts. This evaluation experiment was designed to simulate an IT course, therefore in March of 2009, a

group of 5 undergraduate and 5 postgraduate students were recruited from the computer science and information systems departments at Massey University as the learner and instructor users for both the prototype discussion forum system and the baseline forums. The background information of all participants is presented in the next chapter.

All participants had experiences with using discussion forums and they were requested to post questions or comments in each thread of the discussion forum regarding the corresponding topic and also submitted their feedback on retrieved postings within 30 minutes. After this section was completed, the prototype system's database was ready for the evaluation experiment.

To evaluate the performance of the prototype and baseline systems, the average precision ratio was measured for the list of posts returned to the user. The precision ratio data for the prototype discussion forum system and the baseline systems was collected and analysed to determine whether the evaluation objectives were achieved. The formula to calculate precision ratio is:

$$\text{Precision} = (\text{number of relevant documents}) / (\text{number of retrieved documents})$$

A script was developed to simulate a user posting news questions and comments in the prototype discussion forum and the baseline discussion forums. All submitted questions and comments were designed by researchers and loaded into the system before the experiment was started. The submitted posts were designed to be related with created topic in the discussion forum database for example, one question was 'what is information retrieval and information retrieval system'? The postings

automatically triggered the systems to send the message text on for processing. The test set contained a number of posts, but there was no guarantee that the posts would be answered or even discussed in the corpus. The testing performance for this situation was expected to be considerably worse than for the situation where the post was always answered.

9.4.2 The Second Experiment

The second evaluation experiment was a user-system interaction test where users completed an information seeking task using the prototype and baseline systems. The users were given an information need in a learning scenario and were requested to submit posts to identify information and articles relevant to the need. Users were required to identify as many items as possible within a time frame.

This experiment was designed as a simulated rather than real task completion process to overcome time and resource constraints. A simulated learning scenario was created as the background. The simulated scenario was based on the research objectives of this project to overcome communication limitations of discussion forums in the learning environment.

Volunteer students were recruited to use the prototype and baseline systems. The simulated scenario was learning-based and all user activities were designed to be learning-oriented, for example as users enrolled in an IT course that needed to use the discussion forum to find useful information in order to complete a course assignment. The participants were required to complete an information seeking task using the prototype and baseline discussion forums as the information source. The simulated

scenario involved a group of 30 students who had enrolled in an ‘Information Technology’ course. The sample size of participants was similar with other related studies (Belkin et al., 2004; Turpin & Scholer, 2006; Venturi & Bessis, 2006). These students were given an assignment to design an approach to evaluate the effectiveness of an information retrieval system. There were discussion forums for this Information Technology course, and all students used the discussion forums (prototype forum, eQuake forum and intelligent-bot forum) to seek information and ask questions. The discussion forum contains 950 loaded articles and users’ posts from the first experiment. Some articles and postings were related to evaluation methodology and information retrieval and some were not. All relevant items were identified by the researcher before this experiment started.

In the simulated scenario, participants sought useful postings by submitting one or more posts in each of the prototype and baseline discussion forums.

After users completed the experiment, the list of relevant information found from each participant was collected and those related to the corresponding topic (as determined by the researcher in advance) were identified and the details of actual experiment data are presented in the chapter 10. For each user, the search effectiveness was calculated by the number of postings found related to the task topic per participant:

Search Effectiveness = the number of posts found related to the task topic

The search effectiveness for each participant was analysed to assess the difference in search effectiveness between the prototype system and the baseline systems.

A questionnaire methodology was used to investigate user satisfaction in the evaluation process of the prototype system. The questionnaire allowed evaluators to gather the participants' perceptions and expectations. It also provided useful information on which qualities were regarded as being more important for example, usability or efficiency (Venturi & Bessis, 2006). User satisfaction was measured using the questionnaire shown in Appendix II to elicit: the general background information of participants; the perceived quality of retrieved information; the effectiveness of the system; the complexity of the information problems; the retrieved support; and the influence of the user feedback function. All participants were required to complete the questionnaire after the second evaluation task was finished. The experiment was conducted according to the guidelines of the Human Ethics Committee of Massey University and oral informed consent was received from all participants.

The above outline of the experiment describes the designed simulated learning scenario in the second evaluation experiment. In the actual experiment, the simulated learning scenario had three information seeking tasks to be completed with 30 participants. Twenty four undergraduate students majoring in computer science or information systems at Massey University were recruited as experiment learner participants. Meanwhile 6 masters and doctoral students were invited to be the tutor users for the experiment. The number of participants recruited was limited by the available student resource at the time. The details of participants' information will be presented in the next chapter.

9.5 Summary

The evaluation goal in this study was to determine whether the developed prototype system was able to enhance the effective communication amongst learners and between learners and instructors in a discussion forum. Both objective and subjective system performance were adopted to assess the prototype and baseline systems. The baseline systems were determined and installed before the evaluation approaches were selected.

Two evaluation experiments were conducted to assess both objective and subjective system performance. The first was an automation system test to assess the system precision ratio and the second was a user-system interaction test to evaluate search effectiveness. A questionnaire was used to investigate user satisfaction after the second evaluation experiment was completed.

The experiment details and the collected data are presented in the next chapter which goes on to show data analysis, evaluation of results and conclusions for these experiments.

CHAPTER 10

EVALUATION RESULTS AND ANALYSIS

10.1 Introduction

This chapter describes the experimental procedure and results for the evaluation of the developed prototype system. This formal evaluation was performed at the end of the development cycle. The evaluation experiments primarily involved learner users, but one evaluation in user satisfaction section with just instructors was also performed.

Key findings are presented in this chapter following each of the relevant experiment descriptions in the following two sections. This includes statistical comparison of the data between the prototype system and the baseline systems. The chapter concludes with the discussion on the relationship trends between system performance with parameters of posts (e.g. length of posts) and limitations of the prototype forum.

10.2 The First Experiment

10.2.1 Experiment Procedure

In the first evaluation experiment, both the prototype forum and baseline systems (eQuake forum and intelligent-bot forum) were set up on an apache server and a testing script was developed to automatically simulate users submitting questions on both forums. After question submission, all retrieved postings were automatically saved into a text file and manually analysed by the researcher.

A total of 950 articles published in ACM (Association of Computing Machinery) and IEEE (Institute of Electrical & Electronics Engineers) from 1980 to 2008 were inducted into the test collection. This evaluation experiment aimed to simulate an IT course-used discussion forum so all articles in the test collection were related to IT. They were divided into six topic groups and the content of each article was randomly uploaded as a post into each forum. The six topic groups were as follows:

- Information Technologies
- Learning Management Systems
- Information Retrieval Systems
- Search Engines
- System Evaluation
- User Evaluation

All uploaded postings were mapped to a corresponding thread from a set of 40 threads (topic folders) that were created by the researcher in each forum. Meanwhile a group of 10 student volunteers (table 10-1) were recruited from the Computer Science and Information System Departments at Massey University as learner users for both the prototype and baseline forums. These learners were requested to randomly submit related questions and comments into available topic threads in both forums. For the prototype forum system, all learners also was requested to rate each retrieved relevant post (related or not) when they submitted a post.

Table 10-1: The general information of participants

	Undergraduate	Postgraduate
Role	Learner	Instructor
Total	5	5
Age range	18 - 22	22-30
Experience with discussion forums (year)	1-3	2-8

The system finally comprised 40 threads which contained 950 article posts and 220 user postings as the test collection for each baseline forum (eQuake and intelligent - bot). The distribution of the length of each thread, that is, how many posts were included in each thread, is presented in Figure 10-1. Most threads (nearly 70%) had a length of 20 posts or less. Very few threads contain more than 30 posts.

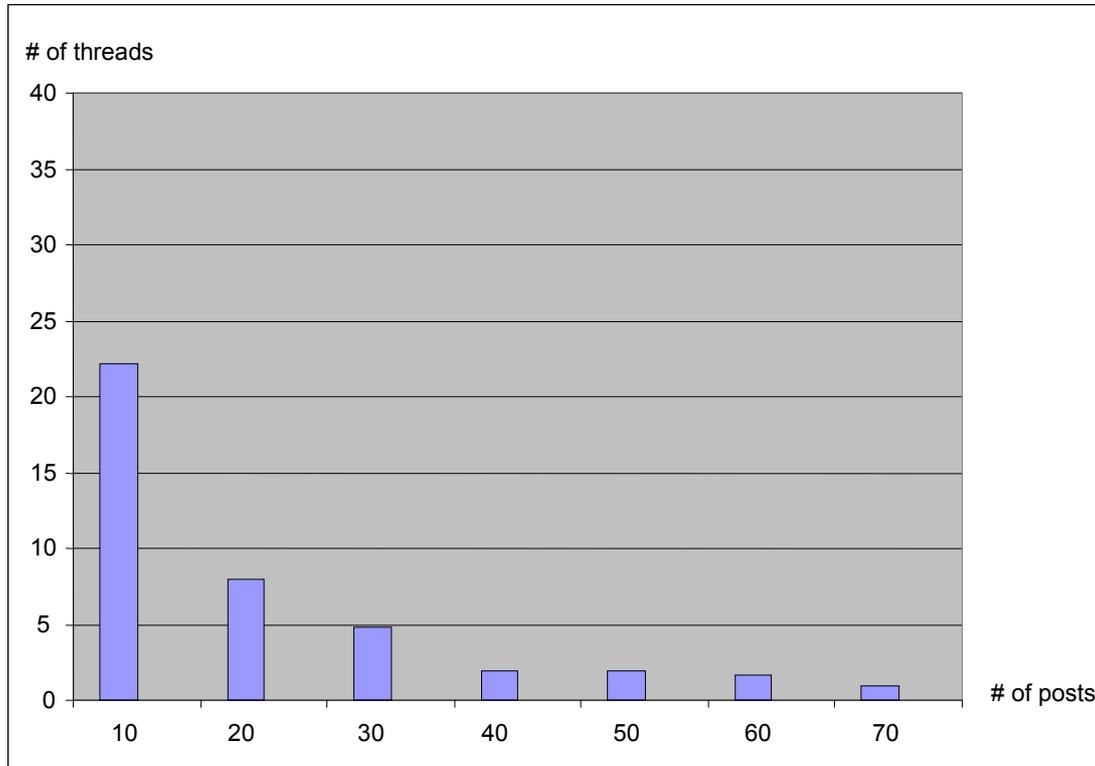


Figure 10-1: Statistics of Thread Length

Once the testing collection had been set up, the testing script was run. This automatically triggered each of 30 simple and complex test questions to be submitted into each forum and all returned posts were recorded to a text file. The threshold for a simple question was a posting containing no more than 50 words. The distribution of the length of each question is presented in Figure 10-2. This figure shows the distribution of the word counts for each question. A total of 5 questions comprised less than 50 words, 5 questions had greater than 500 words, 10 questions were between 50 and 100 words and 10 questions were between 100 and 500 words.

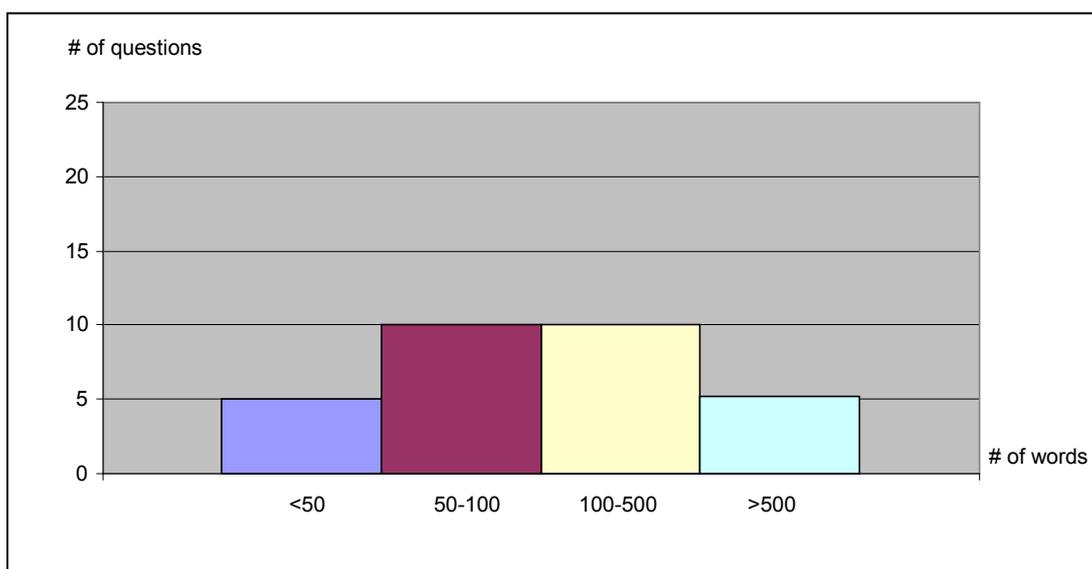


Figure 10-2: Distribution of Testing Question Length

10.2.2 Results and Analysis

When a question was submitted to the forum, the threshold number of retrieved posts could be controlled by forum instructors. In this evaluation process, the maximum number of retrieved relevant posts was restricted to a common threshold number - ten because ten was the threshold also used for baseline systems intelligent-bot forum and eQuake forum.

Table 10-2 shows the results of the first evaluation experiment which aimed to compare the effectiveness of the auto-reply function in the prototype, eQuake and intelligent-bot forum systems. Retrieved information was manually organised into 3 categories by the researcher: 'extract related posts' which contained extracted answers for the submitted questions; 'related posts' which contained some content related to the meaning of the submitted question and 'unrelated posts' which had nothing related to the submitted question. For example, the question 'what is information retrieval?' produced one extracted answer of the post containing '*Information retrieval (IR) is the science of searching for documents, for information within documents and for metadata about*', and one related information was the post containing '*The idea of using computers to search for relevant pieces of information was popularized in an article As We May Think by Vannevar Bush in 1945. First implementations of information retrieval systems were introduced in the 1950s and 1960s.*'

For the prototype discussion forum, 273 retrieved posts were collected and classified, 56 previous posts were identified as extract relevant answers to the submitted questions, 205 retrieved posts were identified as relevant to the submitted questions and 12 previous posts were totally irrelevant.

For the eQuake forum, 278 retrieved posts were collected and classified, 38 previous posts were identified as extracted relevant answers to the submitted questions, 192 retrieved posts were relevant to the submitted questions and 48 previous posts were totally irrelevant.

For the intelligent-bot system, 271 retrieved posts were collected and classified, 52 previous posts were identified as extract relevant answers to the submitted questions, 193 retrieved posts were relevant to the submitted questions and 26 previous posts were totally irrelevant.

Table 10-2: Relevance results in the first evaluation experiment

Retrieved Posts	eQuake Forum	Discussion-bot	Prototype Forum
Total Retrieved	278	271	273
Extract related	38	52	56
Related	192	193	205
Unrelated	48	26	12

These results were used to calculate the percentages of extracted relevant answers, relevant answers and irrelevant answers contained in retrieved posts in the eQuake, intelligent-bot and prototype forums. These percentages are shown in the Table 10-3.

Table 10-3: Ratio of relevance results in the first evaluation experiment

Percentage of retrieved posts	eQuake Forum	Discussion-bot	Prototype Forum
Extract related (ER)	13.7%	19.2%	20.5%
Related (R)	69%	71.2%	75.1%
Extract related and Related (ER&R)	82.7%	90.4%	95.6%
Unrelated (U)	17.3%	9.6%	4%

The table 10-3 shows that the highest percentage of ‘Extract related’ retrieved posts was achieved by the prototype forum (20.5%) and the lowest percentage by eQuake forum (13.7%). In the ‘Related’ row, the highest percentage was still the prototype system (75.1%) and the lowest was eQuake (69%). eQuake forum returned the most unrelated posts (17.3%) and the prototype forum returned the least (4%). This is represented in Figure 10-3.

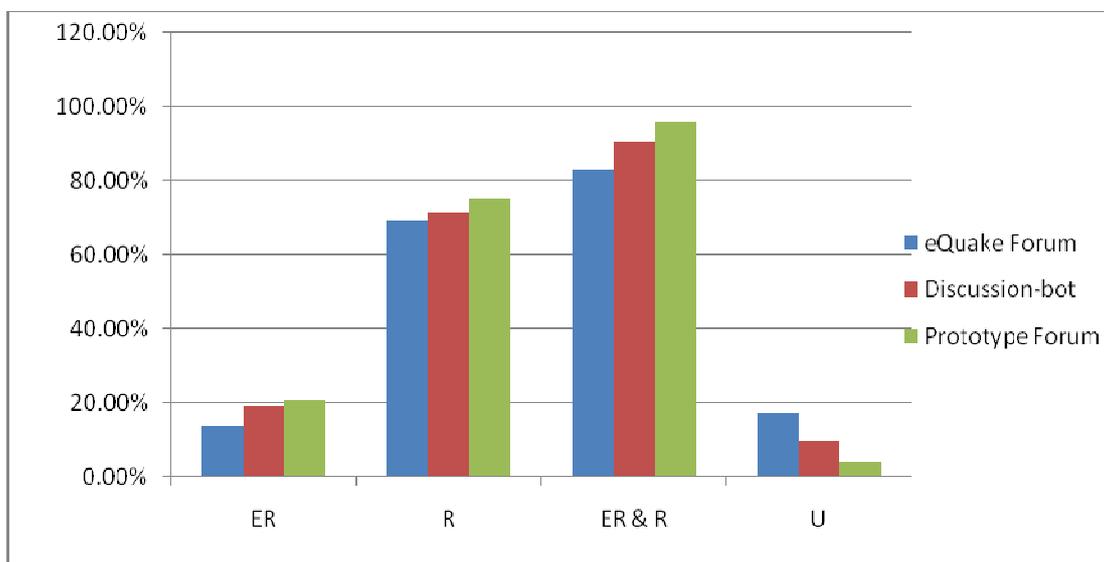


Figure 10-3: The relevance results in the first evaluation experiment

The formula used to calculate precision ratio is:

$$\text{Precision} = (\text{no. of relevant documents}) / (\text{no. of retrieved documents}).$$

Therefore $P_{\text{prototype}}$ is 0.956, $P_{\text{intelligent-bot}}$ is 0.904, and P_{eQuake} is 0.827.

The percentage of ‘Extract Related’ answers rates the prototype forum only 1.3% higher than the intelligent-bot forum but 6.8% higher than the eQuake forum. Similarly the percentage of ‘Related’ posts rates the prototype forum 3.9% higher

than the intelligent-bot and 6.1% higher than eQuake. However from the precision perspective, $P_{\text{prototype}}$ is 5.2% higher than $P_{\text{intelligent-bot}}$ and 7.7% higher than P_{eQuake} .

From the evaluation result, it has shown that the prototype forum had the best precision ratio and eQuake forum had the worst. This is because of different information retrieval approaches adopted in those three forums. Boolean-keyword and TF*IDF were adopted as the main information retrieval approaches for the eQuake forum and the intelligent-bot forum. The eQuake forum and the intelligent-bot forum have been evaluated in other research and have been shown to improve the effectiveness of communication and to enhance the information retrieval performance in discussion forums (Smestad, 2006; Feng et al., 2006).

Simple Boolean-keyword search is the most common information retrieval techniques and it requires users to input the exact words they are looking for into a text field (Salton & Buckley, 1988). Upon submission, a search will return a list of documents that only contain the search terms and Boolean-keyword search gives all words the same value of importance. Previous research had suggested that improvements to Boolean-keyword approach using techniques to weight keywords based on how commonly they appear in the forum like tf-idf (term frequency – inverse document frequency) (Salton & Buckley, 1988; Smestad, 2006; Feng et al., 2006). The IF*IDF approach is a weight often used in information retrieval. This weight is a statistical measure used to evaluate how important a word is to a document in a collection and it is able to enhance the information retrieval performance more than the Boolean-keyword technique (Salton & Buckley, 1988). Meanwhile the prototype forum system adopted a modified Latent Semantic Analysis (LSA) approach which is a

mathematical/statistical technique for extracting and representing the similarity of meaning of words and passages by analysis of large bodies of text. LSA has been proven to significantly improve automatic information retrieval performance more than common a statistical method such as TF*IDF by allowing user requests to find relevant text on a desired topic even when the text contains none of the words used in the query (Deerwester et al., 1990; Landauer, 1997; Foltz et al., 1999; Kintsch et al., 2000).

The results of this experiment support those of previous research., In this evaluation of objective system performance, the prototype forum demonstrated better information retrieval performance than the baseline forums, because the current evaluation experiment demonstrated that the modified-LSA adopted prototype forum had better information retrieval performance (based on measured relevance and precision) than the Boolean-keyword adopted eQuake forum and IF*IDF adopted intelligent-bot forum.

10.3 The Second Experiment

10.3.1 Experiment Procedure

The second evaluation experiment was designed as a user-system interaction to evaluate whether another objective system performance measure, search effectiveness, improved in the prototype forum compared with the baseline forums. In this experiment, a simulated scenario was designed to assess this search effectiveness in the prototype and baseline systems via interaction between users and system. All data

from the first experiment were reused in this experiment. The evaluation experiment was designed as follows:

Simulated Scenario: 30 students were enrolled in an ‘Information Technology’ course. These students were given an assignment to design an approach to evaluate the effectiveness of an information retrieval system. Discussion forums were provided for this Information Technology course, and all students used the discussion forums (prototype forum, eQuake forum and intelligent-bot forum) to seek useful information and ask questions.

Users: 24 undergraduate students majoring in computer science or information systems at Massey University were recruited as experiment participants. Additionally, 6 postgraduate (masters or doctoral) students were invited to be tutor users for the experiment. All 30 users were separated into 3 groups based on three testing forums. There were 2 postgraduate students and 8 undergraduate students in each user group.

Location: The experiment was performed in a classroom at Massey University Albany Campus.

Duration: 90 minutes

Tools: Hardware: five IBM ThinkPad Laptops - Model T400. Software: evaluated systems installed on an Apache Tomcat server networked with the laptops.

Guidelines: The experiment was conducted according to the guidelines of the Human Ethics Committee of Massey University and oral informed consent was received from all participants.

Training Session: Every user was trained for up to 30 minutes to ensure that they were familiar with the task and the user interface before beginning the actual experiment.

User Task: There were 3 information seeking tasks in this experiment and users were requested to complete each task within 15 minutes. The 15 minutes duration was based on a user test which concluded that this was the average time to find relevant information to complete the each information seeking task in a discussion forum. For each task, every user group was required to use a different forum (prototype forum, intelligent-bot forum and eQuake forum). Users could browse the forum to seek relevant materials and they were also required to submit a question or comment in their own words to help them complete their tasks. This question or comment was posted to all three forums. The individual user ability and skill were considered in this evaluation experiment. All recruited users were studying towards computer science or information systems majors so they were familiar with all terminologies and concepts in this experiment.

- **Information seeking task 1:** In 15 minutes, each user needed to find as many possible relevant posts regarding the sub-tasks below and write the retrieved posts' titles down on an answer sheet.
 - What is a search engine?
 - What is information retrieval?

- What is an information retrieval system?
- **Information seeking task 2:** In 15 minutes, each user needed to find relevant posts and write their titles down on an answer sheet for the following questions.
 - What is system evaluation?
 - What are measures of system evaluation?
 - What are measures of information retrieval system evaluation?
- **Information seeking task 3:** In 15 minutes, each user needed to find relevant posts and write them down on an answer sheet.
 - If a new information retrieval system is developed, how can you evaluate whether it has better performance than other similar system and what evaluation approaches are normally used for such systems?

User satisfaction: user satisfaction can be measured through a questionnaire in which the user evaluated the quality of the retrieved information and the effectiveness of the system (Brajnik et al., 1996). The questionnaire was based on previous questionnaires used to investigate user satisfaction for other IR-related systems (Brajnik, et al., 1996, Smestad, 2006; Feng et al., 2006; Venturi & Bessis, 2006). An questionnaire was developed and then handed out to all after they had completed all information seeking tasks. This questionnaire comprised two parts: part A was for all experiment participants and part B was only for tutor users to complete. The questionnaire is shown in Appendix II.

10.3.2 Experiment Results and Analysis

During the first few minutes of the experiment, all participants browsed the threads and then submitted a post (question or comment) regarding the sub-task into the corresponding topic thread of each forum. The average length (word count) of the submitted post per user for each task is shown in Figure 10-4. This shows that the average lengths of submitted posts was 15, 21 and 51 words for the first, second and third task respectively.

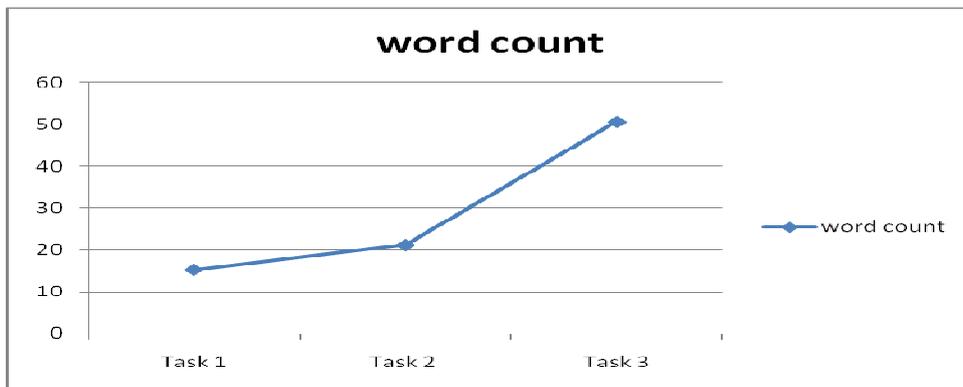


Figure 10-4: Average Length of Submitted Post for Each Task

In this experiment, all users have used the auto-reply functionality to submit questions into corresponding forums and then received relevant information from the retrieved results. All answers from users were manually collected by researchers to verify if the answers were related to each sub-task. To avoid the individual user's ability and skill influencing the experiment result, the average number of answers found in each user group was utilised as a performance trend to be analyzed and discussed.

In the first information task seeking process, all users spent some time (average 4.5 minutes for each task) browsing the forums first but only 3 users wrote corresponding

answers after it. Most users (27 out of 30) started writing down answers after they submitted a question into the corresponding forums. About 93% of the answers that users found for the first task came from retrieved results rather than forum browsing. This result suggests that it was difficult to quickly find relevant information from heaps of posts in the discussion forums, but the auto-reply function helped users quickly and effectively find relevant posts and information in order to achieve the first information seeking task.

The first user group used the prototype forum, the second user group used the intelligent-bot forum and the third user group used eQuake forum. An average of 12 relevant posts was found per user for each sub-task in the first group, 9 in the second group and only 7 in the third group. Table 10-4 shows the average number of relevant posts found per user for each sub-task by forum type during the first task.

Table 10-4: Average number of relevant posts found for each sub-task

Task 1	Prototype Forum	Intelligent-bot	eQuake Forum
Group 1	12	-	-
Group 2	-	9	-
Group 3	-	-	7

In the second information task seeking process, from the observation, all users spent some time (average 4.8 minutes for each task) browsing the forums first but only 2 users wrote down corresponding answers after it. Most users (28 out of 30) started writing down answers after they submitted a question into the corresponding forums.

About of the 95% answers that users found for the first task came from retrieved results rather than forum browsing.

In the second information seeking task, the first user group used the eQuake forum, the prototype forum was used by the second user group and the third user group used the intelligent-bot forum. An average of 7 relevant posts was found per user for each sub-task in the first group, 9 in the second group and only 6 in the third group. Table 10-5 shows the average number of relevant posts found per user for each sub-task by forum type for the second task.

Table 10-5: Average number of relevant posts found for each sub-task

Task 2	Prototype Forum	Intelligent-bot	eQuake Forum
Group 1	-	-	6
Group 2	9	-	-
Group 3	-	7	-

In the third information task seeking process, from the observation, all users spent some time (average 5.1 minutes for each task) browsing the forums first but only 1 user wrote corresponding answers after it. Almost all users (29 out of 30) started writing down answers after they submitted a question into the corresponding forums. About 97% of the answers that users found for the first task came from retrieved results rather than forum browsing.

In the third information seeking task process, the intelligent-bot forum was used by the first user group, the eQuake forum was used by the second user group and the

third user group used the prototype forum. An average of 5 relevant posts was found per user for each sub-task in the first group, 4 in the second group and 9 in the third group. Table 10-6 shows the average number of relevant posts found per user for each sub-task by forum type in the third task.

Table 10-6: Average number of relevant posts found for each sub-task

Task 3	Prototype Forum	Intelligent-bot	eQuake Forum
Group 1	-	5	-
Group 2	-	-	4
Group 3	9	-	-

The above findings show that the highest average number of relevant posts found for each sub-task was achieved using the prototype forum and the lowest number was achieved using eQuake. Figure 10-5 compares these results for the prototype forum, intelligent-bot forum and eQuake forum. This graph clearly shows that the average number of relevant posts found in prototype forum is higher than for both intelligent-bot and eQuake forums for all three tasks. For the first and second task, the average number of relevant returns by prototype forum was 33% and 25% higher than intelligent-bot, and 71% and 50% higher than eQuake. For the third task, the relevant posts found by the prototype forum were much higher than by the intelligent-bot and eQuake forums (80% and 125%).

Search effectiveness is calculated from the number of postings found related to the task topic:

$$\text{Search Effectiveness} = \text{no. of posts found related to the task} / \text{no. of participants}$$

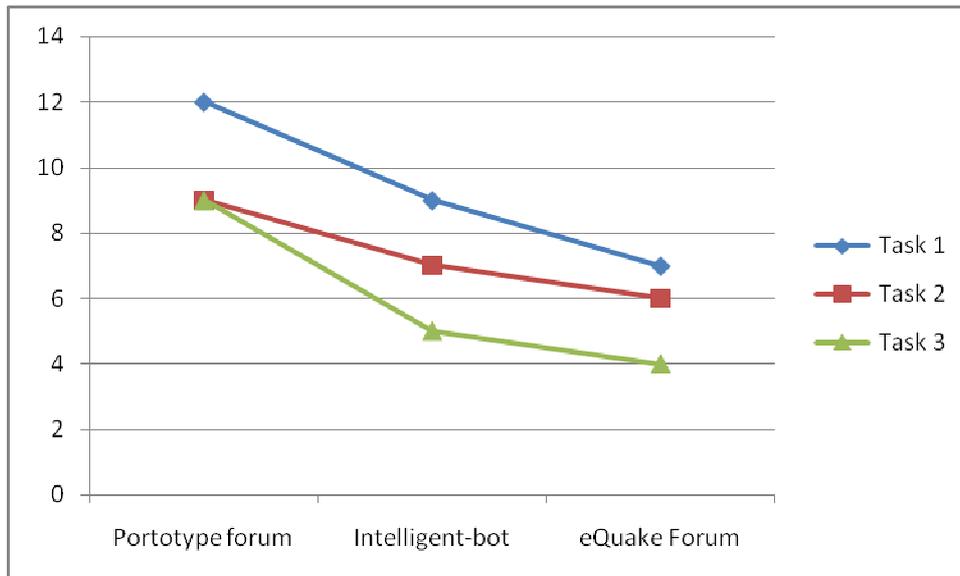


Figure 10-5: Average Relevant Posts Found in Three Forums for Three Tasks

From the experiment results, over 90% answers came from the retrieved results. It suggests that the search functionality was often used as the major approach to find relevant answers in this experiment. Regarding the search effectiveness measure, the experiment results have clearly shown that the search effectiveness of the prototype forum had a higher number of relevant posts found than baseline systems for all three tasks. The better search effectiveness in the prototype forum suggests that the prototype system has a more powerful system performance which provides more relevant information and enhances the communication capabilities of the discussion forum. Figure 10-5 indicates that the best improvement was achieved for task 3. Figure 10-4 shows the average word count of submitted posts for task 3 was double the post length for task 1 or 2. Further discussion on these experimental results is presented in Section 10.4.

10.3.3 User Satisfaction

After the second experiment was completed, all 30 participants (24 learner users and 6 tutor users: table 10-7) were polled for user satisfaction using a custom designed questionnaire. The designed questionnaire comprised two parts. Part A was completed by all participants and part B was only for tutor users.

Table 10-7: Summary of evaluators

Category	Number	Major		Gender	
		Computer Science	Information Systems	Male	Female
Postgraduate	6	4	2	3	3
Undergraduate	24	13	11	16	8
Total	30	17	13	19	11

Part A contained three sections. The first section was designed to investigate user satisfaction and their opinion on the system usability for each function of the prototype forum system. The second section was used to investigate the benefits and limitations of using the prototype forum. Finally, the third section was used to collect the general background information of each participant.

Part B was used to elicit user satisfaction and opinion about the prototype forum from the tutor user perspective.

Section 1: In order to evaluate the usability of the “auto-reply function” and “user feedback function” of the prototype system, four questions were designed to elicit users’ opinion about: how easy the “auto-reply function” was to use; to what extent

the “auto-reply function” was a useful functionality; how easy the “user feedback function” was to use; and to what extent the “user feedback function” was a useful functionality.

Table 10-8: Statistical data for the section 1

	Auto-reply function	User feedback function
Easy to use	25	20
Useful function	25	15

In the completed questionnaires (table 10-8), most users suggested that the “auto-reply function” was easy to use and useful. In detail, 83% (25 out of 30) evaluators felt that it was very easy to use the “auto-reply function” and that the function (synchronously retrieving related information while they submitted a post) was very useful. Sixty-seven percent (20 out of 30) of respondents considered that the ‘user feedback function’ was easy to use, but only 50% (15 out of 30) of participants agreed that the “user feedback function” helped to provide more accurate information to them.

Section 2: The following four questions were designed to evaluate the benefits and drawbacks of the overall prototype system: From your point of view, what are your personal expectations in using the prototype discussion forum? What are the reasons why you would like to use this prototype discussion forum? What are the benefits you believe the prototype discussion forum can bring to learners? What are the negative issues that you find annoying in using the prototype discussion forum?

From the completed questionnaires about 93% of the evaluators (28 out of 30) expected the prototype forum to deliver good quality and highly related post contents. When asked the reason to use the prototype forum rather than other forums, 86% (26 out of 30) of people chose the first option, “more useful and related information could be quickly found with the ‘auto-reply’ function”. Twenty-seven evaluators (90%) considered that more efficient and effective retrieval of useful information was the main benefit in using the prototype system. Meanwhile twenty-four respondents (80%) chose “knowledge being easily shared and used” as the major benefit that the prototype forum could bring to learners. Another seven participants selected “less burdens for teachers to manage the discussions” as the major benefit. As a negative issue, all participants considered that the search approach adopted in the prototype forum could be improved to retrieve answers that more exactly matched the submitted posts.

Table 10-9: General information of evaluators

	Number	Percentage
Familiar with discussion forums	30	100%
Familiar with search engines	30	100%
Often use course discussion forums	29	97%
Discussion forum is first approach to find information	20	67%
Search engine is first approach to find information	9	30%
Text book is first approach to find information	1	3%

Section 3: This section was used to collect general background information on participants such as familiarity with discussion forums; familiarity with search engines; how often they use discussion forums in their own learning and what their first approach is when they have questions? This section was included to identify the effect of knowledge and experience on the results of the evaluations.

In this section (table 10-9), all respondents reported being very familiar with discussion forums and search engines. Almost all participants (97%) indicated that they often used course discussion forums to seek answers and useful information for their study. Sixty-seven percent of evaluators (20 out of 30) stated that the course discussion forum is the first approach that they use when they need to find course related information and useful knowledge, and nine (30%) said that an internet search engine is their first choice. Only one person (3%) selected a text book as their first approach.

Part B (Tutor user only): This part was used to investigate the background and opinions of prototype forum from tutor user's perspective. It included: their familiarity with the instructor role in discussion forums, the major issues for instructors in using discussion forums, the reasons why an instructor might like to use this prototype discussion forum, the benefits instructors believe the prototype discussion forum can bring to learners and instructors and what needs to be improved in the prototype forum system.

All six tutor users were very familiar with the instructor role in the discussion forum. They all advised that the biggest disadvantage of using discussion forums was the

time and effort spent on forum management. Easing the management burden was selected by five people as the major reason to use the prototype forum. All six participants suggested that the major benefit offered by the prototype forum was the improvement in effective communication between learners, and between instructors and learners. For the question what needs to be improved in the prototype forum, four people thought that the information retrieval approach could be improved for accuracy and two suggested that the forum posts could be better categorised into groups such as query, answer and comments. Along with more relevant information provided in the prototype forum from the previous evaluation experiments' result, the questionnaire results also suggested that most users were satisfied with the "auto-reply function" in the prototype forum. However only 50% of people thought the "user feedback function" helped their information seeking process. Nearly all evaluators (93%) considered that good quality and relevant content of posts was the key expectation of discussion forum systems. The main reason for 86% users to choose the prototype forum was that more relevant and useful information could be quickly found with the 'auto-reply' functionality across different topic folders in the prototype forum. Therefore 'knowledge being easily shared and used' and 'efficiently and effectively obtaining useful information' were the most popular (90%) benefits of the prototype forum. Most respondents suggested that effectiveness of retrieving extracted answers needs to be improved in the prototype forum.

From the users' background information, most learners and tutors were familiar with discussion forums and search engine tools. Two thirds of respondents frequently used discussion forums as the first tool to find useful knowledge and relevant information in their learning processes. The other third normally used search engines.

10.4 Discussion and Conclusion

10.4.1 Further Discussion

The user satisfaction survey showed that the discussion forum was the most important communication tool for a course of study and the prototype forum achieved most user expectations with regard to improving effective communication between users and overcoming the limitations of traditional discussion forums. Meanwhile analysis of the two evaluation experiments demonstrated that the prototype forum enhanced effective communication in the discussion forum and achieved better system IR performance compared with baseline systems.

The evaluation experiment showed two trends for the prototype system. The first is that the length (word count) of posts was a very important parameter influencing the relative performance of information retrieval between the prototype forum and the intelligent-bot forum. This is because if a post contains more words (longer length), its content can have more keywords and the keyword density is an important parameter influencing statistics-based information retrieval approaches (Deerwester et al., 1990; Landauer, 1997; Foltz et al., 1999; Kintsch et al., 2000). In the second evaluation experiment, the relevant posts found by the prototype forum were 33%, 28% and 80% higher than by the intelligent-bot forum for the three tasks. This result clearly shows that the prototype forum had much better performance for the third task than intelligent-bot forum. This result was also related the length of submitted posts in the third information seeking task process. From the collected data the average length of each submitted post were 15, 21 and 51 words for the first, second and third task respectively. Around 90% of submitted questions for task 3 had more keywords than

questions for other 2 tasks. Therefore the longer length of the submitted posts represented more keywords in the content of posts.

Figure 10-6 clearly shows a positive proportional relationship between search effectiveness and average length of submitted posts. From the statistical data of submitted questions, about 90% of questions for task 3 had higher keyword density than task 1 and 2. This relationship suggests that the number of keywords in the submitted posts controls the effectiveness of information retrieval performance in the prototype forum system.

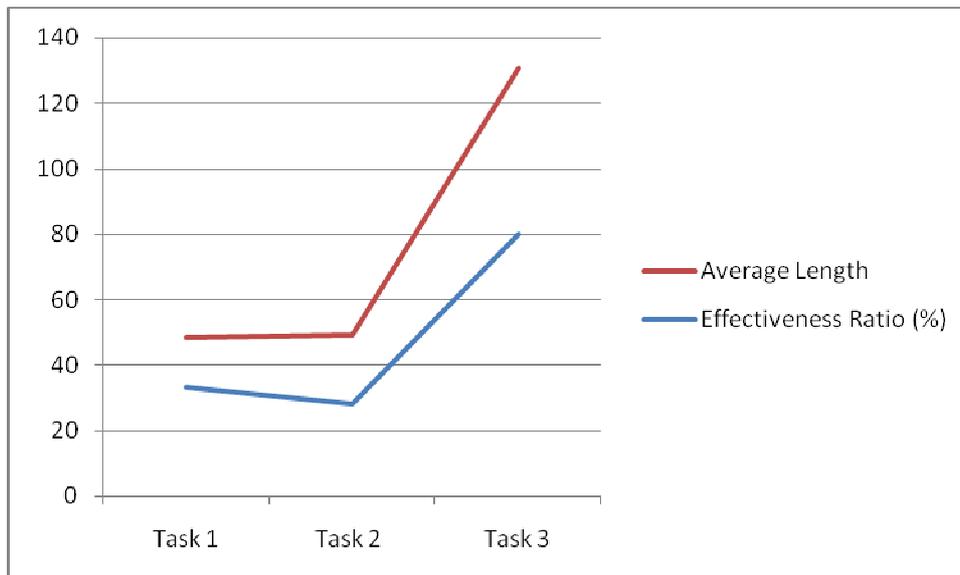


Figure 10-6: Relationship between Effectiveness Ratio and Average Length

In the other words, if a submitted post had fewer keywords, the performance of the information retrieval function of the prototype forum might be affected, increasing the amount of irrelevant information returned. Therefore the benefits of the proposed approach are limited to shorter messages.

The second trend shown in the user satisfaction investigation was that the “user feedback function” was not recognised as an effective function to help users find relevant information. In the prototype forum database, there were only 48 feedback stored against 23 posts. Therefore the likely reason is that not much changed after submitting feedback on the returned information. The prototype forum did not show the back-end data computing process so users could only identify the impact of this function from the retrieved information. The user feedback function was proposed so that when more feedback was submitted from users, more retrieved information will be modified. This is because the proposed concept reclassification approach was based on relevance feedback from users and the more relevance feedback is given from the users, the more postings are accurately classified into semantic concepts to influence the information retrieval result. Moreover, the influence parameter β is also highly dependent on the number of relevance feedback. However, the amount of feedback from users was not sufficient to noticeably influence the performance of information retrieval because the evaluation experiment did not have much user feedback on the existing posts in the prototype forum.

The user satisfaction investigation showed that the biggest limitation of the developed prototype forum system was that it was not intelligent enough to identify all extracted answers for submitted questions, although it was able to provide more effective relevant information to users. To overcome the current limitation, advanced natural language process such as lexical database word-net technology (Miller, 2009) could be used in the prototype forum system to more accurately categorise each submitted message into the appropriate concept. Such an approach can collect and analyse the noun phrases of each message which would enhance the system’s ability to

understand the meaning of the messages. Discourse representation theory has shown that the primary concepts in text are carried by noun phrases, and noun phrases can be considered the conceptual entities in text messages (Wu & Chen, 2004).

Finally, a weak relationship between system performance and user feedback was found because the amount of user feedback was not enough to influence information retrieval results. Therefore the user feedback function needs to be made more useful even when the amount of user feedback is not very high to influence information retrieval. Improving this performance needs to be pursued in future research..

10.4.2 Conclusion

The experiment results show that the prototype forum system fulfilled the evaluation goals defined in Chapter 9. In the first evaluation experiment, the prototype forum demonstrated better effective information retrieval performance than the baseline forum systems enabling it to more effectively capture previously related information and useful knowledge. From the statistical data analysis in the second evaluation, the prototype forum system showed better system performance than the other two baseline system enabling it to find more relevant information within a time frame. According to the feedback from the questionnaires, the prototype system was considered a better solution for improving effective communication amongst learners, more easily capturing previous information and effectively finding relevant information. Furthermore, while improving effective communication between learners and instructors, instructors were able to reduce the time and effort spent on discussion forum management.

From the experimental data and questionnaire analysis, the biggest limitation of the prototype system was that it was not intelligent enough to understand all messages or effective enough to help users find the most exact answers to their questions in the prototype discussion forum. Advanced natural language processing technology and user feedback function improvement are recommended to further enhance the performance and usability of the prototype system. In the next chapter, the conclusion of this thesis and possibilities for future research are presented.

CHAPTER 11

CONCLUSIONS AND FUTURE WORK

11.1 Dissertation Summary

The research objective of this study is to improve the effectiveness of communication between learners and between teachers and learners in online discussion forums. Literature has been reviewed in the areas of limitations of traditional discussion forums in learning communities and technologies used in educational areas including information retrieval, text similarity and the user model concept. Building on the background literature, various challenges were identified in Chapter 3 for using discussion forum systems in learning environments and a requirements analysis is undertaken to identify the ways for overcoming the limitations of traditional discussion forums. These are summarised as follows.

- A forum should be able to be shared by various courses in different institutions and since different institutions use different learning management systems and discussion forums, a key requirement for the proposed system is to be able to be integrated properly with the existing discussion forum systems that are used in different courses and different institutions.
- Queries submitted by users should be automatically responded by the discussion forum.
- Queries submitted by users should be intercepted by the system, the existing knowledge base should be searched, and a list of possible answers and relevant information should be displayed to the users, before the queries are

added to the system. This would allow the users to quickly find relevant information to their queries rather than having to browse or search through the entire forum to see if the information already exists. This feature should help to minimise the creation of duplicate messages in the system.

- Students should be able to give feedback on a searched message to indicate whether the displayed message is related to their queries. The system should then use this information to more accurately link messages with certain types of queries so that more accurate search results could be extracted for future queries.
- Teachers should be able to reduce the time consuming burden of managing the discussion forum and answering duplicated questions when the forum has immediate response functionalities of searching relevant information in the forums.

Various solutions proposed in other research are analysed in this research with regard to the above requirements, and were discussed in the first section of Chapter 4. Additionally, it has been suggested that a system proposed to meet such requirements should include the following functionalities:

- It should have general discussion forum functions (view, post, and reply).
- Knowledge in this forum should be sharable between courses and institutions.
- It should have a synchronised function to respond to users' queries.
- It should have a search function for existing relevant information within the forum.
- Users should be able to rate the information retrieval results to enhance the forum's accuracy.

In the proposed solution architecture, a core component, namely the Dialogue Mechanism, was depicted to perform natural language processing tasks in the new forum system. Chapter 5 summarised various approaches to natural language processing for text similarity calculation, information retrieval and user feedback that have been utilised in other research. Based on analysis of those approaches, an innovative approach was then proposed to design and develop the Dialogue Mechanism component.

This was followed by the development plan in the sixth chapter. Chapter 6 analysed the benefits and limitations of available open source discussion forum systems in order to select an appropriate forum system to be customised in the development process. The proposed development approach, development environment and associated tools were also introduced.

Chapter 7 illustrated the architecture and the implementation of the developed prototype discussion forum system. This prototype was implemented as one instance of the proposed discussion forum system.

The eighth, ninth and tenth chapters presented the evaluation of this research. The eighth chapter first provided a research background to traditional evaluation methods and approaches. The subsequent sections introduced traditional evaluation methodologies for Information Retrieval (IR) and presented evaluation approaches used by other related research. The ninth chapter outlined the evaluation approach chosen and the consequent evaluation experiments designed to determine how well the prototype system met the system design requirements and achieved research

objectives. Chapter 10 presented analysis of data from the evaluation experiments, along with an interpretation of findings and recommendations.

Finally, this chapter presents a summary (above) and highlights major contributions and limitations of this study. Future research directions are also suggested.

11.2 Conclusion: Contributions, Limitations, and Future Research

One of the most important contributions of this research is the solution proposed to improve effective communication and overcome major limitations of discussion forums in learning communities. The proposed solution is a novel approach that effectively finds relevant information from previous postings across various institutions' systems. The prototype forum system has been implemented as one instance of the proposed solution to achieve the main search function of the proposed design but the function that can be integrated with various institutions' system will be developed in the future work.

Traditional discussion forums are used by the majority of on-line courses and learning management systems. Concerns about incorporating them into learning environments have included that:

- it is hard for learners to contribute new ideas and manage their learning from the discussion forums (Sheard et al., 2003);
- automatic responses by the system for duplicate queries can reduce the number of responses required by the instructor or other students (Weller et al., 2005);

- reusing data to support improved learning is a worthy goal that can be extended to assessment of learning; creating a clear structured discussion and automatically identifying similar postings in the forums will help users contribute more meaningful content (Wijekumar & Spielvogel, 2006);
- immediate feedback can eliminate misconceptions, hold the interest of students and prompt them to think about their postings (Wijekumar & Spielvogel, 2006).

Recent proposed solutions include a number of systems and two recent existing solutions such as eQuake and intelligent discussion-bot which have been proved showing a good performance to overcome current limitations of traditional discussion forums used in academia. Both aim to improve communication in and usability of discussion forums in learning environments. Even though both solutions improve forum usability and meet some of the proposed system requirements for solving problems of traditional forums, they have obvious limitations such as ineffective information retrieval processes and limited system architecture.

This research proposed further design improvements by combining an architecture that supports knowledge reuse and interactions between users across multiple institutions with an innovative advanced information retrieval algorithm that enhances the usability of forum systems to improve the effectiveness of communication between users. A prototype was only developed for the search functionality of the proposed system as one instance of the proposed innovation in this study.

The two experiments described in Chapters 8, 9 and 10 were designed and conducted to evaluate system performance of the prototype. From the objective system performance perspective, the prototype forum demonstrated best information retrieval performance than the baseline forums. From the objective system evaluation, the prototype forum is shown to have more effective information retrieval performance (for effectively capturing previously related information and useful knowledge) compared with baseline forums. It was observed that the prototype forum showed better system performance than the baseline systems for finding relevant information within a time frame. Meanwhile from the subjective system performance perspective, most users considered that the discussion forum was the most important communication tool for a course of study and the prototype forum achieved most user expectations with regard to improving effective communication between users and overcoming limitations of traditional discussion forums. The prototype system was considered to be a better solution for achieving effective communication among learners, easily capturing the previously posted information and effectively finding relevant information and useful knowledge. Furthermore, instructors were able to reduce the time and effort spent on discussion forum management. Therefore on the top of existing solutions, the prototype system has been evaluated as a better solution to improve the effectiveness communication in discussion forums.

The evaluation data showed two trends for the prototype system. The first is that the number of keywords in submitted posts controls the effectiveness of information retrieval in the prototype forum. The second is that the user feedback function of the prototype forum is not recognised as an effective function to help users find relevant information. Furthermore, there is a weak relationship between system performance

and user feedback because the amount of user feedback was below the threshold to influence the information retrieval result. Based on the user satisfaction investigation, it is suggested that the biggest limitation of the prototype system is that it is not intelligent enough to understand all messages or effective enough to help users find the most exact answers to their questions.

Although the prototype system achieves all the anticipated objectives and has some additional advantages over traditional discussion forum system and other solutions, the system can be improved in many ways. The most important improvement would be to make the prototype system more intelligent, not only for finding more relevant information but, with a better search methodology, for finding more exact answers to submitted questions. With this improvement, it will be able to enhance the performance of the search functionality of the prototype system so that improving the effectiveness of communication in the prototype forum system.

To weaken the influence of length of posts and user feedback one could improve the system usability and performance. In future research, advance natural language processing technology such as word-net technology could be adopted to enhance the performance and usability of the prototype system. Such technology would be appropriate for message processing in the discussion forum.

With enhancing the search functionality of the prototype system, there is another important aspect for the future work. It is the architecture implementation of the proposed system which can be integrated with different forum systems of multiple institutions but this has not been included within the scope of the development plan in this study. In the future development plan, the current prototype forum will be

integrated with different forum systems of various institutions so that the prototype forum is able to support knowledge reuse and the interactions between users in separate courses and institutions.

REFERENCES

Agosti, M & Smeaton, A.(editor) (1996). Information Retrieval and Hypertext, Kluwer Academic Publishers, 1996.

Ahn, H. & Picard, R.W. (2005). Affective-Cognitive Learning and Decision Making: A Motivational Reward Framework for Affective Agents. The 1st International Conference on Affective Computing and Intelligent Interaction (ACII'05), Beijing, China, October 22-24.

Allee, V. (2003). The Future of Knowledge: Increasing Prosperity Through Value Networks. Butterworth-Heinemann, New York.

Alonso, G., Casati, F., Kuno,H., & Machiraju, V. (2003). Web Services: Concepts, Architecture, and Applications. Springer, Berlin Heidelberg New York (ISBN: 35404440089).

Anderson, M. & Jackson, D. (2000). Computer systems for distributed and distance learning. Journal of computer assisted learning, 2000, vol. 16, No. 3, pp. 213-228 (4 p.1/4).

Apache (2008). 'Apache License', Available from web-site (<http://www.apache.org/licenses>), Accessed: 01. December 2008.

Apache (2008). Software package, Available from web-site (<http://httpd.apache.org>), Accessed: 01. December 2008.

Apache Axis (2008). Software package, Available from web-site (<http://ws.apache.org/axis>), Accessed: 01. December 2008.

Apache Tomcat (2008). Software package, Available from web-site (<http://tomcat.apache.org>), Accessed: 01. December 2008.

Arasu, A., Cho, J., Garcia-Molina, H., Paepcke, A & Raghavan, S. (2001). Searching the Web. *ACM Transactions on Internet Technology*, Vol. 1, No. 1, August 2001, Pages 2–43.

Balabanovic, M & Shoham, Y. (1997). Fab: Content-Based, Collaborative Recommendation. *Communications of the ACM*, 40(3):66–72, 1997.

Balke, T. & Eymann, T. (2008). The conclusion of contracts by software agents in the eyes of the law, *International Conference on Autonomous Agents Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems - Volume 2*, Pages 771-778.

Baraniuk, R. G. (2006). 'Connexions - Building Communities and Sharing Knowledge', Available from web-site (<http://video.google.com/videoplay?docid=6852287090518403675>), Accessed: 30. Dec 2008.

Barker, S. (2002), *Online Discussion Boards: Impacting the Learning Experience*, Australian Computer Education Conference (ACE2003), Adelaide, Australia. *Conference in Research and Practice in Information Technology*, Vol.20

Basu, C., Hirsh, H. & Cohen, W. (1998). Recommendation as Classification: Using Social and Content-Based Information in Recommendation. In *Proc. of the 15th National Conference on Artificial Intelligence (AAAI '98)*, pages 714–720, 1998.

Bauman, M. (1997). *Online Learning Communities*, Available from web-site http://kolea.kcc.hawaii.edu/tcc/tcc_conf97/pres/bauman.html Accessed: 30. Dec 2008.

Beaulieu, S.E. & Robertson, M. (1997). Research and evaluation in information retrieval” , *Journal of Documentation*, Vol.53, No.1, pp.51.

Beehive Website (2008). available from web-site <http://beehiveforum.sourceforge.net>, Accessed: 01.December 2008.

Belkin, N. J. & Croft, W. B. (1987). Retrieval techniques, in Annual Review of information science and technology, M. E. Williams, Ed. Chapt. 4, pp. 109 -145. Elsevier, 1987.

Bellegarda, J. R. (2000). Exploiting latent semantic information in statistical language modeling. Proceedings of the IEEE, 88(8):1279–1296.

Berry, M., Dumais, S. & O'Brien, G. (1995). Using linear algebra for intelligent information retrieval. SIAM Review, 37(4): 573-595, 1995.

Bhagyavati, Kurkovsky, S., Whitehead, C. (2005), Using Asynchronous Discussion to Enhance Student Participation in CS Course, SIGCSE'05, February 23-27, 2005, St. Louis, Missouri, USA.

Bielaczyc, K. & Collins, A. (1999). Learning communities in classrooms: A reconceptualization of educational practice. In C. M. Reigeluth (Ed.), Instructional-design theories and models (volume II): A new paradigm of instructional theory (pp. 269-292). Mahwah, NJ: Lawrence Erlbaum.

Billhardt, H., Borrajo, D. & Maojo, V. (2002). A Context Vector Model for Information Retrieval”, Journal of the American Society for Information Science and Technology, Vol.53, No.3, pp.236 - 249.

Bradner, E., Kellog, W. A., & Erickson, T. (1999). The adoption and use of 'babble': A field study of chat in the workplace. Paper presented at the Sixth European Conference on Computer Supported Cooperative Work, Copenhagen.

Brajnik, G., Mizzaro, S. & Tasso, C. (1996). Evaluating user interfaces to information retrieval systems: a case study on user support. Annual ACM Conference on Research and Development in Information Retrieval archive Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval table of contents Zurich, Switzerland, Pages: 128 – 136 Year of Publication: 1996. ISBN: 0-89791-792-8.

Breese, J. S., Heckerman, D. & Kadie. C. (1998). Empirical Analysis of Predictive Algorithms for Collaborative Filtering. In Proc. of the 14th Conference on Uncertainty in Artificial Intelligence (UAI '98), pages 43–52, 1998.

Brookfield, S.D. (1986). Understanding and Facilitating Adult Learning, Jossey-Bass, San Francisco, CA.

Brown, J. S., Collins, A. & Duguid, P. (1996). Situated cognition and the culture of learning. *Educational Researcher*, 18 (1). 32-41.

Brugge, B. & Houghton, A. (1996). Computer-supported Cooperative Work. *Software Engineering: Education & Practice*, IEEE, Dunedin, New Zealand, 1996, 516-519.

Brusilovsky, P. (1996). Methods and Techniques of Adaptive Hypermedia. In *Proceedings of User Modeling and User-Adapted Interaction*, 1996, pp. 87-129.

Brusilovsky, P. (2001). WebEx: Learning from examples in a programming course. In: Fowler, W. and Hasebrook, J. (eds.) *Proc. of WebNet'2001, World Conference of the WWW and Internet*, (Orlando, FL, October 23-27, 2001), AACE, 124-129.

Brusilovsky, P. (2005). KnowledgeTree: a distributed architecture for adaptive e-learning. *Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters New York, NY, USA SESSION: Adaptive e-learning systems* Pages: 104 – 113 Year of Publication: 2004 ISBN:1-58113-912-8.

Brusilovsky, P., Eklund, J., & Schwarz, E. (1998). Web-based education for all: A tool for developing adaptive courseware. *Computer Networks and ISDN Systems*. 30, 1-7 (1998), 291-300.

Brusilovsky, P. & Maybury, M. T. (2002). From adaptive hypermedia to adaptive Web. *Communications of the ACM*, 45, 5 (2002), 31-33.

Brusilovsky, P. & Miller, P. (2001). Course Delivery Systems for the Virtual University. In: Tschang, T. and Della Senta, T. (eds.): Access to Knowledge: New Information Technologies and the Emergence of the Virtual University. Elsevier Science, Amsterdam, 2001, 167-206.

Budanitsky, A., & Hirst, G. (2001). Semantic distance in Word-Net: An experimental, application-oriented evaluation of five measures. In Proceedings of the NAACL Workshop on WordNet and Other Lexical Resources.

Carenini, G., Smith, J. & Poole, D. (2003). towards more conversational and collaborative recommender systems Proceedings of the 8th international conference on Intelligent user interfaces Miami, Florida, USA SESSION: Full Technical Papers Pages: 12 - 18 Year of Publication: 2003 ISBN: 1-58113-586-6

Chanana, V., Ginige, A., & Murugesan, S. (2005). Improving Information Retrieval Effectiveness by Assigning Context to Documents. ACM International Conference Proceeding Series; Vol. 90 Proceedings of the 2004 international symposium on Information and communication technologies Las Vegas, Nevada SESSION: Information processing Pages: 86 – 91.

Charkrabarti, S., Dom, B., Agrawal, R. & Raghavan, P. (1998). Scalable feature selection, classification and signature generation for organizing large text database into hierarchical topic taxonomies, VLDB Journal, Vol.73, p.163-178, 1998.

Chklovski, T., & Pantel, P. 2004. Verbocean: Mining the Web for fine-grained semantic verb relations. In Proceedings of Conference on Empirical Methods in Natural Language Processing.

Cohen, P. R. (1991). Integrated interfaces for decision-support with simulation. in 1991 Winter Simulation Conference Proceedings (Cat. No.91CH3050-2), (New York, NY, USA, 1991), IEEE, 1066-1072.

Constant, D., Kiesler, S. & Sproull, L. (1996). What's Mine is Ours, or Is It? A Study of Attitudes about Information Sharing. *Information Systems Research*, 5 (4). 400-421.

Cox, I. J., Miller, M. L., Minka, T. P., Papathomas, T. V., & Yianilos, P. N. (2000). The Bayesian Image Retrieval System, PicHunter: Theory, Implementation and Psychophysical Experiments. *IEEE Transactions on Image Processing*, 9(1):20-37, 2000.

Croft, W.B. (1995). Effective Text Retrieval Based on Combining Evidence from the Corpus and Users, *IEEE Expert*, Vol.10, No.6, pp.59-63.

Cunningham-Atkins, H., Powell, N., Moore, D., Hobbs, D. & Sharp, S. (2004). The role of cognitive style in educational computer conferencing. *British Journal of Educational Technology*, Vol. 35, No. 1, pp. 69-80.

Cutts, Q. I., and Kennedy, G. E. (2005), Connecting Learning Environment Using Electronic Voting Systems, *Australasian Computing Education Conferences 2005*, Newcastle, Australia. Conference in Research and Practice in Information Technology, Vol. 42.

CVS (2008). Software package, Available from web-site (<http://www.nongnu.org/cvs>), Accessed: 01. December 2008.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing By Latent Semantic Analysis. *Journal of the American Society For Information Science*, 41, (1990) 391-407.

Dixon, N. M., Allen, N., Burgess, T., Kilner, P. & Schweitzer, S. (2005). *Company Command: Unleashing the Power of the Army Profession*. Centre for the Advancement of Leader Development and Organizational Learning, West Point, N.Y., 2005.

Dumais, S. (1995). Latent semantic indexing (LSI): TREC-3 report. In D. Hartman, editor, *The Third Text Retrieval Conference*, NIST special publication 500-225, pages 219-230, 1995.

Eclipse (2008). Software package, Available from web-site (<http://www.eclipse.org>), Accessed: 01. December 2008.

Eclipse Web Tools Platform (2008), Software package, Available from website (<http://www.eclipse.org/webtools>), Accessed: 01. December 2008.

Errico, B & Micarelli, A. (1994). From student modelling to user modelling. *Systems, Man, & Cybernetics*, 1994. *Humans, Information and Technology*, 1994 IEEE International Conference on Volume 2, Issue , 2-5 Oct 1994 Page(s):1885 - 1890 vol.2.

Feng, D., Shaw, E., Kim, J. & Hovy, E. (2006). An intelligent discussion-bot for answering student queries in threaded discussions. *International Conference on Intelligent User Interfaces Proceedings of the 11th International Conference on Intelligent User Interfaces*. Pages: 171 – 177, 2006.

Fischer, G. (1999). User Modeling: The Long and Winding Road. In *Proceedings of UM99: User Modelling Conference (Banff, Canada, 1999)*, Springer Verlag, Wien New York, pp. 349-355.

Fischer, G. (2001). User Modeling in Human-Computer Interaction. Contribution to the 10th Anniversary Issue of the Journal "User Modeling and User-Adapted Interaction (UMUAI)" (2001), Vol. 11, No. 1/2, pp. 65-86.

Foltz, P. W., Laham, D., & Landauer, T. K. (1999). The Intelligent Essay Assessor: Applications to Educational Technology Interactive Multimedia Education *Journal of Computer enhanced learning On-line journal*, 1(2). <http://imej.wfu.edu/articles/1999/2/04/index.asp> (1999). Accessed: 01. December 2008.

Free Software Foundation (1991). 'GPL – GNU General Public License', 81 Available from web-site (<http://www.gnu.org/licenses/gpl.txt>), Accessed: 01. December 2008.

Frei, H.P. & Wyle, M.F. (1991). Retrieval Algorithm Effectiveness in a Wide Area Network Information Filter, Proceedings of the 14th ACM-SIGIR International Conference on Research and Development in Information Retrieval, Chicago, USA, pp.114-122.

Freiberger, P. & Swaine, M. (2000), Fire In The Valley, The Making of the Personal Computer, McGraw-Hill.

French, J.C., Brown, D.E., & Kim, N.H. (1997). A Classification Approach to Boolean Query Reformulation, Journal of the American Society for Information Science, 48(8): 694-706

Friesen, N., Fisher, S. & Roberts, A. (2004). CanCore Guidelines - for the implementation of Learning Object Metadata (IEEE 1484.12.1-2002), Athabasca University.

Furnas, G. W., Landauer, T. K., Gomez, L. M., Dumais, S. T. (1987), The Vocabulary Problem in Human-System Communication: an Analysis and a Solution. Communications of the ACM. <http://www.si.umich.edu/~furnas/Papers/vocab.paper.pdf>

Gamma, E., Helm, R., Johnson, R., & Vlissides, J. (1995). Design Patterns. Addison Wesley, 1995.

Garlan, D., Siewiorek, D. (2002). Smailagic A. & Steenkiste, P. Project Aura: Towards Distraction-Free Pervasive Computing. IEEE Pervasive Computing, 2002, 1 (2): pp.22-31.

Giurgensohn, A. & Lee, A. (2002), Making Web Sites Be Places for Social Interaction. CSCW'02, November 16-20, 2002, New Orleans, Louisiana, USA.

Glance, N., Hurst, M., Nigam, K., Siegler, M., Stockton, R., and Tomokiyo, T. (2005). Driving Marketing Intelligence from Online Discussion, KDD'05, August 21-24, 2005, Chicago, Illinois, USA.

Glossary Website (2008). www.w3schools.com/site/site_glossary.asp. Accessed: 01. December 2008.

Goldberg, D., Nichols, D., Oki, B. M., & Terry, D. B. (1992). Using Collaborative Filtering to Weave an Information Tapestry. *Communications of the ACM*, 35(12):61–70, 1992.

Goren-Bar, D. & Kuflik, T. (2004). Recommending Restaurants through an Adaptive Mobile System. In: Nunes, Nuno Jardim and Rich, Charles (eds.) *International Conference on Intelligent User Interfaces 2004* January 13-16, 2004, Funchal, Madeira, Portugal. pp. 250-252.

Golub, G. & Reinsch, C. (1971). *Handbook for Automatic Computation II, Linear Algebra*, Springer-Verlag, New York, 1971

Golub, G. & Van Loan, C. (1989). *Matrix Computations*, John-Hopkins, Baltimore, Maryland, second edition, 1989

Goman, K. (2002). Why People Don't Share What They Know *Link&Learn Newsletter*, 2002.

Greenberg, S. (1991). (ed.), *Computer-supported cooperative work and groupware*. Harcourt Brace Jovanovich, San Diego, CA, 1991.

Greisdorf, H. (2000). Relevance: an Interdisciplinary and Information Science Perspective, *Informing Science*, Vol.3, No.2, pp.67-72.

Grudin, J. (1992). Groupware and social dynamics: eight challenges for developers. *Communications of the ACM*, 37 (1). 92-105.

Harasim, L. M. (1995). *Learning networks: a field guide to teaching and learning online*. MIT Press, Cambridge, Mass., 1995.

Harter, S. P. (1986). *Online Information Retrieval: Concepts, Principles, and Techniques*, Academic Press, Inc., London, 2002.

Haythornthwaite, C., Kazmer, M. M., Robins, J., & Shoemaker, S. (2000). Community development among distance learners: Temporal and technological dimensions. *Journal of Computer-Mediated Communication*, 6, 1.

Heaps, H. S, (1978). *Information retrieval, computational and theoretical aspects*, Academic Press, 1978.

Hibernate (2008). Software package, Available from web-site (<http://www.hibernate.org>), Accessed: 01. December 2008.

Hiltz, S.R. & Turoff, M. (1993). *The network nation: Human communication via computer*. MIT Press, Cambridge, Mass., 1993.

Ho, S., Encouraging Online Participation? Proceedings of the Eleventh Annual Teaching and Learning Forum 2002, Perth, Australia, 2002.

Hoadley, C. M. & Enyedy, N. (1999). Between information and collaboration: Middle spaces in computer media for learning. in Hoadley, C. and Roschelle, J. eds. *CSCLE '99: Proceedings of Computer Supported Collaborative Learning 1999*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1999, 242-250.

Hoadley, C. M. & Kilner, P. G. (2005), Using Technology to Transform Communities of Practice into Knowledge-Building Communities, *ACM SIGGROUP Bulletin archive Volume 25, Issue 1 (January 2005) Special issue on online learning communities* Pages: 31 – 40, New York, NY, USA.

Hoadley, C. M. & Linn, M. C. (2000). Teaching Science through online, peer discussion: SpeakEasy in the Knowledge Integration Environment, *International Journal of Science*, 22(8):839-857

Jackson, P. (2002). *Natural Language Processing for Online Applications: Text Retrieval, Extraction and Categorization*, John Benjamins Publishing Company, Philadelphia, PA, USA.

Jforum website (2008). Web-site (<http://www.jforum.net>), Accessed: 01. December 2008

Joachims, T., Freitag, D., & Mitchell, T. M. (1997). WebWatcher: A Tour Guide for the World Wide Web. In *Proc. of the 15th International Joint Conference on Artificial Intelligence (IJCAI'97)*, pages 770–777, 1997.

Kanejiya, D., Kumar, A. & Prasad, S. (2003). Automatic evaluation of students' answers using syntactically enhanced LSA. *Human Language Technology Conference, proceedings of the HLT-NAACL 03 workshop on Building educational applications using natural language processing - Volume 2*, Pages: 53 – 60, 2003.

Kang, M. H. & Byun, P. H. (2001). A conceptual framework for a web-based knowledge construction support system. *Educational Technology*, 41, 1, 48-53.

Kazmer, M. M. (2004). How technology affects students' departures from online learning communities. *ACM SIGGROUP Bulletin* Volume 25, Issue 1 (January 2005). Special issue on online learning communities Pages: 25 – 30.

Kazmer, M. M., & Haythornthwaite, C. (2001). Juggling multiple social worlds: Distance students online and offline. *American Behavioral Scientist*, 45, 3 (2001), 510-529.

Kelly, D. & Teevan, J. (2003). Implicit Feedback for Inferring User Preference: A Bibliography. *SIGIR Forum*, 37(2):18–28, 2003.

Kent, A., Berry, M., Leuhns, F.U. & Perry, J.W. (1955). Machine Literature Searching VIII: Operational Criteria for Designing Information Retrieval Systems, *American Documentation*, Vol.6, No.2, pp.93-101.

Kintsch, E., Steinhart, D., Stahl, G., Matthews, C., & Lamb, R. (2000). The LSA Research Group Developing summarization skills through the use of LSA-backed feedback, *Interactive Learning Environments*, 8 (2), (2000) 87-109.

Kleinberg, J. M. (1999). Authoritative Sources in a Hyperlinked Environment. *Journal of the ACM*, 46(5): pages 604–632, 1999.

Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R., & Riedl, J. (1997). GroupLens: Applying Collaborative Filtering to Usenet News. *Communications of the ACM*, 40(3):77–87, 1997.

Korfhage, R. R. (1997). *Information Storage and Retrieval*, John Wiley & Sons, New York.

Korkea-Aho, M. (2000). Context-aware applications survey. Technical Report, Dept of Computer Sciences, Helsinki Univ. of Technology, April 2000,

Krishnan, V. & Chang, S. G. (2000). Customized Internet Radio. In Proc. Of the 9th International World Wide Web Conference (WWW9), pages 609–618, 2000.

Lapata, M., & Barzilay, R. (2005). Automatic evaluation of text coherence: Models and representations. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence*.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211-240.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes*, 25, 259-284.

Lawson, P. P., Alpert, E., Bland, C. G., Carswell, L., Cizmar, D., Dewitt, J., Dumitru, M., Fahraeus, E. R., Scott, K. (1997). *The Web and Distance Learning: What Is Appropriate and What Is Not*, Working Group Reports and Supplementary Proceedings SIGCSE/SIGCUE ITiCSE'97.

Leacock, C., & Chodorow, M. (1998). Combining local context and WordNet sense similarity for word sense identification. In *WordNet, An Electronic Lexical Database*. The MIT Press.

Lee, D. L., Chuang, H & Seamons, K. (1997). Document ranking and the vector-space model, *IEEE Software*, Vol.14(2), p.67-75, 1997.

Lefoe, G. (1998). *Creating constructivist learning environments on the Web: The challenge in higher education*. ASCILITE'98, Wollongong, NSW.

Lesk, M. (1986). Automatic sense disambiguation using machine readable dictionaries: How to tell a pine cone from an ice cream cone. In *Proceedings of the SIGDOC Conference 1986*.

Liaw, S.-S. & Huang, H.-M. (2000). Enhancing interactivity in Web-based instruction: A review of the literature, *Educational Technology*, 40 (2000) 41-45.

Lieberman, H. (1995) *Letizia: An Agent That Assists Web Browsing*. In *Proc. of the 14th International Joint Conference on Artificial Intelligence (IJCAI '95)*, pages 924–929, 1995.

Lieberman, H. (1997). *Autonomous Interface Agents*. In *Proc. of the Conference on Human Factors in Computing Systems (CHI '97)*, pages 67–74, 1997.

Lin, C., & Hovy, E. (2003). Automatic evaluation of summaries using n-gram co-occurrence statistics. In *Proceedings of Human Language Technology Conference* .

Lowder, J. & Hagan, D. (1999). *Web-based student feedback to improve learning*. *Innovation and Technology in Computer Science Education (ITiCSE'99)*, Cracow, Poland, 1999, 151-154.

Malone, T. & Lepper, M. (1987). Making learning fun: A taxonomy of intrinsic motivations of learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning, and instruction: Conative and affective process analyses*, 1987, Vol. 3, pp. 223-253.

Marwick, A. D. (2001). Knowledge management technology. *IBM Systems Journal*, 40(4), 814-830.

Mazza, R. & Dimitrova, V. (2004). Visualising student tracking data to support instructors in web-based distance education, in 'WWW Alt. '04: Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters', ACM Press, New York, NY, USA, pp. 154- 161.

McKenzie, W. and Murphy, D., "I hope this goes somewhere": Evaluation of An Online Discussion Group, *Australian Journal of Educational Technology*, 16 (2000) 239 -257.

Melis, E., Andrès, E., Büdenbender, J., Frishauf, A., Goguadse, G., Libbrecht, P., Pollet, M., & Ullrich, C. (2001). ActiveMath: A web-based learning environment. *International Journal of Artificial Intelligence in Education*, 12, 4 (2001), 385-407.

Melville, P., Mooney, R. J., & Nagarajan, R., 2002. Content-Boosted Collaborative Filtering for Improved Recommendations. In *Proc. Of the 18th National Conference on Artificial Intelligence (AAAI2002)*, pages 187–192, 2002.

Microsoft Office (2008). Software package, Available from web-site (<http://office.microsoft.com>), Accessed: 01. December 2008.

Microsoft Windows (2008). Software package, Available from web-site (<http://www.microsoft.com/windows>), Accessed: 01. December 2008.

MiddlegenIDE (2008). Software package, Available from web-site (<http://boss.bekk.no/boss/middlegen>), Accessed: 01. December 2008.

Merten C & Conati C. (2006). Eye-Tracking to Model and Adapt to User Meta-cognition in Intelligent Learning Environments. In proceedings of International Conference on Intelligent User Interfaces (IUI'06), Sydney, Australia, ACM press, 2006.

Milne, D. (2007). Computing Semantic Relatedness using Wikipedia Link Structure. Proceedings of the New Zealand Computer Science Research Student conference (NZCSRSC'07), Hamilton, New Zealand.

Molnar, A. (1997), 'Computers in Education: A Brief History', THE Journal (Technological Horizons In Education) 24(11).

Morita, M., & Shinoda, Y. (1994). Information Filtering Based on User Behavior Analysis and Best Match Text Retrieval. In Proc. of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '94), pages 272–281, 1994.

Muthukumar S.L. & Hedberg J. G. (2004). A Knowledge Web Portal For Teacher Professional Development And Dissemination Of Research Into Pedagogy And Practice. Retrieved December 23, 2004, from <http://www.tlinc.com/articl62.htm>

MySQL (2008). Software package, Available from web-site (<http://www.mysql.com>), Accessed: 01. December 2008.

Nagypál, G. (2005). Improving information retrieval effectiveness by using domain knowledge stored in ontologies. OTM Workshops, 2005, LNCS 3762.

Oda, T., Satoh, H., & Watanabe, S., 1998. Searching deadlocked Web learners by measuring similarity of learning activities. In: Proc. of Workshop "WWW-Based Tutoring" at 4th International Conference on Intelligent Tutoring Systems (ITS'98), (San Antonio, TX, August 16-19, 1998), available online at <http://www.sw.cas.uec.ac.jp/~watanabe/conference/its98workshop1.ps>

Palloff, R. M. & Pratt, K. (1999). Building learning communities in cyberspace: Effective strategies for the online classroom. San Francisco, Jossey-Bass Inc.

Papineni, K., Roukos, S., Ward, T. & Zhu, W. (2002). Bleu: a method for automatic evaluation of machine translation. In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics.

Paralic, J. & Kostial, I. (1999). Ontology-based Information Retrieval. Conference Proceedings of 14th International Conference on Information and Intelligent systems (IIS), pp6, Varazdin, Croatia, 2003.

PhpBB Website (2008). Available from web-site <http://www.phpbb.com>, Accessed: 01.December 2008.

Picard, R. W. (1999). Affective Computing for HCI. In Proceedings HCI, Munich, Germany, 1999.

Popescul A., Ungar, L. H., Pennock, D. M., Lawrence, S. (2001). Probabilistic Models for Unified Collaborative and Content-Based Recommendation in Sparse-Data Environments. UAI '01: Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence.

PunPP Website (2008). Available from web-site <http://punbb.informer.com/>, Accessed: 01.December 2008.

Reeves, T. C. (2003). Storms Clouds on the Digital Education Horizon. Journal of Computing in Higher Education, Fall 2003, Vol. 15(1), 3-26.

Resnick, P., Iacovou, N., Suchak, M., & Bergstorm, J. R. P. (1994). GroupLens: An Open Architecture for Collaborative Filtering of Netnews. In Proc. of the ACM 1994 Conference on Computer Supported Cooperative Work (CSCW '94), pages 175–186, 1994.

Rheingold, H. (1994). *The virtual community: homesteading on the electronic frontier*. HarperPerennial, New York, NY, 1994.

Riel, M., & Polin, L. (2004). Online learning communities: Common ground and critical differences in designing technical environments. In Barab, R. Kling & J. Gray H. (Eds.), *Designing for virtual communities in the service of learning* (pp. 16-50). Cambridge: Cambridge University Press.

Rios, A., Millán, E., Trella, M., Pérez, J.L. & Conejo, R. (1999). Internet based evaluation system. In: Lajoie, S.P. & Vivet, M. (eds.) *Artificial Intelligence in Education: Open Learning Environments*. IOS Press, Amsterdam, 1999, 387-394.

Robertson, S. (2002). Comparing the Performance of Adaptive Filtering and Ranked Output Systems, *Information Retrieval*, Vol.5, pp.257 - 268.

Robertson. S. E. (1981). The methodology of information retrieval experiment, *Information Retrieval Experiment*. In K. Sparck Jones, Ed. Chapt. 1, pp. 9-31. Butterworths, 1981.

Robertson, S. E. (1995). In *Overview of the Third Text Retrieval Conference (TREC-3)*. Edited by D.K. Harman. Gaithersburg, MD: NIST, 1995 (NIST Special Publication 500-225).

Rocchio, J. (1971). Relevance feedback information retrieval. In Gerard Salton (ed.): *The Smart Retrieval System — Experiments in Automatic Document Processing*, pp. 313–323. Prentice-Hall, Englewood Cliffs, N.J., 1971.

Rovai, A. P. (2003). Relationships of communicator style, personality-based learning style, and classroom community among online graduate students. *Internet and Higher Education*, 6, 347-363.

Saint-Onge, H. & Wallace, D. (2003). *Leveraging communities of practice for strategic advantage*. Butterworth-Heinemann, Boston, 2003.

- Salton, G. (1983). *Introduction to Modern Information Retrieval*, McGraw-Hill, 1983.
- Salton, G. & Buckley, C. (1996). Term Weighting Approaches in Automatic Text Retrieval, *Information Processing and management*, Vol32(4), p. 431-443, 1996.
- Salton, G. & Lesk, M. (1971). *Computer evaluation of indexing and text processing*. Prentice Hall, Inc. Englewood Cliffs, New Jersey. 143–180.
- Salton, G., Singhal, A., Mitra, M. & Buckley, C. (1997). Automatic text structuring and summarization. *Information Processing and Management* 2(32).
- Sanderson, M., & Zobel, J. (2005). Annual ACM Conference on Research and Development in Information Retrieval. Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval. Salvador, Brazil, SESSION: Evaluation Pages: 162 - 169.
- Saracevic, T. (1995). Evaluation of evaluation in information retrieval. Annual ACM Conference on Research and Development in Information Retrieval. Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval, Seattle, Washington, United States , Pages: 138 – 146.
- Sarwar, B. M., Karypis, G. & Konstan, J. A. (2000). Analysis of Recommendation Algorithms for E-commerce. In Proc. of the 2nd ACM Conference on Electronic Commerce (EC '00), pages 158–167, 2000.
- Schutze, H. (1998). Automatic word sense discrimination. *Computational Linguistics* 24(1):97–124.
- Schafer, J. B., Konstan, J. A. & Riedl, J. (2002). Meta-recommendation Systems: User-controlled Integration of Diverse Recommendations. In Proc. of the 11th International Conference on Information and Knowledge Management (CIKM '02), pages 43–51, 2002.

Shardanand, U. & Maes, P. (1995). Social Information Filtering: Algorithms for Automating “Word of Mouth”. In Proc. of the Conference on Human Factors in Computing Systems (CHI '95), pages 210–217, 1995.

Sheard, J. (2004). Electronic Learning Communities: Strategies for Establishment and Management, ITiCSE'04, June 28 – 30, 2004, Leeds, United Kingdom.

Sheard, J., Miller, J., & Ramakrishnan, S. (2003). Web-based Discussion Forums: The Staff Perspective, ACM SIGCSE Bulletin, Proceedings of the 8th annual conference on Innovation and technology in computer science education, Volume 35 Issue 3.

Sheard, J., Postema, M. & Markham, S. (2001). Time versus utility: What IT staff say about the educational use of the Web. EdMedia 2001, Tampere, Finland.

Slavin, R.E. (1983), Cooperative learning, Longman, New York.

Smestad, O. (2006). Fostering inter institutional knowledge sharing among students. Massey University Palmerston North Campus Library, New Zealand.

Soller, A. & Lesgold, A. (2003). A computational approach to analysing online knowledge sharing interaction. In: Hoppe, U., Verdejo, F. & Kay, J. (eds.) Artificial intelligence in education: Shaping the future of learning through intelligent technologies. IOS Press, Amsterdam, 2003, 253-260.

Sonnenreich, W, & Macinta, T. (1998). WEB DEVELOPER.COM GUIDE TO SEARCH ENGINES. U.S.A.. John Wiley & Sons, Inc.

Subclipse (2008). Software package, Available from web-site (<http://subclipse.tigris.org>), Accessed: 01. December 2008.

Subversion (2008). Software package, Available from web-site (<http://subversion.tigris.org>), Accessed: 01. December 2008.

Sue, H., K, Moon, Jang, M. &H. K. (2005). Identifying peer interaction patterns and related variables in community-based learning, Proceedings of the 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years! Taipei, Taiwan. Pages: 657 – 661. Year of Publication: 2005. ISBN:0-8058-5782-6.

Sugiyama, K., Hatano, K. & Yoshikawa, M. (2004). Adaptive web search based on user profile constructed without any effort from users Proceedings of the 13th international conference on World Wide Web New York, NY, USA SESSION: Query result processing Pages: 675 - 684 Year of Publication: 2004 ISBN:1-58113-844-X

Sun JDK (2008). Software package, Available from web-site (<http://java.sun.com/javase>), Accessed: 01. December 2008.

Terveen, L., Hill, W., Amento, B., McDonald, D., & Creter, J. (1997). PHOAKS: A System for Sharing Recommendations. Communications of the ACM, 40(3):59–62, 1997.

Turney, P. (2001). Mining the web for synonyms: PMI-IR versus LSA on TOEFL. In Proceedings of the Twelfth European Conference on Machine Learning (ECML-2001).

Turpeinen, M. & Saari, T. (2004). System Architecture for Psychological Customization of Communication Technology. In Proceedings of the 37th Annual Hawaii International Conference on System Sciences (HICSS'04), Track 7, 2004.

Turpin, A. & Scholer, F. (2006). User Performance versus Precision Measure for Simple Search Tasks. Annual ACM Conference on Research and Development in Information Retrieval Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval Seattle, Washington, USA SESSION: User behaviour and modelling, Pages: 11 – 18.

Vallet, D., Fernández, M. & Castells, P. (2005). An Ontology-Based Information Retrieval Model. *Lecture Notes in Computer Science : The Semantic Web: Research and Applications (2005)*, pp. 455-470.

Van Rijsbergen, C.J. (1979). *Information Retrieval (2nd ed)*, Butterworths, London, 1979.

Venturi, G. & Bessis, N. (2006). User-centred Evaluation of an E-learning Repository. *ACM International Conference Proceeding Series; Vol. 189 Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles. Oslo, Norway, Pages: 203 – 211.*

Vinay, V., Wood, K., Milic-Frayling, N., Cox, I. J. (2005). Comparing relevance feedback algorithms for web search. *Special interest tracks and posters of the 14th international conference on World Wide Web Chiba, Japan POSTER SESSION: Posters Pages: 1052 – 1053.*

W3C (2008). *Web Services, Web-site (<http://www.w3.org/2002/ws>)*, Accessed: 01, December 2008.

Weber, G. & Brusilovsky, P. (2001). ELM-ART: An adaptive versatile system for Web-based instruction. *International Journal of Artificial Intelligence in Education*, 12, 4 (2001), 351-384, available online at http://cbl.leeds.ac.uk/ijaied/abstracts/Vol_12/weber.html

Weber, G., Kuhl, H.-C., & Weibelzahl, S. (2001). Developing adaptive internet based courses with the authoring system NetCoach. In: Bra, P.D., Brusilovsky, P. and Kobsa, A., 2001. (eds.) *Proc. of Third workshop on Adaptive Hypertext and Hypermedia, (Sonthofen, Germany, July 14, 2001)*, Technical University Eindhoven, 35-48, available online at <http://www.wis.win.tue.nl/ah2001/papers/GWeber-UM01.pdf>

Weiser, M. (1991). *The Computer for the Twenty-First Century*. Scientific American Press, September 1991, pp. 94-100.

Weller, N., Pegler, C. & Mason, R. (2005). Use of innovative technologies on an e-learning course. *Internet and Higher Education* 8 (2005) 61-71.

Wenger, E. C. & Snyder, W. M. (2000). Communities of practice: the organizational frontier..*Harvard Business Review*, (Jan-Feb 2000), 139-145.

Whitsed, N. & Weller, M. (2004). 'Learning and teaching – Developing pedagogy for the Internet', *Health Information and Libraries Journal* 21(1), pp. 74-77.

Wijekumar, K. K. & Spielvogel, J. (2006). Intelligent discussion boards: Promoting deep conversations in asynchronous discussion boards through synchronous support. *Journal: Campus-wide Information Systems*, Year 2006, Volume 23, Issue 3, Page 221 – 232.

Wikipedia contributors. (2008). Learning community. In *Wikipedia, the Free Encyclopaedia*. Retrieved 15:37, March 13, 2008, from http://en.wikipedia.org/wiki/Learning_community

Wilson, B. G. & Ryder, M. (1998). Distributed learning communities: An alternative to designed instructional systems. <http://www.cudenver.edu/~bwilson/dlc.html>

Wilhelm, S., Rodehorst, T. K., Young, S., Hensen, L. & Stepan, M. B. (2003). Student' Perspective of the Effectiveness of an Asynchronous On-line Seminar. *Journal of Professional Nursing*, Vol.19, No.5 (September – October), 2003:pp313-319.

Winograd, T. (1989). Groupware: the next wave or just another advertising slogan? in *COMPCON Spring '89. Thirty-Fourth IEEE Computer Society International Conference: Intellectual Leverage (IEEE Cat. No.89CH2686-4)*, (San Francisco, CA, USA, 1989), IEEE Computer. Soc. Press, 198-200.

Woodruff, E. E. (1999). Concerning the Cohesive Nature of CSCL Communities. in *Hoadley, C. & Roschelle, J. eds. Proceedings of Computer Supported Collaborative*

Learning '99 Conference, Lawrence Erlbaum Associates, Mahwah, NJ, 1999, 677-680.

Wu, Y-F. & Chen, X. (2004). Assessing distance learning student's performance: A natural language processing approach to analyzing class discussion messages. *Information Technology: Coding and Computing*, 2004. Proceedings. ITCC 2004. International Conference on Volume 1, Issue , 5-7 April 2004 Page(s): 192 - 196 Vol.

XFree86 (1996). 'BSD – (Modified) Berkeley Software Distribution License', Available from web-site (<http://www.xfree86.org/3.3.6/COPYRIGHT2.html#5>), Accessed: 01, December 2008.

Yang, J. (2005). *Knowledge Management in Web Communication Technologies*. Massey University Palmerston North Campus Library, New Zealand.

Zelikovitz, S. & Hirsh, H. (2005). Using LSI for Text Classification in the Presence of Background Text. *Conference on Information and Knowledge Management. Proceedings of the tenth international conference on Information and knowledge management*. Pages: 113 – 118, 2001.

APPENDIX

Appendix I – Stop-word List

Stop-word list = ("a", "about", "above", "above", "across", "after", "afterwards",
"again", "against", "all", "almost", "alone", "along", "already",
"also", "although", "always", "am", "among", "amongst", "amongst",
"amount", "an", "and", "another",
"any", "anyhow", "anyone", "anything", "anyway", "anywhere", "are", "around",
"as", "at", "back", "be", "became", "because", "become", "becomes", "becoming",
"been", "before", "beforehand", "behind", "being", "below", "beside", "besides",
"between", "beyond", "bill", "both", "bottom", "but", "by", "call", "can", "cannot",
"cant", "co", "con", "could", "couldn't", "cry", "de", "describe", "detail", "do",
"done", "down", "due", "during", "each", "eg", "eight", "either", "eleven", "else",
"elsewhere", "empty", "enough", "etc", "even", "ever", "every", "everyone",
"everything", "everywhere", "except", "few", "fifteen", "fifty", "fill", "find", "fire",
"first", "five", "for", "former", "formerly", "forty", "found", "four", "from", "front",
"full", "further", "get", "give", "go", "had", "has", "hasn't", "have", "he", "hence",
"her", "here", "hereafter", "hereby", "herein", "hereupon", "hers", "herself", "him",
"himself", "his", "how", "however", "hundred", "ie", "if", "in", "inc", "indeed",
"interest", "into", "is", "it", "its", "itself", "keep", "last", "latter", "latterly", "least",
"less", "ltd", "made", "many", "may", "me", "meanwhile", "might", "mill", "mine",
"more", "moreover", "most", "mostly", "move", "much", "must", "my", "myself",
"name", "namely", "neither", "never", "nevertheless", "next", "nine", "no", "nobody",
"none", "none", "nor", "not", "nothing", "now", "nowhere", "of", "off", "often", "on",
"once", "one", "only", "onto", "or", "other", "others", "otherwise", "our", "ours",

“ourselves”, “out”, “over”, “own”, “part”, “per”, “perhaps”, “please”, “put”, “rather”,
“re”, “same”, “see”, “seem”, “seemed”, “seeming”, “seems”, “serious”, “several”,
“she”, “should”, “show”, “side”, “since”, “sincere”, “six”, “sixty”, “so”, “some”,
“somehow”, “someone”, “something”, “sometime”, “sometimes”, “somewhere”,
“still”, “such”, “system”, “take”, “ten”, “than”, “that”, “the”, “their”, “them”,
“themselves”, “then”, “thence”, “there”, “thereafter”, “thereby”, “therefore”, “therein”,
“thereupon”, “these”, “they”, “thick”, “thin”, “third”, “this”, “those”, “though”,
“three”, “through”, “throughout”, “thru”, “thus”, “to”, “together”, “too”, “top”,
“toward”, “towards”, “twelve”, “twenty”, “two”, “un”, “under”, “until”, “up”, “upon”,
“us”, “very”, “via”, “was”, “we”, “well”, “were”, “what”, “whatever”, “when”,
“whence”, “whenever”, “where”, “whereafter”, “whereas”, “whereby”, “wherein”,
“whereupon”, “wherever”, “whether”, “which”, “while”, “whither”, “who”,
“whoever”, “whole”, “whom”, “whose”, “why”, “will”, “with”, “within”, “without”,
“would”, “yet”, “you”, “your”, “yours”, “yourself”, “yourselves”, “the”);

Appendix II – Questionnaire

Part A (All participants)

Section 1: Search Function

To what extent do you agree with each of the following statements regarding search function on the preview page in this discussion forum system?

1. It is easy to use the auto-reply function. Answer: _	5	4	3	2	1
2. The auto-reply function is a useful functionality. Answer: _	5	4	3	2	1
3. It is easy to use the user feedback function. Answer: _	5	4	3	2	1
4. The user feedback function is a useful functionality. Answer: _	5	4	3	2	1

Section 2: advantages and disadvantages of the system

5. From your point of view, what are your personal expectations in using the prototype discussion forum? (Choose one or more options: 1, 2, 3) **Answer:** _

Good quality and highly related post contents

Faster to get useful and valuable information

Easily share information

Other expectations:

6. What are the reasons why you would like to use this prototype discussion forum? (Choose one or more options: 1, 2, and 3) **Answer:** _

- More useful and related information could be quickly found with the 'auto-reply' function
- Easy to use
- User feedback function is useful

Other reasons:

7. What are the benefits you believe the prototype discussion forum can bring to learners? (Choose one or more options: 1, 2, 3). **Answer:** _

- More efficient and effective retrieval of useful information
- Knowledge being easily shared and used relevant information.
- Less burdens for teachers to manage the discussions

Other benefits:

8. What are the negative issues that you find annoying in using the prototype discussion forum? (Choose one or more options: 1, 2, 3). **Answer:** _

- The search approach adopted in the prototype forum could be improved to retrieve answers that more exactly matched the submitted posts.
- User feedback function is helpless
- Difficult to understand how to use the system

Other drawbacks:

Section 6: General Questions

9. How familiar are you with discussion forums? **Answer:** _

- 1 Never used

- 2 Really used
- 3 Often used.

10. How familiar are you with search engines? **Answer:** _

- 4 Never used
- 5 Really used
- 6 Often used.

11. Which one is your first approach when you have a question regarding your course content in the below options? **Answer:** _

- 7 Textbook
- 8 Discussion forum
- 9 Search Engine.

Part B (Tutor only)

1. How familiar are you with the instructor role in discussion forums? **Answer:** _

- 1 Never
- 2 A little bit
- 3 Very familiar.

2. The major issues for instructors in using discussion forums? **Answer:** _

- Difficult to use
- The time and effort spent on forum management
- Difficult Questions

2. The reasons why an instructor might like to use this prototype discussion forum?

Answer: _

- More valuable information
- Easing the management burden
- Easy to use

2. The benefits instructors believe the prototype discussion forum can bring to learners and instructors? **Answer:** _

- Improvement in effective communication between learners, and between instructors and learners
- More valuable information
- Easily reuse information

3. What needs to be improved in the prototype forum? **Answer:** _

- The information retrieval approach could be improved for accuracy
- The forum posts could be better categorised into groups such as query, answer and comments.
- Not difficult to understand how to use the system