A TACIT HEALTHCARE KNOWLEDGE EXPLICATION INFO-STRUCTURE USING CONTRIVED KNOWLEDGE ACQUISITION AND REPRESENTATION APPROACHES

END OF PROJECT REPORT

USM SHORT TERM GRANT

PROJECT LEADER:
CHEAH YU-N

BAHAGIAN PENYELIDIKAN DAN PEMBANGUNAN

CANSELORI

UNIVERSITI SAINS MALAYSIA

Laporan Akhir Projek Penyelidikan Jangka Pendek

1) **Nama Penyelidik:** Dr. Cheah Yu-N

Nama Penyelidik-Penyelidik Lain *(Jika berkaitan)* : Dr. Vincent Khoo Kay Teong

2) **Pusat Pengajian/Pusat/Unit:** Pusat Pengajian Sains Komputer

3) **Tajuk Projek Penyelidikan:** A Tacit Healthcare Knowledge Explication Info-Structure Using Contrived Knowledge Acquisition and Representation Approaches

4) (a) **Penemuan Projek/Abstrak**

*(Perlu disediakan makluman di antara 100 – 200 perkataan di dalam **Bahasa Malaysia dan Bahasa Inggeris**. Ini kemudiannya akan dimuatkan ke dalam Laporan Tahunan Bahagian Penyelidikan & Pembangunan sebagai satu cara untuk menyampaikan dapatan projek tuan/puan kepada pihak Universiti).*

Projek ini telah menghasilkan suatu Info-Struktur Pengeksplikasian Pengetahuan Kesihatan Tersirat yang mampu memperolehi, menyimpan dan menyebarkan pengetahuan kesihatan tersirat untuk digunakan oleh para pakar dan doktor kesihatan supaya perkhidmatan kesihatan yang berkualiti dapat diberi secara berterusan. Dalam penyelidikan ini, kami telah:

1. Mendefinisikan dan membangunkan suatu Perwalikan Pengetahuan Kesihatan Tersirat menggunakan struktur senario yang diubahsuai.
2. Mendefinisikan dan membangunkan suatu Mekanisma Pengeksplikasian Pengetahuan Kesihatan Tersirat berasaskan teknik pemerolehan pengetahuan secara tak terus, iaitu dengan menjana 'cabaran' di beberapa lokasi persoalan dalam struktur senario.

Struktur senario yang diubahsuai ini lebih mudah berbanding dengan senario asal untuk pemerolehan pengetahuan tersirat dan ia direkabentuk untuk digunakan dalam bidang kesihatan. Struktur senario yang diubahsuai ini dapat menampung binaan-binaan Pembelajaran Berasaskan Masalah manakala Info-Struktur Pengeksplikasian Pengetahuan Kesihatan Tersirat juga memuatkan fungsi-fungsi untuk menyokong pemerolehan pemicu yang kaya dengan pengetahuan tersirat untuk memperkayakan Pembelajaran Berasaskan Masalah.

Project Findings/Abstract

The project has produced a Tacit Healthcare Knowledge Explication Info-Structure that is designed to acquire, store and disseminate tacit healthcare knowledge to be used by healthcare specialists, experts and practitioners to ensure the provision and continuation of expert-quality healthcare services. In this research, we have:

1. Defined and developed a Tacit Healthcare Knowledge Representation Formalism using a modified scenario structure.
2. To define and develop a Tacit Healthcare Knowledge Explication Mechanism based on contrived knowledge acquisition techniques, i.e. by generating challenges at various point of interrogation in the scenario structure.

The modified scenario structure is a more streamlined version of the original tacit knowledge explication scenario and is designed to suit the healthcare environment better. This modified scenario structure has been utilised to incorporate Problem-Based Learning constructs while the Tacit Healthcare Knowledge Explication Info-Structure also incorporates functions to support the acquisition of tacit knowledge-rich triggers to enhance Problem-Based Learning.

(b) **Senaraikan Kata Kunci yang digunakan di dalam abstrak:**

Bahasa Malaysia	**Bahasa Inggeris**
Pengetahuan tersirat	Tacit knowledge
Pemerolehan pengetahuan tak terus	Contrived knowledge acquisition
Pengeksplikasian pengetahuan	Knowledge explication
Pengurusan pengetahuan	Knowledge management
Informatik kesihatan	Health informatics

5) Output Dan Faedah Projek

(a) Penerbitan *(termasuk laporan/kertas seminar)*
(Sila nyatakan nejis, tajuk, pengarang, tahun terbitan dan di mana telah diterbit/dibentangkan).

1. Y.-N. Cheah, F.A. Rashid and S.S.R. Abidi. (2003). An Approach to Enrich Online Medical Problem-Based Learning with Tacit Healthcare Knowledge, Medical Informatics Europe (MIE 2003), St. Malo, France.

2. N.A.A. Alim, W.N.W. Hasan, F.A. Rashid and Y.-N. Cheah. (2003). Scope of Telehealth Research Projects in Universiti Sains Malaysia, International Conference on Advancement in Science and Technology (iCAST 2003), Kuala Lumpur, Malaysia. (Poster Paper)

3. P.K. Lye. (2004). An Enterprise Memory for Healthcare Knowledge Management. M.Sc. (Kerja Kursus) Laporan Projek, Universiti Sains Malaysia.

(b) Faedah-Faedah Lain Seperti Perkembangan Produk, Prospek Komersialisasi Dan Pendaftaran Paten.
(Jika ada dan jika perlu, sila gunakan kertas berasingan)

Tiada

(c) Latihan Gunatenaga Manusia

(i) ***Pelajar Siswazah:*** Seorang pelajar Sarjana (Sains Komputer) melalui mod kerja kursus telah membantu dalam usaha membangun sistem ini sebagai keperluan projek beliau.

(ii) ***Pelajar Prasiswazah:*** Empat (4) orang pelatih industri telah diambil untuk tempoh lebih kurang 5 bulan (15 Mac 2004 hingga 7 Ogos 2004) untuk membantu dalam usaha penyelidikan berkenaan pengetahuan tersirat dan pengurusan pengetahuan.

(iii) ***Lain-Lain:*** Tiada

6) **Peralatan Yang Telah Dibeli:**

Komputer Intel Pentium P4 2.4 GHz

Pencetak Laser Canon LBP 1120

UNTUK KEGUNAAN JAWATANKUASA PENYELIDIKAN UNIVERSITI

T/TANGAN PENGERUSI
J/K PENYELIDIKAN
PUSAT PENGAJIAN

DR. BAHARI BELATON
Timbalan Dekan
Pengajian Siswazah Dan Penyelidikan
Pusat Pengajian Sains Komputer
Universiti Sains Malaysia
11800 USM, Pulau Pinang

APPENDIX A: RESEARCH BACKGROUND AND METHODOLOGY

1 Research Background

Knowledge management (KM) is becoming an important research area especially in terms of how the knowledge capital of healthcare enterprises can be effectively managed and utilised. KM is very relevant in healthcare especially since KM experts have argued that the strategic, intuitive and experiential knowledge of (healthcare) experts are stored tacitly in the minds of the experts.

Traditional and more direct healthcare knowledge acquisition methods are quite effective in procuring explicit healthcare knowledge but they are relatively ineffective in acquiring a healthcare expert's tacit knowledge and do not adequately address the intuitive and innate characteristics of tacit healthcare knowledge. As a result, tacit healthcare knowledge is often unsuccessfully explicated. Furthermore, tacit healthcare knowledge remains elusive due to the lack of adequate tacit knowledge representation schemes.

Related research/publications:

- Baumard, P. (1999). *Tacit Knowledge in Organizations.* London: Sage Publications.
- Nonaka, I. and Takeuchi, H. (1995). *The Knowledge Creating Company.* New York: Oxford University Press.
- Raven, A. and Prasser, S.G. (1996). Information Technology Support for the Creation and Transfer of Tacit Knowledge in Organizations. *Second Americas Conference on Information Systems (AIS '96)*, Phoenix, Arizona.

2 Research Objectives

The objective of the project is to acquire, store and disseminate tacit healthcare knowledge to be used by healthcare specialists, experts and practitioners to ensure the provision and continuation of expert-quality healthcare services. We aim to achieve this through the definition and development of formalisms and applications that can be effectively applied to the area of healthcare. Therefore, in this project, our main aims are:

- To define and develop a Tacit Healthcare Knowledge Representation Formalism
- To define and develop a Tacit Healthcare Knowledge Explication Mechanism based on contrived (or indirect) knowledge acquisition techniques
- To build a Tacit Healthcare Knowledge Explication Info-Structure with application in healthcare Problem-Based Learning

3 Research Methodology

Research efforts for this project was focused on the task of healthcare knowledge creation, in particular, the explication, i.e. the acquisition and representation, of tacit healthcare knowledge from healthcare experts.

The main premise of our methodology is that tacit healthcare knowledge can best be explicated by provoking healthcare experts to act and apply their knowledge, experience and skills to solve novel or atypical healthcare problems. Therefore, contrived or indirect knowledge acquisition techniques are deemed more suitable as compared to direct knowledge acquisition methods. For this purpose, we have defined and developed a knowledge representation formalism that is suited for tacit healthcare knowledge explication based on existing tacit knowledge explicating scenarios.

This streamlined scenario-based tacit knowledge representation formalism was then used as the basis for the definition and development of a contrived knowledge explication tool. Based on our main premise and the original content of the tacit knowledge base, the knowledge explication mechanism would then generate healthcare scenarios and situations, or challenges, to provoke the healthcare experts into using their tacit knowledge to produce the solutions. To do this, specific points of interrogation are applied.

We have also defined a Tacit Healthcare Knowledge Acquisition Info-Structure (TKAI) that integrates the knowledge acquisition tools with other possible components (e.g. a knowledge validation/crystallisation tool) into a framework that utilises the JAVA platform. The TKAI was then experimented on a Problem-Based Learning environment where existing triggers can be enriched with tacit knowledge.

4 Project Timeline

The research was carried out in 4 phases

Phase I – Literature Review
- Knowledge management
- Contrived knowledge acquisition techniques

Phase II – Tacit Healthcare Knowledge Representation
- Formalism of tacit knowledge
- Conceptual design of contrived knowledge acquisition strategy using this formalism

Phase III – Tacit Healthcare Knowledge Explication
- Implementation of contrived knowledge acquisition strategy

Phase IV – Tacit Healthcare Knowledge Explication Info-Structure
- Consolidating implementation with other possible knowledge-based components

5 Project Outcome

Key deliverables include:
- Tacit Healthcare Knowledge Representation
- Tacit Healthcare Knowledge Base
- Tacit Knowledge Explication Tool
- Tacit Healthcare Knowledge Acquisition Info-Structure (TKAI)

APPENDIX B: CONFERENCE PAPERS

1. Y.-N. Cheah, F.A. Rashid and S.S.R. Abidi. (2003). An Approach to Enrich Online Medical Problem-Based Learning with Tacit Healthcare Knowledge, Medical Informatics Europe (MIE 2003), St. Malo, France.
2. N.A.A. Alim, W.N.W. Hasan, F.A. Rashid and Y.-N. Cheah. (2003). Scope of Telehealth Research Projects in Universiti Sains Malaysia, International Conference on Advancement in Science and Technology (iCAST 2003), Kuala Lumpur. (Poster)

An Approach to Enrich Online Medical Problem-Based Learning with Tacit Healthcare Knowledge

Yu-N Cheah[a], Faridah Abdul Rashid[b] and Syed Sibte Raza Abidi[c]

[a]Health Informatics Research Group, School of Computer Sciences, Universiti Sains Malaysia, 11800 Penang, Malaysia
[b]Department of Chemical Pathology, School of Medical Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia
[c]Faculty of Computer Science, Dalhousie University, Halifax B3H 1W5, Canada

Abstract

Existing Problem-Based Learning (PBL) problems, though suitable in their own right for teaching purposes, are limited in their potential to evolve by themselves and to create new knowledge. Presently, they are based on textbook examples of past cases and/or cases that have been transcribed by a clinician. In this paper, we present (a) a tacit healthcare knowledge representation formalism called Healthcare Scenarios, (b) the relevance of healthcare scenarios in PBL in healthcare and medicine, (c) a novel PBL-Scenario-based tacit knowledge explication strategy and (d) an online PBL Problem Composer and Presenter (PBL-Online) to facilitate the acquisition and utilisation of expert-quality tacit healthcare knowledge to enrich online PBL. We employ a confluence of healthcare knowledge management tools and Internet technologies to bring tacit healthcare knowledge-enriched PBL to a global and yet more accessible level.

Keywords:
Problem-Based Learning; Healthcare Scenarios; Tacit Healthcare Knowledge

1. Introduction

Problem-Based Learning (PBL) is a cognitive paradigm in which a student learns through solving a certain problem. The rationale for posing problems is so that the student is motivated to hold discussions and independently learn new knowledge in order to solve the problem at hand [1]. PBL has been utilised to train medical students since the 1960s in view that students found it difficult to apply the relevant skills and knowledge into the real-life practise of medicine [2].

Presently, medical PBL problems are based on textbook examples of past cases and/or cases that have been transcribed by a clinician. We argue that these existing PBL problems, though suitable in their own right for teaching purposes, are limited in their potential to evolve by themselves and to create new knowledge. Furthermore, healthcare institutions are sometimes faced with the problem of procuring PBL problems. We believe that current healthcare PBL problems can be brought to a higher level of abstraction that will allow healthcare experts to use their tacit knowledge to generate PBL problems that are of added-value with the injection of tacit healthcare knowledge and thus alleviating the problem of acquiring new PBL problems.

Tacit healthcare knowledge is non-formalised healthcare knowledge. It is the kind of knowledge that governs the healthcare experts' skills, common sense and intuitive judgement. Due to the abstract and informal nature of tacit healthcare knowledge, the state-of-affairs vis-à-vis healthcare tacit knowledge acquisition does not commensurate with the importance of tacit healthcare knowledge, thereby resulting in the under-utilisation and non-documentation of such a vital component of the overall healthcare education and delivery system [3].

To address the above issues, we present (a) a tacit healthcare knowledge representation formalism called *Healthcare Scenarios* [4], (b) the relevance of healthcare scenarios in PBL in healthcare and medicine, (c) a novel PBL-Scenario-based tacit knowledge explication strategy and (d) an online PBL Problem Composer and Presenter (PBL-Online) – a sophisticated client-server PBL system that combines the effectiveness of PBL-Scenarios and the Internet – to facilitate the efficient acquisition and utilisation of fresh and dynamic expert-quality tacit healthcare knowledge from healthcare experts in order to enrich medical PBL.

2. Tacit Healthcare Knowledge Representation using *Healthcare Scenarios*

A healthcare scenario is a customised, goal-oriented narration or description of a healthcare situation. It includes the specification of actors, events, outputs and environmental parameters. Healthcare scenarios may be composed of four main components [4, 5, 6]: *Meta-Scenario*, *Scenario-Construct*, *Healthcare Episode* and *Healthcare Event* (see Figure 1). Healthcare scenarios can be represented by a four-tier scheme where Meta-Scenarios are placed at the top level followed by Scenario-Constructs, Healthcare Episodes and Healthcare Events at the bottom level. Just as knowledge bases store rules for healthcare expert systems, the healthcare scenario components are stored in *scenario bases* which also adhere to the same hierarchical scheme.

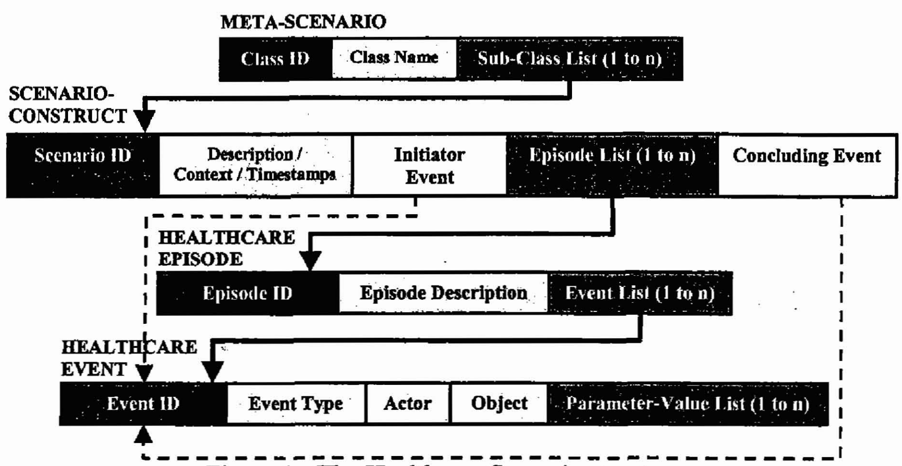


Figure 1 - The Healthcare Scenario structure

Of these four main components, the key components of the healthcare scenario are actually the Scenario-Construct and Healthcare Event components. The Scenario-Construct stores all background information pertaining to individual healthcare scenarios. It also ensures the correct sequencing of Healthcare Episodes and Events. The Healthcare Event stores the details of a healthcare scenario including the Event Type (which could be *Normative*, *Obstacle* or *Action*) and the various event parameters and their values.

Therefore, a healthcare scenario can be viewed as a sequence of real or hypothetical healthcare events encountered by healthcare experts together with their responses or actions. An example of a healthcare scenario is the procedures (investigations, differential diagnoses, treatment, etc.) undertaken to treat a patient with urinary tract infection.

3. The Relevance of Healthcare Scenario in Problem-Based Learning

We have observed that current medical PBL materials are also in the form of problem narrations. Each problem is presented in sections called *triggers* over the course of a certain time period, usually one week. Medical PBL problems/triggers typically begin with some background description of the problem followed by events that happened (e.g. complications, improvement, etc.), actions that were carried out by a clinician, laboratory test results and other details.

With the above details in mind, we are able to adapt our healthcare scenario representation to suit the dynamics of medical PBL. We have noticed that it is common to have two or three triggers pertaining to a particular PBL problem which can be decomposed into a number of sub-problems similar to that of our Healthcare Episodes. Therefore, we introduce another component to our healthcare scenario structure to accommodate these triggers. The resultant structure is what we will call a *PBL-Scenario* with five main components: *PBL Meta-Scenario*, *PBL Scenario-Construct*, *PBL Trigger*, *PBL Episode* and *PBL Event.*

4. Utilising PBL-Scenarios: A Tacit Healthcare Knowledge Acquisition Method

Our tacit healthcare knowledge acquisition exercise distinguishes between three types of PBL-Scenarios:

1. *Solved PBL-Scenarios*: These are PBL-Scenarios that define actual clinical situations/problems that have already been encountered and solved by healthcare experts. Such scenarios are records of complete descriptions of typical or routine clinical situations together with the expert's responses.
2. *Challenge PBL-Scenarios*: These are PBL-Scenarios that represent atypical situations and are posed to healthcare experts as challenges to their expertise. We argue that tacit knowledge is best explicated when experts are required to solve atypical problems in line with contrived knowledge acquisition approaches [7]. Challenge PBL-Scenarios are derived from existing Solved PBL-Scenarios or Solved Challenge PBL-Scenarios by way of selecting a *Point of Interrogation (POI)*, i.e. a distinct point in the PBL-Scenario between two PBL Event of type *Obstacle* or *Normative* and a PBL Event of type *Action.* The result is a Challenge PBL-Scenario which is then presented to a healthcare expert for the explication of his/her tacit healthcare knowledge (see Figure 2).
3. *Solved Challenge PBL-Scenarios*: These scenarios originate from Challenge PBL-Scenario that have been completed or solved by a healthcare expert and are deemed as the encapsulation of the healthcare expert's tacit knowledge (see Figure 2).

	Trigger	Episode	Event	Event Description	
PBL-Scenario 5 month old boy with history of fever for 6 days prior to admission	Trigger		Initiator Event EV0001 Obstacle	Patient admitted with fever, decreased appetite and vomiting several times a day. Frequency of passing urine is decreased.	Challenge
		Episode	EV0002 Action	Perform anthropometry.	
			EV0003 Obstacle	Weight = 4.8 kg (p10 = 6kg). Length = 55cm (p10 = 61cm). Head circumference = 40cm (p10 = 40.5cm).	
			Action	Check condition.	POI
			Obstacle	Patient is ill looking, lethargic and showing signs of mild to moderate dehydration.	Expert's Response + Tacit Knowledge
	. . .				
			Concluding Event Normative	Patient responding well to treatment and is discharged. Arrange a review in 1 month.	

Figure 2 - The expert's response to the Challenge in a Solved PBL-Scenario

The Solved PBL-Scenarios are then subjected to a process of peer evaluation. The discussion on this process is beyond the scope of this paper. However, suffice to say that this process allows healthcare experts to assign ratings to the PBL-Scenario components based on their relevance and usefulness. This leads to the crystallisation of healthcare tacit knowledge [4].

5. PBL-Scenario-based PBL-Online System

To enrich current PBL problem repositories and to facilitate the explication of tacit healthcare knowledge, we are in the midst of developing an online *PBL Problem Composer and Presenter (PBL-Online)* – an intelligent web-based system that allows healthcare experts to systematically respond to atypical PBL-Scenarios. To achieve this, PBL-Online presents web-based knowledge elicitation forms (see Figure 3) which contain attributes that correspond to the PBL-Scenario structure. These prompt healthcare experts to provide information or suggest probable and realistic values to the various PBL-Scenario attributes presented in the forms thus creating new PBL problems that are of added-value and enriched with expert-quality tacit knowledge. PBL-Online is a step-forward compared to the stand-alone version of our legacy Healthcare Scenario Composer [4, 8].

PBL-Online also includes healthcare ontology/thesaurus-based consistency-checking mechanisms to standardise the input of the healthcare experts because, as it turns out, there are no restrictions on the terms used by healthcare experts [9]. For our purpose, we will incorporate the MeSH thesaurus.

In addition to facilitating the explication of tacit healthcare knowledge, PBL-Online functions as a full-fledged medical PBL problem presentation system that allows coordinators, facilitators and students to prescribe and view the scheduled PBL problems and triggers.

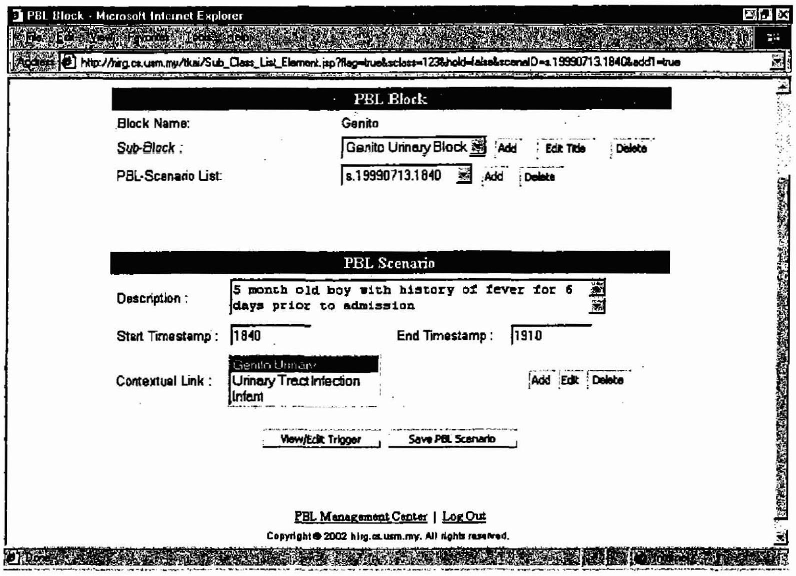


Figure 3 - Sample PBL Scenario-Construct form

5.1 PBL-Online for Students

Upon logging-in to PBL-Online, the students would first be presented with the PBL trigger for the session in narrative format. The students can then input notes in the form of key points, hypotheses and learning issues in the text boxes provided. Students can also view the PBL trigger's recommended key points. The key points are derived directly from the hierarchical Solved PBL-Scenario for the corresponding PBL trigger narration. Finally, students are allowed to answer a number of discussion questions based on the PBL trigger. Their answers will be evaluated by the PBL facilitator and students are allowed to proceed to the next PBL trigger only if they achieved a predetermined score.

5.2 PBL-Online for Coordinators and Facilitators

PBL coordinators are allowed to prescribe PBLs to the students. They can also edit the narrations, recommended key points (i.e. the Solved PBL-Scenario) and discussion questions. Facilitators monitor the PBL sessions and evaluate the students answers once the deadline for answers has passed (see Figure 4).

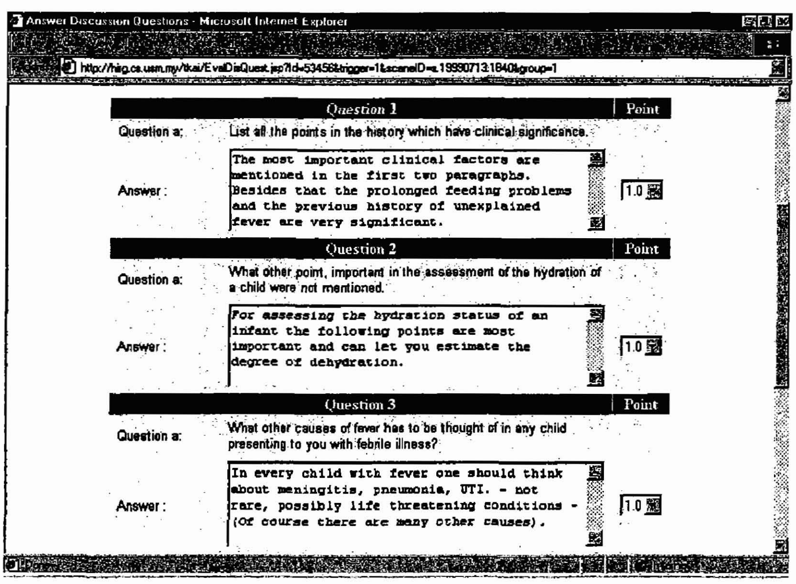


Figure 4 - Student evaluation form

6. Concluding Remarks

We believe that the PBL-Online system presented here does provide an all-rounded facility to enriching the overall PBL initiative through the addition of explicated tacit healthcare [redacted] healthcare knowledge with the aim of extracting the essence of the healthcare experts' problem-solving methodology. We also presented the PBL-Online system and its components that function as both a PBL presentation system as well as an authoring tool to enrich the PBL-Scenario base with tacit knowledge-enriched PBL problems. PBL-Online employs a confluence of healthcare knowledge management tools and Internet technologies to bring tacit healthcare knowledge-enriched PBL to a global and yet more accessible level.

7. References

[1] School of Medicine, Queen's University, Ontario, Canada. Problem-Based Learning at Queen's: Introduction. http://meds.queensu.ca/medicine/pbl/pblhome1.htm.

[2] Gooding K. Problem Based Learning Online. *Net*Working 2001* (Workshop presentation), Brisbane, Australia, 2001.

[3] Cheah Y-N and Abidi SSR. A Scenarios Mediated Approach for Tacit Knowledge Acquisition and Crystallisation: Towards Higher Return-On-Knowledge and Experience. In: Reimer U, ed. *Proceedings of the Third International Conference on Practical Aspects of Knowledge Management (PAKM 2000),* Basel, Switzerland, 2000.

[4] Cheah Y-N and Abidi SSR. The Role of Information Technology in the Explication and Crystallization of Tacit Healthcare Knowledge. *Health Informatics Journal* 2001: 7 (3/4), pp. 158-167.

[5] Schultz AC, Grefenstette JJ and De Jong GA. Learning to Break Things: Adaptive Testing of Intelligent Controllers. *Handbook on Evolutionary Computation*, IOP Publishing Ltd. and Oxford University Press, 1997, Chapter. G3.5.

[6] Potts C, Takahashi K and Anton A. Inquiry-Based Scenario Analysis of System Requirements. *International Conference on Requirements Engineering (ICRE'94)*, Colorado Springs, Colorado, 1994.

[7] Shadbolt NR and Burton AM. Knowledge Elicitation. In: J.R. Wilson and E.N. Corlett, eds., *Evaluation of Human Work: A Practical Ergonomics Methodology*, London: Taylor and Francis; 1990, pp. 321-345.

[8] Cheah Y-N and Abidi SSR. Health Expert's Tacit Knowledge Acquisition and Representation Using Specialised Healthcare Scenarios. In: Hasman A, Blobel B, Dudeck J, Engelbrecht R, Gell G and Prokosch HU, eds. *Medical Infobahn for Europe (MIE 2000)*, Amsterdam: IOS Press, 2000.

[9] Chandrasekaran B and Josephson JR. What Are Ontologies, and Why Do We Need Them? *IEEE Intelligent Systems* 1999: 14 (1), pp. 20-26.

8. Address for correspondence

Yu-N Cheah
Health Informatics Research Group, School of Computer Sciences,
Universiti Sains Malaysia, 11800 Penang, Malaysia
E-mail: yncheah@cs.usm.my

SCOPE OF TELEHEALTH RESEARCH PROJECTS IN UNIVERSITI SAINS MALAYSIA

Nor Azlina A.B.A., Wan N.W.H., Faridah A.R., Cheah Y.N.

Telehealth Research Group, School of Medical Sciences,
Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia

INTRODUCTION

Telehealth in the broadest sense, refers to the application of any Information and Communications Technology (ICT) to provision of healthcare when distance separates the participants. Telehealth research and development at the School of Medical Sciences, USM, focuses on an electronic health system that is equitable, affordable, efficient, technologically appropriate, environmentally adaptable and consumer friendly. It is in line with the MSC Telehealth Flagship application and includes projects such as Mass Customised, Personalised Health Information and Education (MCPHIE), Lifetime Health Plans (LHP), Continuing Medical Education (CME), and Teleconsultation.

Currently, there are 6 on-going research projects under Telehealth. Each project is led by a lecturer who acts as Project Leader. One or more Research Officer (RO) or Research Assistant (RA) assist each Project Leader. All projects are carried out in collaboration with the researchers of the Health Informatics Research Group at the School of Computer Sciences, USM Main Campus, in Penang.

Project 1: An Info-Structure for the Acquisition of Expert Tacit Knowledge: Development of PBL Online for Phase 2 Medicine, USM

Project 2: A Clinical Case Acquisition and Transcription Info-Structure: The Development of Multimedia XML-based EPR for the Acquisition, Cleaning and Transformation to Clinical Cases

Project 3: Development of an Agent-Based Data-Mining Info-Structure: The Development of Multimedia XML-based EPR for The Acquisition, Cleaning and Transformation to Clinical Cases

Project 4: Web-Based Personalized Wellness Maintenance Info-Structure: Towards An Intelligent Tele-Healthcare Environment Offering Person-Centric and Wellness-Oriented Decision Support

Project 5: An Intelligent Healthcare Assistant for Knowledge Adaptation and Retrieval from Healthcare Knowledge Repositories: A Web-Based Radiological Assistant In Clinical Practice.

Project 6: An Info-Structure for the Delivery of PHI for Respiratory Problems.

OBJECTIVES

Two major objectives of present Telehealth research are:

1. To create the necessary systems which can transfer the manual health information and records to the electronic form.

2. To save and record healthcare expertise and knowledge by using IC_.

METHODOLOGY

Two main research methods are used. These are data analyses and key-in data.

Data Analyses
There are two data sets the project collects and utilizes for analyses. These are manual medical records and online surveys.

Key-in Data
This is the manual key-in of contents into the prepared key-point template of the various systems created.

RESULTS

Six novel web-based templates (end-user interfaces) have been created and they are at various stages of refinement.

1. PBL Online using the PBL-Scenario System.

   We are at the second stage of further refining the PBL-Scenario System that has been developed. We propose to perform a test run and evaluation of the PBL-Scenario System some time during the Genitourinary Block in January-February 2004

APPENDIX C: PROJECT REPORT

P.K. Lye. (2004). An Enterprise Memory for Healthcare Knowledge Management. M.Sc. (by coursework) Project Report, Universiti Sains Malaysia.

Universiti Sains Malaysia
Pulau Pinang

Final Report

CCS500 M.Sc. COMPUTER SCIENCE PROJECT

An Enterprise Memory for Healthcare Knowledge Management

Lye Poh Khim
S-COM063/02
Part-Time

Dr Cheah Yu-N
Prof. R.K. Subramaniam
Dr Vincent Khoo Kay Teong

Semester I
2003/2004 Academic Session

Acknowledgements

I would like to express my sincere appreciation to my project supervisor, Dr Cheah Yu-N, for his guidance and support. This work would not have been possible without Dr Cheah's guidance and advice.

I would like to express my gratefulness to my family and friends, who are always by my side to support me. Thank you for them.

Table of Contents

Acknowledgements i
Table of Contents ii
List of Figures iv
List of Tables vii
List of Equations viii
Declaration ix
Abstract x

CHAPTER 1 INTRODUCTION 1
1.1 BACKGROUND 1
1.2 DEFINITION OF ENTERPRISE MEMORY 2
1.3 ENTERPRISE MEMORY SYSTEMS (EMS) 3
1.4 PROBLEM STATEMEN 4
1.5 PROBLEM SOLUTIONS 4
1.6 OBJECTIVES 5
CHAPTER 2 RELATED WORKS 6
2.1 ENTERPRISE MEMORY FRAMEWORKS 6
2.1.1 ANSWER GARDEN 6
2.1.2 ANSWER GARDEN 2 8
2.1.2.1 SYSTEM FLOW 9
2.1.2.2 MAIN SYSTEM COMPONENTS 9
2.1.3 HYPERMAIL ARCHITECTURE 11
2.2 RELATED APPROACHES/ALGORITHMS 13
2.2.1 SCENARIOS: STRATEGY FOR TACIT KNOWLEDGE EXPLICATION 13
2.2.2 CONNECTING PEOPLE TO KNOWLEDGE 14
2.2.3 TACIT KNOWLEDGE CRYSTALLIZATION 15
CHAPTER 3 PROPOSED DESIGN METHODOLOGY 18
3.1 PROPOSED METHODOLOGY 18
3.2 ARCHITECTURE OF THE PROPOSED SYSTEM 19
3.2.1 THE MAIN COMPONENTS IN THE KNOWLEDGE OBJECT LAYER 20
3.2.2 THE MAIN COMPONENTS IN THE APPLICATION LAYER 22
3.2.2.1 KNOWLEDGE ACQUISITION & REWARDS 22
3.2.2.2 KNOWLEDGE BROWSING & RATING 23
3.2.2.3 UPDATE REMINDER 24
3.2.2.4 SCENARIO CRYSTALLIZATION 24
3.2.2.5 Q-A CONNECT 25
CHAPTER 4 IMPLEMENTATION AND RESULTS 28
4.1 INTODUCTION 28
4.2 IMPLEMENTATION AT THE KNOWLEDGE DESCRIPTION AND APPLICATION LAYERS 30
4.2.1 IMPLEMENTATION DETAILS OF KNOWLEDGE REPRESENTATION IN THE KNOWLEDGE DESCRIPTION LAYER 30
4.2.2 IMPLEMENTATION DETAILS FOR THE APPLICATION LAYER 35
4.2.2.1 KNOWLEDGE ACQUISITION & REWARDS 35
4.2.2.2 KNOWLEDGE BROWSING & RATING 36
4.2.2.3 UPDATE REMINDER 38

4.2.2.4 SCENARIO CRYSTALLIZATION..... 40
4.2.2.5 Q-A CONNECT..... 42
CHAPTER 5 SUMMARY AND FUTURE WORK..... 47
5.1 REVISITING OUR CONTRIBUTIONS..... 47
5.1.1 REVISITING THE FIRST OBJECTIVE..... 48
5.1.2 REVISITING THE SECOND OBJECTIVE..... 49
5.1.3 REVISITING THE THIRD OBJECTIVE..... 51
5.2 COMPARISON WITH OTHER SIMILAR FAQ SYSTEM..... 52
5.2.1 FAQ FINDER SYSTEM..... 52
5.2.2 ANSWER GARDEN..... 55
5.3 FUTURE WORKS..... 55
5.4 CONCLUSION..... 58
REFERENCES AND BIBLIOGRAPHY..... 59

APPENDIX A..... 62
APPENDIX B..... 64
APPENDIX C..... 78

List of Figures

Figure 2.1	Answer Garden 2 (AG2) Architecture	9
Figure 2.2	Two possible escalations for a question	10
Figure 2.3	Enhanced Email (HyperMail) Architecture	12
Figure 2.4	The Scenario Structure Outline	14
Figure 2.5	Knowledge Crystallization algorithm	16
Figure 3.1	5-layer For Healthcare Enterprise Memory Framework	19
Figure 3.2	Modified Scenario Structure Outline	21
Figure 3.3	Q-A Structure Outline	21
Figure 3.4	Knowledge Acquisition Process	25
Figure 3.5	Q-A Connect Workflow	26
Figure 4.1	The two important layers of Healthcare-EMS framework	29
Figure 4.2	Procedural knowledge representation in XML file (Part of *meta_scenario.xml*)	32
Figure 4.3	Procedural knowledge representation in XML file (Part of *scenario.xml*)	33
Figure 4.4	Procedural knowledge representation in XML file (Part of *event.xml*)	33
Figure 4.5	Problem-solutions knowledge representation in XML file (Part of *Q_AList.xml*)	34
Figure 4.6	Scenario Item Knowledge Browsing Interface	37
Figure 4.7	Q-A Item Knowledge Browsing Interface	37
Figure 4.8	Q-A Item Knowledge Rating Interface	38
Figure 4.9	Personal Contributions with Update Reminder	39
Figure 4.10	Outdated answers in *garbage.xml*	39
Figure 4.11	Email notification sent by Update Reminder	40

Figure 4.12	Evaluation Phase and Calculation of Crystallization Factors	41
Figure 4.13	Question interface with keywords and weights	44
Figure 5.1	Heterogenous repository (Document Base(s)) and Hybrid representation	56
Figure B1	Knowledge Editor-Scenario Structure, Q-A Structure and Personal Contributions	63
Figure B2	Scenario Structure -View Scenario Details Interface	63
Figure B3	Scenario Structure - View Event's Details Interface	64
Figure B4	Scenario Structure - View All Events' Details for a Scenario (s.20000613.0540) Interface	64
Figure B5	Scenario Structure - Add New Scenario Interface	65
Figure B6	Scenario Structure – Vote for Event Interface	65
Figure B7	Scenario Structure - Add New Scenario Interface	66
Figure B8	Q-A Structure – Browse Q-As Interface	66
Figure B9	Q-A Structure – View Q-A Structure's Details Interface	67
Figure B10	Q-A Structure – Add New Answer (With Keyword and Weight) Interface	67
Figure B11	Q-A Structure – Vote for Answer Interface	68
Figure B12	Q-A Structure – Create New Question Interface (With Keyword and Weight)	68
Figure B13	Personal Contribution - Update Reminder highlights outdated answer with red star	69
Figure B14	Personal Contribution – Modify Personal Contributed Answer Interface	69
Figure B15	*login.xml* – Records All The Users (Employees in an Healthcare Enterprise) Details	70
Figure B16	A Clearer Illustration of *Q_AList.xml*	70
Figure B17	A Clearer Illustration of *scenario.xml* – Scenarios'Details are kept in this file	71
Figure B18	A Clearer Illustration of *sub_class_list_element..xml*	72

Figure B19	A Clearer Illustration of *event.xml*	73
Figure B20	A Clearer Illustration of *garbage.xml*	73
Figure B21	A Clearer Illustration of *context.xml*	74
Figure B22	Asking question of FAQFinder	75
Figure B23	Choosing a FAQ Files to match against	75
Figure B24	Choosing an answer	76

List of Tables

Table 1	Description of five categories of EMS	61
Table 2	Four general activities of authoring process	11
Table 3	Simulation of SCF, EpCf and EvCF values based on Users' Votes	62

List of Equations

Equation 4.1	User Evaluation	35
Equation 4.2	Event Crystallization Factor	41
Equation 4.3	Scenario Crystallization Factor	41
Equation 4.4	Scenario Attraction Factor And Keyword Similarity Measure	43
Equation 5.1	Overall match similarity for FAQFinder System	54

Declaration

Abstract

Enterprises and organizations are facing difficulties in keeping the knowledge of their employees especially staff turnover of an enterprise is quite high nowadays. However, information technology has enabled organizations to generate and retain huge amounts of information by using the underlying concept of an Enterprise Memory (EM). Generally, Enterprise Memory is used widely in most organizations due to its capability of decreasing redundancy by reusing the past experience and knowledge of their employees. Therefore, it is important to healthcare organizations have to deal with big chunks of new discoveries and past records everyday.

In this report, we are going to commence with the background, definition of EM, Enterprise Memory Systems (EMS) and state the problems as well as the solutions in Chapter 1. Following this chapter, we reviewed about a few related works and approaches/algorithms that reveal some points of commonality with our proposed methodology. Chapter 3 is an essential part of this report as it describes and explains the details of the proposed methodology for our *Enterprise Memory System for Healthcare Knowledge Management (Healthcare-EMS).* Subsequently, Chapter 4 presents the implementation details for *Healthcare-EMS* with justification by some example results. These expected results are captured in the form of screenshots of user interfaces, calculation by applying proposed equations along with simulated results. Finally, evaluations, comparison with other similar systems and possible future work are further discussed in Chapter 5.

CHAPTER 1

INTRODUCTION

They have the information they need, but they don't know they have it.
Or, knowing they have it, they can't find it.

– Mark S. Ackerman

1.1 BACKGROUND

Enterprises and organizations are facing difficulties in keeping the knowledge of their employees and therefore staff turnover of an enterprise is quite high nowadays. However, information technology has enabled organizations to generate and retain huge amounts of information by using the underlying concept of an Enterprise Memory (EM). Generally, Enterprise Memory is used widely in most organizations due to its capability of decreasing redundancy by reusing the past experience and knowledge of their employees. Thus, there are many organizations attempting to build their own frameworks for knowledge reuse purposes. The knowledge of employees will persist in the enterprise, as the knowledge will be kept in the core of EM. Therefore, this repository will maintain a group of both formal and informal knowledge.

The enterprise memory, however, has become a hot topic recently due to the growing recognition that it appears to be missing in many organizations today [1]. These organizations have serious limitations in transferring previous learning experiences to current problems. Thus, enterprise memory must exist for many aspects/in many forms within an organization, including organizational culture, processes, and structures. It must exist not only in archival forms but as a vital ingredient embedded within the many organizational processes and tasks [2] without disturbing the flow of work of employees. Enterprise memory has been used in both profit and non-profit organizations or enterprises including healthcare organizations. Healthcare organizations have to deal with big chunks of new discoveries and past records everyday. Consequently, healthcare enterprises will make use of EM concepts to manage its intellectual resources and therefore minimize the outflow of limited resources.

1.2 DEFINITION OF ENTERPRISE MEMORY

Enterprise Memory is sometimes written as Corporate Memory, Group Memory, Organizational Memory, Community Memory and many more. The term used solely depends on the context of use for the memory or repository being maintained and preserved by an organization. Although Organizational Memory (OM) is the term most widely used nowadays, we refer to the collection of memory as Enterprise Memory (EM) throughout this mini thesis rather than those mentioned before.

Even though, there is little consensus in organizational literature about how to define organization memory, we will mention some definitions proposed by experts. As defined in [3], Enterprise Memory (EM) is the "explicit, disembodied, persistent representation of knowledge and information in an organization which preserve(s) reasoning, capture, modeling and storage of organization knowledge". Conklin [1], defined enterprise memory as "the record of an organization that is embodied in a set of documents and artifacts" whereas OM is delineated as long-term vision at the core of learning organizations, supporting sharing and reuse of individual and corporate knowledge and lessons learned (past experience) in [4]. In defining the Enterprise Memory, Ackerman [5], suggested it as " an organization's ability to benefit from its past experience in responding more effectively (i.e. faster or more accurately) in the present. Apart from that, "Enterprise Memory" refers to stored information from an organization's history that can be brought to bear on present decision [6].

For our purpose here, enterprise memory is a repository that allows knowledge input, retrievals and access by employees in a healthcare enterprise where this repository stores the past experience of experts and lessons that can be learned. Enterprise memory deals mostly with "authoritative" knowledge about how to do things (such as using devices or following organizational procedures).

1.3 ENTERPRISE MEMORY SYSTEMS (EMS)

Whilst EM is a conceptual term, an EMS tries to support this concept with information technology. The ambition to create an Organizational Memory Information System (OMIS) (referred as OMS in [7]) is not realistic as shared knowledge spaces are spanning all over the entire knowledge of an organization [6]. The main problem is the intricacies to interpret the stored data that has a large knowledge domain correctly.

Organizational learning and knowledge management have been the two most important underlying approaches of enterprise memory. Organizational learning and knowledge management approaches in these EMSs were supported by the application of advanced database and network technologies. Fundamentally, an EMS is a system, which realizes parts of the EM (also called organizational knowledge base) with the help of information system and processes closely related to the use of the EM.

Enterprise Memory System (EMS) can be grouped into five categories based in approaches as defined in [7]. Basically, these five categories of EMS can be distinguished and classified as *corporate knowledge repository* or *meta-knowledge system*, *knowledge agents*, *knowledge bases expert systems*, *knowledge and communication integration platforms* and *knowledge creation and knowledge structuring* respectively. Further descriptions on these five categories can be referred to in Appendix A, Table 1.

1.4 PROBLEM STATEMENT

Nowadays, many organizations are attempting to abolish the redundancies in term of information and data after the recognition of enterprise memories (EM). Generally, there are some disadvantages of current EM system as stated below:

1. Current knowledge representation technique does not support too fine-grained and complex knowledge and at the same time, provided another more comprehensible form of knowledge representation.
2. Acquisition mechanism being used does not reflect how experts solve problems especially too fine-grained knowledge representation technique (i.e. scenario) is less convenient for showing how a problem is being handled. Manual knowledge acquisition mechanism is less helpful in acquiring experience or knowledge.
3. Current knowledge acquisition technique does not link back to the performance evaluation on employees in the enterprise.
4. Current knowledge repository does not evolve with knowledge while the repository is growing.

Hence, multiple precautionary steps must be taken to create a resourceful and prevailing EM system.

1.5 PROBLEM SOLUTIONS

To address the above issues, we present an Enterprise Memory System (EMS) for healthcare knowledge management. For this purpose, we would like to put forth the following solutions: -

1. Using *scenario-based* and *QA-based* approaches to represent the procedural and problem-solution knowledge in healthcare enterprises.

2. If users cannot find the desired answers or knowledge in the scenario-based knowledge repository, they can post their questions in the forum-like session and experts will be invited according to their previously submitted experience/knowledge to give their answers.
3. Making use of a *reward system* to motivate employees to store their respective knowledge in the EM system.
4. Informing the owner of a particular knowledge to do appropriate modification from time to time via a *update reminder* to prevent answer or knowledge from becoming out-of-date or inaccurate. In addition, employees in a healthcare enterprise should be allowed to rank their favorite knowledge in the forum via a *ranking system*. Consequently, the irrelevant or less popular knowledge can be removed permanently from the core repository.

1.6 OBJECTIVES

The main objectives of this research project are: -

- Propose suitable knowledge representations for a healthcare enterprise memory system.
- Introduce knowledge acquisition algorithm to enable the repository of a healthcare enterprise memory system to grow continuously.
- Produce a new methodology to maintain a healthcare enterprise memory system.

CHAPTER 2

RELATED WORKS

2.1 ENTERPRISE MEMORY FRAMEWORKS

There are lots of organizations and enterprises that are making use of the underlying concepts of enterprise memory to build their own framework for reusing and even trading their employees' knowledge. Generally, in this chapter, we discuss four main frameworks, which are Answer Garden (AG), Answer Garden 2 (AG2), and HyperMail architectures.

2.1.1 Answer Garden

One of the popular enterprise memories widely used is the Answer Garden (AG). It is a tool for growing organizational memory as well as to facilitate a cooperative work that helps improve an enterprise's memory. The AG permits enterprises to develop a database of commonly asked questions, which grows "organically" as new questions arise and are answered [5]. It is designed such that there exists a stream of questions, in which it may occur over and over but along with those that the system might not have seen before. AG incorporates three key ideas as stated below:

1. ***A branching network of diagnostic questions:*** This branching network assists users to find the answers they requested. At the end of the path of the branching network of diagnostic questions are a stream of questions people have asked in that situation and the answer to the particular questions by experts.

2. ***Automatic routing of new questions to appropriate experts and inserted into the branching network:*** In case users cannot find the answer they want by following through the guided branching network, they can create a new question and this question will be routed to the appropriate expert. The answer for the question will be sent back to users and at the same time inserted into the branching network where question was initiated. Consequently, the information database grows over time.
3. ***Modifications of the diagnostic branching network in response to user's feedback:*** When experts conclude (or from usage statistics for the branching network as a whole) that a branching network is misleading or inefficient for users, they can add to or change the diagnostic questions in the branching network.

As described before, there are three key ideas that play a crucial role in the Answer Garden. Nevertheless, [5] does not clearly indicate how these three ideas work in the system. The main way users locate answers in a branching network is by answering a branching series of multiple- choice questions. These questions are more likely to be obvious symptoms users can observe, not underlying causes or subtle differences that experts can distinguish. Nevertheless, it is preferable that a view tree is used where some users (particularly experts), may know roughly where their question is answered and so, will not have to "click" their way down the long branching chain to find it. By using the view tree, users can jump instantaneously to any required node in the tree.

However, we can anticipate that by representing knowledge using this approach my have some disadvantages. The previous scheme only works well when Answer Garden has correct answers to all the questions people might ask and a well-structured branching network. There are several problems that create intricacies to the Answer Garden, for example when users cannot understand the questions; the knowledge-based answer is incomplete or out-of-date; answer for questions being asked is "buried" at the end of the long branching path, etc. The Answer Garden has solved them by utilizing two primary tools, which are New Questions and usage feedback respectively. The first tool enables users to create new questions and routed questions to appropriate experts as well as an additional notification list. Ultimately, answer for the questions will be sent back to users and inserted into the branching network. Usage feedback provides users with another kind of indication of how the system is being used and locations where the branching network is ambiguous, confusing and unclear.

2.1.2 Answer Garden 2

In the earlier section, we discussed about how Answer Garden allows organizations to develop information databases that grow when new questions arise and are answered [8]. Answer Garden 2 continues its work to yield a second-generation architecture of Enterprise Memory by investigating some of the issues encountered in the original system. Thus, there are several advantages to this architecture:

1. The design cleanly separates the front-end of Answer Garden from the back-end needs.
2. Answer Garden's functionality is decomposed into a set of distributed software services that provides a high level of organizational flexibility along with services that can be mixed and matched for additional flexibility.
3. The new architecture is generalizable as it makes much of the help functionality possible from any information system.

In this new architecture, Answer Garden 2 incorporates two underlying support systems: The Café ConstructionKit (CafeCK) and Collaborative Refinery (Co-Refinery). CafeCK provides a set of reusable objects that include message transport for synchronous and asynchronous communication; parsing a variety of semi-structured protocols; private and public channels for narrowcast communication; message filters; and message retrieval by a variety of semi-structured methods. Co-Refinery provides mechanism for handling individual and joint information spaces [8].

2.1.2.1 System Flow

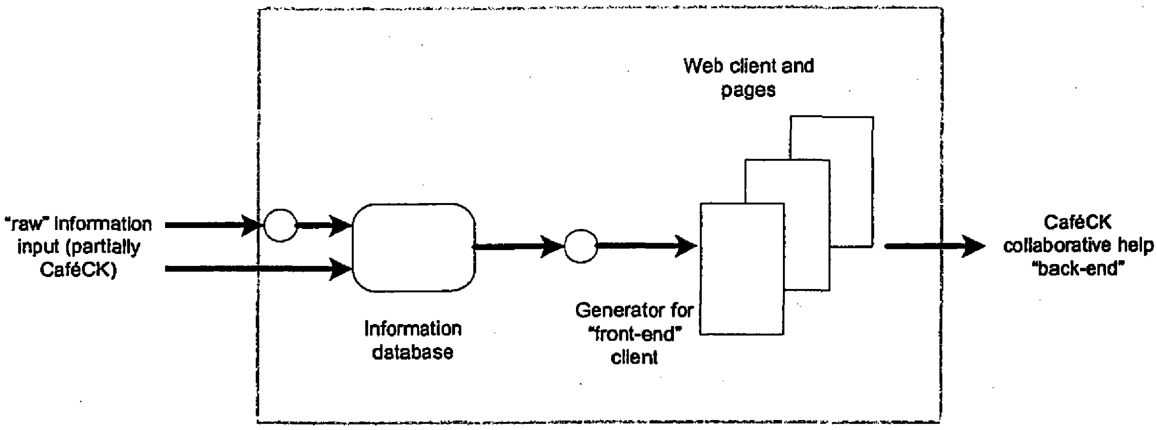
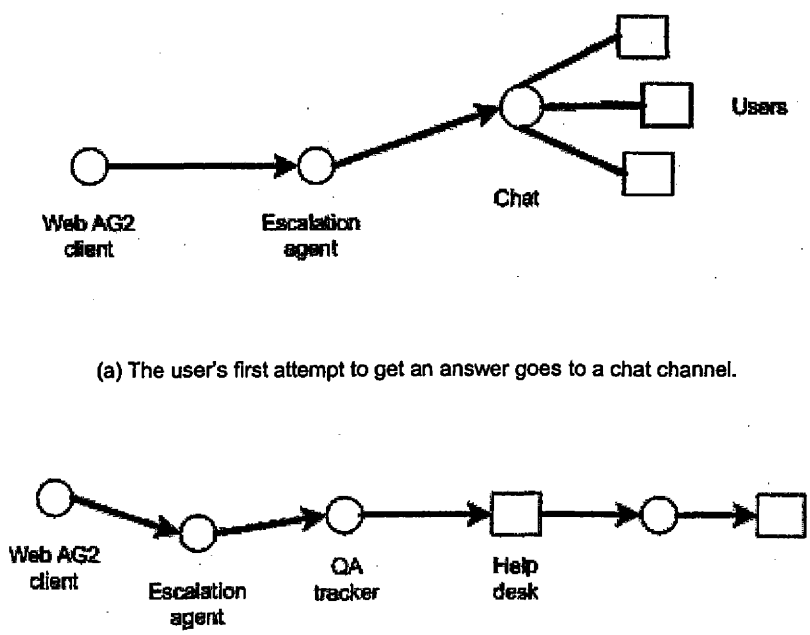


Figure 2.1: Answer Garden 2 (AG2) Architecture [8]

The working of the Cafe ConstructionKit (CafeCK) and Collaborative Refinery (Co-Refinery) is shown in Figure 2.1. First of all, raw information (partially processed before it is moved to the information database) comes into the collection archive through CafeCK processes (such as news filters) as an input to the Co-Refinery, by being explicitly sent to the archive through e-mail or filtering agents. Then, upon explicit queries (depending on a site's tailoring of AG2), materials are built into Web Pages, notes documents, or flat files. Sequentially, the AG2 Web or notes clients can send mail to CafeCK's back-end processes to handle the details of obtaining help.

2.1.2.2 Main System Components

AG2's back end can be viewed either as a collective memory system or as a collaborative help system. (The term *collaborative help* is used to denote those help systems that use people as information sources). Each of these views is dual of the others [8]. The term of duality in this context means that the language of linear programming is invoked when two forms exist for each particular problem. Users can choose form, which with the most analytical tractability to be solved. By having this duality characteristic, it is believed that a mechanism for reducing the context problem is found. Furthermore, its functionality also provides help to users at their own explanation level and potentially with iterative diagnosis. This collaborative help system provides escalation services (module) to users by its escalation agents. Users were able to simply construct an escalation agent for questions in AG2 using the facilities of CafeCK. This module allows users to decide what to do if the question is not answered. They are permitted to get answers from chat systems, bulletin boards, software agents, or other people. It is envisaged that the system being used can escalate the help request until it can be answered. This process can be quite flexible because the escalation agent makes use of a CafeCK process. This agent is currently programmed to follow organizational rules on the order of escalation (although this is under user control). The possible escalation processes are shown in Figure 2.2.



(a) The user's first attempt to get an answer goes to a chat channel.

(b) The user's *j*th attempt to get an answer gets escalated to a help desk

Figure 2.2: Two possible escalations for a question [8].

The Co-Refinery tool is another main component of AG2, which supports an authoring process that includes four general activities, which are collecting, culling, organizing, and distilling. Each activity is clearly important and may be done iteratively or in any order. The table below shows those four activities:

Table 2: Four general activities of authoring process

General activities	Explanation
Collecting	❑ The phase in which information is gathered. ❑ Automatic collecting can be set up for certain types of information streams that occurs in AG2, such as NetNews, synchronous chat channels, or distribution lists. ❑ Manual collecting allows individual items to be submitted through the system directly or by e-mail. ❑ It places items into the archive.
Culling	❑ This activity will cull the collection for interesting material. ❑ It is a selection mechanism, identifying themes or threads that occur within a collection to make subsequent organizing and distilling easier. ❑ It causes a sizable reduction of material as well as the noticeable size of the archive.
Organizing	❑ Enhance retrievability and understandability by allowing one to group materials according to some classification schemes.
Distilling	❑ The most important part of refining – uncover the answers or knowledge by boiling down the existing materials (and culled). ❑ This results in a more concentrated and concise form of the original information.

2.1.3 HyperMail Architecture

HyperMail Architecture is an enhance email architecture. Figure 2.3 presents the overview of the HyperMail architecture. Fundamentally, the software components of the HyperMail architecture comprises of an email client, a parser that analyses the email text, a relational representation of meta-knowledge, an indexing search engine, a HTML generator that recomposes the email message with embedded links into OM and an Enterprise Memory (EM). Besides these, there are a few types of actors involved in the system flow, i.e. the email author, the email recipient, the contributor of any portion of the Enterprise Memory referred to by the enhanced email. Each type of actor utilizes different functionalities and different interfaces when interacting with HyperMail.

Basically, the process of the HyperMail (see Figure 2.3) is initiated when users intend to send an email. If users choose to enhance an email when sending out an email, a dialog box will present three choices: Enhance and Send; Enhance and Preview; and Just Send it. Firstly, the enhanced email will parse the email text to identify any concepts that is located in the Organizational Memory. The selection and ranking of the identified concepts are based on correlation between the user-specific meta-knowledge about the sender, and similar knowledge about the recipient. The identified text is then augmented with a link to the appropriate location in the OM. Finally, a HTML version of the email message is created wherein the OM entries are presented as links [10].

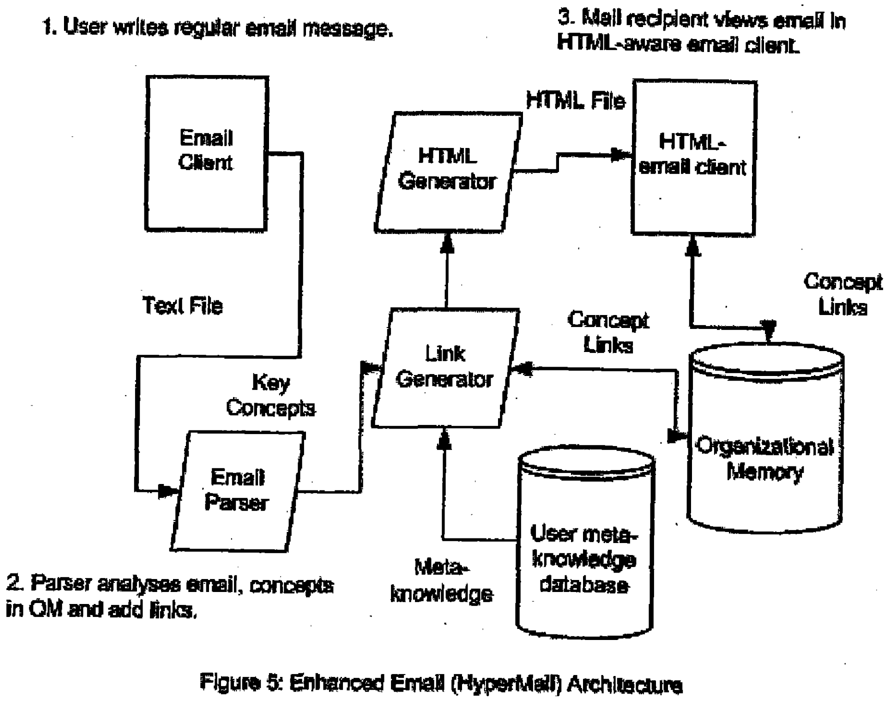


Figure 2.3: Enhanced Email (HyperMail) Architecture [10]

Every key concept identified by HyperMail appears as a link in the email text when the recipient receives it. For viewing concept information, one can choose to ascribe to a given meaning, disassociate himself/herself from the explanation or view the representation of other participants who did not do so, or add new entry to represent his/her understanding of the concept.

2.2 RELATED APPROACHES/ALGORITHMS

Several useful approaches and algorithms were being examined. These approaches and algorithms were previously employed as a strategy and structure for tacit knowledge explication and tacit knowledge crystallization.

2.2.1 Scenarios: Strategy for Tacit Knowledge Explication

The tacit knowledge strategy suggested by Cheah and Abidi (2001) utilizes 'hypothetical' scenarios for systematic manipulation of innate problem-solving skills in response to complex and/or novel problem situations. It is asserted that domain experts will make use of tacit knowledge when they are required to solve atypical problems.

The essential strategy of tacit knowledge explication is the notion of 'hypothetical' scenarios and therefore a novel knowledge structure called *scenarios* was devised. It manifests dual functionality: (a) situation descriptor, and (b) tacit knowledge accumulator. Scenarios are depicted as a goal-oriented narration or sequenced description of situation, together with the entities that constitute the situation, i.e. the actors, events, inputs, outcomes, environment and so on. With a scenario knowledge structure, explicated tacit knowledge can be accumulated as it is captured in terms of (a) a sequence of distinct actions that might be undertaken to accomplish a particular task; and (b) details of the sequence of interactions performed or experience by the scenario's entities to solve the problem [11].

Fundamentally, this knowledge representation structure comprises four main components and they are organized in a hierarchical taxonomy as illustrated in Figure 2.4. The functionality and description of these main components are further explained in the following chapter.

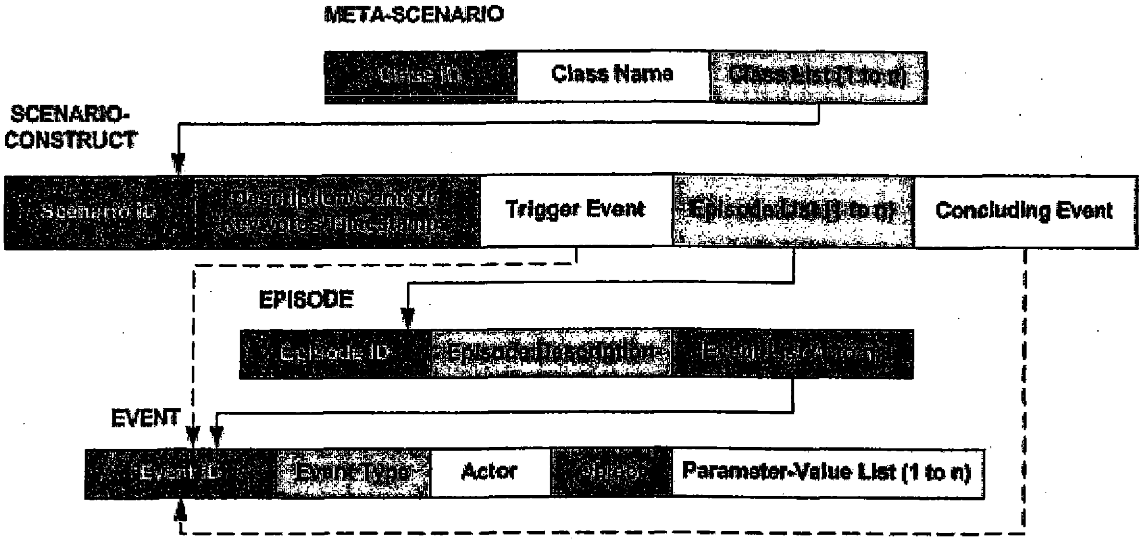


Figure 2.4: The Scenario Structure Outline [11]

The strategy for tacit knowledge explication using the scenario approach does have its own weaknesses. This scenario structure is too structured and less comprehensive from a user point of view. Training as well as a help tool might be useful to provide explanation but these may not to be a good way to solve the problem. Therefore, some other useful ways of representing tacit knowledge should used to complement this knowledge explication strategy in order to reduce any shortcomings of this scenario structure.

2.2.2 Connecting People to Knowledge

The second approach that might be useful in knowledge management is connecting people to knowledge [12]. In addition to search engines, unique tools are generated to help connect users and information. These tools are utilized to summarize available knowledge and suggest other places with the same and similar knowledge. In this approach, the "pull" technology is applied-search engines and intelligent agents.

Intelligent agents (e.g. InfoFinder developed by Arthur Andersen) offer another approach by generating user's profiles based on user's information interests in a document repository (Lotus

Notes). This agent learns profiles from documents and guides the users to other new and existing documents in the repository.

2.2.3 Tacit Knowledge Crystallization

Knowledge crystallization would prove to be a vital process in an EM. This process follows the chemical interpretation of crystallization based on structural/conceptual similarity and/or specified business rules. The underlying concept being used is the establishment of relationship between similar scenarios, leading to the ontological classification of knowledge within the scenario base [11]. Knowledge crystallization comprises of two sub-processes: *Nucleation* and *Growth* (Cheah and Abidi 2000).

The *nucleation* sub-process is the process where *knowledge seeds* are created. A knowledge seed is illustrated as a specification for synthesis of scenario items, which also served as catalyst for the formation of knowledge crystal in addition the nuclei around which the knowledge crystal is to be created. These knowledge seeds are released into the scenario base for following the growth sub-process. Thus, a domain expert, knowledge engineer and knowledge manager are responsible to design these knowledge seeds by following the knowledge requirements.

Automatic attraction of scenarios towards scenarios seeds is the main process of the *growth* sub-process. A priori screening is performed on candidate scenario items based on their user acceptance, correctness, appropriateness and completeness and as a subjective evaluation by enterprise-wide knowledge practitioners towards the scenario items they used. The scenario items that are voted with accumulated high points will be considered as stable and assigned with high *Crystallization Factor* (CF) values. Ultimately, the growth process compares all the knowledge seeds that have CF values greater than a predefined threshold and yields a *Scenario Attraction Factor* (SAF) that would determine if growth could proceed. Knowledge crystallization comprises of a two-phase algorithm: Evaluation (Phase I) and Attraction (Phase II). The process flow of these phases is illustrated in Figure 2.5.

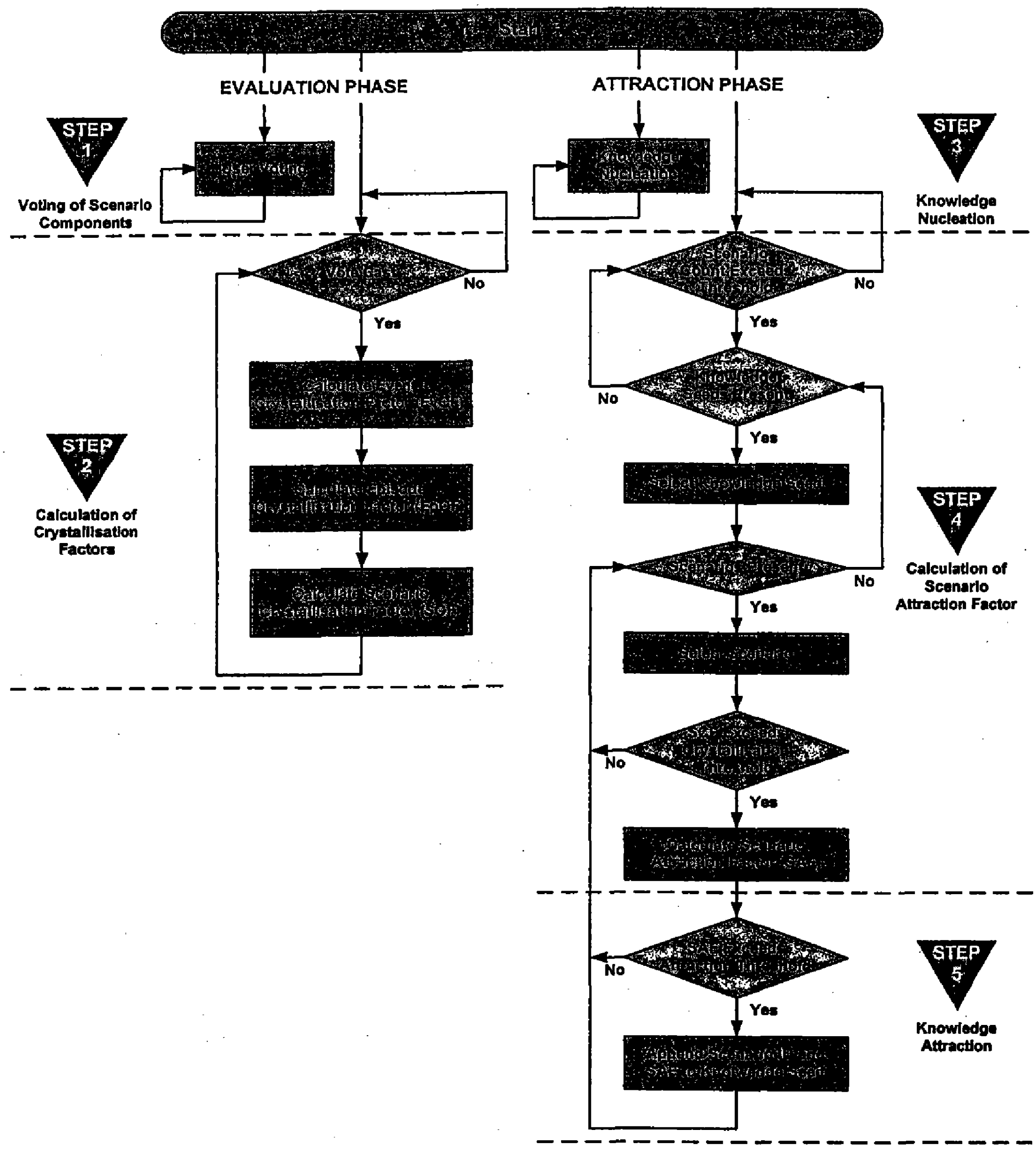


Figure 2.5: Knowledge Crystallization algorithm [11]

For Phase I, there are two important steps to accomplish its goals whereas there are three steps for Phase II. These steps will be performed in the sequence as shown in Figure 2.5. The steps involved in these 2 phases are voting of scenario components, calculation of Crystallization

Factor, knowledge nucleation, calculation of Scenario Attraction Factor and knowledge attraction respectively.

CHAPTER 3

PROPOSED DESIGN METHODOLOGY

3.1 PROPOSED METHODOLOGY

In this chapter, we will describe the details of a proposed methodology for a Healthcare Enterprise Memory. The proposed method was designed to achieve following objectives:

- To store/retrieve knowledge in/from the Enterprise Memory.
- To augment the knowledge repository.
- To encourage the process of knowledge sharing in an enterprise.

In this project, the proposed methodology mainly makes use of the algorithms and approaches being reviewed in Chapter 2. This proposed methodology would be used to implement a system called the *Enterprise Memory System for Healthcare Knowledge Management (Healthcare-EMS)*. Fundamentally, this system works as a core repository for knowledge acquisition and dissemination. Hence, there must be ways of augmenting the knowledge repository to enable it to learn from past experiences of experts. In addition to that, this system should facilitate the browsing, searching as well as editing of the past knowledge. Ultimately, the process of sharing knowledge in an enterprise will be accelerated and the members of the enterprise will start to gain knowledge through their own experiences. The proposed methodology is designed such that all the objectives mentioned earlier can be achieved.

3.2 ARCHITECTURE OF THE PROPOSED SYSTEM

For solving the problems associated with the current Enterprise Memory systems, we introduce a 5-layer framework (See Figure 3.1) to implement the system described in section 3.1.This framework reveals some points of commonality with those related works being reviewed in Chapter 2. The five layers are as follows:

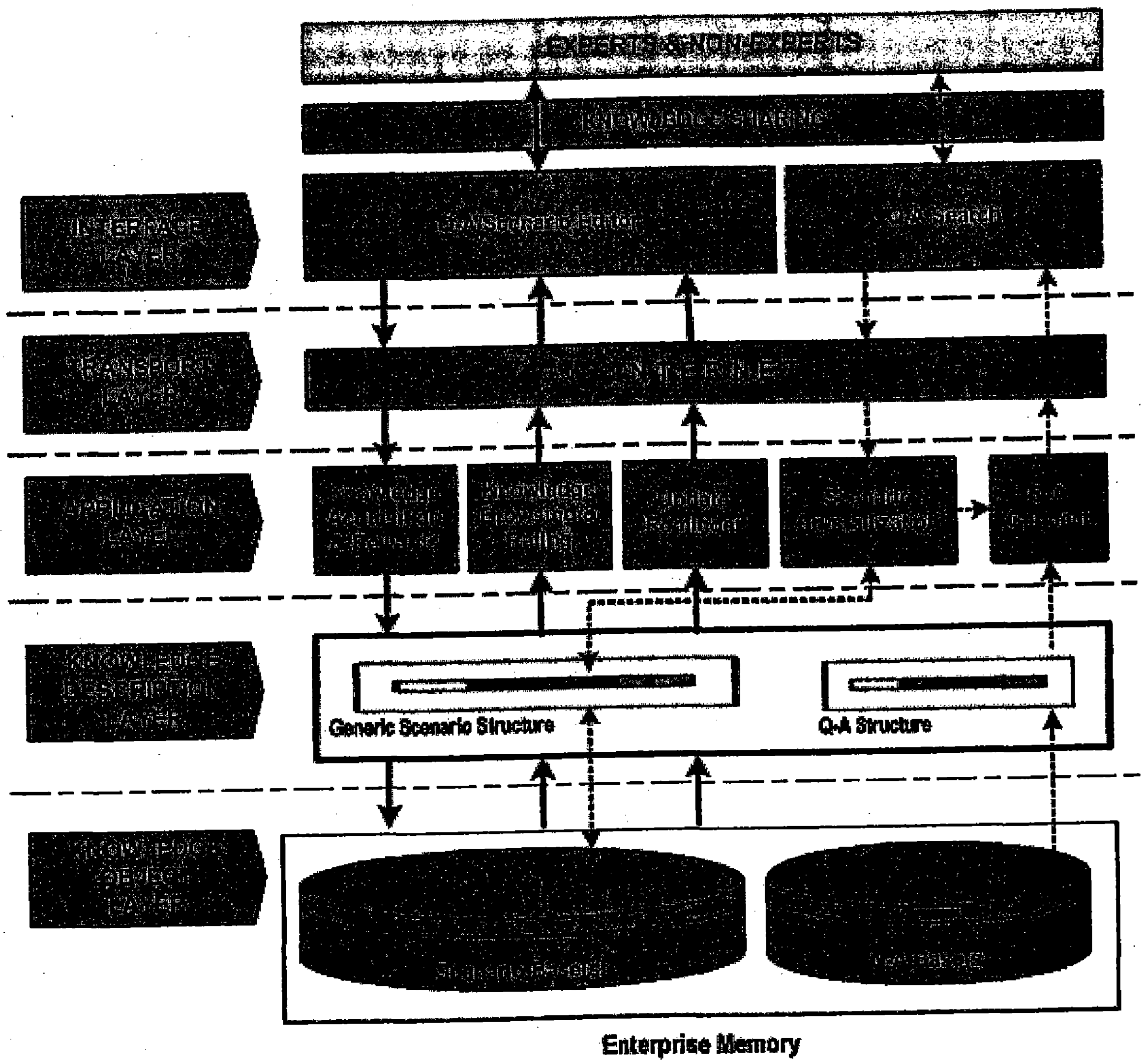


Figure 3.1: 5-layer Healthcare Enterprise Memory Framework

1. ***Knowledge Object Layer***: The lowest level of the proposed framework, which consists of both scenario base(s) and Q-A base(s). The scenario structure is described in section 2.2.3

whereas Q-A structure is working by pairing a question with a list of answers given by experts. These structures are quite direct and simple yet comprehensible in knowledge explication processes. These two distinctive knowledge structures are utilized to represent procedural knowledge and problem-solutions knowledge.

2. ***Knowledge Description Layer***: Both the generic scenarios and Q-A structures reside here. This layer describes the organization and formatting details of the scenario structure as well as the Q-A structures. Hence, it provides a structured and standard access to sources of knowledge at the Knowledge Object Layer.
3. ***Application Layer***: The main subsystems of the Healthcare-EMS reside in this layer. These subsystems comprise the Knowledge Acquisition, Knowledge Browsing & Rating, Update Reminder, Scenario Crystallization and Q-A Connect applications.
4. ***Transport Layer***: This is the medium through which the applications on the server interact with the users at the client side and vice versa. For our purpose, we propose to extend the functionality of this system over the Internet. However, other possible media could be the local area network (LAN) or the wide area network (WAN).
5. ***Interface Layer***: This is the layer where clients (users) interact with the respective applications in the server through the transportation layer. The layer provides user interfaces for sending requests and inputs to, and receiving outputs and replies from the server. This may be in the form of the Q-A/Scenario Editor and Q-A Search interfaces.

3.2.1 The Main Components in the Knowledge Object Layer

There are two types of distinctive knowledge objects in this layer. These knowledge objects make up the Healthcare-EMS system and they are called scenario structures and Q-A structures. The scenario base(s) consists of generic scenario structures whereas Q-A base(s) comprises of Q-A structures.

The scenario structure is explained at Chapter 2 (section 2.2.1) and it is illustrated in Figure 2.4. But, it is modified from the original structure (scenario in tacit knowledge strategy suggested by Cheah and Abidi (2001)) such that this knowledge representation structure contains only three main components namely Meta-Scenario, Scenario-Construct and Event (Figure 3.2) in the Healthcare-EMS. This modification can reduce the problem of scenarios being too fine-grained as described at section 2.2.1.1.

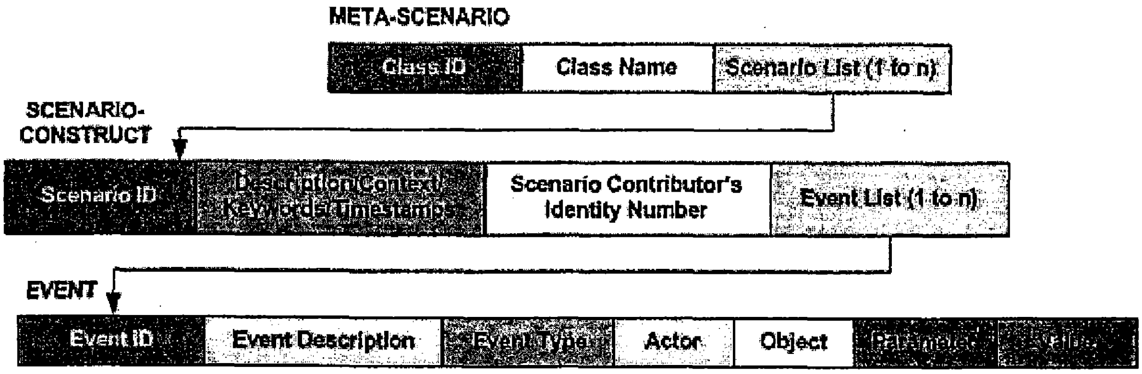


Figure 3.2: Modified Scenario Structure Outline

Q-A structure is another knowledge representation in the Healthcare-EMS (Figure 3.3). This knowledge structure is composed by pairing a question with a list of solutions/answers provided by experts based on the degree of matched keywords. These two knowledge objects work together to enable knowledge management processes in Healthcare-EMS.

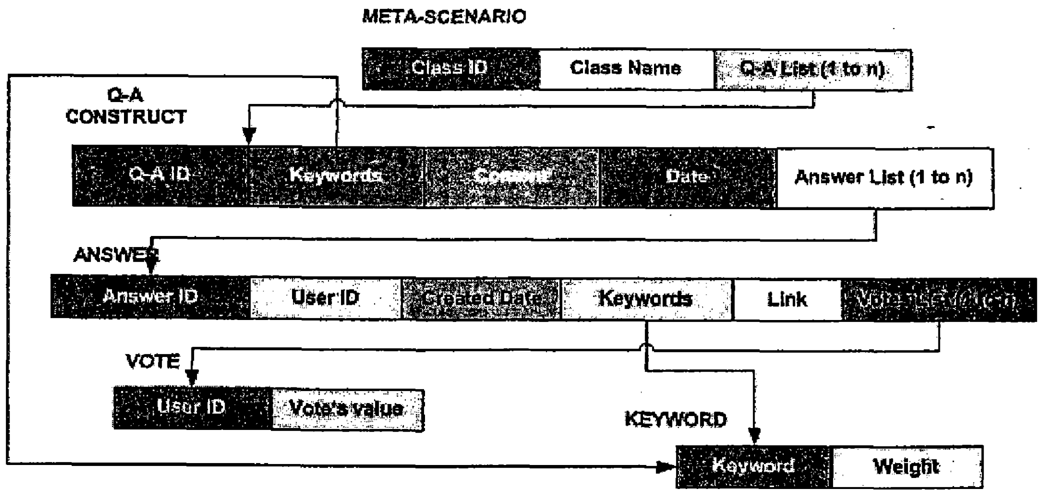


Figure 3.3: Q-A Structure Outline

Q-A structure has 3 main components, which play an important role in knowledge representation. These components comprise of Meta-Scenario (sharing same component with scenario structure), Q-A Construct and Answer. Q-A structure's components have their respective attributes used for knowledge maintenance as well as knowledge acquisition.

3.2.2 The Main Components in the Application Layer

There are various knowledge-centric and regulatory components in the application layer. Fundamentally, there are five components used to support this system: Knowledge Acquisition, Knowledge Browsing & Rating, Update Reminder, Scenario Crystallization and Q-A Connect. Their functionalities as well as their respective algorithms or approaches used will be explained in detail in the following section.

3.2.2.1 Knowledge Acquisition & Rewards

There exists two ways of knowledge acquisition in the Healthcare-EMS because the system employs two different knowledge representation structures. For a particular scenario structure, users must fill information for three components as shown in Figure 3.2 in order to create a complete scenario. This information comprises of class name, descriptor/context/keywords/timestamp, scenario contributor's identity number and an episode list.

The Q-A structure is utilized when users cannot find their solutions/answers of their problems in the scenario-constructs. Therefore, they will be allowed to create their questions/problems in a forum. Those questions/problems can either be answered by anyone who has experienced/encountered the similar/same problems before or a number of experts, who will be invited to contribute their knowledge/experience in the forum. Answering questions in the forum-like approach will be further explained in Q-A Connect section (see section 3.2.2.5).

Along with knowledge acquisition, Knowledge Rewards application will assign bonus points to those who contribute their knowledge to the Enterprise Memory repository. Contributors to both the scenarios base(s) as well as the Q-A base(s) will be given points, which may contribute to their performance appraisal at the end of the year. Consequently, employees in an enterprise will be motivated to share their experience or knowledge in the healthcare enterprise. The rewards system will assign points to the knowledge contributors based on the users' acceptance level and the type of knowledge representation they will use for knowledge explication (scenarios base(s) or Q-A base(s)). This users' acceptance level refers to how helpful users found the knowledge in solving their own problems. Knowledge contributors will be assigned with higher bonus points if they contribute their knowledge using the scenario approach, which requires a greater number of steps to yield a complete scenario.

3.2.2.2 Knowledge Browsing & Rating

Knowledge Browsing is an application that permits users to navigate through the Healthcare-EMS repository. This application is particularly useful when users are seeking for solutions/answers for their respective problems. Scenario structures are categorized into classes (differentiated by class name) and therefore, users can browse through scenario items under different classes. For Q-A structures, users can find their questions, which are categorized under distinctive class name (same like scenario browsing). A list of solutions/answers will be presented to users for every single question.

Along with the browsing process, users will be forced to rate the solutions/answers provided by experts. These rated values reveal the user acceptance, completeness, correctness, and appropriateness for scenario or Q-A structure. Knowledge Rating for scenario items requires users to evaluate usefulness of every single component for a scenario item whereas users are requested to rate the answers for a question.

3.2.2.3 Update Reminder

The Update Reminder is an important application for the Knowledge Creator/Editor interface. This application helps to remind owners of a particular knowledge residing in the core repository (Enterprise Memory) to update their knowledge periodically. This is done to maintain the "freshness" of the knowledge in the knowledge repository. The reminder is based on other users' acceptance towards the composed knowledge as well as the period between the last modification date and the current date, D. If the other users' acceptance value is greater than a predefined threshold then value D will have negligible influence on activating the update reminder. These users' acceptance values will be highly dependent on the users voting for Q-A items.

Update Reminder is being practiced for Q-A items only. It is utilized for the purpose of maintaining the Q-A base(s) that will grow continuously over the time.

3.2.2.4 Scenario Crystallization

Scenario Crystallization is an underlying application of Q-A Search interface that utilizes knowledge crystallization algorithm. However, crystallization algorithm is performed on scenario items only where the final result of the algorithm will be employed by the Q-A Connect application to connect people to knowledge. This crystallization algorithm is important for keywords matching (sub-process numbered 3 in Figure 3.4). It collaborates with the Q-A Connect application to permit users to specify their questions and a list of keywords with their respective keyword attraction points (KAPs) (see explanation in Chapter 4). This application uses knowledge crystallization algorithm along with these KAPs values and keywords in order to invite experienced experts to contribute/share their knowledge.

The knowledge explication using the scenario approach enables the utilization of the knowledge crystallization algorithm, i.e. it allows the computation of various crystallization factors that reflect user acceptance, correctness, appropriateness and completeness and as a subjective evaluation by enterprise-wide knowledge practitioners. It works towards organizing the knowledge into clearly defined sub-repositories or 'knowledge crystals'.

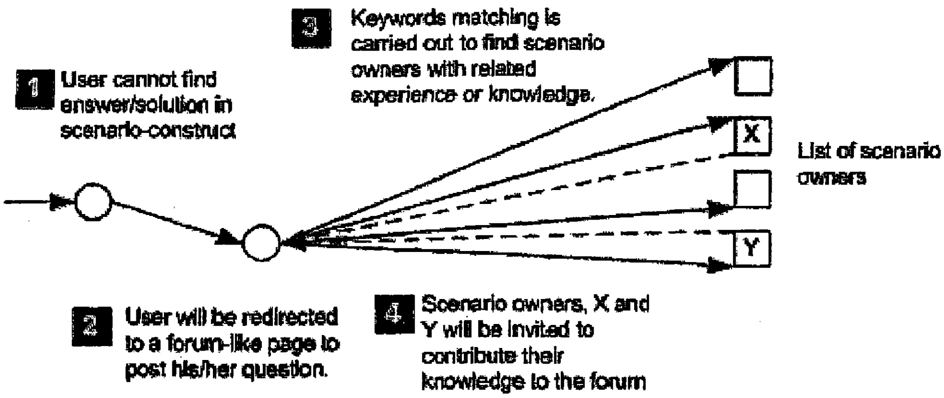


Figure 3.4: Knowledge Acquisition Process

The Attraction phase utilizes knowledge seed as a query mechanism for attracting knowledge owners to share their experience/solutions in the forum. A knowledge seed is defined by a questioners contains a question together with a list of keywords and their respective weights (KAPs values). These keywords and KAPs values of the knowledge seed will be used to calculate a scenario attraction factor (SAF) (see explanation in Chapter 4). The calculation of SAF is computed based on the criteria stated in Equation 3.4. If this SAF value is greater than a pre-determined threshold, then the scenario owner will be invited to share their knowledge to the Healthcare-EMS system. The invitation information (e.g. email address, scenario contributor's identity number) will be sent to Q-A Connect application for further processing.

3.2.2.5 Q-A Connect

The Q-A Connect application is used to connect all the experienced experts to share their knowledge in the forum. It only works with Q-A items in Q-A base(s). If a computed SAF value is greater than a predefined threshold, this denotes that the expert has previously submitted a similar scenario and is assumed to have similar/same knowledge/experience for a particular problem or answers provided by other experts (an answer may activates an attraction process). Therefore, an email will be sent to the expert in order to invite him/her to share their knowledge/experience in the forum-like session. This email will be attached with a link that will guide users to view the question in the session. The Q-A connect application will iteratively navigate through the scenario base(s) to connect all qualified experts to the session.

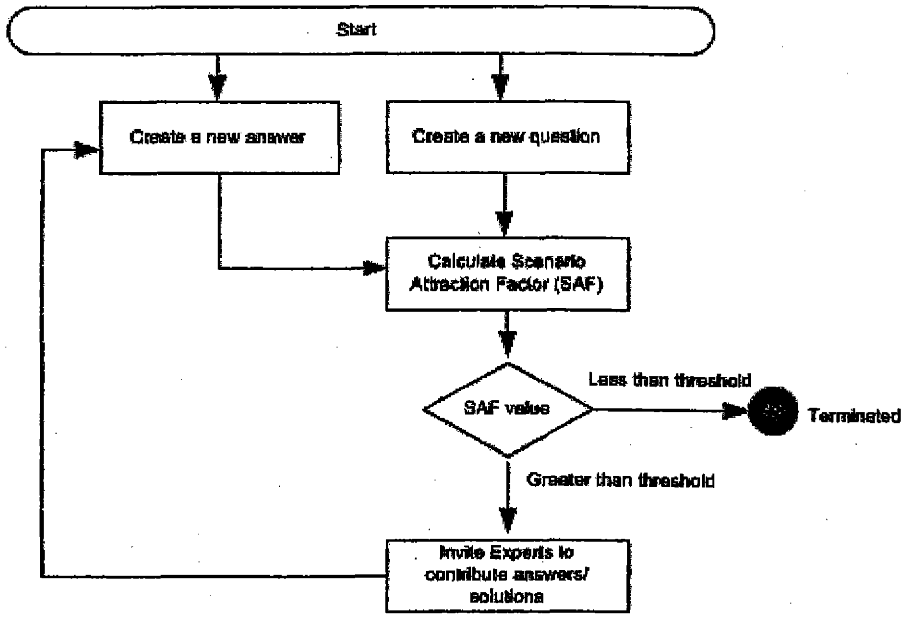


Figure 3.5: Q-A Connect Workflow

Fundamentally, Q-A Connect application is a tool that assists the Healthcare-EMS knowledge repository to grow continuously. This module is provided as a solution due to the lack of employees' initiatives to share their knowledge or experience. In another words, it cultivates a knowledge-sharing culture in an organization. The workflow of the Q-A Connect module is depicted as shown in Figure 3.5.

When less experienced experts who encounter a problem that they might not be able to diagnose, they will come and post their question in a forum-like session. Subsequently, Q-A Connect will request for a few keywords which can best describes the problem that has been submitted to the system. These keywords will be used as inputs for the calculation of the SAF value which determines to whom the invitation emails should be sent (refer to Equation 3.4(a)). A link to the question is attached with the email to inform the experienced experts to give solutions to the question. When these experts contribute their answers/solutions to the question, they are, again required to specify a few keywords than can relate to the answer provided by them. Consequently, this process creates a chain of invitations to the experts from scenario base(s).

CHAPTER 4

IMPLEMENTATION AND RESULTS

In the previous chapter, we had discussed the proposed methodology and algorithms being used. Now, we will proceed with implementation details for two important layers in the Healthcare-EMS framework, i.e. the Application Layer and Knowledge Description Layer (Figure 3.1 in Chapter 3). In this chapter, we will discuss and explain in details the workflows, equations and the AI approaches/algorithms in the following sections by using some sample results (i.e. screen shots) that we expected to obtain.

4.1 Introduction

We have designed and developed a Healthcare-EMS (Healthcare Enterprise Memory System), which provides the functional and operational framework for knowledge creation. Both the Knowledge Description Layer and Application Layer in the proposed framework (highlighted using dotted rectangular box as in Figure 4.1) are two vital layers in the Healthcare-EMS system. The knowledge description layer comprises of two distinctive knowledge structures for knowledge representation and knowledge acquisition purposes. These knowledge structures are the generic scenario structure and Q-A structure. The Application Layer consists of the underlying applications which support the operation of the Healthcare-EMS system. These essential applications are the Knowledge Acquisition and Rewards, Knowledge Browsing and Rating, Update Reminder, Scenario Crystallisation and Q-A Connect applications.

The generic scenario structure is a modified structure of the 'hypothetical' scenarios taken from the tacit knowledge explication strategy suggested by Cheah and Abidi (2001). This knowledge structure has been evolved such that it becomes a simplified version of the original 'hypothetical' scenario (see Figure 3.2). This generic scenario knowledge structure breaks down a scenario into a sequence of events that is used to describe the actions carried out in a healthcare scenario. Therefore, composition of a scenario requires an experts to explain/describe their treatment/ experience (procedural knowledge) with a great numbers of steps (events) in this system. In contrast to this knowledge structure, the Q-A structure allows experts to give both their problems and solutions in free-text format (see Figure 3.3). Both these knowledge representations maintain a list of keywords, which permits the Q-A base(s) to grow continuously. In the framework, the automatic knowledge acquisition mechanism only works with Q-A base(s) but it depends greatly on scenario base(s)'s crystallization values to grow.

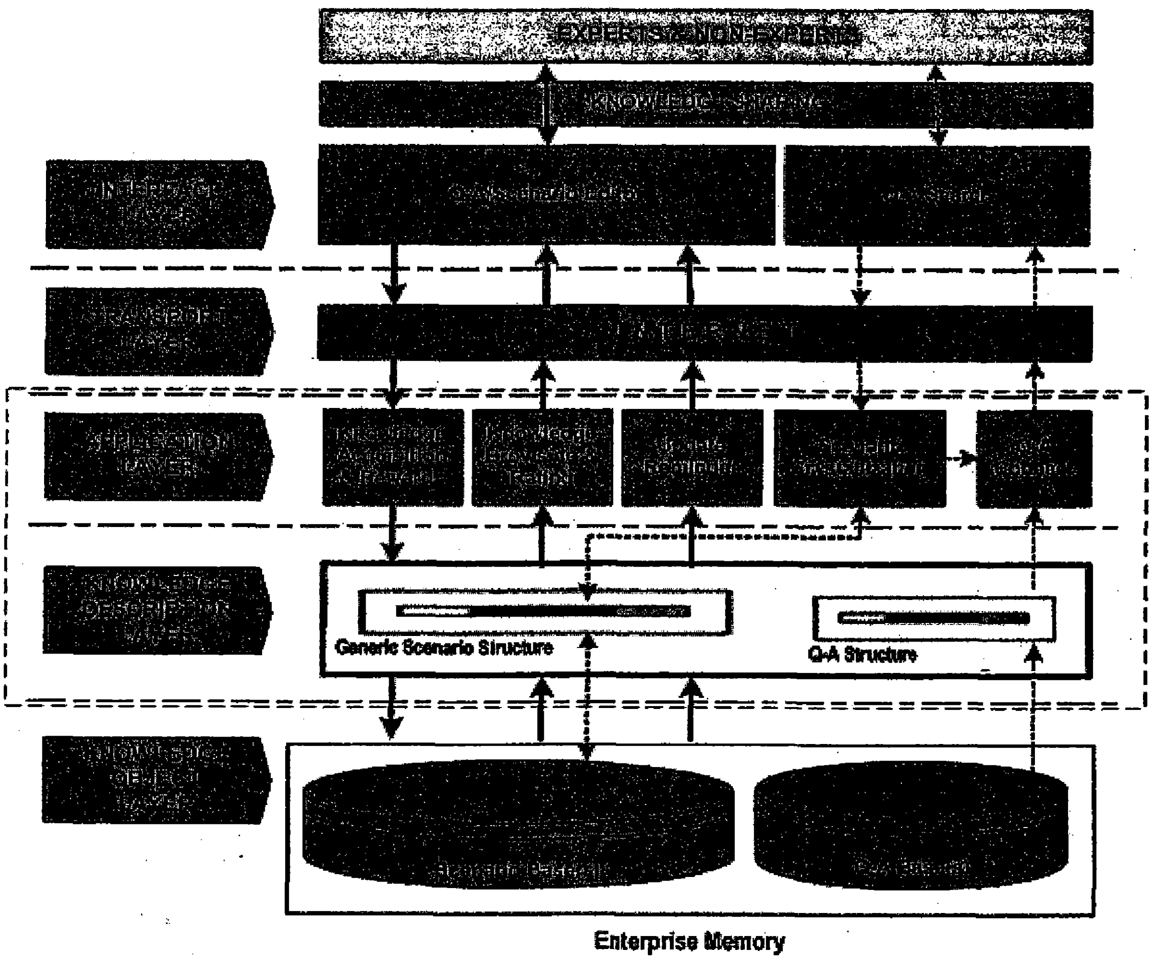


Figure 4.1: The two important layers of Healthcare-EMS framework (indicated by dotted box)

The applications of the Healthcare-EMS reside on the Application Layer. The Knowledge Acquisition and Rewards and Q-A Connect applications are primarily utilized as tools for knowledge acquisition to cultivate a knowledge sharing culture in an organization whereas the Update Reminder application helps to maintain the 'freshness' of the information/knowledge in the Q-A base. Thus, outdated/inapplicable knowledge in Q-A base might be eliminated after a particular period. Besides that, the Scenario Crystallization application is an essential tool to calculate scenario crystallization factors used for the attraction phase of the Q-A Connect application. The explanation and implementation details of the applications will be further discussed in the following sections in this chapter.

As mentioned in Chapter 1, the objectives of the Healthcare-EMS cover the knowledge representations, knowledge acquisition and knowledge maintenance. Thus, the explanation of implementation details and example results will reflect my contributions towards the system.

4.2 Implementation at the Knowledge Description and Application Layers

In this section, we will discuss the functional and algorithmic details of the applications in the Knowledge Description and Application Layers which we have briefly described in Chapter 3.

4.2.1 Implementation Details of Knowledge Representation in the Knowledge Description Layer

The Knowledge Description Layer consists of two distinctive knowledge representation structures namely generic scenario structures and Q-A structures. These knowledge structures are represented using XML files in a tree structure.

Procedural knowledge (scenario structures) is stored in five XML files in which three of them are representing the main components in a scenario structure. These components are Meta-Scenario, Scenario-Construct and Event (see Figure 3.2) respectively. These components' attributes are kept in *meta_scenario.xml, scenario.xml* and *event.xml* (see Figure 4.2, Figure 4.3 and Figure 4.4). XML files are treated as the storage medium of the Healthcare-EMS system. Besides, it allows storing of knowledge predicates by using custom tags (names that represent the knowledge item's elements and attributes). The procedural knowledge representation structure can be defined in a number of predicates as stated below:

- **meta_scenario/3**

  `meta_scenario(Class_Id, Class_Name, Sub_Class_Instance_List)`

  `Class_Id`, `Class_Name` and `Sub_Class_Instance_List` are fundamentally the elements to represent the scenario/problem domain for the Healthcare-EMS system. For instance, the Class_Name can be "Heart", "Choking", "Back" and etc. `Sub_Class_Instance_List` is used to keep all the scenarios that are categorized under the same `Class_Id` (see Figure 4.2, *meta_scenario.xml*).

- **scenario/12**

  `scenario(Sub_Class_Id, Scenario_Id, Contextual_Instance_List, Start_Timestamp, End_Timestamp, Scenario_Description, Event_Instance_List, Keywords, Vote_List, Scenario_Crystallisation_Factor, Voter_List, Scenario_Contributor_ID)`

  A scenario relation has twelve elements as mentioned above (see examples in Figure 4.3, *scenario.xml*). Apart from the details of a scenario (for instances `Sub_Class_Id`, `Scenario_Id`, `Contextual_Instance_List`, `Start_Timestamp`, `End_Timestamp`, `Scenario_Description`, `Event_Instance_List` and `Keywords`), information of the votes obtained and scenario contributor's information are kept for use during the scenario attraction phase. The scenario attraction phase is highly dependent on the `Scenario_Crystallisation_Factor` and `Keywords`.

- **event/7**

event(Event_ID, Event_Name, Event_Type, Event_Actor, Event_Object, Parameter, Value)

This is the predicate used to keep the information of an event. Basically, it keeps the details of an event (i.e. Event_ID, Event_Name, Event_Type, Event_Actor, Event_Object, Parameter and Value). This information describes the treatment/action by a medical officer (see examples in Figure 4.4, *event.xml*).

```
<meta_scenario>
    <class_id id="c11">
        <class_name>Heart</class_name>
        <sub_class>2</sub_class>
        <sub_class>368</sub_class>
    </class_id>
    <class_id id="c12">
        <class_name>Choking</class_name>
        <sub_class>70</sub_class>
    </class_id>
    <class_id id="c13">
        <class_name>Bleeding</class_name>
        <sub_class>101</sub_class>
    </class_id>
    <class_id id="c14">
        <class_name>Back</class_name>
        <sub_class>200</sub_class>
    </class_id>
</meta_scenario>
```

Figure 4.2: Procedural knowledge representation in XML file (Part of *meta_scenario.xml*)

Apart from the scenario structure for procedural knowledge representation, the representation for problem-solutions knowledge representation is also stored in XML file. This knowledge representation structure is called the Q-A structure. The predicates of this structure are as follows:

- **question_Instance_List/2**

  question_Instance_List(Domain_ID, Question_List)

- **question_answer_List/3**

  question_answer_List(Question_ID, Keyword_List, Answer_List)

- **answer/7**

  answer(Answer_ID, Answer_Creation_Date, Answer_Composer_ID, Answer_Content, Keywords, Reference_Link, Vote_List)

```
<scenario>
    <scenario_construct list_id="71" scenario_id="s.20000228.1038" userID="4">
        <scenario_desc>Choking adult male, 35 years of age. Location: Restaurant</scenario_desc>
        <contextual_link context_id="75" />
        <start_timestamp>2005</start_timestamp>
        <end_timestamp>2010</end_timestamp>
        <episodes episode_id="77" />
        <episodes episode_id="133" />
        <episodes episode_id="134" />
        <episodes episode_id="135" />
        <episodes episode_id="85" />
        <voter id="2">0.74</voter><avgSc>0.74</avgSc><ScCrysFactor>0.645</ScCrysFactor>
    </scenario_construct>
    <scenario_construct list_id="134" scenario_id="s.20000613.0540" userID="4">
        <scenario_desc>Mild heart attack on adult male</scenario_desc>
        <contextual_link context_id="145" />
        <contextual_link context_id="149" />
        <start_timestamp>1520</start_timestamp>
        <end_timestamp>1538</end_timestamp>
        <episodes episode_id="150" />
        <episodes episode_id="152" />
        <episodes episode_id="153" />
        <episodes episode_id="158" />
        <episodes episode_id="159" />
        <episodes episode_id="160" />
        <episodes episode_id="161" />
        <episodes episode_id="166" />
        <voter id="1">0.78</voter><avgSc>0.86</avgSc><ScCrysFactor>0.7909375000000001</ScCrysFactor><voter
        id="2">0.94</voter>
    </scenario_construct>
</scenario>
```

Figure 4.3: Procedural knowledge representation in XML file (Part of *scenario.xml*)

```
<episode_construct>
    <episode id="77">
        <name>ep6</name>
        <desc>Verify</desc>
        <event id="79">
            <name>ev18</name>
            <type>Action</type>
            <actor>First-aider</actor>
            <object>Patient</object>
            <parameterValue id="80">
                <name>pv27</name>
                <parameter>Ask</parameter>
                <value>Patient is choking</value>
            </parameterValue>
        </event>
        <voter id="2">0.56</voter><avdEp>0.56</avdEp>
    </episode>
    <episode id="133">
        <name>ep7</name>
        <desc>First-aid</desc>
        <event id="82">
            <name>ev19</name>
            <type>Obstacle</type>
            <actor />
            <object>Patient</object>
            <parameterValue id="83">
                <name>pv28</name>
                <parameter>Choking</parameter>
                <value>Yes</value>
            </parameterValue>
        </event>
        <voter id="2">0.34</voter><avdEp>0.34</avdEp>
    </episode>
</episode_construct>
```

Figure 4.4: Procedural knowledge representation in XML file (Part of *event.xml*)

```
<trigger_list domain="c13">
    <question id="1" date="22/02/2004">How Are Fibroids Treated?
        <keywords>
            <keyword weight="0.67">menstrual bleeding</keyword>
            <keyword weight="0.89">Fibroids</keyword>
        </keywords>
        <answer>
            <ans id="a" date="22/01/2004" userID="1">
                <cont>If fibroids are not causing problems, they may need no treatment apart from regular medical
                checkups. Surgery is the standard treatment for fibroids that are causing pain, heavy or lengthy
                menstrual bleeding, or other problems. The two kinds of surgery most commonly performed are
                hysterectomy and myomectomy. &lt;br&gt;&lt;br&gt;Hysterectomy is the surgical removal of the uterus
                (and usually of the cervix as well). It is the most common treatment for fibroids. Three out of
                every 10 hysterectomies in the United States are performed because of fibroids. Currently,
                hysterectomy is the only permanent cure for fibroids. However, a woman cannot become pregnant or
                carry a baby after having a hysterectomy.&lt;br&gt;&lt;br&gt;Myomectomy is the removal of fibroids
                without removing the uterus. This operation preserves a woman's ability to bear children. However, a
                successful pregnancy is not guaranteed. Only 4 or 5 out of 10 women become pregnant and give birth
                after a myomectomy.
                </cont>
                <keywords>
                    <keyword weight="0.80">Hysterectomy</keyword>
                    <keyword weight="0.80">Myomectomy</keyword>
                </keywords>
                <link>http://yourmedicalsource.com/library/fibroids/FBR_treatment.html</link>
                <voter id="2">0.95</voter><avVote>0.9533333333333333</avVote><voter id="1">0.98</voter><voter
                id="4">0.93</voter>
            </ans>
        </answer>
    </question>
</trigger_list>
```

Figure 4.5: Problem-solutions knowledge representation in XML file (Part of *Q_AList.xml*)

In Figure 4.5, there is a domain "c13" with a question created on "12/01/2004", question content is "How Are Fibroids Treated?" along with two related keywords and weights, which are "menstrual bleeding" and "Fibroids" respectively. As depicted in Figure 4.5, there is an answer provider (userID "1") who gave his/her solution on "22/02/2004" along with keywords "Hysterectomy" and "Myomectomy" and weights, a reference link, with a list of votes' values from other experts. Some other predicates will be illustrated as examples in Appendix B.

These knowledge representation structures provide a clear and comprehensible illustration of both the procedural and problem-solution knowledge. On the other hand, the system needs not create a connection object when accessing the XML file. Therefore, this can be considered as one of the advantages of using XML to represent the knowledge. Every element or attribute of the XML files plays a crucial role for example element *keyword* and its respective *weight* contribute towards the knowledge acquisition in the Q-A base(s) of Healthcare-EMS.

4.2.2 Implementation Details for the Application Layer

We now discuss the implementation details for the applications in the Application Layer. A few example results are shown with some justifications to illustrate the efficacy of the Healthcare-EMS in the next sections.

4.2.2.1 Knowledge Acquisition & Rewards

The calculations of the user acceptance level for the scenario structure and Q-A structure are quite different from one and another. The Q-A structure makes use of the users' evaluation value (see Equation 4.1) whereas the scenario approach makes use of the scenario crystallization factor (see explanation in section 4.2.2.4) although both values are greatly dependent on users' votes. This is because users need to evaluate every single fine-grained level (i.e. events) in the scenario strategy but they only need to give their feedbacks (vote values) on Q-A structures for every item in the list of answers/solutions. Thus, the calculation of scenario crystallization factor (see Figure 4.11) requires more computations. The knowledge rewards in term of bonus point is calculated by multiplying the users' acceptance level for a particular knowledge with a pre-defined number (i.e. 10) to get its nearest integer value.

$$\text{Users evaluation, } E_i = \frac{\sum_{j=1}^{N} V_i}{N}$$

where V_i = Total votes of answer numbered i^{th}
N = Number of registered users in Healthcare-EMS

Equation 4.1: Users Evaluation

For the example shown in Figure 4.5, a Q-A solution/answer composer (*userID* "1") will be given 9.5 points as bonus for his/her contribution if all the experts who registered with the Healthcare-

EMS voted for this answer. Scenario crystallization factor (SCF) equation will be discussed later (see section 4.2.2.4). Thus, the calculated SCF value will be employed to produce bonus points for the scenario contributor (example in Figure 4.3). For the calculations below, we assume that all the registered experts have already voted for the Q-A item and scenario item below. The calculations involved are: -

(i) Q-A item bonus point's calculation

By applying Equation 4.1,

$$\text{The user evaluation, } E_a = \frac{\sum_{j=1}^{N} v_i}{N} = \left\{ \frac{0.95+0.98+0.93}{3} \right\} = 0.953$$

Bonus points = 0.953 X 10 **points**

= 9.5 (approximation) **points**

(ii) Scenario item bonus point's calculation

Scenario Crystallization Factor for s.20000613.0540 = 0.7909

Bonus points for the scenario contributor = 0.7909 X 10 **points**

= 7.9 (approximation) **points**

4.2.2.2 Knowledge Browsing & Rating

The Knowledge Browsing application allows experts to navigate the knowledge/experience in the Healthcare-EMS system. Knowledge Browsing enables experts to view both scenario items and Q-A items in a more comprehensible and organized interfaces. This knowledge are loaded from the

XML files (as mentioned in the previous section) and displayed in a well-organized manner. Procedural knowledge is depicted as a sequence of events (see Event List in Figure 4.6) that have was carried out in a particular scenario. In contrast, problem-solution knowledge is displayed as a question with a list of answers shown in a forum-like session (see Figure 4.7).

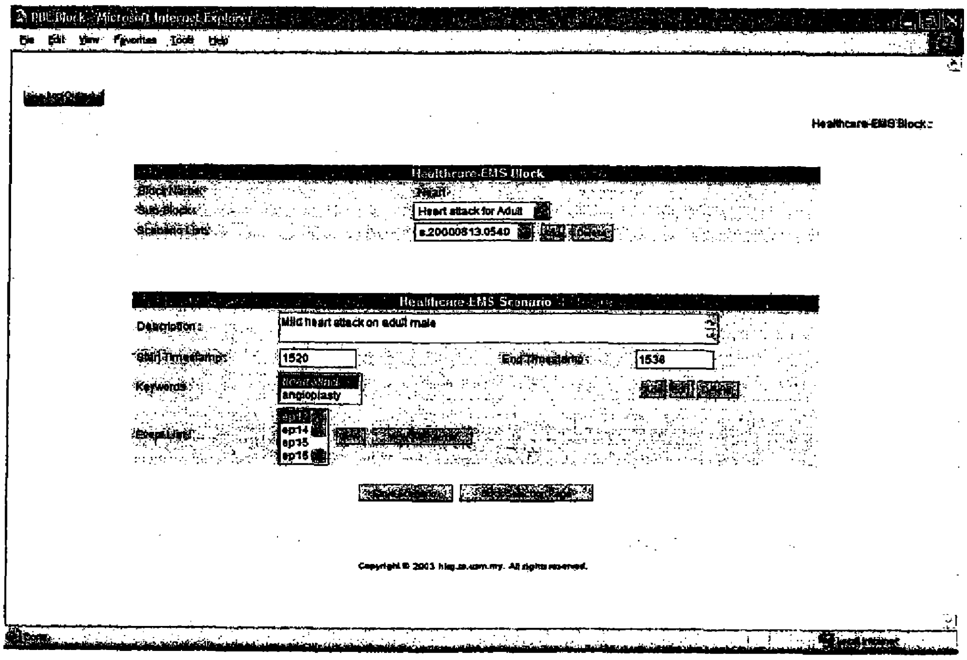


Figure 4.6: Scenario Item Knowledge Browsing Interface

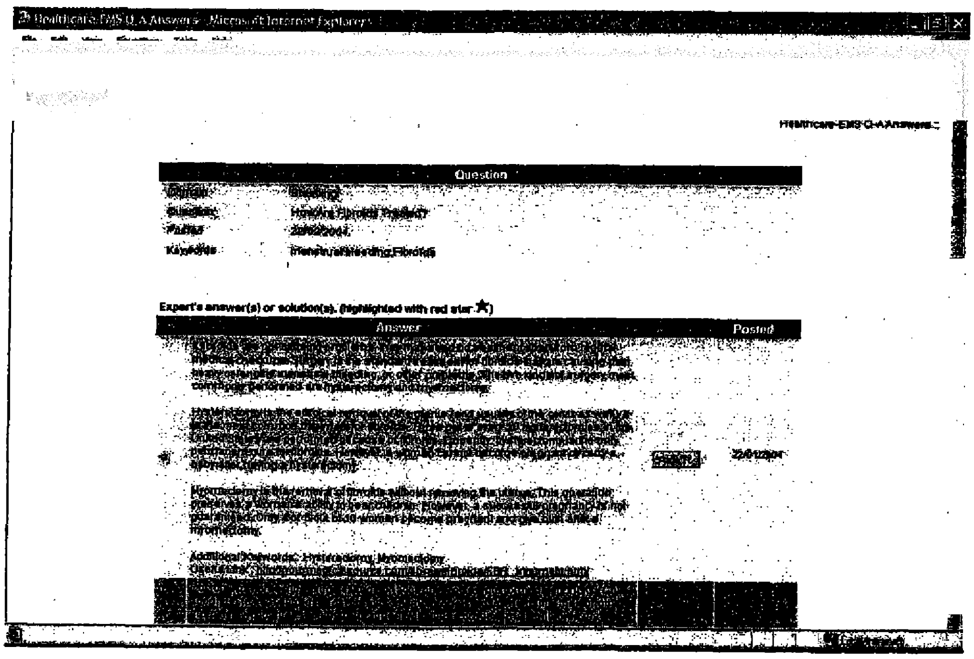


Figure 4.7: Q-A Item Knowledge Browsing Interface

The Knowledge Rating application has similar rating interfaces for both the scenario items and Q-A items. The vote values of scenario items are more important than Q-A items because these values will ultimately used to yield the scenario crystallization factor (see its functionality in section 4.2.2.4) that will be used for automatic knowledge acquisition process later. Therefore, experts will be forced to rate an event after they viewed each event whereas experts can volunteer to rate the answers of a question for Q-A items (see the "Rate It" button in Figure 4.7). However, for both knowledge items, each expert only can rate once for each scenario component or each answer to a Q-A item. Each rate/vote is a value between zero and one.

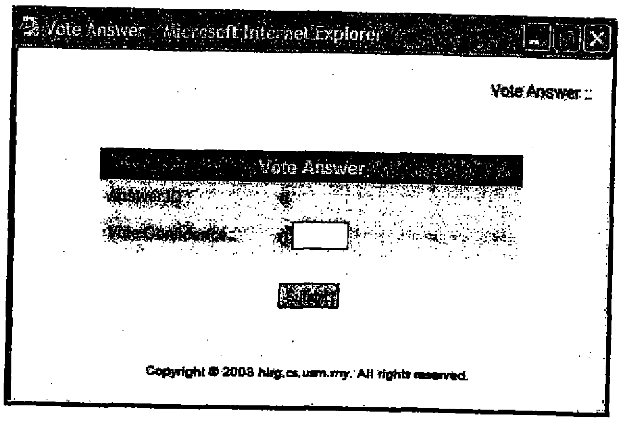


Figure 4.8: Q-A Item Knowledge Rating Interface

4.2.2.3 Update Reminder

The Update Reminder is a tool to keep all knowledge in the Healthcare-EMS "fresh". Experts who contributed knowledge/experience that is rated as inapplicable (determined by a pre-defined system threshold) will be notified to update them. If they do not update their contribution in a specified period, then their knowledge will be removed permanently from the repository. The weakness of this application is its inapplicability on the scenario items because alteration of a scenario's content (i.e. the information for events) will greatly falsify the voted/rated values previously done by other experts (since an expert only can vote/rate once).

There are two kinds of reminders. One of them is just to notify experts (highlighted with a red star, see Figure 4.8) to update their contribution when they visit their Personal Contributions page. There is another type of reminder, which is the result of the Update Reminder performing daily inspections on the outdated Q-A items. If these Q-A items are still not modified after more than one month, then an e-mail notification (see Figure 4.10) will be sent (information of these outdated answers will be kept in *garbage.xml*) and it will permanently eliminated if the answer composer does not modified the answer after three days of being informed. When these answers are eliminated, they will be subsequently removed from *garbage.xml* (see Figure 4.9) .

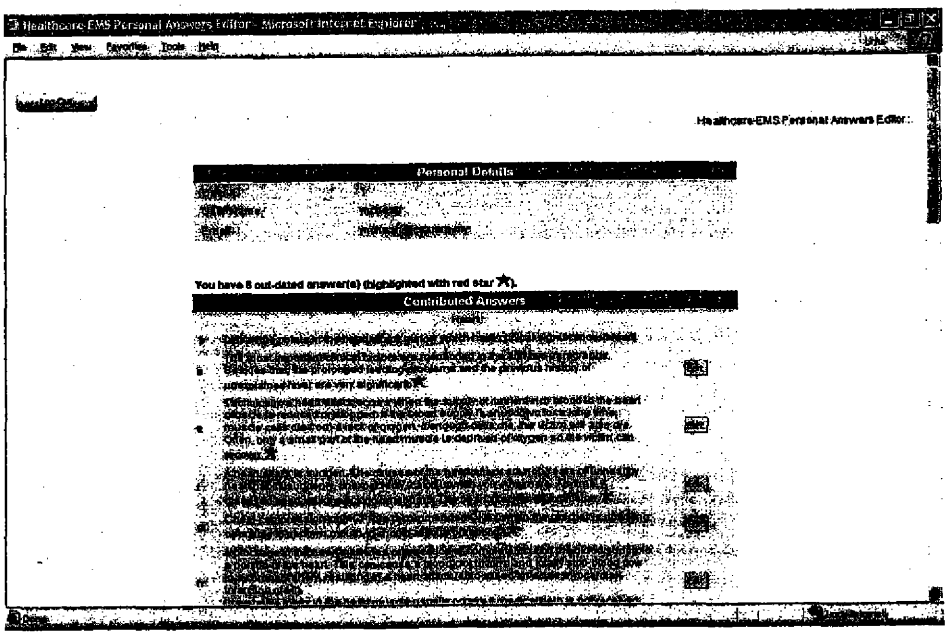


Figure 4.9: Personal Contributions with Update Reminder

```
<garbage>
    <garbageElem domain="c11" Qid="1" dateInsert="09/04/2004">c</garbageElem>
    <garbageElem domain="c11" Qid="1" dateInsert="09/04/2004">f</garbageElem>
    <garbageElem domain="c11" Qid="3" dateInsert="09/04/2004">b</garbageElem>
    <garbageElem domain="c12" Qid="1" dateInsert="09/04/2004">b</garbageElem>
</garbage>
```

Figure 4.10: Outdated answers in *garbage.xml*

The task of the Update Reminder is performed based on User Evaluation values (see Equation 4.1) for a composed knowledge as well as the period between the last modification date and the current date, *D*. If the User Evaluation of an answer is greater than a predefined threshold, then value *D* will have negligible influence on activating the update reminder.

If modification of an answer is done, then the composed date field will be set to the present date. Therefore, the Update Reminder tool will help to ensure all the knowledge items in the Q-A base(s) are applicable and "fresh".

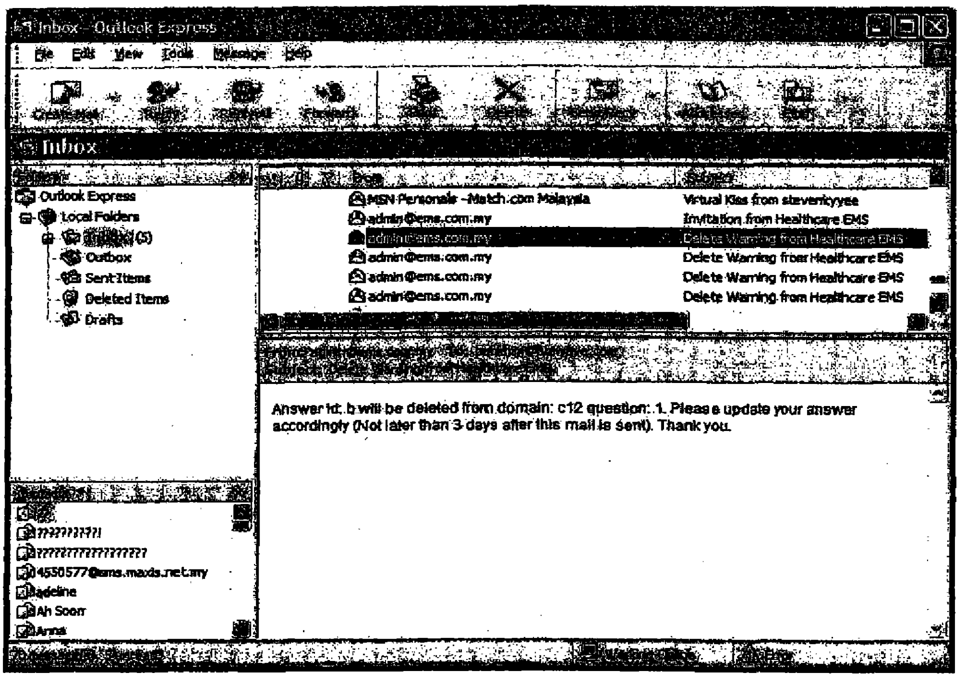
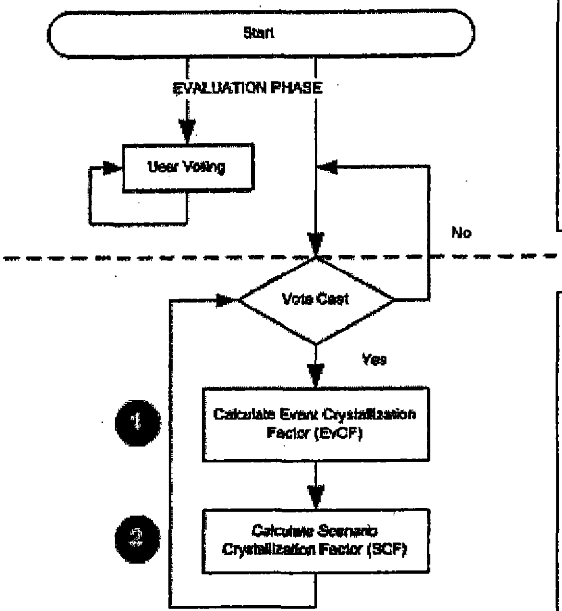


Figure 4.11: Email notification sent by Update Reminder

4.2.2.4 Scenario Crystallization

There are two vital phases in knowledge crystallization algorithm, namely the evaluation phase (phase I)(same process as for the evaluation in Knowledge Rating) and attraction phase (phase II) (Cheah and Abidi, 2000). In the evaluation phase (see Figure3.5), scenario crystallization factors and event crystallization factor will be computed based on users voting/rating process. The crystallization factors are basically defined in Figure 4.11 (Equation 4.2 and Equation 4.3).



Calculation 1:

Event Crystallization Factor, $Y = \overline{B}$,

where $\overline{B}$ = Average score of votes cast for each event.

Equation 4.2: Event Crystallization Factor

Calculation 2:

Scenario Crystallization Factor, $Z = \dfrac{\overline{Y} + \overline{D}}{2}$

where $\overline{Y}$ = The average of the Event Crystallization Factors of a particular scenario-construct, and

$\overline{D}$ = The average score of votes cast for the scenario-construct.

Equation 4.3: Scenario Crystallization Factor

Figure 4.12: Evaluation Phase and Calculation of Crystallization Factors

For an example, we will show the calculation of the scenario crystallization factor (SCF) and event crystallization factor (EvCF) for scenario item "s.20000613.0540". This scenario item has five events and their EvCF values are 0.56, 0.34, 0.74, 0.85 and 0.26 respectively. Before we can calculate the SCF value, we must first collect all EvCF values the event of this scenario (see Figure 4.11). Unlike the calculation for event crystallization factor suggested Cheah and Abidi, this equation will not take into consideration the sharing of same event because the event and episode proposed in the knowledge explication strategy has been combined into a single component named event. The following are the calculation for both the event crystallization factor and scenario crystallization factor by applying Equation 4.2 and 4.3 in Figure 4.11: -

By applying Equation 4.2

$$\text{The average of event crystallization factor} = \frac{0.56 + 0.34 + 0.74 + 0.85 + 0.26}{5}$$

$$= 0.55$$

By applying Equation 4.3,

$$\text{The scenario crystallization factor} = \frac{\overline{Y} + \overline{D}}{2}$$

$$= \frac{0.55 + 0.74}{2}$$

$$= 0.645$$

These calculations will be performed whenever there are new votes/rates for the event or the scenario. New votes will affect the value of these crystallization factors for a scenario.

Due to the simplification of the scenario structure, the calculation of the scenario crystallization factor and the event crystallization factor has become less complicated as compared to the calculation discussed by Cheah & Abidi, 2000. The scenario crystallization factor will later be employed in the attraction phase. A greater SCF value will have more influence in 'attracting' a scenario expert to contribute an answer for a question posted in the forum-like session (Therefore, forming a Q-A item).

4.2.2.5 Q-A Connect

As mentioned in Chapter 3, the Q-A Connect application is utilized to link Q-A questions to the relevant experts who have contributed scenarios, or in other words to 'attract' experienced scenario composers to contribute their knowledge to the questions and hence, form Q-A items. The scenario attraction process will be initiated when an less experienced medical officer who encounters a problem, post his/her question with a list of keywords and their respective weights (also known as Keyword Attraction Potential (KAP)). These keywords and their respective weights (Keyword Attraction Potential) describe how important these keywords can relate to the question (see Figure 4.12).

$$\text{Scenario Attraction Factor} = \frac{P + S}{2},$$

where P = Total keywords similarity measure degree divided by TW,

S = Scenario Crystallisation Factor of the same scenario,

TW = Total weights for Q-A's keywords.

(a) Scenario Attraction Factor

$$\text{Keyword Similarity Measure, } P_i = \frac{\sum_{i=1}^{K} \left\{ 1 - (1 - \frac{1}{K})^f \times \frac{W}{K} \right\}}{W}$$

where P_i = Keyword similarity measure for a particular keyword i,

K = Number of words for a particular keyword i,

f = Frequency of word for a particular keyword i that can match with keywords for a scenario,

W = Keyword Attraction Potential (KAP) for keyword i.

(b) Keyword Similarity Measure for a keyword i

Equation 4.4: Scenario Attraction Factor And Keyword Similarity Measure

The KAPs values will be employed to calculate Keyword Similarity Measure (KSM), which shows the degree of keyword similarity between an existing scenario item and the present Q-A item. The KSM equation takes into consideration the frequency of word for a particular keyword *i* that can match with keywords for a particular scenario item. This equation makes an assumption that every word for a particular keyword i has the same confidence value (e.g. keyword "Heart Attack" with weight "0.98"; Therefore, the word "Heart" and "Attack" have "0.49" as their own confidence). The calculations of both the Scenario Attraction Factor (SCF) and Keyword Similarity Measure (KSM) for a keyword i are explained in Equation 4.4. The SAF value is greatly dependent on the average of the total Keywords Similarity Measure and SCF value for a scenario.

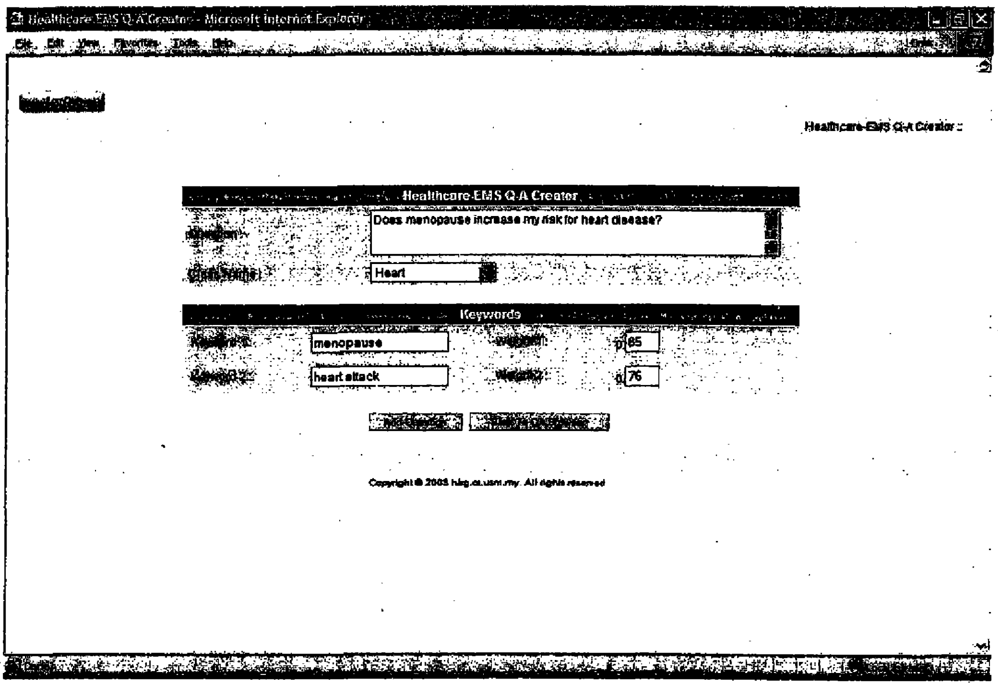


Figure 4.13: Question interface with keywords and weights

The same process will be performed when 'attracted' experts post their answers. The answer submission interface is almost the same as the one shown in Figure 4.12 except that there is extra information needed (e.g. reference link). The keywords specified for the answer is used to 'attract' yet scenario contributors with the knowledge/experience denoted by the keywords, to give their solutions to the problem and give further information for the answer (which would invite them to give their answer to form Q-A items). As an example (see Figure 4.12), the keywords stated are "menopause" and "heart attack" with weights "0.86" and "0.76" respectively. The calculations involved will take scenario item "s.20000613.0540" as an example and this scenario item has two keywords, which are "angioplasty" and "heart attack".

By applying Equation 4.4 (a) and Equation 4.4 (b),

Let Keyword 1 = menopause,

Keyword 2 = heart attack,

Scenario crystallization factor for s.20000613.0540 = 0.7909

$$P_1 = 1-\left(1-\frac{1}{1}\right)^0 \times \frac{0.86}{1}$$

$$= 0$$

$$P_2 = \left\{1-\left(1-\frac{1}{2}\right)^1 \times \frac{0.76}{2}\right\} + \left\{1-\left(1-\frac{1}{2}\right)^1 \times \frac{0.76}{2}\right\}$$

$$= 0.38$$

$$P_{s.20000613.0540} = \frac{0.38+0}{0.86+0.76}$$

$$= 0.234$$

$$\text{Scenario Attraction Factor} = \frac{0.234 + 0.7909}{2}$$

$$= 0.512$$

Thus, the scenario attraction factor for scenario item "s.20000613.0540" is 0.512. As a conclusion, the scenario attraction factor not only dependent on keyword similarity measure alone. Scenario crystallization factor (value determined by its applicability, correctness, appropriateness and completeness) also plays a crucial role in determining the attraction process. If the attraction threshold is predefined as 0.5, then an invitation email will be sent to expert with *userID* "4". From the calculations above, we can observe that if the frequency of a word is higher then the computed

Keyword Similarity Measure will be greater. Therefore, the possibility of experts' attraction will be higher as well.

CHAPTER 5

SUMMARY AND FUTURE WORK

Ultimately, we have implemented an enterprise memory system for healthcare knowledge management system, namely Healthcare Enterprise Memory System (Healthcare-EMS). After providing the implementation details with some example results, we are now going to evaluate the Healthcare-EMS. This evaluation would not be quantitative in nature, i.e. in terms of speed/performance in view that contributions are not algorithmic in nature or process oriented. Rather, we would present comparisons (advantages and/or disadvantages) of the Healthcare-EMS with other similar systems.

Before we do so, we are going to revisit the objectives, which we stated at the beginning of this report (see section 1.6) and hence, we will then evaluate our contributions for this project. In view that more can be done in the area of enterprise memory and that the Healthcare-EMS can be expanded further, we will conclude with some suggestions for future work.

5.1 REVISITING OUR CONTRIBUTIONS

Earlier in Chapter 1, we stated three main contributions of this research project: -

1. The development of knowledge representations formalisms called *scenarios* and *Q-A* that models both the procedural knowledge and problem-solution knowledge as two distinctive hierarchical structures.

2. A hybrid knowledge acquisition methodology, which makes use of both the scenario crystallization with scenario attraction strategies (combines scenarios and Q-A in the sense that

they are linked via keywords) to 'provoke' the experts to give their solutions for a particular problem via notification through email.

3. A knowledge maintenance mechanism to preserve the 'freshness' and applicability of the knowledge, which resides in the repository of the Healthcare-EMS. This knowledge maintenance mechanism is solely dependent on the rating process, which reflects the applicability, correctness and completeness of the knowledge items.

5.1.1 Revisiting the First Objective

The first objective as stated early in this report (see section 1.6 in Chapter 1) is the definition of suitable knowledge representation formalisms for two distinctive forms of knowledge, which are procedural knowledge and problem-solution knowledge.

Our scenario representation formalism for procedural knowledge is a simplified hierarchical structure, simpler when compared with the *scenarios* in the tacit knowledge strategy suggested by Cheah and Abidi (2001). We argue that the simplified scenario representation formalism facilitates the navigation and exploration of problem situations at a less fine-grained descriptive level (event level) which is favourable rather than being too fine-grained.

The Q-A representation for problem-solution knowledge can be viewed as hierarchical too and hence, it provides an economic and scalable formalism.

As a conclusion,

1. ***The scenario representation is simple but multi-purpose:*** Its simplified structure illustrated in Figure 3.2 enables the situational experience to be stored as a *scenario.* This simplistic knowledge representation structure facilitates the scenario crystallization and scenario attraction processes, which ultimately permit the knowledge acquisition methodology.

2. ***The Q-A representation is simple and scalable:*** Scalable in term of its ability to grow continuously whenever a question or an answer is added to the knowledge repository. Definition of keywords and weights (explained in section 4.2.2.5) allows the process of scenario attraction hat invites experts who have previously submitted similar knowledge in the form of scenarios to give their solutions in Q-A format. In addition, the *Q-A* representation enables experts to explicate their knowledge in free-text format that is different from *scenarios*. Knowledge explication can be performed for both procedural knowledge and problem-solutions knowledge. However, their scalability allows for the expansion of the knowledge representation to accommodate more sophisticated knowledge creation functions in future.

5.1.2 Revisiting the Second Objective

The second objective of the system is to introduce knowledge acquisition methodology to enable the repository of a healthcare enterprise memory system to grow continuously. As we can see, the traditional knowledge acquisition techniques seem to be ineffective in both knowledge explication and acquisition. This is due to the fact that knowledge manifests itself in situational descriptions especially those knowledge that are novel, atypical and incomplete situations.

In addressing this difficulty, both the knowledge representation formalisms have been manipulated to explicate knowledge in procedural and problem-solution format, which provides alternatives to explicate tacit knowledge.

Knowledge sharing is not a common culture in an enterprise because knowledge is considered as the intellectual property of an expert. Therefore, knowledge acquisition methodology must be used as a way to 'populate' the knowledge items in the Healthcare-EMS and motivating employees in an enterprise to share their knowledge by having incentives scheme (i.e. in terms of reward based on the assigned bonus points as explained in the Knowledge Rewards section). In the knowledge acquisition process, new knowledge will be generated when experts who previously submitted similar knowledge/experience (in the form of scenarios) are informed to give their answers/solution in Q-A format. The attraction of experts to contribute their solutions is basically dependent on the keyword similarity measure and scenario crystallization factor. However, this knowledge acquisition process needs plenty of scenario items to have keywords in order to find similar scenarios.

Overall, we can conclude that:

1. ***Knowledge acquisition can be achieved under certain conditions:*** Knowledge acquisition can be achieved at an optimum level if there are plenty of *scenarios* items in the Healthcare-EMS because knowledge acquisition depends on the keywords similarity degree and the scenario crystallization factor to grow. Keywords similarity degree will not be able to find their matched keywords with scenarios if there are only a limited number of *scenarios* in the system and would result in the knowledge acquisition process being less effective. Users' rating for the events of a scenario also affects the knowledge acquisition process. Thus, the performance of the knowledge acquisition methodology also greatly depends on the wisdom of user to rate this knowledge.

2. ***Knowledge explication in a systematic manner involves a trade-off in tacit knowledge content:*** When experts are provoked to explicate their knowledge in a systematic manner, it does not remain tacit because any form of formalism is explicit in nature. Therefore, to soften the impact of loosing too much tacit content caused by the structured nature of scenarios, another knowledge representation formalism is introduced to permit knowledge to be captured in a free-text format, i.e. in the form of Q-As Therefore, we can argue that *scenarios* and *Q-A* are collaborating with one another, more like a symbiosis rather than hybrid, to achieve better knowledge explication and knowledge acquisition.

3. ***Experts' wisdom determines the optimum operationality and functionality of the system:*** Knowledge acquisition of Q-A knowledge items requires definition of keywords and weights. Experts' wisdom in determining keywords and weights are very crucial at this stage. Thus, they should be very aware in specifying the keywords and weights for the knowledge item, as this will determine the practicability of the system.

5.1.3 Revisiting the Third Objective

By looking back to the third objective, we can argue that the knowledge maintenance is a vital process in a knowledge management system. The applicability and 'freshness' of the knowledge residing in the system will affect the diagnosis and treatments performed by the experts (e.g. medical officers) in a healthcare enterprise. Inaccurate or outdated knowledge may cause fatal treatments to be prescribed or ambiguities in diagnosis. Consequently, outdated knowledge (which already notified because of its inapplicability but no actions are taken) must be removed from the knowledge repository to prevent incorrect, incomplete and inapplicable knowledge/experience to be shared among the experts.

With regards to the knowledge maintenance process, we conclude that:

1. ***Knowledge maintenance only practiced in the Q-A base(s):*** We argue that knowledge maintenance process should determine the applicability of the knowledge in the system because it takes a number of ratings to assess the applicability of a *scenario.* Incompleteness of user's rating will affect the ultimate evaluation and therefore, the possibility of *scenarios* to be removed is much higher than a Q-A answer/solution (knowledge). Since the knowledge acquisition relies solely on the 'population' of the

scenarios, it is advisable not to practice the knowledge maintenance process with scenario base(s). Thus, the knowledge maintenance process only caters for the Q-A base(s).

2. ***Knowledge applicability and 'freshness' can be maintained with the scheduled update reminder:*** Since the evaluation of knowledge keeps on changing everyday, there exist a update reminder that is scheduled to keep track of the applicability, correctness (based on users' evaluations) and the difference between the last modification date and present date. The update reminder will remove the knowledge permanently if necessary actions are not taken after the outdated knowledge notification is sent. This agent collaborates with the personal update reminder to ensure that experts can benefit themselves with this useful knowledge.

5.2 Comparison with other Similar FAQ System

5.2.1 FAQ Finder System

There exists a FAQ system, namely FAQ Finder System that is quite similar to the Healthcare-EMS system from the knowledge representation structure. FAQ Finder system uses AI approaches to automatically find the right answer. The FAQ Finder system will return the right answer taking into account the following assumptions [13]:

- QA Format: All the information in the FAQ file is organized in question/answer format.
- Locality of information: All the information needed to determine the relevance of a question/answer pair could be found within that question/answer pair.
- Question relevance: The question portion of the question/answer pair is the most relevant for determining the match to a user's question.
- General knowledge: Broad general knowledge is sufficient for question matching.

The similarity between the Q-A knowledge structure and the QA Format for FAQ Finder system is that they are made up of questions and answers. However, our Q-A knowledge structure allows multiple answers to be associated with one question.

The answer acquisition approach used is different from one and another. The FAQ Finder system makes use of the Case-Base Reasoning (CBR) approach where the question serves as an index to the knowledge contained in the answer. The use of CBR is based on an assumption that there exist a single QA pair that best matches the user's query [13]. This system will find the right answer by comparing the question with the question in FAQ files. But in our system, we are using keywords matching to measure the similarity of keywords for two different knowledge representations, i.e. linking the Q-A with scenarios, and therefore invite the scenario knowledge contributor to give a solution to the question. However, there is a possibility that no experts are invited to share their knowledge/experience. The CBR approach is not suitable for our system because there can be many different medical terms that refer to the same thing. Nevertheless, both of the systems also rely on another file (data) in order to perform at an optimum level. In another words, the FAQ Finder system rely on a FAQ files in order to get an answer for a frequently-ask-question whereas our system rely on scenario's keywords to indirectly get an answer to form Q-A knowledge.

Both of the systems employ words similarity calculation to measure the similarity between the questions but in our case, these are keywords. We utilize equation 4.4 to calculate the degree of similarity whereas the FAQ Finder system makes use of the equation (see Equation 5.1) below:

$$m = \frac{tT + wW + cC}{T + W + C}$$

where:

T, W, C are weights

t - statistical term-vector similarity score (TFIDF)

w - shallow level of semantic analysis (variations in lexical content), knowledge taken from WORDNET(semantic network of English words - words and synonyms, is-a links)

c - degree of coverage of user terms by the FAQ question (penalty, if Q doesn't correspond to words in user request)

Equation 5.1: Overall match similarity for FAQFinder System [13]

One of the advantages of this system is its ability to apply for a wide range of generic questions. In addition, it does not require user to specify the relevancy (in term of weights) of question with those Qs in FAQ files (FAQ Finder syetem's knowledge base). The FAQ Finder system employ the marker passing technique to compute a similarity score for two questions. Marker parsing is performed to compare each word in the user's question with each word in the FAQ files. Therefore, this technique requires a lot of computations especially there are plenty of QA pairs in the FAQ files. It will be even worse if these questions in the FAQ file are lengthy. Unlike the FAQ Finder system, Q-A structure in Healthcare-EMS only compares each keyword for Q- with each keyword for the scenario in scenario base and the weights for each keyword is defined by a user's question/answer. However, these two systems bear the risk that they might not be able to find the correct answers in their knowledge base. Overall, the approaches used by the FAQ Finder system are not suitable for Healthcare-EMS as each question/problem in healthcare domain need distinctive treatment/solution from medical officers.

Some of the interfaces of the FAQ Finder system are shown in Appendix B to give a better illustration of how this system operates.

5.2.2 Answer Garden

As mentioned before, Answer Garden is a tool for growing organizational memory as well as to facilitate a cooperative work that helps improve an enterprise's memory. However, this system does reveal some similarities and distinctions with the Healthcare-EMS. First of all, Answer Garden is using a branching network of diagnostic questions. The AG permits enterprises to develop a database of commonly asked questions, which grows "organically" as new questions arise and are answered [5]. Each question is associated with a few simple answers or in another words, the knowledge is represented using Q-A format too but each answer leading to different subsequent questions. Healthcare-EMS allows discription of possible problem and solutions in a free-text format where these solutions can be fairly complicated answers. This approach permits experts to explicate their knowledge using their natural language without too much restrictions (which may result in loosing of knowledge's tacitness).

These two systems are routing questions/problems to appropriate experts and the answers provided by them will be inserted in the knowledge base ultimately. The approaches used by these systems is able to overcome the possibility of questions which cannot be found in the system's knowledge base. Therefore, they make it possible for the population in knowledge base to grow over time and both knowledge acquistion processes are initiated by the users. The routing of questions to experts in AG does not based on the knowledge which the experts previously provided. In contrast, Healthcare-EMS 'attracts' experts to give solutions by comparing scenario's keywords with the question's keywords. Besides, an answer of Q-A item too can initiate

In term of knowledge maintenance, modifications of diagnostic barnching in Answer Garden is performed in response to user's feedback. Similarly, Healthcare-EMS will inform knowledge owner to update/modify their contributions based on the users' evaluation. Both the approaches work to maintain the incomplete or out-of-date answers and hence, making sure that the knowledge bases in both systems are applicable and correct.

Apart from that, the type of questions in AG's repository is commonly-asked (general) questions whereas every question in the Healthcare-EMS's Q-A base can be very unique, along with a list of keywords and weights to describe the relevancy of keywords with the question.

5.3 Future Works

Overall, we might argue that our system still imperfect because of a few weaknesses which may be improved/enhanced in future. To conclude, we would like to present possible improvements or extensions to our work that we believe would extend the functional limits of the proposed Healthcare-EMS:

1. ***Distributed knowledge crystallization and scenario bases:*** If there are plenty of scenario bases in a distributed network, it would be advantageous to permit the system to capture all the *scenarios* from individual scenario base in every personal computer in the Local Area Network (LAN). Thus, the system will not have to rely on the experts to log on to the system to fill in the knowledge acquisition form or the centralized scenario base(s). In addition, this will provide the Healthcare-EMS with enough *scenarios* for the Q-A knowledge acquisition process.

   In a client-server environment, we believe that the distributed crystallization of scenario bases too could be performed effectively. It would prove to be a more efficient means to enforce knowledge crystallization and knowledge acquisition. When a question or an answer is created, a knowledge acquisition task can be replicated at other scenario bases to attract more experts to participate in giving their solutions/answers. This saves time and effort as the need to performed knowledge acquisition process at the other scenario bases individually is eliminated.

2. ***Inclusion of Document Base and a Contextual Document Pre-processor component into the Healthcare-EMS system:*** The system will perform better by including a document base at the knowledge object layer because this transformation of unstructured documents in document base is believed to enrich the 'population' of scenarios. This document base is made up with unstructured documents, which will need necessary preprocessing to extract the contextual knowledge and restructure them to become *scenarios*.

Hence, the hybrid knowledge representation formalism comprises of the structured scenario representation (supplemented by preprocessed unstructured documents from document base(s)) and Q-A representation will become the core repository of the Healthcare-EMS. The inclusion of document base will enrich the knowledge repository of the system. The illustration above (see Figure 5.1) shows the heterogeneous repository (document base(s)) and hybrid representations (scenario and Q-A). From Figure 5.1, there is a preprocessor (Contextual Document Pre-processor component) used to extract the contextual knowledge from unstructured documents. Extraction of the knowledge will turn them into scenario representation structures.

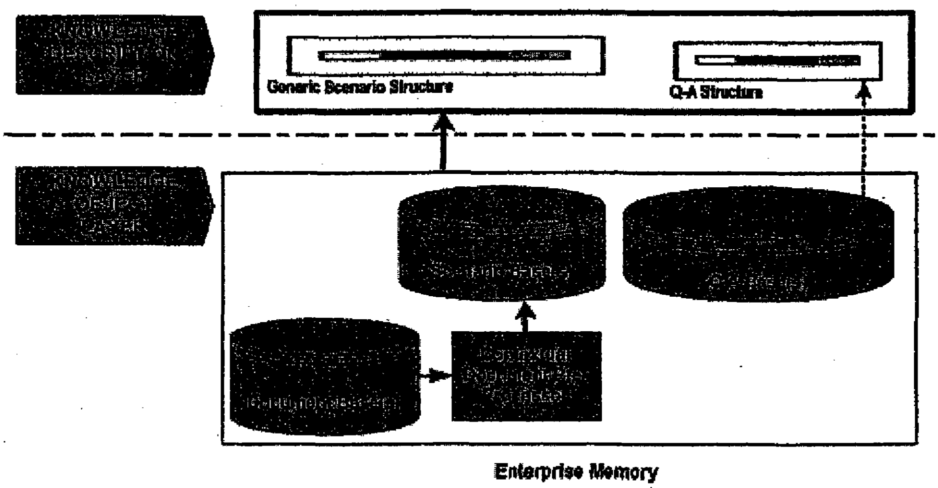


Figure 5.1: Heterogenous repository (Document Base(s)) and Hybrid representation

3. ***Incorporation of Knowledge Standardization and Knowledge Repair applications:*** Knowledge standardization is a crucial issue in healthcare domain because there exists many different medical terms (synonyms) that actually refers to one single thing. The computation of keywords similarity degree is very sensitive towards these synonyms as it may get different final result if different medical terms are specified as keywords for a Q – or A- item. Hence, knowledge standardization must be carried out to 'unify' the medical terms used by experts in a healthcare enterprise.

   Besides, a Knowledge Repair application can be utilized to repair the scenario knowledge items because *scenarios* have the tendency to be inapplicable and incomplete. This application can adopt case-base reasoning approach as explained is the in FAQ Finder system earlier. We believe a Knowledge Repair application can help to maintain the 'freshness' and applicability of *scenarios* in the Healthcare-EMS, as users' perception itself might not determine precisely the correctness and applicability of knowledge items in the system.

4. ***Extension of Knowledge Acquisition for scenarios:*** In addressing the weaknesses of Healthcare-EMS, it would be better if knowledge acquisition methodology can cater for both scenarios and Q-As. Knowledge acquisition for scenario needs experts to manually key in their knowledge into the Healthcare-EMS. Therefore, an automatic knowledge acquisition methodology must be designed such that it will populate the number of scenarios in the system. This knowledge acquisition might be in the form in an e-mail that will request experts to fill in their knowledge. The knowledge can be the follow ups (events) of a patient and finally form a scenario. Therefore, they need not to log on to the Healthcare-EMS to fill in the scenario form. The procedural knowledge can be created locally in a personal computer that has no Internet services available temporarily. After this completed form is sent back to Healthcare–EMS, the system will keep the knowledge into

the scenario base. This methodology is attempting to solve the difficulties encountered by experts because they might not be free to fill the form when they are connected to the system.

5.4 Conclusion

Healthcare-EMS is an enterprise memory for knowledge management system. This system comprises of scenario base and Q-A base as their knowledge bases. Knowledge acquisition has been using these knowledge items to expand the 'population' of knowledge base but in my system, it only caters for Q-A base. This automatic knowledge acquisition makes use of the keywords and their respective weights to attract experts who have submitted their similar scenarios previously. However, we can argue that this methodology is better because users can describe the relevancy of the question they wanted with the scenario in terms of keywords and weights.

However, the power of our approach rises, as we are not using a pure question and answer pair. The Q-A structure is designed such that this question and answer are not purely in free-text format since some useful information is included (i.e. useful reference link). Knowledge maintenance is making use of e-mail notifications method to prevent outdated or inapplicable knowledge to stay in the knowledge base under certain conditions.

Ultimately, Healthcare-EMS will appear to be a powerful system by incorporating some of the applications, which might be improving its functionality. We believe, the enrichment of the enterprise memory (knowledge base) is the most important factor overall. All the applications are attempting to keep the repository such that it contains all the applicable and constructive knowledge. This knowledge is important as it allows people to learn from past and finally the knowledge sharing culture in an enterprise. Thus, the manipulation of this knowledge might benefits both the employers and employees in an organization.

REFERENCES AND BIBLIOGRAPHY

[1] E.Jeffrey Conklin, PhD, "Capturing Organizational Memory", GDSS:Working Papers (Samples of relevant research by GDSS associates), http://www.gdss.com/wp/COM.htm

[2] Ackerman, Mark S., and Eric Mandel. 1995. Memory in the Small: An Application to Provide Task-Based Organizational Memory for a Scientific Community. *Proceedings of the Hawaii International Conference of System Sciences (HICSS 95)*: 323-332.

[3] Anne-Marie Alquier – Marie-Hélène Tignol and Jean-Pierre – Anne Hernandez-Jones, "Corporate Knowledge and Memory: Combining Human Extraction And Machine Storage. How to Enhance Corporate Productivity".

[4] Andreas Abecker, Ansgar Bernardi, Knut Hinkelmann, Otto Kühn and Michael Sintek, "Toward a technology for Organizational Memories", *IEEE Intelligent System and their application*, May/June, 1998.

[5] Mark S. Ackerman and Thomas W.Malone, "Answer Garden : A Tool for Growing Organizational Memory", *Proceedings of ACM Conference on Office Information System* (Cambridge, MA, April, 1990), pp 31-39.

[6] C. Wargitsch, "WorkBrain: Merging Organizational Memory and Workflow Management Systems", *Knowledge-Based Systems for Knowledge Management in Enterprises Workshop, 21st Annual German Conference on AI '97 (KI-Jahrestagung '97)*, 1997.

[7] Ronald K.Maier and Oliver W.Klosa, "Organizational Memory Systems to Support Organizational Information Processing".

[8] Mark S.Ackerman, David W.McDonald, "Answer Garden 2: Merging Organizational Memory with Collaborative Help", *Proceedings of the ACM Conference on Computer-Supported Cooperative Work* (CSCW'96).

[9] M.S. Ackerman, "Augmenting the Organizational Memory: A Field Study of Answer Garden", *ACM Conference on Computer-Supported Cooperative Work* (CSCW94), Nov. 1994, pp. 243-252.

[10] David G. Schwartz, "Towards the Use of User-Centric Meta-Knowledge in Applying Organizational Memory to Email Communications ".

[11] Cheah, Y.-N. and Abidi, SSR , "A Knowledge Creation Strategy To Enrich Enterprise Information Systems With Enterprise-Specific Tacit Knowledge", *ICEIS 2001*.

[12] Daniel E.O'Leary, "Knowledge-Management Systems: Converting and Connecting", *IEEE* Intelligent Systems and their applications, May/June 1998.

[13] Robin D.Burke, Kristian J.Hammond, & Vladimir Kulyukin, Steven L.Lytinen, Noriko Tomuro, & Scott Schoenberg, "Question Answering from Frequently-Asked Question Files: Experiences with the FAQ Finder System", December 20, 1996.

Appendix A

Table 1: Description of five categories of EMS [7]

Category of EMS	Description
Corporate Knowledge Repository/ Meta-knowledge System	These systems provide fast retrieval and easy access to existing information within the company [11]. Besides, they also give an overview of the existing knowledge throughout the whole organization with help facilities.
Knowledge Agents	In these systems, knowledge is used actively and autonomously for intelligently searching information in other information systems (internal as well as external to the organization). The ability to learn is one of the typical prerequisite to be considered as knowledge agent.
Knowledge Bases, Expert Systems	These systems have been around for a long time and stored as well as process knowledge.
Knowledge and Communication Integration Platforms	These systems integrate several independent information systems, information types or communication systems.
Knowledge Creation and Structuring	These systems support groups in the creative process of generating and structuring knowledge as well as linking the newly found knowledge to existing knowledge.

Table 3: Simulation of SCF, EpCf and EvCF values based on Users' Votes

<u>Scenario Construct: s.19990713.1520</u> Votes: 0.64, 0.58, 0.66,0.55, 0.70, 0.71 Average Votes: 0.64 SCF = 0.69	EP0001 EpCF: 0.74	Votes: 0.60, 0.66,0.81, 0.88, 0.58, 0.77 Average Votes: 0.72 Average Events: 0.76		
	EP0002 EpCF: 0.65	Votes: 0.85, 0.73, 0.50, 0.60, 0.62 Average Votes: 0.66 Average Events: 0.64		
	EP0003 EpCF: 0.61	Votes: 0.42, 0.50, 0.45, 0.55 Average Votes: 0.48 Average Events: 0.74		
	EP0004 EpCF: 0.73	Votes: 0.84, 0.67, 0.82, 0.78, 0.79 Average Votes: 0.77 Average Events: 0.68	[redacted]	[redacted]
			EV0013 EvCF: 0.53	Votes: 0.50, 0.65, 0.50, 0.47 Average: 0.53
			EV0014 EvCF: 0.56	Votes: 0.61, 0.57, 0.55, 0.55, 0.47, 0.61 Average: 0.56
	[redacted]	[redacted]	[redacted]	[redacted]
			EV0017 EvCF: 0.80	Votes: 0.84, 0.76 Average: 0.80
<u>Scenario Construct: s.20000802.1928</u> Votes: 0.55, 0.81, 0.76, 0.82, 0.76 Average Votes: 0.74 SCF = 0.77	[redacted]	[redacted]	[redacted]	[redacted]
			EV0017 EvCF: 0.80	Votes: 0.84, 0.76 Average: 0.80
	EP0006 EpCF: 0.60	Votes: 0.51, 0.58, 0.62, 0.53, 0.56 Average Votes: 0.56 Average Events: 0.64		

Appendix B – User Interfaces

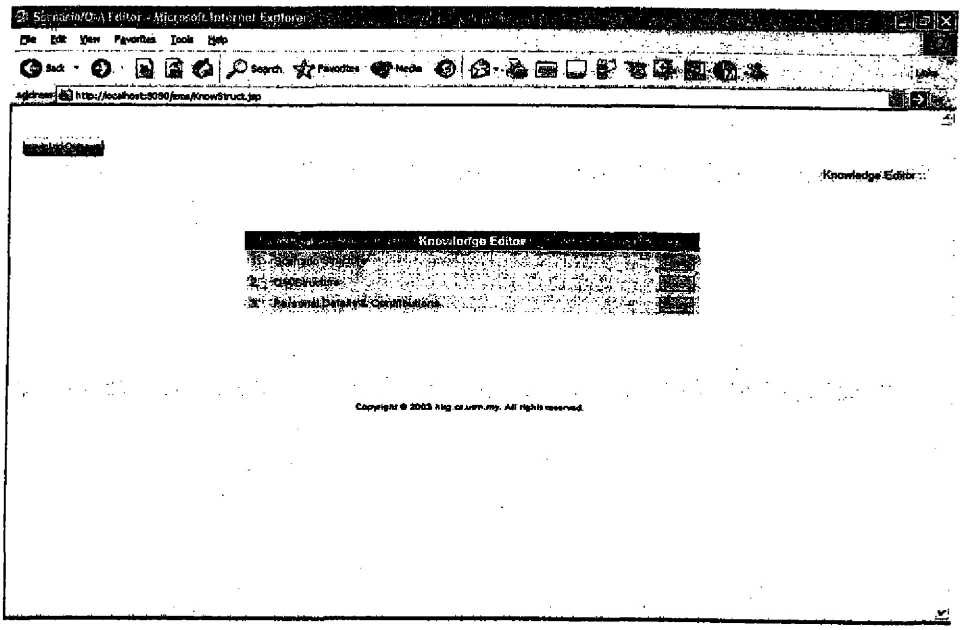


Figure B1: Knowledge Editor-Scenario Structure, Q-A Structure and Personal Contributions

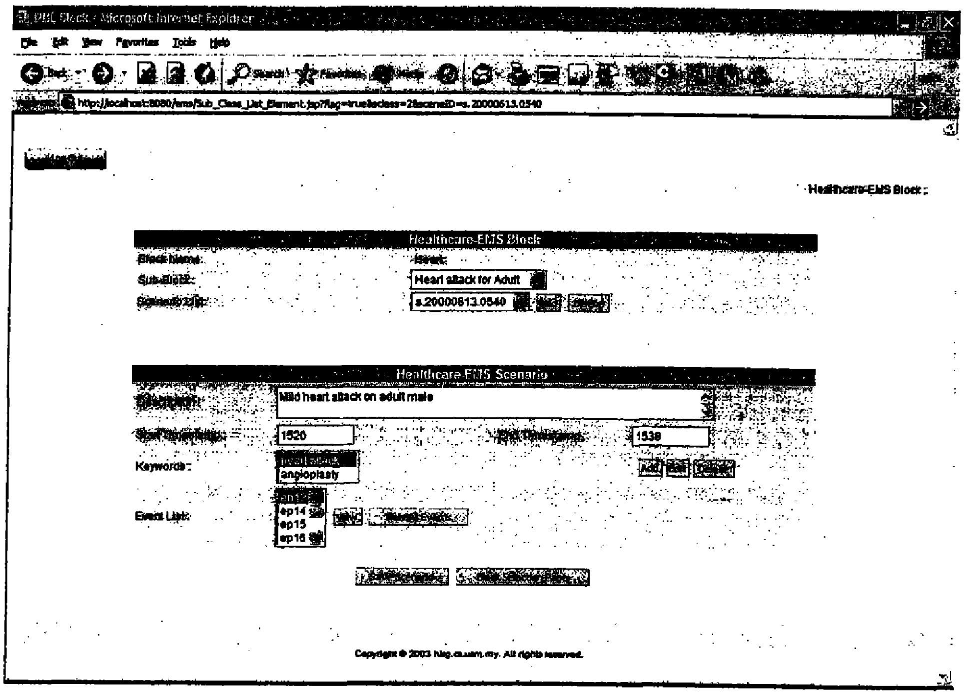


Figure B2: Scenario Structure -View Scenario Details Interface

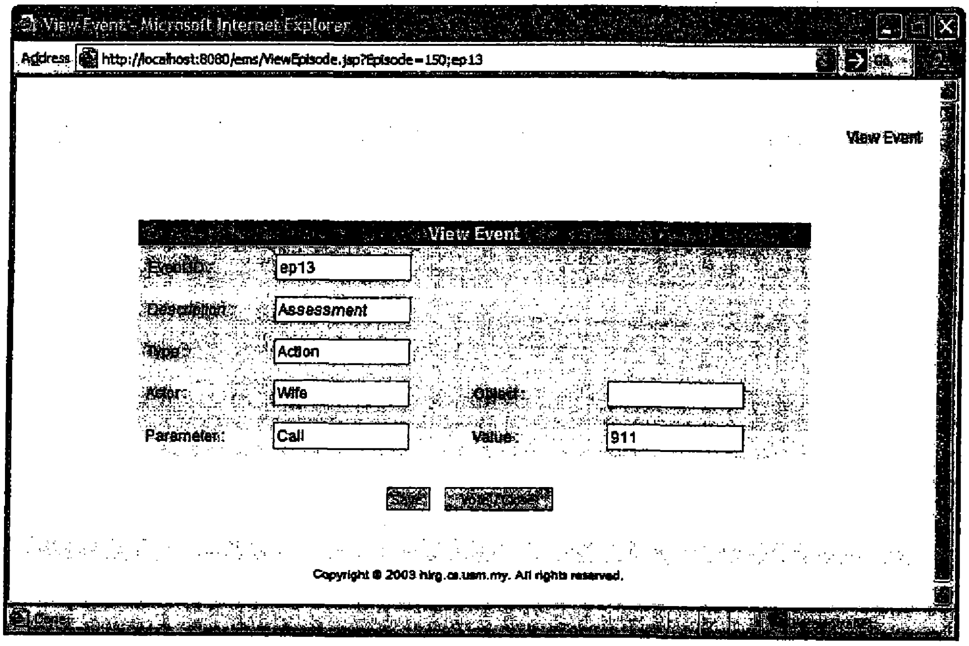


Figure B3: Scenario Structure - View Event's Details Interface

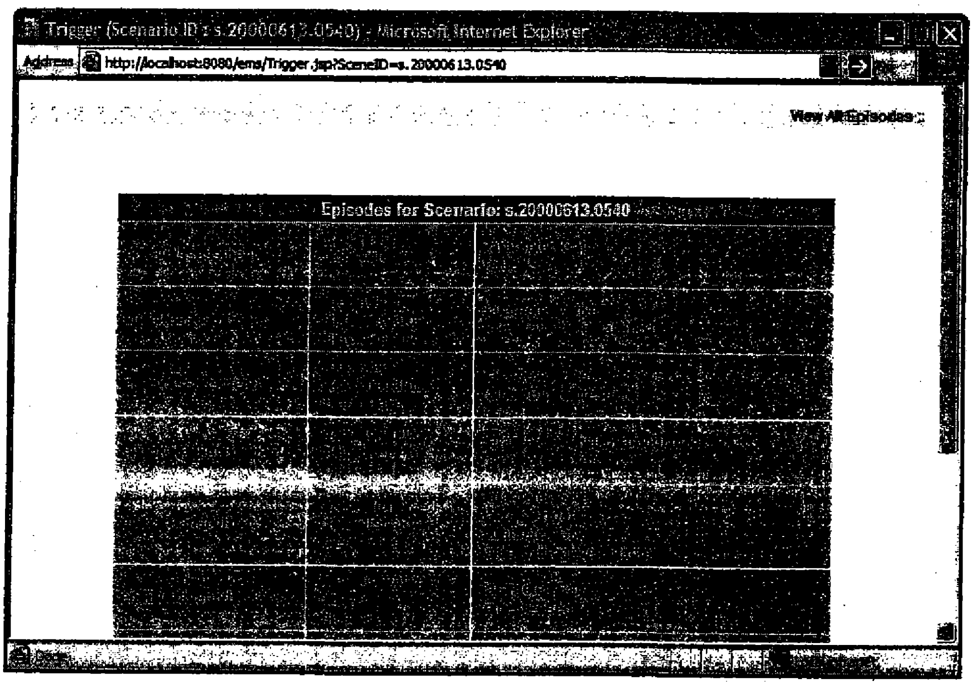


Figure B4: Scenario Structure - View All Events' Details for a Scenario (s.20000613.0540) Interface

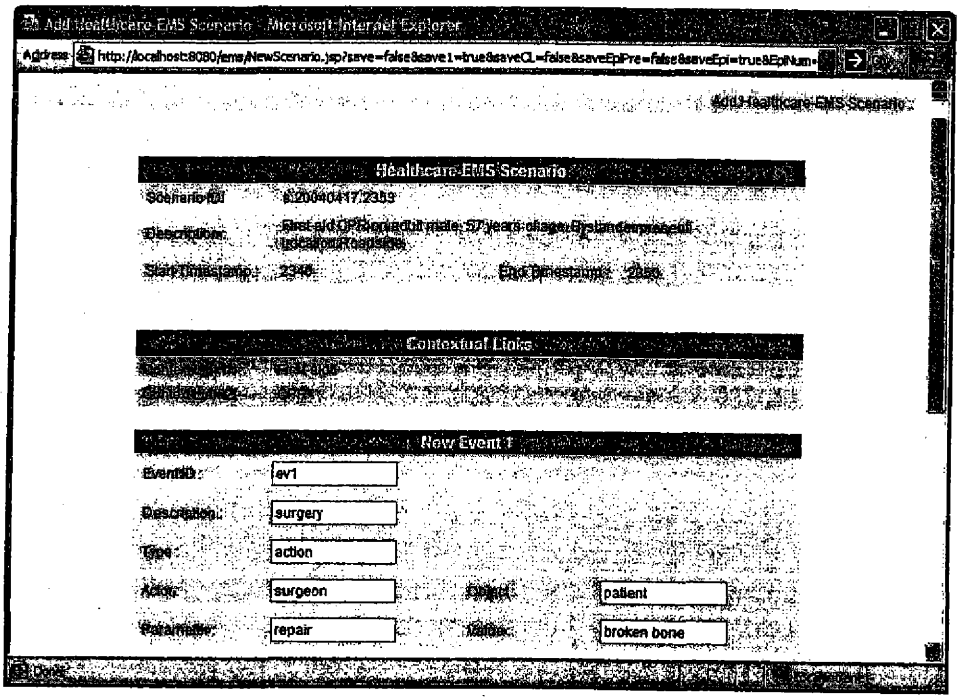


Figure B5: Scenario Structure - Add New Scenario Interface

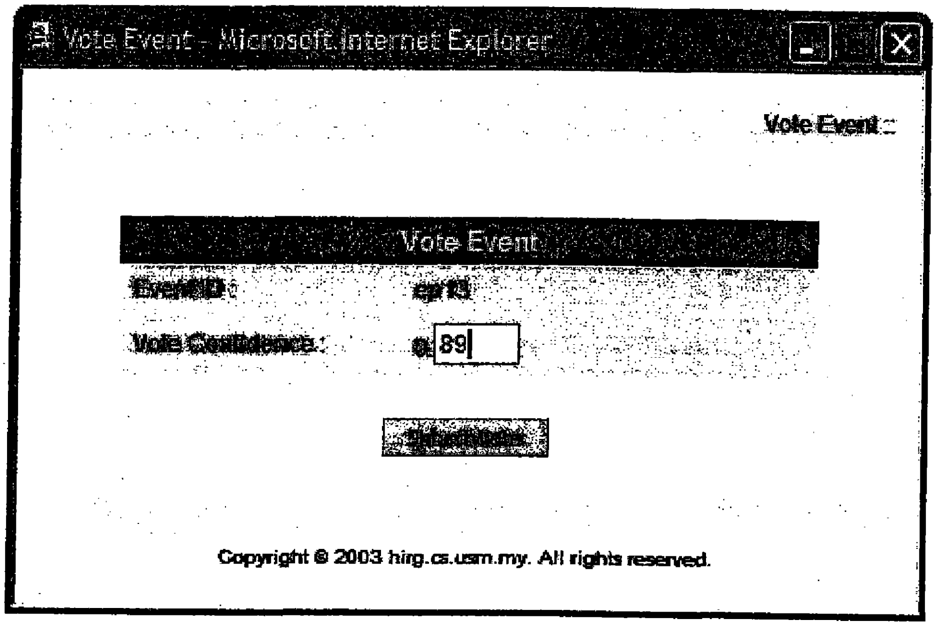


Figure B6: Scenario Structure – Vote for Event Interface

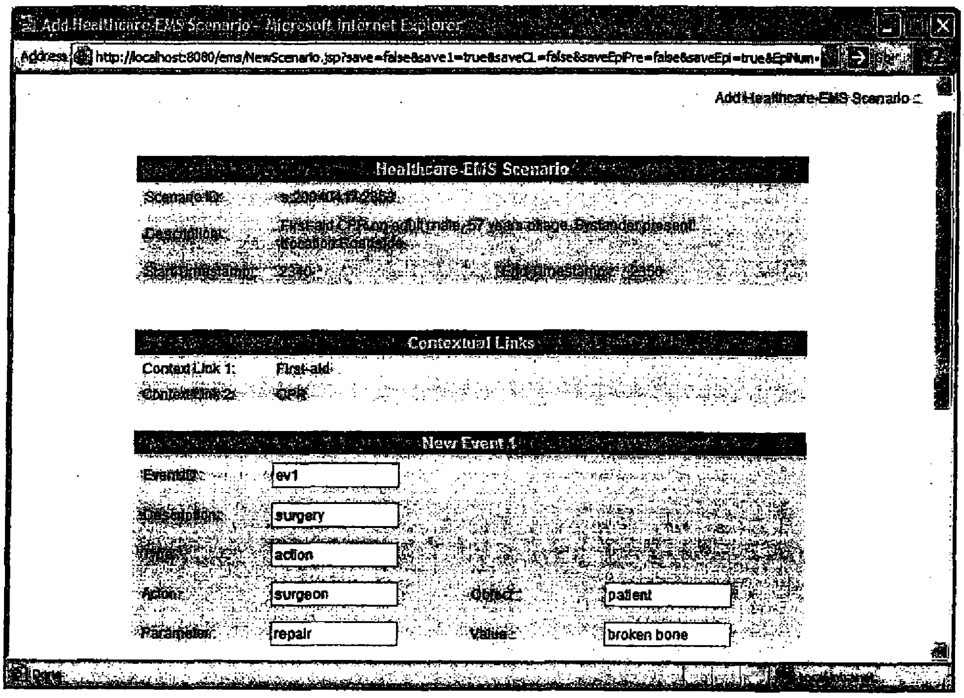


Figure B7: Scenario Structure - Add New Scenario Interface

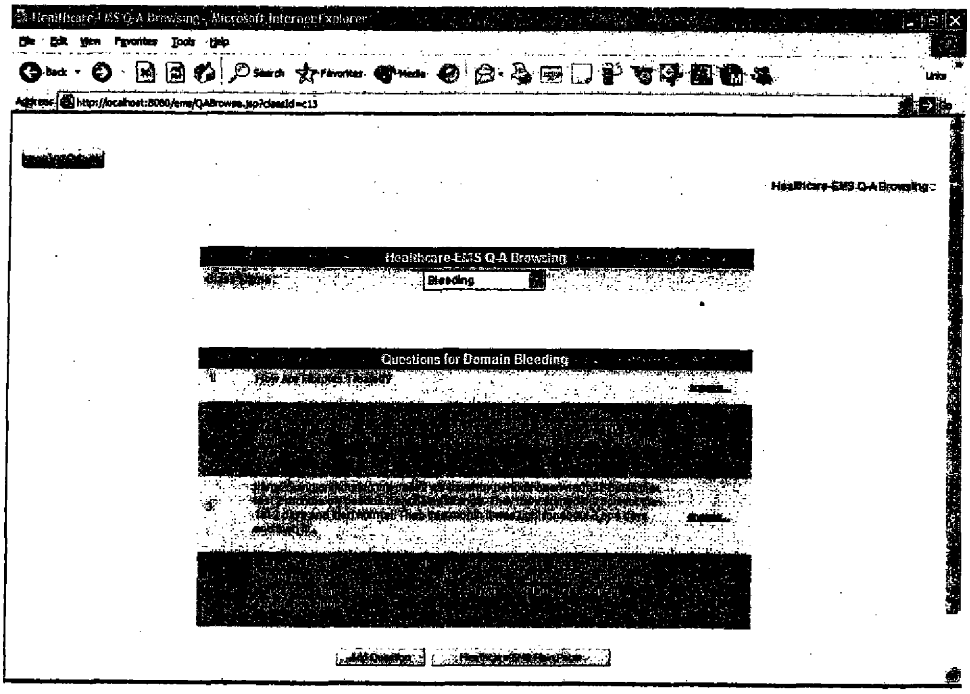


Figure B8: Q-A Structure – Browse Q-As Interface

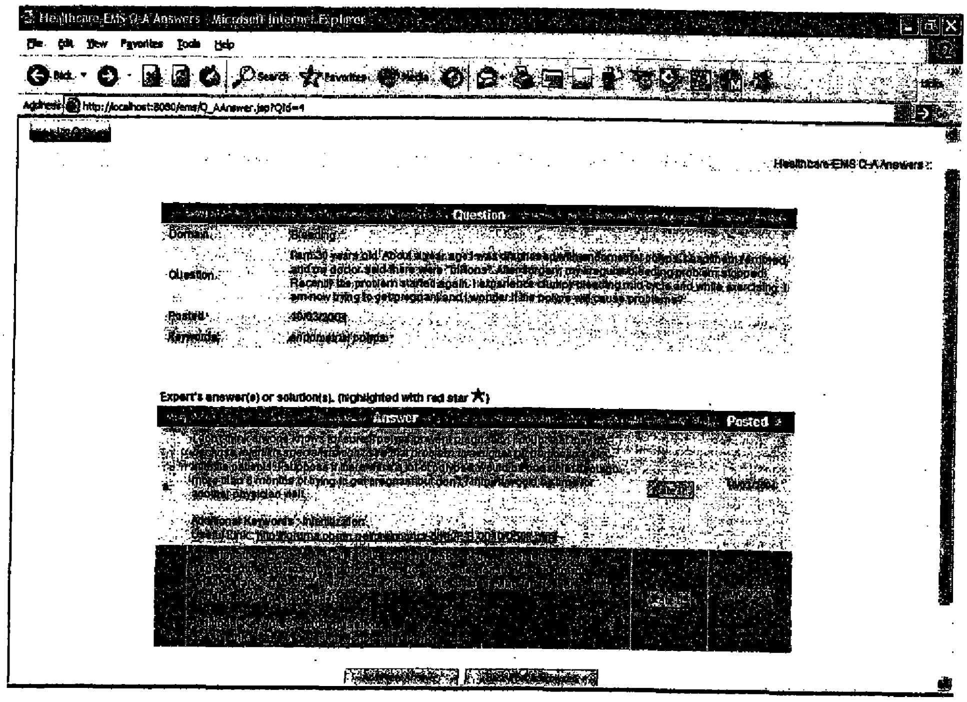


Figure B9: Q-A Structure – View Q-A Structure's Details Interface

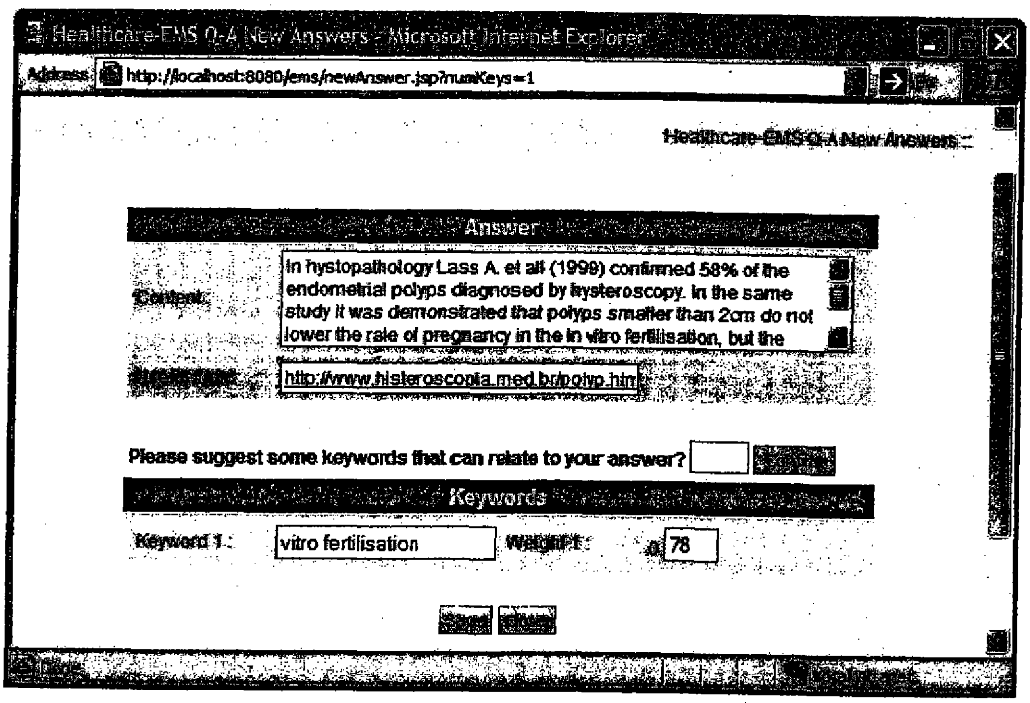


Figure B10: Q-A Structure – Add New Answer (With Keyword and Weight) Interface

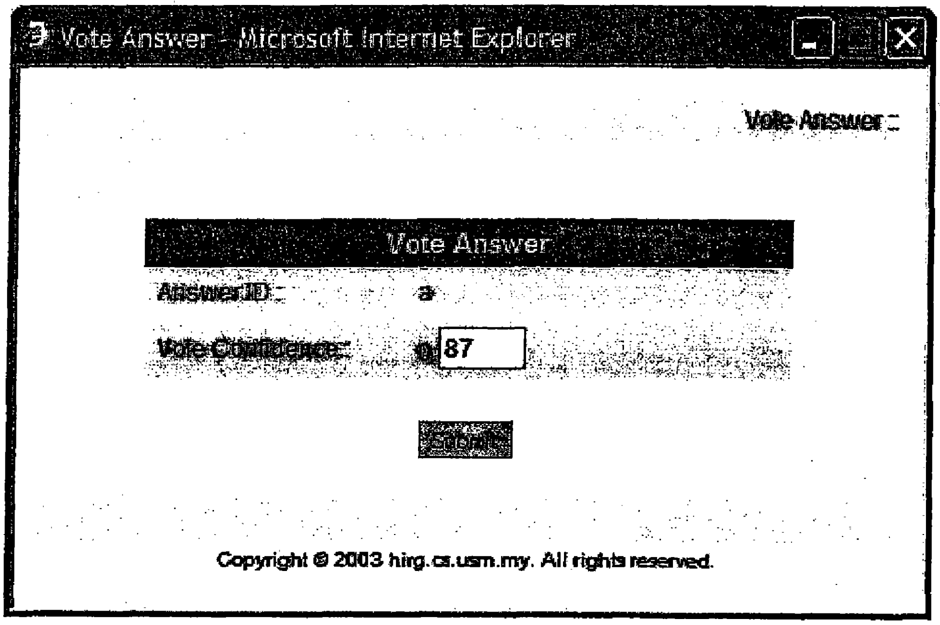


Figure B11: Q-A Structure – Vote for Answer Interface

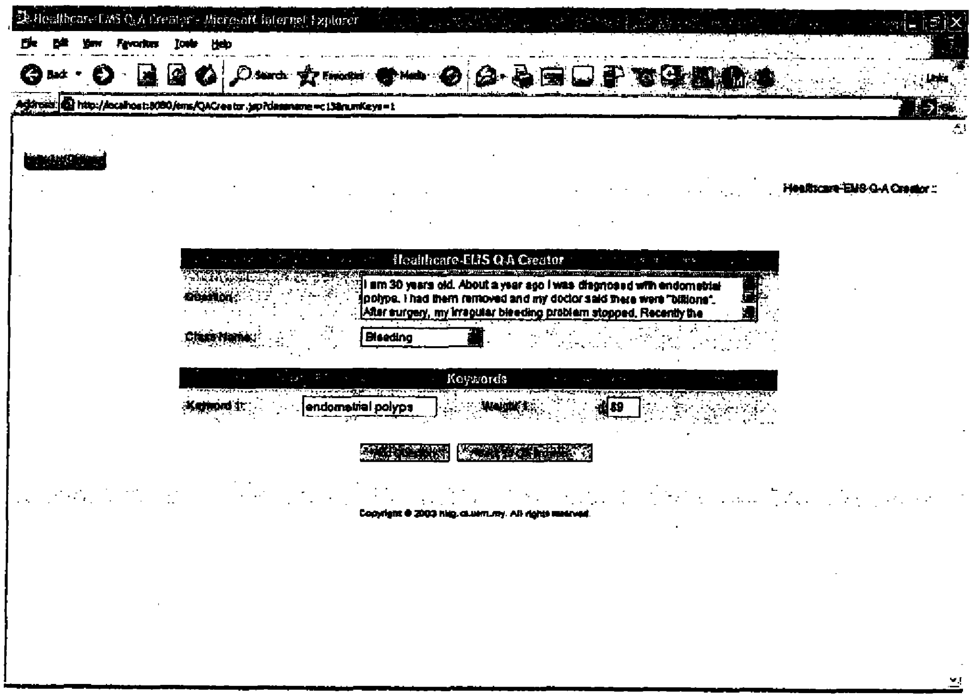


Figure B12: Q-A Structure – Create New Question Interface (With Keyword and Weight)

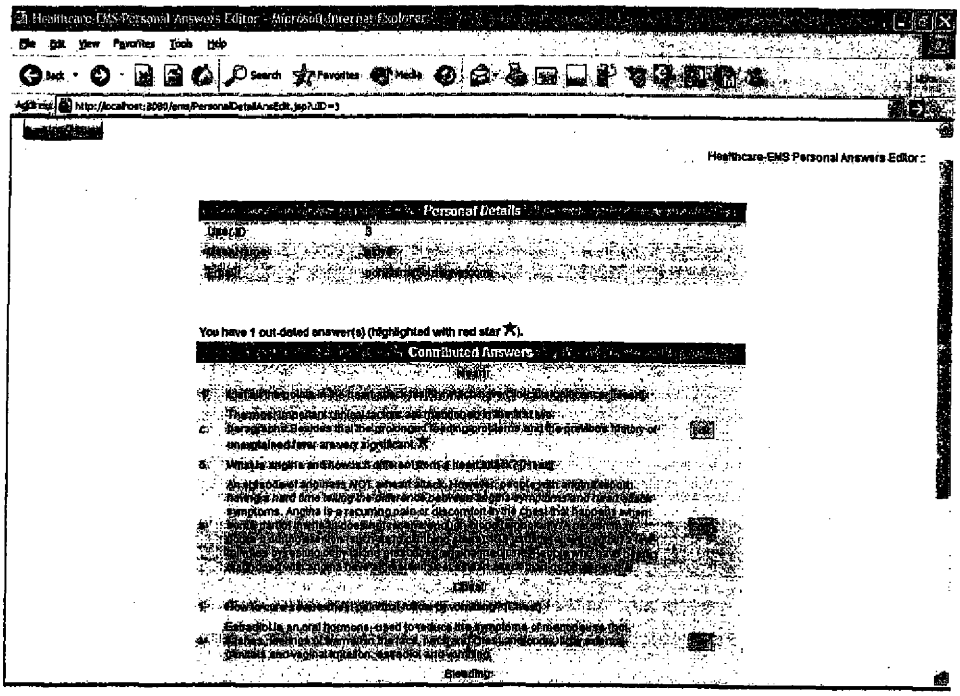


Figure B13: Personal Contribution - Update Reminder highlights outdated answer with red star

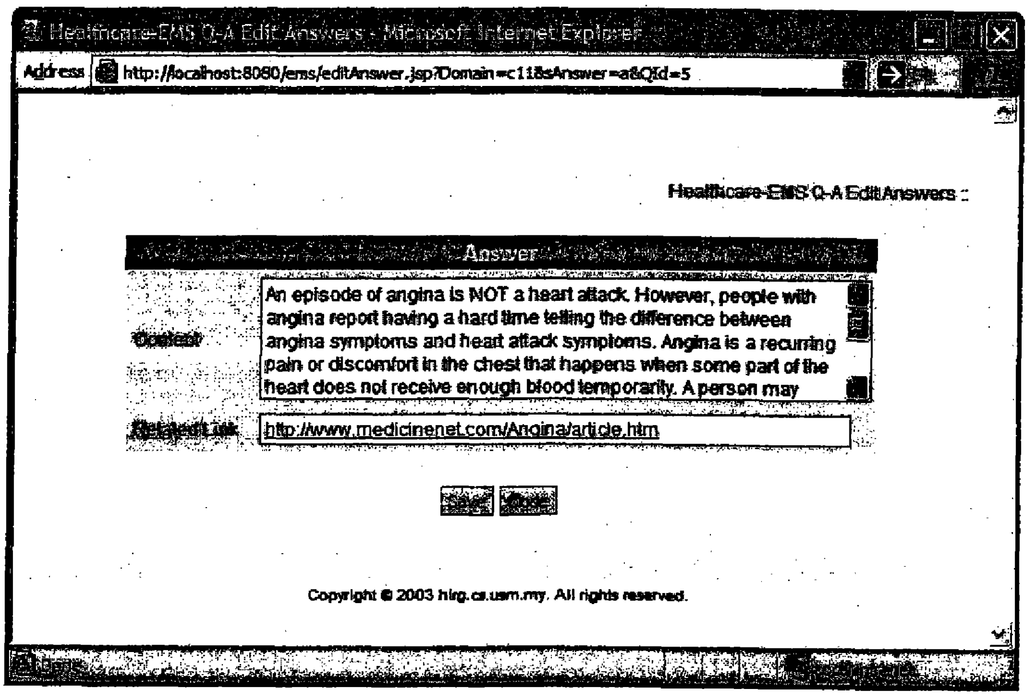


Figure B14: Personal Contribution – Modify Personal Contributed Answer Interface

Appendix B – XML files

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<login_construct>
        <login id="1">
                <name>yncheah</name>
                <password>yncheah</password>
                <emel>yncheah@cs.usm.my</emel>
        </login>
        <login id="2">
                <name>kwbeh</name>
                <password>kwbeh</password>
                <emel>kwbeh@cs.usm.my</emel>
        </login>
        <login id="3">
                <name>pklye</name>
                <password>pklye</password>
                <emel>pohkhim@biziwave.com</emel>
        </login>
        <login id="4">
                <name>phteoh</name>
                <password>phteoh</password>
                <emel>phteoh@cs.usm.my</emel>
        </login>
</login_construct>
```

Figure B15: *login.xml* – Records All The Users (Employees in an Healthcare Enterprise) Details

```
<question_list>
  <trigger_list domain="c11">
   <question id="1" date="02/10/2003">List all the points in the heart attack history which have clinical
   significance.
        <answer>
          <ans id="a" userID="1" date="22/02/2004">
              <cont>The most important clinical factors are
              mentioned in the first two paragraphs. Besides that, the prolonged feeding problems and
              the previous history of unexplained fever are very significant.
              </cont>
              <keywords>
                    <keyword weight="0.56">breathing</keyword>
                    <keyword weight="0.56">hand</keyword>
                    <keyword weight="0.56">compression</keyword>
              </keywords>
             <link>http://www.takeheart.co.uk</link>
             <voter id="1">0.56</voter>
             <voter id="2">0.98</voter>
             <avVote>0.6125</avVote>
             <voter id="4">0.67</voter>
             <voter id="3">0.24</voter>
          </ans>
      <keywords>
             <keyword weight="0.56">adult</keyword>
             <keyword weight="0.56">death</keyword>
             <keyword weight="0.56">compression</keyword>
      </keywords>
      </question>
  </trigger_list>
</question_list>
```

Figure B16: A Clearer Illustration of *Q_AList.xml*

```xml
<scenario>
        <scenario_construct list_id="3" scenario_id="s.19990713.1520" userID="1">
                <scenario_desc>First-aid CPR on adult male, 57 years of age. Bystander present.
                 Location:Roadside.</scenario_desc>
                <contextual_link context_id="66" />
                <contextual_link context_id="67" />
                <contextual_link context_id="68" />
                <start_timestamp>1520</start_timestamp>
                <end_timestamp>1538</end_timestamp>
                <trigger event_id="4" />
                <episodes episode_id="8" />
                <episodes episode_id="9" />
                <episodes episode_id="10" />
                <episodes episode_id="11" />
                <episodes episode_id="12" />
                <episodes episode_id="13" />
                <episodes episode_id="22" />
                <episodes episode_id="23" />
                <episodes episode_id="24" />
                <episodes episode_id="25" />
                <episodes episode_id="26" />
                <episodes episode_id="27" />
                <episodes episode_id="41" />
                <episodes episode_id="42" />
                <episodes episode_id="47" />
                <episodes episode_id="48" />
                <episodes episode_id="49" />
                <episodes episode_id="58" />
                <episodes episode_id="200" />
                <episodes episode_id="201" />
                <episodes episode_id="202" />
                <concluding event_id="5" />
                <keyWord>breathing</keyWord>
                <keyWord>hand</keyWord>
                <keyWord>compression</keyWord>
                <voter id="1">0.67</voter>><voter id="2">0.45</voter><avgSc>0.56</avgSc>
                <ScCrysFactor>0.6138095238095238</ScCrysFactor
        </scenario_construct>
</scenario>
```


Figure B17: A Clearer Illustration of *scenario.xml* – Scenarios'Details are kept in this file

```
<sub_class_list_element>
        <sub_class id="101">
                <sub_class_name>Severe bleeding</sub_class_name>
                <scenario_intance>102</scenario_intance>
        </sub_class>
        <sub_class id="70">
                <sub_class_name>Choking adult male</sub_class_name>
                <scenario_intance>71</scenario_intance>
        </sub_class>
        <sub_class id="200">
                <sub_class_name>Lower back pain, adult</sub_class_name>
                <scenario_intance>201</scenario_intance>
        </sub_class>
        <sub_class id="257">
                <sub_class_name>Gun-shot wound</sub_class_name>
                <scenario_intance>258</scenario_intance>
                <scenario_intance>323</scenario_intance>
        </sub_class>
        <sub_class id="222">
                <sub_class_name>Knife wound</sub_class_name>
                <scenario_intance>223</scenario_intance>
        </sub_class>
        <sub_class id="368">
                <sub_class_name>Heart rapture for Adult</sub_class_name>
                <scenario_intance>369</scenario_intance>
        </sub_class>
        <sub_class id="2">
                <sub_class_name>Heart attack for Adult</sub_class_name>
                <scenario_intance>3</scenario_intance>
                <scenario_intance>134</scenario_intance>
                <scenario_intance>174</scenario_intance>
                <scenario_intance>278</scenario_intance>
                <scenario_intance>431</scenario_intance>
                <scenario_intance>440</scenario_intance>
                <scenario_intance>445</scenario_intance>
                <scenario_intance>600</scenario_intance>
                <scenario_intance>607</scenario_intance>
        </sub_class>
</sub_class_list_element>
```


Figure B18: A Clearer Illustration of *sub_class_list_element.xml*

```
<episode_construct>
    <episode id="77">
        <name>ep6</name>
        <desc>Verify</desc>
        <event id="79">
            <name>ev18</name>
            <type>Action</type>
            <actor>First-aider</actor>
            <object>Patient</object>
            <parameterValue id="80">
                <name>pv27</name>
                <parameter>Ask</parameter>
                <value>Patient is choking</value>
            </parameterValue>
        </event>
        <voter id="2">0.56</voter>
        <avdEp>0.56</avdEp>
    </episode>
    <episode id="133">
        <name>ep7</name>
        <desc>First-aid</desc>
        <event id="82">
            <name>ev19</name>
            <type>Obstacle</type>
            <actor />
            <object>Patient</object>
            <parameterValue id="83">
                <name>pv28</name>
                <parameter>Choking</parameter>
                <value>Yes</value>
            </parameterValue>
        </event>
        <voter id="2">0.34</voter>
        <avdEp>0.34</avdEp>
    </episode>
</episode_construct>
```

Figure B19: A Clearer Illustration of *event.xml*

```
<garbage>
    <garbageElem domain="c11" Qid="1" dateInsert="09/04/2004">c</garbageElem>
    <garbageElem domain="c11" Qid="1" dateInsert="09/04/2004">f</garbageElem>
    <garbageElem domain="c11" Qid="3" dateInsert="09/04/2004">b</garbageElem>
    <garbageElem domain="c12" Qid="1" dateInsert="09/04/2004">b</garbageElem>
</garbage>
```

Figure B20: A Clearer Illustration of *garbage.xml*

```
<context_construct>
	<context id="601">
		<text>diabetes</text>
	</context>
	<context id="602">
		<text>severe bleeding</text>
	</context>
	<context id="605">
		<text>heart attack</text>
	</context>
	<context id="606">
		<text>chest pain</text>
	</context>
	<context id="608">
		<text>First-aid</text>
	</context>
	<context id="609">
		<text>CPR</text>
	</context>
</context_construct>
```


Figure B21: A Clearer Illustration of *context.xml*

Appendix B – The FAQFinder System User Interfaces

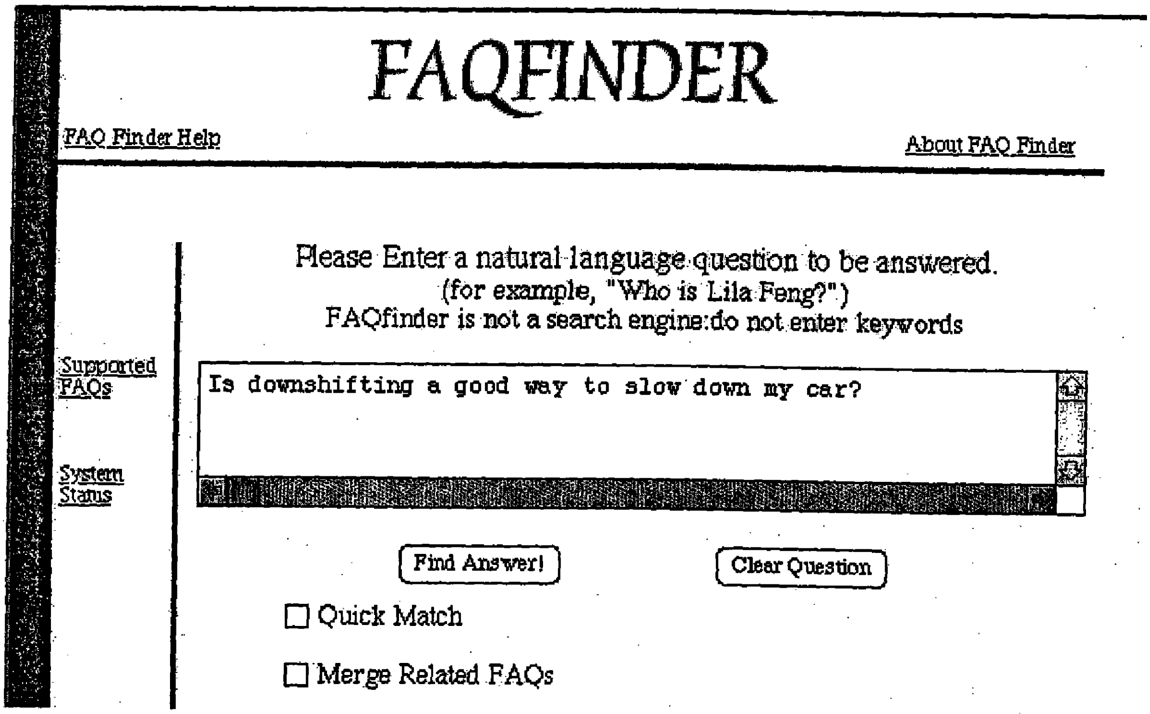


Figure B22: Asking Question of FAQFinder [13]

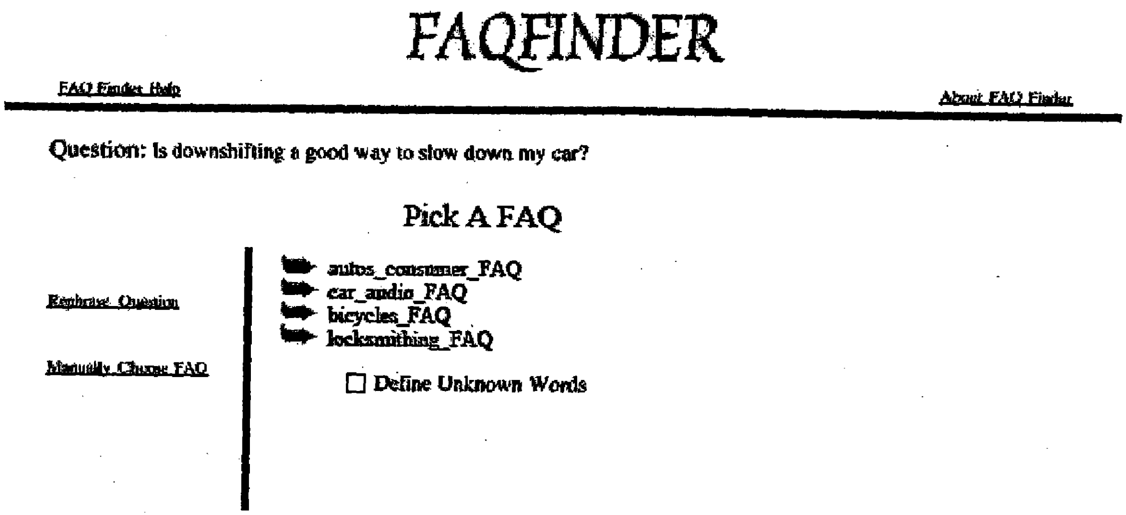


Figure B23: Choosing a FAQ Files to match against [13]

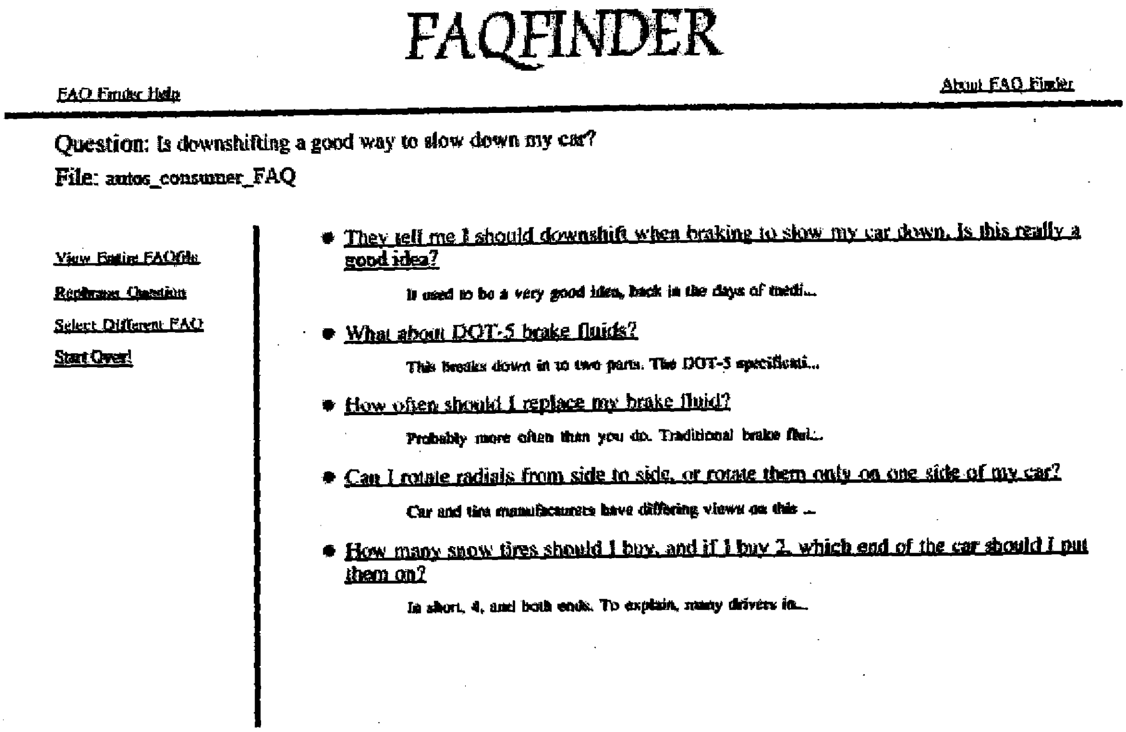


Figure B24: Choosing an answer [13]