Knowledge Management
in Construction Project Management

by

Wong Chi Fai Eric

A Thesis Submitted to The Hong Kong Polytechnic University

for the Degree of

Master of Philosophy

under the Supervision of Prof. Heng Li

Department of Building and Real Estate
The Hong Kong Polytechnic University
2001

ABSTRACT

Abstract of thesis entitled: 'Knowledge Management in Construction Project Management' submitted by Wong Chi fai, Eric for Degree of Master of Philosophy at The Hong Kong Polytechnic University in December 2001.

The construction industry generates and retains a large amount of information during construction project life cycle. It is a hidden treasure that is buried inside and can become valuable knowledge for the future. Currently, this kind of knowledge accumulated from the past is not well organized and documented.

This study presents a Critical Issue knowledge management system to capture the project events by the impact on time, cost, progress and quality. Issues range from Site Instructions, Works Suspensions to works co-ordination clashes. By using the case-based reasoning approach, the cases can be weighted and calculated for similarity with existing records. Meanwhile, the case adaptation resembles the buried knowledge in case base. The Critical Issue therefore, facilitates the project information collection from project management system as a case that can become a precedent for future reference and re-use. This process resembles the essence of human reasoning and simulates the philosophy of knowledge management.

The result proves that transforming project data to retrievable knowledge is vital to improve project management and decision-making. Moreover, it is feasible to run together with every project management system already developed.

ACKNOWLEDGEMENT

I would like to express my deep appreciation and gratitude to Professor Heng Li, my project supervisor, for his assistance, support and guidance throughout my dissertation research. Also, Mr. Conrad Wong, the CEO of VHSoft Technologies Co. Ltd, subsidiary of Yau Lee Holdings Ltd. With his richest experience in construction industry, his valuable suggestion and idea, it benefits throughout the research studies, design and system implementation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Background of Study

The erosion of traditional business demarcation is probably most acute in the construction sector and to introduce strategies aimed at making them more cost-effective. The need to improve cost efficiency ratios will increasingly necessitate that management becomes more profitable and risk conscious. The important issue facing the construction industry in the 21's century will continue to be declining profitability, as measured by return on assets. Building and construction, however, has already emerged as an important sector in improving the profitability. Continued improvement will be dependent on combinations of strategies which concentrate primarily on both "products" and "costs" in an endeavor to improve market position and the overall cost-effectiveness of the business. Competition and profitability factors will mean that the construction industry, in terms of its management style, use of technology, internal and external infrastructure and internal and external use of information technology, must be increasingly responsive to market considerations.

"Knowledge is a fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information." By Thomas H. Davenport, Author of Working Knowledge (Thomas, 1999).

Former Oracle president Ray Lane described the new successful enterprise as a group that could "automate the process of gathering around a table, with the customer sitting in one seat, the product design people in another, manufacturing and HR, partners and suppliers in theirs. And they have a conversation that changes the current way business is done, from static, top-down planning to true collaboration. And the company that does that is going to win." (Pflaging J., 2001)

Chairman of Microsoft, Bill Gates elaborate that "Business is going to change more in the next ten years than it has in the last fifty." (Bill Gates, Oct 2001)

Nokia CEO, Jorma Ollila said "We are moving towards a situation where we are continually refreshing the whole product range. Underlying this new strategy is knowledge management." (Jorma Ollila, 2000)

Butler Group associate Paul Strassmann, his firm has developed a knowledge capital metric to objectively measure the capacity of a firm to generate profits by other means than through the utilization of financial capital. He says that knowledge management must become a priority on companies' executive agenda. He said: "If society is to successfully evolve from an industrial model to an information-based economy, knowledge asset and the methods of protecting them will become increasingly important. (KMM, Oct 2001)

Knowledge is at the center of the wealth-producing process, it explains corporate growth, productivity improvement and innovation. In fact, it is a most valuable asset of every corporation. It's different from physical assets, since knowledge never depreciates or gets used up. With the more knowledge is used, the more value it has.

The philosophy of knowledge management is to enable people to do the right thing when the right information is given to people at the right time. IT can extend knowledge networks and make it easier to find people who have the knowledge most suitable. Our goal is to develop new approaches for managing knowledge that is useful in construction practices.

The informational assets of a of company possesses and the staff continual interaction with situation which increase the value of small pieces of information by adding experience, knowledge and association in order to produce other, even more valuable pieces of information (KMM, Oct 2001).

1.2 Research Objectives

We have witnessed the arrival of information age in the final decades of the last century. Information technology has encouraged the empirical breakthrough and improved work efficiency. In the construction industry, handling a project for example, involves thousands of document exchange, design agreement, etc. We can see that knowledge is being exchanged all the time. Thus construction process throughout its engineering life cycle requires many experts in different fields to cooperate with each other. A training process for a novice to become on expert requires tremendous cost and time. When an expert leaves or retires, knowledge that is accumulated through years of experience has the risk of losing as well. In many cases, knowledge may be stored in piles of documents that are unorganized and hard to find. Human planners are one type of experts that require such a learning process, which slow and inefficient (Dzeng *et al.*, 1997). Construction experience, such as planning, structure design and project management are the key factors that determine whether the construction can proceed smoothly and be completed in time.

It would seem that the construction industry always has problems in achieving required levels of quality whilst remaining within acceptable budgets and completing by agreed dates (Moore *et al.*, 1998). The major problem concerning construction management is how to apply the resources available to the manager in a most efficient way. Typically, when speaking of resources for construction, we think of the four "M"s: manpower, machines, materials, and money. Management involves the timely and efficient application of the four "M"s during a project. Many issues must be considered when managing a project and successfully applying the four "M"s.

Because the construction industry typically consists of fragmented entities with differentiated technology development levels, the control function is done through the communication of standards, becomes increasingly complex. Such complexity may potentially cause project failure. When something fails we experience a temporary loss of certainty. Why did it fail? Will it happen again? Lots of questions come out. Until determining why it fails, we cannot be certain if it will fail again. Therefore the sooner the reasons of failure can be determined, the sooner certainty will return (Halpin *et al.*, 1998).

While manipulating a construction project, project manager often spent a lot of time in processing information such as data collection, enquiry report, labor knowledge exchange, etc. It is crucial to have an effective management system for these processes. At current stage, the building engineer, quantity surveyor and project manager are often frustrated by classifying building information, presentation of data format and utilizing human knowledge. It is therefore clear that systemically implementing those processes provides a strong financial incentive to improve design and building processes (Dzeng *et al.*, 1997).

1.2.1 Cumulating Experience and Managing the Knowledge

Knowledge is the distilled form of information and knowledge management is the systematic process of finding, selecting, organizing, distilling and presenting information in a way that improves an employee's comprehension in a specific area of interest. Knowledge management helps an organization to gain insight and understanding from its own experience. Specific knowledge management activities help focus the organization on acquiring, storing and utilizing knowledge for such things as problem solving, dynamic learning, strategic planning and decision making. It also protects intellectual assets from decay, adds to firm intelligence and provides increased flexibility (Kolodner, 1993).

If KM is "the process through which organizations gain value through their knowledge-based assets" then the key to teaching KM is to understand what constitutes a knowledge-based asset, and to maximize return from those assets. Any process can be taught. In common with teaching other organizational processes, the key lies in defining the process, its perceived relevance within the organizational culture, demonstrating tangible benefits, and providing useful toolsets. (KMM, 2001)

Useful knowledge is the information that can be applied to a context or problem. A train time table may be considered information, identifying that the last train leaves at 10pm is knowledge. To extract the most relevant items from the information base though, often requires the identification of a person who can provide real insight or intelligence relevant to the problem: the station master who can tell you that the train departs from platform 2 at 10pm, and that there is a spare seat in coach C. KM is more than just information and processes: critically it's about people. Understanding KM is predominantly about understanding the motivations of people and influencing the processes they adopt.

We can identify two types of knowledge assets: tacit (for example, know-how - the expertise in people's heads) and explicit (for example, research, plans, patents, best practice information etc). Effective KM involves making it as quick and easy as possible to locate and use knowledge assets, whether they are tacit, explicit or a combination. The tools that will facilitate the implementation and teaching of KM will recognize both these types of knowledge asset.

KM is a discipline that applies to every level of an organization. For KM to be effective, it requires all staff to implement it, in fact staff shouldn't just implement it, they should with live it! KM isn't a process that can be adopted for a few weeks/months/years - that would ignore the fact that knowledge can be archived, only to become relevant once again several years down the line. KM processes require ongoing training and support, with the KM approach being a key part of the induction of new employees and of team members from outside the organization.

KM processes can be implemented by reference to sets of guidelines, and by the adoption of a KM culture. When we are taught to drive a car, the Highway Code is used as our 'process manual'. However, road safety is only maximized if every driver adheres to the rules. The extent of process adoption is dependent upon the culture for accepting such rules, and in turn on whether the rules are perceived as valid, reasonable and relevant. For example, in France the culture has led to a set of traffic laws that are quite different from those in the UK. Additionally, the extent of adoption is quite different, and this is a direct result of our cultural differences.

Cultural change is possible. For example, not that many years ago, drinking and driving was an accepted part of a night out, and the wearing of seatbelts was not ubiquitous. Today, anyone is guilty of drink driving, or of not belting up their children in a car would be labeled a social pariah. Equally the cultural changes are necessary for the successful implementation of KM can be accomplished, but the process of culture changes takes time.

Two approaches can be implemented to achieve KM processes. Given that cultures evolve, an evolutionary approach can be used, with organic growth in the use of KM processes - perhaps facilitated by the implementation of appropriate technology. An alternative would be a much more structured approach. However, cultural change cannot be enforced. The solution is to put frameworks in place to facilitate KM, with a clear identification of tangible benefits of KM. Evolutionary cultural change will then follow with bottom-up adoption enabled by the technologies and processes that are already in place. These processes should include individual reward processes to recognize performance, approaches to personal recognition, and processes to make KM adoption easy.

Although technology is said to be as little as 10 per cent of a KM implementation, the intelligent use of KM tools can support the key elements of the processes required to effectively introduce and develop a successful KM strategy.

Technology has a clear role to play not only in the KM solution, but also in the influencing of organizational culture. The effective KM model for a particular organization will be one that is culturally acceptable to the workforce. By integrating KM with other aspects of the team-based working model (with team formation, intra- and extra-organizational collaboration support, document management and repository searching), such solutions will more rapidly become the accepted way of working.

1.2.2 IT Revolution, The Knowledge Management System (KMS)

The KMS is the on-going, persistent interaction among agents within a system that produces, maintains, and enhances the system's knowledge base. This definition is meant to apply to any intelligent, adaptive system composed of interacting agents. An agent is a purposive, self-directed object.

In saying that a system produces knowledge we are saying that the system (a) gathers information and (b) compares conceptual formulations describing and evaluating its experience, with its goals, objectives, expectations or past formulations of descriptions, or evaluations. Further, this comparison is conducted with reference to validation criteria. Through use of such criteria, intelligent systems distinguish competing descriptions and evaluations in terms of closeness to the truth, closeness to the legitimate, and closeness to the beautiful. (Nicholas, 1997)

1.3 Organization of Thesis

Firstly, how to manage the knowledge that's worth knowing? This means having the ability to have a knowledge plan that can ensure the knowledge is needed for everyday processes and innovation processes is consistently captured and built into a generic process structure that can be easily accessed by teams. It also means having the ability to identify the knowledge that is worth commoditising and outsourcing.

Secondly, how to learn faster than the competition and embed the knowledge into appropriate media? This requires the ability to view the industries from outside the industrial box and by wearing the predator's shoes, to continually question the life cycle of industrial knowledge, running premature visualization exercises and learning from them.

Thirdly, how to create new forms of knowledge that deliver new market value? This means anticipating the boredom of the market and constantly visualizing something new and different that means you are in control of the market's attention-span and perhaps like MTV, determining the attention-deficiency of a generation through the continual reduction of the duration of camera-shots. This is about having a sequence of timely products that succeed and reinforce each other.

The three Knowledge Imperatives mentioned above require a keen focus on clarity within a context that is in continual transition. The two constants are the decaying value of applied knowledge and the realization that the acquisition of enabling technologies without visualizing their part in delivering new value, means that whilst the organization's physiology may grow, its psychology and perception remains fixed.

If KM is to deliver, it needs to be the complementary mirror of the business strategy, like integrating the knowledge strategy into business strategy, so that the connection is seamless. IT needs to be built into the perception and muscles of the organization and reviewing competitive knowledge in terms of what it was, is and will be, and also building explicit actions around knowledge and capability gaps.

1.3.1 How Knowledge Management Benefit the Construction Industry

Construction is an information rich industry. However, since information is highly fragmented in nature, the ineffective and slow transmission of information between workgroup leads to slow response time and delay overall scheduling. Inside the process cycle, which typically consists of planning, procurement, design, construction and maintenance, large percentage of time was spent on retrieving information. Involving parties such as architect, structure engineer and project manager are unable to effectively monitor and track their works. The operational tasks, such as paper copying, document duplication and report distribution are done repetitively. Slow delivery of or non-synchronization and miscommunication of information has caused excessive rework and labor wastage. By transferring the information to knowledge, work processes can be improved and streamlined in an industry that is inherently inefficient and labor intensive. Structured knowledge management can effectively benefit communications, workflow, information delivery and retrieval, inter-company and inter-departmental collaboration.

A consideration for the construction industry is that knowledge management problem exists within construction firms, irrespective of what sorts of the projects being carried out. This kind of the knowledge problem typically presents itself within an individual firm as a tendency to "re-invent the wheel" causing by low levels of information sharing between workers. This is evidenced by a firm that is not achieving group-thinking for its operations and is functioning solely on the basis of the less desirable team coordination. While, in the maintenance of group-thinking is suggested as being the traditional split between the development of knowledge and the manner in which that knowledge is subsequently stored with in a firm. The standard mode of knowledge storage is to store it within the workers who develop it. In other words, workers store it in their brains. Unfortunately, knowledge that is stored in this manner produces two significant problems: it is not always easily accessible when required, particularly by others; and it is highly portable when the worker and their brains move to work for the other company. It is therefore obvious that storing knowledge in a proper manner is one of the main concerns of knowledge management system.

Information is accumulated by continuously resolving errors and problems from past experience. Our approach is to collect and classify past experience and represented as easy to retrieve, semantics and syntactic searchable knowledge base, plus using sophisticated algorithm to analyze the information pattern and develop useful knowledge. The design of a knowledge-based systems exclusively dedicated to construction process continues to present a major challenge. The role of such system is to automate the performance of complicated workflow and tasks during retrieving and referencing to past knowledge. Information retrieval and knowledge based system is a crucial terms of academic field. The research and the goal of this project is to convert and carry the pass experience of skillful persons to a computer based system. Therefore, we can retrieve their knowledge anytime when similar problems occur (Jiang *et al.*, 2000).

By developing a software solution, it is possible to increase the efficiency of processing construction information systematically through providing a knowledge base that architects and engineers can acquire knowledge from past experience. With the use of integrated and centralized construction information querying mechanism, the ambiguity between different constructors would be eliminated.

1.4 Outline of Thesis

Following this introductory chapter, Chapter Two presents a literature review on knowledge management and the methodologies of applying knowledge management. It surveys the available techniques to support knowledge management. These include artificial neural network (ANN) and case based reasoning (CBR). It introduces the technical terms, types of applications, and the implementation methodology. These techniques are compared. Based on the comparison, case based reasoning integrated with the artificial neural network is considered the most appropriate for the problem domain and task type. Chapter three presents the development of a prototype system for knowledge management embedded in a construction project management system. Finally, Chapter Four concludes and summarized the study by discussing the limitation of the present study and suggests further research that should be further carried out.

CHAPTER 2

LITERATURE REVIEW

2.1 Knowledge Management: Case-based reasoning (CBR)

Case-based reasoning (CBR) is a problem solving strategy founded upon the re-use of past solutions to address new problems. Instead of representing knowledge as rules in rule-based expert systems, a case-based reasoning system maintains case precedents that retain previously solved issues. Upon encountering a new problem, the system retrieves similar cases from its database, selects the most suitable one, and modifies the old solution for solving the new problem. This technique is appealing as human often make decision in a similar fashion. Several attempts have been made to address CBR as a scientific cognitive model (Althoff, 1995).

There are two important operators in case-based reasoning system: select and adapt. Select is the means of retrieving the most suitable case from the database according to the features of the new problem. A weighted count of matched features provides one method to select the best case; however, this does not take into account the fact that the case itself.

2.1.1 CBR Leverages Prior Experiences to Solve Current Problems

Diagnostic decision making and determining how to fix things might be supported by a CBR system that:

- Helping to identify the problem at hand by comparing it to previously encountered problems.
- Once identified, adapts the solutions of previous occurrences of the problem to the current situation.

A system that supports underwriting decisions might use CBR to compare a current application to previously processed or example applications, and then recommends whether to approve, deny, request further information, or include contingencies (Kolodner, 1993).

In the course of its deliberation, a CBR system would retrieve previously processed applications that were similar to the application under consideration (Cercone, 1999). Given similar applications and the justifications for those decisions that management agreed was correct for those applications, a CBR system would:

- Estimate the relevance of each retrieved case to the case at hand,
- Exclude retrieved cases that business policies indicate are logically distinct,
- Focus on the most relevant of the remaining cases,
- Determine whether any crucial differences that interfere with conclusive approval or denial exist,
- And make a final conclusion, perhaps including remedial action, such as increasing collateral.

2.1.2 The Functionality of Case-Based Reasoning

Case-based reasoning can mean adapting old solutions to meet new demands, using old cases to explain the new situations, using old cases to critique the new solutions, or reasoning from precedents to interpret a new situation or create an equitable solution to a new problem. A case-based reasoner learns as a byproduct of its reasoning activity. It becomes more efficient and more competent as a result of storing its experiences and referring to them in later reasoning. Because case-based reasoning integrates reasoning and learning, it is not enough for a case-based reasoner to stop reasoning after it derives a solution. Rather, it must continue by collecting feedback about its solution and evaluating that feedback (Kolodner, 1993).

There are two major types of case-based reasoning: interpretation or classification case-based reasoning and problem solving case-based reasoning. Through case-based reasoning, a solution to the new problem is proposed by extracting the solution from a retrieved case. This is followed by adaptation, the process of fixing an old solution to fit a new situation; and criticism, the process of critiquing the new solution before trying it out. Through interpretation of the case-based reasoning, desired result is proposed, sometimes based on retrieved cases, sometimes imposed from the outside. This is followed by justification, the process of creating an argument for the proposed solution, done by a process of comparing and contrasting the new situation with prior cases, looking for similarities between the new situation and others that justify the desired result and differences that implying other factors must be taken into account (Kolodner, 1993).

Case-based reasoning as a learning paradigm has several advantages; firstly, CBR doesn't require causal model or a deep understanding. In addition, many domains have existing case bases that can seed the use of case base system.

There are also disadvantages in using cases to reason:

1. A case-based reasoner might be tempted to use old cases blindly, relying previous experience without validating it in the new situation.
2. A case-based reasoner might be biased by previous cases in solving a new problem.
3. Often people, especially novices, are not able to use the most appropriate set of cases when they are reasoning.

Relying on pass experience without taking validation can result in inefficient or incorrect solution and evaluations. Retrieved inappropriate cases can cost precisions problem-solving time and lead to costly errors that can be avoided by more incremental method.

2.1.3 Processes of Case-Based Reasoning

1. Recall cases form memory

   Good cases are those that have the potential to make relevant predictions about the new case. Retrieval is a process that calculating the features of a new case stored in case library. Cases labeled by subsets of those features or by features that can be derived from those features are recalled.

2. Form collection cases retrieved in step 1

   The purpose of this step is to winnow down the set of relevant cases to a few most-on-point candidates that worthy to consider intensively. Sometimes it is appropriate to choose one best case; sometimes a small set is needed.

3. Construct solution

   In problem solving, it normally involves selecting the solution to the old problem, or some piece of it. In interpretation, this step involves partitioning the retrieved cases according to the interpretations or solutions they predict and, based on that, assigning an initial interpretation to the new problem. There are several issues that arise in constructing the solution. First is the question of how appropriate portion of an old case can be selected for focus. Another issue has to do with how much work to do in this step before passing control to adaptation or justification processes. Third issue has to do with the choice of an interpretation in interpretive reasoning.

4. Test and criticize

   Criticism may require retrieval of additional cases and may result in the need for additional adaptation, which is called repair. Major issues here include strategies for evaluation using cases; strategies for retrieving cases to use in interpretation, evaluation, and justification; the generation of appropriate hypothetical for using them; and the assignment of blame or credit to old cases.

5. Result evaluation

   The result of reasoning is tried out in the real world. Feedback about the real things that happened during or as a result of executing the solution are obtained and analyzed. If results are as expected, further analysis is not necessary in this step, but if they are different then expected, explanation of the anomalous results is necessary. Explanation requires figuring out what caused the anomaly and what cold have been done to prevent it. It can be done by the CBR and reapplying previous explanation.

6. Update memory

   The new case is stored appropriately in the case memory for future use. A case comprises the problem, its solution, plus any underlying facts and supporting reasoning that the system knows how to make use of. The most important process at this time of memory update is choosing the ways to "index" the new case in memory. Choose appropriate indexes for the new case using the right vocabulary, and at the same time make sure that all other items remain accessible after a new case is added to the case library's store (Kolodner, 1993).

In summary, CBR incorporates a six-step process:

(1) Accepting a new experience and analyzing it to retrieve relevant cases from case memory;

(2) Selecting a set of best cases from which to craft a solution or interpretation from the problem cases;

(3) Deviating or interpreting the solution by supporting arguments in the case of precedent based CBR and with implemented details in the case of problems solving CBR;

(4) Testing of the solution or interpretation with an eye to assessing its strengths, weakness, generality, etc.

(5) Executing it in the world and analyzing the feedback; and

(6) Storing the newly solved or interpreted case into case memory and appropriately adjusting indices and other mechanisms.

2.2 Indexing

The biggest issue in case-based reasoning is retrieval of appropriate cases. To let the computer doing the same things as human remember the right ones at the right time, we have to indexing the problem in computer understanding manner. The indexing problem has several parts. First is the problem of assigning labels to cases at the time that they are entered into the case library to ensure they can be retrieved at appropriate times. Second is the problem of organizing cases so that search through the case library can be done efficiently and accurately. Table 2.1 sums up the steps of choosing Indies (Kolodner, 1993).

1. Determine what the case could be useful for, by designating its points with respect to the set of tasks the reasoner is being asked to carry out.
2. Determine under what circumstances its points would be useful for each of these tasks.
3. Translate the circumstances into the vocabulary of the reasoner.
4. Massage the circumstances to make them as recognizable and generally applicable as possible.

Table 2.1 Step in assigning Indies to cases

2.3 Retrieval Algorithms

Retrieval algorithms shown as Figure 2.1 can be parallel or serial. They can have hierarchical or flat organizational structures. Depending on the indexing method they use, they can interact with dimensional matching function, aggregate matching function, or both. Serial search on a flat memory has the advantage of being easy to implement and doing a full search of the case library, but it gets slow as the case library grows. When the library gets large, some means of partitioning must be used to make search algorithms efficient. Shared-feature networks partition the case library according to the sizes of the sets of features shared by the cases. Searching such a network is more efficient than serial search, but well-matching cases can be missed if the network isn't also prioritized. In addition, as a case library gets large, it is hard to keep a shared-feature network optimal, and the complex matching that must be done in searching such a network. Consequently, it causes an inefficient searching process.

Prioritized discrimination networks are more efficient. The case library is subdivided one dimension at a time, dividing first on the most important dimensions.

Redundant discrimination networks deal with the missing features problem. In a redundant discrimination network, multiple discriminations are done at each level of the network.

Retrieval algorithms rank the usefulness of retrieved cases, insertion algorithms invoke insertion algorithms to insert the case in the case library and reorganize its organizational structure as necessary (Lambrix, 1998). The update and retrieval processes can be summarized as following Unified Modulation Language:

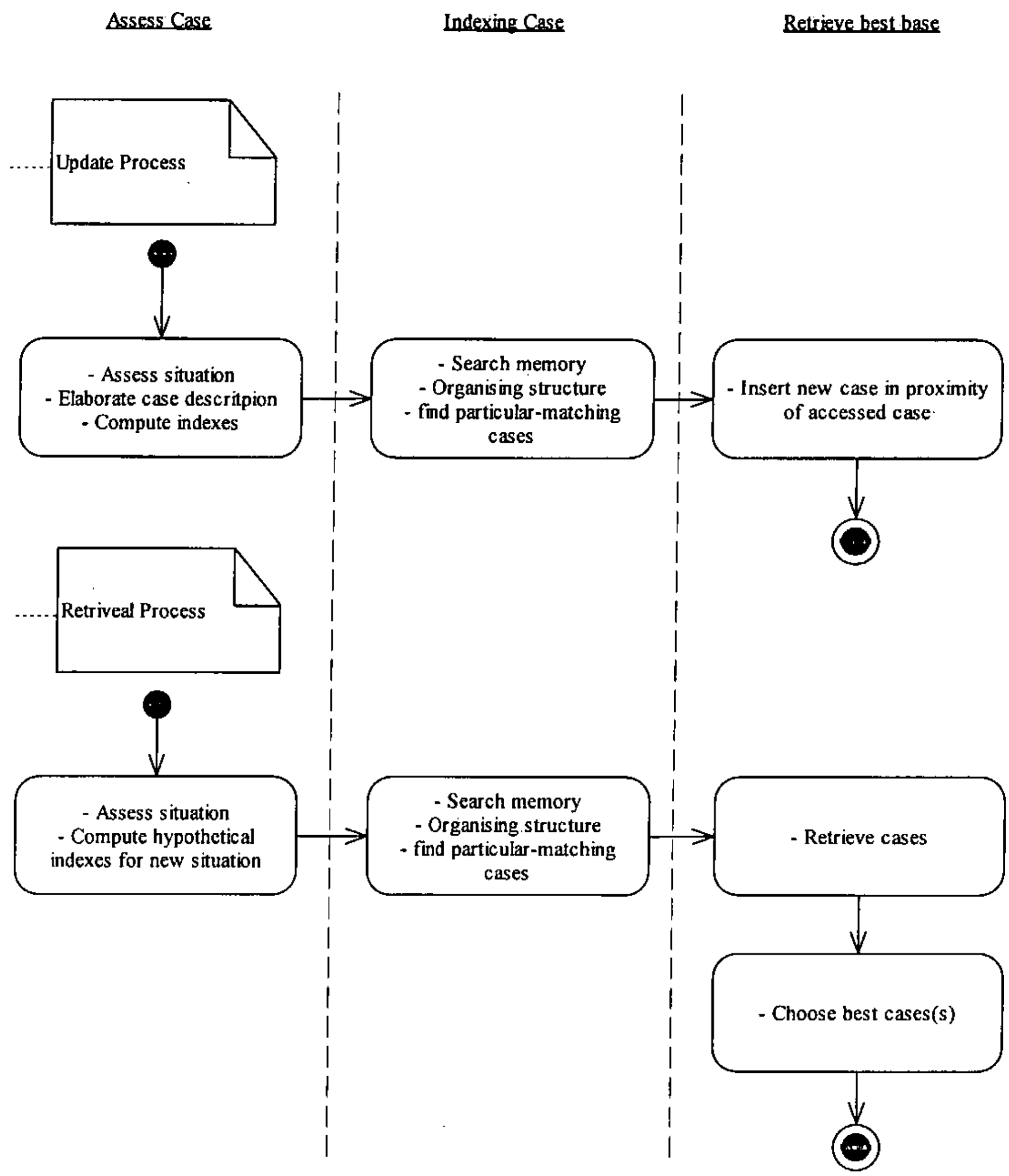

Figure 2.1 CBR transaction diagram

2.3.1 Matching and Ranking

A process to select the best, most useful and good enough cases among the case set. Building blocks include the finding of correspondences between cases, computing degree of similarity of corresponding features, and assigning importance values to dimensions of cases. Methods of matching and ranking are only as good as the available building blocks. If a system tries to match values that are not in correspondence, it won't compute good match scores.

Methods of matching and ranking are shown as follows:

1. Nearest-neighbor matching, it combines importance of each dimension with its degree of match and sums the scores to create a match score for each case. One chooses the case with the best score. Figure 2.2 shows the scoring function of the matching. Such a scheme can be made more or less sensitive by assigning importance values in more or less context-sensitive ways.

$$\frac{\sum_{i=1}^{n} w_i \times sim(f_i^I, f_i^R)}{\sum_{i=1}^{n} w_i}$$

where w_i is the importance of dimension (slot) i,

sim is the similarity function for primitives,

and f_i^I, f_i^R are the values for feature f_i in the input and retrieved cases, repectively

Figure 2.2 A numerical evaluation function

2. Heuristic procedures

Evidence rules identify a good-enough match. Preference heuristic shift the set of available cases according to specified preferences.

3. Combined heuristic and numeric procedure

Exclusion and ranking by either numeric or heuristic procedures. Heuristic choice includes choosing a case that shows sufficient evidence that can be useful and using domain preference to compare through the set of potentially matching cases according to a set of preference condition. Table 2.2 demonstrates the relationship and degree of match between cases.

			Degree of Match	
	Importance of dimension	*Case 1*	*Case 2*	*Case 3*
feature 1	*0*			
feature 2	*0.8*	*0.8*	*0.6*	*0.8*
feature 3	*0.4*	*0.8*	*0.8*	*0.8*
feature 4	*0.4*	*0.8*	*0.2*	*0.2*
feature 5	*0*	*1*	*1*	*1*
feature 6	*0*	*0.6*	*0.8*	*0.7*
feature 7	*0*			
feature 8	*1*	*0.7*	*1*	*0.5*
feature 9	*1*	*1*	*0.9*	*0.8*
feature 10	*1*	*0.4*	*0.8*	*0.8*

feature 11	*0.8*	*0.2*	*0.2*	*0.2*
feature 12	*0.8*	*0*	*0.2*	*0.4*
feature 13	*1*	*0.2*	*0.1*	*0.2*
feature 14	*0.8*	*0.4*	*0.2*	*0.3*
feature 15	*0.4*	*0*	*0.3*	*0*
feature 16	*0.4*	*0.5*	*0.5*	*0.3*
Match score	*8.8*	*4.26*	*4.48*	*4.18*
Normalized match score		*0.484*	*0.509*	*0.475*

Table 2.2 Importance of dimension and degree of case matching

2.4 Case-based Reasoning in Construction Application

CBR has been widely used in different domain. In construction industry, for examples, an expert prequalification system called *EQUAL* which help decision-maker to produce more reliable and expeditious decision for contractor prequalification. Many researchers have observed that contractor prequalification relies heavily on expert judgment. Experiential judgment is required in decision situations when it would be too costly to measure a large amount of information quantitatively and it is also invaluable in providing shortcuts for complex tasks (Thomas Ng, 2000). In order to reduce the subjectivity of current prequalification, research scientists has been examined a range of approaches, including statistical, artificial intelligence method and operations research. These methods include cluster analysis (Holt, 1996), multiple attribute analysis (Russell *et al.*, 1990), rule-based expert system (Russell *et* *al.*, 1990), and fuzzy set models (Nguyen, 1985).

EQUAL is a CBR model developed to capture and reuse experiential knowledge pertinent to decision making of contractor prequalification. It consists of five interrelated case bases; algorithm, screen, score, finance and project. By adopting and adapting past experience, *EQUAL* can assist inexperienced clients throughout the process of contractor prequalification. Throughout the process, two scenarios have been developed: demonstrating the formulation of decision criteria for a construction client and the other showing the determination of a contractor's eligibility to enter into an approved list of contractor (Thomas Ng, 2000).

Another example used in construction area, a hybrid representation of architectural precedents: Support Early Building Design (SEED) System, it aims to provide computational support for the early phases in building design (Ulrich *et al.*, 2001). It combines elements from both Case and Model Based approaches. During the early stages of building design, the system provides support for analysis, evaluation and rapid generation of design representations. SEED consists of modules, each of which addresses a specific task within preliminary design. The tasks supported are architectural programming, schematic layout design and the generation of a fully three-dimensional configuration of physical building components form schematic layouts.

The SEED database stores precedent under a practical/generative perspective. The retrieved precedent is classified by two roles:

1. As a prototype with standard properties that apply across projects, even building types. An example is a residential kitchen and the space requirements typically associated with it. The architectural programming module is expected to make heavy use of such prototypes;
2. As a case (applied in CBR) representing a specific solution for a given set of requirements or a solution that comes closest to satisfying the requirements.

The episodic memory is represented by object database of precedents, which instantiates a class in underlying object-oriented model, can be searched by a set of attribute values combination. Since SEED is a hybrid system, the case base is only one of the three 'bases' it contains. The most robust is the object-oriented database, it provided and shared object model to all modules. The object membership as defined by the inheritance relations in the object class model. For example, the design units allocating the constituents of a floor and the sub-layout, since design units, layouts and floors are classes in the object model and can be queried the database by class membership and relational attributes.

SEED decomposed the design problems into sub-problems, the decomposition carried recursively until user satisfactory. Because real-world design problems are barely decomposable, the original problem should be improved and enhanced during the decomposition process. This was one of the main reasons to apply CBR to design, as cases can provide the glue to hold the different components of a design together (Kolodner, 1993).

2.4.1 Using Case-based Reasoning in Managing Construction Knowledge Management

In a construction project, there are many reports, memorandums, photographs and design drawings that describe project events. Human brain can easily analyses such data, but it is not easy for a computer system to perform similar analytic procedures. Also, the intelligence of human can grow and accumulate. But it is very difficult for a computer to grow. In the real world, a project event or issue can provide experience that cope with different operation conditions. Under different conditions, the expert system will select different solution to suit the operation and other conditions. Case-base reasoning can store pass knowledge and learn from current cases. Its knowledge case can grow and accumulate. In summary, the advantages of case-based reasoning CBR are:

CBR allows the reasoner to propose solutions to problems quickly avoiding the time necessary to derive the answer from scratch.

CBR allows a reasoner to purpose solutions in domains that are not completely understood by the reasoner.

Remembering previous experiences is particularly useful in warming the potential for problems that have encountered in the past. Alerting a reasoner to take action to avoid repeating past mistakes.

Cases help a reasoner to follow it reasoning on important part of a problem by pointing out what features of problem are important.

2.5 CBR Software Tools

There are many CBR products. Some of them are listed below:

ART * Enterprise: It is designed for Management Information System (MIS) developers and offers a variety of representational paradigms, including procedural programming language, objects supporting multiple inheritance, encapsulation, and polymorphism. Also it provides rule sets and cases representation. As an impressive ability, it can link to data repositories in most proprietary DBMS formats for developing client-server applications. (Lan, 1997)

CASE-1: It is obviously developed with interference the CBR in mind. Cases are represented as free-form text describing a problem, a set of weighted questions that can confirm or reject a case, and a set of solutions. Cases are stored in a relational database, and the interface is developed using Visual Basic.

CaseAdvisor: It is a CBR product with inference capacity. There are six components to the software: CaseAdvisor Authoring, case-authoring environment, Case AdvisorProblem Resolution, runtime case-retrieval engine, CaseAdvisor WebServer, web utility to use case-bases. It enables the user creation of diagnostic decision trees. These can be used in conjunction with the CBR system at any point during a diagnosis to lead the user through a series of diagnostic steps.

CasePower: It previously called Induce-it, it builds the cases within the spreadsheet environment of the Microsoft Excel. Symbolic data can be represented as ordered hierarchies that are mapped to numerical ranges. However, for more complex non-numerical applications, other CBR tools would be preferable. CasePower uses nearest-neighbor retrieval and it reduces search time by calculating an index for each case off-line. This can be a lengthy process for a large case-base, but it does reduce retrieval times considerably.

CBR3: It includes CBR Express, CasePoint, Generator, Tester, and CasePoint WebServer. It is specifically tailored to the vertical market of customer support help-desks, as the customer help-desks often deal with technical faults. To implement a rule-based expert system, the knowledge of faults and their solutions are required. But if the manufacturer knows there was a fault, it probably wouldn't release the product. Thus, help-desks often deal with faults that designers and engineers have not envisaged, and consequently the solutions could not be in rule-base. The CBR3 offers a solution. Through the application of the CBR-cycle, which includes *case retention* (learning), the CBR3 can learn from the cases reported by customers. CBR3 applies the cycle very successfully to the help-desk scenario and is currently the market leader for knowledge-based help-desk software.

Eclipse: It is the forward chaining implementation of ART, was re-implemented in C by NASA, entering the public domain as the language CLIPS. It offers object, fully compatible C++ objects, and optimized forward chaining using the Rete algorithm. It uses nearest-neighbor and inductive retrieval of records in the database. Once records have been retrieved, they asserted as Eclipse object for adaptation by its rule-base. Eclipse supports the usual range of variable types and offers similar text-handling facilities to ART, which means it ignoring noise words and using trigrams to cope with the spelling mistakes. In addition, it offers two Class libraries for C++:

1. CPR C++ Class library. It provides case-based problem resolution in an embeddable C++ class library. It has embedded-problem resolution, diagnostic knowledge authoring and publishing, end user problem resolution, and information retrieval.
2. Help!CPR. It is a window application program interface calling the CPR C++ class library and import third-party case-bases from ServiceWare and knowledgeBrokers. It can also import the case-bases authority using CBR Express form Inference Corporation.

ESTEEM: It uses Kappa's inference engine (Kawamura, 1994), thus enabling developers to create adaptation rules. It supports applications that access multiple case-bases and nested cases. It enables one case base to reference another case-base through an attribute slot in a case. It supports various similarity-assessment methods, including feature counting, weighted feature computation, and inferred feature computation. It also can automatically generate feature weights, either using an *ID3* weight-generation method, or a gradient descent weight-generation method. User can incorporate their own similarity functions into ESTEEM if they wish.

KATE: Including KATE-INDUCTION, KATE-CBR KC, KATE-EDITOR, and KATE-RUNTIME.

1. KATE-INDUCTION is an *ID3*-based induction system that supports an object-oriented (OO) representation for cases. Cases can be imported from many database and spreadsheet formats. The induction algorithm is tolerant of missing data and can make use of background knowledge. Retrieval using trees generated by the induction algorithm is extremely fast.
2. KATE-CBR is the nearest-neighbor component of the suite. User can customize similarity assessments and, since it supports the same object hierarchies as KATE-INDUCTION, the two techniques can be combined in a single application.
3. KATE-EDITOR is a set of C DLLs that are integrated with Tool-Book to form a customizable developer's interface. In particular, forms can be developed to assist case entry.
4. KATE-RUNTIME is a set of interface utilities can be customized and deliver as embedded C code.

ReCall: This tool offers a combination of nearest-neighbor and inductive case retrieval. ReCall is coded in C++ and uses an object-oriented representation with taxonomies, inheritance mechanisms, typed descriptor, facets, demons, and relationships between objects. ReCall provides multiple hierarchical indexes, which are used for organization purposes and for efficient case retrieval. Similarity functions take both the properties and values of descriptors into account, as well as structural differences between cases. ReCall uses a variant of a nearest-neighbor algorithm that improves similarity computations.

ReMind: Cases are represented as attributes: value pairs, it offers template, nearest-neighbor matching, inductive, and knowledge guided inductive retrieval. The template retrieval supports simple SQL like queries, returning all cases that fall within set parameters. The nearest neighbor is informed by weighting that can be placed in case features. Inductive retrieval involves building a decision tree that indexes the cases. User can also create a qualitative model to guide the induction algorithm with background knowledge.

When the cases have been retrieved inductively, ReMind is able to explain why the cases were retrieved. The explanation also indicates how the retrieved cases may be adapted. Case adaptation provides and creates adaptation formulas that adjust values based on the difference between the retrieved and the new case. These are also created graphically using a visual programming technique. It provides extremely close typing of case feature combines with the close typing of the operators does reduce syntax errors. ReMind can import case-bases from existing databases, making case entry easier.

The limitation of ReMind is its retrieval speed. Nearest-neighbor retrieval is quite slow even on a relatively small case-base. Inductive retrieval on the other hand is fast, but creating the inductive index is slow.

2.6 CBR and Databases

Database systems are designed to do exact matching between queries and stored information, while the goal of CBR is to retrieve a 'most similar' case or set of similar cases. The most similar cases may include conflicts with some of the attributes that were specified in the retrieval query. In CBR, whether a particular case should be retrieved depends not only on the case itself, but whether there are better competitors. (Leake, 1996)

However, similarly to the relation between CBR and Information Retrieval, the difference is only true as long as 'traditional' relational database management systems are considered. Moreover, this comparison focuses on the retrieval result and neglects other main differences between the two fields. A number of differences, such as transactions, recovery, etc, can be ignored for the moment, but the following two aspects are relevant for the retrieval:

- Firstly, databases operate on data set that the order of magnitude larger than the case bases of CBR. Hence, the retrieval and access method include the management of secondary storage.
- Secondly, the access methods are fully dynamic, i.e., insertion and deletion of data items can be arbitrarily intermixed with queries, without performance loss.

2.6.1 Using Databases as a CBR Backend

The most prominent current use of database techniques within CBR is to use database management systems (DBMS) as a backend for the CBR system. Two different forms can be distinguished (WESS, 1995)

1. DBMS as a persistent case store. Databases are only used as a persistent store for the cases. They are bulk loaded while starting the CBR system. The index structures are computed and storied in main memory, they contain only pointers to the cases stored in the database. Throughout the CBR session, these index structures are used for the retrieval and, the original case is retrieved from the database.
2. Realizing a CBR system with database means. In this scenario, the database system plays an active role. The CBR system is realized with database means. If relational system is used, then similarity retrieval will perform by generating tables encoding the similarity measures as well as the appropriate SQL-queries (Kitano, 1996). The difficulties then are the generation of adequate queries and the integration of this query generation within a typical CBR architecture.

The situation is different when object-oriented database management system (OODBMS) is selected. Since most object-oriented databases are lacking a descriptive query language, they are tightly coupled with a programming language or they provide their own. This conceptual deficit of OODBMS makes it easier to realize

CBR system, since the retrieval process can be directly realized in the programming language of the database.

2.6.2 Multi-dimensional Access Methods

Traditional database management systems (DBMS) normally use some variation of the B^+-tree for single-attribute indexing. However, various applications need to deal with multi-attribute, i.e., multidimensional, data; Geographic information system (GIS) operating on spatial objects (either 2D or 3D) is the oldest and most established application area. Newer application areas include multimedia databases (images, audio, video), time series, string and DNA databases, as well as data mining and online analytical processing (OLAP). These application areas typically operate on higher dimensional data (typically $10 < d < 100$)[4]. The main problem in designing multi-dimensional access methods is that there exists no total ordering among spatial objects that preserves spatial proximity. Hence there is no easy and obvious way to extend the well understood and efficient one-dimensional access methods from traditional databases in order to handle multidimensional data.

Queries that is typically supported by multidimensional access methods:

1. Exact match queries. Given a query object find all the objects in the database with the same geometry.
2. Range queries. There are several variants. The window query finds all objects in the database that shares at least one point with the query object, a *d*-dimensional

interval. The point query is a specialization of this query type, where the query interval is reduced to a point in *d*-dimensional space.

3. Nearest-neighbor queries. These are the well known nearest neighbor queries that, given a query item, find the item that is closest or most similar to the query item. Variants are k-nearest neighbor queries that find the k element closest to a query, and fixed-radius-near-neighbor queries that retrieve all objects with a distance of at most *r* from the query item.

All these query types can be used for CBR. The nearest-neighbor queries can be directly used to determine the most similar cases. (Kamp *et al.* 1996)

2.6.3 Integration via Extensible Databases

The multidimensional access models were mainly of academic interested, only few of them were deployed in commercial database servers. However, the world has changed radically. The application like Internet and Intranet, OLAP, etc., most database vendors are working on extensible database server. These so called object-relational database management systems (ORDBMS) can be extended to incorporate multidimensional data types like spatial information, texts, etc. Microsoft, IBM, Informix and Oracle are already providing such servers. Multidimensional access structures for geographic information systems as well as multimedia data and texts are already available.

CHAPTER 3

IMPLEMENTATION

3.1 Background of the Critical Issue System

Study of real world model to understand and study the actual case is necessary. With the assistance of a construction firm, Yau Lee Construction Ltd., completed past buildings were collected as the cases for developing the case base.

Figure 3.1 illustrates the data and document exchange between architectural, engineering and construction professional with current computer operation, paper work and human operation. Parties involve in data exchange are inherent as system functional entities, such as authorization unit and operator of critical issue. Every entity performs its operation. For example, file linkage / attachment from server's file system, will be shared with related parties. Also, the unfilled box: like result condition, solution of events, etc. should be completed by appropriate entity directly involved in such cases.

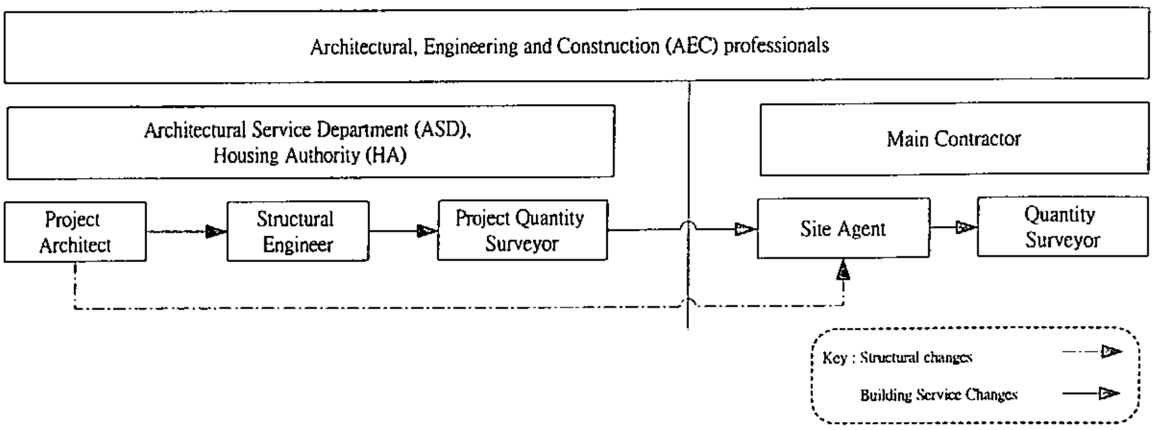

Figure 3.1 Data & Document exchange between AEC professionals

3.2 Overview of System I/O and Object Operation

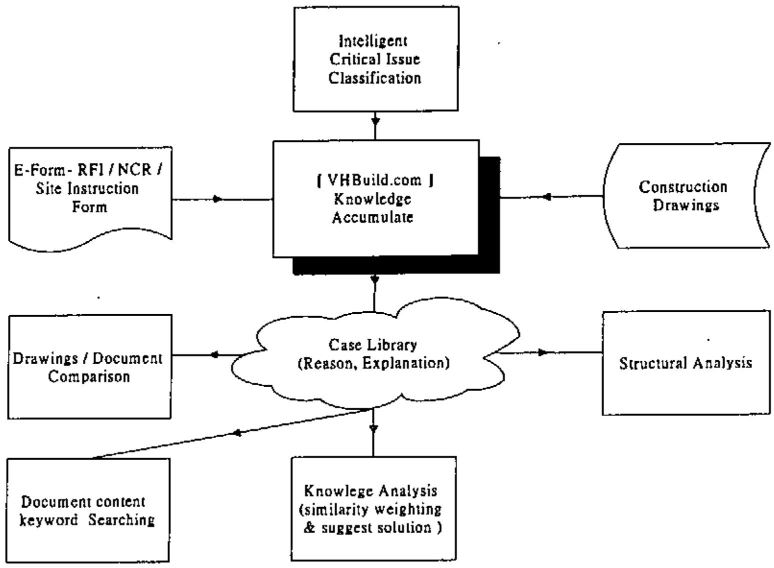

Figure 3.2 System functional diagram

3.3 System Functionality

The knowledge management system is named as the "Critical Issue", which is one of the modules in the web base project management system VHBuild.com. The system structure of the Critical Issue system is divided into 4 parts:

- *Issue*, which describes the fundamental information of the problem or the event, such as the construction stage or event description.
- *Cause*, which describes the reason or cause of the event. The reason may be related to design, human or procedure factors, etc.

- *Effect*, which shows the date range of the event and the area affected, including time, environment, etc.
- *Solution*, in which the domain expert or the information provider can provide the solution to the problem or the suggestion on how to solve the problem.

By formulating and inserting the elements of a critical issue, the case base consists of two portions: problem and solution. Each portion contains various features to represent the contents of the case.

In summary, a case is represented by: Issue section, cause section, affect and solution section. These sections are further described below.

3.3.1 Issue Section

The issue section describes the background information of construction event. The descriptive information would be obtained from VHBuild.com construction project management system. This information include:

- *Issue name*

  As a memorable name that represents and indexes the critical issue easily.

- *Description*

It describes some background element of the event. For example, if a case involves changing and assigning a new subcontractor for fitting-out works, the action and procedure taken to submit safety plan and quality plan is the description of this case.

- *Stage type*

It affects the ease of formulating cases. It indicates the stage in which the critical issue occurs. Basically the stages include:

- Construction, the event occurs in the construction stage, such as constructions sub-structure, super structure, service, finishes, utilities.
- Maintenance, the event occurs in the maintenance stage.
- Handover, the event happens in the handover stage, which includes certification and site investigation.

- *Problem solver*

The person solved the problem.

- *Responsible party*

Who may carry out the problem solving procedure, or problem occurs due to their fault. For example, quality surveyor team, out budget of fitting-out work and lends to subcontract the fitting-out process to other subcontractor.

3.3.2 Cause Section

The section captures the reasons of the event and issues that formulate the problem. From the Case Base Reasoning point of view, it is also the preprocessing stage of a case base, the different type of causes are calculated into a fuzzy set, in order to evaluate the textual classification during comparison or searching. Causes of a critical issue include the following:

- *Human cause*

  It captures the reason related to the human factor. For example, construction cost being under estimated by a quality surveyor.

- *Design cause*

  It concerns design related causes. For example, during the architectural design, the request of design change can cause problems in the construction stage.

- *Material cause*

  The cause is materials related. For example, for the reason of inadequate supply of metal ceiling, it increases the cost of fitting-out work.

- *Plant cause*

  This type of cause relates to malfunction of machine and may lead to construction accidents. For examples, evaluator malfunction causes worker injuries, fuel cost inflation affects the operation cost of power shovel, lorry, and other machinery.

- *Procedure cause*

  This type of cause relates to scheduled process or workflow rearrangement, or modification of construction procedure that incur errors during the actual performance.

- Legal / Contractual cause

  Due to the change of construction ordinance or building regulations, safety law for example, the degree of fulfillment may not be guarantee or even breaking the law.

3.3.3 Effect Section

The effect section captures the duration of the event affecting the project operation. Also it describes the aspects affected, which includes quality, time, safety and environmental impact. The effect section is elaborated by:

- *Date from*

  The starting time of a Critical Issue

- *Date to*

The ending time of a Critical Issue

- *Area affected: Quality*

  It measures the effect of quality by the Critical Issue. For example, the quality of fitting-out work will be affected by the bad quality of material supply.

- *Area affected: Time*

  It measures the effect of time by the critical issue. For example, extension of construction time due to higher typhoon signal is hoisted.

- *Area affected: Safety*

  It measures the effect of a critical issue on the safety. Inefficient or incorrect planning performed by a safety officer may affect the working environment.

- *Area affected: Environmental*

  It measures the effect of a critical issue on the environmental impact. For example, water pollution due to incorrect design of drainage system, noise pollution due to overtime work of site activities are some aspects affected by a critical issue.

3.3.4 Solution Section

This section describes the solution provided by domain expert or related user, which helps the case-based reasoning system to identify a new solution and adapt it for the new problem.

- *Design solution*

  It describes the complementary answer of the problem described in the "Design Cause" at the previous section.

- *Material solution*

  It describes the complementary answer of the problem described in the "Material Cause" at the previous section.

- *Plant solution*

  It presents the complementary answer of the problem described in the "Plant Cause" at the previous section.

- *Procedure solution*

  It presents the complementary answer of the problem described in the "Procedure Cause" at the previous section.

- *Human solution*

  It presents the complementary answer of the problem described in the "Human Cause" at the previous section.

- *Contractual solution*

It presents the complementary answer of the problem described in the "Legal / Contractual Cause" at the previous section.

3.3.5 Result Section

The result section presents the consequences of the critical issue. The results can be categorized into different aspects, such as time, quality, safety and environment, which are the major factors that influencing future development of construction project.

By providing the suggestion and accumulating useful experience, the conclusion is given by the domain expert or related user after completing the critical issue. During the process of case base retrieval and adaptation, the solution and the result section play an important role in formulating the new problem solution. The result section can be characterized by:

- *Time*

  It is the result used to evaluate the appropriateness of time planning.

- *Quality*

  It concludes the different quality levels before and after the solution is applied to that event. It is useful when performing quality evaluation, like validating the construction project using ISO 9001.

- *Safety*

Also, after applying the solution, results of safety walkthrough can then be concluded and given in that section.

- *Environment*

Environmental impact caused by the events is measured as pollution or damage to the surrounding environment.

3.3.6 Attributes Used in a Case Representation

Having reviewed the influencing factors, a case is therefore represented by the following attributes or features summarized in Table 3.1:

Attribute	*Candidate content*	*Local similarity measure*
Issue name	Textual base	Fuzzy set
Description	Textual base	Fuzzy set
Stage type	SubStructure/ SuperStructure/ Service/ Finish/ Utility/ External/ Maintenance/ Testing/ Handover/ Certification/ SiteInvestigation	Exact
Raised person	Record explicitly defined in record set	Exact
Human cause	Textual base	Fuzzy set
Design cause	Textual base	Fuzzy set
Material cause	Textual base	Fuzzy set
Plant cause	Textual base	Fuzzy set
Procedure cause	Textual base	Fuzzy set
Legal / Contractual cause	Textual base	Fuzzy set

Date from	Date format	Exact
Date to	Date format	Exact
Quality affect	Textual base	Fuzzy set
Time affect	Textual base	Fuzzy set
Safety affect	Textual base	Fuzzy set
Environmental affect	Textual base	Fuzzy set
Design solution	Textual base	Exact
Material solution	Textual base	Exact
Plant solution	Textual base	Exact
Procedure solution	Textual base	Exact
Contractual solution	Textual base	Exact
Time result	Textual base	Fuzzy set
Quality result	Textual base	Fuzzy set
Safety result	Textual base	Fuzzy set
Environment result	Textual base	Fuzzy set

Table 3.1 Attributes representation in Case Based Reasoning System

3.3.7 Case collection and representation

A collection of past cases is essential to the development of a CBR system. An exhaustive list of project events is drawn up by discussion with experienced engineers, and by studying the official instruction and assessment reports. Cases are classified into architectural, structural, electrical engineering, air-conditioning and fire services. Information in the VHBuild.com project management system are translated or mapped into the case representation structure as shown in Table 3.1. To protect sensitive information, which replaced by project identifiers. The followings is the conditional summery of case that being studied and described as follows:

1. **Design development and improved design standards**

   New design standards may be adopted thus increasing the project cost. For example, expansion of provision sum for external lighting builders work.

2. **Special temporary works**

   Include situations of bridge foundations in navigation channels, bridges spanning roads and railway lines, etc.

3. **Tight program for phased possession/handover**

   Part of the work site may need to be handed over to another party for other projects early, restricting the flexibility of the working program.

4. **Land resumption delay site possession**

   Part of the site may not be available due to problems of land resumption.

5. **Construction has serious drainage impact**

   Relevant for highway works in rural areas susceptible to flooding.

6. **Drainage improvement required**

   Drainage improvement works within or adjacent to the works site may be added to the project scope.

7. **Tight construction program with heavy penalty**

   There is political pressure to complete the project in a certain time frame. This could result in a very tight construction period, which would drive up the cost.

8. Special contract arrangements

This includes the deletion of contract price fluctuation clause, entitlement to extension of time due to inclement weather, etc. in the contract documents.

9. Pricing data for inaccurate base estimation

It caused by lack of recent contracts of similar nature at similar location.

10. Variation orders

This is related to design changes during construction, mostly to suit unforeseen site conditions.

11. Complicated works coordination

In contracts that involve large number of sub-contractors and suppliers, complicated works coordination would increase the cost of site management.

12. Contractual claims

Contractual claims can be expected where unforeseen site conditions would have a large impact on cost and project progress.

13. Tight construction and labor market at time of tender

Construction materials can be imported easily, labor supply cannot. The supply of labor at time of tender is a risk factor. The supply of construction labor is closely related to the state of the construction market. Tender prices fluctuate with the state of the construction market. Contractors demand and get a high profit margin during the boom time.

14. Sub-contractor bankrupt

Accidental contingency should be carried out, so that it can avoid the tragedy in which may affect the project schedule and the followed large contracts.

3.3.8 Case Similarity Measurement

The essence of CBR is to retrieve similar past cases for solving new problems. The system searches the case base using the similarity measure. The purpose of this is to select cases that have nearly the same solution to the current problems, or that can be adapted easily to the current problem. As many cases may be more or less useful to a query, similarity must be quantitatively measured. Start from beginning, the system compares two cases in terms of local and global similarities. Local similarity deals with the values of a single feature, whereas global similarity represents a holistic view of the cases. The global similarity measure combines local similarity measures, we select weighted sum with different weights of relative importance of different attributes or features. The intuitive interpretation for the combination of local similarity is the collection of arguments in support of a particular case.

Similarity measures can have different forms according to the underlying domain, which may be symbolic, ordered, or numerical. We use numerical representation to simulate discrete measurement of attributes and feature, Textual notions such as cause-description, effect, can be evaluated using approaches from fuzzy and neural network theory.

The local similarity measures the attribute and categories into 2 types, exact match and fuzzy set. The exact match is required for symbolic variables except the linguistic variables that describe the conditions of stage type, raised person, date range, design solution, material solution, plant solution, procedure solution and contractual solution. For attribute values a and b of two cases, the similarity measure or function for exact match is:

Sim(a,b) = 1 if equal, 0 if otherwise

A similarity measures provided by the system "fuzzy set feature matching" is used for the numeric variable of human cause for example after weighted features. It measures the closeness of match between numeric values. A similarity measure in the interval of 0 and 1 is calculated based on how close the values are from each other. For example, 100 and 80 differ by 20 percent, therefore, if a tolerance of 40 percent is specified, the similarity would be 50 percent. Processing details and functional specification in the CBR are provided in the following sections.

3.3.9 Case Feature Selection

We introduce the feature selection method, it is a process to obtain the case value with individual attributes and estimate the total sum within a single case, so that the similarity function can be performed in a numerical adjustment.

A traditional CBR systems, all the features or fields will be used to draw the solution. However, some features may have different degrees of importance for solution determination. In some situations, unnecessary fields may affect the performance and accuracy of the entire CBR system. Feature selection can help to identify the degrees of importance so that the case becomes more useful. We propose the alternative techniques for feature selection. These include, null feature selection, random selection and feature weight generation, in which it depends on the user input area and volume of information provided.

3.3.10 Feature Selection Method – Null feature selection

The traditional CBR systems do not have any feature selection. It should be noticed that in some situations, all the fields or attributes involved would affect the accuracy at the same degree. In most situations, it is necessary to apply the process of feature selection. A standard back-propagation NN can perform the process of case classification.

Let CL = $\{e_1, e_2, \ldots, e_N\}$ denote our case library. Each case in the library can be identified by an index of corresponding features. In addition, each case has an associated action. More formally we use a collection of features $\{F_j\,(j = 1, \ldots, n)\}$ to index the cases and a variable V to denote the action. The *I-th* case e_i in the library can be represented as $n+1$ – dimensional vector, i.e. $e_i = (x_{i1}, x_{i2}, \ldots, x_{in}, v_i)$ where x_{ij} corresponds to the value of feature F_j $(1 \leq j \leq n)$ and v_i corresponds to the action (I=1, …, N). v_i $(I=1, \ldots, N)$ are considered to be class symbol and the total number of classes is supposed to be M. The M classes are denoted by $CM=\{C_1, C_2, \ldots, C_M\}$ (Baluja *et al.*, 1995).

Then, we model the classification using a neural network with three layers. The classification result will be a fuzzy set on cluster space CM. The key structure of the constructed neural network is described as Figure 3.3

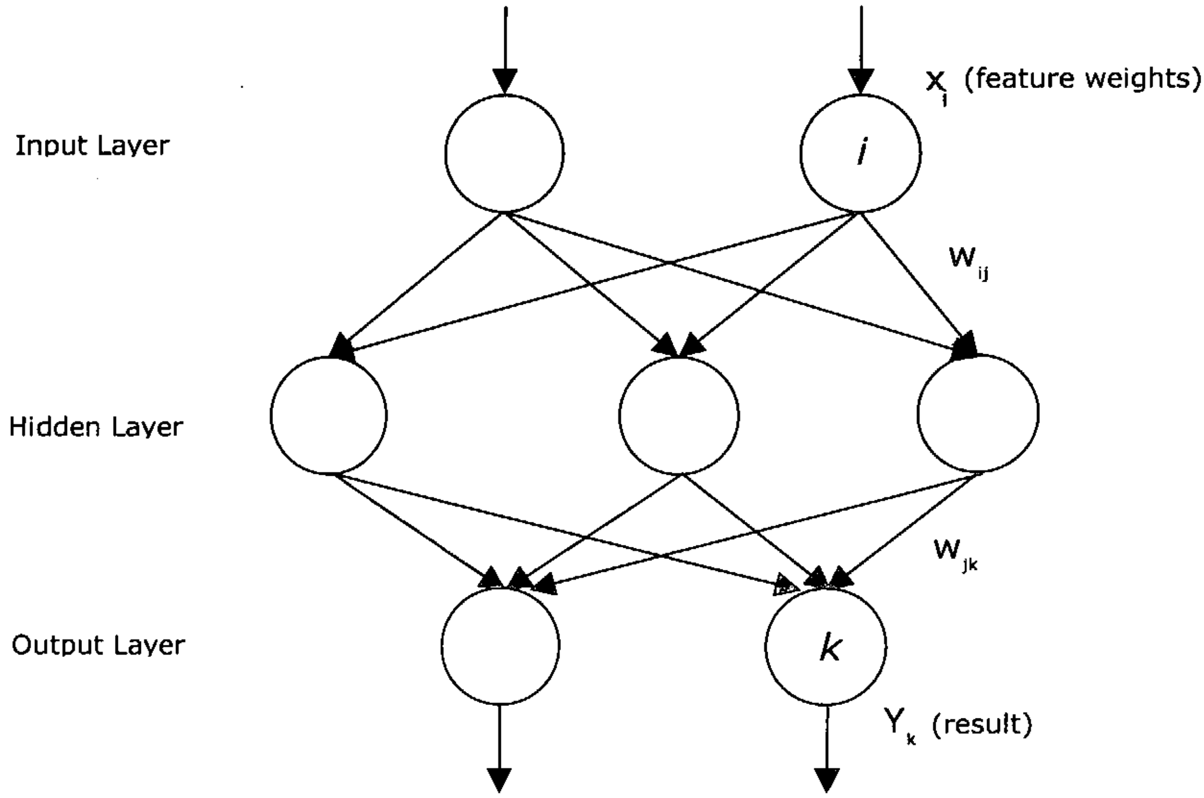

Figure 3.3 Feature Weights in Neural Network

Input Layer

The number of nodes in this layer is equal to the number of attribute/feature of the case base. Each node represents an attribute.

Hidden Layer

The number of nodes in this layer is determined according to real applications. The number h is ($n < h < 2n$), where n is the number of nodes in input layer.

Output Layer

This is the classification layer that contains M nodes where M is the number of clusters. Each node represents a fuzzy cluster. The training result will be form of fuzzy vector (discrete fuzzy set defined on the cluster space *CM*). The meaning of each output value is the membership value indicates to what degree the training case belongs to the cluster corresponding to the node.

The popular Sigmoid function is selected as the activation function. According to assumptions, there are n, L, M nodes in the input, the hidden and the output layers respectively (Kuo *et al.*, 2001). For a given input case (the *m-th* case, $1 \leq m \leq N$), the forth-propagation process of the input vector is described as follows.

The input layer : $\{ x_{mi} \mid i = 1,2,\ldots,n\}$ (the given input vector);

The hidden layer : $y_{mj} = f\left(\sum_{i=1}^{n} u_{ij} \quad x_{mi}\right) \quad j=1,2,\ldots,L$ Equation (1)

The output layer : $\mu_{mk} = f\left(\sum_{j=1}^{L} v_{jk} \quad y_{mj}\right) \quad k=1,2,\ldots,M$ Equation (2)

Where u_{ij} and v_{jk} are the connection weights of the neural network, and the notation f represents the Sigmoid function defined as $f(x) = \dfrac{1}{1+e^{-x}}$

This is a traditional full connection network with three layers. The standard Back propagation algorithm can be used to train this network. In other words, the popular gradient descent technique can be used to find the values of weights u_{ij} and v_{jk} such that the error function

$$E = \sum_{m=1}^{N}\left(\frac{1}{2}\sum_{k=1}^{M}(\mu_{mk} - c_{mk})^2\right) = \sum_{m=1}^{N} E_m \qquad \text{Equation (3)}$$

achieves a local minimum, where c_{mk} taking either 0 or 1 corresponds to the action of the *m-th* case, e.g., $(c_{m1}, c_{m2,\ldots,} c_{mM}) = (1,0, \ldots, 0)$ if the *m-th* case belongs to the first cluster.

After finishing this training, a fuzzy set on the cluster space $\{c_1, c_2, \ldots, c_M\}$ can be given for each case according to equations (1) and (2). Denoting the fuzzy set by $(\mu_{m1}, \mu_{m2}, \mu_{mM})$ in which each component μ_{mj} represent the degree of the *m-th* case belonging to the *j-th* cluster will be calculated.

3.3.11 Feature Weight Generation

We introduce the parameter of feature weight, which indicates that different features have different degrees of importance to the training classification.

Let the training results be $\{(\mu_{m1}, \mu_{m2}, \ldots, \mu_{mM}), m = 1,2, \ldots, N\}$ for N cases $\{ e_m, m = 1,2, \ldots, N\}$ where $\mu_{mk} (\in [0, 1])$ represents the degree of the *m-th* case belonging to the *k-th* cluster. An index evaluation is defined as (Basak 1998):

$$E(w_1, w_2, \ldots, w_m) = \sum_{m=1}^{N} \sum_{k=1}^{M} \left(\mu_{mk}^{(w)}(1 - \mu_{mk}) + \mu_{mk}\left(1 - \mu_{mk}^{(w)}\right)\right) \quad \text{Equation (4)}$$

in which μ_{mj} is computed according to equation (1) and (2), and $\mu_{mj}^{(w)}$ according to the following equations (5) and (6):

$$\mu_{mk}^{(w)} = f\left(\sum_{j=1}^{L} v_{jk} y_{mj}^{(w)}\right) k = 1,2,\ldots,M \quad \text{Equation (5)}$$

$$y_{mj}^{(w)} = f\left(\sum_{i=1}^{n} w_i u_{ij} x_{mi}\right) k = 1,2,\ldots,L \quad \text{Equation (6)}$$

where u_{ij} and v_{jk} are the weights trained already in the previous phase, and f represents the Sigmoid function defined as $f(x) = \frac{1}{1+e^{-x}}$

In equation (4), $\{ w_i : w_i \in [0,1], i = 1,2,\ldots,n \}$ are called feature weights which remain to determine during case interpretation. I.e. influence during new case enter to system. They indicate that, for the trained neural network in previous phase, different features have different degrees of importance to the training-classification.

The evaluation function (6) is designed according to a simple function g(x,y) = x(1-y)+y(1-x) ($1 \le x, y \le 1$).

Nothing that $\frac{\partial g}{\partial x} = 1 - 2y, \frac{\partial g}{\partial x} > 0\ if\ y < 0.5, \frac{\partial g}{\partial x} < 0\ if y > 0.5$, one can easily find that equation (6) has the following characteristics:

(a) if $\mu_{mk} < 0.5\ and\ \mu_{mk}^{(w)} \to 0,\ then\ E \to 0$ (minimum);

(b) if $\mu_{mk} > 0.5\ and\ \mu_{mk}^{(w)} \to 1,\ then\ E \to 0$ (minimum);

(c) if $\mu_{mk} = 0.5\ and\ \mu_{mk}^{(w)} = 0.5,\ then\ E\ attains\ maximum$;

The main task of this phase is to minimize the index function $E(w_1, w_2, \ldots, w_n)$ with respect to the weights w_1, w_2, K, w_n. More formally, we attempt to find (w_1^*, w_2^*, K, w_n^*) such that

$$E(w_1^*, w_2^*, K, w_n^*) = Min\{E(w_1, w_2, K, w_n) \mid w_i \in [1,0], i = 1,2, K, n\} \qquad \text{Equation (7)}$$

Minimizing the index function E is regarded as a process of refinement. If we consider $\mu_{mk}\ and\ \mu_{mk}^{(w)}$ as the membership degrees of the *m-th* case belonging to the *k-th* cluster before refinement and after refinement respectively, the minimization of equation (4) attempts to make the membership degree after refinement. That is, we expect that the minimization of equation (4) can make the membership degree after refinement being close to 0 if the membership degree before refinement is less than 0.5; and the membership degree after refinement being close to 1 if the membership degree before refinement is bigger than 0.5.

The process of Feature Weight Training can be summarized as in Figure 3.4

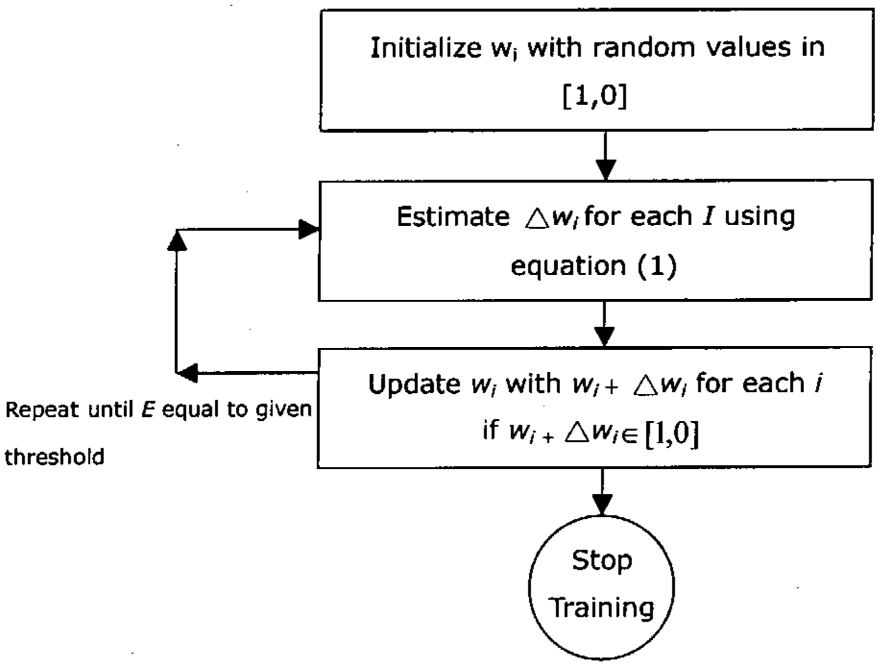

Figure 3.4 Feature Weight Training

After the training process, the function $E(w_{1,} w_{2,} ..., w_{n,})$ attains a local minimum. We assume that, the membership degrees $\{\mu_{mk}^{(w)} : m = 1, K, M \}$ with trained weights close to 0 or 1 than $\{\mu_{mk} : m = 1, K, N; k = 1, K, M \}$ without trained weights.

each case to a new critical issue. The case base is indexed by the keywords of each case. When a new critical issue is defined as a set of issues and goals, the system traverses the cases according to the new definition of critical issue, and identifies the similarities between the cases and the new events. The case with the highest similarity is then selected as a basis for generating a solution for the critical issue. To model this selection process, a weighted count of matching keywords can be applied. User interaction allows the set of matching keywords and their relative importance to be modified.

The membership degree has obtained by training a neural network. The selection of the case representatives can then be performed. The selection strategy is mainly based on the case-density and membership degree. This selection makes use of the membership degree and the concept of case density (Smyth 1998), which is defined as follows:

$$\text{CaseDensity}(c,G) = \frac{\sum_{C' \in G-\{c\}} sim(c,c')}{|G|-1} \qquad \text{Equation (8)}$$

$$\text{Sim}(c,c') = 1 - \frac{\left[\sum_i w_i (c_i - c'_i)\right]^{1/2}}{n} \qquad \text{Equation (9)}$$

In Equation (8), G denotes a class of cases, |G| is the number of cases in class G, c and c' are two cases, c_i and c_i' are the *i-th* component of cases c and c' respectively, n is the number of features, w_i is the feature weight learned in phase 2 (all w_i will be equal to 1 if phase 2 is skipped), and Sim(c,c') represents the similarity between cases c and c'.

Our selection algorithm is described as follows:

Step 1. Select two thresholds α and β for membership degree and case density.

Step 2. Compute the case density for every case in each class.

Step 3. Select representative cases from every other class if their membership degree of belonging this cases are greater than or equal to α and their case densities are greater than or equal to β.

Once the similarity between cases exceeds a pre-determined threshold (we set at 0.75), the case with the highest similarity is selected as the most similar case, and the solution presented in the selected case is adapted to suit the specific situation presented in the problem at hand. This process is known as case adaptation. Case adaptation forms the essence of solution generation. As the case selection provides a case that is close to an acceptable solution and adapts those aspects of the case that are inconsistent need to be adapted. The selected case should contain most issues and goals that define the problem of a critical issue to be resolved. Knowledge used in case adaptation includes heuristics, common senses and issues, and proportional relations. Modifying reservation values, introducing new issues/goals, and selecting additional cases are three techniques of case adaptation.

3.3.14 Case Adaptation

Case with global similarity above a user-defined threshold value are retrieved and listed alongside their global similarity measures. The user is to visually compare retrieved case feature values to the problem case feature values and determine if a retrieved case is a good match to the problem case. Usually but not necessarily, the case with the highest global similarity measure is selected for adaptation.

Adaptation of the selected case depends largely on user interpretation and intervention to account for the different between the selected cases and the problem case.

3.3.15 Learning

A CBR system learns by incorporating new cases, adjusting similarity measures, and adaptation rules. The case base of Critical Issue is constructed using historical experience. New cases can be added with new experience data, and existing cases can be deleted when case selection.

3.3.16 Case Study

Case 1: Remove Asbestos Containing Materials in Kwong-wah Hospital

Attribute	*Candidate content*	*Local similarity measure*
Issue name	Remove asbestos containing materials in Kwong-wah hospital 19960401	Fuzzy set (Keywords: Remove asbestos)

Description	Remove asbestos containing materials in Kwong-wah hospital's wards, operation theatre, general office and machinery room. Parts mainly made before the mid 1980s Contract no: KWH77 Form no: TA 7765 OW	Fuzzy set (Keywords: electrical accident)
Stage type	Maintenance	Exact
Raised person	Real Estate Officer: C.W. Kwong	Exact
Human cause	Medical report about asbestosis and lung cancer posted after construction project of Kwong-wah hospital.	Fuzzy set (Keywords: asbestosis and lung cancer)
Design cause	-empty-	Fuzzy set
Material cause	Decoration, fitting out made and suspected asbestos containing materials in hospital ward and operation theatre before the mid 1980s: * thermal insulation : lagging materials for steam and hot water pipe, heater, boiler, furnace, chimney, flue duct * electrical insulation : switch box arc chutes, cable tray * sound absorption : acoustic plaster sprayed on ceiling and wall	Fuzzy set (Keywords: materials in hospital)

	* fireproofing : filling material in wall and floor penetration, laboratory bench top, fire blanket, fire curtain * building materials : corrugated cement sheet for roofing, wall board, cement roof tile, asphalt roofing felt, vinyl floor covering, cement soil stack, cement pipe, refuse chute, ceiling tile * friction products : brake lining, clutch facing * building services : air duct flexible joint, cable trough and conduit, cistern * sealing and jointing : gasket, gland packing materials for pump and valve, putty, adhesive	
Plant cause	-empty-	Fuzzy set
Procedure cause	-empty-	Fuzzy set
Legal / Contractual cause	-empty-	Fuzzy set
Date from	1995/02/10	Exact
Date to	1995/07/28	Exact
Quality affect	-empty-	Fuzzy set
Time affect	Work out time is necessary:	Fuzzy set

	Finishing deadline: 6 months	(Keywords: finishing deadline)
Safety affect	6 medical personnel get asbestosis, 2 technicians get lung cancer. Asbestos can split into very fine fibres that can remain airborne for long periods of time after release. When inhaled, asbestos fibres can remain in the body for many years. The following diseases caused by exposure to asbestos fibres: * lung cancer * mesothelioma, a cancer of the lining of the chest or of the abdominal wall * asbestosis, scarring of the lung with fibrous tissue The symptoms of asbestos related diseases is not appear until about 10 to 40 years after first exposure to asbestos.	Fuzzy set (Keywords: lung cancer)
Environmental affect	Air Pollution	Fuzzy set (Keywords: air pollution)
Design solution	-empty-	Exact
Material solution	1. Replace the asbestos with heat resistance PVC plastic & fabric mixed ceiling. 2. Replace the floor tile with unbreakable plastic 3. Remove spray insulation with	Exact

	pressured paper fabric sheet.	
Plant solution	-empty-	Exact
Procedure solution	Problem procedure: 1. Registered asbestos consultant to conduct an asbestos investigation. 2. Prepare an asbestos investigation report and an asbestos abatement plan. 3. Submit the asbestos investigation report and the asbestos abatement plan to the Environmental Protection Department at least 28 days before start. 4. Work on the asbestos containing materials. 5. Notify the Environmental Protection Department of the date of commencement of work. 6. Asbestos containing materials at least 28 days in advance. 7. Registered asbestos contractor to carry out the asbestos work in accordance with the asbestos abatement plan. 8. Registered asbestos consultant to supervise the implementation of the asbestos abatement plan and the work of the registered asbestos contractor. 9. Registered asbestos laboratory to conduct sampling and analysis for the asbestos work.	Exact
Contractual solution	-empty-	Exact

Time result	Replacement time for whole hospital: 6 year	Fuzzy set
Quality result	-empty-	Fuzzy set
Safety result	The asbestos waste is packed in the approved containers as soon as it is produced, and handled and transported as recommended in this Code of Practice, there will be little risk of asbestos fibres being emitted. However, as a further precautionary measure, safety and emergency handing equipment should be carried on every delivery vehicle including - * approved face masks for all workers. * appropriate heavy duty rubber gloves. * protective clothing. * working shoes. * spare containers/bags and seals. * shovel. * amended water (water mixed with wetting agents) in airless spraying equipment - about 50 L. * emergency and first-aid kit. The workers should wear the protective clothing and other safety equipment during all loading and unloading operations. Immediate action must be taken to prevent asbestos fibres from being	Fuzzy set (Keywords: precautionary measure)

	released to the air in the event of an accident during transportation, or at a disposal site, resulting in the bursting or splitting of containers and the release of asbestos waste. This is done by immediately wetting the waste with amended water applied in a fine mist. Protective clothing, appropriate gloves, working shoes and an approved dust mask should be worn for this operation in order to prevent contamination of hair, skin and personal clothing. The gloves are used for protection against friction and abrasion during handling of drums as well as for prevention of contamination. The shovel and spare bags or drums together with seals carried on the vehicle are used for repacking and resealing the split waste. In case of extensive contamination of the environment, EPD should be notified immediately. Contaminated vehicles should be thoroughly washed. The washings should be disposed of into the asbestos waste disposal trenches at the disposal sites. All contaminated equipment should be thoroughly washed before reuse. However, if they are difficult to clean, they should be treated as asbestos waste and placed in suitably labeled plastic bags for proper disposal.	
Environment result	-empty-	Fuzzy set

Case 2: New Fire Services Safety Regulation, which affect and require amendment of fire service design.

Attribute	*Candidate content*	*Local similarity measure*
Issue name	Fire Services Inspection 19971008	Fuzzy set (Keywords: Fire Services Inspection)
Description	Air Traffic Control Complex and Tower at Chek Lap Kok Airport, fire services inspection reports that modification required for new safety regulation. Contract no: SSC307 Form no: FS-22	Fuzzy set (Keywords: Modification required)
Stage type	Maintenance: Testing	Exact
Raised person	Chief Architect: L.W. Chan	Exact
Human cause	Inadequate communication with Fire Service Department.	Fuzzy set (Keywords: Inadequate communication)
Design cause	Outdated design specification cannot satisfy Fire Service regulation	Fuzzy set (Keywords: Outdated design)
Material cause	-empty-	Fuzzy set
Plant cause	-empty-	Fuzzy set
Procedure cause	-empty-	Fuzzy set
Legal / Contractual cause	-empty-	Fuzzy set
Date from	1996/10/10	Exact

Date to	1997/10/01	Exact
Quality affect	Textual base	Fuzzy set
Time affect	Time adjustment is necessary: Extension of time: 5 working days	Fuzzy set (Keywords: Extension of time)
Safety affect	Ensure the delivery process is under monitored. Due to relocation of FM200 panel c/w conduit and wiring. and followed by architectural layout revised.	Fuzzy set (Keywords: Monitor delivery process)
Environmental affect	-empty-	Fuzzy set
Design solution	-empty-	Exact
Material solution	-empty-	Exact
Plant solution	-empty-	Exact
Procedure solution	Installation procedure: 1. Add 2 nos. of smoke detector at Booster Pump Room. 2. Add 4 nos. of smoke detector under false ceiling at lift lobby. 3. Add 4 nos. of smoke detector at protected lobby. 4. Add 2 nos. of smoke detector at genset room. 5. Add 3 nos. of smoke detector under ceiling soffit/false ceiling at control cab staircase. 6. Add 3 nos. of smoke detector at lift machine room on basement of Tower. 7. Add 2 nos. of fire hydrant at 5/F of Complex. 8. Add 1 nos. of FM200 cylinder at	Exact

	room T Room due to change the building plan. 9. Add 8 nos. of genset signal from genset room to AFA panel. 10. Add and relocate 1 no. of interfacing unit from lift supervisory panel to 4 nos. of lift machine room. 11. Add 3 nos. off indication lamp at genset room and transformer room. 12. Relocate 2 nos. of FM200 panel c/w conduit and wiring at room G23. 13. Modify FM200 system due to architectural layout revised, including re-design the FM-200 system of 3 cylinders and associated piping, re-route the conduit and 25mm dia. 14. Add 2 nos. flashing lamp, 1 nos. alarm bell & siren at room MT Add 2 nos. of pressure trip for fire curtain at fuel oil tank room.	
Contractual solution	-empty-	Exact
Time result	Delay project construction schedule by extension of time: 5 working days	Fuzzy set (Keywords: delay schedule)
Quality result	-empty-	Fuzzy set
Safety result	Introducing new firing system avoid and increasing the ability of fire resistance.	Fuzzy set (Keywords: fire resistance)
Environment result	-empty-	Fuzzy set

Case 3: Disaster Recovery by Electrical Accident

Attribute	*Candidate content*	*Local similarity measure*
Issue name	Disaster Recovery by Electrical Accident at 19970120	Fuzzy set (Keywords: Disaster Recovery)
Description	Air Traffic Control Complex and Tower at Chek Lap Kok Airport, disaster recovery of electrical accident due to explosion of FM200 system. Contract no: SSC301 Form no: EE-42	Fuzzy set (Keywords: electrical accident)
Stage type	Utility	Exact
Raised person	Engineer's Representative: C.M. Chan	Exact
Human cause	Improper electrical maintenance: explosion and fire due to short circuit of broken wire.	Fuzzy set (Keywords: Improper maintenance)
Design cause	-empty-	Fuzzy set
Material cause	Wire quality below standard. Inadequate resistance of temperature variation.	Fuzzy set (Keywords: below standard)
Plant cause	-empty-	Fuzzy set
Procedure cause	-empty-	Fuzzy set
Legal / Contractual cause	-empty-	Fuzzy set
Date from	1997/01/20	Exact
Date to	1997/01/29	Exact
Quality affect	Site management procedure under ISO 9001 standard. Accident	Fuzzy set

	happened affects scoring grade.	(Keywords: ISO 9001)
Time affect	Time adjustment is necessary: Extension of time: 9 working days	Fuzzy set (Keywords: Extension of time)
Safety affect	3 technicians and 1 engineer injured. Due to short circuit explosion.	Fuzzy set (Keywords: engineer Injured)
Environmental affect	-empty-	Fuzzy set
Design solution	-empty-	Exact
Material solution	1. Add adaptable box at every electrical room. 2. Replace conduct wire with double fire resist plastic clothing and fulfill CG & FCC international standard	Exact
Plant solution	-empty-	Exact
Procedure solution	Setup procedure: 1. replace 2 nos. 13A fused spur unit electrical room. 2. revise lighting layout in Lockers Room and Equipment room. And add 8 nos. SFC 4 in FM200 room. 3. replace 200mm(W) cable tray in Transformer room. 4. reposition 24 nos. 13A fused spur unit at upper level. 5. reposition 16 nos. fused spur unit and add 22 nos. fused spur unit. 6. add water leakage detection system in lift lobby B at second floor and CAD/RO duct. 7. add 1 nos. adaptable box at top of	Exact

	door and 25mm dia. Conduit at T1 floor.	
Contractual solution	-empty-	Exact
Time result	Extension of time: 9 working days	Fuzzy set
Quality result	Increase the maintenance efficient by replacing all wire and electrical equipment.	Fuzzy set (Keywords: maintenance efficient)
Safety result	Improve safety condition by enhanced electrical system	Fuzzy set (Keywords: improve safety condition)
Environment result	-empty-	Fuzzy set

3.3.17 Case Presentation

As mentioned in process of CBR, a new case will compare to each record in case library (*CL*), it maintains a feature scoring table which records the compared feature field from individual case. For example,

> if Rec(x) is the incoming case,
>
> 1,000 records maintain in case library,
>
> then, 999 records will generate and belong to record Rec(x), which contains the numeric value of compared feature matching score with others case in CL.

Step 1 Feature selection

Since now we have totally 28 attributes / features, we decide to use 14 features (half of the total) as bottom lines of feature method selection.

E.g. For estimating the weight value of Fuzzy Set attribute

IF (*incoming case provide more than 14 attributes*) THEN

Propagate Neutral Network to estimate the trained weight of each feature

ELSE IF (*incoming case provide more than 5 but less than 14 attributes*) THEN

Random selection of weighted value, within 0 or 1

ELSE IF (*incoming case provide less than 5 features*) THEN

Null feature selection applied

ELSE

Do nothing

END IF

Step 2 Estimate the feature value of each attribute

IF (*Fuzzy Set attribute*) THEN

1. Occurrence counting of every attribute

2. Calculate the average occurrence value among all case in Case Library (CL)

ELSE

Exact matching with relative attribute, return value either 0 or 1

END IF

Step 3 Formulate the match score

Using the feature value result in **step2** and times the result in **step 1**, the match score can be estimated.

3.4 System Architecture

There are two versions of CBR system that has been developed. The first one is in the client server structure, which responsible to integrate the domain expert opinion, for example, structure engineer and site agent can provide idea during creating the site instruction and daily report. Since executable program can be distributed and run at site computer. It facilitates the adoption versus changing environment because of the system functional requirement can be effectively defined.

After the stage of system design and raw data is obtained by client server program, the Internet version can than be developed to centralize the case storage and data retrieval for different site users and project participants.

Data collected in client server version is always unstructured. The data normalization should be taken to maintain data consistency. The following example demonstrates the procedure of data normalization.

First Normal Form

The objective of first normal form is to ensure that every cell in the table means only one thing and that it contains only one value. Consider the following data is in zero normal form:

Field name	*Case Description*	*Responsible people*	*Attached Document*	*Effective date*
Data value	Project no. CHK001, redesign the substructure.	MK Chan	RFI/01/VVK001.dwg	From 1997/02/02 to 1997/06/20

This data is not in normal form because it violates the rules of first normal form. At the column Attached Document, "RFI/VVK001.dwg." This column contains two meanings: file type and filename. Because they are in the same column, it is impossible for the computer to locate a drawing file because the computer considers the table cell to be a single value (i.e., *filename.dwg*, not drawings file). The following table, corrects these problems and is in first normal form:

Case name	*Case Description*	*Responsible people*	*Document Type*	*Document ID*	*Date from*	*Date to*
CHK001/CI001	Redesign of block 1 to block 10	MK Chan	Drawing	RFI/01/VVK001	1997/02/02	1997/06/20

	substructure					
CHK031/CI002	Maintain FM200 panel	MK Chan	Drawing	RFI/01/VVK002	1997/06/02	1997/06/30
CHK031/CI002	Enhancing superstructure	MK Chan	Document	RFI/01/VVK001	1997/06/02	1997/06/22

Second Normal Form

The objective of second normal form is to eliminate redundant data. Significant problems associated with inconsistent data arise when data is repeated inside a computer. For instance, if *Case Name: CHK031/CI002* were to change people from MK Chan to KW Lee, 2 records require updating rather than just one. Inconsistencies would result, if not all records are updated.

Second normal form is accomplished by eliminating partial dependencies. Because the primary keys can be composed of multiple fields, it is sometimes possible for data to depend on only a part of the primary key for its access. This condition violates second normal form. Data fields that depend on only a portion of the primary key must be put in separate tables.

3.4.1 Architecture of the Client Server Module

The development language and computer platform used in this study are listed as follows:

- Development Language: Visual Basic 6, and Visual Interdev 6
- Deployment of dynamic link library DLL: ActiveX COM+ Component Object Model
- Database Engine: Microsoft SQL Server 7.0
- Development Platform: Microsoft Windows NT 4.0 Workstation with Service Pack 6

Figure 3.5, is the screen dump of the client server version of Critical Issue browser. The upper part of the screen organizes the Critical Issue as root of the tree structure, and the sub tree is used to represent the category of the case. The supportive document and design drawings are listed at right hand side when category is selected and displayed as open cabinet at left side.

The lower part of the screen provides the properties of highlighted object selected in critical issue.

- In the first tab, it preview the file contents, for example, a file *tested4.dwg* is selected, it will displays the drawings preview, project name in which the drawings belongs to, file creation date, revision or version of drawings file, etc.
- The second tab displays the file history, usually in construction project when a design change is requested by site instruction (SI), the CAD drawings file need to be modified as condition listed in SI. It represented by the revision form A to Z, This function facilitate the history tracking upon design variation.
- The third tab lists the reference materials of the object selected in category area,

for example, a list of revised site instruction is displayed whose contains information and file related to that critical issue.

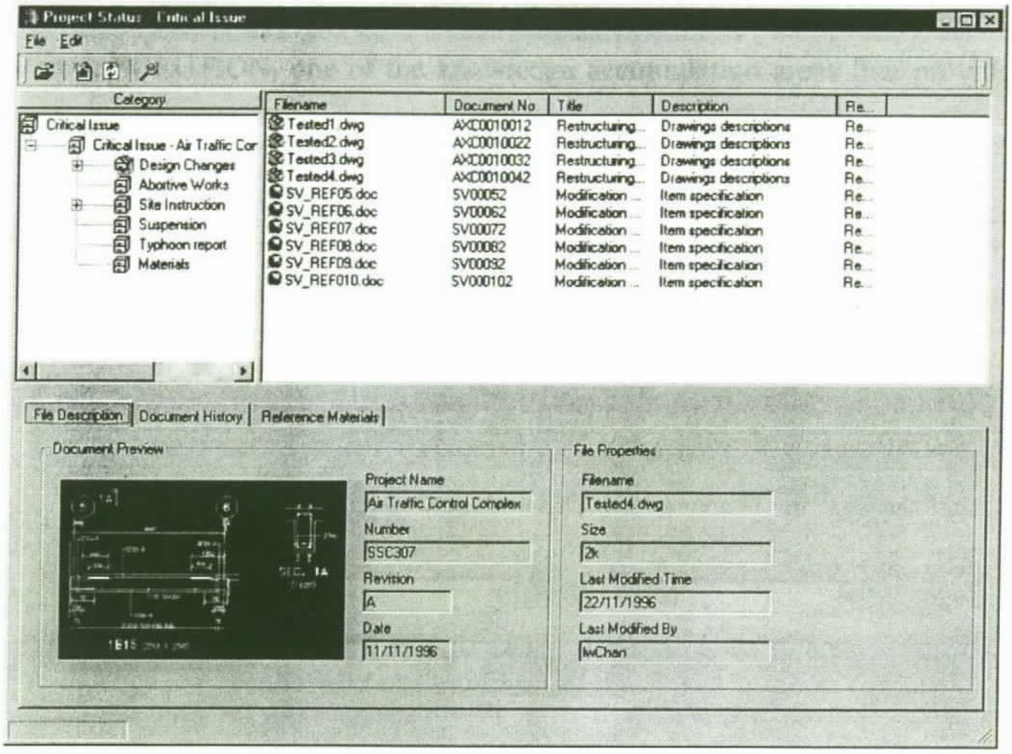

Figure 3.5 Client Server version of Critical Issue structure browser

By filling informational items in the e-forms shown in Figure 3.6, a case represented by critical issue is established and stored in the case base which immediately becomes available for retrieval and re-use by other users. As shown in the Figure, there are four tabs represent four components that compose of a single critical issue.

Part 1. ISSUE, it describes the background of the event. User can attach site instruction, safety report and other project documentation that as an evidence describing the event.

Part 2: CAUSE, it provides the reason why the event happened. For example, human cause, design cause and environmental cause, etc.

Part 3: AFFECT, it describes area affected by the events, including additional incurred cost, time cost, quality affected to final product, etc.

Part 4: SOLUTION, one of the knowledge accumulation areas that provide domain expertise to contribute their experience or actual practice about how to solve the case. This part may leave to blank when the project is still under construction or waiting to find out the solution.

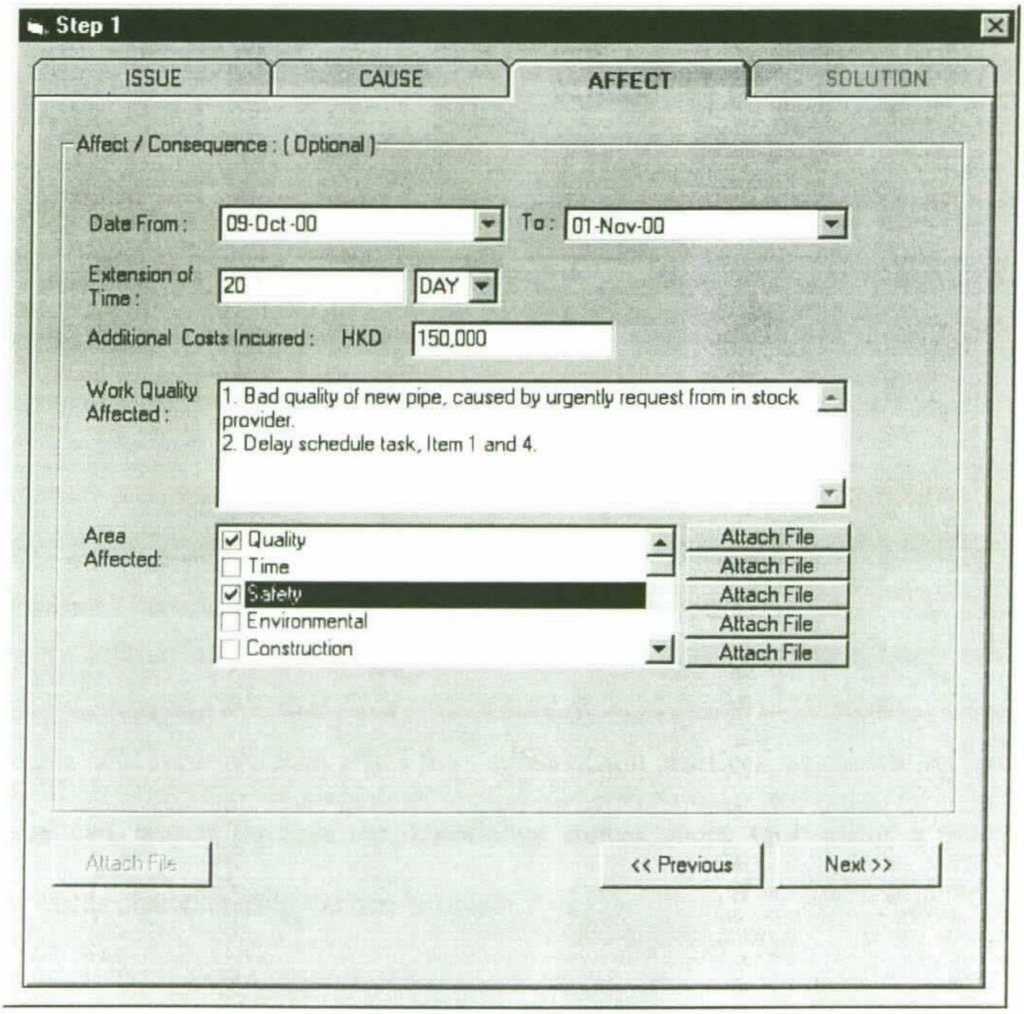

Figure 3.6 Components of Critical Issue

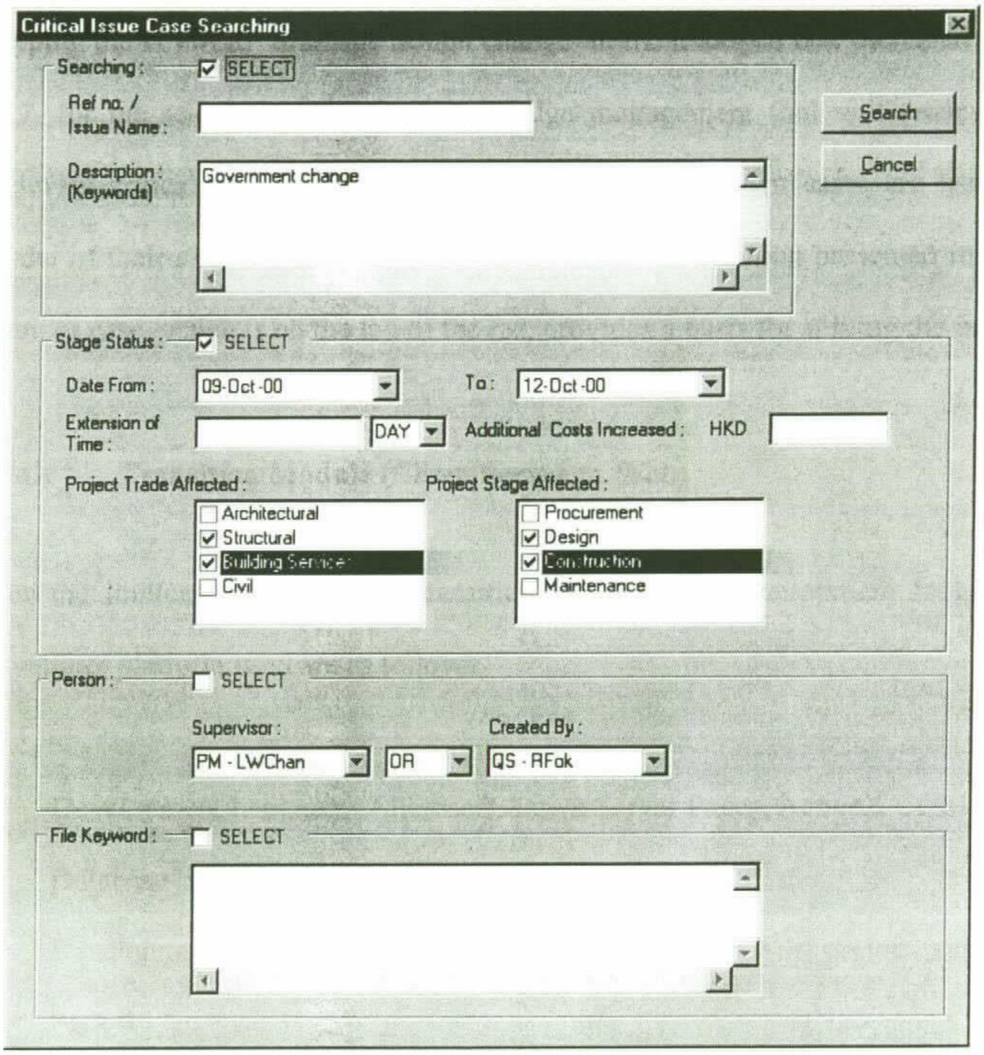

Figure 3.7 Dialogue box for searching (change of drainage design)

Once a new event/problem arises from construction practices, as shown in Figure 3.7, users can search through the knowledge management tool using a number of keywords characterizing the new problem.

For example, in a public housing project commissioned by the government, the client requested to alter part of the drainage design.

Typing the keyword 'drainage design change' in the dialogue box indicated in Figure 3.7, for the search function, the knowledge management tool will retrieve all the relevant critical issues stored in the case base. The relevant cases are listed in the order of their similarities to the problem at hand. The solution presented in the most similar case, which is on the top of the list, provides a basis for solving the problem.

3.4.2 Transition Module (Client Server to Web)

For the implementation of the transition module, the development language and computer platform used are as follows:

- Development Language: Microsoft Active Server Pages, ActiveX COM objects
- Database Engine: Microsoft SQL Server 7.0
- Development Platform: Microsoft Windows NT 4.0 Workstation with Service Pack 6
- Web Server: Microsoft Internet Information Server 4.0
- Process Transaction Control: Microsoft Transaction Server 2.0

3.4.3 Architecture of Internet Module

For the implementation of the Internet module, the development language and computer platform used are as follows:

- Development Language, Server Side:
    - Enterprise Java Beans
    - Java Component Object Model
    - Java Server Pages 1.2 – JDK 1.2.2
    - Java Servlet 2.2
    - Hyper Text Mark-up Language HTML
- Development Language, Client Side:
    - Java Script 1.3 with version compatible with Netscape 4.0 or Internet Explorer 4 or above.
- Server Side Scripting Engine: Jakarta-Tomcat V3.2
- Deployment Platform: Microsoft Windows 2000 Advance Server
- Web Server: Apache Server 1.3.14
- Running platform: Internet Explorer 4.0 and Netscape Navigator 4.07 or later.

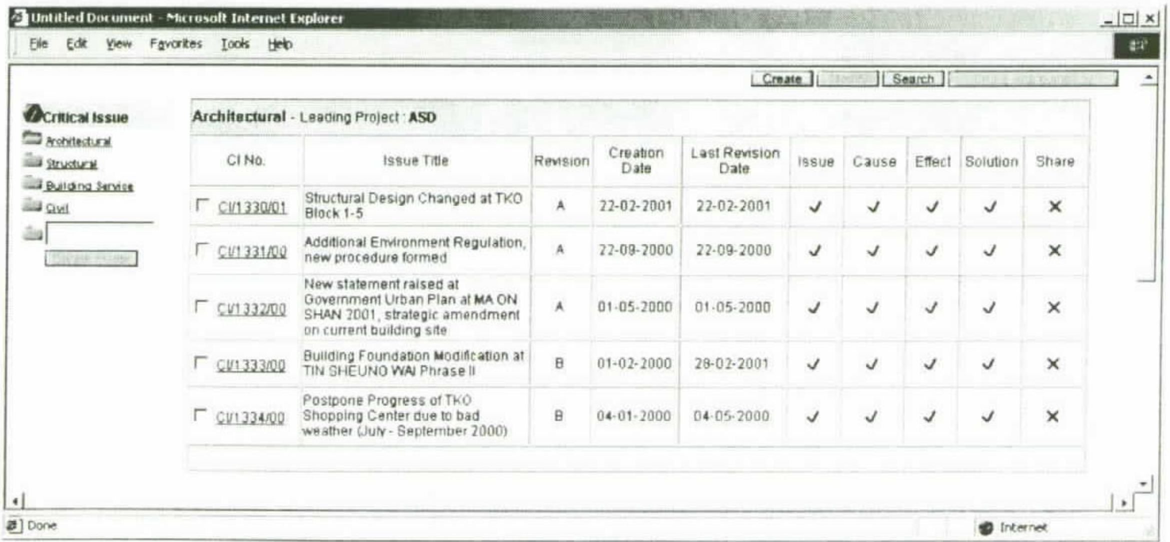

Figure 3.8 Web version of Critical Issue structure browser

Figure 3.8 demonstrates the structure browser of critical issue at web version. The cases represented as critical issue are listed and categorized in a drop down tree structure. It acts as knowledge management system as one of the module inside the VHBuild.com. VHBuild.com is a web-based project management system developed by the sponsor, Yau Lee Construction Ltd.

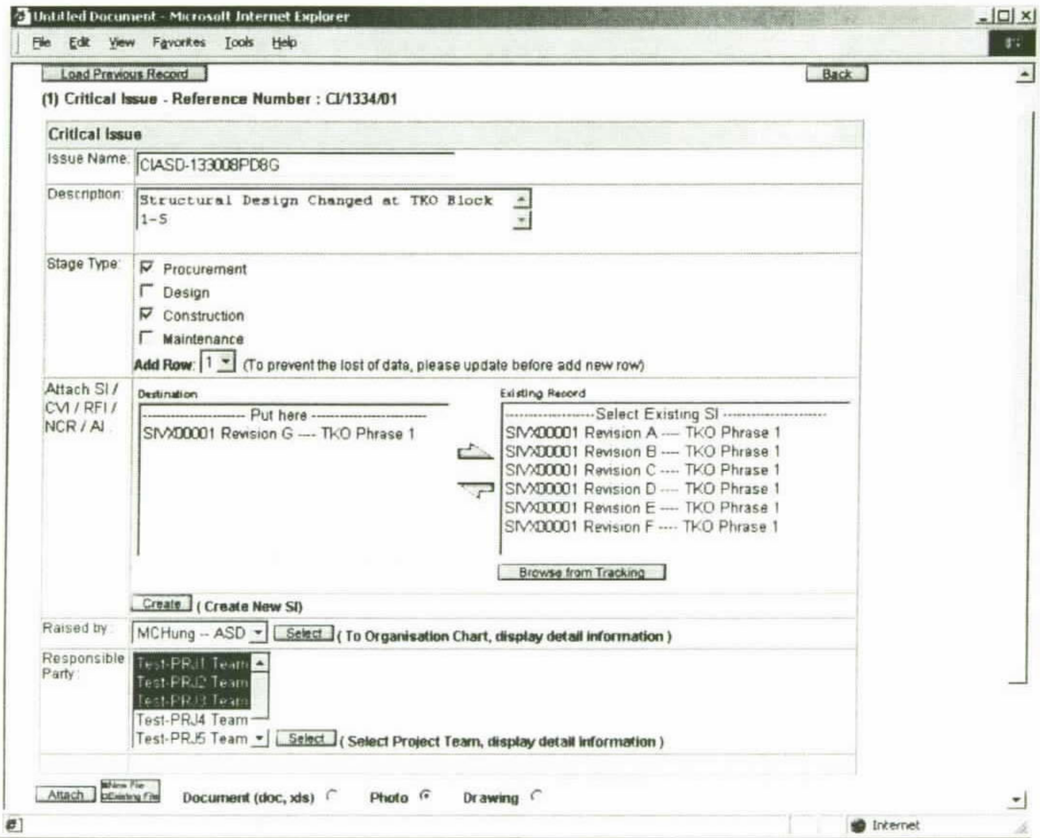

Figure 3.9 Web version of Critical Issue case descriptor

By attaching the descriptive construction E-Form as shown in Figure 3.9, which is optional and is captured from the construction procedure, for example, project daily report, safety walk report, Non Conformance Report (NCR), Request for Information

(RFI) provided by VHBuild.com, and followed by filling the informational items in the case descriptor, a critical issue case is established and stored in the case base which immediately becomes available for retrieval and re-use by other users.

CHAPTER 4

CONCLUSION

4.1 Contribution of the Research

In this study, we designed a Critical Issue knowledge management system, which consolidates construction project information in a case or issue bases and using the Case-Based Reasoning approach, analyses such information for re-use to solve problems in the future. Such a knowledge management system enables the buried knowledge inside cases to be obtained and applied. By maintains a record of project critical issue, it centralizes the project data in an efficient and retrievable manner. We have applied the system to a real business case study, VHBuild.com, a web based project management system for the construction industry. The system not only facilitates the retention of a clear history of events and documentation, but also provides a storybook of the construction project. There are four main components of Critical Issue: Issue, Cause, Effect and Solution. The Issue describes the background information of construction event; the Cause captures the reasons of the event and issues that formulates the problem; the Effect captures the duration and impact the issue has on the project operation; the Solution describes the solution provided by the domain expert and enables the case based reasoning system to identifies a new solution and adapt it for the new problem. These components utilize the electronic-form function in VHBuild, for example, "site instructions" and "request for information" are developed in an electronic format in VHBuild system that providing rich information of the project background and supporting documentation. Furthermore, with the linkage of the project organization chart, the responsible user of the Critical Issue can be clearly identified and activities tracked in any related project.

We have discovered that this system has the following advantages:

(1) Utilize case precedent to establish a knowledge library for experience sharing, learning, evaluation and analysis.

(2) Using past solutions to solve new problems, through two main operators of CBR, Select and Adapt:

  i. By using searching and matching of keywords to select previous case precedents based on similarities.

  ii. By identifying similarities and differences to adapt previous cases to address new problems.

(3) Facilitates a more flexible solution method compared to other rule-based system.

(4) Cross project data referencing, and eventually, industry integration.

(5) Can be applied at enterprise level, not just on single project basis, thus allowing knowledge to be shared by the entire company and potentially industry.

For us the knowledge base is primarily a domain specific, emergent outcome of the interaction among agents, both engineer and clerical workers. Part of this knowledge base may be our aggregations of the knowledge base properties of individual project personnel, and our structural analyses and measurements of the relationships among individuals with respect to the properties of their knowledge bases. With the efficient retrieval of past projects critical issue, many potential benefits, notably the training of new and inexperience engineers who can now draw on the experience contained in the previous cases can be derived.

The effectiveness of a CBR system performing knowledge management depends very much on the completeness of the case base. While an advantage of CBR systems over rule based systems is that CBR systems can be delivered with incomplete case bases, if a case base is under-populated, the system may not be able to deal with a lot of problems it may encounter. Therefore, it is believed that more cases for the construction project case base would lead to a more sophisticated knowledge management system.

4.2 Further Studies and Development

The study on knowledge management system pointed out that there are a few threads worth further investigation. For example, the refinement of feature weightings used in similarity assessment. The features used in the critical issue system are determined by trained and weighted process, and inevitably it will be based on intuition and judgment. More effective approaches, such as using statistical technique can be applied to analyze case data. The use of statistical technique to fine tune the case bases would produce a more reliable case bases. Thus, a more accurate similarity measurement would produce, based upon sound principles rather than intuition and judgment.

Statistical analysis of case data would also help to develop general adaptation heuristic for modifying previous cases to fit new cases. It is recognized though, that automatic adaptation is unlikely to be feasible for critical issue, because of the unique nature of projects, specific solution are hard to generalize. Meanwhile, the improvement of the speed of the CBR system is necessary to function as a professional system. Moreover, to organize a committee of construction expert in order to collect and evaluate domain specific problems and solutions would improve the accuracy of the result generated from the system.

REFERENCES

Althoff C., (1995) *A Review of Industrial Case-Based Reasoning Tools*, AI Intelligence, Oxford, Great Britain.
URL: http://wwwagr.informatik.uni-kl.de/~lsa/CBR/toolreview.html

Alfs T., (2000) *Knowledge and Workflow System*, IEEE transactions on knowledge and data engineering, pp. 1102-1116

Baluja, S. and Pomerleau, D. A. (1995). *Using the representation in a neural network's hidden layer for task-specific focus of attention.* In Proceedings of the International Joint Conference on Artificial Intelligence, IJCAI-95 Montreal, Canada, pp. 133-139

Basak, J., De, R.K. and Pal, S.K., (1998) *Unsupervised Feature Selection Using a Neuro-Fuzzy Approach*, Pattern Recognition Letters 19 (1998), pp998-1006

Bergmann, R., Wilke, W. & Schumacher J., (1997) *Using Software Process Modeling for Building a Case-Based Reasoning Methodology: Basic Approach and Case Study*, International Conference on Case-Based Reasoning, Springer-Verlag: ICCBR'97.

Bill Gates, Chairman of Microsoft, (2001) *Editor's Note* October 2001 issue of MSDN Magazine.

Cercone Nick, Aijun An, Christine Chan, (1999) *Rule-Induction and Case-Based Reasoning: Hybrid Architectures Appear Advantageous,* IEEE transactions on knowledge and data engineering, pp.1041-4347

Cercone Nick, Senior Member, Aijun An, Christine Chan, (1999) *Rule-Induction and Case-Based Reasoning: Hybrid Architectures Appear Advantageous,* IEEE transactions on knowledge and data engineering pp.1041-4347

Daniel E, Peter S, (1999) *Knowledge Management for Best Practices*, Don Bishop. LLC. Winter 1999. intelligence, pp.13-23

Dzeng R.J., Chen W.T., (1997) *Learning Construction-Planning Knowledge from Experience*, IEEE Proceedings of the 1997 IASTED International Conference on Intelligent Information Systems, pp. 321-325

Eiter T., Lukasiewicz T., (2000) *Default reasoning from conditional knowledge bases: Complexity and tractable cases*, Elsevier Science Artificial Intelligence 124 (2000), pp.169-241

Eyke Hüllermeier, (2001) *Similarity-based inference as evidential reasoning,* International Journal of Approximate Reasoning 26 (2001), pp.67-100

Fan X., Sheng F., Ng P.A., (1998) *DOCPROS: A Knowledge-Based Personal*

Document Management System, IEEE Proceedings of the 10th International Workshop on Database & Expert Systems Applications, pp.527-531

Graham W, (2000) *Knowledge Management and Competitive Manufacturing*, Engineering Management Jornal, June 2000, pp.130-134

Hage J., (1993) *Monological Reason-Based Logic, A Low Level Integration of Rule-based Reasoning and Case-based Reasoning,* pp30-39, ACM

Halpin W. Daniel, Wookhead W. Ronal, (1998) *Construction Technology and Construction Management*, John Wiley & Son Inc., 1998

HKG, (2001) *Chemical Waste Control, Code of Practice on the Handling, Transportation and disposal of Asbestos Waste*, Hong Kong Government, Part 7. Storage of asbestos waste

Holt, G. D., (1996) *Applying cluster analysis to construction contractor classification*, Building and Environment 31 (6) (1996), pp. 557-568

Holt G. D., Olomolaiye P. O., Harris F. C., (1994) *Evaluating performance potential in the selection of construction contractors*, Engineering, Construction and Architectural Management, pp. 29-50

Ivo V, Wolfgang W, Ralph B, (1998) *Intelligent Sales Support on the Web*, Proceedings of the 6th German Workshop on Case-Based Reasoning (GWCBR'98), pp.1-18

Jiang M.F.and Tseng S.S., (2000) *Structural Cases Representation in Case-based Expert Systems (CBES),* 4th International Conference on knowledge-based Intelligent Engineering System & Allied Technologies, pp. 432-435

John B, (2001) *Defining Knowledge: an Epistemological Foundation for Knowledge Management*, 2001 Institute of Electrical and Electronics Engineers, pp.1-7

Jorma Ollila, (2000) *Most admired knowledge enterprise*, The Know Network, URL: http://www.knowledgebusiness.com

Joyee L, Rarnes S, Mark W, (2000) *Coordinating Technologies for Knowledge Management in Virtual Organizations*, 2000 Institute of Electrical and Electronics Engineers, pp.21-26

Jurisica, I. and Glasgow J, (1996) *A case-based reasoning approach to learning control,* In 5th International Conference on Data and Knowledge Systems for Manufacturing and Engineering, DKSME-96, Phoenix, Arizona.

Kawamura M. and Kawamura T., (1994) *Parallel Database Management System: Kappa*, Proc. FGCS'94, ICOT, Tokyo, Dec.1994.

Kamp, G. (1996) *Using Description Logics for Knowledge Intensive Case-based Reasoning,* ISBN 3-540-63233-6, pp. 223-234

Kitano, H. and Shimazu, H., (1996) *The Experience-Sharing Architecture: A Case Study in Corporate-Wide Case-Based Software Quality Control,* ISBN 0-262-62110-X, pp. 420

KMM, (Oct 2001) *Knownledge Management Magazine*, 2001 Bizmedia Ltd.

Kolodner Janet, (1993) *Case-based reasoning,* Morgan Kaufmann, ISBN 1-55860-237-2

Kuo R.J., Chen C.H., Hwang Y.C., (2001) *An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network,* Elsevier Science Fuzzy Sets and System 118 (2001), pp.21-45

Lambrix P., Shahmehri N. and Aberg J., (1998) *Towards Creating a Knowledge Base for World-Wide Web Documents*, IEEE Proceedings of the IASTED International Conference on Intelligent Information, pp. 507-511

Lan Watson, (1997) *Applying Case-based Reasoning: Techniques for Enterprise Systems*, University of Salford, U.K. Morgan Kaufmann Publishers, Inc

Leake d. b., (1996) *Case-Based Reasoning: Experiences, Lessons, and Future Directions*. AAAI Press, Menlo Park, CA. ISBN 0-262-62110-X

Lenz M. and Burkhard H., (1996) *Case Retreival Nets: Basic Ideas and Extensions,* in KI-98:Advances in Artificaial Intelligence, Springer-Verlag, Berlin, 1996

Li, H.(1996) *Case Based Reasoning for Intelligent Support of Construction Negotiation.* Information and Management Journal. 30, pp. 231-238.

Li, H (2000) *VHBuild: Transferring Project Management into Knowledge Management.*

Marcel H., Kai U.L, Thomas W, (1998) *A Design Process for Embedding Knowledge Management in EveryDay Work,* IEEE transactions on knowledge and data engineering, pp. 296-305

Markus R, Klaus T. (2000) *Business Models for Coordinating Next Generation Enterprises,* IEEE Proceedings of the Academic/Industry Working Conference on Research Challenges, pp.163-168

Menzies T., Cukic B., (1998) *On the Sufficiency of Limited Testing for Knowledge Based Systems*, Proceedings of the 11th IEEE International Conference on Tools with Artificial Intelligence, pp.431-440

Mitchell T., (1997) *Decision Tree Learning,* in T. Mitchell, Machine Learning, The McGraw-Hill Companies, Inc., pp. 52-78.

Moore D.R. & Haque D.J., (1998) *Building Production Management Techniques, Longman & The Chartered Institute of Building*

Nguyen V. U. (1985) *Tender evaluation by fuzzy sets,* Journal of Construction Engineering and Management, ASCE 111, pp. 231-243

Nicholas R., (1997) *The Obligations of Impersonal Reason;* University of Notre Dame Press, 1997.

Oehler, P., Vollrath, I., Conradi, P., Bergmann, R., Wahlmann, T. (1998) *READEE - Decision Support for IP Selection using a Knowledge-Based Approach*, Proceedings of the 3rd World Conference on Soft Computing: Advances in Soft Computing - Engineering Design and Manufacturing, Springer

Pflaging J., (2001) *Collaborative commerce – Everyone's doing it but does anyone understand it ?,* The journal of work process improvement, October 2001, pp.18-21

Ralph B, Gerd P, Wolfgang W, (1993) *Explanation-based Similarity for Case Retreival and Adaptation and its Application to Diagnosis and Planning Tasks,* Proceedings of the 1st European Workshop on Case-based Reasoning (Preprints) University of Kaiserslautern, pp.1-17

Ralph B, Wolfgang W, Ivo V, Stefan W, (2000) *Integrating General Knowledge with Object-Oriented Case Representation and Reasoning*, Göker, M. H.: Proceedings of the 8th German Workshop on Case-Based Reasoning, GWCBR 2000, Lämmerbuckel, DaimlerChrysler Research and Technology, March 2-3

Russell, J. S., (1990) *Contractor prequalification model*, Journal of Computing in Civil Engineering, ASCE 4 (1), pp.77-90

Russell J. S., Skibniewski M.J., (1990) *Qualifier 2: knowledge-based system for contractor prequalification*, Journal of Construction Engineering and Management, ASCE 116 March, pp.157-171

Schmitt S., Maximini R., Landeck G, Hohwiller J., Blanzieri L. Portinale J., (2000) *Advances in Case-Based Reasoning.* Proceedings of the 5th European Workshop on Case-Based Reasoning, EWCBR2k, in Trento (Italien), LNAI 1898, Springer, pp. 6-9.

Smyth, B and McKenna E. (1997), *Modeling the Competence of Case-Bases*, in Proceedings of 4th European Workshop, EXCBR-98, pp.207-220

Stahl, A. & Bergmann, R. (2000) *Applying recursive CBR for the customization of structured products in an electronic shop*, 5th European Workshop on Case-Based Reason-ing, Springer Verlag. pp.1-12

Stephenson, John, (1995) *Building regulations explained: 1995* revision, London ; New York : E & FN Spon, c1995. Edition 5th ed.

Thomas H. Davenport, (1999) *Working Knowledge: How Organizations Manage What They Know*, Bus Sch. Press, pp.5

Thomas Ng,(2000) *EQUAL: a case-based contractor prequalifier*, Automation in Construction (10), pp.443-453

Trent J, Atul P, (1995) *Management and Utilization of Knowledge for the Automatic Improvement of Workflow Performance*, Association for Computing Machinery, pp.32-43

Ulrich F., Zeyno A., (2001) *A hybrid representation of architectural precedents,* Automation in Construction (10), pp. 687-699

Vollrath I. , Wilke W. and Bergmann R., (1998) *Case-Based Reasoning support for online catalog sales,* IEEE Internet Computing, Vol. 2, No. 4, 47-54

Wess S., (1995) *Using kd-Trees to Improve the Retrieval Step in Case-based Reasoning.* pp.167-181.

Wilke W. and Bergmann R., (1998) *Techniques and Knowledge Used for Adaptation During Case-Based Problem Solving,* in Tasks and Methods in Applied Artificial

Intelligence, Springer-Verlag, Berlin

Wilke W., Cunningham P., and B.Smith, (1998) *Using Configuration Techniques for Adaptation,* in Case-Based Reasoning Technology, Springer-Verlag, Berlin

Winston P., (1992) *Learning by Building Identification Trees, in P. Winston, Artificial Intelligence,* Addison-Wesley Publishing Company, pp. 423-442.