Case-Based Reasoning Integrating with Direct-Case-Linkage for Tacit Knowledge Management[1]

L. Lin[a], H. Ling[a], J. Lu[b], C. Zhang[a], L. Song[a], H. Xue[a]

[a]Fudan University, Shanghai, 200433, China
lmlin@fudan.edu.cn

[b]Faculty of IT, University of Technology, Sydney, P.O. BOX 123, Broadway, 2007, Australia
jielu@it.uts.edu.au

Abstract

Tacit knowledge is the source of knowledge creation, but the organization and retrieval of tacit knowledge using IT has not been discussed much. In this paper, we defined Internal Knowledge Evolution Network (IKEN) to describe the knowledge evolution process from tacit to explicit, and applied the Case-Based Reasoning (CBR) integrating with Direct-Case-Linkage technique to support the IKEN management. To argue that the technique can better support the retrieval of tacit knowledge, we first implemented a prototype system, and then evaluated the system by carrying out a performance test.

Keywords

Case-Based Reasoning, Tacit Knowledge Management

1 Introduction

The knowledge that owned by an organization can be classified into two categories. One is the knowledge that is formally documented and widely accepted, such as the reports and publicized papers. Nonaka (1991) defined this category of knowledge as 'explicit knowledge'. Another category is the knowledge that is applied when producing the explicit knowledge, such as the experience owned by experts. Nonaka defined this category as 'tacit knowledge', and claimed that tacit knowledge is the true source of knowledge creation. Nonaka's definition is not in accordance with the original definition proposed by Polanyi (1966) who stated that all knowledge having a tacit dimension that cannot be separated from explicit knowledge. Nonaka's definition has already been accepted by researchers on knowledge management and this paper takes Nonaka's definition.

[1] In support by NSFC No. 70201009

There are literatures discussing about the management of explicit knowledge (Ralf 2003, Maedche 2003) and tacit knowledge (Roy 2001, Armbrecht 2001, Noh 2000, Stenmark 2002, Shimazu 1999). About applying IT for tacit knowledge management, Noh, etal. (2000) discussed about representing tacit knowledge using cognitive map and applying Case-Based Reasoning technique to implement the management. Stenmark (2002) discussed about the application of Internet/Intranet for the communication of tacit knowledge. But for the organization and retrieval of tacit knowledge using IT, it has not been addressed before.

This paper defined Internal Knowledge Evolution Network (IKEN) to describe the knowledge evolution process from tacit to explicit, and applied the Case-Based Reasoning (CBR) integrating with Direct-Case-Linkage technique to support the IKEN management. To argue that the technique can better support the retrieval of tacit knowledge, we first implemented a prototype system, and then evaluated the system by carrying out a performance test.

Section 2 introduces the IKEN concept. Section 3 motivates the CBR technique integrating with Direct-Case-Linkage. Section 4 describes the implementation of the technique. Section 5 presents a performance evaluation.

2 Internal Knowledge Evolution Network

We define the **Internal Knowledge Evolution Network (IKEN)** as a directed network of knowledge fragments collected/created by members of the enterprise, see Fig.1, where along the direction of the arrows knowledge fragments are created based on previous knowledge fragments.

It is obvious that with IKEN any sub segment of the tacit-to-explicit process can be retrieved by tracing down the directed route.

For instance, for an university research group which has an 18-year history on Decision Support System, has 10 National Projects finished and over 50 papers and books published, her potential in competition lies on the research knowledge and experience of the members, which are tacit knowledge. The tacit knowledge are far richer than the collection of the publications which are explicit knowledge that can be retrieved anywhere. The tacit knowledge should include the internal knowledge created and used during the process of publication composing.

A typical query problem for tacit knowledge could be one of the followings,

There is a working paper in the group titled 'Case-based Reasoning Integrating with Direct-Case-Linkage for Tacit Knowledge Management'. A student read it and is very interest the idea of the application of CBR, she wonder how this idea came into being and if it is possible for other application.

It is a question that can not be answered without understanding the CBR researching process of the group on previous applications. There are at least two ways for the student to identify the answer. One is to collect papers by 'hard force' -looking for all relative information such as the references, the papers by the same author(s), etc.. The second way is to interview with the author for direct explanation. The first one is an old fashion one but too rough to be efficient, and the second one is too ideal to be always true.

If there is a system support the IKEN of the research group, then the student has a third way to find the answer – tracing backward down the route she can find all application research papers and all the intermediate results – even the failed but creative ideas.

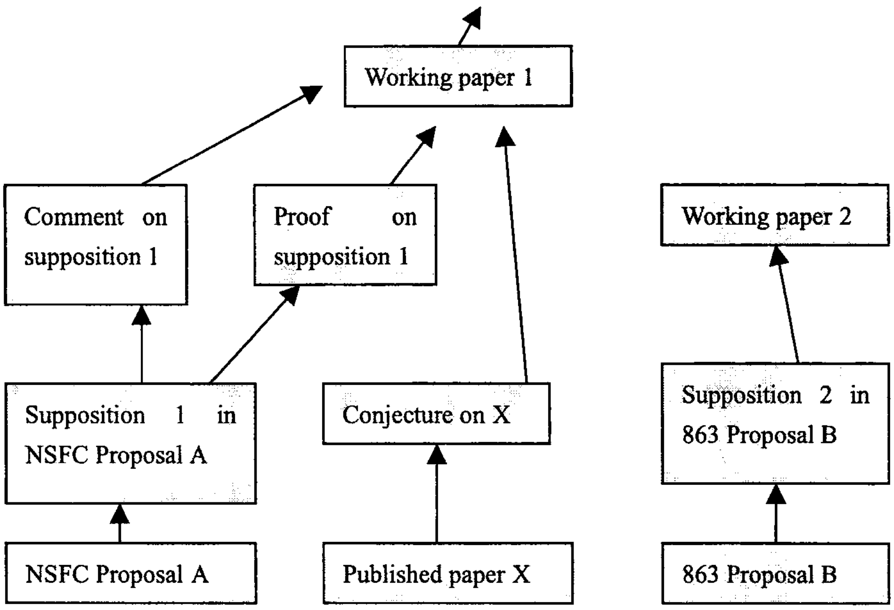

Fig.1 Internal Knowledge Evolution Network

3. Case-Based Reasoning with Direct-Case-Linkage

Case-Based Reasoning (CBR) is a decision support technique in AI with 'analogy' as its cognitive basis. When solving new problems, CBR retrieves the past problems by analogy to find the most similar case, adapts the solution of the most similar case according to the new problem. The new problem and its solution are stored as new case in the case base for future reuse. The CBR process is often described as a recycle of '4Rs', Retrieve→Reuse→Revise→Retain→Retrieve → ...(Ian,1997).

By applying CBR knowledge fragments can be stored as cases, and can be retrieved using the typical Nearest Neighbor Algorithm. When Applying CBR on documents management, the major difference from other Decision Support System CBR systems is that for documents management, the result of retrieval is the final result, whereas for Decision Support purpose, the retrieve result might need revise before the solution can be used.

To support IKEN, the relationship between knowledge fragments need to be recorded. But CBR does not support this feature. In CBR, cases are related to each other only indirectly when they have similar feature values, and the similarity relationship can only be explored by

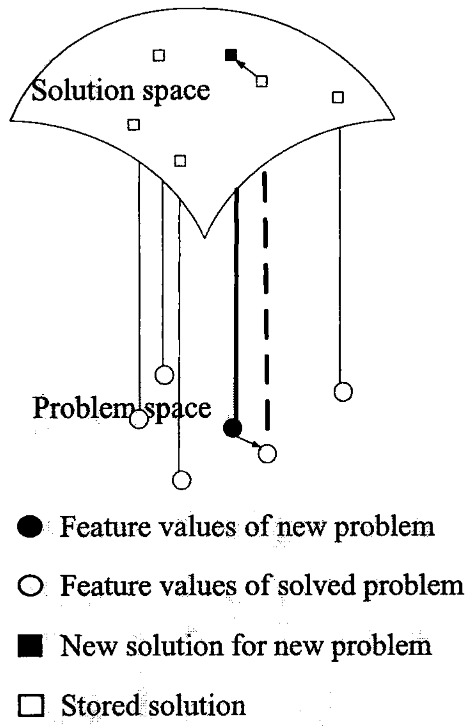

Fig.2 CBR explore the similar case by comparing the feature values

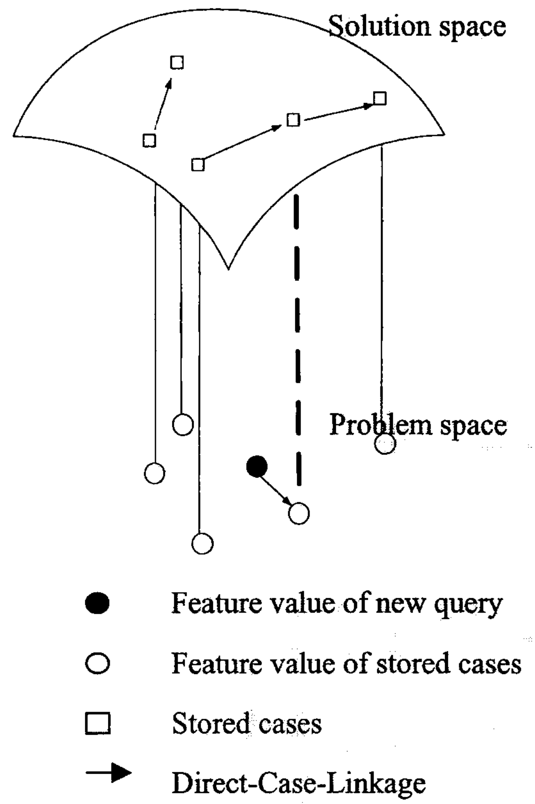

Fig.3 Direct-Case-Linkage in CBR

comparing the feature values. If in terms of the problem space and the solution space, see Fig. 2, '...Retrieval identifies the case with the most similar problem features, and its stored solution is found....'(Ian,1997).

See Fig.3, we build direct linkage between cases having context sequential order.

When query tacit knowledge, the exist information are feature values of a concrete case. Using the same example as in section 2, the information that the student knows about is there is a working paper focusing on the application of CBR to tacit knowledge management, and that the application of CBR has been a topic for the group for some years. CBR can retrieve the case by feature-value comparison. This case is very important since it is the start point for 'dragging' out the following knowledge fragments along the creation process.

From the start point, and with the construction of Direct-Case-Linkage, the trace of knowledge creation process can be 'redrawed' by following the forward /backward direction. The student may find out that the research group used to apply CBR on State-Enterprise Reform Decision Support System, and their research results on the knowledge representation were further extended to apply on tacit knowledge management study.

4. Case Structure and the Direct Linkage between cases

Section 3 introduced that by constructing Direct-Case-Linkage on knowledge cases the temporal sequence of knowledge fragments of IKEN can be tracked down. In this section we explain how to implement the technique.

Each knowledge fragment is stored as a case in case base and two pieces of information are stored for each case. One is the original document, the other is a set of feature values extracted from the document for each knowledge fragment. For example, a working paper titled 'Case-Based Reasoning integrating with Direct-Case-Linkage' is stored as a case in case base, with the original word file 'LINLiminIECB2003.doc' saved in a specified directory, and a set of feature values saved in the database which include:

Material ID = 1 (auto number by database)
Material Type ID = 7 (which means the knowledge fragment is a working paper)
Title = Case-Based Reasoning integrating with Direct-Case-Linkage
Author = LIN Limin
KeyWords = Case-Based Reasoning, Tacit Knowledge Management
PubDate = N/A
Extract Date= 2003-2-21
Extractor ID = 123
Publish Information = ICEB2003
Context = N/A
FileTypeID = 3 (which means the file is .doc format)
FileName = LINLiminIECB2003.doc
......

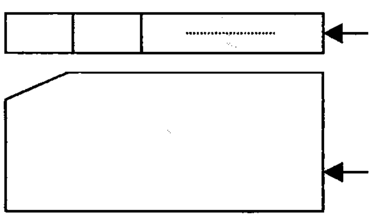

Fig.4 Case representation for knowledge fragment

Fig.4 is a illustration of the case representation of single knowledge fragment. The Internal Knowledge Evolution Network can be illustrated in Fig.5, in which the directed edge between cases shows the reference relationship between knowledge fragments. The referring knowledge fragment is produced earlier than the referred knowledge fragment, and the directed edge starts from the referring case and points to the referred cased.

The set of directed edge constructs the direct-linkage between cases, which makes it possible

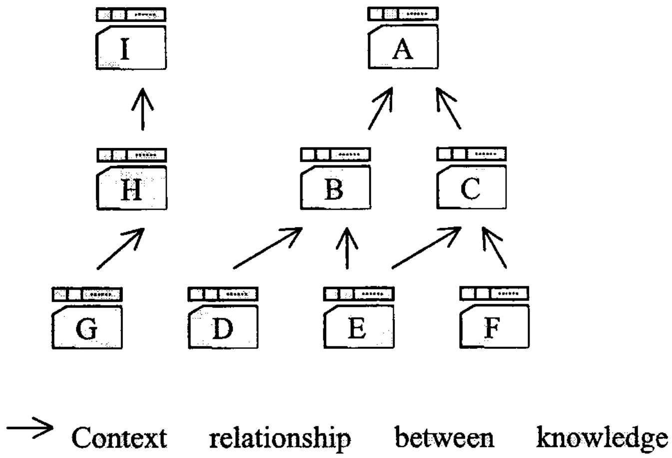

Fig.5 Internal Knowledge Evolution Network

that the sequence of knowledge fragments can be retrieved forward or backward along the link route.

The direct-linkage can be stored in a database table shown in Table 1. In the first column are the referring fragments, and in the second column are the referred fragments, and in the third column are the reference content.

Referring fragment	Referred fragment	Reference content
A	B	A uses the XX of B
A	C	A refutes the XX of B
B	D	B applies the X method of D on XX
B	E	……
C	E	……
C	F	……
…..	……	…..

Table 1 database table recording the linkage between cases

5. Evaluating Direct-Case-Linkage CBR

This section describes an experimental study designed to evaluates the performance cof Direct-Case-Linkage CBR on the retrieving of both explicit and tacit knowledge, where the explicit knowledge fragments are a set of documents with various query conditions, and the tacit knowledge are several sequences of temporal knowledge fragments with various query conditions.

For comparison, we choose another system – Windows Resource Manager (WRM) to implement the same experiment. The reason for choosing WRM is that the commercial Knowledge Management Software (KMS) such as the K-Station by IBM-Lotus and the Sharepoint by Microsoft are not yet widely used, and the most often used system is File Directory System which is accessory to each popular Operating System such as Windows, Unix, and Macintosh etc. and among which the most popular on is Windows.

In the experiment, we will ask the volunteers to finish the proposed retrieval questions by operating on both systems, and calculate the rate of error and the average time.

5.1 Experiment Data Preparation

Let N represent the number of documents. When n is small, for example N<10, the performance of the two systems might not be drastically different from each other, and when n increases the performance difference might increase either. To confirm this conjecture, we design three sets of documents with N=10, N=100, and N=500.

The structure of the two systems are different and the complexity of the structure might have influence on the performance. To avoid, we set the major structure parameters of both systems to be similar with each other. See table 2.

		KMS		WRM		
		The number of projects	The max number of documents in each project	The number of directories	The max depth of directories	The max number of documents in each directory
The total number of documents	N=10	2	7	2	3	7
	N=100	1	100	1	3	100
	N=500	4	180	4	3	180

Table 2 The complexities of both systems are similar with each other

5.2 Query Questions

When preparing the query questions, we follow the following rules,

- Both tacit knowledge and explicit knowledge query questions are needed.
- The query questions may not have results to avoid the tester use traversing retrieval.
- For explicit knowledge query, the number of questions is set with different number of conditions and different number of resulting documents. See Table 3.a.
- For tacit knowledge query, the number of questions is set with different sequential directions and different number of resulting documents. See Table 3.b.

		The number of condition for each query	
		1	>1
The number of resulting documents	0	1	1
	1	1	1
	>1	1	1

Table 3.a The number of testing questions for explicit knowledge query

		The retrieving direction	
		forward	Backward
The number of resulting documents	0	1	1
	>1	1	1

Table 3.b The number of testing questions for tacit knowledge query

5.3 Results

Volunteers are college students who participated in the coding of the KMS system. Therefore they are familiar with both KMS and WRM usage. The number of tester is 9.

The error rate and average time for KMS and WRM with different number of documents are listed in Table 4. Fig. 6a illustrates the error rate for explicit query, Fig.6b illustrates the average time for explicit query. Fig.7a illustrates the error rate for tacit query, and Fig.7b illustrates the average time for tacit query.

	[illegible]											
	Error rate			Average time (second)			Error rate			Average time(second)		
N	Explicit query	Tacit query	average	Explicit query	Tacit query	average	Explicit query	Tacit query	average	Explicit query	Tacit query	average
10	0.167	0.389	0.256	18.722	25.528	21.444	0.352	0.722	0.467	44.722	61.083	51.267
100	0.426	0.722	0.544	32.630	30.944	31.956	0.593	0.833	0.689	51.222	50.750	51.033
500	0.185	0.111	0.156	21.056	22.667	21.700	0.574	0.556	0.567	41.546	62.347	49.867

Table 4 The Error rate and Average time for KMS and WRM with different number of documents

5.4 Analysis

From the figures we can see that KMS has overall performance better than WRM.

There are still several phenomena in the figures need to be analyzed.

One is that the performance of both systems for N=100 is the worst, see Fig.6a, Fig.6b, Fig.7a. The reason could be that the sample documents for N=100 are grouped into one directory/project, please refer to the figures in Table 4. The complexity increases if the queries happen to be in the directory/project of too much documents.

The second phenomenon is that the average time for explicit query when n=10 is longer than when n=500, see Fig.6b. The reason could be that with the fewer documents, testers were confident that the traverse could work without realizing its time consuming.

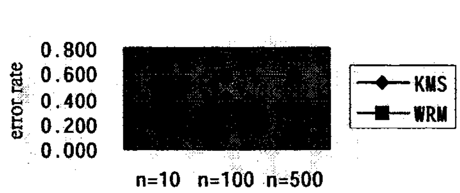

Fig. 6a Error rates for explicit query

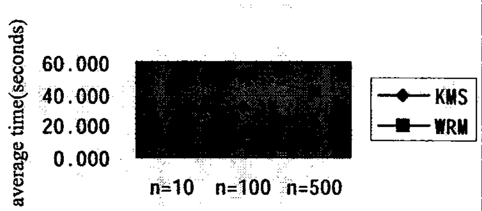

Fig. 6b Average time for explicit query

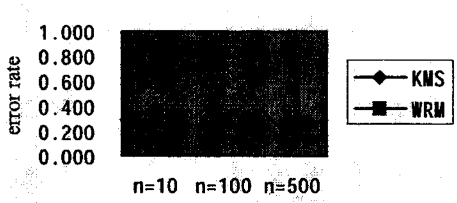

Fig.7a Error rates for tacit query

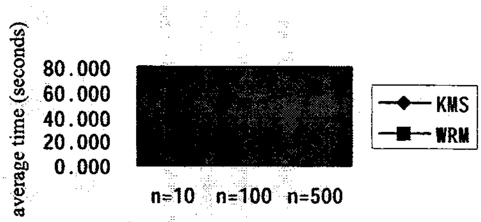

Fig. 7b Average time for tacit query

The third phenomena is that when n=100, the error rate for tacit query is the highest, but the average time is the lowest, see Fig.7a and Fig.7b. The reason could be that for WRM, if the documents are stored without any obvious information on the file name, the tacit query where the testers need to identify the sequence of related documents could be too much difficult. Therefore they tended to give up traversing the documents one by one. The error rate went up but the average time shortened.

6 Conclusion

In this paper, we addressed the problem of tacit knowledge management. The knowledge transformation process from tacit to explicit was described using Internal Knowledge Evolution Network, and the IKEN could be supported by Case-Based Reasoning technique integrating with Direct-Case-Linkage. A prototype system KMS was implemented and the performance evaluation was carried out.

References

Ada (2001), Enabling Knowledge Creation: How to unlock the mystery of tacit knowledge and release the power of innovation, *The Academy of Management Executive.*

F M Ross Armbrecht Jr; Richard B Chapas; Cecil C Chappelow; George F Farris; et al (2001), Knowledge management in research and development, *Research Technology Management*, Washington.

Hideo Shimazu (1999), Translation of Tacit Knowledge into Explicit Knowledge: Analyses of Recorded Conversations between Customers and Human Agents, *proceedings of the 1999 International Conference on Case-Based Reasoning: Practical Case-Based Reasoning Strategies for Building and Maintaining Corporate Memories.*

Ian Watson (1997), *Applying Case-Based Reasoning: Techniques for Enterprise Systems.* Morgan Kaufmann Publishers, Inc..

J. B. Noh, K. C. Lee, J. K. Kim, J. K, Lee, S. H. Kim (2000), A case-based reasoning approach to cognitive map-driven tacit knowledge management, *Expert Systems with Applications,*Vol. 19,pp. 249-259.

K. Sycara &D. Navinchandra (1991), "Influences: A Thematic Abstraction for the Creative Reuse of Multiple Cases," *Proc. Case-Based Reasoning Workshop*, pp.133-144.

Maedche, A.; Motik, B.; Stojanovic, L.; Studer, R.; Volz, R.;(2003), Ontologies for enterprise knowledge management, *Intelligent Systems, IEEE* [see also IEEE Expert], Vol.18, No.2, pp. 26 -33.

M.A. Redmond (1990), "Distributed Cases for Case-Based Reasoning: Facilitating the Use of Multiple Cases," *Proc. Eighth International Conference of Artificial Intelligence*, pp.304-309.

M.L. Maher and D.M. Zhang (1993), "CADSYN: A Case-Based Design Process Model, " Artificial Intelligence for Eng. Design, Analysis, and Manufacturing, vol. 7, no.2, pp.97-110.

Polanyi, M. (1966), *The tacit dimension* , London: Toutledge and Kegan Paul.

Ralf Klischewski, (2003), Towards an ontology for e-document management in public administration - the case of Schleswig-Holstein, *Proceedings of the 36th Annual Hawaii International Conference on System Sciences*, pp.136 -144.

Roy L.(2001), Tacit knowledge and knowledge management: The keys to sustainable competitive advantage, *Organizational Dynamics;* New York.

Stenmark, D. (2002),"Information vs. Knowledge: The Role of intranets in Knowledge Management". In *Proceedings of HICSS-35,* IEEE Press, Hawaii, January 7-10.