



UNIVERSITY OF GOTHENBURG

Case Based Reasoning in Practice:
Investigation of The Practical Importance of Case Based Reasoning
Bachelor's Thesis in Software Engineering and Management

YUECHUAN WANG
YANG SHI

University of Gothenburg
Chalmers University of Technology
Department of Computer Science and Engineering
Göteborg, Sweden, 16-03-30

The Author grants to Chalmers University of Technology and University of Gothenburg the non-exclusive right to publish the Work electronically and in a non-commercial purpose make it accessible on the Internet.

The Author warrants that he/she is the author to the Work, and warrants that the Work does not contain text, pictures or other material that violates copyright law.

The Author shall, when transferring the rights of the Work to a third party (for example a publisher or a company), acknowledge the third party about this agreement. If the Author has signed a copyright agreement with a third party regarding the Work, the Author warrants hereby that he/she has obtained any necessary permission from this third party to let Chalmers University of Technology and University of Gothenburg store the Work electronically and make it accessible on the Internet.

Case Based Reasoning in Practice

Investigation of The Practical Importance of Case Based Reasoning

YUECHUAN WANG

YANG SHI

© YUECHUAN WANG, 16-03-30.

© YANG SHI, 16-03-30.

Examiner: Imed Hammouda

Supervisor: Anthony Anjorin

University of Gothenburg
Chalmers University of Technology
Department of Computer Science and Engineering
SE-412 96
Göteborg

Telephone + 46 (0)31-772 1000

Department of Computer Science and Engineering
Göteborg, Sweden, 16-03-30

Case Based Reasoning in Practice: Investigation of The Practical Importance of Case Based Reasoning

Abstract

Software products are taking and increasingly important part in everyday life and all kinds of industry. In order to provide a great environment for software development and increase the productivity and effectiveness, various types of approaches have been suggested. Case-based reasoning (CBR) has its potential to be one of best methods for knowledge management due to its apparent merits in the software industry. But there has not been enough empirical evidence showing to what extent CBR is used in practice and the challenges involving the difficulties of applying case-based reasoning, personal capacity and so on. In order to find out how case-based reasoning is applied in practice in current software development industry, we conduct a research, which applies literature review with interview study as its main research method. The results shows that due to the multiple difficulties of applying case-based reasoning, some companies choose to use other types of knowledge management techniques, and other companies have not applied CBR applications.

Key words: Case-based reasoning, knowledge management.

1. Introduction

The software industry is constantly making a difference in people's life. In order to meet the increasing requirements of users and to preserve their competitiveness in the market, companies apply various methods to increase their productivity and effectiveness. Case-based reasoning (CBR) has been growing fast and has attracted more attention in past few years. As the name suggests, CBR is the process of solving new problems based on the best practice of similar past problems [4]. CBR is frequently applied in the area of Artificial intelligence (AI). Some existing studies addressed that CBR is beneficial to apply in other kinds of software products [1]. The function of knowledge management (KM) is to allow an organization to leverage the information resources it has and to support purposeful activity with positive definable outcomes [2]. Ian Watson [7] pointed out that CBR could be considered a methodology for knowledge management.

Although CBR has shown its potentiality for being an effective method that can contribute to the software development industry, there exist some problems that increase the difficulty widely applying it. For instance, the case-based reasoner has to select a successful layout that closely matches the parts waiting to be fixed and adapting it to the current situation substituting similar parts that had similar characteristics [3]. To select a successful layout that perfectly matches the current problem requires an efficient system and also a strong knowledge base.

One of our purposes of this research is to make a systematic analysis of the possible results and factors in the projects, which applies CBR during the development period. To complete this research, literature review and interviews are applied. In the literature study, we went through much research in the area of CBR theory and its application. In the interview, we

designed a series of questions asking people currently working in different positions in the software industry. The answers of those interviews are the data we analyze.

This report is composed of seven sections. The first section is this introduction. The second is the detailed introduction to the background of the study, including the introduction of case-based reasoning and knowledge management. The third section is composed of research questions. The fourth section shows the methods we applied in this study. The fifth section demonstrates the results. The sixth section is a discussion about results we got. The seventh and the final part is our conclusion.

2. Theoretical Background

2.1.1 Literature study

We had doubt that CBR would fit in other software engineering domains besides Artificial intelligence (AI) due to the stereotypical impression one gets of CBR. According to the research conclusion of Smyth and Cunningham, there is one promising area of application for CBR in software engineering [4]. Some scholars like Janet L. Kolodner has conducted research showing that CBR could be often use for solving problems on many issues in software engineering, such as design, planning, diagnosis, and explanation [5].

Besides the benefits described in the previous paragraph, CBR also has practical significance in real industry. Ian Waston (1997), has shown that organizations as diverse as IBM, VISA International, Volkswagen, British Airways, and NASA have already made use of CBR in applications such as customer support, quality assurance, aircraft maintenance, process planning and decision support. And many more applications are easily imaginable [6].

2.1.2 Search strings

- Case Based Reasoning AND Knowledge Management
- Case Based Reasoning AND Software Engineering
- Knowledge Management AND Software Engineering

The database we used to collect the scientific research papers is:

- Engineering village: <http://www.engineeringvillage.com/>
- Google Scholar: <https://scholar.google.se/>

The inclusion criteria of the papers were:

- In relation to case-based reasoning in software engineering domain, in terms of knowledge management.
- Scientific papers which were peer-reviewed, written using empirical methods.

The exclusion criteria of the papers were:

- No focus on our search area
- Not considered to be scientific
- Considered to be biased

2.2 Related work and literature review

2.2.1 Case-Based Reasoning

Case-based reasoning can be seen as the process of solving new problems associated with the solutions of similar past problems. A short classical definition of case-based reasoning is “a case-based reasoner solves problems by using or adapting solutions to old problems [4].”

Generally, a case-based reasoner will be presented a new problem according to another user or program, and then the case-based reasoner will search into the past cases in order to find a case that have the same requirement specifications as the current case. The solution proposal will be directly applied into the new case if the case-based reasoner can find the same one, otherwise the case-based reasoner will attempt to find a case or cases that are most close and similar to the current one.

Originally, the case-based reasoning cycle contains four phases, which are retrieve, reuse, revise and retain [5] [6]. Researcher Ian Waston developed this case-based reasoning cycle into a more detailed model that consists of six activities, which are retrieve, reuse, revise, review, retain and refine respectively [7]. The diagram he draw for this new model of the case-based reasoning cycle is revealed in figure 01. But there is no big difference between these two models because the last three activities Ian Waston mentioned, can be concluded into retain phase in the old theory, but we will represent case-based reasoning cycle by following his steps.

The retrieve activity is to find and decide which case in the case base is same or similar to the current one, which could either transfer the solution proposal to the current case or inspire the case-based reasoner. When the case has been found in the case base, the CBR system will retrieve the targeted case that provide the detailed solution for the solved problems either by completely applying it to solve the new problem in the current case, or will bring up the similarities to inspire the case-based reasoner in order to conduct a potential solution to solve the new problem.

The reuse activity means applying or adapting the targeted case that has been found in the case base to solve the new problem in the current case. At the reuse phase, the solution in the targeted case can be used to solve the new problem without modification, and adaption might be applied to establish a potential solution that is suitable for solving the new problem. The purpose of case adaptation in reuse phase is to increase and improve the overall ability of problem solving in CBR system by using the newly adapted cases for future needs.

The goal of revise activity is to verify that case adaption do match the requirements of solving the new problem. The activity revise also prepare the newly adapted case to be peer reviewed on its own values whether they are worth to be retained as a new case in the case base or not. And these are the two activities after revise in the CBR cycle.

The last activity refine is to make sure that the case base index and feature weights are accurate.

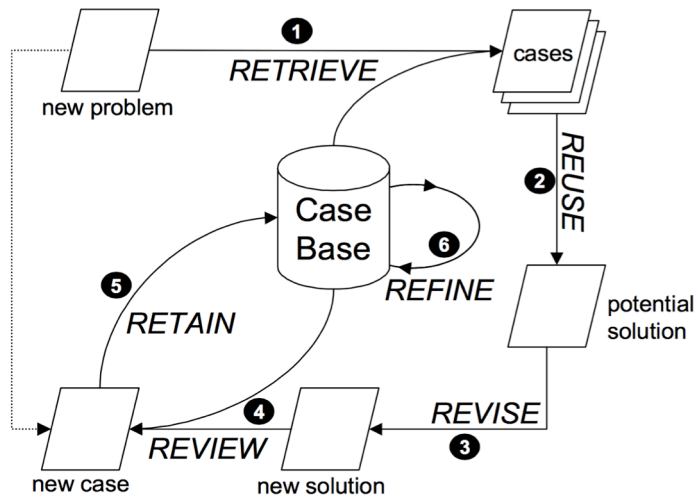


Figure 01 (Sourced from Ian Waston)

2.2.2 Knowledge Management

Knowledge management can be seen as a business discipline involving the identification, capturing, organization, analysis, and processing of data and information to create new knowledge, which is made available to others in order to use and create more knowledge [9]. Knowledge management attempts to address one big problem that is that large amounts of

knowledge in an organization are not reused, which has been mentioned in one of our interviews for this project. The advantage of implementing knowledge management includes improving decision-making, creating a better condition for information sharing and exchange, improving the ability of change management on flexibility and effectiveness, and to preserving the knowledge in order to decrease the knowledge loss. The key factors of knowledge management are creation, sharing and reuse. Knowledge management solutions mostly consist of computer networks (e.g. intranets, extranets), storage facility (e.g. relational database), capture and collection systems (e.g. document management systems), dissemination technologies (e.g. data warehouse), and knowledge processing and analysis technologies (e.g. data mining, data visualization) [10].

There is another element that has a strong relationship with knowledge, and that is information. When it comes to making a distinction between information and knowledge, researcher Ian Waston explained it as “should be seen in terms of a dynamic and interactive relationship” [11]. He quoted the phrase from the earlier study that Nonaka and Takeuchi stated, “information provides a new point of view for interpreting events or objects, which makes visible previously invisible meanings or sheds lights on unexpected connections. Thus information is a necessary medium or material for eliciting and constructing knowledge [12]”, this reveal that information facilitates the development of knowledge, which creates more information that deepens knowledge [11]. Figure 02 shows the dynamic relationship between information and knowledge.

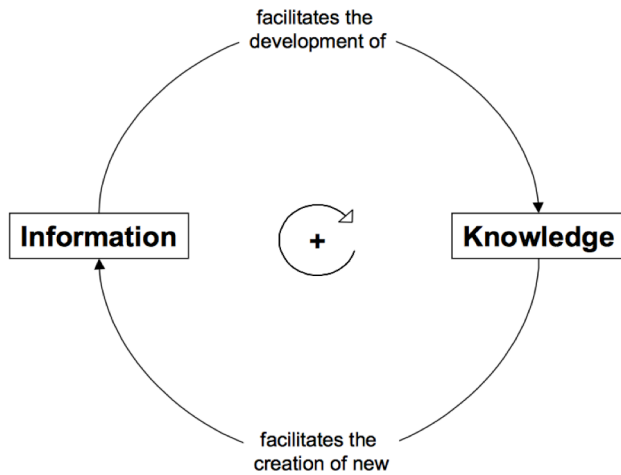


Figure 02 (sourced from Ian Waston)

Therefore data, information and knowledge have a dialectical relationship that can be visualized as in figure 03 below. Why looking at the information purely in terms of the degree to which it has happened, it is important to remember that the data, information, and knowledge hierarchy oversimplifies the complex relationship between the three intangibles [13]

[14]. As Stewart mentioned earlier in the research, “the idea that knowledge can be slotted into a data-wisdom hierarchy is bogus, for the simple reason that one man’s knowledge is another man’s data” [15].

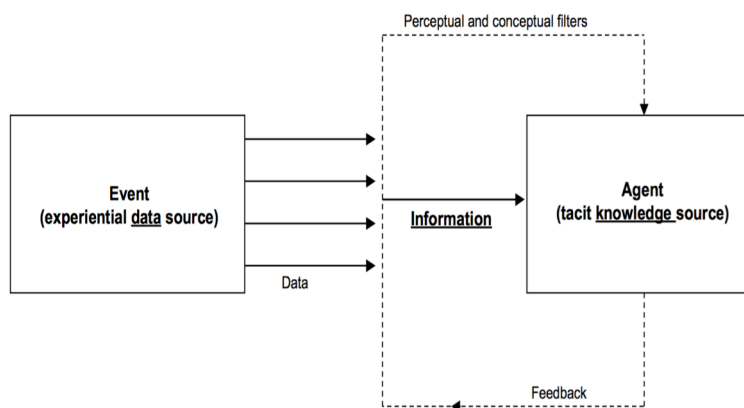


Figure.03 (sourced from Boisot)

2.3 A example of applying case-based reasoning into the knowledge management

In order to provide a clear view of applying CBR into knowledge management in real-life, to give a hint of the working process and development scheme, an example will be introduced to briefly explain the CBR in knowledge management.

Assume there is a company W now doing business in language training, the new project is to establish an online platform (website) that helps the student and the teacher/tutor to get connected, and making the language learning process more easier and less-boring. The company team consists five people who are responsible for three different core roles, they are the team leader/Scrum master, developers (3 individuals) and the product owner respectively. The team follows the development process of Scrum in order to accomplish their goal.

Case-based reasoning has been brought to the team for their actual development. In the early stage of the development, all team members collect relevant information in order to reach the final achievement and make it the software requirement specification. Each one of the team members has been assigned to a task called preparation, which means everyone needs to go through the previous cases in the internal system in order to find cases sharing similar goals

and conditions. All the similar cases will be collected together and they will be judged whether they are useful for the new project development. The qualified cases will be reused for the new project development, including re-applying computer language, developing timeline design, delivery goals, quality measurements, and emergency reaction plan. All those re-applied factors have one thing in common, for being highly suitable for the new project development and being adaptable. All measurements and standards will be revised according to the new project's software requirements specification, and constant developmentally adjustment will be applied based on the rule of sustainable development for CBR refine. Each phase of the development process will be documented in order to track the record if something happened, and all these documents will be revealed again when it comes to making this new project as a new case in the case base in their internal system after the project has been accomplished. This new case will be retained in the internal system for the further use. This is a complete cycle of case-based reasoning.

Case-based reasoning in this case can be seen as an approach of knowledge management in the whole working process. It helps the developer to track the record of software development, helps the project leader/Scrum master to manage the change during the time period of the project development, helps the company to regulate the timeline of product delivery and helps the company to store the knowledge for further use.

3. Research Question

3.1 Main Research Question

To what extent of case-based reasoning used in practice?

3.2 Sub-questions

- What are the advantages of CBR?
- What are the motivations of applying CBR and Knowledge Management?
- What is the practical importance of applying CBR and Knowledge Management?
- Under what kind of circumstances would you apply CBR?
- What are the challenges of applying CBR and Knowledge Management?
- How do you see the future of CBR?

Sub-questions are the supplementary factors to the main research area, the aim of these questions is to produce a thorough knowledge for participants in order to complete the main research goal. Such knowledge as the advantages of CBR, the future of CBR not only to educate the participants on the concept of theory, but also to inspire the participants to a better management methodology on increasing the efficiency and effectiveness of the process.

4. Methodology

The main type of research methods that we have applied to this study, namely, interview study. Several theoretical studies and related research papers built a good theoretical

foundation in our study. Literature review was the very first step for us to conduct this research that gave us a broad view of knowledge management and CBR in both theoretical and practical sense. One of the reasons why we choose to apply interviews to collect relevant data is that by choosing the subjects of the interviews we could select more professional and qualified subjects than random people, that means, we were more likely to get useful and reliable data from them being professionals in their fields of experience. Another of the merit of doing interviews is that face-to-face interviews reduce the risk of misunderstanding. Also, during interviews, we can get even more information than we expected. We continued the literature review, which helped us doing essential comparison, and helped us update our knowledge at the same time.

4.1 Data collection

4.1.1 Literature study

We had doubt that CBR would fit in other software engineering domains besides Artificial intelligence (AI) due to the stereotypical impression one gets of CBR. According to the research conclusion of Smyth and Cunningham, there is one promising area of application for CBR in software engineering [4]. Some scholars like Janet L. Kolodner has conducted research showing that CBR could be often use for solving problems on many issues in software engineering, such as design, planning, diagnosis, and explanation [5].

Besides the benefits described in the previous paragraph, CBR also has practical significance in real industry. Ian Waston (1997), has shown that organizations as diverse as IBM, VISA International, Volkswagen, British Airways, and NASA have already made use of CBR in applications such as customer support, quality assurance, aircraft maintenance, process planning and decision support. And many more applications are easily imaginable [6].

4.1.2 Search strings

- Case Based Reasoning AND Knowledge Management
- Case Based Reasoning AND Software Engineering
- Knowledge Management AND Software Engineering

The database we used to collect the scientific research papers is:

- Engineering village: <http://www.engineeringvillage.com/>
- Google Scholar: <https://scholar.google.se/>

The inclusion criteria of the papers were:

- In relation to case-based reasoning in software engineering domain, in terms of knowledge management.
- Scientific papers which were peer-reviewed, written using empirical methods.

The exclusion criteria of the papers were:

- No focus on our search area
- Not considered to be scientific
- Considered to be biased

4.1.3 Interview Study

The method we use to collect information and data consist of interview studying. This method has been used frequently in a lot of researches and reports. We have visited three

different types of IT companies and did interviews with employees in different working positions. In order to complete this research, we designed a series of questions (shown in Appendix). The focus of this set of questions was on how CBR performs in real industry and the real outcome. Most the dialogues between interviewers and interviewees were documented on text.

The companies and individual we choose to interview were all in the IT industry, more specifically all within software engineering domain. They were different in size measured by business area and number of employees, and on different levels of business development. One is a major international company, another is a medium sized local company and the third other one is a start-up small local company.

4.1.4 Interview design

In conducting the interview study, we selected three different companies that are very different in size. By doing interviews with those three companies, we got data that presents the different knowledge manage approaches used by companies of different size. The interviewees who participated in the research are managers in the target companies. They are familiar with the managing system of the company they work for. It was unknown if they are applying case-based reasoning. Then how they are managing their existing knowledge became the most essential topic in the interview. Based on the assumption that the target company is currently not applying CBR in their project, the following questions we designed are about the basic knowledge of CBR, in order to find out the reasons why they applying current method rather than CBR. The series of questions that we designed are flexible, they change and facilitates further questions in accord with a real dialogue.

Company A

Company A is a worldwide leading provider of communication technology and services. From our research, we found out that the method of case-based reasoning is frequently apply to various types of software products in the company A. Moreover, because company A is running a huge number of projects at the same time from different locations all over the world, they need a very systematic way to manage a huge knowledge database. Also, to meet the changing requirements, the quality and efficiency of change management for them is significant as well.

The interviewee of this company is the manager of the test department in Gothenburg, Sweden; managing several test teams, who are dealing with different projects.

Company B

Company B is a star-up IT internet company which is aiming at building a open international language learning platform, mostly for the Chinese market. As a start-up company, it does not yet have a clear understanding of the target group; the coming challenge is to meet the demands of the potential customers. That means that they are facing the problems in knowledge management. To apply case-based reasoning in their development procedure is a priority idea. For us, interviewing company B brings interesting result.

The interviewee is the company B's current CTO (Chief Technology Officer) and also a developer.

Company C

Company C is a small sized company that has existed for over 30 years. It currently has 6 employees working. The main business company C run is to supply administrative software to municipalities around Sweden. Company C owns much successful experience and technology as well as a strong and a mature develop team.

The interviewee is currently the CTO of company C. The main responsibilities include software developing, concept generation, sales and process improvement. Except for the role of CTO, this interviewee is also a course responder in Gothenburg University with solid knowledge and experience in knowledge management and other software engineering fields.

Company D(Tangiamo)

Company D designs, develops and manufactures multi-player gaming systems for casinos and electronic arcades. The company is based in the city of Gothenburg on the West coast of Sweden, and sells its products in over 25 countries worldwide.

The interviewee is an experienced software developer who is working for company D in Sweden.

Company E(volvo)

Company E is an international conventional company located in Gothenburg, Company E has many sites located all over the world with over 20,000 employees.

The interviewee is an experienced software specialist who mainly develops features about safety and has a great overview about the knowledge management in Company E.

Company F

Company F was established by a small group of software developers in the mid 90s, the main business that they deal with is game development in Europe. Nowadays, company F has around 70 employees in three countries in total.

The interviewee is a software developer of this company who works in the site of Växjö, Sweden.

Interviewee G

Interviewee G is working at company E, because interviewee G has not worked long for company E, he does not have enough experience on knowledge management in company E. Based on the fact, interviewee G shared his knowledge management ideas from his PhD experience in the Netherlands as an individual developer.

4.1.5 Data Analysis

Most of the dialogues are recorded in electronic devices, then transformed in to text. Based on the information we collected from interviews. In each interview, we have one or several question to let interviewees describe the current knowledge management method they are using and we analyze and differ the knowledge management extent into 4 different levels:

1. *Company or individual does not have any textual knowledge stored, knowledge only exist in brain and exchanges manually.*
2. *There are some data or previous projects(code) on third party, for example, GIT.*
3. *Has an exclusive system or platform to store knowledge and exchange knowledge.*
4. *Has an exclusive Case-based Reasoning (CBR) system which obeys the CBR circle.*

Another issue we concerned is how do people think of applying Case-based reasoning on knowledge management. In other words, we selected information about interviewees hold either positive or negative opinion about the method of CBR.

4.2 Validity threats

4.2.1 Internal validity threats

Internal validity threats have the most impact on the reliability of the data we collect. There are many internal factors that influence the results. The main internal validity threats can be seen as following.

- The design of the interview questions is not comprehensive enough. There might still be something that is not clear or lead to misunderstanding.
- Last minute cancellations from other companies that we intend to interview.
- The data privacy that limited the answers to our interview questions.
- The interviewee's capacity.
- Misunderstanding of the concept of CBR and KM.
- The level of knowledge management is not clear enough.

4.2.2 External validity threats

There are several external validity threats to our research but the biggest one is that we only had three interviews in the IT industry. Therefore it is difficult to generalize our research results.

4.3 Conclusion

Due to the fact that the study sample we had is not big enough, the result cannot reflect the overall use of current knowledge management methods in the software industry. Also, the concept of CBR is not as well-known as we thought. During the interviews, we still need to explain, we cannot guarantee that interviewees really understand it after explanation. Moreover, before the interviews, the interview questions have not been tested, thus the questions that we designed might be unclear and increase the difficulty to collect relevant data from interviewees.

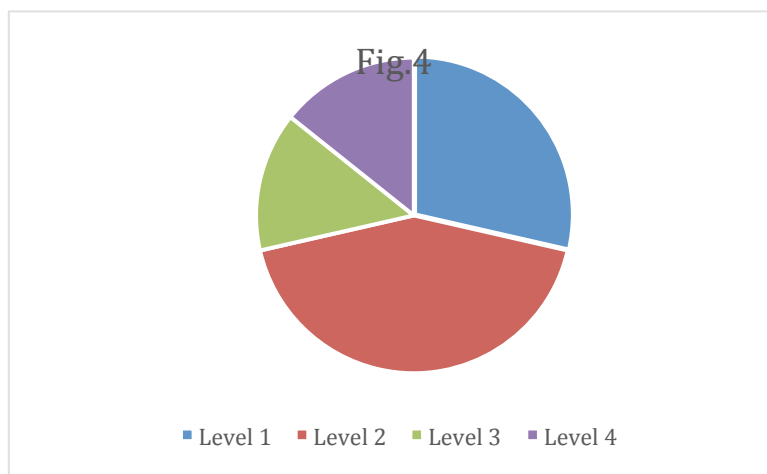
5. Results

According to the information data we collected from the interviews for this project, it has been shown that there are some companies now using the theory concept of CBR even though without themselves noticing it, for example company A. It shows that the advantages of CBR have potential to benefit the knowledge management in an organization no matter the size of the organization. The results shows that the CBR application has been adapted differently due to company size and number of employees, and the difficulty of CBR implementation is strongly dependent on many factors, such as budget, human resource, time control, etc.

As we mentioned before regarding validity threats, considering that the number of company participants in this project is limited, the results does have its limitations and hard to generalize from. We planned to interview more companies and collect more data but we couldn't manage to it.

One of the biggest findings in the results is that the size of the company does have a tremendous impact and influence on the CBR application implementation, in particular on challenges. Another big issue we noticed, which might have affected the result is that different companies have their different focus on different areas, which could have resulted in the data we collected being not 100% accurate for answering our main research question. But in general, the data we collected is useful and helpful.

Another fact that we found from our data collection is the big/international companies are more likely to apply some methods for knowledge management due to their tremendous amount of intellectual properties. A system which knowledge can be kept in order to be reused and developed is required and there is a trend of the knowledge management extent. According to our degree of standardization, big/international companies are more likely at the level of 3 & 4; small size companies are more likely at the level of 1 & 2 ,shown fig4 .



We also found that CBR has not been widely used and recognized in real industry and there are some hard facts that make the CBR implementation more complex. The disadvantage of CBR application, such as requires a large space to store the cases, takes long time finding similar and related cases -- will be explained later. The more detailed result will be revealed in the followings.

5.1 CBR could be a perfect match for knowledge management

After research and investigation of interviews from companies in real industry, we found that CBR could be a perfect match for knowledge management just like its advantages has been suggesting. According to the advantages of CBR and the data we collected from the interviews, 6 out of 7 interviewees declare that CBR has its potential to bring convenience

and effectiveness to organizations and could lead to the best solution to maximizing benefits in a complex system.

The first result we noticed is that the six-REs of the CBR cycle can be mapped to the activities required by the knowledge management system, in accordance with the results as Ian Waston [7], described earlier. A simplified mapped system that displays the mapping of six-REs of the CBR cycle into the knowledge management system is shown in figure 04 below.

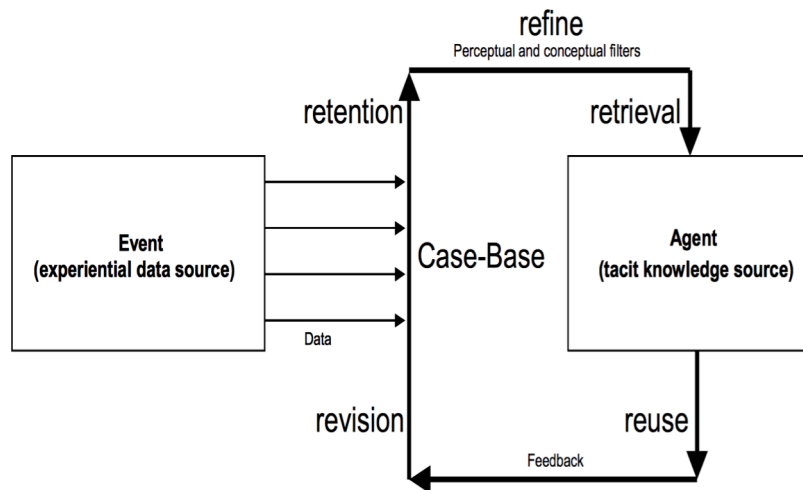


Figure.04 (sourced from Ian Waston)

Comparing with Boisot’s original diagram, it’s easy to notice that the event can be seen as an episodic data that constitute a case, which is retained in the case base but undergoes pre-processing and filtering at the same time. The case based reasoner or the agent retrieves the case in order to find a solution to solve a new problem and reuses the case. This may bring revision or adaptation of the case’s solution and may result in creating a new episodic data that constitutes a new case. The new case will be reviewed in order to determine if it’s useful and worth to be retained in the case base. In addition, the use of the case will be refined according to the case-base’s indexing scheme and the feature weights.

5.2 The advantages of CBR application and its practical importance

According to the findings in literature study, CBR can offer various advantages when used under proper condition [15] [16] [17] [18]. Besides being a theory, CBR also has its practical importance, something that has been potentially confirmed after we have interviewed three industrial cooperators for this thesis research project. In summary, CBR has its potential to improve the efficiency and the effectiveness of knowledge management in an organization, especially in a big international transaction company like company A. The most discussed advantages and the practical importance of CBR will be listed as following, some of which will be described in a more detailed way and some of them will be introduced by referring to the theory.

1. CBR is an intuitive performance - this is how everyone works.

According to the definition of CBR shown in the theory, CBR is a way of solving a new problem by adapting a similar solution from previous cases. The way CBR works is actually the way of how humans react in order to solve new problems. There is no guarantee that CBR is the best solution in problem solving, but CBR has been found to be a suitable approach in problem solving due to its advantages.

2. CBR does potentially improve the productivity by increasing efficiency and effectiveness. Firstly, CBR potentially makes the development easier. According to the findings in the literature review, there are some factors which show that CBR makes development easier -- such as avoiding repeating all the unnecessary steps that need to be taken to arrive at a solution, or potentially reducing the time spent on actual development work.

Secondly, CBR potentially makes the maintenance easier due to its own benefit from the process cycle.

3. CBR creates the possibility for avoiding repeating the same mistakes that have been made in the past.

Due to the fact that a CBR system stores the record of the failures as well as the successes, the system can hand out the knowledge to the case-based reasoner in order to predict the potential risk of having to apply some particular rules or disciplines. Previous cases could be used to identify possible factors for failure and be avoided in the new case.

4. CBR may be used to predict the success of the new solution.

As well as having the advantage of being applicable to new solutions, CBR could also be used to avoid the same mistake that has been made in the past. CBR could be used to predict the success of the new proposed solution due to the fact that all factors in the case will be stored in the case base, which could be revised and reused for new solutions to new problems.

5. CBR makes the system learn over time.

When CBR systems are used, they encounter more situations and create more solutions. CBR systems constantly grow via absorbing and adjusting the new solutions into the case base in order to provide a better opportunity continuously for a case-based reasoner to find the similarity in case difference, and hence drive the new solution towards success.

6. CBR can be used in different ways.

The number of ways a CBR system can be implemented is almost unlimited [18]. The way of applying CBR can depend on the purpose of creating a solution, or justifying a point of view, or making some suggestions for plan adaptation.

7. CBR can be applied and extended to various domains.

Due to the fact that it has almost unlimited ways of representing, indexing and adapting cases, CBR can be applied and extended to various domains. CBR originally arose from artificial intelligence (AI) but nowadays its capacity has been discussed and extended to other domains like software engineering -- as we revealed from the literature study.

5.3 To what extent is CBR application used in practice and why

After the investigation of the use of CBR application in real industry, we found that CBR application has not been so widely used. Only one of our interviewees had heard about CBR theory before. There are many reasons that can be addressed regarding the disadvantages of the CBR application, and the following listed areas can be seen as the disadvantages.

- CBR needs a large storage in order to contain all the cases.
- CBR could increase the processing time while the case-based reasoner is about to find similar and relevant cases.
- The difficulty of the case adaption in CBR process could be high and there is no guarantee for success.
- CBR could be able to create a suitable solution for problem solving but it is not guaranteed that this is the best solution to solve the problem.
- CBR application cost a lot of investment such as human resources, crossover functionality, time resource, budget investment, etc.

5.4 Findings of the most mentioned issues by the interviewees regarding the use of CBR

There are several facts that we discovered when we were implementing our investigation for our research project, some of them we had already predicted when we were doing the literature review, and some of them we noticed when we were conducting our interview study.

First of all, CBR is still not a well-known concept in the real industry, which was confirmed by our interviewees. We have been pondering about how many people out there in the industry have heard about CBR, the reason why we had this in mind is due to its similarity with common sense in knowledge management.

Secondly, CBR is not irreplaceable by other methods in knowledge management. According to the interview 2, the international well-known company A is currently using some other method in order to manage their knowledge sharing within the company.

Thirdly, CBR seems too hard to be fully applied alone without support of any other method in the knowledge management. According to interview 1, interview 2 and interview 5, the interviewees revealed either that there is no CBR applied in their company or that CBR could be seen as partly applied in the knowledge management.

The next most mentioned issue is the budget. According to the data we collected from our interviews, budget has been mentioned repeatedly. Lot of interviewees believe that budget is one of the most obvious and crucial factors when it comes to apply CBR. No matter what size of the company, budget is and will always be one of the most important thing that need to be considered before the company take any kind of action.

In the end, the most important and exciting issue which we have discovered in the interviews, is that CBR has a bright future and that it seems to be necessary for knowledge management in every IT company.

5.5 Challenges and research evaluation

During the time when we were implementing our research, one of the biggest challenges was whether the interviewees would be able to answer our questions. Based on the different backgrounds of our interviewees, their daily work focus could be irrelevant from what we want to investigate. Even though CBR has its capacity to be extended to various domains, the un-awareness of the CBR application in real industry could be a common issue.

Our interviewees have discussed the challenges of CBR application; the biggest one is how to implement CBR application in the real industry. Almost all of our interviewees agreed that CBR application is important and necessary in the knowledge management in any IT companies, but how to make it work and maximize the benefit out of it is still uncertain and undiscovered. Time manner has been discussed several times according to different interviewees, one think that CBR may not be good because there is no guarantee for finding a solution by applying CBR and CBR could possibly increase the processing time due to its working procedure. And at the same time, the CBR application does not have irreplaceability in the knowledge management. Suggestion to those challenges is has to do more with action research on CBR application with the purpose of how to implement CBR in order to accomplish the knowledge management in the real industry.

6. Discussion

Theoretically, applying CBR to software development project brings many advantages. In academic field, applying CBR as a method of knowledge management has not been frequently talked about. But according to our research, we found that the approach of case-based reasoning has not been widely applied in software industry. The biggest challenge is to have an advanced knowledge system. For start-up companies or small sized companies, because of their lack of project experience, successful project samples are not able to be gathered. Also, copyright issue may occur if information gathering is well done from another company or individual developer. Having a running CBR system requires a huge knowledge foundation as well as manpower. In the CBR circle, the stage of retrieve, revise and retain, requires people with solid development background to evaluate and manage them properly. Thus, the budget to build up infrastructure for CBR is not always affordable, which becomes a significant reason for small sized companies to not have CBR application in their project. Another fact in small companies is that few individuals hold the knowledge. They share experiences and techniques in a natural way. For instance, sharing successful experience by face-to-face discussion.

Many big companies that have enough financial support choose not to have the exact CBR application due to the complex procedure for sorting and selecting all the relevant successful cases. Also, software industry changes gradually, many think that to include all the successful samples in a database requires too much manpower, in other words, a big amount of successful samples might not be reused again. It wastes storage, manpower and time.

Big company that are applying CBR has already got great benefit. One of the most evident advantage is that all the coding data is recorded in the system. However, in current industry there are still many essential missing in the database of system. There is a fact that many employee has their own technique or special skill to work, but they will not share if nobody ask, thus, new employees may spend a long time to figure it out rather than read it in document. Documentation has always been a weak part in industry, things like advice, tips to work should also record into system. If all those knowledge has been stored and reuse by other employees, the budget for training could be reduced.

From the research result, we found that in ordinary software product industry, the penetration of the CBR application has not been going on as quickly as might be thought, due to the seeming similarity between CBR and ordinary information and knowledge exchange. The reason we could think of has been mentioned already. But another reason why CBR application has not been widely used could be that there is no well developed CBR application model yet.

One approach to make CBR more widely known could be, firstly, to getting more people know about the CBR concept and theory. The more people who know about it, the more chance for CBR application to be used in the future. Secondly, getting more people involved in developing a model of CBR application, in order to making the CBR application more adaptable into the knowledge management system in different companies. In the end, CBR must extend its capacity and ability in functioning with other knowledge management approaches in order to create a better working environment.

7. Conclusion

To summarize our research project, we can say that CBR has been widely discussed in theory according to different purposes, CBR is no longer just a theoretical issue, it does have its practical importance in making it an application in real industry. The CBR method has been known in the real industry of course, in order to supplement the knowledge management with a purpose of creating a sustainable development environment for the IT industry.

According to our research purpose, we choose to apply a mixed methodology combining literature review and interview. We firstly conducted the literature review by going through many relevant research and study cases. It gave us the knowledge in our research area and helped us designing the interview questions. Thus we could be able to implement the interviews in a good way and helped us in data collection and result analysis. And, based on our result, some of our predictions have been confirmed by the interviewees: it shows that there is a gap between the CBR in theory and the CBR application in real industry. The CBR in theory suggested the advantages of the CBR application in real industry and the actual use of the CBR application in real industry displayed there is a room for further study and development of the CBR theory. The widespread lack of knowledge of the CBR theory, combined with what seems to be a very widespread use of profitable CBR application in real industry suggest a fruitful research area for the future of CBR theory and its conscious application in real industry, which is also the initial motivation for us to conduct this research project.

According to our research, the future of CBR application is bright and there are many other aspects that wait to be discovered or accomplished. CBR is really important and necessary to the knowledge management in any IT companies due to its potentiality of improving productivity by increasing the efficiency and effectiveness of the development process. Due to its limitation and disadvantages shown in our search, CBR may not be suitable when it comes to finding the best solution for fixing a problem, but it does have a potential to develop a business model and to be generalized.

8. Acknowledgement

We wish to express our sincere thanks to Imed Hammouda, program manager of Software Engineering and Management, Department of Computer Science and Engineering, who helped us in organizing the thesis project. We wish to express our sincere thanks to Anthony Anjorin as well, the supervisor of our thesis project, Post-doc, Department of Computer Science and Engineering, who provided us with important essential knowledge and guidance for the thesis project. Also, we are grateful to the company A, company B, company C and the interviewees there who graciously accepted our interview inquiry, and provided us with essential data and documents.

References

- [1] Kai Bartlmae & Michael Riemenschneider, Case Based Reasoning for Knowledge Management in KDD-Projects, DaimlerChrysler AG
- [2] Checkland, P. & Scholes, J. (1990). *Soft Systems Methodology in Action*. Wiley, UK.
- [3] Deepak Khemani, Knowledge Management with Case-based Reasoning, Department of Computer Science & Engineering IIT Madras.
- [4] Reisbeck, C.K. & Schank, R. (1989). *Inside Case-based Reasoning*. Erlbaum, Northvale, NJ, US.
- [5] Aamodt, A. & Plaza, E. (1994) 'Case-Based Reasoning: Foundational Issues, Methodological Variables, and System Approaches', *Artificial Intelligence Communication*, vol. 7, no. 1.
- [6] Watson, I. (1997) *Applying Case-Based Reasoning: Techniques for Enterprise Systems*, Morgan Kaufmann Publishers, Inc., San Francisco
- [7] Watson I. Knowledge Management and Case-Based Reasoning: A Perfect Match? [C]//FLAIRS Conference. 2001: 118-122.
- [8] Kang S H, Lau S K. A framework for case-based reasoning integration on knowledge management systems [J]. 2003.
- [9] Radding A. (1st ed.). 1998. *Knowledge Management: Succeeding in the Information-based Global Economy*. In *Report of Computer Technology Research Corp.*
- [10] Dubitzky W, Büchner A G, Azuaje F J. Viewing knowledge management as a case-based reasoning application [C]//AAAI Workshop Technical Report WS-99-10. 1999: 23-27.
- [11] Watson I. Knowledge Management and Case-Based Reasoning: A Perfect Match? [C]//FLAIRS Conference. 2001: 118-122.
- [12] Nonaka I, Takeuchi H. *The knowledge-creating company: How Japanese companies create the dynamics of innovation* [M]. Oxford university press, 1995.
- [13] Davenport, T.D. (1997). *Information Ecology, Mastering the Information and Knowledge Environment*, Oxford University Press, Oxford, UK.
- [14] Checkland, P. & Howel, S. (1998). *Information, Systems and Information Systems, Making Sense of the Field*, John Wiley & Sons, UK.

[15] Janet L. Kolodner. An introduction to case-based reasoning. *Artificial Intelligence Review*, 6(1):3-34, 1992.

[16] Janet Kolodner and William Mark. Case-based reasoning. *IEEE Expert*, 7(5):5-6, October 1992.

[17] Kevin D. Ashley. Case-based reasoning and its implications for legal expert systems. *Artificial Intelligence and Law: An International Journal*, 1(2-3):113-208, 1992.

[18] Main J, Dillon T S, Shiu S C K. A tutorial on case based reasoning[M]//Soft computing in case based reasoning. Springer London, 2001: 1-28.