Making Use of a System's Treasury of Experiences

Mirjam Minor
Department of Informatics, Humboldt University Berlin
Unter den Linden 6, 10099, Berlin, Germany
minor@informatik.hu-berlin.de

Abstract

In this article, we discuss ideas for a new model for a case-based reasoning (CBR) system's treasury of experiences (ToE). Like human brains, CBR and other experience management systems learn not only new knowledge contents during their life cycles but develop also immanent knowledge structures depending on the system's work history. Systems that communicate their experiential knowledge improve the transparency for the users and allow computer programs to integrate and use other programs' know how. We give an overview what knowledge sorts belong to the treasury of experiences besides of the actual knowledge contents. We propose a simple representation framework and recommend communication principles and forms. Finally, we show some sample applications for the framework and discuss the benefits and further application areas of this approach.

1 Introduction

The crucial problem of experience management (EM) is to make experiential knowledge accessible in that cases and on that places where it is needed [Bergmann02]. Recent EM systems provide solutions for this by means of intelligent techniques like case-based reasoning (CBR). We observed that the systems are not only mediators of knowledge but become also treasures of knowledge and experiences themselves during their life cycles. They are frequently confronted with new knowledge, new persons, and new situations. In analogy to the human brain where every new experience, every input from the sense organs modifies the neural structure [Eliot99], a system's experiences have of course an impact on the immanent knowledge (see also [RothBerghofer02], [NickAlthoff01]). The aim of this paper is to investigate how experience management systems can actively use and communicate their treasury of experiences. We concentrate our focus on CBR systems but still with the idea in mind that the model is possibly also applicable to other experience management systems, for instance ontology-based systems. The first steps of this ongoing work are (1) to define clear terms to identify and describe a system's treasury of experiences (ToE) and it's parts, (2) to find appropriate representation and communication forms, and (3) to investigate with the help of sample applications whether it is worth-while to make use of the whole treasury of experiences in practice.

First, we will describe which sorts of knowledge belong to a system's treasury of experiences. In Section 3, we will investigate a net-like structure to represent items of all knowledge sorts and their relationships. In Section 4, we will propose principles and different ways to communicate this knowledge to both human and artificial receivers, and finally, present some application examples in Section 5 and give a summary with an outlook in Section 6.

2 A System's Treasury of Experiences

A long-lived EM system updates its knowledge in different ways. New human experiences are made in a certain application domain and have an impact on the system's current knowledge contents. Besides of the stored actual contents, the systems gather plenty of further knowledge during their usage. This knowledge is needed to execute or improve the systems' core tasks. By and by, the systems become treasuries of experiences themselves. A system‘s treasury of experiences consists of but is not restricted to the following sorts of knowledge:

- Knowledge contents
- Work history of knowledge pieces
- Indexing and other background knowledge
- Valuating knowledge
- Social knowledge
- Knowledge about users

The *knowledge contents* are the pieces of knowledge that the system manages for the users and that are accessible for them. In case-based reasoning (CBR), the knowledge contents are the cases.

We call the knowledge besides of the actual contents the own immanent experiences of the system.

The *work history of knowledge items* stores knowledge that has emerged during the life cycle of this item. This includes, for instance, different revisions, knowledge authors and knowledge sources respectively, eventually also the expiration date.

The *indexing and other background knowledge* is the knowledge that is necessary for the core tasks of the system. In case-based reasoning, those tasks are mostly retrieval and adaptation. The indexing and background knowledge for CBR is the vocabulary, domain ontologies and other descriptions of similarity relationships, as well as adaptation rules.

Valuating knowledge qualifies knowledge, for instance, by an assigned degree of maturity or received user feedback. Valuating knowledge can concern particular knowledge pieces, as well as arbitrarily chosen subsets of the treasury of experiences, e.g. the knowledge contents. Of course, the whole treasury of experience can be valuated, too. In CBR, valuating knowledge can be used to improve the retain and maintenance processes.

Social knowledge adds a social dimension to the knowledge pieces, e.g. the reliability of persons, systems, groups of both, or knowledge sources, the responsibility for knowledge pieces or tasks, and roles a person or system can play. Current research on social knowledge in computer systems is done in the field of socionics (see [Whitley96], [LindemannMuench01]).

Knowledge about users can be long-term or short-term knowledge: Some systems work with persistent user profiles; others observe or request information just for the current session.

Further sorts of knowledge in a system's treasury of experiences are knowledge about contexts in which certain pieces of knowledge are valid, or maintenance specific knowledge like maintenance strategies, meta knowledge and so on. Recent CBR systems in the literature store particularly just a subset of the above mentioned most common parts of the treasury of experiences. They use these own experiences to improve the core tasks. For instance, in [MinorHanft00] the work histories of cases are used to support current retain processes (see short description in Section 5).

3 Knowledge Representation

The knowledge contents of an EM system are usually represented in well-defined forms. In case-based reasoning, the cases are stored in a uniform case format. For the other knowledge types of EM systems, we found a variety of different representation forms in the literature. It is an intriguing question to look how human beings internally represent their knowledge and to learn from them for the EM systems, but there has not yet been found a widely accepted and practicable answer. To investigate the human knowledge representation is out of our focus and we let the psychology and the neuro sciences [Eliot99] work on it. Let us rather discuss some known representation forms and, secondly, think about a suitable way to represent an EM system's treasury of experiences.

We know several precise and well-understood representation forms like static data structures or logic languages. Many AI approaches extend the exact representation forms with the aim to model a bit more flexibility and human inexactness. Good examples for this are extended logics, semantic nets, ontologies, or cases. These representation forms are still readable and understandable for human beings, at least for those with a mathematical background. For non-mathematics, they can be made quite understandable by clever graphical user interfaces. Beside them, we have some useful black box approaches like neural networks that cannot be understood in detail. Such forms are especially necessary to process visual and audio data.

Instead of focussing on one of those representation forms or even develop a new one, we recommend a net-like structure which connects arbitrarily chosen representation forms for different sorts of knowledge. The only facts that we really know and would like to consider in our representation form, are the following: Some parts of the own experiences are strongly related to the actual knowledge contents, e.g. indexing knowledge to particular cases, while others are still independently, e.g. new vocabulary gathered from an electronic dictionary with the expectation that according cases will probably be acquired in future. The connection between immanent experiential knowledge of the same or of different sorts can be quite complex, as the source as well as the target of such a link can be both single pieces of knowledge or sets of knowledge pieces within a sort or even over the boundaries of sorts (like evaluating knowledge). We decided to simplify this by the following model which can be regarded a network of single knowledge pieces.

Definition

We describe a system's treasury of experiences ToE as a tupel:

$$ToE = (KC, EXP, EB, ECC).$$

It consists of the ground sets of the actual knowledge contents KC, the own experiences EXP, the function EB (experiential background) which links the knowledge contents with the system's own experiences to a certain degree:

$$EB : KC \times EXP \rightarrow \Re$$

and the function ECC (experiential cross connections) which links pairs of own experiences:

$$ECC : EXP \times EXP \rightarrow \Re$$

In the case some of the own experiences are made isolated from particular knowledge contents and are not yet linked to any matching content, the values of EB for the isolated experiences with all knowledge contents are zero. The same holds for ECC and experiences that are isolated from other experiences in EXP.

In case-based reasoning, KC is the set of cases in the case base while EXP consists of all the technical and organizational pieces of knowledge which Richter calls the structure S over a case base [Richter00]. The representation of a case with a set of information entities is expressed by the according values of EB, while the local similarity relationships between information entities are expressed by values of ECC.

4 Communication of Knowledge

To use a system's treasury of experiences requires the communication of knowledge. This means that the system should not only use its own experiences for the core tasks but also make them accessible from outside the system. The receivers can either be human beings, e.g. users and administrators, or other systems. Communicating internal knowledge to users improves the transparency of the system and motivates users to keep the system in use and, hopefully, to participate in maintenance processes (see the RoboCup scenario in the outlook section). Anyhow, the administrators normally have access on all knowledge items of the system. If common users are supposed to insert new knowledge into the system, they need not necessarily to know all the internal structures. But we made the experience that it can be de-motivating if they have the feeling that they are not taken seriously enough to know what will happen with their knowledge. They should have at least the opportunity to look behind the curtain. Communicating the experiential knowledge to other systems allows them to make use of it and possibly integrate it with their own experiences. In both cases, of course, the privacy is not to be hurt.

First, we will have a look again on the way how human beings solve this problem. We assume that the most of the human knowledge is communicable in principle. How is it communicated in reality? Which media are used by human beings, which of them are adaptable to robots or programs? People write letters or e-mails, talk with their voices either directly or by means of a phone. They use prosodic and body language. Of course, they use the body also to communicate body-dependent knowledge like playing the piano or riding a bike. Books and slides can be regarded a very body-independent form of communication of knowledge, even the particular receiver is unknown, at best the target audience. Systems communicate with other systems usually by means of protocols, with the users by means of a user interface.

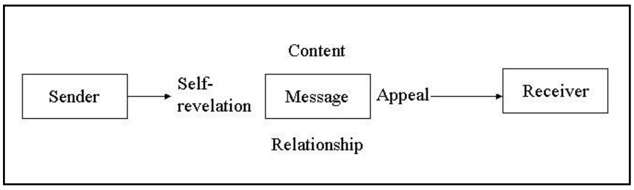

Fig. 1. The four aspects of a message following Schulz von Thun

A psychological model of communication between people from Schulz von Thun [SchulzvonThun81] (see Fig. 1) helps to find a model of communication for KM systems: Schulz von Thun defines four aspects of a message: The objective content, the relationship between sender and receiver, the self-revelation, and the appeal. When human beings communicate, they tell always something about themselves and their model of the receiver. Independently of the medium, it leads often to misunderstandings to focus only on the content. Applied to the communication of knowledge in an EM system, the system does not necessarily need an extra medium to transport additional information to the knowledge content, it can append it to the message itself. This method guarantees the principle of voluntariness: The additional knowledge is sent in a way that the receiver can decide either to skip it, or to take it into account, or to integrate it into the own treasury of knowledge. The communication works without a model of the receiver and the skills of the receiver: The parts of the received experiential background that are semantically not understandable can simply be skipped.

We learn from this model the following aspects. In case of a human receiver, the own experiential knowledge can be communicated together with the actual knowledge contents. The system attaches links to the knowledge origin, the knowledge history, executable programs, responsible persons, and so on. So, the user has not to deal with information overload but can easily have a look on interesting details. In case of an artificial agent as receiver, the exchange of own experiences has to be a bit more sophisticated. Some agents observe users to gather knowledge about their behavior and to learn from them. In a similar way, it could be possible to learn from other agents by observing how to kick a ball, how to perform an internet research, how to find a solution for a problem. It could also be possible to receive textual information and to interpret it by means of own techniques. Both seems to be quite difficult and will probably be a future task. But let us return to recent agent theory: In a multi-agent system, agents cooperate with other individuals to extend the own abilities. If an agent sends a service with the according part of its thesaurus or ontology, it reveals information about its own experiential knowledge. Of course, additional information needs to be integrated with the overall communication protocol.

5 Sample Applications

Some parts of this work have been implemented in sample applications. First, we present three applications where the receivers of the messages are human beings, and then we describe a multi-agent system.

The *ExperienceBook* is a textual CBR system for the support of system administrators and common computer users. It manages descriptions of problem cases resulting from the daily work with computers, e.g. how to delete Windows' ^M symbols in UNIX files:

```
<CASE>
  <CASE_NUMBER>
    74
  </CASE_NUMBER>
  <TITLE>
    remove Windows' ^M from file
  </TITLE>
  <ATTRIBUTES>
    AUTHOR = M. Minor
    DATE = 14.03.2003
    MODE = lay and expert
  </ATTRIBUTES>
  <SYMPTOMS>
    If one transfers files from Windows to
    UNIX, e.g. via ftp, sometimes ^M replaces
    the usual line break.
  </SYMPTOMS>
  <SOLUTION>
    open file with emacs, mark a ^M and use
    Search->query replace
    copy ^M to the command line, let it re
    place with nothing (just press newline
    twice), and give the answer ! to the
    question 'replace query…'
  </SOLUTION>
</CASE>
```

The system has been in use at our AI Lab for nearly five years now. The technical part of the system works properly: It is online accessible and selects the most similar cases from the case base concerning a query in natural language. For details of the mapping of texts and attribute-value pairs on sets of information entities and of the similarity function, we refer to [LenzEtal98]. The development of the case base, the requirements of the users, and the experiences that the system has made during the last years have inspired us to the current work on a system's treasury of experiences: Some of the solved problem cases have got new revisions with alternative or newer solutions, but the users are still interested in the old revisions. The case base has been extended with links to scripts and other programs according to particular cases. Astonishingly, the system has 'autonomously' enhanced its focus and includes some cases about administrative processes within our institute now. These cases refer to external forms for proposals etc. The ExperienceBook has self-developed to a system with an open net-like structure beside of the actual knowledge contents. In terms of the above defined treasury of experiences, we identify the following issues in the current implementation:

KC: set of cases C
EXP: indexing knowledge (set of information entities IE to represent the cases internally), work history (external sources like executable programs, administrative forms etc.)

$$EB(c,e) = \begin{cases} rel(e,c), & e \in IE \\ link(c,e), & else \end{cases}$$

where $c \in C$ and $e \in EXP$, $rel: IE \times C \rightarrow [0,1]$ describes the relevance of an IE representing a case and $link: C \times (EXP \setminus IE) \rightarrow \{0,1\}$ describes the linkage between a case and an external source.

$$ECC(e_i,e_j) = \begin{cases} sim(e_i,e_j), & e_i,e_j \in IE \\ 0, & else \end{cases}$$

where $sim: IE \times IE \rightarrow [0,1]$ describes the degree of local similarity from one IE to another. The overall similarity between a query and a case is computed by the following sum formula:

$$\sum_{e_i \in Query} \sum_{e_j \in Case} sim(e_i,e_j)$$

We omitted the relevance value, as we simplified the formula in the way that rel's value is 0 for not relevant and 1 for relevant IEs. The sets of information entities as well as the links to further sources have not yet an impact on the retrieval results. They are presented to the users as links from the cases in hypertext form. The principle of voluntariness is here especially important, as the users are often under time constraints when they query the system. They are glad to read the answers in a compact form. The oral feedback concerning the additional information in a 'second layer' is positive. For the future, we are planning to implement more parts of a treasury of experiences, e.g. further parts of the work history of cases. The new terms of the treasury of experiences will hopefully help us to keep the structure of the system easily comprehensible.

In an industrial cooperation with the software company PSI AG, Berlin, we have developed a case-based system for the *authoring support of software tests* [MinorHanft00]. The system stores test cases in all degrees of maturity and provides access on them when people have to write test reports. Copy and paste of similar cases is very useful to do the boring work of test specification. The cases have a life cycle, i.e. are persistent but develop from test ideas to complete test cases with testing environments. The system communicates its own historic experience intensively to the users. To store this historic dimension, we use a repository like in a revision control system. It manages different views on the testing knowledge for querying and editing purposes particularly.

KC: set of test cases C
EXP: indexing knowledge (set of information entities IE), work history (former revisions FR, authors, and dates), valuating knowledge (degree of maturity)

$$EB(c,e) = \begin{cases} rel(e,c), & e \in IE \\ link(c,e), & else \end{cases}$$

where $c \in C$ and $e \in EXP$, $rel: IE \times C \rightarrow [0,1]$ describes the relevance of an IE representing a case and $link: C \times (EXP \setminus IE) \rightarrow \{0,1\}$ describes the linkage between a case and its former revisions, the current author or the date.

$$ECC(e_i,e_j) = \begin{cases} sim(e_i,e_j), & e_i,e_j \in IE \\ link(e_i,e_j), & e_i \in FR, e_j \in EXP \setminus FR \\ 0, & else \end{cases}$$

where $sim: IE \times IE \rightarrow [0,1]$ describes the degree of local similarity from one IE to another and $link: FR \times (EXP \setminus FR) \rightarrow \{0,1\}$ describes the linkage between a former revision of a case and its current author or date.
In this project, it is part of the core task 'authoring support' to communicate the work history of the cases to the users. The current retain processes use experiential knowledge from former case revisions.

The *SocionicsBook* stores definitions of important terms which occurred during a study of socionical literature. To deal with the multiple definitions in the literature, the additional dimensions of origin and responsibility is integrated: Each definition has a link to the original source in literature plus a link to the persons who are responsible for the concerned sub-area of socionics. The practical use has shown that it is very useful to get into contact with benevolent members of such communities of practice. In this project, we integrated the own experiences directly with the knowledge contents. So, we have only the following issues:

KC: set of cases C
EXP: indexing knowledge (set of information entities IE)

$EB(c,e) = rel(e,c)$

where $c \in C$ and $e \in EXP$, $rel : IE \times C \to [0,1]$ describes the relevance of an IE representing a case.

$$ECC(e_i, e_j) = \begin{cases} sim(e_i, e_j), e_i, e_j \in IE \\ 0, else \end{cases}$$

where $sim : IE \times IE \to [0,1]$ describes the degree of local similarity from one IE to another

At our AI Lab, *a multi-agent system for the users' personal assistance* has been implemented [Kuehnel99] [MinorWernicke03]. The agents can exchange services to solve tasks. Recently, the user interfaces of the agents have been extended - according to the above presented ideas of the communication between virtual agents - by a natural language user interface using textual CBR technique. The linguistic and domain-specific knowledge about relationships between terms etc. can be exchanged as background knowledge together with the services.
KC: service descriptions SD (including the according executables and the source code if permitted by the access rights)
EXP: indexing knowledge (set of information entities IE), social knowledge (the social type of an agent, e.g. egoist or altruist, is expressed by the access rights AR given to the particular services)

$$EB(sd, e) = \begin{cases} rel(e, sd), e \in IE \\ ar(sd, e), else \end{cases}$$

where $sd \in SD$ and $e \in EXP$, $rel : IE \times SD \to [0,1]$ describes the relevance of an IE representing a service description and $ar : SD \times AR \to \{0,1\}$ describes the linkage between a service description and an access right.

$$ECC(e_i, e_j) = \begin{cases} sim(e_i, e_j), e_i, e_j \in IE \\ 0, else \end{cases}$$

where $sim : IE \times IE \to [0,1]$ describes the degree of local similarity from one IE to another.

The communication of the particularly relevant parts of indexing, similarity, and social knowledge is integrated with the communication protocol of the service descriptions. The social knowledge controls parts of the communication. New indexing terms and new similarity relationships are merged with the receiver agent's own treasury of experiences. For further details, we refer to [MinorWernicke03].

Summary and Outlook

In this paper, we discussed a model how case-based Experience Management systems can represent and communicate their treasury of experiences to other systems and to human receivers. We presented some ideas for a net-like representation of the different knowledge sorts belonging to a system's treasury of experiences, recommended principles and ways to communicate it, and, finally, investigated some sample applications. The main benefits of using and communicating a system's treasury of experiences are more transparency for the users and more capabilities for participating computer programs. To proof that this framework really improves the human-computer and the computer-computer interfaces of CBR systems requires further research. We are planning to implement and evaluate some results of the above described theory in an internet encyclopaedia for the RoboCup project [RoboCupFederation]. The project deals with football robots and consists of several national and international research teams. It follows the idea of open source code what leads to some knowledge management problems increased by the frequent introduction of new members. As a solution, we will make available an internet encyclopaedia with the help of experienced staff members. The articles will define important issues of the RoboCup world and will refer to according entries in the RoboCup newsgroup. The system will use textual case-based reasoning to autonomously link newsgroup contributions. Furthermore, it will ask the users for help to maintain the system's knowledge, e.g. to assess a particular link, to insert a new term into the vocabulary, or to decide to what degree a term is similar to another. It is really intriguing which experiences the system will make and whether it will collect useful knowledge additionally to the manually edited entries of the encyclopaedia. The frequency and quality of the users' contributions will hopefully provide a good evaluation scenario for the model of a system's treasury of experiences.

References

[Bergmann02] R. Bergmann: *Experience Management - Foundations, Development Methodology, and Internet-Based Applications*, LNAI 2432, Springer Verlag, 2002.

[Eliot99] L. Eliot: *What's Going on in There? How the Brain and Mind Develop in the First Five Years of Life*, Bantam Books, New York, 1999.

[Kuehnel99] R. Kühnel: *Assistant Agents that Distribute How-To-Knowledge*, Proceedings of AOIS'99, 1999.

[LenzEtal98] M. Lenz, A. Hübner, M. Kunze: *Textual CBR*. In: Case-Based Reasoning Technology - From Foundations to Applications, Lenz M., Burkhard HD., Bartsch-Spörl B., Wess S. (Eds.), LNAI 1400, Springer Verlag, Berlin, 1998.

[LindemannMuench01] G. Lindemann, I. Münch: The Role Concept for Agents in Multi-Agent Systems. In Sozionik aktuell, Ausgabe 3/2001, ISSN 1617-2477, 15-30, 2001.

[MinorHanft00] M. Minor, A. Hanft: *The Life Cycle of Test Cases in a CBR System.* Proc. EWCBR-2000, LNAI 1898, S. 455-466, Springer Verlag, 2000.

[MinorWernicke03] M. Minor, M. Wernicke: *The Exchange of Retrieval Knowledge about Services between Agents*, Proceedings of the GWEM 03, to appear.

[NickAlthoff01] M. Nick, K.-D. Althoff: Engineering Experience Base Maintenance Knowledge, In: Advances in Learning Software Organizations, K.-D. Althoff, R.L. Feldmann, W. Müller (Eds.), LNCS 2176, p. 222 ff., Springer Verlag, 2001.

[Richter00] M. M. Richter: *Fallbasiertes Schließen*, In: Handbuch der Künstlichen Intelligenz, G. Görz, C.-R. Rollinger, J. Schneeberger (Eds.), 2. Auflage, p. 407 - 430, Oldenbourg Verlag, München, 2000.
.
[RoboCupFederation] The RoboCup Federation http://www.robocup.org

[RothBerghofer02] Th. Roth-Berghofer: *Knowledge maintenance of case-based reasoning systems. The SIAM methodology*, PhD thesis, University of Kaiserslautern, 2002.

[SchulzvonThun81] F. Schulz von Thun: *Miteinander reden, Störungen und Klärungen*, Hamburg 1981.

[Whitley96] E. A. Whitley: *Confusion, Social Knowledge and the Design of Intelligent Machines*, Journal of experimental and theoretical artificial intelligence, 8 (3/4), pp. 365-381, 1996.