Case-Based User Modeling in an Experience Management System

Andreas Jedlitschka & Klaus-Dieter Althoff

Fraunhofer Institute Experimental Software Engineering (IESE),
Sauerwiesen 6, D-67551 Kaiserslautern
{jedl|althoff}@iese.fhg.de

1. Introduction

Experience-based continuous learning is essential for improving products, processes, and technologies in emerging as well as in established areas of business and engineering science. It can be facilitated by case-based organizational learning, meaning that relevant experience is captured in the form of cases for reuse in a corporate experience repository (case base; experience base; EB). For obvious reasons, learning from experience needs to be a permanent endeavor. Thus, an organization has to handle a "continuous stream of experience". For this purpose, a learning organization, called "Experience Factory" (EF; Basili et al. 1994, Althoff et al. 2000), was established at Fraunhofer IESE with the Corporate Information Network (CoIN) initiative (Althoff et al. 2001). On one side, CoIN grants access to intranet-based information like organizational processes and news and, on the other side, case-based access to project management experiences. The objective of the next development step of CoIN is to provide users with valuable information/knowledge at the right time, in an adequate representation, and within the actual context ("just-in-time").

In this article, we describe ongoing and planned work to enhance our existing experience management system. We introduce new strategies to capture, process, disseminate, and exchange knowledge. Case-based reasoning (CBR) is used for both knowledge modeling/retrieval/adaptation as well as for a "learning from example" based approach to user modeling. A brief outlook concludes this article.

2. New Strategies to Capture, Process, Disseminate, and Exchange Knowledge

Knowledge has actually been identified as the "fourth factor of production"[1]. Therefore, unstructured, non-personalized flooding with information can be counterproductive for building up and exchanging knowledge (Fischer & Ye 2001; Jameson 2001). To better support our employees, we are (a) moving from a "pull" to a "push" strategy in the sense of providing the right information at the right time (context-sensitive), (b) developing more flexible and faster mechanisms for capturing and sharing information by introducing Communities of Practice (Jedlitschka & Althoff 2001), and (c) developing a method for aggregating and adapting information to users' contexts and needs by using CBR. Within his doctoral dissertation Tautz (Tautz 2000) evaluated COIN and concluded that the technology-based approach provides more valuable information than the human-based approach. After employing the initial CoIN for almost two years we interviewed representative users.

[1] Besides work, capital, raw material (e.g., Stewart 1997)

One result of this survey was the fact that with increased complexity of the system the needs of the users increased as well. This has been expressed with the demand for features like “myCoIN”, i.e., assistance in form of “personalization”.

The main challenge is to convince users of the system’s helpfulness. The user should notice a personal gain. Only this will bring him to spend some voluntary effort. As a minimum, a break-even with respect to time/effort spent and time/effort gained should be reached (Kluge 1999).

2.1. “Push” of Information/Knowledge

We do not want to burden users with overhead for searching information or asking for experience. Our solution grants a single point of access, admission to all knowledge and information produced in an organization, only restricted by access rights defined by (a) the organization in the form of the employee’s role within it, (b) the projects and the corresponding role the employee plays, and (c) the owner of a piece of information. Therefore, a user interface has to be developed corresponding to the presentation layer further described in Jedlitschka & Althoff (2001).

With his login in combination with stored but also dynamic user data (organizational role, project roles, skills, and interests (Fig. 1)) and a chosen view (e.g., concrete project, information channel), the user provides the actual context, for example: “project: x; role: developer; task: code testing” (the task is determined from the project plan). The given context is, on the one hand, used to build his individual navigation bar (e.g., below the topic projects, only those projects that he is a member of are listed). On the other hand, the context is necessary for the delivery of knowledge (knowledge is gained within a context, anyway). If the context (including the user model) is treated as a case, it can be compared with other contexts. CBR helps to identify similar contexts (see 2.2). Thus, it is possible to deliver knowledge gained within former similar contexts without an explicit user query (“push” of information). The user can ignore the delivery but, hopefully, he will at least evaluate the utility of the delivered information within his actual context. The evaluation is used, on the one hand, to “educate” or “edge” agents for users’ business and personal information needs and, on the other hand, to get more accurate evidence in accordance with the usability of this information for other users as well. The agents observe users’ behavior (i.e., navigation), and they are also “responsible” for discovering desired information. Personal needs can be context-sensitive and/or free of user’s choice.

In the case of new, improved, or changed content, the user will be informed automatically, if he has registered for this service. This automatic information can take place either by mail or directly in the news window of the application. It is thought of as a multi-step news channel structure: The most important news, categorized by the author, are presented directly within the main window (not more than five), whereas the others are presented within their context. The user registers only once for those components he wants to be informed of.

Additionally, he can send specific queries to the EB (“pull”). For this “manual” query, the user has two options: he can either use the context provided by the system or he can specify his own. For that individual specification, he first restricts the resources and then posts a free text.

We plan to support this approach using a combination of structural and textual CBR. While case-based retrieval is used to find the most similar knowledge items based on the login/context information and/or the specific query, a user case base is

used for bridging the gap between the known information about the user and the knowledge the user is "really" currently looking for (Weibelzahl & Weber 1999).

2.2. Aggregation and Adaptation of Information

Every member of an organization or, more abstractly, every role, has different needs with regard to the granularity of information. Stepping higher on the organizational or project level, information has to be aggregated and adapted more and more with respect to urgency and criticality. Therefore, different information pieces have to be fragmented. The relevant fragments have to be selected, brought together in a convenient way (aggregation), and prepared for the user (adaptation) in accordance with his needs. This concept is well known in data warehouse methodologies. Extending these approaches, we are dealing with experience in the form of un-/structured documents. Text mining shall support the gaining of valuable information to confirm/reject experience.

The user gets standard information in addition, with an attribute telling him about the degree of utility (personalized or evaluated experience) and the name of the author. Highly aggregated and adapted information can usually not be assigned to one unique input source. The level of aggregation and adaptation is then given to the user, so he is able to comprehend the outcome. Detailed information (source information) is available on demand, which is especially of interest if a state is detected as being critical.

While project members need specific and in-depth information about their status within the project, the project leader is more interested in an overview of all project activities. For him, knowing that a deviation will occur (e.g., because of illness of a project member) is valuable information. Experiences from similar cases enriched with input from the risk plan can assist him in evaluating the critical potential of this state. If he detects a business-critical state, the information is forwarded on a "red-phone" channel to the respective persons. This channel will also be used if an addressed person does not react after a pre-defined period of time.

The approach of retrieving parts of information from different sources and configuring (Wilke 1999) and adapting (Bergmann 1996) these items to one item to be presented is analogous to the case adaptation step in the basic CBR process model (Aamodt & Plaza 1994). Techniques available here can be found in Althoff (1997) and Bergmann (2001). The technical realization of the aggregation and adaptation component resides within the representation layer further described in Decker & Jedlitschka (2001).

It is planned to offer various different information services (IS) through CoIN, roughly grouped into organization (management), business (projects, business areas), research (core competencies), service (help), employee, and up-to-date (news, absence list, lunch plan) topics. Every user can subscribe to some IS: whereas some are specific to the roles (business area manager, project manager in project x) he plays in the organization, others are optional (competence areas like knowledge management). Each IS is realized by a query case base (initially one query case) and a content case base (provides content of the IS). The cases are characterized by attributes describing the context (project, process) or by keywords. This approach

will be realized by a combination of structural and textual CBR using a commercial CBR tool (orenge from empolis knowledge management[2]).

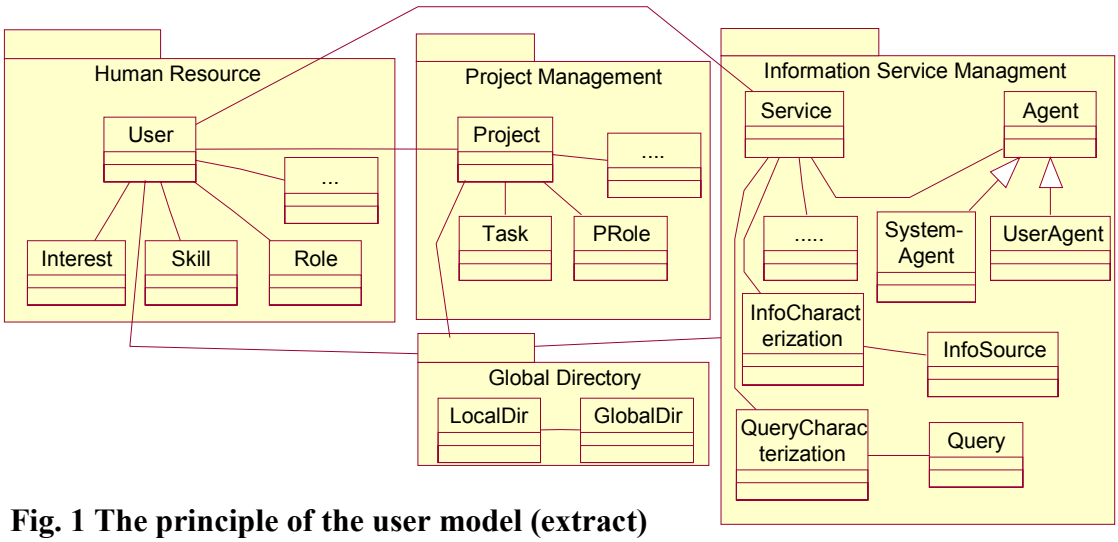

Fig. 1 The principle of the user model (extract)

The user model is partially kept within an extended human resource system (Fig. 1), where each user is represented by his organizational (roles, skills) and individual (interests) data. Additional information comes from the projects the user is involved in and from stored data about earlier behavior (e.g. navigation, information services that the user has registered for). Thus, the user model is physically distributed and has, together with the actual view, dynamically joined to the actual context.

Because users cannot know which information sources are available, it can be useful to use queries that have been used successfully by users with similar interests (Weibelzahl & Weber 1999). In CBR so called problem-solution pairs are stored and cases are used for case-based retrieval. With this terminology, the combined object context is treated as a "problem". When a new context is obtained, for example, because the user changes his view, this context is handled as a new case, which has no solution yet (Fig.2). Similarities are used to retrieve a most appropriate matching case, which then includes a solution representing a query successful used in a former context. In the sense of user modeling terminology, this is a collaborative learning approach, because former contexts are used for estimating the actual context and here the user model is part of the context.

[2] www.km.empolis.com

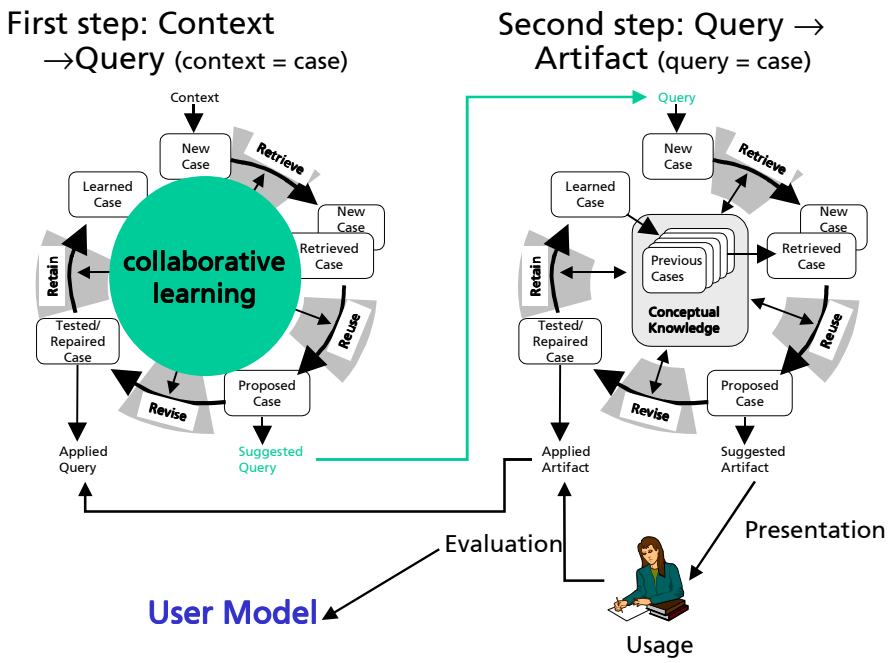

Fig. 2 Two-step case-based reasoning

The query gained in the first CBR cycle is the input for the second CBR cycle, resulting in a similar case with a query-artifact pair. This artifact is presented to the user. The evaluation of delivered information by the user is necessary for future evolution of the whole system. If the user agrees with the information, the value of the query that led to the content and the content itself are increased (the same will happen to a successfully reused context-query pair). In case of rejection, the value of the query for the next retrievals is decreased. Additionally, the rejection is stored as a hint for the maintenance of the case base. In particular, advanced users can pose a new query, which is then stored in the query case base. With his collaboration, the user also contributes to his own model. The evaluation data delivers data about the actual interests and refines the context. So we also have the content-based approach to user modeling. The results of the two-step CBR and the user evaluation are used for better forecast of users' needs with regard to information.

To clarify the functionality, we provide an example for an IS named "What is interesting to read". For instance, a business area manager reads a study on knowledge management tools. She evaluates it as useful and offers it for this IS. The service asks her to characterize the study briefly. The characterization is needed for similarity-based access to the study in the future. She notes that it may be useful for business area managers and colleagues who are interested in knowledge management. If another business area manager logs into the CoIN system later (with respect to his role he has also subscribed to this IS), the link to the study is offered automatically, because there is increased similarity due to the match in the role. If someone logs in who is interested in knowledge management, which is described as part of his individual user model, the link is also offered. Further on, if a person managing a knowledge management project, which is described in the project information, logs in, increased similarity would again cause the link to be provided.

3. Related Work

This work is related to other efforts taken in the following areas: Experience Management (Tautz 2000), CBR used for user modeling (UM) (Weibelzahl&Weber '99, Waszkiewicz et.al.'99, and Bradley et.al.'00), UM itself (Jameson '01), and Workflow Management (Elst et.al.'001). Our approach employs CBR for two purposes: to learn from users' past behavior (content-based) and to learn in a collaborative way by assuming that users behave in similar ways as other users before (collaborative) (Zukerman 2001).

4. Outlook

Managers tend to find information more valuable when they are convinced of the reliability of the respective source (Traphöner 2001). Solving this problem is a major topic in the work on future knowledge management technology. We try to solve it by using CBR as the starting point, and by using an open environment that will be extended with other techniques as appropriate (e.g., from machine learning, knowledge management, etc.):

- Contributing information *and* sources to the case base, on the one hand, and evaluation points that provide information on successful application, on the other hand.
- Forecasting of user's agreement with delivered information based on user's history stored in the case base. If a similar person (i.e., role) within a similar context considers a similar information item useful, hopefully the addressee has the same opinion. For how to improve similarity assessment based on a knowledge discovery approach, see Rech et al. (2001).

Currently we are working on virtual competence centers for software engineering. For these projects, knowledge/experience repositories have to be constructed. The difference with respect to CoIN is that in the beginning we do not know anything about the user. Therefore, we are developing a system – again based on CBR technology – to acquire informational requirements from the user, with the help of the approach introduced in this article.

5. References

Aamodt, A. and Plaza, E. (1994). Case-Based Reasoning: Foundational Issues, Methodological Variations and System Approaches. *AI Communications Vol. 7*, *No. 1*, 39-59.

Althoff, K.-D. (1997). *Evaluating Case-Based Reasoning Systems: The INRECA Case Study*. Postdoctoral thesis (Habilitationsschrift), Department of Computer Science, University of Kaiserslautern.

Althoff, K.-D., Birk, A., Hartkopf, S., Müller, W., Nick, M., Surmann, D. & Tautz, C. (2000). Systematic Population, Utilization, and Maintenance of a Repository for Comprehensive Reuse. In G. Ruhe & F. Bomarius (Eds.), *Learning Software Organizations - Methodology and Applications*, Springer Verlag, Lecture Notes in Computer Science, LNCS 1756, 25-50.

Althoff, K.-D., Decker, B., Hartkopf, S., Jedlitschka, A., Nick, M. & Rech, J. (2001). Experience Management: The Fraunhofer IESE Experience Factory. In P. Perner (ed.), *Proc. Industrial Data Mining Conference*, Leipzig, July 24-25, 2001, Institut für Bildverarbeitung und angewandte Informatik.

Bartsch-Spörl, B., Jargon, C. & Althoff, K.-D. (2001). Wissensmanagement in der Praxis – Erfahrungen aus erfolgreichen und weniger erfolgreichen WM-Projekten. Workshop auf der 1. Konferenz Professionelles Wissensmanagement: Erfahrungen und Visionen, Baden-Baden Mar. 14-16, 2001, (http://demolab.iese.fhg.de:8080/AK-Praktisches-Wissensmanagement/WM2001-WS5/).

Basili, V.R.; Caldiera, G.; Rombach, D.: Experience Factory; In Marciniak, J.J. ed., Encyclopedia of Software Engineering, vol 1, 469–476; John Wiley & Sons; 1994.

Bauer, M., Gmytrasiewicz, P. J. & Vassileva, J. (eds.) (2001). *User Modeling 2001*. Springer Verlag, LNAI 2109.

Bergmann, R. (1996). *Effizientes Problemlösen durch flexible Wiederverwendung von Fällen auf verschiedenen Abstraktionsebenen*. PhD Thesis, Department of Computer Scienece, University of Kaiserslautern.

Bergmann, R. (2001). *Experience management - foundations, development methodology, and internet-based applications*. Postdoctoral thesis (Habilitationsschrift), Department of Computer Science, University of Kaiserslautern (submitted).

Bradley, K.; Rafter, R.; Smyth, B.: Case-Based User Profiling for Content Personalization; In Proc. of the Int. Conf. on Adaptive Hypermedia and Adaptive Web-based Systems, (AH2000), Springer, LNCS 1892

Decker, B. & Jedlitschka, A. (2001). The Integrated Corporate Information Network iCoiN: A Comprehensive, Web-based Experience Factory. In: K.-D. Althoff, R. Feldmann & W. Müller (eds.), *Advances in Learning Software Organizations*, Springer Verlag, 192-206.

Elst, L.v.; Abecker, A.; Maus, H.; Exploiting User and Process Context for Knowledge Management Systems; In Proc. of Workshop on UM for Context-Aware Applications at UM2001.

Fischer, G. & Ye, Yunwen (2001). Personalizing Delivered Information in a Software Reuse Environment. In Bauer et al. (2001) *U M 2001*, 178-187.

Jameson, A. (2001). *Personalization for E-Commerce*. Tutorial presented at UM2001.

Jedlitschka, A. & Althoff, K.-D (2001). Using Case-Based Reasoning for User Modeling in an Experience Management System. *Proc. Workshop "Adaptivität und Benutzermodellierung in Interaktiven Systemen (ABIS01)", GI-Workshop-Woche „Lernen – Lehren – Wissen – Adaptivität (LLWA01)*“, Universität Dortmund, Oct. 8-12, 2001.

Kluge, A.: Erfahrungsmanagement in lernenden Organisationen; Verl. für angewandte Psychologie; Göttingen 1999.

Nick, M.; Althoff, K.-D. & Tautz, C. (2001). Systematic Maintenance for Corporate Experience Repositories. Computational Intelligence 17(2), 364-386.

Rech, J., Decker, B. & Althoff, K.-D. (2001). Using Knowledge Discovery Technology in Experience Management. *Proc. Workshop "Maschinelles Lernen (FGML01)", GI-Workshop-Woche „Lernen – Lehren – Wissen – Adaptivität (LLWA01)“,* Universität Dortmund, Oct. 8-12, 2001.

Stewart, Thomas A.: “*Der vierte Produktionsfaktor: Wachstum und Wettbewerbsvorteile durch Wissensmanagement*“, München: Hanser 1998.

Tautz, C.: Customizing Software Engineering Experience Management Systems to Organizational Needs; Ph. D. diss., Dept. of Computer Science, University of Kaiserslautern, Germany; 2000; Stuttgart: Fraunhofer IRB Verlag.

Traphöner, R. (2001). Knowledge Sharing Medium. In H. U. Buhl et al. (eds.), Information Age Economy, *Proc. 5th International Conference on Business Information Science (WI 2001*), Physica Verlag, 223-234, 2001.

Waszkiewicz, P.; Cunningham, P.; Byrne, C.: Case-based User Profiling in a Personal Travel Assistant; In Proc. of the Int. Conf. on Adaptive Hypermedia and Adaptive Web-based Systems, (AH2000), Springer, LNCS 1892

Weibelzahl, S. & Weber, G. (1999). Benutzermodellierung von Kundenwünschen durch Fallbasiertes Schließen. In Proc. 7. GI-Workshop "Adaptivität und Benutzermodellierung in interaktiven Softwaresystemen", LWA-1999, Otto-von-Guericke-Universität Magdeburg, Germany, Sep. 29-30, 1999; 295-300.

Wilke, W. (1999). *Knowledge Management for Intelligent Sales Support in Electronic Commerce*. PhD Thesis, Department of Computer Science, University of Kaiserslautern.

Zukerman, I. & Albrecht D.W. (2001). Predictive Statistical Models for User Modeling. User Modeling and User-Adapted Interaction (UMUAI) 11(1-2): 5-18 (Ten Year Anniversary Issue), 2001. Kluwer Academic Publisher, Netherlands.