Maintaining Experience to Learn: Case Studies on Case-Based Reasoning and Experience Factory

Klaus-Dieter Althoff*, Jens Mänz* & Markus Nick**

*University of Hildesheim, Intelligent Information Systems
Marienburger Platz 22, D-31141 Hildesheim, Germany
althoff|maenz@iis.uni-hildesheim.de

**Fraunhofer Institute for Experimental Software Engineering (IESE)
Sauerwiesen 6, D-67661 Kaiserslautern, Germany
markus.nick@iese.fraunhofer.de

Abstract

In the past many experience factory case studies and experiments have been carried out. We summarize some development steps and research results that, from our perspective, are important. We especially focus on the integration of experience factory and case-based reasoning and report on the respective benefits and impacts of such a seamless integration for building (more) autonomous and automated knowledge-based information systems, which will be of increasing importance in the future. It is our goal to build software-agent-enacted experience factories that improve case bases using maintenance and learning methods.

Introduction

Knowledge management (KM) provides promising approaches to make software engineering (SE) more “knowledge-based”. Experience factory (EF) is such an approach for which meanwhile many lessons learned are available. This contribution especially addresses the overlapping area of research of both artificial intelligence (AI) and SE. One of our goals is to enable a closer cooperation between researchers and practitioners of both fields. According to Kirn [Kir04] and Rombach [Rom04] this is a necessary and important step but requires, among others, to learn how researchers/practitioners of the respective “other” field think and work. With case-based reasoning (CBR) and EF an integration is presented that exemplifies which kind of synergies can be achieved.

It is our goal to develop information systems that have a high degree of automation and autonomy because of built-in learning capabilities. Knowledge about EF as a human-enacted learning organization is used to develop software-agent-enacted EFs that improve case bases using maintenance and learning methods. A collection of thus improved CBR systems is then used to build these future information systems.

In this contribution we focus on the integration of EF and CBR and the motivation for such future information systems.

In the following we shortly survey EF as an approach with relevance for KM in SE focusing on the above mentioned integration perspective. After a short introduction of CBR we explain the seamless integration of these AI and SE approaches that has been achieved so far. We report on the benefits and impacts of such an integration that include the use of systematic evaluation and maintenance to improve an EB together with some relevant evaluation results. Finally, we mention some still open problems including the implementation of fully automated, software-agent-enacted EFs that enable learning through maintaining experience.

Project Organization

Plan
1. Characterize
2. Set Goals
3. Choose Model

Perform
4. Execute Project

Project Team

Project N

Cases
General Knowledge
Feedback

Experience Engineer
Experience Base

Experience Factory

Learn
5. Analysis
6. Packaging

Knowledge from past projects

Fig. 1. Experience factory
Experience Factory

EF is a logical and/or physical infrastructure for continuous learning from experience. It includes an experience base (EB) for knowledge storage and knowledge reuse. EF is primarily an organizational approach to learning from experience (organizational learning), that is, learning is achieved by humans. There are a number of different roles and responsibilities that have been invented to support this. A characteristic feature of an EF is the separation of learning and project organization, which is motivated by practical experience from software organizations [BCR94a]. EF bases on the quality improvement paradigm (QIP), a goal-oriented learning cycle for the EB-based improvement of project planning, execution, and analysis/learning [BCR94b].

Figure 1 describes the basic "working cycle" of an EF. The project organization, shown in the upper part of the figure, is responsible for project planning and project execution. These projects provide case-specific and general knowledge including feedback. The EF then is responsible to analyze such knowledge and to package it for reuse, that is, to represent the knowledge in a way that makes it useful and easy to reuse for the (projects in the) project organization.

The EF approach was already introduced by Basili, Rombach et al. From an EF perspective in the mid 1990s the basic approach was already introduced by Basili, Rombach et al. With NASA SEL a very successful and established application was available. In addition, there were also some other positive examples. Important problems in the mid 1990s were how to implement an EB, how the necessary processes for developing an EF/EB should look like in detail, as well as how experiments about implementation issues could be carried out.

Case-Based Reasoning

CBR is an approach to learning and problem solving based on past experience. A past experience is stored in the form of solved problems ("cases", case-specific knowledge) in a so-called case base. A new problem is solved based on adapting solutions of known similar problems to this new problem. For solving a new problem a query is submitted to a CBR system to retrieve the solutions of the most similar problems/cases in the case base. The query is considered as a potential new case. The classical view on a case imposes that it consists of a problem description and a described solution to this problem.

Recent applications have shown that this view is too restrictive. A more general view on cases is that a case consists of a characterization (a more or less structured set of information entities) and - optionally - one or more artifacts. The characterization part is mandatory and usually contains both formal parts, which can be interpreted by the CBR system, and informal parts, which can be understood by the user only. The characterization can be completed by links to any kind of artifact. Examples for artifacts are web pages, pictures, audio-visual media, or formally described pieces of knowledge. During the retrieval a part of the formal characterization of the new case is specified. The CBR system uses this for matching
against the characterizations of the other cases in the case base. The most similar cases are retrieved and parts of their characterizations and associated artifacts are combined to yield a proposed case. Applying the suggested artifact(s) associated with the proposed case results usually in a repaired case and validated and/or revised artifact(s). The repaired case with its applied artifact(s) can be stored as a learned case, merged with previously available cases, and/or only extracted parts of it are stored. These problem solving steps are also known as the CBR cycle and/or CBR process model [AP94]. They represent the first level of the CBR task-method decomposition model [AP94], which models on a certain level of abstraction how CBR systems work. It is based on components of expertise from [Ste90]. Figure 2 shows the CBR (sub)tasks arranged in a hierarchy. The respective CBR methods either define in what order to perform the sub-tasks of a task or how to perform a task that is not further decomposed (see [Alt97] for a detailed description). In addition, access to knowledge is required that is needed to perform some task including both contents (domain knowledge) and pragmatic constraints imposed on this domain knowledge.

CBR was introduced in the late 1970s and early 1980s in the area of cognitive science and AI by Schank and Kolodner [Sch82, Koh93] as an approach for modeling problem solving and learning of both humans and machines. In the late 1980s and beginning 1990s this led to a focus of knowledge-based systems on experience, that is, on case-specific knowledge [AP94, Bar87, Alt89, Aha99, Alt01]. Basing on the dynamic-memory-idea of Schank guaranteed the situatedness of the approach, which often led to a good user acceptance. Accordingly a number of commercial tools and many real-life applications were developed (e.g., [Alt95, Ber03, Wat03]).

Important problems in the mid 1990s were how to systematically develop a CBR system, how to operate it, how to integrate it into an industrial environment as well as how to evaluate it.

Integrating Case-Based Reasoning and Experience Factory

Since both approaches EF and CBR deal with learning from experience they are obviously related. To our knowledge the first who mentioned that CBR could be used for EF was Henninger [Hen95]. However, the key idea how to seamlessly integrate EF and CBR was first mentioned by Althoff and Wilke, and then further detailed by Tautz and Althoff [TA97, TA98, Alt98, Tau00]. The integration idea between EF and CBR is very simple but, nevertheless, was also very helpful: Interpret the methods within the CBR task-method decomposition model organizationally. That is, in the first place consider them as tasks to be done by some human being, and which maybe partially supported by a software system. One important aspect here is that CBR, as originally been introduced for modeling human problem solving and learning, can also be carried out by human beings. So, in principle it was possible to enact the CBR task-method decomposition model by human beings. As a consequence, CBR and EF/QIP became comparable (see Fig. 3). By this, the detailed knowledge on CBR processes elaborated in the CBR community provided a very good starting point for describing the EF processes ([AP94], [AA96], [Alt97]). Thus, human-based execution of tasks was the natural means to integrate CBR and EF/QIP. With components of expertise a framework for developing knowledge-based systems – here instantiated for CBR systems – was used for carrying out organizational learning. Of course, CBR was a first natural implementation technology for this framework, that is, could easily be used to automatically support the execution of the one or the other (sub)task.
For using this framework, Tautz [Tau00] developed the DISER method, a systematic, detailed method for building integrated CBR/EF systems supporting both human-based and computer-based task execution. Figure 4 shows the four main development phases (vision, version 1, pilot, running) as well as the main aspects of developing a CBR/EF system: Goals/existing knowledge and infrastructure, subjects, usage and filling, modeling, implementation, “going online”, and maintenance. Based on its EF part, DISER provided knowledge about how to organizationally embed an EF/CBR system in commercial environments. This included the EF roles that base on experiences with tasks and responsibilities from existing EF approaches:

- The **manager** provides resources, defines strategic goals and initiates improvement programs. He determines the structure and content of the case base and controls its quality.

- The **supporter** is responsible for documenting new experiences and supporting the project team. He collects and qualifies artifacts from the projects in accordance with the reuse criteria and the goals of the engineer. Upon request, he supports the project team in retrieving and modifying the experience knowledge.

- The **engineer** is responsible for packaging and analyzing existing experiences. Together with the manager, he identifies new reuse criteria and, based on that, acquires new cases. He analyzes the case base in order to detect (further) improvement potential.

- The **librarian** is responsible for technical aspects like setting-up and maintaining the case base, storing, and publishing new cases.

Again based on its EF part, DISER also included goal-oriented measurement and evaluation, an approach that is easily usable for evaluating CBR systems [NAT99, NF00]. From an AI perspective, the use of a systematic SE method like DISER led to an extension of the technology domain, that is, besides CBR also other search and repository technologies became usable. This new technology domain is called “Experience-based Information Systems” (EbIS). An EbIS is an information system that includes experiences (case-specific knowledge). In addition, the potential use of other (than CBR) AI technologies was identified [e.g., text/data mining: RA04]. Some of the mentioned advantages have meanwhile been confirmed by other research groups through the use of CBR in EF/SE as well as EF in CBR/AI (e.g., [Hen95], [She03], [Gom04] as well as [Kal00], [Bar02], [Ber03]).

From a SE perspective, the use of mature technology like CBR resulted in the introduction of systematic reuse into the EB development process, namely the software product-line approach based on some product-line architecture [Nic05] (see Fig. 5). Based on CBR technology a number of projects were carried out that brought to light some deficiencies of DISER concerning [Nic05]:

- phase models and development strategies for a better integrability on the software process side;
- solutions for “feedback loops” as well as experience life cycle models;
- solutions for relating different types of knowledge/experience each represented on a different level of granularity;
- rapid application development approaches for a „cheap start“;
- knowledge modeling approaches/guidelines for “scaling up”;
- scalability of the underlying knowledge technology;
- integrability of knowledge technology with traditional software system technology;
- supporting maintenance as a knowledge-intensive task;
- maintenance process;
- decision support for maintenance;
- acquisition method for maintenance knowledge;
- maintenance enactment support (for optimizing the maintenance process);
- business goal oriented method for running an EbIS;
- relating maintenance to the goals of an EbIS to guide maintenance with evaluation;

Fig. 5. Product-line architecture for EF/EB systems
• availability of an evaluation plan and maintenance knowledge already for the beginning of regular use for handling the to be expected continuous stream of experience.

Fig. 6. EbIS development process with DILLEBIS [Nic05]

Based on numerous projects the DISER successor DILLEBIS [Nic05] has overcome these deficiencies. Figure 6 describes, on a certain level of abstraction, how an EbIS can be developed using DILLEBIS. Based on learned “development patterns” rapid application development becomes possible. The knowledge acquired through the continuously running evaluation is directly used for maintenance [Nic05]. Figure 7 details this knowledge capturing process: In parallel to the deployed system a monitoring evaluation program is running. Then both is supporting maintenance the knowledge gathered for preparing evaluation and the achieved evaluation results. The deployed EbIS is developed based on the resulting tailored EbIS concept and the available reference architecture, schema guidelines, and components.

The main advantages of DILLEBIS include its integration into the respective work processes, feedback-based learning, evaluation-guided maintenance (see Fig. 7) as well as enabling a low-cost start in customer projects using EbIS patterns and a product-line architecture (see Fig. 5). The product-line architecture helped to introduce a number of variabilities with respect to search, storage, and maintenance techniques.

Benefits

Integrating EF and CBR resulted in various advantages. DISER/DILLEBIS generalized from CBR Systems to EbIS, a newly defined type of information system that contains experiences and is able to handle a (more or less) continuous „stream of experience“.

DISER/DILLEBIS explicitly considers that an EbIS is embedded into a “lived” process. Here DISER/DILLEBIS goes beyond other methods for CBR system development/maintenance (e.g., Inreca [Ber99], SIAM [RB00]). Such approaches strictly stay within the “CBR domain”. They do not open up for other AI technologies, and do not explicitly consider the task of embedding the developed software system into the respective business processes (e.g., maintenance approaches remain on the technical level only). DISER/DILLEBIS treats this business process with much more precision than these other CBR system development methods do. In contrast to other CBR approaches that mainly use classification accuracy oriented evaluation approaches from machine learning, DISER/DILLEBIS includes a goal-oriented measuring and evaluation approach for knowledge bases that is more flexible and more appropriate for evaluating complete software systems/components [NAT99, NF00, BCR94b].

Since DISER/DILLEBIS generalized from the “CBR domain” to the “EbIS domain”, one natural question is which technologies can contribute to EbIS development and how? One idea is using supplementing AI technologies that naturally can extend CBR, e.g., using machine learning techniques as part of the DISER method [Rec02]. Another idea is using other AI technologies that naturally can substitute CBR, e.g., ontology based knowledge modeling and retrieval, model based reasoning, and/or information retrieval.

Evaluation

In [Tau00] an experiment was carried out that proved that DISER basically worked. The EbIS approach (”using the EbIS”) was scientifically compared with the human-based approach (”talking to colleagues”).

The experiment focused on a typical scenario within a Fraunhofer institute. Someone is responsible for a new customer project and has to collect all important information about the customer as well as about similar projects and/or customers.

The experiment showed:

• Efficiency: The EbIS approach finds more useful guidelines and observations per time period (in terms of both effort and duration).
• Effectivity: The EbIS approach finds useful guidelines and observations not obtained by the human-based approach.
The experiment validated this in a statistically significant way. Thus, the result of the experiment was to combine the human-based and the EbIS approach. The participants in the experiment agreed: 28 out of 29 participants would apply both approaches in combination.

In the scope of [Nic05] a series of case studies was carried out where it was analyzed how well the modeling related parts of DISER/DILLEBIS were, especially with respect to maintenance. The goal was to evaluate the initial concepts of the EbIS development method. Two role playing games were carried out with seven students each with nearly identical design. The context was a company for developing CBR applications called X-CBR that is a few years old and can be viewed as an SME.

The task was following: X-CBR management decided that an EF on CBR projects should be developed, because this would enable X-CBR to better manage their knowledge on their core competence „CBR applications“.

Lecturers and students took over the roles of the EF and the X-CBR organization. Within the lecture the initial modeling of the EbIS was carried out. At the end of the lecture(s) a final questionnaire had to be filled out. In addition, a feedback meeting was carried out.

The following results were achieved:

- The initial modeling was completed within 180 min (2 lectures of 90 min each)
- The case study provided a lot of valuable experience for an initial workshop with a financial services company.
- During the case study it came out that DISER/DILLEBIS needed a phase concept which was introduced directly after the case study.

Further evaluation results can be found in [Tau00], [Nic05], and [Rec04]. The following overview in Figure 8 on relevant EbIS application projects is due to [Nic05].

<table>
<thead>
<tr>
<th>Status</th>
<th>#</th>
<th>Size of EbIS (#User, #Cases)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Small</td>
</tr>
<tr>
<td>Successful</td>
<td>8</td>
<td>SkE-Pilot, ISI, SLI-EB, IPGM, KM-FEB, CBR-FEB</td>
</tr>
<tr>
<td>Implementation</td>
<td>1</td>
<td>T-Com SR</td>
</tr>
<tr>
<td>Status unknown</td>
<td>1</td>
<td>Project A</td>
</tr>
<tr>
<td>#successful</td>
<td>6</td>
<td>1</td>
</tr>
</tbody>
</table>

Fig. 8. Overview on EbIS projects [Nic05]

The above mentioned EbIS methods showed their broad applicability through successfully occurring in numerous real-life projects. The listed projects are different in size and project type and thus show a certain “breadth”. An EbIS is considered to be successful if it is in usage for more than one year (status “accepted”) or its status is “deployable” and it is tightly integrated. An assumption for tight integration is the acceptance and correct usage of the tool that supports the business process.

Impact and Outlook: Planned Research Activities

The authors are convinced that the integration of SE and AI methods and techniques will be a very important future research task, and that maintaining a case/experience base is an important example for this.

We shortly reported about the integration of CBR and EF and about the application of software product-lines to EF/CBR systems (EbIS). The authors believe that software product-lines and EF/CBR have to be seamlessly integrated with a specific focus on the knowledge aspects. The more general case here is to integrate software product-lines with knowledge-based systems, which leads to what we call a “knowledge-line” ("Wissenslinie"). We plan to realize software architectures for such knowledge-lines based multi-agent system technology because this allows to have the flexibility that will be required for future information systems. CBR/EF will be used for implementing the respective agents, which leads to a hierarchical, mixed software-agent-human-based EF, where each agent is embedded in a “digital EF” and the multi-agent-based software product-line will be an EF of (digital) EFs. Such agents will learn from experience that is collected in their case base. Other software agents that take over the classical EF roles, continuously try to improve the content of the case base, for example adding/deleting new cases, improving the similarity assessment, or generating new/improved adaptation rules.

One major idea behind this approach is using knowledge management processes as a kind of specification of “learning processes“. By this we mean processes that contribute to achieve a higher degree of autonomy and automation through experience based improvement. These are then implemented as software-agent-enacted EF processes. Thus, for each software agent the EF processes would be completely automated.

How to integrate more knowledge into software product-lines and their supporting processes has been addressed by various authors [DM05, Sch04, Mut05, Bay05]. Some first ideas about knowledge-lines can be found in [DA04]. An example for using multi-agent system technology for CBR is given by [Pla05].

1 Further role playing games took place with similar design in other lectures/semesters.

2 For a motivation of such an integration see, for example, [Bib04] and/or [RA04].
References


