Structured Case Base Organizational Memories to Manage Experiential Knowledge

Stefania Bandini, Sara Manzoni, Paolo Mereghetti, and Carla Simone

Department of Computer Science, Systems and Communication
University of Milano-Bicocca
Via Bicocca degli Arcimboldi 8 20126 Milan - Italy
tel +39 02 64487835 fax + 39 02 64487839
e.mail:{bandini, manzoni, mere, simone}@disco.unimib.it

Abstract. The paper proposes a solution to one of the main topics related to the use of Case Based Reasoning (CBR) in the organizational learning context, that is the problem to define a suitable and effective representation of cases and Case Based Organizational Memory (CB-OM). The proposed solution focuses on case complexity related to the huge amount of involved information and to the need of incremental knowledge representation. Moreover, it allows to deal with another important aspect that must be taken into account: nowadays organizations are distributed environments. The proposed approach will be exemplified with an ongoing Knowledge Management project that exploits CBR to support the CB-OM management of the Business Unit Truck of Pirelli Tyres.

1 Introduction

Organizational Learning (OL) can be defined as the process of “detection and correction of errors” [1, 2]. Supporting the Organizational Learning means to enable the organization ability to learn faster than its competition. OL is focused on the way in which knowledge becomes embedded in network of relationships, information technology, business processes, performance management systems or product and services offering. In this view successful organizations are those that can assimilate and transform new ideas into action faster than a competitor.

Within this framework learning means to transform important experiences into knowledge, experiences that can belong to the entire organization or a part of it. In the latter case, information sharing inside the organization favors the learning process of the whole organization. This happens also when the related information is subjected to different interpretations and a local experience provides knowledge to the organization in different forms.

Besides, the usual phases (i.e. knowledge acquisition, information distribution, and information interpretation), another important concept must be related to the Organizational Learning process that is Organizational Memory [3]. The organizational memory can be considered as a set of heterogeneous structures within an organization that hold knowledge in different forms. For instance,
databases and information stores, work processes, procedures and product or
service architecture. The learning process can also be considered as a process
to develop the organizational memory [4], and the organizational memory is the
place where the experiences made by the organization reside.

The management of the organizational memory must focus on those com-
ponents strictly related to experiences. In particular, it must focus on those
experiences that could be considered a key factor for the success of the orga-
nization. Experience representation can be facilitated by case based organiza-
tional learning, meaning that relevant experiences can be captured in the form
of cases for reuse in a corporate experience repository, called Case Based Or-
ganizational Memory (CB–OM). Case Based Reasoning (CBR) [5, 6] can be ef-
ectively adopted in order to manage the organizational memory [7], and allows
experiential knowledge to be captured when models cannot or are hard to be
formalized. CBR has been applied to Knowledge Management (KM) in many
areas and there is a growing interest in this kind of approach (for instance, to
support organizational learning in software development [8], project planning [9],
and local government regulations [10]). Using the CBR paradigm in KM is still
a challenge from many points of view. In fact the challenge is to characterize
cases in the appropriate way to capture the dynamic aspects of innovation, to
define retrieving strategies that fit the needs of skillful actors involved in the
innovation process, and finally to collect cases from the every day activities and
the related technological supports.

The paper proposes a solution to one of the main topics related to the use
of CBR in the organizational learning context, that is the problem to define a
suitable and effective representation of cases and CB–OM. As it will be better
explained in the following, the proposed solution focuses on case complexity re-
lated to the huge amount of involved information and the need of incremental
knowledge representation. Moreover, it allows to deal with another aspect that
must be taken into account in the development of an organizational memory:
nowadays organizations are distributed environments. Organizational memory,
and thus CB–OM, could be distributed both logically (i.e. inside a company)
and geographically (i.e. networks of department or companies). The importance
to consider this aspect as central is stated by the huge number of researches and
applications dealing with the adoption of the CBR approach in distributed envi-
ronments (e.g. Distributed CBR [11], Collaborative CBR [12, 13] and Cooperative
CBR [14]).

The proposed solution to represent and manage cases and Case Bases for
organizational memory will be described in Section 2. Then the approach will
be exemplified with an overview of an ongoing KM project that exploits CBR
to support the CB–OM management of the Business Unit Truck of Pirelli Tyres
(Section 3).
2 Case and Case Base Management

In this section the result of researches focussed on effective case and Case Base representation for organizational memories will be summarized.

2.1 Structured Cases

One of the main aspects to be considered in the representation of cases for CB–OM is related to the heterogeneity of the involved information. A case is composed by a set of attributes that are organized in different categories and sub–categories according to their meaning. This organization leads to a tree structure where the children nodes can be viewed as a refinement of the concept expressed by the father node [15]. For instance the concept of “geometrical figure” (i.e. father node) can be refined in “triangle”, “rectangle” and “pentagon” (i.e. children nodes).

This case structure is particularly suitable for the knowledge representation task in the organizational learning context, where the information that builds up a case in the OM could be located in different places. Hence information sources can be both logically and geographically distributed. For example, in Figure 1 information concerning different sources (e.g. organization departments) is represented by different colors and it is organized as a different branch of the tree.

A hierarchical structure of cases has already been used and validated in other works [16]. However the peculiarity of this approach to structure cases is related to the flexibility of case representation that, when needed, can be easily modified (e.g. in order to include new acquired information). This is particularly useful when facing with OM where long knowledge acquisition campaigns can be needed (incremental knowledge acquisition) and the acquired knowledge must be incrementally represented in the CB–OM. Figure 2 shows two possible ways to update the case structure. A new child node can be added to an existing father node (Figure 2(b)) or a leaf node can be refined adding new nodes under it (Figure 2(c)).

Another important aspect related to Case Bases with huge amount of information concerns the level of information details needed to fill each case. Even if in our proposal each case is structured according to the same tree structure, not all cases have necessarily to contain information at the same level of details (e.g. some attribute values can be unspecified). This can be useful when the case structure is dynamic, some pieces of information are not available or too difficult to be provided. In the first circumstance, cases created before a case structure updating, may have no information associated to the new added attributes. On the other hand, the value associated to a father node can sometimes be omitted since the children nodes contain more detailed information (Figure 4(a)). On the contrary, sometimes information related to children is not available but the one associated to the father node can be enough.

Since this case structure allows the Case Base to contain cases described at different level of details, a suitable similarity function has to be studied. The
Fig. 1. An example of tree structure to represent cases in an OM. Information concerning different sources (e.g., organization departments) is represented by different colors.

easiest way to compute the similarity between two cases with different level of details, is to use only those attribute values that are present in both cases. Obviously this is not an effective solution because when a case is newer than another, all the knowledge acquired and represented in the newer case is lost. Similarly it happens when different information is available for the two cases description. As shown in Figure 3, if it has been decided that the nodes involved in the similarity computation are the circled ones, a lot of available knowledge is not used because only the two black circled attributes have associated values in both cases.

A suitable solution that has been investigated and used in [17] proposes to derive the set of required node values. In this proposal, since the children nodes are refinements of the concept expressed by the father node, the idea is to encapsulate in the father the knowledge available in the children nodes. This could be done associating to each father node a function that derives its value using, for example, the values of the children and a set of weights. The effort and difficulty to design this function depends on the concepts expressed by the node. As a simple example it is possible to think to the concept of “sweetness”
of a cake. There are different ingredients in a cake recipe that make it sweet; for instance sugar, saccharine, jam and so on. As shown in Figure 4(b), in this situation it is possible to derive the value associated to the sweetness simply calculating the weighted sum of all the ingredients that bring sweets to the cake. In this example it has been assumed that the sugar brings a sweetness equal to the hundred percent of its weight, the saccharine the eighty percent and the jam the twenty.

### 2.2 Structured Case Base

Since in many circumstances a new case is derived by adaptation from others already present in the Case Base, a hierarchical structure of the Case Base can be considered as an appropriate choice that allows to efficiently store them. The user of the CBR system can choose to append the new retained case under the case it derives from or to create a new tree root (see Figure 5).

The main advantage of this approach is to allow to track the sequence of adaptation phases that had lead to the solution achieved.

A hierarchical Case Base structure also maximize the efficiency of case retrieval and update. The set of cases contained in the Case Base has to be sufficiently wide to guarantee the retrieval and reuse of previously stored and useful cases, but at the same time the case base organization has to be sufficiently compact to allow the storage of a large number of cases with a limited memory occupancy. In a set of tree structures in fact only differences from the father node
Fig. 3. Case similarity computation. The circled nodes represent the ones that have to be involved in the similarity computation. The values of the grey nodes can be derived in order to maximize the number of nodes to be involved.

and information about type of adaptation must be stored in order to represent son nodes derived from it.

3 Application in a KM Project

The work described in this paper takes its motivations from the P–Truck project, whose main aim is the development of a KM system to support the Business Unit Truck of Pirelli Tyres in the design and manufacturing of truck tyres. In this paper we take tyre manufacturing as the reference domain, since we adopted the CBR paradigm to build a computer system supporting the design of rubber compounds in the production of tyres dedicated to car racing and trucks (P–Race [18] and P–Truck [19], respectively). However, the approach we are going to illustrate could be applied to other manufacturing domains.
The design approach of the solution proposed to Pirelli Tyres was deeply knowledge based. A knowledge acquisition campaign and direct observation of Business Unit working activities have been conducted. The result of this design phase is an integrated KM system in which different knowledge based and service modules exploit the Pirelli organizational memory to support the main design and productive activities of the Business Unit Truck of Pirelli Tyres. In particular a knowledge based module, based on a previously introduced model for the design of rubber compounds [20], supports compound designers in the definition of product specification. Moreover, rule based reasoning and CBR are exploited to support the design of some tyre production processes for rubber compounds (i.e. mixing of rubber compound ingredients and tyre vulcanization, respectively). As reported in [19], the CBR paradigm has also been applied to support the tuning of production processes, a fundamental task strictly related to the experience of people involved in manufacturing.

### 3.1 The KM Context

The life–cycle of a product in manufacturing companies consists of several phases that can be divided into two main steps: design and production. The design step is usually triggered by the need of innovation defined by marketing strategies oriented to answer requests emerging from the market or from new needs induced in it [21]. The goal of the design step is to define product and production process specifications. Product and process specifications define the constraints the production process has to respect in order to make products pass the quality control and be successfully distributed on the market. In the reality of manufacturing another step has often to be considered in the design–production cycle (i.e. production tuning). This need is due to the possibility to encounter unpredictable problems during the production step. Like any sort of specifications, product and process specifications cannot be exhaustive of all the possible details characterizing the instantiations of the production
A case to be retained to the structured Case Base can be added as a child node to the one it derives from, or it can originate a new tree and be added as root node.

process. In fact, design is usually a process located where the core design competencies reside, while production is increasingly distributed in different plants to take advantage of local facilities or to make production scheduling more flexible to the market needs. Each instantiation of the production process happens in a different contingent situation where, for instance, raw materials can show even slightly different properties. Different production context might lead to variable or unexpected results on the final product and to the consequent modification of the product structure or composition [22].

The tuning of product and process specifications in order to adapt them to take care of structural properties of the local plant, of unanticipated events or non uniformity in raw materials may be necessary in order to guarantee the necessary uniformity of production results. Thus, besides standard process design, process tuning has sometimes to be performed during manufacturing. Although this critical situations (also referred as, production anomalies) are not too frequent, they are extremely crucial since the time available to solve them and to continue the production under the new circumstances is very short. Search for a quick solution justifies to record anomalies connected to problems that have already been experienced and solved in the past.
3.2 The CBR Solution

Knowledge engineering with Pirelli technicians involved in product and process tuning (i.e. people working in the production line, typically experienced workers who are however very busy or involved in off-line production activities, like product testing) revealed that the nature of their problem solving activity is mainly episodic, that is based on previous similar situations: they often reason about past cases in order to solve the current one. Thus, CBR has been chosen as a suitable approach in order to captures the episodic knowledge characterizing most of the reasoning activity of people involved in production tuning, and to support the dynamical process of experience growth and knowledge creation through incremental learning. A dedicated Case Based Organizational Memory (CB-OM) has been designed for the Business Unit Truck of Pirelli Tyres. Interested readers can refer to [19] for more details about the content and structure of the Pirelli CB-OM.

Flexibility of the case structure was an fundamental requirement for the Pirelli CB-OM related to the fact that a huge amount of information is usually involved in the decision making process of people that perform this crucial activity. Knowledge involved in the tuning process is related to products (e.g. in terms of marketing requirements, product specification, composing subassemblies, sub-
assembly specifications), production processes (e.g. timing, number and order of each production phase), and to production context (e.g. peculiarities of the local production plant, set of available machineries and raw materials). The heterogeneous nature and huge amount of involved knowledge have obliged P–Truck system designers to plan an incremental knowledge acquisition campaign that is being performed with a team of about twenty Pirelli technicians that are dedicated to different activities for tyre design and production.

According to this incremental approach to knowledge acquisition, the chosen case representation had to be flexible enough in order to allow an incremental enrichment of case structure when new information is collected. The approach presented in Section 2 has been adopted to this aim. The CB–OM that has been developed and integrated into the overall P–Truck system architecture collects a set of production cases. As shown in Figure 6, each case content has been organized into product, production processes, production context, and outcome. Then, each node has been specialized into more details (for instance, the product part into product use and recipe).

The advantage of this approach to case structuring has been validated within the P–Truck project, during the development of a CBR module to support the design of the vulcanization phase that characterizes tyre production (vulcanization, also referred as curing process is a chemical process of treating crude rubber to give it useful properties, such as elasticity, strength, and stability). A CBR module has been developed to support Pirelli technicians in the vulcanization design. The tree structure of vulcanization cases has allowed to integrate this Case Base to the P–Truck CB–OM (i.e. the vulcanization process node has been detailed according to the structure of vulcanization cases) and thus experiential knowledge about curing processes can be shared among P–Truck modules devoted to other tasks (e.g. production tuning).

Moreover, as in the most of manufacturing contexts, Pirelli design and production activities are the result of interactions among multiple and heterogeneous competence (i.e. functional distributedness) that can sometimes be also geographically distributed. Thus, the Pirelli CB–OM has been designed in order to allow the management of a distributed Case Base: data sources for the CB–OM are located in different places inside the Pirelli information system and the resulting CB–OM can be remotely accessed, retrieved, updated, created, and so on.

3.3 P–Truck Technical Solutions

The P–Truck system has been designed and developed according to a Component Based approach. P–Truck modules have been designed in order to satisfy all the identified user and non–user requirements (i.e. integration of heterogeneous and interacting components, independent development of components, extendible architecture, separation between business and presentation logic, portability, integration into the existing information system and Web orientation).

System requirements lead to the choice of the Java 2 Enterprise Edition (J2EE [23]) platform as technological solution to develop the P–Truck system.
Currently the J2EE technology supports component-based development (Enterprise Java Beans, EJB), integration between systems (it complies with CORBA and Remote Method Invocation, it supports XML, and so on), database accessibility and web-based user interfaces development. Figure 7 shows an example of the web based user interface of a module of the P-Truck system. Moreover, the designed solution allows the integration of several knowledge based modules based on different approaches (e.g. rule based and Case Based Reasoning) into an already existing and used company information system (made of a huge number of databases and applications).

Fig. 7. A screenshot of the web based user interface of the P-Truck system. The shown interface allows the access to the knowledge based module to support the design of rubber compounds (P-Truck Formulation module).

4 Future Work

The proposed approach to structure cases in CB-OM that in this paper has been exemplified within the P-Truck project will be applied to manage OM to other manufacturing contexts. However other activities have to be done before the end of the P-Truck project. For instance, suitable security policies have to be
studied and integrated into the P–Truck system in order to protect the privacy of involved information.

Moreover, the design and development of a training portal dedicated to new employees of the Business Unit Truck has been planned. P–Truck knowledge based modules will be exploited within this portal as simulators of design and tuning activities.

From a software engineering point of view, possible advantages in moving from the Component Based to the Agent Oriented approach will be investigated. Moreover, the Peer–to–Peer model will be investigated as a possible solution to manage the dynamic nature of OM, and if suitable and effective, it will be adopted with the chosen architectural approach.

References