Case Base Engineering for Large Scale Industrial Applications

Kalyan Moy Gupta
Atlantis Aerospace Corporation
1 Kenview Blvd, Brampton, ON L6T 5E6
Canada
kgupta@atlantis.com

Abstract

Case-based reasoning (CBR) has emerged as a promising technology for decision support. However, its application to large scale industrial problems has been negligible. We believe that the lack of usable problem-solving cases is the major factor limiting the wide spread use of CBR in complex industrial environment. We propose a framework of methodologies called case base engineering (CBE) that can be used for the development of high quality case bases. The framework is presented in light of our CBE experience in application to plastics extrusion facility and aircraft maintenance. Finally, we identify issues for future research.

Introduction

In recent years, case-based reasoning (CBR) has emerged as a promising technology for decision support (Allen 1994). CBR systems aid a decision maker (DM) by retrieving past experiences and using their outcome to solve new decision problems (Kolodner 1993). Each nugget of experience can be encoded as a problem description and solution pair, that is, a case.

CBR technology is attractive to the industry because it is easily understood, there is apparent availability of problem-solving cases throughout the organization, and its capability to reason with partial input and grow by simply adding cases. Consequently, a number of firms have adopted the technology and successfully developed diagnostic and planning applications (Allen, 1994). Despite a number of vendors marketing CBR products, application development for industrial equipment maintenance support, for example, has been negligible. What factors limit CBR system development? We believe it is the unavailability of case data in a readily usable form. The problem-solving data is scattered throughout an organization and needs to be structured, distilled and synthesized to produce usable problem-solving cases. This observation is inconsistent with the notion that CBR systems hardly require any knowledge engineering (e.g., Simoudis 1992).

In this paper, we describe a framework called "Case Base Engineering" (CBE) that can be used to build case bases. We discuss CBE in the context of CBR application to support the maintenance of complex industrial equipment. The remaining paper is structured as follows: characteristics of industries operating complex equipment is presented followed by the CBE framework. Finally, we discuss results of our CBE experience and identify issues for further research.

Industrial Equipment Maintenance Environment

A typical industrial facility operating and maintaining complex equipment has the following characteristics:

1. *Domain Complexity:* A production-line is a complex assemblage of components from different manufacturers and is highly configural. For example, a plastics extrusion facility comprises flake transfer system, blenders, conditioners, an extruder, transfer lines, dies, and winders. Despite the overall similarity of production lines, no two production lines are entirely alike with regard to their components, processes, and products.

2. *Problem frequency:* The average operating life of components in a production line is long. The equipment is highly reliable. It breaks down infrequently and in unforeseen ways. Therefore, it is very difficult for any one individual to gain significant expertise throughout the domain. For example, a plastics extrusion facility operates equipment that is 10-35 years old and has been retrofitted to keep up with the technological changes..

3. *Time Constraint:* The problems are solved under strict time constraints because production shutdown or loss of service is extremely costly.

4. *User Expertise:* The equipment is operated and maintained by a pool of well-trained users with a significant level of expertise in their respective areas.

5. *Collaborative problem solving:* Often, problems transcend different areas of expertise and require collaborative problem-solving. For example, extrusion equipment problem may require the collaboration of

operators, mechanics, electricians, and process engineers.

Due to the complexity of large scale equipment domain and the interdependencies between various users, the problem-solving information is scattered around the organization. Furthermore, the scattered information sources may not contain the problem-solving information needed by the users. The key reason for this deficiency is that data are not collected with the intention of re-using it for problem-solving. Consequently, a well defined process of case-base engineering is required to build cases that brings together the necessary information from different sources.

Case Base Engineering

The effectiveness of the decision support offered by a CBR system is dependent to a large extent on the quality of problem-solving information contained in cases. Good quality problem-solving information requires a synthesis and reconstruction of problem-solving cases from dispersed but related data sources.

We propose a process model of CBE that comprises elements from knowledge engineering (Scott et al., 1991) and publishing. The process includes the following four phases (See Figure 1):

1. *Analysis*: The preparatory phase when the scope of case base is defined, information sources are identified, and the case base size is estimated.

2. *Design*: The case structure is designed.

3. *Development*: Cases are developed via the processes of extraction, representation, verification and testing.

4. *Deployment*: The case base is put into active usage and maintained there after.

Analysis

Define Scope. The scope statement specifies the kinds of problems that the case base will cover. For example, a scope statement in a plastics extrusion facility could be simply stated as "Extruder problems", or alternatively, "Extrusion line problems". Limiting the scope to extruder problems may simplify the CBE activity, however, may turn out to be too restrictive for a mechanic who attends to all the problems in a line, and furthermore, problems with the extruder often lie in components associated with it. Therefore, case base scope statement is incomplete without the specification of intended users. For example, a plastics extrusion problem may require users from the technical services, mechanics, electricians, and process operators to work together. Methods such as hierarchical task analysis, information requirement analysis, work flow analysis can

be used to determine the potential users and their information needs. It is important to assess technical skill level of the user groups to avoid including overly simple problems in the case base. This can be done by structured questionnaire surveys. Users view problem-solving as integral to their task. They interact with a variety of information systems to perform these tasks. Therefore, the requirement analysis should also determine the integration needs of the CBR system.

Figure 1
Case Base Engineering: Process, Methods, and Resources
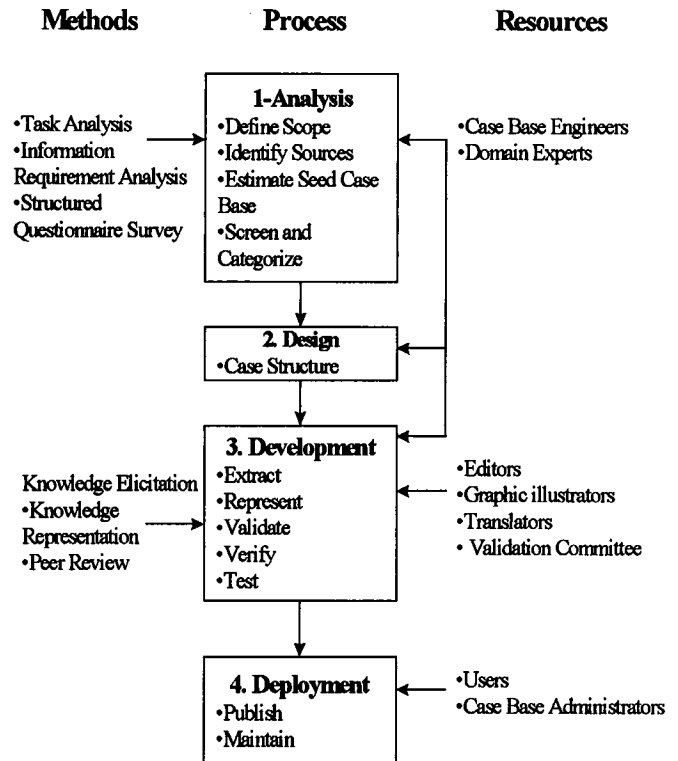
Identify Sources. The work flow analysis and the information requirement analysis reveal kinds of information that users require to solve problems and where details of any such activity might be recorded. For example, mechanics in an extrusion facility record service calls in logbooks and work-orders. The events of a complex problem are recorded in a total quality investigation report. Similarly, in the airlines, problem-solving activity is recorded in documents such as the pilot report, maintenance activity report, and in-service activity report. In addition, standard procedures and specific instructions are available in documents such as safety manuals, overhaul instructions, and operating manuals.

Two attributes of these potential information sources are critical to CBE. These are: (1) storage format and (2) quality of information. The information source records can be electronic or paper documents. When records are electronically stored, they lend themselves to a variety of searching and clustering tools to aid the case extraction processes, while case extraction from paper records is entirely manual. Of greatest importance is the quality of information recorded. The problem solving elements such as observations, actions, decision outcomes are the most important. For example, a total quality investigation report in the plastics domain contains the relevant details. However, most often the relevant problem-solving information is not recorded. For example, an airline maintenance record may include part removals without any significant observations.

Estimate Seed Case Base. The case base is a growing pool of problem-solving knowledge. New cases are acquired when the system is unable to provide adequate decision support. However, in an industrial application, to begin with, the CBR system needs to include a minimum number of cases so that the user finds it attractive enough to consult it. The case base with this minimum number of cases is called the seed case base. An important issue is determining the size of the seed case base.

Screen and Categorize. In industrial applications the seed case base may be quite large. To manage the CBE effort, the problems need to be categorized. Natural categories may be available depending on the domain. For example, airlines has problem categories such as avionics and landing gear. The source records are then screened and categorized for further processing.

Design Case Structure

The information elements of the case can be determined by the requirements analysis and task analysis. For example, an extrusion facility mechanic identifies the faulty part or a process upset based on observations and plans the repair. The repair plan may include resource assignments such as a pipe fitter or control systems engineer. Furthermore, when the cases are used for training, additional items need to be included, for example, a detailed explanation of an obscure problem.

Development

Extract. Extraction is the process of completing all the information elements in the case structure. This process is performed by case base engineers with close cooperation with domain experts. As described earlier, the source documents may not include the key items of problem-solving and include unrelated or discordant information. The case base engineer consults domain experts to separate the information from the noise and ensures the logical consistency and completeness of each case. The missing items need to be identified and elicited form domain experts by an interview process. The industrial domain experts such as mechanics and engineers have difficulty communicating and articulating their observations, conclusions and instructions. Therefore, the case base engineer re-expresses the contents to improve the clarity of information presentation. The engineer also ensures that the case includes references to additional problem-solving information such as technical drawings and manuals. Additional graphical illustrations may be needed to explain the events in a case. Therefore, besides domain experts, this step requires editors, translators, and graphic artists. This is the step where knowledge is created, synthesized or distilled from various raw document sources and human problem solvers with particular attention to its reuse and communication.

Represent. Representation involves transformation of extracted cases and encoding them to ensure their proper retrieval by the CBR system. Consequently, the process is dependent on the representation and retrieval methods used by the CBR system. Nonetheless, general principles of representation can be borrowed from the area of knowledge representation. For instance, cases need to be generalized to the extent possible. For a twin engine aircraft, a left engine startup problem is applicable to the right engine, and therefore, the case attributes need not include the engine location. Such scenarios occur throughout the maintenance application domain. Therefore, suitable representation methods and guidelines can be identified for these scenarios.

Validate, Verify and Test. As the extraction and representation process involve a transformation of raw information by different resources such as case base engineers, domain experts, and graphic illustrators, it is essential that the case be reviewed for its validity and consistency with the events in the domain. Methods such as inspection testing and peer review, as practiced in software engineering and publishing, can be used. Setting up of a validation committee comprising experts from different user groups is critical for ensuring the high quality of cases. Next, cases are tested for retrieval and representational errors.

Deployment

After a case has been represented and tested, it is published for active usage in problem solving. At this stage, the users may provide feedback about the validity and the content of cases.

A case base is an evolving collection of cases. New cases may be added as a result of equipment modifications and the stored cases may need to be modified or removed

as a result of these changes. The case base administrator manages the requests for changes to maintain the high quality of the case base.

Discussion

Thus far, we have used the CBE process model to build three seed case bases, all in the area of complex equipment maintenance. There are two principal benefits derived from the use of this model. First, like the software engineering process models, the CBE process model enhances the communication between those undertaking the CBE activity and the users of the case base. Second, as the model reveals the need for various resources such as validation committee and graphic illustrators at different stages of the CBE activity, it serves as a basis for resource planning and allocation.

The observations from our CBE experience in the area of complex equipment maintenance can be summarized as follows:

1. Although organizations have means and procedures to collect problem-solving information, much of these information sources do not contain re-usable good quality problem-solving information. The two principal reasons for this deficiency are: a) The problem-solvers function under strict time constraints and as a result do not record them, and 2) Problem-solvers have considerable difficulty verbalizing and articulating the relevant information. Consequently, much of the important problem-solving information does not get recorded in these information sources.

2. The case extraction step of the development phase is critical, where the quality knowledge is created. It is also the most effort intensive. A forum of experts comprising individuals from different user groups is essential to the overall success of CBE. Organizational process changes and directives are needed to enable this collective problem-solving and knowledge creation activity.

3. Representation rules can be identified that apply to the maintenance of complex equipment.

Various methodologies can be used to improve and support each phase of CBE. The potential areas of research include the following:

1. Methodologies to assess the domain complexity in order to determine the size of seed case base.

2. Develop measures of information quality.

3. Techniques such as datamining and knowledge discovery to support the extraction and representation steps of the case base development phase (e.g., Fayyad et.al, 1996).

4. Methodologies to support the knowledge elicitation process. In particular, investigate the frameworks available in the area of linguistic communication such as structure of argumentation and explanation (e.g., Maybury,1992).

Acknowledgments

References

Allen, B., 1994, Case-Based Reasoning: Business Applications, *Communications of the ACM*, 37(3): 40-42.

Fayyad, U., Shapiro, P.G., Symth, P., 1996, The KDD Process for Extracting Useful Knowledge From Volumes of Data, *Communications of the ACM*, 39(11):27-34.

Kolodner, J.L, 1993, *Case-Based Reasoning*, San Mateo, CA:Morgan Kaufman.

Maybury, M.T., 1992, Communicative Acts for Explanation Generation, *International Journal of Man-Machine Studies*, 37(2):135-172.

Scott, A.C., Clayton, J.E., and Gibson, E.L., 1991, *A Practical Guide to Knowledge Acquisition*, New York, New York: Addison-Wesley.

Simoudis, E., 1992, Using Case-Based Retrieval for Customer Technical Support, *IEEE Expert*, 7(5):7-11.