An Architecture for Reusing Problem-Solving Components

Dieter Fensel\(^1\) and V. Richard Benjamins\(^2\)

\(^1\) University of Karlsruhe, Institute AIFB, 76128 Karlsruhe, Germany, dieter.fensel@aifb.uni-karlsruhe.de, http://www.aifb.uni-karlsruhe.de/WBS/dfe

\(^2\) Dept. of Social Science Informatics (SWI), University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands, richard@swi.psy.uva.nl, http://www.swi.psy.uva.nl/usr/richard/home.html

Abstract. Developing software by selecting, adapting, combining, and integrating existing components instead of starting the system development process from scratch has become a key factor in economic software development. However, such a development process has to deal with four problems: First, components must be selected. Second, components must be adapted because they neither fit precisely the task that should be performed nor do they necessarily fit well to other selected components. Third, components must be combined and their interaction must be established. Fourth, it may be necessary to decompose complex problems into smaller subtasks for which components can be found. In this case, a general system frame must be established that enables to form an integrated systems out of separate components. In this paper, we present our means to deal with these problems: broker, adapter, connectors, and task structures. Although we discuss our approach in the context of problem-solving methods, some concepts are also applicable to software components in general.

1 INTRODUCTION

Knowledge-based systems are computer systems that deal with complex problems by making use of knowledge. This knowledge is mainly represented declaratively rather than encoded using complex algorithms. However, human experts can exploit knowledge about the dynamics of the problem-solving process and such knowledge is required to enable problem-solving in practice and not only in principle. Making this knowledge explicit and regarding it as an important part of the entire knowledge contained by a knowledge-based system is the rationale that underlies problem-solving methods. Problem-solving methods describe this control knowledge independently from the application domain thus enabling reuse of this knowledge for different domains and applications. It describes which reasoning steps and which types of knowledge are needed to perform a task. Problem-solving methods are used in a number of ways (see e.g. [Chandrasekaran et al., 1992]): as a guideline for acquiring problem-solving knowledge from an expert, as a guideline for decomposing complex tasks into subtasks, as a description of the essence of the reasoning process of the expert and knowledge-based system, as a skeletal description of the design model of the knowledge-based system, and as a means to enable flexible reasoning by selecting methods during problem solving. Meanwhile, problem-solving methods are used by nearly all frameworks for knowledge engineering. Libraries of problem-solving methods are described in [Benjamins,
1995], [Breuker & Van de Velde, 1994], [Chandrasekaran et al., 1992], [Motta, 1997], and [Puppe, 1993]. Some of these libraries provide textbook-style descriptions of reasoning strategies and others provide implemented software components. The former support the system development process during conceptual modeling and specification while the latter also provide support during implementation.

However, serious bottlenecks exist that hamper the usefulness of most of these libraries because they provide very limited support in selecting, adapting, combining and integrating problem-solving methods. That is, they are collections of methods but not sophisticated libraries. Therefore, trying to use these libraries produces a number of problems: the selection of a method from a library is only weakly supported and remains a kind of “magic trick”. Often, problem-solving methods fit only approximately to particular domain- and task-specific circumstances. They fit in principle, however, significant adaptation would be necessary to actually use them as an implemented component. Problem-solving methods hardwire assumptions concerning the way in which they interact with their environment (their communication style) and this way may not match the software or task environments they are applied in (an interoperability problem). Complex problems may require the combination of different methods, possibly from different libraries, e.g., if libraries are incomplete.

Our contribution deals with these problems. We provide a general architecture of a library of problem-solving methods that supports the reuse process by four different means. A broker is described that enables the selection of methods from different libraries assuming that (1) these libraries are made available via wide-area networks like the WorldWideWeb and (2) are annotated with ontological information that enables their informed retrieval. Adapters are presented as means to enable the adaptation process of methods to tasks and domains. They are the key concept for keeping libraries manageable (i.e., limiting the number of elements) and useful (i.e., providing elements that either fit perfectly to given domain- and task-specific circumstances or provide significant support in their generalization or specialization). Connectors are —in software architectures— a means to implement the interaction of components (i.e. their interoperability) and we propose their use for problem-solving methods. They provide an abstraction mechanism in cases where assumptions on the communication style may conflict. Finally, we show how task structures serve a twofold purpose: they decompose complex functionalities into functionalities that can be handled by components, and they provide the integrated system frames for loosely coupled components.

This paper is organized as follows. A broker that handles the selection process of components is described in Section 2. Adapters are discussed in Section 3 as means to enable simple and reusable adaptations of components. Connectors are briefly introduced in Section 4 to provide the connections between components. Section 5 shows task structures as means to decompose problems and to provide integrated system frames. Finally, conclusions are provided in Section 6.

2 BROKER: COMPONENT SELECTION

“You must find it before you can reuse it!” [Will Tracz]

Software and knowledge reuse via networks is becoming an increasingly popular topic. Developing intelligent agents that provide reasoning services based on problem-solving methods (cf. [Benjamins et al., 1998]) introduces an interesting application framework and raises interesting research issues. In the following, we will sketch two key issues in designing such reasoning services. First, we discuss the role of ontologies and second we provide the core of a brokering service.

Ontologies have been proposed to enhance knowledge sharing and reuse (cf. [Fridman Noy &
An ontology provides a conceptualization, which can be shared by multiple reasoning components communicating during a problem solving process. Using ontological engineering for describing problem-solving methods provides two important benefits with respect to reuse. The resulting method specification is grounded on a common, shared terminology and its knowledge requirements are conceptualized as ontological commitments [Gruber, 1995]. Currently, ontologies are mainly used to formalize domain or common sense knowledge. However, there are recent proposals and initiatives ([Benjamins et al., 1998], [Benjamins & Fensel, 1998], [Chandrasekaran & Josephson, 1998]) that employ ontologies to formalize problem types and problem-solving knowledge (i.e., methods). For example, the Knowledge Annotation Initiative of the Knowledge Acquisition Community (KA)² [Benjamins & Fensel, 1998] aims at building a consensual ontology which can be used to describe the various research groups and their results in knowledge acquisition. Part of this ontology is concerned with problem types (i.e., tasks) and problem-solving methods. In such initiatives, where consensus is of major importance, choosing appropriate names is essential.

Another example is to provide axiomatic descriptions of problem-solving methods to support (semi-) automatic classification and selection of methods. Examples of such efforts include [Fensel et al., 1997] where ontologies are used to describe the task parametric design, the problem-solving method propose & revise, and their mapping resulting in a task-specific ontology for the method; and [ten Teije, 1997] who defines a task ontology for diagnostic problems.

Providing different libraries of problem-solving methods via the net requires a broker that mediates between customers and providers of problem-solving methods (cf. [Benjamins, 1997], [Fensel, 1997a]). The general layered architecture of such a broker is depicted in Figure 1. The broker has to communicate with clients via clients systems. A client is someone that has a complex problem but can provide domain knowledge that describes it and that supports problem-solving. The providers are developer teams of problem-solving methods. Their provider systems are annotated libraries of problem-solving methods. The problem-solving methods are implementations that solve complex tasks by using domain knowledge for defining and solving the problem. Their annotations are necessary to support their selection process and the communication process with the methods. The core of an intelligent broker for problem-solving methods consists of an ontologist that supports the selection process of problem-solving methods for a given application. Basically it has to provide support in building or reusing a domain ontology and in relating this ontology to an ontology that describe generic classes of application problems. This problem-type ontology has to be linked with problem-solving method-specific ontologies that allow the selection of a method.

The broker uses three different ontologies: domain ontologies (D_O), problem type ontologies (PT_O), and problem-solving method ontologies (PSM_O). These ontologies provide the contents of the communication processes between the different elements of the broker (cf. Figure 1). Because this communication includes the exchange of knowledge, KQML [Finin et al., 1994] is suitable as communication protocol. The different information flows (see Figure 1) are concerned with the process of finding appropriate problem-solving methods, and can be distinguished by the contributing agents and the direction of the communication. In total, four different communication flows can be identified (see Figure 1):

- **Sending a request from the client to the ontologist.** Terms of the domain and problem ontologies are the content of the message. The client interface uses a domain ontology to guide the interaction process with the client and it sends selected expressions to the ontologist.

- **Sending a negotiation from the ontologist to the client interface.** The ontologist might need further clarification from the client before it can finish the selection process of
problem-solving methods. This clarification may ask for more precise definitions of terms and their relationships necessary to derive an element of the problem ontology from an element of the domain ontology.

- **Sending a query from the ontologist to a provider interface.** After having translated the domain-specific terminology into a problem-specific terminology the ontologist has to derive an expression in a problem-solving method-specific ontology. Then, this expression is passed as a query to the provider interface.

- **Sending a response from a provider interface to the ontologist.** The response of a provider interface may have two forms: providing a simple yes that the required service can be provided or the wish to introduce further assumptions on the problem that make it tractable and/or the introduction of requirements on domain knowledge that has to be provided by the client.

Most of these information flows can be automated. However, the translation of a domain ontology in an ontology describing problem types has to be done semi-automatically for each new domain. The introduction of assumptions that introduce requirements on domain knowledge or that weaken the task (cf. [Benjamins et al., 1996], [Zaremski & Wing, 1997]) cannot be done without user interaction. The mapping between problem type and problem-solving methods has to be done by the providers of these method libraries. Therefore it is already established before the broker uses these libraries. The mappings needed during runtime of the method can be derived from the links between the different ontologies. These mappings are implemented by adapters that will be discussed next.

### 3 ADAPTER: COMPONENT ADAPTATION

Adapters are used to mediate between problem definitions, domain knowledge, and problem-solving methods. In the following, we will explain with an example why adapters are required for this purpose and what their implications are.

![Fig. 1. The layered architecture of the broker.](image-url)
3.1 Adaptation for Reuse

Chronological backtracking defines a very generic problem-solving method (i.e., a search strategy) that does not make any assumptions on the structure of nodes in the search space nor on the elementary transitions between states. [Eriksson et al., 1995] describe several adaptations of this generic search procedure for specific tasks. Chronological backtracking is adapted to a method for one-player board games by adding code that refines state transitions and data structures of the method. The resulting boardgame method adds two refinements which introduce commitments on the task and the required domain knowledge:

- the initial, intermediate and final states describe board positions, and
- the knowledge that is required to define state transitions is defined in terms of moves in games.

However, this adaptation of chronological backtracking is not kept separate in [Eriksson et al., 1995]. Therefore, the implementation of chronological backtracking is no longer available for defining a problem solver for, e.g., planning problems. Moreover, the adaptation to the board-game problem cannot be used to adapt, e.g., a local search procedure to board games. Separating the descriptions of search procedures and their refinements provides reusability in two dimensions:

- search procedures remain available for different tasks and domains, and
- task- and domain-adaptations remain applicable for different search strategies.

In general, different search methods can be applied to the same task, and [Beys et al., 1996] conclude that problem-solving methods should be represented in a task-independent way. However, their task-specific (i.e. adapted) versions (for domains and tasks) provide much more support in developing systems and in acquiring the domain knowledge than the task-neutral versions. An easy way to overcome this dilemma is to externalize their adaptations through (external) adapters. Adapters were originally proposed in [Fensel & Groenboom, 1997] to link problem-solving methods and tasks. However, they can also be used to externalize the specialization of problem-solving methods. The use of adapters in this way provides three variants of reuse that overcome the shortcomings above mentioned:

- The search procedure can be reused for different problems because the adaptation is kept separate.\(^1\)
- The adaptation can be reused for different search methods because it is kept separate.
- The combination of search procedure and adapter can be reused as a strong problem-solving methods (i.e., as a method with many task- and domain-specific commitments) in cases where its hardwired assumptions fit.

The adapter concept is therefore essential in preventing combinatorial explosion. Already a textbook style description of problem-solving methods in [Breuker & Van de Velde, 1994] provides hundreds of methods. Implementing these methods adds additional level of details and therefore distinctions. Externalizing adaptation is the key factor in component-based development of problem-solving methods. A simple illustration should clarify the point. Assuming \(n\) search strategies and \(m\) problem types. Without adapters \(n \times m\) components would be necessary to cover all these combinations. By using adapters only \(n+m\) components are necessary. Without adapters one ends up either

- with too many different components \((n \times m)\) leading to too much implementation and retrieval effort, or

---

\(^1\) This implies more than just keeping a copy of the old implementation of the search procedure because during maintenance only one search procedure need to modified and not each of its copies.
• with a small set of too generic components (n generic search procedures), or
• with a small set of useful components which, however, are often too specific for given
problems (x problem specific refinements with $m \leq x < n\cdot m$).

The effect of adapters becomes even more obvious when taking a closer look at search
methods on the one hand, and task- and domain-refinements on the other hand.

• A local search algorithm has four main parameters that determine its search character
 [Graham & Bailor, 1996]: the selection of start nodes, the generation of successors nodes,
 the selection of better nodes, and the definition of the preference relation. Different values
 for these parameters distinguish between, e.g., best-first search, hill-climbing and beam
 search. Keeping the more precise definitions of these parameters externally of the core
definition of the method, enables to provide a large variety of search methods with a small
number of components.

• The definition of a task may be less or more specific. A parametric design task (cf.
 [Schreiber & Birmingham, 1996]) can be defined by a design task that is refined to
configurational design (where the elements of the artifact are given) and further on to
parametric design (where in addition the structure of the artifact, i.e. the selection of the
components is given). [Fensel, 1997b], [Fensel & Motta, 1998] discuss the idea of adapter
piling to express this refinement process. A core definition of design problems (which is
already an adapter for global-optimum problems) is enriched with an adapter for
configurational design and another adapter adds commitments for parametric design.

Implementing a component for each variation of a search method or each task- and domain-
specific refinement is intractable. A tractable and structured approach for reusing (usable)
components can only be achieved by performing refinements via adapters and implementing
different aspects or degrees of refinement by different adapters. Thus, a refined method is
achieved by connecting to it a pile of adapters.

3.2 Deriving adapters

The adapters that are necessary to connect a problem-solving component with a problem or
with other components, can be derived from the results of its selection process. During
selection of a component, the broker has to establish ontological links between the input and
output of a component and the problem definition. By saving this result, the adapter hardwires
the outcome of the ontological engineering that was necessary to establish the appropriate links
between the problem, the domain knowledge and the problem-solving method via a problem
ontology. The adapter can be automatically derived from these ontological links and it is used
to mediate between the domain knowledge and problem-solving method during runtime of the
problem-solving process. Again four types of information flows can be identified during
problem-solving (see Figure 2):

• Sending domain knowledge and case data from the client interface to the adapter.
• Sending a solution from the adapter to the client interface.
• Sending an input from the adapter to a provider interface.
• Sending an output from a provider interface to the adapter.

One additional communication type is an internal communication between the ontologist and
the adapter. It transfers the ontological information derived by the ontologist to construct or
select an adapter that can provide the runtime mapping of terms as required by the clients and
the providers. The ontology mapping information (cf. Figure 2) between the ontologist and the
adapter is exchanged for this purpose.
3.3 Related Approaches in Software Engineering

Adapters are of general importance for component-based software development. [Gamma et al., 1995] introduces an adapter pattern in his textbook on design patterns for object-oriented system development. An adapter enables reusable descriptions of objects and allows to combine objects that differ in their syntactical input and output descriptions. Adapters for problem-solving methods extend this approach in several aspects:

- Adapters introduce assumptions necessary to close the gap between a problem definition (task) and the competence of a problem-solving methods (cf. [Fensel & Benjamins, 1996], [Fensel & Straatman, to appear]).
- Adapters express the task-specific refinement of a problem-solving method (cf. [Fensel & Motta, 1998]). Therefore, an adapter also specifies reusable knowledge. It does not only provide some application-specific glue, but specifies refinements relying on a problem type.
- Adapters could get piled on top of each other to express stepwise refinement of problem-solving methods (cf. [Fensel, 1997b]). This refinement may reflect stepwise adaptation to a problem or stepwise refinement of a generic search strategy according to refined requirements on domain knowledge.

In software engineering, each adaptation is viewed as a (nonreusable) application-specific one. Reusable adapters in knowledge engineering are possible because of the amount of work on reusable task and problem descriptions (cf. [Chandrasekaran et al., 1992], [Breuker & Van de Velde, 1994]). Refining a problem-solving method for a specific type of problems remains reusable for applications that are instances of the same problem or task type.

3.4 Related Approaches on Intelligent Information Integration

Adapters manipulate the syntactical structure of the input and output of components. In that
way, they enable the application of components to different tasks and the combination of components that otherwise could not interact properly. Note, that input of a component may be case data or more complex domain knowledge. In the latter case, adapters fulfill a similar purpose as mediators for heterogeneous information and knowledge systems. Instead of assuming a global data schema, heterogeneous information and knowledge systems have a mediator [Wiederhold, 1992] that translates user queries into sub-queries on the different information sources and integrates the sub-answers. E.g. in the projects Infomaster [Genesereth et al., 1997], Information Manifold [Levy et al., 1996], SIMS [Arens et al., 1993], and TSIMMIS [Papakonstantinou et al., 1995] mediators are provided to integrate heterogeneous information sources. Therefore, adapters that modify components may be combined with mediators to enable different components to access the same knowledge source in different ways or to provide homogeneous access to distributed and heterogeneous knowledge sources for a component. The former minimize the adaptation effort because the same knowledge is presented in different styles. The latter minimize the adaptation effort because heterogeneous knowledge sources are presented homogeneous and uniform.

4 CONNECTOR: COMPONENT COMBINATION

Adapters are a mechanism for coupling components that helps to overcome syntactic and semantic differences. However, further problems may arise from different communication styles of the component. For example, a component $C_1$ may continuously send data to another component $C_2$ and wrongly assume that $C_2$ stores the received data during problem solving. Therefore, data may get lost. Similar, $C_1$ may assume that $C_2$ process its input with the LIFO principle (last in first out, i.e., its input store is a buffer), however, it may apply LILO (last in last out, i.e., its input store is a queue). Such communication mismatches require additional means that properly implement the communication between components. In software architectures, connectors are introduced to fulfill this purpose (cf. [Garlan and D. Perry, 1995], [Shaw & Garlan, 1996]). Connectors provide intermediate stores for data and organize these data in a way that fulfills the assumptions of the sending and receiving components. A variety of connectors and their according architectural styles is given in [Shaw & Garlan, 1996]. Because we have never encountered any difference in the communication styles of “ordinary” software components and knowledge-based components—actually most times the latter assume simple batch mode processing—these connectors of software architectures can immediately be applied to connect knowledge-based components.

5 TASK STRUCTURES: PROBLEM DECOMPOSITION AND SYSTEM INTEGRATION

Task structures [Chandrasekaran et al., 1992], Generalized Directive Models (GDM) [Terpstra et al., 1993] and KADS inference structures [Schreiber et al., 1994] are different means to decompose a complex task into more tractable subtasks. Such an approach is necessary when the problem cannot be solved by directly available components. In this case the task of the broker becomes more complex. It has to select an appropriate task structure that decomposes the entire problem into smaller pieces for which either directly a component can be selected, or for which recursively a new refinement must be selected.

Besides being a means for decomposing a problem, task structures also provide the definition of an integrated system. They can be considered as a system frame for components by associating to each of its subtasks either a component or, recursively, a more refined task structure. A task structure defines an integrated system by defining the control flow and
dataflow between its subtasks. Therefore, such system frames are not only means to
decompose problems, but express vice versa how a complex system can be configured out of
its components. Such task structures or system frames corresponds to software architectures
mentioned above.

In a network of several provider of problem-solving components such an integrated system
frame may refer to components of different providers. The integrated system only exists
virtually in that case. Executing one of its subtasks may imply a call for reasoning service via
the network. [Gennari et al., 1996] describe the use of the CORBA protocol for this purpose. It
allows the distributed execution of reasoning tasks where the substeps may be performed at
different servers.

6 CONCLUSIONS

In this paper we presented a principled approach to reusing problem-solving components.
Components provide reasoning services for problems or fragments of problems. Adapters glue
compONENTS to other components and to domain- and problem-specific circumstances they are
otherwise not applicable. Connectors manage communication aspects. Task structures
decompose problems and integrate loosely coupled components into a system frame. Finally,
a broker provide support in selecting, adapting and combing components (and system
frames). We will briefly summarize these different aspects in component reuse.

At a technical or architectural level, implementing integrated reasoning systems — either
directly or virtually in distributed networks — does not differ from implementing arbitrary
software systems with connectors and software architectures suitable for such environments.
The main extension to existing work in software engineering stems from the fact that research
in knowledge engineering has accumulated a significant number of problem-specific task
structures that can be used as problem-specific architectures. These architectures are not
specific for a domain —like domain architectures in software engineering— but are specific
for a class of problems and a specific decomposition paradigm that assumes specific types of
knowledge for efficiently executing its substeps.

Adapters are also present in component-based software engineering (cf. [Gamma et al., 1995]).
However, the notion of reusable adapters and the use of several adapters to achieve stepwise
refinement of components is rather non-standard. In knowledge engineering, problem-solving
methods and problem types (i.e., tasks) are well-studied. This enables the reuse of adaptations
of reasoning components and of problem-specific refinements. In this paper, we argued that
this use of adapters is the key issue in tractable libraries (i.e., the effort in implementing the
required components and in searching for them is low) and in usable libraries (i.e., by
combining components and adapters one finds combined elements that fit to domain- and task-
specific circumstances).

The key-effort required for providing advanced brokering services for reusing problem-
solving methods is ontological engineering. A broker that enables problem-solving method
reuse requires that problem-solving methods are annotated by formal and comprehensive
descriptions of their assumptions and competence. Ontologies of problem types and tasks
could support the selection process of components and help to establish a link from the generic
vocabulary of problem-solving methods to the domain-specific terminologies of case data and
domain knowledge. Using ontologies to annotate problem-solving methods and problem types
is getting increasingly more attention ([Mizoguchi et al., 1995], [Fensel et al., 1997],
[Benjamins et al., 1998], [Chandrasekaran & Josephson, 1998]) and establishes reuse of

---

2 Not all approaches to task structures provide control. Therefore, such approaches miss an important aspect of our needs.
problem-solving methods as a special branch of knowledge reuse supported by ontologies (cf. [Fridman Noy & Hafner, 1997]). The method specifications are grounded on a common, shared terminology and their knowledge requirements are conceptualized as ontological commitments. Therefore, efforts like the Knowledge Annotation Initiative (KA)$^2$ [Benjamins & Fensel, 1998] tackle a major bottleneck of knowledge and problem-solving components reuse.

Defining such ontologies proper and consensual is essential for the success of component reuse because component retrieval and adaptation cannot not be done fully automatically (cf. [Schumann & Fischer, 1998] for the merits and limitations of theorem proving for component retrieval). Therefore, the human has to be kept in the loop which introduces strong requirements on understandability and user guidance through these ontologies.

Acknowledgement

Richard Benjamins was partially supported by the Netherlands Computer Science Research Foundation with financial support from the Netherlands Organisation for Scientific Research (NWO).

References


