CAHIER DU LAMSADÉ

Laboratoire d’Analyse et Modélisation de Systèmes pour l’Aide à la Décision
(Université Paris-Dauphine)
Unité de Recherche Associée au CNRS n° 825

KBS LIFE-CYCLE AND VALIDATION

CAHIER N° 133
janvier 1996

C. HAOUCHE 1
C. ROSENTHAL-SABROUX 1


1 LAMSADÉ, Université Paris-Dauphine, Place du Maréchal De Lattre de Tassigny, 75775 Paris Cedex 16.
# CONTENTS

<table>
<thead>
<tr>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Résumé</td>
<td>1</td>
</tr>
<tr>
<td>Abstract</td>
<td>1</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>2</td>
</tr>
<tr>
<td>2. Classical main steps in a KBS life-cycle</td>
<td>3</td>
</tr>
<tr>
<td>2.1 Acquiring requirements</td>
<td>3</td>
</tr>
<tr>
<td>2.2 The analysis phase</td>
<td>3</td>
</tr>
<tr>
<td>2.3 Building the KBS</td>
<td>4</td>
</tr>
<tr>
<td>2.4 Validating the KBS</td>
<td>4</td>
</tr>
<tr>
<td>2.5 Installation and maintenance</td>
<td>5</td>
</tr>
<tr>
<td>3. A proposal for validation steps in a KBS life-cycle</td>
<td>5</td>
</tr>
<tr>
<td>3.1 CM validation</td>
<td>7</td>
</tr>
<tr>
<td>3.2 The validation of the CM representation</td>
<td>7</td>
</tr>
<tr>
<td>3.3 Inference paths validation</td>
<td>8</td>
</tr>
<tr>
<td>3.4 Validating the links between the KBS and the CM</td>
<td>9</td>
</tr>
<tr>
<td>4. KBS validation</td>
<td>9</td>
</tr>
<tr>
<td>5. Conclusion</td>
<td>13</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>14</td>
</tr>
<tr>
<td>References</td>
<td>14</td>
</tr>
</tbody>
</table>
 Cyc le de vie et validation d’un SBC

Résumé

Nous nous intéressons aux étapes de validation dans le cycle de vie d’un système à base de connaissances (SBC). Nous nous sommes focalisées sur l’utilisation d’un modèle conceptuel, élaboré pendant la phase d’acquisition des connaissances, pour valider un SBC. Cela implique de définir de nouvelles étapes dans le cycle de vie du SBC et de définir les processus de validation inhérents à ces étapes. Plus pratiquement, nous proposons une méthode de test utilisant en partie un modèle conceptuel à la KADS, plus précisément la structure d’inférence, pour tester le SBC. Nous nous efforçons de préciser, pour chaque étape de validation considérée, quand cette validation doit avoir lieu, qui sont les acteurs de la validation et, enfin, comment cette validation doit être menée.

KBS Life-Cycle and Validation

Abstract

This research deals with validation steps within the KBS life-cycle. Our paper is centered on the use of the conceptual model, elaborated during the Knowledge Acquisition phase, for the KBS validation. We will, on the one hand, show the implications of such a use, that is, the need to specify new steps in the life-cycle as well as validating their outputs. On the other hand, we will propose a procedure to use one part of a KADS conceptual model - the inference structure - to test the KBS. We will, as far as possible, and for each considered step of the life-cycle, precise the following points: When the system will be validated, that is at which step of the life-cycle, who will be validating the system and how this validation will be achieved.
1 Introduction

"The primary functions of a software process model are to determine the order of the stages involved in software development" as Boehm said in (Boehm, 1988). For the KBSs, there are essentially two kinds of development methodologies (Dieng, 1990). The first one is the incremental development technique. This methodology is also called rapid prototyping (Hayes-Roth, et al, 1983). The main steps of this methodology are: identification, conceptualization...
2 Classical main steps in a KBS life-cycle

Some studies have already dealt with KBS life-cycle (Halker & Welz, 1989; Chabaud et al., 1991; Wieginga et al., 1992; Dieng, 1990). The main steps that are identified are as follows (see Figure 1):

2.1 Acquiring requirements

This step is an informal one. It is conducted by means of rough interviews to identify the user’s needs. During this step, the expert, the knowledge engineer and the final user have to define these needs. Then, a decision is taken about the KBS appropriateness to solve the problem.

2.2 The analysis phase

The analysis phase aims at building a conceptual model (CM) that is high level and...
(KADS (Wielinga et al., 1992), COMMET \textsuperscript{1} (Canamero & Geldof, 1993), MACAO (Aussenac, 1989), etc.) in order to structure the knowledge that has previously been acquired and to build a conceptual model (Krivine & David, 1991; Wielinga et al., 1992). Let us notice that the CM building can lead to acquire new data and vice-versa, and this, until the CM is valid.

**The KADS modeling layers** KADS\textsuperscript{2} offers structures for building a conceptual model to guide the knowledge acquisition process. A KADS conceptual model is four-level organized\textsuperscript{3}, from the least abstract, which describes the domain concepts, to the most abstract, which represents the behavior of the system.

* The *domain layer* is a description of the domain concepts and of the relations between these concepts.

* The *inference layer*. In this layer, rules are attributed to the domain concepts. In addition,
Testing approaches

Two main approaches have been used to test KBSs: the black box approach and the white box approach.

a - The black box approach

A test case is composed of a set of data and of the results that are expected from these data. These data are used to run the system. The results obtained are then compared to the expected ones. Notice that this approach deals only with the inputs of the system and with the outputs it provides. It does not consider the structure of the system.

b - The white box approach

The white box approach focuses on the structural aspects of the KBS. It handles a KB structure to study if all the parts of this KB are exercised (Preece et al., 1993). For instance, if the KBS is rule-based, this approach allows to detect the rules which are not used on a given set of test cases, the ones that are often used, etc. Notice that some tools aim at generating test cases automatically; this is the case for SYCOJET (Vignollet, 1991).

2.5 Installation and maintenance

Two possibilities are to be considered. (i) The KBS is completely independent. In this case, it is used with no further effort. (ii) The KBS is to be integrated into a more complex information system. Then, some problems concerning interactions with other modules could occur. The maintenance phase may lead to review the conceptual model.

3 A proposal for validation steps in a KBS life-cycle

Our opinion is that the conceptual model which constitutes a set of specifications for the future KBS, must be used during the validation step. We also think that it is important to link these specifications to the KBS in order to build a KBS that actually corresponds to the initial requirements. Consequently, we propose three additional steps. The first one deals with representing the conceptual model under an implemented form (see (A) on Figure 2). The second one aims at building the inference paths (see (B)) on Figure 2). The third one aims at establishing links between the implemented from of the CM and the KBS (see (C) on Figure 2). The steps (A) and (B) are introduced between the steps building the conceptual model (2) and building the KBS (3), whereas the step (C) is introduced after the step (3). As we do not intervene on steps installation (5) and maintenance (6), they are not represented on Figure 2.

(A) Representing the conceptual model under an implemented form

The conceptual model constitutes a set of specifications and as it must be linked to the
the KADS conceptual model (the domain layer, the inference layer and the task layer) is represented. This step is performed by the knowledge engineer and/or by a programmer.

(B) Building the inference paths

We use the inference structure representation provided within the conceptual model representation, to derive all the inference paths. An inference path is a sequence of inferences between an initial inference and a final inference. An initial inference has only initial roles as inputs, and a final inference has a final role as output.

(C) Linking the implemented form of the CM and the KBS.

These links are used to build a high level trace during the testing phase. Up to now, there is no, to our knowledge, a method to establish such links. In our work, it is realised by the knowledge engineer who can use his knowledge of both the CM and the KBS.

Each step $s$ of the KBS life-cycle is expected to produce outputs $o(s)$ from inputs $i(s)$. Some of these inputs can be seen as a kind of specifications. The outputs $o(s)$ are represented using a given representation. In our opinion, two kinds of validations are thus to be performed on $o(s)$. (i) A validation, which we call internal validation, which aims at verifying that the output $o(s)$ is represented in accordance with the representation formalism used at step $s$, and
(ii) an external validation which aims at assessing that the outputs $o(s)$ verify the specifications expressed in $i(s)$.

We will focus on the validation of the conceptual model (2), the validation of the CM representation (A), the validation of the inference paths (B), the validation of the links (C). The KBS validation will be studied in the next section. We will specify the inputs and outputs for each of these steps as well as the kinds of validation which are necessary.

### 3.1 CM validation

For this phase, the inputs correspond to the CM and to the requirements which have been identified in phase (1). The outputs are a valid conceptual model. The representation formalism on which we focus is the one given by the KADS methodology.

**The CM internal validation** is performed by the knowledge engineer. As we said above, a KADS conceptual model is structured in 3 layers, namely: the domain layer, the inference layer and the task layer. Consequently, we identify two kinds of validation:

- An intra-layers validation that is a validation of each of the layers in the KADS conceptual model. For example, each inference used in the inference structure must have roles as inputs and one role as its output, according to the inference layer definition. For instance, on Figure 8, the inference *complete* has two input roles which are domain model 0 and initial specification 1 and has one output role which is initial specification 2.

- An inter-layers validation. It concerns the validation of the links between the different layers. For instance, each inference used in the task layer must be defined in the inference layer (Haouche, 1993) and every role of the inference layer must be linked to a concept defined in the domain layer. On Figure 8, the concepts that are linked to the roles are written in italics next to these roles. As an example, the concept *l-description* is linked to the role *complete specification*.

**The CM external validation** is performed by the expert, the knowledge engineer and the user. This validation is an informal one, and its goal is to state that the conceptual model satisfies the requirements. The acquired concepts must be suitable, the task hierarchy must be sufficiently detailed, ... For instance, considering the inference structure represented on figure 8, the expert is asked on the sequencing of the inferences. That is, does the inference *complete2* really follows the inference *complete*, etc.

### 3.2 The validation of the CM representation

The inputs for this step are the conceptual model and the CM representation. The output is a valid CM representation. Let us recall that our aim is to be able to handle the “paper” conceptual model under an implemented form according to the KADS methodology. Thus, we use a frame-based formalism in which every component of each layer is described by a frame. For instance, each inference is represented by the following attributes: input-roles, output-roles,
linked-to-domain, linked-to-kbs). Each role is also described by a set of attributes (input-for, output-for, linked-to-domain) as we can see on Figure 3:

The CM representation internal validation is performed by the knowledge engineer and/or the programmer, and is a syntactic one regarding the representation language. For instance, each frame must be described by the appropriate attributes depending on its type (concept, inference, role, ...).

The CM representation external validation is performed by the knowledge engineer and/or the programmer and is an informal one. Its goal is to assess that the implemented form of the conceptual model corresponds to its paper form.

3.3 Inference paths validation

The inputs for this step are a subpart of the conceptual model representation, that is, the inference structure representation and the inference paths, which are represented by lists, as for example the inference path (complete, complete2, create). The outputs are a set of Valid Inference Paths (VIP).

The inference paths internal validation. This phase aims at assessing that the outputs of the step building the inference paths (B) are effectively syntactically correct inference lists. This validation is performed by the programmer/knowledge engineer.

The inference paths external validation. This validation consists, on the one hand, of verifying that the order of the inferences in the lists corresponds to the one given in the inference structure. This is a syntactic validation performed by the programmer/knowledge engineer. On the other hand, an informal semantic validation consists (i) in suppressing semantically incorrect paths, and/or (ii) in adding some “parallel paths” (see the example in Section 4.1). We notice
The KBS internal validation consists of verifying whether the KB is represented in accordance with the formalism. Many tools have been developed to detect redundancies, subsumptions as well as incoherencies in such knowledge bases. This validation is mainly performed by the knowledge engineer and/or the programmer.

The KBS external validation
We focus here on the KBS test using the IP and we propose the following procedure:

1) The KBS is run on a data case and a code level trace is built.

2) The current inference path (CIP) is built from this trace using the links that are established between the CM and the KBS. This CIP constitutes a "high level trace"
Figure 5: The lexical inputs *nodule* and *gauche*

Figure 6: The lexeme instances *nodule-13* and *gauche-15*

Figure 7: The l-description *l-descr25* created for the concept *tumeur*
Figure 8: An inference structure for the sub-task Building L-descriptions

1. (complete complete2 create)
2. (complete complete2 create assemble1)
3. (complete complete2 create assemble2)
4. (complete complete2 create assemble2)
5. (complete complete2 assemble1)
6. (complete complete2 assemble2)
7. (complete complete2 assemble3)
Using the VIP during the testing phase Let us consider the following example: “The thyroid palpation reveals a left lobar nodule”. The KBS tracer yields the following set of fired rules:

<table>
<thead>
<tr>
<th>Rule</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>!regle-instanciation-composantes</td>
<td>x 7 firings</td>
</tr>
<tr>
<td>!create-L-description</td>
<td>x 5</td>
</tr>
<tr>
<td>!transfer-comp-on-l-descr</td>
<td>x 1</td>
</tr>
<tr>
<td>!specify-links-on-l-descr-<em>with-polarity</em></td>
<td>x 5</td>
</tr>
</tbody>
</table>

The rule “!regle-instanciation-composantes” has thus been fired seven times in order to complete the description of the initial lexemes. The rule “!create-L-description” has been fired 5 times to create five L-descriptions, etc.

Without giving the whole rules, we give below the links that have been established between

The rule “!regle-instanciation-composantes” is linked to the inference !complete.
The rule “!create-L-description” is linked to the inference !create.
The rule “!transfer-comp-on-l-descr” is linked to the inference !assemble1
The rule “!specify-links-on-l-descr-*with-polarity*” is linked to the inference !assemble3

These links permit to build the Current Inference Path which is given below:

(complete create assemble1 assemble3).

This path is not a member of the VIP. Generally speaking, if a CIP is not a member of the VIP, the expert has to decide if this CIP is a valid path or not. In the first case (i.e. the CIP is valid), as the CIP does not match the inference structure, the latter has to be modified to take it into account. In the second case, the code level has to be corrected. In our case, the expert considers that the CIP is valid, thus leading to modify the inference structure (Figure 9). Notice that this may lead to add new paths which will have to be validated.

5 Conclusion

Many previous researches in Knowledge Acquisition have shown the importance of the modeling phase in a KBS life-cycle. Our study tends to show that an effective use of the conceptual model implies to introduce new steps in the KBS life-cycle, especially steps dealing with its explicit representation under an implemented form. On the other hand, we propose a procedure to test the KBS using this “implemented” conceptual model and, more precisely, its inference structure. Our approach clearly depends on the KADS methodology. Our future work will deal first with the extension of our experiment to the whole conceptual model and KBS, and second, we will investigate the use of the two other layers of the KADS conceptual model: the domain layer and the task layer. We will try for instance to build the inference paths using the task layer in order to limit the number of inference paths generated during the phase building inference paths (B) of our life-cycle.
Figure 9: The new inference structure for the sub-task Building l-descriptions

Acknowledgments

We would like to thank Jean Charlet and André Molia for their helpful advices.

References


## Contents

1 Introduction  

2 Classical main steps in a KBS life-cycle  
   2.1 Acquiring requirements  
   2.2 The analysis phase  
   2.3 Building the KBS  
   2.4 Validating the KBS  
   2.5 Installation and maintenance

2  Classical main steps in a KBS life-cycle

2.1 Acquiring requirements  
2.2 The analysis phase  
2.3 Building the KBS  
2.4 Validating the KBS  
2.5 Installation and maintenance