Knowledge Acquisition as Constructive Modelling

B.LEROUX, P.LAUBLET

Congrès international de systémique,
Prague, Octobre1993

An approach to knowledge acquisition combining alternate steps of constructive modelling and elicitation

B.LEROUX, P.LAUBLET

East-West Artificial Intelligence Conference, EWAIC 93
Moscou, Septembre 1993

KNOWLEDGE ACQUISITION AS CONSTRUCTIVE MODELLING

Bernard LE ROUX [1,2]

1 ONERA
29, Avenue de la Division Leclerc
BP 72
92320 CHATILLON Cédex
FRANCE
E-mail: leroux@onera.fr

Philippe LAUBLET [1]

2 LAFORIA-CNRS
Université Pierre & Marie Curie
Tour 46-0 - 4, place Jussieu
75252 PARIS Cédex
FRANCE
E-mail: laublet@onera.fr

Abstract

Model-driven knowledge acquisition is now a common view. KADS-based VITAL methodology emphasizes on knowledge acquisition as constructive modelling. We advocate the break off with the transfer approach and justify the modelling approach. In this framework, the Generalised Directive Models' approach that we recall, has been a first attempt to provide a context sensitive use of the models. We are here mainly concerned either by the double problem of model's selection and refinement and by the problem of knowledge elicitation , which is sometimes the blind point of model driven knowledge acquisition. We present some improvements of the GDM' approach. Our approach to model-driven knowledge acquisition considers it as a combined and recurrent process of model's refinement and knowledge elicitation. Our conclusion emphasizes on some relationships between our approach to knowledge acquisition and some statements of the systems' science approach.

Résumé

L'acquisition des connaissances guidé par les modèles est maintenant une vue commune. La méthodologie VITAL basée sur KADS insiste sur la considération de l'acquisition des connaissances comme modélisation constructive. Nous défendons la rupture avec l'approche transfert et justifions l'adhésion à l'approche modélisation. Dans ce cadre, les "generalised directive models" ont constitué une première tentative de modélisation contextuelle. Nous sommes ici principalement intéressés, d'une part, par le double problème de la sélection et du raffinement des modèles, d'autre part, par le problème de l'explicitation des connaissances, qui est parfois le point aveugle de l'acquisition guidée par les modèles. Nous présentons des améliorations de l'approche GDM. Nous considérons l'acquisition guidée par les modèles comme processus combiné et récurrent de modélisation constructive et d'explicitation des connaissances. En conclusion nous soulignons les relations entre notre approche de l'acquisition des connaissances et certains traits de l'approche systémique.

Introduction

Knowledge Acquisition for Knowledge Based Systems, as a separate field of research, is about ten years old. Many among the more actual works in this field refer to a model-driven approach. This is illustrated by the following and non-exhaustive list: Role Limiting Methods **[Marcus 88]**, KADS **[Schreiber 93]**, PROTEGE-2 **[Puerta 91]**, VITAL **[O'Hara 92]**... These works are different from many aspects, more particularly about their ability to provide an executable knowledge base (Role Limiting Methods or PROTEGE-2), nevertheless all of them carry a break off with an older approach to building knowledge-based systems, the so-called transfer approach.

KADS methodology arises now as a possible European standard. The models that it advocates are useful to design knowledge based systems and also effective to help the work of the knowledge engineer. The main model is the expertise model, that specifies the knowledge embodied by the system and its problem-solving method. Our work is mainly concerned by this model. In contrast with approaches that provide an executable knowledge base, the expertise model is not executable. It is the specification of the requirements of the target system. A further work is needed to implement the system, this work is not here under consideration. Some on going works try to reduce the gap between conceptual models and implementation, see **[Balder 92]** or **[Jonker 92]**.
To provide guidance and help to the knowledge engineer, KADS proposes models of problem-

solving methods called interpretation models. The scheduling is first to specify a suitable interpretation model according to the application under consideration, this work being related to the selection of a generic model into a library and to its refinement, then to organise knowledge acquisition according to the model.

Our research is partially funded by the Esprit project VITAL[1] which is developing a KADS-based methodology and workbench. This project advocates a different approach, more particularly about model selection and refinement. We assume that the strong separation, that holds KADS methodology, between two steps, the former to select and refine the model, the latter to organise knowledge acquisition according to it, is not completely realistic. We propose an approach, the *"Generalised Directive Models"*, whose purpose is a better integration of the two activities. The first idea of the generalised directive models' approach has been introduced in the Esprit project ACKnowledge **[Van Heijst 92]**, and is a significant part of the VITAL methodology **[O'Hara 92]**. Our purpose is to give an effectual content to this idea.

In the sequel, we will first present the general background of our approach, then in the second section we will develop some statements about our view of knowledge acquisition as a combined process of constructive modelling and knowledge elicitation, and our conclusion will bear on the connection between our approach and system science.

Knowledge Acquisition and Modelling

The so-called transfer approach, contemporaneous with the first developments of expert systems, saw the acquisition of expert's knowledge as its expression into a formalism executable by the inference engine of the target expert system. A first set of problems occurs: such systems were very difficult to maintain, their improvement was often unmanageable, the understanding of their performance was very fuzzy. We faced up another problem, of main importance, expert's knowledge is not spontaneously available, then a direct re-expression of this knowledge is often impossible, always difficult. As a consequence of the last point, many attempts occur in the framework of the transfer approach to design elicitation tools and techniques based on cognitive psychology **[Boose 88]**. We would like now to briefly present a paper, often quoted in the field of knowledge acquisition.

Some points of reference

During the last years, in the field of knowledge acquisition, many references have been made to the well-known Newell's article about *"the knowledge level"*. **[Newell 82]**

The main point usually emphasized by the knowledge acquisition community is the separation between symbol level and knowledge level, as two different levels of description of computer systems. This separation is seen as the possibility of working on knowledge far away from the constraints of the implementation, semantic constraints of implementation languages being too strong to allow an effective expertise modelling. The emphasis on different levels of description for a complex system may be related to other statements made, for instance, in the field of biology: *"des systèmes conceptuels particuliers ont été choisis pour chaque niveau, afin de décrire convenablement et simplement les phénomènes observés"* [2] **[Delattre 84]** or also *"bien souvent l'équipement en concepts et en techniques qui s'applique à un niveau ne fonctionne ni au dessus, ni en dessous"* [3] **[Jacob 70]**. Newell underlines the fact that knowledge itself must be focused more accurately[4] in the field of Artificial Intelligence. The importance of this point is noted by Newell himself in his conclusion: *"Splitting what was a single level (symbol) into two (knowledge plus symbol) has immense long-term implications for the development of AI. It permits each of the separate aspects to be adequately developed technically. "* **[Newell 82]**

---

[1] This article binds the authors but not the members of the consortium: Andersen Consulting(E), Bull (F), Nokia (SF), Onera (F), Open university (UK), PTT Nederland (NL), Syseca (F), University of Nottingham (UK).

[2] "some particular conceptual systems have been chosen for each level, to achieve a suitable and simple description of observed phenomena" (we translate)

[3] "fairly often the kit of concepts and techniques which is applied at a particular level doesn't work neither at the above level, nor at the level below" (we translate)

[4] *"our notion of knowledge has precisely a competence-like character"* **[Newell 82]**

Many of the most current works in knowledge acquisition correspond to this conception. Some of these works are on the symbol level side, languages and representations, some others are on the knowledge level side, improving conceptual models, ontologies and models of reasoning processes. This second part of the work, KADS interpretation models for instance, is a first answer to Newell's hope that *"the separate knowledge level may lead to constructing a generative class of rational systems..."* **[Newell 82].**

We want now to stress two other points in Newell's paper. We consider that these two points are at least as important as the former presented just above, but not always emphasized by the knowledge acquisition community.

The former is about intentionality[5]. In his paper, Newell underlines the fact that intelligent systems are intentional systems, and that accordingly knowledge is related to intentionality. So, the Newell's view of these systems is that they are composed of goals and knowledge, and directed by a principle of rationality. It is worth noting that Newell considers that his knowledge level corresponds to Dennett's intentional stance **[Dennett 90]**, and symbol level to sub personal stance. This point is important according to Dennett's and Newell's formulations about interpretation. As, for an observer, the sub personal stance of an agent is unreachable, the understanding of the behaviour of this agent entails interpretation and simulation. Interpretation because the observer ascribes goals and knowledge to an agent to explain its behaviour: the observer constructs a theory of this behaviour. Simulation because, to construct such a theory, the observer makes the hypothesis that there is a functional equivalence between the agent and him: both are teleological systems with rationality principle.

The latter is about modelling. Newell stresses on simulation: *"the knowledge attributed by the observer to the agent is knowledge about the external world. If the observer takes to itself the role of the other, assuming the agent's goals and attending to the common external environment, then the action it determines for itself will be those that the agent should take. (...) Thus, simulation turns out to be a central mechanism that enables the knowledge level and makes it useful."* **[Newell 82]** We think that this emphasis is clearly related to modelling. Simulation is here a mean to create a model allowing to understand the behaviour of an agent, with no claim that it is the model of the agent itself. As stated by Dennett, the only hypothesis necessary to allow such a modelling is that the observer must consider the agent as functionally equivalent to him. "One *way of viewing the knowledge level is as the attempt to build as good a model of an agent's behaviour as possible based on information external to the agent, hence permitting distal prediction."* **[Newell 82]**

The three points noted here corroborate the relevance of Newell's paper for knowledge acquisition. Splitting between symbol and knowledge levels, as two different levels of description of a same system, emphasising on intentional systems and on interpretation/simulation based modelling, corresponds to many of the more current views in knowledge acquisition. The last idea to forsake is that an expert system is not a copy of the expert himself[6], but rather an artificial system that embodies a theory of some systems in the world that the expert is only a useful informant to build. Our KADS-based methodology is an attempt to provide *"adequate technical development"* for knowledge level models whose purpose is to construct *"a generative class of rational systems"*. The three epistemological categories of knowledge embodied in KADS models (task inference domain) support the modelling of teleological system with extended rationality principle[7].

A recall about KADS models

KADS, to break off with the transfer approach of knowledge acquisition, emphasizes on modelling

---

[5] "An important connection appears to be with the notion of intentional systems.../...In the hands of Daniel Dennett, a philosophers who has concerned himself rather deeply with AI, the doctrine of intentional systems has taken a form that corresponds closely to the notion of the knowledge level, as developed here. He takes pains to lay out an intentional stance, and to relate it to what he calls a subpersonal stance. His notion of stance is a system level, but ascribed entirely to the observer.../... The subpersonal stance corresponds to the symbolic or programming level.../... The intentional stance corresponds to the knowledge level" **[Newell 82]**

[6] this may be related to Bertalanffy's "la dé-anthropomorphisation progressive de la science" **[Bertalanffy 73]** "the progressive dis-anthropomorphisation of science" (we translate)

[7] in our view this principle is clearly related to Simon's *"bounded rationality"* **[Simon 91]**

[Wielinga 92a and b]. The central model in KADS is the model of expertise (organised initially in four layers) that is the result of knowledge acquisition. We will not consider anymore the fourth layer, the strategic one, which had no real development. The three other layers play different roles: the domain layer embodies knowledge describing a theory of an application: concepts, properties of concepts, relations between concepts, attributes of concepts and so on... - the inference layer is a functional view of the target system, a dependency diagram where roles (abstract entities that refer to pieces of domain knowledge, defined according to problem-solving method) and inferences (which use of domain relations to perform) are linked together - the task layer embodies a goal decomposition and control view (a sequence on inferences, or loops and so on) of the target system.

In KADS, interpretation models, which are expertise models without a domain layer, are proposed to provide guidance for knowledge acquisition. The inference layer, as a **sight** on the domain layer, has to name the different roles of knowledge, and the different inferences (specifying inputs and output and abstract function grounding the system dynamism). The task layer, as a **sight** on the inference layer, has to organise the path between inferences according to goal achievement and control.
Thus, roles plus inferences (making a partition on knowledge to acquire), as well as goal decomposition and control, must be an effective guidance for knowledge acquisition because they indicate, at least partially, what knowledge chunks are needed.

The process of model-driven Knowledge Acquisition

To privilege top-down knowledge acquisition, KADS has interpretation models, which are tools in a model-based approach **[Le Roux 93b]**. Thus, the first problem, when working on a particular application, is to select an appropriate interpretation model.

In KADS, this choice is related to some features of the application allowing to traverse a decision tree; When a leaf is reached, we have an interpretation model. The borderline between features of an application and domain knowledge seems difficult to locate and the amount of information required to reach a leaf may be important. If a leaf is unreachable, no interpretation model is available, so model-based knowledge acquisition is impossible for a while, and it remains a bottom-up approach.
Interpretation models are defined once for all (even if the library is extensible). The leaf reached may provide an interpretation model that suits more or less the application. No guidance is provided to further refine the model, whose grainsize or other features may be an impediment.

Thus, in KADS, model-driven knowledge acquisition is seen as a process embodying two strongly separate steps, the former is to select/refine the model, the latter to organise knowledge acquisition according to the model. We assume that this two-step process, analytic step plus knowledge acquisition step, is not realistic, and we propose a practical issue for model-driven knowledge acquisition: to design it as a combined process of constructive modelling and knowledge elicitation **[Le Roux 93c]**.

A first idea: the generalised directive models' approach

The idea of generalised directive models' approach **[Van heijst 92]** is that it is possible and useful to supply a context sensitive grammar able to transform a current inference structure into a more specified one, according to the application knowledge already acquired. Knowledge acquisition is seen as a recurrent process of refinement where knowledge is gradually acquired and inference structure is progressively refined. A simple presentation **[Van Heijst 92]** for a classification application task starts with the following model:

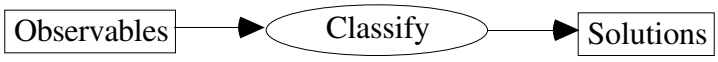
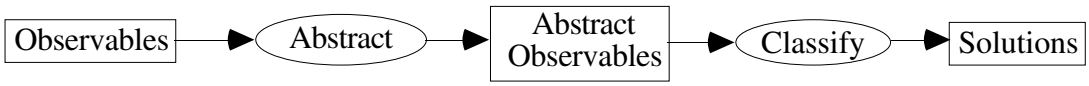


Figure 1: an initial directive model

It provides a partition of domain knowledge into observables and solutions. Different conditional rules should be available to expand this first abstract model. For example, if the elicited knowledge about observables shows that they are hierarchically organised and that there is a lot of observables, the refinement could be:



RULE APPLIED TO OBTAIN THIS MODEL FROM THE FIRST ONE:	
Conditional part:	• *IF input has a hierarchical structure* • *AND IF size(input) > 10 THEN*
(input Classify output)	**- (input Abstract abstract-input Classify output)**

Figure 2 a refined directive model and related rule

Some further refinements are then possible about Classify or Abstract, then begins another step of the combined knowledge acquisition and modelling process.

Our view of model-driven Knowledge Acquisition

From the general idea of the GDM, we have improved and fleshed out the approach with the aim to build up a tool providing advice and guidance both for modelling and elicitation **[Le Roux 93 a]**.
The improvements are:

- to take into account the modelling of the task structure
- to consider the modelling of the domain layer
- to supply an effective context sensitive grammar.

Our work is to develop a synthetic method of application modelling, in contrast with the analytic approach in KADS. Then we consider the knowledge acquisition process as a recurrent and combined process of knowledge elicitation and model construction/refinement.

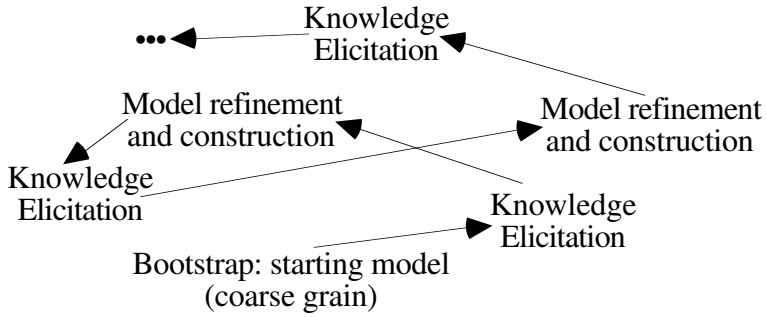


Figure 3: The spiral of the recurrent process of model construction and knowledge elicitation

The refined models running must help to acquire new domain knowledge and this new acquired domain knowledge must help to refine the model in a cyclic way.
In this framework, we have developed different scenarii to experiment with our improved generalised directive models' approach on different tasks, diagnosis, assessment and so on...**[Le Roux 93c]** This work has been accomplished to produce several classical models of problem solving methods presented through the knowledge acquisition literature. As our purpose is to construct a generative method to construct models, it may appear quite strange to develop examples leading to classical problem-solving methods. We consider that a great experience and rationale are underlying these methods and that the minimal expectation for a generative approach is to be able to construct such methods. Moreover, it is a kind of empirical inquiry about existing models and a "Gedanken" experiment about modelling in knowledge acquisition.

The framework of our constructive approach

From our project as well as from our experiments, it is possible to draw some features of our approach and some pragmatic and methodological lessons. The pragmatic problems of model's selection and refinement, and of knowledge elicitation are mainly focused.

A least commitment strategy

This constructive approach is grounded on a least commitment strategy, i.e. an application of Occam's razor to modelling: *"all things being equal, parsimonious models must be preferred."* The

generalised directive models' approach provides the most simple model according to the information available about an application, this is a first answer to the problem of model's selection. This has also significant consequences on some problems of model driven knowledge acquisition: granularity, genericity, re-usability.

Our approach provides a model with a suitable grain (i.e. a satisfying level of specification) according to the available amount of information about an application: it allows a **variable granularity for the models**. The genericity problem is now on the naming of the inferences and roles because names are indices for the rewriting grammar, and also on the abstract function that embodies each inference and related possible model schema: **genericity of basic modelling components.** Then, we have to evidence some primitive elements of artificial reasoning. As well the problem of re-usability changes, the specification of the grammar is mainly focused: our approach brings a relevant framework to think about the problem of re-use. The main concern is here on the rationale of the models, re-use is mainly concerned by the fact that different applications may share the same goals or sub-goals, may share the same model schema (causal models, or decomposition models, or hierarchical conceptual models...), and may share the same pieces of artificial reasoning primitives.

A context sensitive grammar

The conditional part of the rules of the grammar allow to choose the next refinement of the model. According to our experiments, the following elements are concerned:

- a character of the domain knowledge corresponding to a static role (mainly about the structure of the knowledge it refers to: model schema)
- a character of the domain knowledge corresponding to a dynamic role (mainly about the size and the structure of the knowledge that this role refers to)
- a failure in acquiring inference knowledge (often related to an under-differentiation of roles or inferences)
- the possibility for an inference to have a certain static role as input (knowledge about modelling)
- a specific feature of the application at hand (mainly related to goal decomposition).

Neglecting the task structure, as in the Acknowledge version of the GDM, seems to us as a problem, because of the importance of goal decomposition for modelling (furthermore as we build intentional systems) and control acquisition (acquiring control is often underlined as a problem for knowledge acquisition). So, the context sensitive grammar will take into account both inference and task structures.

The emphasis on domain modelling is, from our point of view, important to understand the rationale of the models (e.g. a causal model for a role is related to a class of abductive reasoning systems and accordingly to specific inferences possible in this framework), and to provide an effective guidance for knowledge elicitation. This has strong consequences on the conditional part of the rules of our context sensitive grammar.

An effective context sensitive grammar must take into account the domain knowledge already acquired as in the Acknowledge version of the GDM, but also goal decomposition (as far as we assume that the necessity of an intermediate subgoal may appear during the modelling process), model schema of the domain (as far as we consider that there is a link between a reasoning method and the knowledge structure), control and so on...

A scheduling on Knowledge Acquisition

We assume here that a scheduling on knowledge acquisition is possible and useful. The models provided in the framework of a methodology are supposed to permit the knowledge acquisition task's organisation. One way to organise this task is to know what is the most important to acquire at a time according to the information provided by the inference structure (roles plus inferences and the dependencies between pieces of knowledge to acquire), and by the task structure (specifying the roles that have a control function).

As in our approach the model is itself progressively refined/constructed, either the dependencies between pieces of knowledge, and a function of control for a role, come into view gradually. This gradual process provides a manageable scheduling for knowledge acquisition. As our approach combines the constructive modelling of a reasoning model, and the related model schema and model ontology, it is possible to evidence information about the structure of knowledge, providing a guidance to elicit this knowledge accordingly.

About limited interaction hypothesis

Two hypothesis ground the KADS approach: "*it is possible and useful to distinguish between several generic types of knowledge according to the different roles that knowledge can play in reasoning processes,*" and "*that these types of knowledge can be organised in several layers, which have only limited interactions.*" **[Wielinga 92a]** The different layers under consideration are the task layer, the inference layer and the domain layer.

The former hypothesis is a straightforward continuation of Clancey's works **[Clancey 85]**. This work, through an analysis (reverse engineering) of ten existing expert systems, has evidenced the common reasoning model that these systems were embodying. This common underlying reasoning model has been called by Clancey *"heuristic classification"*. The conclusion of the author was that *"knowledge content has structure"* [8] **[Clancey 85]**.

The latter is a foundation stone authorising the interpretation models' approach: as layers have only limited interactions (the task layer as a sight on the inference layer, the inference layer as a sight on the domain layer), it's possible and sensible to construct abstract reasoning models that are expertise models without a domain layer.
This hypothesis finds here a confirmation while it is also weakened: as application domain knowledge is used to refine the model, that means obviously that there are interactions between layers, and as it entails alone only a modification of the inference structure, it shows the importance of interactions that imply a weakening of the hypothesis (domain to inference) • as the choice of a refinement may imply a choice on a model schema, it confirms limited interactions (inference to domain) • as a new goal decomposition (task structure) relies on a purpose of the application task and only potentially on domain knowledge, it confirms the limitation of interactions (task to inference) • as control in the task structure is related to domain knowledge through a role in the inference structure, it implies again a weakening of the hypothesis (task to domain via inference).

About modelling and models

The use of models in science is mainly related to the purpose of providing sense and understanding to some systems in the world. These models need to be specified as well as the underlying assumptions that ground them. The actual works in the field of knowledge acquisition have evidenced some grounding assumptions and specifications of the models that we use. In the framework of KADS based methodologies, we have for instance models providing:

- an intentional view of the target system (task structure as goal decomposition)
- a control view of the target system (task structure as control on the reasoning process)
- a functional view (inference structure: a dependency diagram between roles and inferences)
- an ontological view (model schema and related model ontology as organising the domain layer, as providing a structure to the domain knowledge and the relations that use the inferences).

Nevertheless, it is sometimes underlined **[Aben 92]** that KADS interpretation models, for example, according to KADS ontology, use of primitive inferences in a quite strange and sometimes inconsistent way, and that the related ontology itself is too poor and incomplete.
On the other hand, approaches as Role limiting methods or Protege-2, as they embody a hard-wired problem-solving method and ontology, appear as effective and useful for knowledge elicitation. As a counterpart, they have a strongly limited scope to applications that perfectly suit the method and ontology they embody. Thus, they are lacking flexibility.

We assume that our approach is a relevant framework to improve a more complete and coherent ontology, and to provide more consistent models, while keeping a great flexibility. We hope that it will have for consequence to authorise a better understanding of models, more suitable to the technical requirements of our scientific work. We claim that it is possible in this framework to evidence some syntactic, semantic and pragmatic constraints on knowledge acquisition models, providing therefore some theoretical advances for model-driven knowledge acquisition.

---

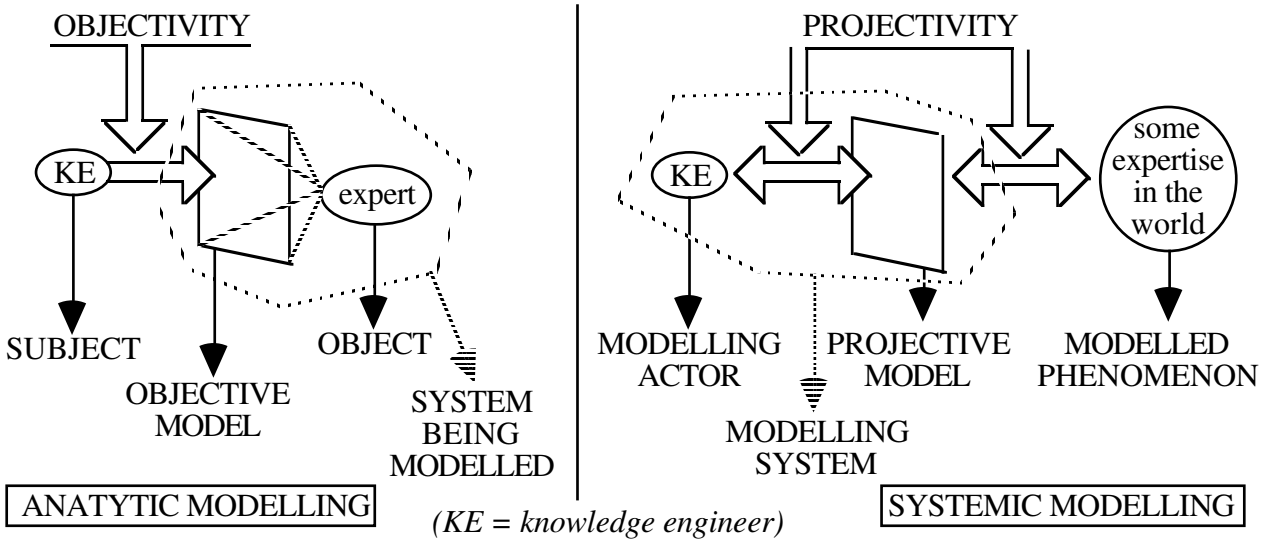
[8] We may here paraphrase the aphorism of Henri Poincaré **[Poincaré 05]** that *"an accumulation of facts is no more a science than a heap of stones is a house"* (we translate) and a heap of knowledge a knowledge based system

System Science & Knowledge Acquisition

We want now to underline the relationship between system science and our view of knowledge acquisition. The existence of such relationship is not surprising. Knowledge acquisition is a particular field of research and activity of Artificial Intelligence. We have previously quoted Newell, but in the field of AI, the name of Newell is strongly associated with Herb Simon, whose works are considered both as improving AI and developing system science **[Le Moigne 90]**.

The first point we want to emphasize on is about complexity. It is worth noting that complexity is a character of expertise and more generally that it is a character of reasoning. This is a first reason to consider the solutions supported by the system science to modelling complex systems.

The emphasis on modelling is shared by system science and knowledge acquisition. Accordingly, we agree with J.L. Le Moigne writing *"un système complexe on doit le modéliser pour construire son intelligibilité (compréhension)"* [9] **[Le Moigne 90]**. We consider that the current works in the field of knowledge acquisition must be considered as an attempt to provide suitable models and a modelling framework able to authorise a better intelligibility of systems embodying as a complex behaviour as expertise. Moreover, the emphasis of Le Moigne on the fact that *"la connaissance est construite par le modélisateur qui en a le projet, dans ses interactions permanentes avec les phénomènes qu'il perçoit et qu'il conçoit."* [10] **[Le Moigne 90]** is very sound in our domain[11]. Particularly on two points that we have presented previously, the former is about the break off with the transfer approach, the latter about the fact that we do not consider our model as a model of the expert (a psychological model). In this way, we may reinterpret the following figure from **[Le Moigne 90]**:



(KE = knowledge engineer)

In the left part of the above figure we could read a representation of the transfer approach. In this framework, the objective model is the language (syntax plus semantic) of the inference engine. The system being modelled is the expert, whose knowledge is represented in this language and whose reasoning is carried out by the inference engine. In this framework, the cognitive psychology-based elicitation may be seen as related to the strong hypothesis of a functional identity between expert systems and experts. As noted earlier, two problems are under consideration, the strong syntactic and semantic constraints of the languages of inference engines, and the assumption that expert's knowledge and expert's way of reasoning are immediately available. Another problem is the underlying assumption of such an approach: the possibility to build a universal inference engine.

In contrast, we could read the right part of the above figure as a representation of the model-driven approach. The projective model is for us, in our domain, the interpretation model that we have

[9] "a complex system has to be modelled to construct its intelligibility (understanding)" (we translate)

[10] "knowledge is constructed by the modelling actor who plans it, within a permanent interaction with the phenomena that he perceives and that he designs" (we translate)

[11] also "The systemic modelling assumes that modelling is not neutral and that phenomenon's representation is not disjoined from the action of the modelling actor" (we translate) **[Le Moigne 90]**

presented along this paper. The concept of projectivity embodies both project and projection (as geometric analogy). The main characters of this model is to provide different views on a same phenomenon: intentional view, control view, functional view, ontological[12] view. No assumption is made on the fact that we are modelling the expert. The strong hypothesis noted above about functional identity is weakened into the hypothesis of functional equivalence. The models we use are specified either according to some assumptions about the kind of phenomenon we want to model (thus related to projection), and also according to the goal of building computer systems (thus related to project). This project is reflected into the kind of models we use and remains as a bias. As noted previously, the models used to provide guidance for knowledge acquisition are called **interpretation** models: the projectivity supported by such models appears clearly in their naming.

According to the presentation of systemic modelling by **[Le Moigne 90]**, the main question about the system to model is "What does the system does?" (functions, operations, transformations...) and not "Of what elements the system is made?". We think that the models that are presented in this paper as useful for knowledge acquisition are clearly organised by the same kind of questions:

- the task layer answers to the question "What are the goals that it is possible to attribute to the system according to his behaviour?"
- the task and the inference layers answer to "How does a (computer) system can do to achieve such goals?"
- and the domain layer answers to "What does the system need to perform?"

We have enough underlined, according to the length of this paper, the relationship between systemic modelling and model-driven knowledge acquisition. Of course, we do not consider that this short contribution has made an exhaustive assessment of the problem. Many other developments should be done according to different aspects of systemic modelling, among others: the particular role of simulation, the three axioms of systemic modelling as presented in **[Le Moigne 90]**, the works of Delattre on models and modelling **[Delattre 84]**, and so on...

As often emphasised by Bachelard, *"un discours sur la méthode scientifique sera toujours un discours de circonstance, il ne décrira pas une constitution définitive de l'esprit scientifique"* [13] **[Bachelard 34]**. Then, it appeared interesting to us to present some views of this particular field of artificial intelligence that is knowledge acquisition, as far as the present technical developments in this domain are significant to improve an epistemology for today, when science has more than ever to face up the problem of complexity in any domain.

Many discussions in the research team ACASA[14] directed by J-G Ganascia have been helpful for one of the authors (BLR). For so many discussions about knowledge acquisition and the "knowledge level" we thank Jérôme Thomas, with no claim to his agreement with this paper.

[12] all along this paper we use the terms ontology or ontological with its usual meaning in the field of knowledge acquisition, that is at a meta-level the organisation of knowledge and particular terms according to this organisation, and also in the sense of the peculiar vocabulary of a domain; We don't refer to this term as it is used in classical philosophy.

[13] "a discourse about the method of science will always be a circumstantial discourse, it will never describe a definitive constitution of the scientific mind" (we translate)

[14] Acquisition des Connaissances et Apprentissage Symbolique Automatique: knowledge acquisition and symbolic automatic learning

References

[Aben 92] Manfred ABEN: On the specification of knowledge model components - 1992 - proceedings of BANFF Knowledge Acquisition Workshop

[Bachelard 34] Gaston BACHELARD: Le nouvel esprit scientifique - 1934 - ed. 1987 - Quadrige Presses Universitaires de France

[Balder 92] J.R. BALDER & J.M. AKKERMANS eds: Formal methods for knowledge modelling in the commonKADS methodology, a compilation - 1992 - Netherlands Energy Research Foundation ECN Service Unit General services P.O. Box 1 NL-1755 ZG Petten The Netherlands

[Bertalanffy 73] Ludwig Von BERTALANFFY: Théorie générale des systèmes - 1973 - Dunod in Wetter & al. eds, Current developments in knowledge acquisition - EKAW'92 - Lecture notes in AI 599 Berlin - Springer

[Boose 88] John BOOSE, Brian GAINES eds: Knowledge acquisition tools for expert systems - 1988 - Academic Press Harcourt Brace Jovanovitch, Publishers

[Clancey 85] William CLANCEY: Heuristic classification - 1985 - Artificial Intelligence N°27 - North-Holland

[Delattre 84] Pierre DELATTRE: Système structure fonction évolution - 1984 - Maloine

[Dennett 90] Daniel DENNETT: La stratégie de l'interprète - 1990 - Editions GALLIMARD

[Jacob 70] François JACOB: La logique du vivant - 1970 - éditions Gallimard

[Jonker 92] Willem JONKER, Jan Willem SPEE: Yet another formalisation of KADS conceptual models - 1992 - in Wetter & al. eds, Current developments in knowledge acquisition - EKAW'92 - Lecture notes in AI 599 Berlin - Springer

[LeMoigne 90] Jean Louis LE MOIGNE: La modélisation des systèmes complexes - 1990 - AFCET systèmes DUNOD

[Le Roux 93a] Bernard LE ROUX, Philippe LAUBLET: From simple diagnosis to Cover & Differenciate: a trial for a conceptual and compositional approach - 1993 - Proceedings 3rd KADS user meeting - Christiane Lockenhoff & al. editors - SIEMENS AG ZFE BT SE 21 Munich

[Le Roux 93b] Bernard LE ROUX, Philippe LAUBLET: Modèles de méthodes de résolution de problèmes et acquisition des connaissances - 1993 - Proceedings Journées d'Acquisition des Connaissances St Raphael

[Le Roux 93c] Bernard LE ROUX, Philippe LAUBLET: A conceptual and compositional approach for knowledge acquisition: examples and lessons - 1993 - Rapport N°93/06 Laforia Institut Blaise Pascal - 4, place Jussieu 75252 Paris Cédex 05

[Marcus 88] Sandra MARCUS editor: Automating Knowledge Acquisition for Expert Systems - 1988 - Kluwer Academic Publishers

[Newell 82] Allen NEWELL: The knowledge level - 1982 - Artificial Intelligence N°18 - North-Holland

[O'Hara 92] Kieron O'HARA, Nigel SHADBOLT, Philippe LAUBLET, Manuel ZACKLAD, Bernard LE ROUX, Mandy MEPHAM: Knowledge Acquisition Methodology in VITAL - 1992 - VITAL Report DD 2.1.2 Esprit Project 5365

[Poincaré 05] Henri POINCARE: La science et l'hypothèse - 1905 - ed. 1968 - Flammarion

[Puerta 91] Angel PUERTA, John EGAR, Samson TU, Mark MUSEN: A multiple-method knowledge-acquisition shell for the automatic generation of Knowledge Acquisition tools - 1991 - Knowledge Acquisition volume 4 N° 2 - Academic Press Harcourt Brace Jovanovitch, Publishers

[Schreiber 93] Guus SCHREIBER, Bob WIELINGA, Joost BREUKER: KADS a principled approach to knowledge based system development - 1993 - Academic Press Harcourt Brace Jovanovitch, Publishers

[Simon 91] Herbert SIMON: Sciences des systèmes sciences de l'artificiel - 1991 - AFCET Systèmes DUNOD

[VanHeijst 92] Gertjan VAN HEIJST, Peter TERPSTRA, Bob WIELINGA, Nigel SHADBOLT: Using Generalised Directive Models in Knowledge Acquisition - 1992 - in Wetter & al. eds, Current developments in knowledge acquisition - EKAW'92 - Lecture notes in AI 599 Berlin - Springer

[Wielinga 92a] Bob WIELINGA, Guus SCHREIBER, Joost BREUKER: KADS: A Modelling Approach to Knowledge Engineering - 1991 - Knowledge Acquisition volume 4, n°1 - Academic Press Harcourt Brace Jovanovitch, Publishers

[Wielinga 92b] Bob WIELINGA, Walter VAN DE VELDE, Guus SCHREIBER, Hans AKKERMANS: The KADS knowledge modelling approach - 1992 - proceedings Japan Knowledge Acquisition Workshop

An Approach to Knowledge Acquisition Combining Alternate Steps of Constructive Modelling & Elicitation

Bernard LE ROUX [1,2]

1 ONERA
29, Avenue de la Division Leclerc
BP 72
92320 CHATILLON Cédex
FRANCE
E-mail: leroux@onera.fr

Philippe LAUBLET [1]

2 LAFORIA-CNRS
Université Pierre & Marie Curie
Tour 46-0 - 4, place Jussieu
75252 PARIS Cédex
FRANCE
E-mail: laublet@onera.fr

Abstract

Model driven knowledge acquisition is now a common view. KADS-based VITAL methodology emphasises on knowledge acquisition as constructive modelling. In this framework, the Generalised Directive Models' approach that we recall, has been a first attempt to provide a context sensitive use of the models. We are here mainly concerned by the double problem of model's selection and refinement. We show the possible use of the GDM's approach to construct a model of cover & differentiate. We advocate some improvements of the GDM approach. Our approach to model driven knowledge acquisition considers it as a combined and recurrent process of model's refinement and knowledge elicitation.

Keywords

Elicitation - Generalised Directive Models - Interpretation Models - KADS - Knowledge Acquisition - Model - Modelling - VITAL.

Introduction

Knowledge Acquisition (KA) developed different approaches, the more actual ones refer to a model driven approach: Role Limiting Methods **[Marcus 88]**, KADS **[Wielinga 92a]**, PROTEGE-2 **[Puerta 91]**, VITAL **[O'Hara 92a]**...
KADS models are used as an effective mean to specify KBS, providing an effectual help. Moreover, the KADS approach proposes generic models as guides for knowledge acquisition itself. Roughly speaking, the general idea is to choose a model of the problem-solving method (named interpretation model) from the task features, then to use it to acquire and to interpret the domain knowledge. In this paper, we propose a related approach, the Generalised Directive Models' approach (GDM), the purpose is to integrate better the two activities of building the interpretation model and acquiring domain knowledge in a recurrent refinement process. The idea of the GDM has been first introduced in the ACKnowledge project **[Van Heijst 92a b]** and is a significant part of the methodology advocated by VITAL Esprit project **[O'Hara 91]** that funds the work of the authors. In the sequel, we will recall elements about KADS and GDM. Then we will explain how we give a pragmatic content to the initial idea of the GDM approach and we will draw some lessons to conclude.

1. Modelling and Knowledge Acquisition

1.1 A Recall about KADS Models

KADS, to break off with transfer approach of knowledge acquisition, emphasises on modelling **[Wielinga 92a]**. We agree with this emphasis. The central model in KADS is the model of expertise (organised initially in four layers) that is the result of KA. We will not consider anymore the fourth layer, the strategic one, which had no real development. The three other layers play different roles: the domain layer embodies knowledge describing a theory of an application: concepts, properties of concepts, relations between concepts, attributes of concepts and so on...,the inference layer is a functional view of the target system (a dependency diagram), the task layer embodies a goal decomposition and control view of the target system.
In KADS, interpretation models, which are expertise models without a domain layer, are proposed to provide guidance for KA. The inference layer, as a **sight** on the domain layer, has to name the different roles of knowledge, and the different inferences (specifying inputs and output and abstract function grounding the system dynamism). The task layer, as a **sight** on the inference layer, has to organise the path between inferences according to goal achievement and control. Thus, roles plus inferences (making a partition on knowledge to acquire), as well as goal decomposition and control, must be an effective guidance for KA because they indicate, at least partially, what knowledge chunks are needed.

1.2 The KADS Process of Model Driven Knowledge Acquisition

To privilege top-down KA, KADS has interpretation models, which are tools providing guidance in a model based approach **[Le Roux 93b]**. Thus, the first problem, when working on a particular application, is to select a suitable interpretation model. In KADS, this choice is related to some features of the application allowing to traverse a decision tree; When a leaf is reached, we have an interpretation model. The borderline between features of an application and domain knowledge seems difficult to locate and the amount of information required to reach a leaf may be important. If a leaf is unreachable, no interpretation model is available, so model based KA is impossible for a while, and it remains bottom-up approach. Interpretation models are defined once for all (even if the library is extensible). The leaf reached may provide an interpretation model that suits more or less to the application. No guidance is provided to further refine the model, whose grainsize or other features may be an impediment. Thus, in KADS, model driven KA is seen as a process embodying two strongly separate steps, the former step is to select/refine the model, the second step to organise KA according to the model. We assume that this two steps' process, analytic step plus KA step, is not realistic, and we propose a practical issue for model driven KA: to design it as a combined process of constructive modelling and knowledge elicitation.

1.3 A First Idea: the Generalised Directive Models

The idea of GDM **[Van heijst 92b]** is that it is possible and useful to provide a context sensitive grammar able to transform a current inference structure into a more specified one, according to the application knowledge already acquired. KA is seen as a recurrent process of refinement where

knowledge is gradually acquired and inference structure is progressively refined. A simple presentation **[Van Heijst 92b]** for a classification application task starts with the following model:

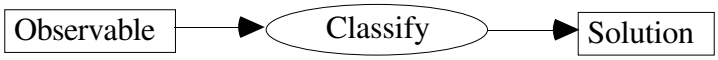
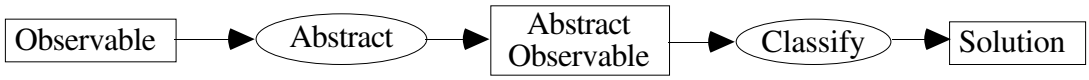


Figure 1: an initial directive model

It provides a partition of domain knowledge into observables and solutions. Different conditional rules should be available to expand this first abstract model. For example, if the elicited knowledge about observables shows that they are hierarchically organised and that there is a lot of observables, the refinement could be:



Rule applied to obtain the model above from the former:	
Conditionnal Part:	• *IF input has a hierarchical structure* • *AND IF size(input) > 10 THEN*
(input Classify output)	**- (input Abstract abstract-input Classify output)**

Figure 2 a refined directive model and related rule

Some further refinements are then possible about Classify or Abstract, then begins another step of the combined KA and modelling process.

1.4 Our View of Model Driven Knowledge Acquisition

From the general idea of the GDM, we have improved the approach with the aim to build up a tool providing advice and guidance both for modelling and acquisition. The improvements are: - to take into account the modelling of the task structure - to consider the modelling of the domain layer - to provide an effective context sensitive grammar.
Neglecting the task structure, as in the Acknowledge version of the GDM, appears to us as a problem, because of the importance of goal decomposition for modelling and control acquisition (acquiring control is often underlined as a problem for KA).
The emphasis on domain modelling is, from our point of view, important to understand the rationale of the models, and to provide an effective guidance for elicitation.
An effective context sensitive grammar must take into account domain knowledge already acquired, but also goal decomposition, model schema of the domain, control and so on...

Our work is to develop a synthetic task of application modelling, in contrast with the analytic approach in KADS. Our view is then to consider the KA process as a recurrent and combined process of knowledge elicitation and model construction/refinement.

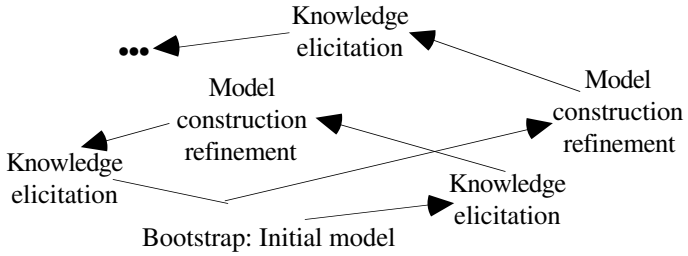


Figure 3: The spiral of the recurrent process of model construction and knowledge elicitation

2. Scenario for a Constructive Approach

We have developed different scenarii to experiment with the improved GDM approach on different tasks, diagnosis, assessment and so on...**[Le Roux 93c]** In this section, we present an example for an hypothetical application. It shows how a knowledge engineer could refine an abstract diagnosis model into the well-known Cover & Differentiate model **[Eshelmann 88a][Eshelmann 88b][Schreiber 92][Le Roux 93a]**. As our purpose is to improve a generative method to construct models, it may appear quite strange to develop examples leading to classical problem solving methods. We consider that a great experience and rational are underlying these methods and that the minimal expectation for a generative approach is to be able to construct such methods. Moreover, it is a kind of empirical inquiry about modelling.

The following figures are all organised as the Figure below, showing a rewrite rule and the result of its application, except the figure concerning the final model:

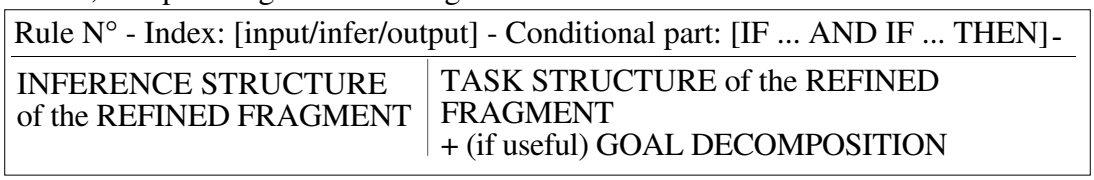


Figure 4: Structure of a rewrite rule of our grammar

Looking at the modelling task in all its bearings implies the consideration of different types of roles. We distinguish two main types of roles, Static and Dynamic ones **[Wielinga 92b]**:

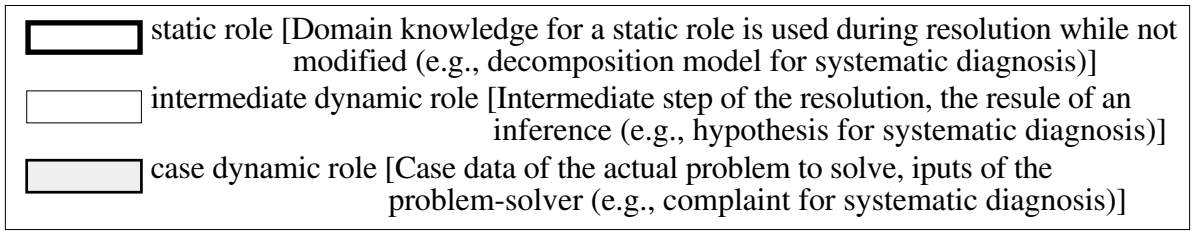


Figure 5: Notations for types of roles,

2.1 From Simple Diagnosis to Cover & Differentiate

Initially the only available character of our hypothetical application is that it is a diagnosis task where solutions are a-priori known and not constructed, the model provided is then the result of the application of rule R3:

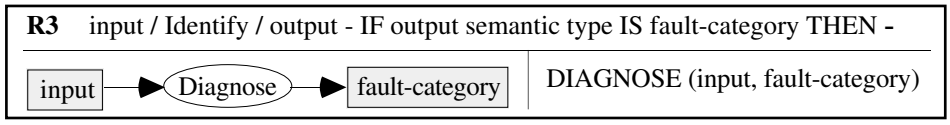


Figure 6: A rewrite rule and the result of the rewriting

This initial model to begin KA is very coarse grain. The amount of information allowing to select the model may be very slight, it's nevertheless possible to begin model based KA. The context sensitive grammar will ease and will lead the refinement of the model. When more information is initially available, it's possible to begin with any model more specified.
The question occurring for the knowledge engineer could be: What *kind of inference is made by Diagnose to build fault-category from input and what is input?* The possibility exists at least of three kinds of relations between input and fault-category: a heuristic relation (leading to a kind of heuristic classification), a causal relation, a part-of relation (leading to systematic diagnosis). For our hypothetical application, let us suppose that, for our hypothetical application it is a causal relation, then the rule R4 is applied:

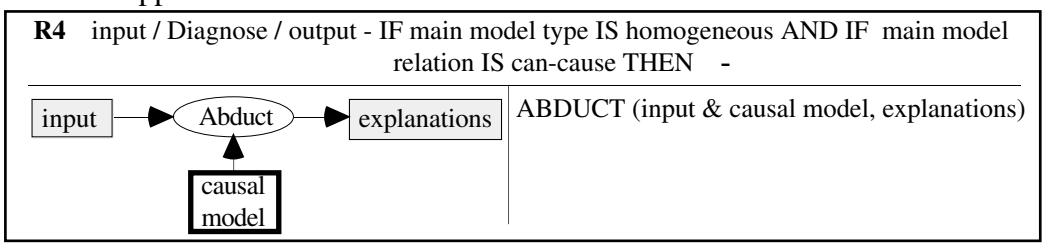
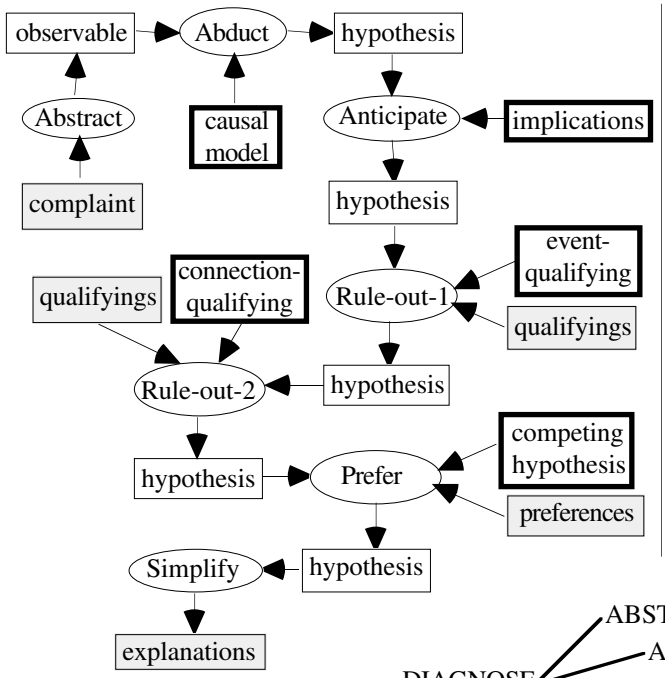


Figure 7: a refined directive model for diagnosis application

According to our view of KA as a recurrent process of knowledge elicitation and modelling, we have to acquire the domain knowledge for causal model. The model schema (and related model ontology) corresponding to this role is a causal network with nodes and can-cause links between nodes. Only a part of the domain knowledge has been acquired, and the following questions occur: *Is it possible for any peculiar set of observables to get many explanations ?* (this question is strongly related with some characters of the causal model: a highly connected one entails many explanations to occur), *Is the application purpose to have all the explanations or only as few as possible and the more general ?* The causal model of our hypothetical application is highly connected and the goal is to have only the

more general explanation. So, the modelling tool will provide advice and guidance to refine/construct the model according to this goal (to eliminate a part of the explanations...), evidencing reasoning methods and kind of knowledge to achieve it.

.../...We continue this process of refinement/acquisition to obtain at the end the following model through seven steps of refinement supported by seven different rules of our grammar:



WHILE complaint
ABSTRACT (complaint, observable)
ABDUCT (observable & causal model, hypothesis)
WHILE implications
ANTICIPATE (hypothesis & implications, hypothesis)
WHILE qualifyings1
RULE-OUT-1 (hypothesis & event-qualifying & qualifyings1, hypothesis)
WHILE qualifyings2
RULE-OUT-2 (hypothesis & connection-qualifying & qualifyings2, hypothesis)
WHILE preferences
PREFER (hypothesis & competing hypothesis & preferences, hypothesis)
SIMPLIFY (hypothesis, explanations)

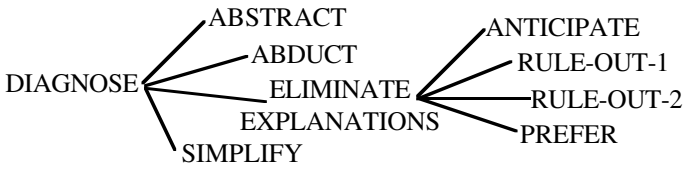


Figure 8: a final model of cover & differentiate (see **[Schreiber 92]***)*

3. Some Lessons as a Conclusion

A Least Commitment Strategy

This constructive approach is grounded on a least commitment strategy, i.e. an application of Occam's razor to modelling: *"all things being equal, parsimonious models must be preferred."* The GDM provides the most simple model according to the information available about an application, this is a first answer to the problem of model selection. This has also significant consequences on some problems of model driven KA: granularity, genericity, reusability.
Our approach provides a model with a suitable grain according to the available amount of information about an application: it allows a **variable granularity for the models**. The genericity problem is now on the naming of the inferences and roles because names are index for the grammar, and also on the abstract function that embodies each inference and related possible model schema: **genericity of basic modelling components.** As well the problem of reusability changes, the specification of the grammar is mainly focused: our approach brings a relevant framework to think about the problem of reuse.

A Context Sensitive Grammar

Our focus here is on the conditional part of the rules of the grammar. According to our experiments, the following elements are concerned: a character of the domain knowledge corresponding to a static role (mainly about its structure: model schema) - a character of the domain knowledge corresponding to a dynamic role (mainly about size and structure of this role) - a failure in acquiring inference knowledge - the possibility for an inference to have a certain static role as input (knowledge about modelling) - a specific feature of the application at hand (mainly related to goal decomposition).

A Scheduling on Knowledge Acquisition

We assume here that a scheduling on KA is possible and useful. The models provided in the framework of a methodology are supposed to permit the KA task organisation. One way to organise this task is to know what is the most important to acquire at time: inference structures provide the dependencies between pieces of knowledge to acquire. In our approach, the dependencies between pieces of knowledge appear gradually, as the model is itself progressively refined/constructed, providing a manageable scheduling for KA.

About Modelling and Models

We assume that our approach is a relevant framework to improve a more complete and coherent ontology, and to provide more consistent models, while keeping great flexibility. We hope that as a consequence it will authorise a better understanding of the models that we use in Knowledge Acquisition. Hence, these models could become more suitable to the technical requirements of our scientific work. We claim that it is possible in this framework to evidence some syntactic, semantic and pragmatic constraints on Knowledge Acquisition models, providing therefore some theoretical advances for model-driven Knowledge Acquisition.

References

[Aamodt 92] Agnar AAMODT, Bert BREDEWEG, Joost BREUKER, Cuno DUURSMA, Christiane LOCKENHOFF, Klas ORSVARN, Jan TOP, Andre VALENTE, Walter VAN DE VELDE: The common KADS library - 1992 - *Esprit project P5248*

[Aben 92] Manfred ABEN: On the specification of knowledge model components - 1992 - *BANFF*

[Breuker 87] Joost BREUKER, Bob WIELINGA, Maarten VAN SOMEREN, Robert de HOOG, Guus SCHREIBER, Paul de GREEF, Bert BREDEWEG, Jan WIELEMAKER, Jean Paul BILLAULT, Massoud DAVOODI, Simon HAYWARD: Model driven KA: interpretation models - 1987 - *Esprit project 1098*

[Eshelman 88a] Larry ESHELMAN: MOLE : A Knowledge-Acquisition Tool for Cover-and-Differentiate Systems - 1988 - *Automating KA for Expert Systems, ed Sandra Marcus, Kluwer Academic Publishers*

[Eshelman 88b] Larry ESHELMAN, Damien EHRET, John McDERMOTT, Ming TAN: MOLE : A tenacious KA tool - 1988 - *Knowledge based systems Vol 2*

[Hoog 92] Robert de HOOG, Rob MARTIL, Bob WIELINGA, Rob TAYLOR, Clive BRIGHT: The common KADS model set - 1992 - *Esprit project P5248*

[Le Roux 93a] Bernard LE ROUX, Philippe LAUBLET: From simple diagnosis to Cover & Differenciate: a trial for a conceptual and compositional approach - 1993 - *Proceedings 3rd KADS user meeting Munich*

[Le Roux 93b] Bernard LE ROUX, Philippe LAUBLET: Modèles de méthodes de résolution de problèmes et acquisition des connaissances - 1993 - *Proceedings JAC 93 St Raphael*

[Le Roux 93c] Bernard.LE ROUX, Philippe LAUBLET, A Conceptual and Compositional Approach For Knowledge Acquisition: Examples and Lessons, Laforia Technical Report N° 93_6 Mars 1993

[Marcus 88] Sandra MARCUS ed.: Automating KA for Expert Systems - 1988 - *Kluwer Academic Publishers*

[Newell 82] Allen NEWELL: The knowledge level - 1982 - *Artificial Intelligence*

[O'Hara 91] Kieron O'HARA, Nigel SHADBOLT: Generalized directive models supporting the KA process - 1991 - *Report VITAL/T2.1.2, Esprit project, VITAL*

[O'Hara 92a] Kieron O'HARA, Nigel SHADBOLT, Philippe LAUBLET, Manuel ZACKLAD, Bernard LE ROUX, Mandy MEPHAM: KA Methodology in VITAL - 1992 - *Report VITAL/DD2.1.2, Esprit Project , VITAL*

[O'Hara 92b] Kieron O'HARA: A representation of KADS-1 interpretation models using decompositional approach - 1992 - *Internal Report VITAL/2.1.2, Esprit Project*

[Puerta 91] Angel PUERTA, John EGAR, Samson TU, Mark MUSEN: A multiple-method knowledge-acquisition shell for the automatic generation of KA tools - 1991 - *KA* v 4 n° 2 (also BANFF 91)

[Schreiber 92] Guus SCHREIBER, Bob WIELINGA, Hans AKKERMANS: Differenciating problem solving methods - 1992 - *EKAW*

[VanHeijst 92a] Gertjan VAN HEIJST, Peter TERPSTRA, Bob WIELINGA, Nigel SHADBOLT: construction of abstract models of problem solving processes - 1992 - *ECAI*

[VanHeijst 92b] Gertjan VAN HEIJST, Peter TERPSTRA, Bob WIELINGA, Nigel SHADBOLT: Using GDM in KA - 1992 - *EKAW*

[Wielinga 92a] Bob WIELINGA, A.TH. SCHREIBER, J.A. BREUKER: KADS: A Modelling Approach to Knowledge Engineering - 1991 - *KA* v4, n°1

[Wielinga 92b] Bob WIELINGA, Walter VAN DE VELDE, Guus SCHREIBER, Hans AKKERMANS: The KADS knowledge modelling approach - 1992 - *JKAW*