PLANFORM-KA tool: Towards a Methodology of Knowledge Acquisition in AI Planning

Christophe Doniat
Tel.: +(00 44) (0) 161 295 2922
E-mail: c.doniat @salford.ac.uk

Ruth Aylett
Tel.: +(00 44) (0) 161 295 2112
E-mail: r.s.aylett @salford.ac.uk

University of Salford
Centre for Virtual Environments
Business House
Salford M5 4WT
Fax: +(00 44) (0) 161 295 2925

Keywords
Knowledge engineering for planning, knowledge acquisition for planning.

Abstract

Our work, embodied in the PLANFORM-KA tool, aims to support domain experts in the knowledge acquisition phase. The tool applies the PLANFORM-KA model supporting the encoding of interviews with domain experts or their direct interaction. A skeleton ontology is generated in which we configure the tool of knowledge acquisition. This ontology focuses explicitly on a particular domain and implicitly on the background and knowledge domain of a domain expert. We illustrate the methodology by applying it to an example domain called 'Entertaining a foreign visitor'. We summarise the results and provide conclusions.

1. Introduction

In the field of AI Planning, as in other knowledge-based applications, knowledge acquisition is recognised as a critical element of the knowledge engineering (KE) process and is also known to be the most problematic. By KE we mean the process that deals with (i) the acquisition, validation and maintenance of a planning domain model, (ii) the selection of appropriate planning machinery and heuristics for a particular application. The work discussed here consists of steps to bring into KE for planning an approach, embodied in a knowledge acquisition tool, based on specifying classes of problems and formalising both an ontology and a model of expertise for a domain.

A feature of current approaches to the specification of planning domain models is that relevant knowledge concerning acquisition, validation and maintenance, used in the process, is never explicitly recorded. Thus such knowledge is effectively lost to the community, only being retained in the head of an experienced individual expert. Creating a domain model and managing it efficiently then never gets any easier since there is no way of reusing past experience. The approach discussed here draws on work in the KBS (knowledge-based systems) community, which sees the combination of ontologies and generic problem-solving methods as a way of addressing this problem [1]. It supports the capture and structuring of relevant knowledge about a domain and its intelligent behaviours [2] because they play an important role in the choice of an appropriate problem-solving method, possibly configured from complex components stored in a library [3].

1.1. Our contribution

Figure 1 shows a knowledge acquisition process with two different extraction processes. The first (bottom) moves from protocol to problem specification. Here by protocol we mean raw domain knowledge - transcripts, documents, interviews, observations. By problem specification we mean a definition or description of an application domain represented as a set of choices at a particular level of abstraction in an ontological hierarchy. Thus 'Entertaining a foreign visitor', the domain used for the experiment reported later, is a specification.

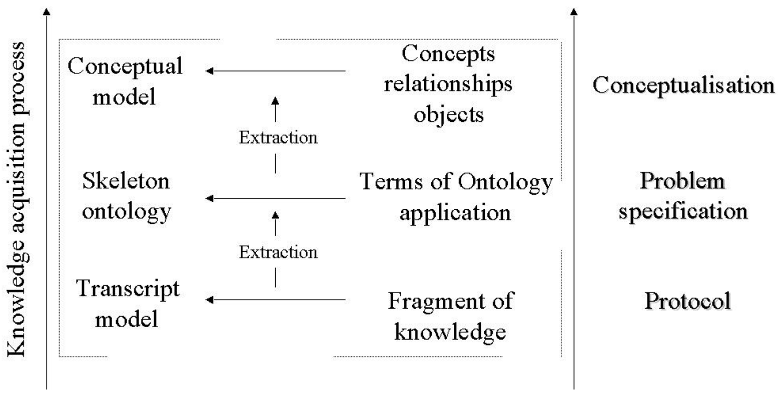

Figure 1. Knowledge acquisition process.

The second level of extraction (top) moves from specification to conceptualisation. By conceptualisation, we mean a stable and restricted formal representation (concepts, relationships and objects) with defined structure and behaviour. It is worth noting that a given specification can produce a number of different solutions, justifying our interest in the extraction process. Clearly there is an interaction or cooperation between these levels during the knowledge acquisition process as discussed in more detail in section 2.

The conceptual model (top of figure) can be represented very naturally using the object-oriented approach. In fact, the generic entities used in these models embody both structure and behaviour in the same object.

An ontology is defined [4] as a specification of a 'domain of discourse' within which conceptualisations can be created. Intuitively, it involves the semi-formal or formal knowledge representation of a domain expertise mainly in hierarchies of concepts and relationships. This approach is required when knowledge must be shared between heterogeneous knowledge systems or different problem-solving systems.

A skeleton ontology (middle of figure) may be seen as the description of an application domain restricted by a problem specification. Using the experimental approach, it is tricky to make an exhaustive analysis of all domain objects. Nevertheless, it is possible to analyse the problem specification and use it to define the relevant objects and relationships, using macroscopic properties that support appropriate choices. Broadly, the skeleton ontology is associated with (i) a particular domain, (ii) specification of a problem or problems that we want to solve, (iii) the reasoning that belongs to the studied domain and allows the specified problem to be solved.

1.2. Architecture

Our work focuses on knowledge acquisition for AI Planning using generic models (ontology, conceptual model). This research is a part of a wider project called PLANFORM [5] – which is constructing an open environment for building planning systems –. A component of this project is a methodology to specify and formalise the domain models. Figure 2 shows the main architecture of the PLANFORM-KA tool that contains the knowledge acquisition process. The conceptual model of the example domain discussed in this paper is the EVentus model as we will see in the next section.

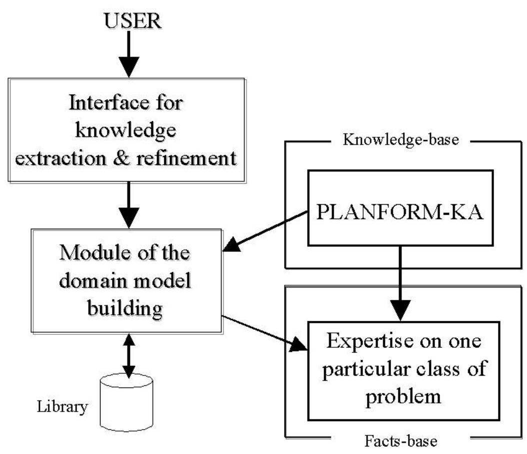

Figure 2. Architecture of the PLANFORM-KA tool.

On the left-hand side of figure 2, the user interacts via the *interface for knowledge extraction*, i.e., the user applies the *module of domain model building* to the study of a particular problem specification. The building of a new conceptual model might be carried out with or without an existing problem specification from the *library*. The result is recorded in this library. On the right-hand side, the overall knowledge base consists of the conceptual model of the knowledge acquisition process itself, called PLANFORM-KA and the *KA-Expertise* that is a part of this particular conceptual model that is instantiated for the current study.

Section 1 presented the architecture of our contribution. Section 2 shows PLANFORM-KA, a knowledge acquisition tool. Section 3 shows a case study called 'Entertaining a foreign visitor' and the using of PLANFORM-KA to build EVentus model. Section 4 discusses some related work. The conclusion summarises the results and further directions.

2. Knowledge acquisition: PLANFORM-KA

PLANFORM-KA is aimed at the generation of a valid conceptual model for the formalisation of a domain model from transcript to the conceptual model via a skeleton ontology. We consider several assumptions that restrict the knowledge acquisition process such as:

- It can be seen as a set of refinement processes carried out by the domain expert on a basic generic specification.
- It might use (i) a generic specification called ATO (Triangle model) and (ii) a definition of a conceptual model (see section 3.3).
- It will focus on two kinds of refinement process: (i) specialisation (ii) aggregation for which rules can be expressed to guide this process.

2.1 Conceptual model of PLANFORM-KA

The conceptual model of PLANFORM-KA (Figure 3) contains five major concepts. The *domain expert* is the subject acquiring the knowledge model, *KA-expertise* is the knowledge required to build the knowledge model, *model* depicts the knowledge of the expert domain, *KA-Task* describes the abstract process carried out by the domain expert and *KA-Process* describes the expert's particular behaviour. The KOD method (Knowledge Oriented Design) [6] was chosen to generate this model.

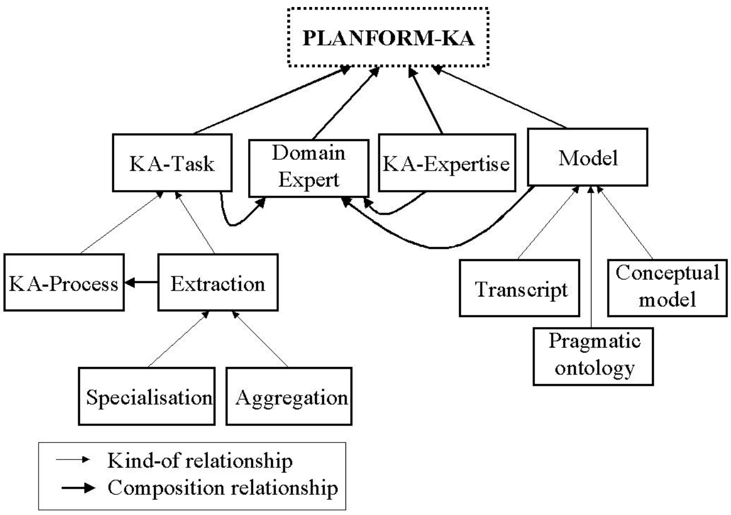

Figure 3. The PLANFORM-KA conceptual model (representation).

2.2. Concepts

The domain expert (DE) is the cognitive agent carrying out the process of knowledge acquisition. DE has a mental model of the real world expressed in concepts. The domain expert (concept) depicts the properties of this agent in relation to the carrying out of the KA-Process. This concept is central to this conceptual model since there are composition relationships with concepts like model, KA-Task and KA-expertise. The **intrinsic** properties are *Author* (name of this agent), and *Category* (expert, non-expert) because while an expert might want to create a new skeleton ontology for the library, the non-expert needs to reuse an existing skeleton ontology and conceptual model. The **extrinsic** properties are *agent knowledge* (*model*, *KA-Task* and *KA-expertise*). The DE's **behaviour** consists of the creation of a particular expertise choosing a combination of models according to the category of this agent. Then, the DE carries out a specific task called *KA-Process* to produce a domain model for a particular problem.

Model represents the properties of a part of agent knowledge. In this representation, the model is specialised according the particular relationships in three models: transcript, skeleton ontology and conceptual model. This **intrinsic** property is its name.

Transcript represents the properties of documents such as free-texts or graphics. The **intrinsic** properties are inherited from the Model concept. The **extrinsic** properties depict a set of texts, a set of graphics and a set of transcript since the transcript is a composed document. The **behaviour** consists of creation of texts, graphics and composition.

Skeleton ontology represents the properties of a description of an application domain restricted by a set of choices as discussed above. The **intrinsic** properties are inherited from Model concept and there is also a description of the problem. The **extrinsic** part is a set of non-ordered terms that belong to the studied problem and allow it to be solved. The **behaviour** consists of the creating of the list of terms.

Here, we introduce a particular ontology, ATO model (Triangle[1] domain [7] in figure 4), for AI Planning. With this ontology, planning domains can be specified in terms of (i) objects (from the Environment) can exist that through their state changes elaborated by (ii) Task(s) carried out by an (iii) Agent. In the other words, Agent can plan Task, which specifies Environment (actualisation of states or structures of objects). Then Agent(s) execute(s) the Task(s) and the Environment supplies new stimuli (transformation).

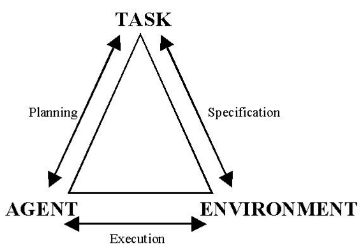

Figure 4. The ATO model as a particular skeleton ontology

[1] Hereafter, we use the ATO acronym for Agent-Task-Object model.

This model, as discussed in [7], is generic and allows the specification of a class of problem. This model is open since it supports the definition of a partially characterised TASK, i.e. without a *desired end-state*. This ATO model is a particular skeleton ontology containing a set of terms {Agent, task and object}. It has to be instantiated according the specification of the given problem. Then it may be used in the first extraction phase to generate a set of scenarios.

Conceptual model represents a stable and restricted formal representation of a set of concepts, relationships and objects according to a set of assumptions, which restrict this class of problem. This concept (Figure 5) allows generating of a particular domain model. The **intrinsic** properties are its title and purpose. The **extrinsic** part is described in terms of assumption, concept and relationship concepts. The **behaviour** consists creating an organised set of relationships from the instances of these three concepts.

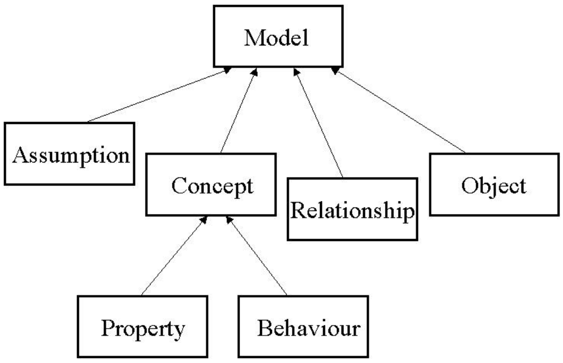

Figure 5. Composition of a conceptual model

Assumption represents a constraint on an idea that solves a given class of problem. The **intrinsic** property is its label. The **extrinsic** property is a set of constraint statements and the **behaviour** consists of the creating of this set.

Concept depicts a set of specific properties of an entity with respects the given assumptions. The **intrinsic** property is its label. The **extrinsic** properties are a relationship with a set of parents (multiple inheritance), an intrinsic structure (set of property about the nature of the concept), an extrinsic structure (set of property about composition relationships), a set of behaviours and a set of its objects (instances). The **behaviour** consists of the creating of these properties according to the specification of the given problem.

Property represents a specific feature of a concept. A property might be either intrinsic or extrinsic. The **intrinsic** properties are its label and category {intrinsic, extrinsic}. The **extrinsic** properties are a state of the property, a kind of type and a relationship that is a dynamic attribute that translate a composition relationship with another concept. The **behaviour** consists of the creating of these properties according to the specification of the given problem.

Behaviour represents properties in a typical way, which translates into a problem-solving method. The **intrinsic** properties are its label, a proposition that describes the use of any properties of a concept in a typical way, a category depicts the nature of the behaviour {static, dynamic}. Static behaviour is used when it manages the transformation of a set of properties. Dynamic behaviour is used when it manages a modification of a concept structure. The **extrinsic** property is a set of composition relationships with the existing properties that play a role in this behaviour. The **behaviour** consists of the creating of these properties according to the specification of the given problem.

Object represents the properties of an instance of a concept. The **intrinsic** property is its label. The **extrinsic** property is an instantiation relationship with its parent. The **behaviour** consists of the creating of the link with the concept, which is its parent.

KA-Expertise represents a particular composition of the managed object according to the category of a particular domain expert. The **intrinsic** property is its label. The **extrinsic** properties concern a set of composition relationships according to the specification of the given problem with transcript, skeleton ontology, instances of skeleton ontology and conceptual model. This **behaviour** consists of the creating (i) a transcript, (ii) a particular skeleton ontology either from transcript or an existing skeleton ontology from the library, (iii) a set of scenarios and (iv) a particular conceptual model.

KA-Task represents the notion of the task of knowledge acquisition used by the domain expert to generate his/her models. We use the term of KA-Task at the conceptual level because, in our process of knowledge acquisition, a KA-Task is defined as a set of tasks that concern the building of a conceptual model, called KA-Expertise, using two main behaviours: specialisation and the aggregation.

KA-Process represents the particular task that organises a knowledge acquisition and refinement process. The **intrinsic** property is its label. The **extrinsic** property is an KA-Expertise-to-transform, the chosen expertise with respect to the categories of the domain expert. The **behaviour** is to construct the structure and the content of the KA-expertise concept triggering the following list of these tasks: (1) create a transcript, (2) generate a skeleton ontology model (creates or customises existing one), (3) create a new set of scenarios (organisation of knowledge), (4) create a new conceptual model.

The **behaviour** (1) consists of the creation of texts and/graphics representing the field of a new problem.

The **behaviour** (2) may be triggered in to two ways. The first way is if a new skeleton ontology model must be created. From the transcript, the domain expert chooses the generic terms to build this skeleton ontology. This new ontology depicts a new concept that will allow the creation of new scenarios. Alternatively, the second way is if the domain expert wants to use an existing skeleton ontology. The expert chooses a problem specification from the library (see figure 2 in section 1.2) and uses it as a template to acquire new scenarios. This point is illustrated by the ATO model (see above). At this step, the expert constrains the domain with some choices.

In the **behaviour** (3), the expert defines a set of new scenarios that contain the contexts for the defined problem specification. Each scenario is created through a particular skeleton ontology. The expert uses a specialisation task to form a set of sub-classes of problems for the main class of problem.

The **behaviour** (4) is the final step that consists of the creation of the generic scenario model, also called conceptual model. This model depicts a stable and restricted class of problem. The first step is the factorisation of the scenarios to form a generic scenario model. This process consists of collecting together all objects that share the same semantics within the same concept. There must be at least two elements to form a concept. The second step involves the specialisation/aggregation of concepts to form the conceptual structure. This process supports the definition and description for each concept of its the list of properties:

- specialisation to add/modify its attributes;
- aggregation to add/modify its relationships.

Then, the domain expert describes the set of behaviours in which the properties play a role.

Finally, the expertise on this particular class of problem is recorded in the library as shown in figure 2 (section 1.2).

3. A case study and methodology: Entertaining a foreign visitor

In this section, we present an application of the methodology described above to the problem specification 'Entertaining a foreign visitor to your lab at the weekend'. We conducted this experiment with five people who verbalised their knowledge about how they would solve this problem during interviews. We chose this problem because (i) people knew about it (drew on general rather than specialised knowledge) and it was not difficult to capture it, (ii) it was an example of a planning domain and (iii) it demonstrated a knowledge acquisition (KA) process. The interviews contained the unstructured knowledge (discourse) and sometimes some notes such as graphics, plans and other material describing knowledge and activity (explicitly/implicitly) both about the case study and the KA process itself.

It is important to understand the level of abstraction at which such a sample problem must work. The Planform toolkit as a whole will be used to create a domain model within which a number of specific tasks can be planned. Thus the experiment does not start with such a specific task, but with the generic problem specification. Subjects were asked to explore the generic domain model that would be needed to plan within the domain of the problem specification and to support the solving of a number of specific tasks. Note that a more abstract version of this problem would be to replace 'your lab' with 'a lab' where this might be anywhere in the world potentially. An instance of a specific task would be something like "Professor Stein from GMD Germany is to be entertained on Saturday May 9th'.

3.1. Building of a skeleton ontology

The first extraction phase gives us a skeleton ontology (figure 3) that isolates four essential terms: thing, activity, context and visitor. We then built five scenarios with the shared knowledge of these domain experts to find out how each expert defines and organises different knowledge to obtain a part of the domain model for this problem specification. We have used this technique (KOD method) to obtain an accurate process for knowledge acquisition and to build the conceptual model through the set of scenarios.

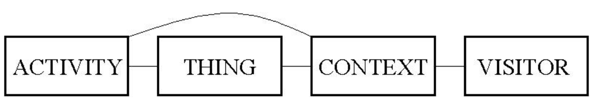

Figure 3. Terms of skeleton ontology.

We measured the coverage of skeleton ontology. Five scenarios were built with the four general terms: Thing, Activity, Context and Visitor. In table 1, we present the number of objects and properties given by each domain expert. We see that the knowledge about this particular domain varies between domain experts giving different numbers of objects and properties for each term. Note also we checked that the skeleton ontology is completely filled out.

Scenario[2]	***Objects***				***Properties***			
	T	***A***	***C***	***V***	***T***	***A***	***C***	***V***
1	9	4	1	1	7	4	4	1
2	5	6	1	1	3	1	1	3
3	8	7	2	2	1	1	2	2
4	5	5	3	1	1	1	2	2
5	13	7	1	2	2	1	2	3

Table 1. The coverage of the skeleton ontology

3.2. Building of conceptual model: EVentus model

The second extraction gives us the conceptual model (part of the domain model) called EVentus[3]. Thus is a conceptual model for the class of problem called 'Entertaining a foreign visitor' which considers a generic foreign and entertainment rather than a particular individual. Five essential concepts were defined: visitor is a locus of the EVentus model and describes a real visitor, activity and context concepts describe the behaviour of a visitor, plan describes a set of alternative plans and thing describes places and events used during the activity. Figure 4 shows the simplified conceptual model with the five major concepts and the additional concepts (in grey) that allow the specification and definition of a domain model from this problem specification.

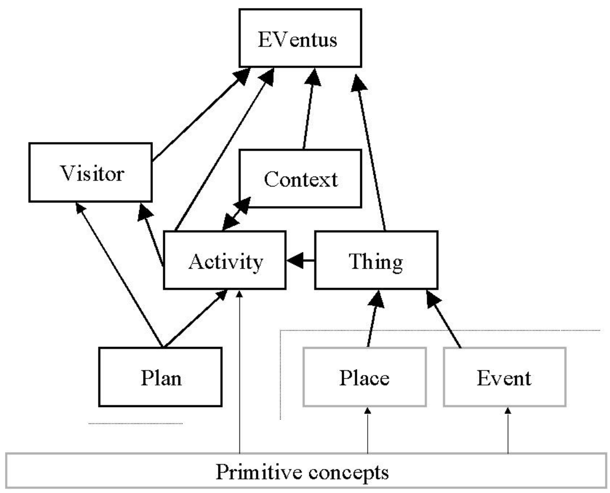

Figure 4.The simplified conceptual model EVentus

[2] Each scenario is designed through the four given general terms: Thing (T), Activities (A), Context (C) and Visitor (V).
[3] EVentus means 'to happen' in latin. EV means Entertaining Visitor.

We measured the stability of our conceptual model with the degree of specialisation and aggregation. Table 2 shows the implicit behaviour behind each concept through their specialisation and aggregation and orders each in a hierarchy of generality as a result, with the most general concept first in the table.

Concept	Specialisation	Aggregation
Eventus	-	4 (Thing, Activity, Visitor, Context)
Thing	2 (Place, Event)	3 (Thing, Dish, Drink)
Activity	1 (Plan)	6 (Resource, Thing, Date, Plan, Pre/Post-condition)
Visitor	-	4 (Activity, Plan, Visitor, Capability)
Context	2 (Pre/Post-condition)	-
Pre-condition	-	4 (Weather, Time, Resource, Relationship)
Post-condition	-	1 (Date)
Plan	-	4 (Date, Activity, Pre/Post-condition)
Place	11	-
Event	3	-

Table 2. The degree of the abstract level

3.3. Summary

A knowledge acquisition process has been carried out to capture knowledge and build a domain model for a particular problem specification. Figure 5 shows the implementation of this process for a problem specification through the KA-Expertise concept. On the left-hand side, point 1 is when the expert (DE) describes some knowledge about the problem. Point 2 is when the DE defines some generic terms, which are the elements of a particular skeleton ontology. Then DE customises the generic skeleton ontology into the specific skeleton ontology called ATO (see definition in section 2.2). Again, it is also customised with the specification from the given problem. Point 3 is the first extraction to capture some structured scenarios relating to the background of DE and the given problem specification. Point 4 is the second extraction to build a conceptual model from the scenarios. On the right-hand side, the KA-Process is running with further hypotheses about the nature of an AI Planning problem. Point 1 is the same. Point 2 is when the DE wishes to customise a skeleton ontology into a specific skeleton ontology and populate it. Point 3 is the same. Point 4 is when DE customises a generic conceptual model into a specific conceptual model.

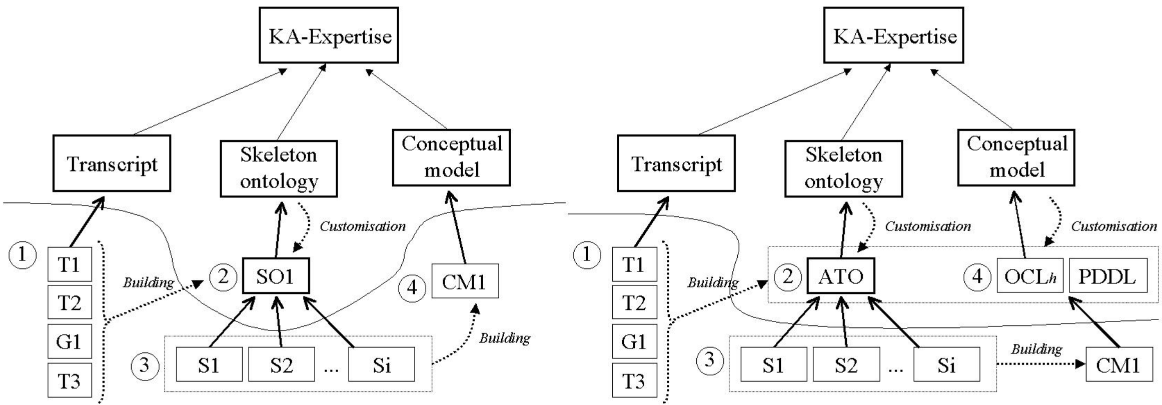

Figure 5. Expertise part of PLANFORM-KA
(Left - specific view for a general case. Right – generic view for a planning case)

T1, T2, T3 are instances of transcript. G1 is a graphic and instance of transcript. SO1 and ATO are instances of a skeleton ontology. S1, S2 …Si are instances of SO1 or ATO. OCLh is an instance of a conceptual model. CM1 would be either an instance of conceptual model (left-hand side) or an instance of OCLh (right hand-side). OCLh [8] is in this case a specialised conceptual model embodied in an enriched language to solve planning problems. PDDL [9] is another example of the conceptual model.

4. Related work

Many specific approaches propose a set of solutions for the acquisition, the representation and the sharing/reusing of knowledge in general, since this topic has been studied extensively in the KBS community since the 1980s. Some of them are more specialised in the first extraction of knowledge proposing a generic way to capture knowledge. For instance, Protégé [10] is an approach supported by a tool that captures new ontologies, and offers a library of problem-solving methods – For example propose-and-revise – to combine with them. EXPECT [11] is a framework and knowledge based system to acquire and represent problem solving method capabilities. PLANET [12] is an ontology for the representation of plans in the AI Planning field and is very relevant to the more extended framework discussed here. In other approaches, the answer for a given problem is built through a combined set of different techniques (AI methodologies, for example KOD [6], KADS [13]) according the major aim (diagnosis for example [14][15]).

5. Conclusion and further work

Surprisingly, given the amount of work in the KBS community in general, knowledge acquisition has not been widely studied in AI planning. Yet applying planning systems to real-world problems requires a systematic approach to knowledge acquisition and a methodology supporting reuse rather than ad-hoc adaptations of specific planning systems by particular individuals whose expertise remains private and invisible. The work discussed here represents some steps in this direction.

5.1. Conclusion

Our work consisted in demonstrating the value of the KA model called PLANFORM-KA, supporting an understanding the KA process. We have presented the basic steps of a simplified methodology to build a skeleton ontology and then a conceptual model, and to configure a typical knowledge acquisition phase so that the domain model can be acquired for a given problem specification.

Second, we have validated our KA process through the building of the case study ‘Entertaining a foreign visitor' and shown some results, such as the coverage of the skeleton ontology and the degree of specialisation and aggregation of the conceptual model. The coverage shows whether the ontology (i) has enough richness to cover the problem, (ii) is completely filled by the scenarios. The degree of specialisation and aggregation measures the generality of concepts in a particular problem. Work will now continue to develop the computer-based tool to support the methodology.

5.2. Further work

Several directions are possible for further work. It seems wise to think about the interaction/cooperation with the domain expert during the different steps of acquisition (extractions). The exploration of the results may uncover relevant knowledge acquisition rules based on the types of ontology coverage and the degrees of specialisation and aggregation of conceptual model. Such results might provide a basis for a stronger dialogue between the domain expert and our tool. We note though that making use of degrees of specialisation and aggregation will imply also a formal representation for the behaviours of the studied problem.

These new developments would allow a global trial of the PLANFORM-KA tool. Finally, we believe that this tool affords an excellent platform for storage of domain models.

References

[1] Musen M. Modern Architectures for Intelligent Systems: Reusable Ontologies and Problem-Solving Methods. In Chute (Eds), AMIA Annual Symposium, 46-52. 1998.

[2] B. Hayes-Roth and F. Hayes-Roth. A Cognitive Model of Planning. Representation and Reasoning. Readings in Planning. Morgan Kaufman Publishers. 1990.

[3] A. Valente. Planning models for the CommonKADS librairy. ESPRIT Project KADS-II. 1993. Available at http://www.swi.psy.uva.nl/usr/andre/publications.html.

[4] T.R. Gruber. A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition, **2**(5), 1993.

[5] Planform.An Open environment for building planners. Available at http://helios.hud.ac.uk/planform. 1999.

[6] Vogel, C. Le genie cognitif. Masson (Eds). 1988.

[7] R.S. Aylett, S. Jones. Planner and Domain: Domain Configuration for a Task Planner. Int. Journal of Expert Systems, **9**(2), 279-318, 1996.

[8] McCluskey T. L. and Porteous J. M. Engineering and compiling planning domain models to promote validity and efficiency, Artificial Intelligence , pp.1-65. 1997.

[9] McDermott D., et al. PDDL --- The Planning Domain Definition Language. In Machine Intelligence 4. D. Michie, ed., Ellis Horwood, Chichester (UK). 1998.

[10] Fridman-Noy N., Fergerson R.W. and Musen M. A. The knowledge model of Protégé-2000: combining interoperability and flexibility. 2th Int. Conf. on Knowledge Engineering and Knowledge Management (EKAW). Juan-les-Pins (France) 2000.

[11] Gil Y. and Blythe J. How Can a Structured Representation of Capabilities Help in Planning? Proceedings of the AAAI – Workshop on Representational Issues for Real-world Planning Systems. 2000.

[12] Gil Y. and Blythe J. PLANET: A Shareable and Reusable Ontology for Representing Plan. Proceedings of the AAAI – Workshop on Representational Issues for Real-world Planning Systems. 2000.

[13] Anjewierden A., Wielemaker J., and Toussaint C. The KADS approach to knowledge engineering. Special issue. Shelley - computer aided knowledge engineering. Knowledge Acquisition Journal, **4**(1), 1992.

[14] Mercantini, J.M., Lochmann, R. and E. Chouraqui. Safety previsional analysis method of an urban industrial site. Scientific Journal of the Finnish Institute of Occupational Health, serie: People and Work, safety in modern society, pp. 105-109, **33**. 2000.

[15] Mercantini, N. Belanger, R. and E. Chouraqui. Etude d'un systeme d'aide au diagnostic des accidents de la securite routiere. IC'99. Palaiseau (France). 1999.