Formal explorations of knowledge intensive tasks

Jansen, M.G.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
This concluding chapter will provide answers to the research question formulated in chapter 1. The results of the previous chapters will be reviewed and put into perspective.

9.1 A review of the previous chapters

In the first chapter we have formulated the main research question as follows:

How can we give a knowledge-level characterization of the properties of knowledge intensive tasks?

This question was then refined into three research questions:

1. What does a conceptualization of a task look like?
2. How do we represent these knowledge-level conceptualizations in a formal way?
3. How can we use these representations to acquire a better understanding of the task?

In this section we will answer these questions on the basis of the results of the previous chapters. As a whole the chapters of the thesis can be seen as a series of explorations, aimed at answering the main research question, each of these can be matched to one of the three research questions. Three types of explorations can be distinguished:

1. Explorations regarding the nature of the conceptualizations of the task. (Chapters 2, 3 and 7.)
2. Explorations regarding the nature of the representation for the conceptualizations. (Chapters 6, 7 and 8.)
3. Explorations regarding the use of task ontologies and representations to specific problems and problem criteria. (Chapters 2, 4, 5 and 7.)
9.1.1 What does a conceptualization of a task look like?

A first analysis of a conceptualization of a task has been presented in chapter 2. The participation in the Sisyphus experiment can be seen as a 'pre-formal' study of the classification task. The approach taken was to focus on ontology construction rather than solving the problem instance itself.

Several domain ontologies were presented, as well as a task ontology for classification. The construction of such a task ontology proved to be an important step in the process of capturing the nature of the task. In addition a criterion and a task method (pruning) were identified. Hence, all elements of a knowledge-level conceptualization of a task were presented.

The task ontology for classification in chapter 2 differs from the one presented in chapter 3 (Adapting tableaux for classification). The first ontology gives a more detailed account of the conceptualizations of the classification task. The ontology in chapter 3 focuses on essential conceptualizations of the task and is therefore less detailed.

The main reason for this difference is that in chapter 3 the aim was to describe the problem solving process (or rather the development of knowledge in that process) of weak and strong classification. Hence, the task ontology here only plays the role of providing initial conceptualizations of the task.

A third set of conceptualizations for classification was presented in chapter 4, where a spectrum of classification criteria was identified. Again there are differences between the conceptualizations described here and the previous ontologies for classification. In this chapter the aim is not to describe a problem solving process but to identify the different solution criteria for the task. This change of viewpoint leads to a refinement of certain ontological elements, particularly those concerning the matching relation.

The ontology in chapter 2 is rich and detailed but informal and serves as input for the other chapters. The two following chapters focus in more detail on different aspects of the task and for that reason focus on essential elements in the ontology, which are then more accurately described.

A similar process takes place for the assignment ontologies in chapters 7 and 8. The ontology presented in the latter is very short, since it only focusses on grouping problems, which are identified as subproblems of assignment.

In chapter 7 the conceptualizations of the task are presented first in an informal way and are then refined into a formal representation of the task. There is much emphasis on the different problem variations within the task itself. Here we try to identify those ontological elements which are judged essential in describing the various problems which are covered by the task. The ontology presented in chapter 8 is very short, since it only focusses on grouping problems, which are identified as subproblems of assignment.

Hence, when describing task ontologies we have first selected those elements which are used extensively in the description of some aspect of the task. Then those elements were more refined and given a more detailed representation. As a consequence our ontologies for the same task differ from chapter to chapter as they are used to describe different aspects of the task.
9.2 How do we represent these knowledge-level conceptualizations in a formal way?

In this thesis we have used a variety of representations to formalize conceptualizations. The rationale behind this is that the choice of representation depends to a large extent on the intended use of the analysis of the task. For example, the choice of formal representation of classification in chapter 3 (Adapting tableaux for classification) was motivated by its intended use in the description of weak and strong classification problem solving. Propositional logic was chosen to show how analytic tableaux can be used to mimic problem solving behaviour of several task methods. The use of tableaux offered some insights into the nature of the problem solving for the task. One of the results was that the task method for strong classification can be described as an *abductive* method. In addition weak classification can be seen as a check for consistency which can be elegantly represented by analytical tableaux.

These results prompted some reflections about the adequacy of propositional representations in abductive background theories in chapter 5. The result was an investigation whether strict implication instead of material implication could be used. We presented a form of "strict abduction" as a possible alternative.

The representations in chapters 4 and 7 are also influenced by problem-solving aspects, namely the description of classification and assignment problems respectively. The aim of both these chapters was to give an overview of the different problem variations within the task. The choice of a representational formalism was here inspired by the analysis of those ontological elements which feature prominently in different problem descriptions. Hence the notions of "matching" (chapter 4) and "preferences" (chapter 7) are given formal interpretations in such a way that they can be used to describe a wide variety of classification problems and assignment problems respectively.

The choice of formalism in chapter 8 is an illustration of the same principle, as it was inspired by the use of *abstractions* in grouping, a subproblem in many knowledge-intensive assignment problems. Since abstractions can be associated with equivalence relations and partitions, one can represent the problem terms of a partition algebra. Hence, in this case the choice of the formal representation was motivated by a general characteristic of an ontological notion: *i.e.* abstraction.

At this point one may ask whether one can somehow generalize the task-specific representation developed in chapter 3, 4, 7 and 8. The answer given in 6 is affirmative. The key point underlying this chapter is that all problem/solution formalizations involve a notion of *knowledge state*. The use of this notion was inspired by the approach taken in chapter 3, where with the help of tableaux it is shown how knowledge about the solution changes during the execution of weak and strong classification. This insight prompted the development of a more general description of knowledge-level problem solving in terms of knowledge states and state transitions. As explained in chapter 6 the representational framework was motivated by ideas on information and knowledge as presented by Dretske [23], and the update semantics of Veltman [95]. This choice of formalism was meant to unify the several chapters into one general framework. Both the use of tableaux for classification and the use of partitions in grouping can be described in it. It should be noted that the framework does not provide a *syntactic* unification of the representations. Its purpose is solely to provide a *semantic* account of knowledge-level conceptualizations.
9.3 How can we use these representations to acquire a better understanding of the task?

Generally speaking the various formal representations were used in two ways: statically and dynamically. The static way of using the various representation involved a description of the problem variations within the task. This is most clearly illustrated in chapter 4 in which a spectrum of classification criteria is systematically generated on the basis of some elementary conceptualizations. The approach here is similar to spectra for diagnosis criteria in the literature [21, 89].

Similarly, chapter 7 offers an overview of different assignment problems. No such spectrum as in the case of classification criteria was presented due to the greater complexity of the task. Still, one gets a good idea of what different problems fall within the range of the task of assignment.

In summary, the static use of the ontological representation consists of the description of problem variations within a task. The dynamic use consists of the description of changing knowledge during problem solving.

The dynamic use of the representations was illustrated by various descriptions of problem solving. In chapter 3 propositional analytical tableaux (a decision method for propositional logic) were used for classification. The manipulation of these tableaux can be interpreted as modelling the dynamics of the problem solving process. The ‘semantic’ nature of tableaux illustrates that classification problem solving can be described as the search for a right model. This approach was made more explicit and taken a step further in chapter 6. The correspondence between tableaux and knowledge states was explained here in more detail.

In chapter 8 the partition lattice was used to formulate methods for grouping problems. It was shown that formal properties of this structure can be used in the description of methods to solve grouping problems. Here again the dynamics of problem solving can be seen as the manipulation of a formal structure, in this case the partition lattice.

As a unifying framework (not as a representational language) we presented a description of problem solving in terms of states and updates. This framework presented in chapter 6 is also aimed at a dynamic description of problem solving. In this case the underlying representational language (update semantics) is itself dynamically motivated. The framework is an attempt to capture the dynamics of knowledge-level reasoning and offers a particular view on knowledge. Problem solving for some tasks can be described as increasing knowledge about possible solutions.

9.4 Synthesis and discussion

The results of the thesis can be summed up in the form of an answer to the main research question: How can we give a knowledge-level characterization of the properties of knowledge intensive tasks?

1. By a good and extensive specification of the ontology of the task.
2. By a characterization of the type and criterion of a solution to problem instances of the task.
3. By the construction of a spectrum of solution criteria for the problem.
4. By a characterization of the solution space.
5. By identifying a solution with a model in some logical system and the search for a solution as the identification of that model.

6. By giving a knowledge level account of the reasoning of the task in terms of acquiring knowledge about possible solutions, and moving through the solution space.

Identification of the vocabulary of the task, some criteria for solutions and an initial high-level description of a task-template make up an initial description of a task. The vocabulary of the task is usually captured in a task ontology. Its construction can be seen as a prerequisite to task analysis.

The main contribution of the thesis in this respect is the fact that we presented a more rigorous specification of task ontologies than is usual in KE. Often ontologies give a specification on a higher level of abstraction and leave notions implicit which we have explicitly formulated. For example, comparing the classification ontology in chapter 4, with ones presented by Motta in [65] or Wielinga et al. [97], our spectrum of criteria describes much more criteria, which are all generated on the basis of a few ontological notions. The main difference in approach between those ontologies and the one compared here is that we focussed much more on the systematic generation of alternative classification criteria.

A point which features heavily in this thesis is the description of problem variation. Systematic descriptions of variations within a task have not been presented very often. As mentioned above, the construction of a spectrum of diagnostic criteria has been presented in the literature by Console et al. [21] and TenTeije et al. [89]. Since classification is similar to diagnosis, an interesting topic for further study would be to investigate in which respects these spectra can be compared to each other. One could also ask the question whether the systematic approach of chapter 4, when applied to diagnosis, would reveal more criteria than hitherto described.

The task ontology for assignment also provided a detailed account of the conceptualizations used in the description of assignment problems. In KE the task of assignment has been relatively ignored and no good task ontology seems to be available in the literature. The proposed ontology can be regarded as a first step to fill this gap. However, providing a spectrum of criteria for assignment is not that easy. Assignment problems have much more variations compared with diagnostic and classification problem types. Furthermore the preconditions and assumptions on the representation of domain knowledge can vary much more, and it is hard to generate all possibilities in a systematic way. However we have shown how one can compare and order different problems, for example by refining constraints and preferences. This goes well beyond the description of assignment problems as given by Puppe [71].

The principle goal of describing a spectrum of criteria is to be able to compare and structure different problems which fall within the range of the task. Another way of looking at a spectrum of criteria is to see it as the description of all post- and preconditions of the problems of the task. In that sense a spectrum can be seen as a functional characterization of the task. However, this only describes the static nature of the task. In addition one would also like a dynamic characterization of the individual problems and the associated problem solving processes.

Therefore a next step is to use the conceptualizations in such a way that the underlying formal structure of a typical problem becomes visible. We identified this structure every time with the solution space. As an example consider the use of partitions in chapter 8. Here the partition lattice consists of an ordering of possible solutions (partitions) in which problem solving can be described. Also when describing tableaux for classification (chapter 3) the solution space can be identified as the set of classes which are still considered viable.
Hence, where a spectrum of criteria offers a description of what a solution consists of, for different problem variations, the underlying structure of problems is given by the set of possible solutions - the solution space. It is this space which is reduced during problem solving as it becomes known which candidates do, or do not qualify any longer as possible solutions.

In chapters 3, 5 and 8 solutions were identified quite explicitly with logical models and the problem space with a formal structure. For grouping (chapter 8) we identified a partition lattice as the solution space, and a partition as a model of a candidate solution.

This adaptation of the use of tableaux in chapter 3 should not be seen as an attempt to describe new, better or faster methods for classification with analytical tableaux. In fact, our use of tableaux has been quite elementary. Analytic tableaux offer a nice example of how semantics and syntax can be combined. Seen as syntactic structures tableaux can be used to build effective theorem provers [30]. Inspecting the branches of tableaux one finds that they correspond to models of the represented theory. Hence, one can interpret our use of analytic tableaux as an attempt to bridge the gap between a symbol-level and knowledge-level description. The symbol-level being the syntactic operations on the tableaux, and the knowledge-level roughly corresponds to the semantic interpretation of the models depicted in the tableau.

In the chapters were we used tableaux, solutions were associated with formal models. This "model-based" approach was made explicit in chapter 6. This approach of starting the analysis of a problem, or task, with the identification of a semantic structure differs from more traditional approaches in KE. Usually one starts with the choice of some, often rich, representational formalism and then represents the domain theory in it. Then problem solving is performed with the help of some accompanying inference engine.

The approach taken here, differs from this tradition in that it is "model-based". By analysing the problem in terms of searching for a model, the semantic structure can be identified. When this has been done an appropriate language can be selected which is not overly expressive and in which the problem can be accurately and intuitively described. Hence, our approach can be summarized by the slogan "semantics first, syntax later".

As a consequence this thesis does not present "yet another representational language". No "universal problem solving language" or a "one-fits-all" representation was chosen or developed. Instead every task was given a formalization which seemed to match it on the basis of a formalization of the ontology and the typical reasoning of a task method.

A dynamic, model-based approach was explicitly formulated in chapter 6. Problem solving is described as acquiring knowledge about solutions, and knowledge in this respect is strongly linked to the reduction of possibilities. We think that a model-based representation of knowledge intensive problems and tasks is a good way of providing a knowledge-level understanding of them.

In summary, we have used the conceptualizations in two ways. First, to construct a static structure aimed at the description of problem variations and a spectrum of solution criteria. The systematic description of problem variations allows one to order and compare different problems. This can be very beneficial for the development of knowledge based systems. This approach is, at least in spirit, similar to the one advocated by ten Teije [88]. In this work parameters for diagnosis were identified which could be varied in the description of several problem variations. In that way various diagnostic problems could be systematically generated.

Second, the conceptualizations were used to construct a structure of possible solutions which was used in the dynamic description of problem solving on the knowledge-level. This has been described in chapter 6 where problem solving behaviour was described in terms of changing states
of knowledge about possible solutions. This dynamic, semantic account of problem solving for knowledge intensive tasks has some important advantages. It allows for the easy extensions to other, non-classical forms of reasoning and it provides a good illustration of our view of knowledge.

Knowledge in our view is a semantic concept and therefore a semantic account of knowledge-level reasoning is called for. Hence, instead of focussing on the development of knowledge based systems, or problem solving methods, we have given an analysis of task-oriented problem solving and representation. We feel that an analysis of knowledge intensive tasks should start with an account of what knowledge is required and how it changes during the problem solving process. The explorations presented in this thesis are means to that end.