Using Description Logics for Managing Medical Terminologies

Ronald Cornet, Ameen Abu-Hanna

Dept. of Medical Informatics, Academic Medical Center, University of Amsterdam
P.O. Box 22700 1100 DE Amsterdam, The Netherlands
{r.cornet, a.abu-hanna}@amc.uva.nl

Abstract. Medical terminological knowledge bases play an increasingly important role in medicine. As their size and complexity are growing, the need arises for a means to verify and maintain the consistency and correctness of their contents. This is important for their management as well as for providing their users with confidence about the validity of their contents. In this paper we describe a method for the detection of modeling errors in a terminological knowledge base. The method uses a Description Logic (DL) for the representation of the medical knowledge and is based on the migration from a frame-based representation to a DL-based one. It is characterized by initially using strong assumptions in concept definitions thereby forcing the detection of concepts and relationships that might comprise a source of inconsistency. We demonstrate the utility of the approach in a real world case study of a terminological knowledge base in the Intensive Care domain and we discuss decisions pertaining to building DL-based representations.

1 Introduction

Medical terminological knowledge bases (TKBs) represent knowledge about medical concepts, relationships and terms. For example, a concept may be defined as “inflammation of the membranes of the brain or spinal cord”, and described by the synonymous terms “cerebrospinal meningitis” and “meningitis”. TKBs provide an invaluable source of structured medical knowledge, serving a range of purposes.

A frame-based representation is commonly used to express definitions of concepts. This formalism supports an intuitive way of knowledge modeling but it lacks explicit semantics, making it hard to automate reasoning. Examples of services expected from the utilization of the TKB include the classification of concepts and consistency checking of the TKB. To perform this automatically, a formal basis is needed for the knowledge representation formalism.

A seemingly attractive formalism to consider is that of Description Logics (DLs), a family of formal languages that are subsets of First Order Logic (FOL) and that provide for an object-oriented like structure of concept definitions.

In this paper we explore a way for deploying DLs for supporting the reasoning services of classification and consistency checking of a medical TKB. Our starting point is that the TKB at hand is specified or implemented in a frame-based language. This is the case in the great majority of TKBs available today. In our approach we migrate the frame-based KB to a DL-based one. Because the frame-based representation is ambiguous, this migration requires making its semantics explicit. We have developed a method to perform this migration by posing explicit assumptions on semantics e.g. of a frame slot. The idea is to start with strong assumptions about definitions in order to force the reasoning system to identify potentially inconsistently defined concepts. This identification is realized by exploiting the satisfiability services of a DL. Each unsatisfiable concept may indicate a too strong assumption but may also indicate errors in the original frame-based definition. Our hypothesis is that going through the migration process and performing satisfiability testing provides a serious contribution for maintaining the contents of medical TKBs. To assess this hypothesis, we have applied our method to a real world knowledge base of Reasons for Admission in Intensive Care, which has been developed in recent years at our department.

This paper is organized as follows. In Section 2 we provide preliminaries on Frame-based representation, Description Logics, and the differences between them. We describe our method in Section 3 and focus on error detection in Section 4. Section 5 reports on the results of this case study. We conclude with observations on application of our method, and on modeling medical terminological knowledge bases.

2 Frame-based and Description Logic-based representations

Frames (Minsky 1981) provide a means of describing classes and instances, with slots of frames representing either relations to other classes, or properties of the represented class. Frames can represent subclasses by means of a KindOf relation, allowing slots (and any slot-fillers) to be inherited from the superclass by the subclass.

As an example of a medical TKB, we will use the DICE knowledge base, which is developed at our department (de Keizer, Abu-Hanna et al. 1999). The DICE system (Diagnoses for Intensive Care Evaluation) represents knowledge in the domain of Intensive Care, with a focus on reasons for admission. Like many medical TKBs, it is organized around health problems, which are defined according to their anatomy, abnormality, etiology, and system (e.g. vascular system, digestive system), as shown in Figure 1. The model is implemented using class frames only.

The model provides the possibility of specifying two special facets of slots, namely transitivity (for example the “part of” slot is transitive), and refinability (for allowing choices of slot-fillers). Figure 2 shows an example of refinability, where the etiology of viral meningitis is indicated by our notation as OR(Virus), meaning that any subclass of virus is accepted here. The application will in that case present to the user the possible values (i.e. all viruses) and request the user to specify one or more viruses that caused the patient’s meningitis.

Description Logics (DLs) (Baader, Calvanese et al. 2003) provide fragments of FOL for formal definition of concepts. These definitions can either be primitive (specifying only necessary conditions), or non-primitive (specifying both necessary and sufficient conditions). For example, consider the following two axioms

$\text{Mother} \sqsubseteq \text{Parent}$; $\text{Mother} \equiv \text{Woman AND Parent}$

The first states that a mother is necessarily a parent, whereas the second states that a

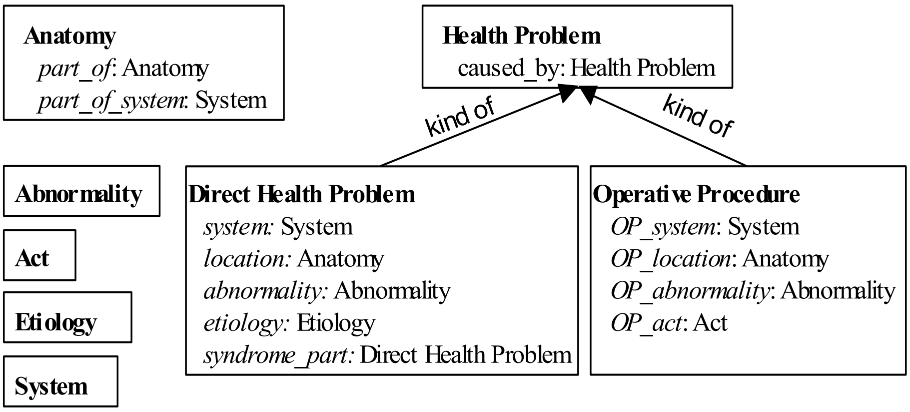

Figure 1: Domain model of the ontology of DICE. Two types of health problems are distinguished, direct health problems and operative procedures. The domains of the slots are represented in Italics. Various examples of subclasses are shown in Figure 2.

mother is necessarily both a woman and a parent, and that anyone who is a woman and a parent is necessarily a mother.

The formal, set-theoretic semantics of DLs provide statements with an unequivocal meaning, which makes reasoning with DL-based knowledge reproducible and application independent. Each DL is characterized by the concept and role constructors it allows for. Examples of concept constructors are AND ($\sqcap$), OR ($\sqcup$), NOT ($\neg$), SOME ($\exists$), ALL ($\forall$), AT-LEAST ($\geq$). For example: Happy Father ≡ Father AND (Rich OR At-Least 3 Children).

Examples of role constructors are transitivity (e.g for the "part of" role: if A part of B and B part of C, then A part of C), inverse roles (e.g. "is_caused_by" is the inverse role of "causes"), or role taxonomies (e.g. "has sister" is a kind of "has sibling" role).

DL-based knowledge bases generally consist of a TBox (Terminology box) containing axioms (such as the above-mentioned examples), and an ABox (Assertion box) containing assertions (e.g. Mary is a Mother; Betty is a child of Mary).

The foremost reasoning tasks with DLs are subsumption (classification) and satisfiability checking. Reasoning is based on the open world assumption, basically meaning that the set of given individuals is not assumed to be complete.

2.1 Differences between Frames and Description Logics

Frames and Description Logics both provide means of representing concepts, relations, and instances. There are however a number of significant differences, which need to be taken into account in the process of migration from frames to DL.

Classes versus concepts. As DL-based reasoning makes it possible to infer subsumption, the resulting taxonomy will be a combination of stated and inferred subsumption (e.g. consider the "Mother" example above). Class frames, in constrast, need to be explicitly defined as subclasses of all applicable superclasses.

Meningitis
Kind of: Brain Disease
Anatomy: Meninges
Abnormality: Infection
Etiology: OR(Virus, Bacterium, Fungus)

Viral Meningitis
Kind of: Meningitis
Etiology: OR(Virus)

Meninges
Kind of: Body Part
Part of: Brains

Microorganism
Kind of: Etiology

Fungus
Kind of: M.-organism

Bacterium
Kind of: M.-organism
Aerobe: XOR(true, false)

Virus
Kind of: M.-organism

Figure 2: Examples of frame-based class definitions. The "Kind of" slot defines direct superclasses. Slot facets "XOR" and "OR" specify whether instances can be defined with exactly one (XOR), or more than one (OR) value from the slot fillers.

Disjointness and covering. As opposed to most Frame-based representations, DLs allow to formally specify that concepts are mutually exclusive (disjoint), by stating that one is subsumed by the complement of the other: Virus $\sqsubseteq$ $\neg$Bacterium.

This axiom renders any concept defined as both a Virus and a Bacterium as unsatisfiable. In addition one could specify that there are no other microorganisms, by: Microorganism $\equiv$ Virus $\sqcup$ Bacterium

Slots versus roles. Without additional constructs, Frame slots and any slot-fillers may be interpreted in various ways. For example, a slot *cause* with slot-filler "(Virus, Bacterium)" may mean that both virus and bacterium are an actual cause, or both are possible causes (possibly combined), either with or without other possible causes, etc.

Description Logics leave no room for such ambiguity. Role quantification is used to express the required meaning. For example (Disease $\sqcap$ $\exists$ cause Virus), uses existential quantification ($\exists$) to denote diseases that have a cause, which is a virus. Universal quantification ($\forall$) is used to limit possible role-values. E.g. (Disease $\sqcap$ $\forall$ cause Virus) denotes diseases of which all causes (if any) are viruses. Combining existential and universal quantification makes it possible to precisely define the semantics of roles.

Slot facets versus role constructors. Semantics of slot facets are often unclear and application-dependent. Examples of such facets are both the refinability and the transitivity facet as described above. In contrast, the semantics of role constructors are explicitly defined, and taken into account by DL reasoners.

3 Migration from Frame-based to DL-based representation

The first step in our method is the translation of a Frame-based representation to a Description Logic-based representation. Because of the loose semantics of frames, assumptions will be made about their semantics. We will focus on disjointness, role quantification and role values, and part-whole reasoning, as these are believed to have the greatest impact on inconsistency detection.

Disjoint definitions. In order to detect as many potential inconsistencies as possible, maximally stringent definitions were assumed, explicitly stating disjointness of siblings. We have defined all concepts subsumed by Act, Abnormality, System and Etiology as mutually disjoint to each of their siblings. In Figure 2 for example, Virus, Fungus, and Bacterium are defined as disjoint. In this way, we can express meningitis caused solely by a virus as:

$\text{ViralMeningitis} \equiv \text{Meningitis} \sqcap \exists\, \text{cause Virus} \sqcap \forall\, \text{cause Virus}$.

An attempt to define viral meningitis caused by a bacterium will result in an unsatisfiable concept, as disjointness of Bacterium and Virus is now explicitly stated.

Role Quantification and Role values. As discussed earlier, semantics of slot-fillers are unclear, allowing multiple interpretations. The assumptions we have posed on the semantics are shown in Table 1, where we present the frame-based representation and its DL-based counterpart, where the slot "cause" and the fillers are taken as examples. In the case of DICE, also the refinability facet of slots needed to be taken into account. Fillers of regular slots are assumed to represent both existentially and universally quantified roles. Fillers of slots with an OR facet (used in DICE to specify zero, one or more of the values when creating an instance) represent only universal quantification. Fillers of slots with an XOR facet (to specify at most one value) are represented as a number restriction (at-most 1) and a universal quantification. As the assumption of universal quantification is too stringent in numerous cases, a special purpose facet has been added to the slots to explicitly specify whether a slot should be considered to represent universal quantification or not. This facet can be updated during the migration process to override the default assumption.

Part-whole relations. Partitive relations play an import role in medical knowledge bases but may demand great expressiveness of Description Logics. This can be overcome by the use of Structure-Entity-Part triplets (SEP), as suggested by (Schulz, Romacker et al. 1998). Motivation for SEP triplets was is the avoidance of the use of transitive roles and role chaining, but comes at the cost of having to define every anatomical component in three ways (as an entity, a part, and a structure). Also for the aim of detecting inconsistencies we found SEP representation to be very useful.

Table 1: Frame-based slot-fillers and their assumed DL-based counterparts

Frame-based representation	**Assumed DL-based equivalent**
cause: (Virus, Bacterium)	$\exists\, \text{cause Virus} \sqcap \exists\, \text{cause Bacterium}$ $\sqcap\ \forall\, \text{cause (Virus} \sqcup \text{Bacterium)}$
cause: OR(Virus, Bacterium)	$\forall\, \text{cause (Virus} \sqcup \text{Bacterium)}$
cause: XOR(Virus, Bacterium)	$\leq 1\ \text{cause} \sqcap \forall\, \text{cause (Virus} \sqcup \text{Bacterium)}$

4 Detecting Errors

In order to detect errors one needs an automatic classifier. A standard Description Logic classifier such as FaCT (Horrocks, Sattler et al. 2000) or RACER (Haarslev and Möller 2000) can be used to find unsatisfiable concepts in the DL-based knowledge base. Unsatisfiability of a concept however does not necessarily imply incorrect definition of the concept. Generally, there can be three explanations for unsatisfiability:

1. The concept itself is correctly defined but refers to an unsatisfiable concept (e.g. it is a child of an unsatisfiable concept)
2. The concept is correctly defined, but the semantics assumed during migration of that concept or any of its subsumers do not represent the intended semantics (e.g. a role is incorrectly assumed to represent universal quantification)
3. The concept is semantically incorrect (e.g. a kind of hepatitis which is defined as located in the kidneys instead of the liver).

In the first situation one unsatisfiable concept can cause a large number of unsatisfiable concepts. As finding such a concept is non-trivial, research is ongoing to develop methods to support this (Schlobach and Cornet 2003). One approach to sort out such situations is to start with concepts that are used as role-values for other concepts. For example, in the case of the Intensive Care knowledge base, subsumers of Anatomy, Act, Etiology, and System are such concepts, hence it is expedient to first address unsatisfiable concepts subsumed by those concepts.

5 Results

We have applied the method described above to the DICE knowledge base, in order to gain insight into the feasibility of this approach. The DICE knowledge base consists of about 2500 concept frames, with over 3000 filled slots (other than "kind of" slots). We used RACER to process the DL-based representation of the knowledge base and check the consistency of the TBox. As mentioned earlier, assumptions posed on the semantics of the frame-based representation may turn out not to be justified. The facet to overrule default interpretation of role quantification made it possible to iteratively migrate from frames to DL, find unsatisfiable concepts, and determine whether the unsatisfiability stemmed from an incorrect assumption or from a modeling error. In either case, the frame-based representation could be changed accordingly, and a new DL-based representation emerged iteratively.

Below we will make a distinction between unsatisfiability introduced by the migration method, and unsatisfiability caused by modeling errors. As the actual migration process is still ongoing, the results are not yet fully quantified. Moreover, the analysis presented here is specific for the DICE knowledge base, and may differ significantly for other TKBs. It does however provide insight in the possibilities of using our method.

5.1 Unsatisfiable concepts caused by the migration method

The stringent assumptions put on the frame-based representation resulted in two types of assumption errors: errors caused by incorrect assumption of disjointness, and errors caused by incorrect assumptions on quantification.

Disjointness errors were found in the descendants of etiology. For example, the (false) assumption was made that “addictive drug” and “analgesic” are disjoint, but “Morphine and Opioids” is (correctly) defined as a descendant of both.

This unsatisfiability could be overcome by removing the assumption of disjointness. It needs to be noted that we have not posed disjointness on the descendant of “health problem”. This is motivated by the fact that the axioms defining them should make it possible to distinguish between them, which is not possible by most of the other concepts, such as descendants of etiology, which lack specification of distinguishing properties.

A large number of unsatisfiable health problems were found, which could be explained by the stringent assumptions posed on the quantification of roles. Universal role quantification was frequently falsely assumed. For many cases, this could be explained by the fact that a frame-based representation requires explicit classification. This led to a large number of grouper concepts, such as “lung disease”, which (falsely) assumed the location to be lungs, and nothing else. This led to unsatisfiability of all diseases that were defined as a “lung disease”, but that also involved a location different from lungs. In these cases, the frame-based representation was altered by tagging the relevant slots as “not universal”.

5.2 Unsatisfiable concepts due to incorrect definitions

Various types of modeling errors were found in the process of migration. We will categorize them as: misclassification, false quantification, missed slot-fillers, and incorrect relations.

Misclassification. A small number of misclassifications have been found, i.e. concepts that were misplaced in the taxonomy. This mainly involved concepts that were placed as siblings where one of the concepts should have been subordinate to the other (be its child). Another example of misclassification is illustrated by a concept that was defined as both a health problem and an abnormality, which are disjoint. Instead of being subsumed by abnormality, it should have been related to abnormality by a slot-filler. The most notable case of misclassification was found in the anatomy taxonomy, where a part of the hierarchy was defined incorrectly by switching subsumers and subsumees. This involved the concept “laryngo tracheo bronchitis” which was defined as the subsumer of laryngitis, tracheitis and bronchitis, whereas it should be defined as a subsumee of these three concepts.

False Quantification. A number of incorrect quantifications were found. These mainly involved concepts for which the OR or XOR facet was not (correctly) specified (see Table1). As this is very specific for the DICE KB, we will not go into further detail. It is however important to realize that correct specification of universal and/or existential quantification is necessary to be able to detect incorrect role values, (or slot-fillers in the frame representation).

Missed slot-fillers. A number of concepts were found that were lacking slot-fillers. Two typical situations were found, of which examples are given below.

Table2: Example of missing slot-filler

Frame-based representation	**Assumed DL-based equivalent**
Brain disease *system* : nervous system	Brain_disease $\sqsubseteq$ $\exists$ system nervous_system
Acromegaly *kind of*: Brain disease *system*: endocrine system	Acromegaly $\sqsubseteq$ brain_disease $\sqcap$ $\exists$ system endocrine_system $\sqcap$ $\forall$ system endocrine_system

Table 2 shows how the slot-filler "endocrine system" overrides the inherited "nervous system", instead of being an additional system. Hence, acromegaly should also involve the nervous system, stating *system*: (endocrine system, nervous system). In the migration process, it turned out that the "system" role in brain disease should not be defined as universal, as brain diseases can involve other systems.

The other typical case was related to OR slots. Frequently, classes had OR slots without slot-fillers that were defined for some of their subclasses. In such cases, these slot-fillers were added to the OR slot of the superclass.

Incorrect slots. The anatomy taxonomy revealed a number of concepts for which a part-of relation was accidentally mixed up with a kind-of relation. This is an error that has been found in other systems as well, and for which DL reasoning provides a powerful means for detecting it (Schulz and Hahn 2001).

5.3 Observations from the case study

During the process of error detection and resolving them, a number of issues came to light that require further investigation. We only have made changes needed to resolve inconsistencies in the original knowledge base. However, studying the definitions indicated that in some cases a more rigorous redefinition would be justified. Also more attention should be paid to the computational properties of the resulting TBox.

Groupers and Patterns. As mentioned earlier, a frame-based representation requires classes to be defined as subclasses of all superclasses involved. As DLs make inference possible on subsumption, a better way of modeling would be to define concepts based on their actual properties, without referring to the grouper concepts. For example, hepatitis would be defined as a disease located in the liver instead of as a "liver disease", as the latter can be inferred from the definition of hepatitis.

Other concepts were found that indicated inconsequent modeling rather than incorrect definition of concepts. For example, both a "part-of" relation and the concepts "body part" and "organ part" are present in the knowledge base. This makes it possible to define a concept by means of either "kind-of organ part" or "part-of organ". Whereas these definitions are logically equivalent, preferably only one of them should be used throughout modeling a knowledge base. Guidelines or modeling patterns might need to be developed to stimulate standardized modeling.

TBox properties. The language that was used for the DL-based representation was $\mathcal{ALCQ}$, which allows the constructors $\sqcap$, $\sqcup$, $\neg$, $\exists$, $\forall$, $\geq$, $\leq$. As we have represented anatomy using SEP triplets, no role hierarchies or transitive roles were required, keeping the language relatively simple.

As the frame-based representation did not contain any axioms other than frame-definitions, and no cycles, the migration resulted in an unfoldable TBox. This means that all definitions are simple (defining only atomic concepts), unique (only one definition for each atomic concept exists), and acyclic (meaning the definition of a concept has no reference to the definiendum, either directly or indirectly). Reasoning on this type of TBox generally has a lower complexity than reasoning on arbitrary TBoxes with cycles and general concept inclusion axioms (Baader, Calvanese et al. 2003).

6 Discussion and Conclusion

We have devised a method for the semi-automated migration from a frame-based representation to a DL-based representation and demonstrated how it helps focusing on weaknesses of a medical terminological knowledge base in Intensive Care. As this knowledge base is modeled in a way comparable to other medical knowledge bases (for example Clinical Terms Version 3 (Read, Sanderson et al. 1995)), it is expected that the methods described here will prove useful in general. There are however a number of remarks to be made.

It is important to realize that although these methods may support detection of incorrect definitions, it cannot be assumed that definitions in a satisfiable knowledge base are correct. For example, if viral meningitis would be defined as hepatitis (instead of meningitis) caused by a virus, this could result in a satisfiable concept, although it is obviously incorrect.

As Description Logics enable automatic subsumption, it can be argued whether or not concepts should be modeled using grouper concepts such as 'liver diseases'. This is in line with the discussions about compiled versus model-based knowledge. In a Frame-based representation, grouper concepts are necessary in order to assure that a disease is considered a liver disease. Using Description Logics, it seems appropriate to define a disease according to its actual properties (e.g. $\exists$ hasLocation liver) and infer the fact that such a disease is a liver disease. Likewise, in a Frame-based representation a concept such as "body part, organ or organ part" would be defined preferably as a disjunction of the constituent concepts in DL.

Application-specific slots or facets, of which the semantics are unclear or non-definitional, cannot be represented using Description Logic. This means that these elements (such as the facets to support post-coordination that allows for the creation of new concepts based on combining existing ones) are lost in the process of migration. Therefore, parts of the functionality provided in the original frame-based representation will have to be realized outside of the DL-based environment. Although this seems to be a drawback at first, it may well turn out to be advantageous as it leads to better understanding of the various aims for which knowledge modeling is being performed.

Admittedly, the DL-representation would include a large number of too strict assumptions. These are mainly concerned with the universal quantification and disjointness. However, the approach provides an automated reasoning tool to identify areas for focusing human attention. Still, a weakness of our approach is that there is no support for tracing or explaining DL-based unsatisfiability. As a consequence, pinpointing and resolving conflicts in definitions is a time-consuming task. Working on explanation facilities comprises important further work that we are planning to address.

References

Baader, F., D. Calvanese, et al. (2003). The Description Logic Handbook: Theory, Implementation, and Applications. Cambridge, University Press.

de Keizer, N. F., A. Abu-Hanna, et al. (1999). "Analysis and design of an ontology for intensive care diagnoses." Methods of Information in Medicine **38**(2): 102-12.

Haarslev, V. and R. Möller (2000). High Performance Reasoning with Very Large Knowledge Bases. International Workshop in Description Logics 2000 (DL2000), Aachen, Germany.

Horrocks, I., U. Sattler, et al. (2000). "Practical reasoning for very expressive description logics." Logic Journal of the IGPL **8**(3): 239-263.

Minsky, M. (1981). A framework for representing knowledge. Mind Design. J. Haugeland, The MIT Press.

Read, J. D., H. F. Sanderson, et al. (1995). "Terming, encoding, and grouping." Medinfo **8 Pt 1**: 56-9.

Schlobach, S. and R. Cornet (2003). Non-Standard Reasoning Services for the Debugging of Description Logic Terminologies. to be published in : International Joint Conference on Artificial Intelligence, Acapulco, Mexico.

Schulz, S. and U. Hahn (2001). "Medical knowledge reengineering--converting major portions of the UMLS into a terminological knowledge base." Int J Med Inf **64**(2-3): 207-21.

Schulz, S., M. Romacker, et al. (1998). "Part-whole reasoning in medical ontologies revisited--introducing SEP triplets into classification-based description logics." Proc AMIA Symp: 830-4.