

Two DL-based Methods for Auditing Medical Terminological Systems

Ronald Cornet MSc, Ameen Abu-Hanna PhD
Dept. of Medical Informatics, Academic Medical Center, University of Amsterdam
Amsterdam, The Netherlands

ABSTRACT

Medical terminological systems (TSs) play an increasingly important role in health care by supporting recording, retrieval and analysis of patient information. As the size and complexity of TSs are growing, the need arises for means to audit them, i.e. verify and maintain (logical) consistency and (semantic) correctness of their contents. In this paper we describe two methods based on description logics (DLs) for the audit of TSs. One method uses non-primitive definitions to detect concepts with equivalent definitions. The other method is characterized by stringent assumptions that are made about concept definitions, in order to detect inconsistent definitions. We discuss the possibility of applying these methods to the Foundational Model of Anatomy (FMA) to demonstrate the potentials and pitfalls of these methods. We show that the methods are complementary, and can indeed improve the contents of medical TSs.

INTRODUCTION

During the last decade, the department of Medical Informatics at the University of Amsterdam has been carrying out research and development on medical terminological systems (TSs) and services. As modeling knowledge in very large TSs and evaluation of their contents are complicated processes, the need arises for systematic, reproducible methods to support these processes. Modeling and evaluating TSs concern various aspects, ranging from ontological decisions to the comprehensiveness of the medical contents of a TS. Ideally, a TS should satisfy four requirements: (1) it should have the necessary knowledge (completeness), (2) the knowledge should be faithful to the real world (correctness), (3) the knowledge should not be self-contradictory (consistency), and (4) the system should have efficient algorithms to perform the inferences needed for the application (competence). Auditing is the process of assessing the fulfillment of these requirements. A number of approaches have been designed, and applied in the field of medicine, for example in¹⁻⁴. The application of a description logics (DL) as representation formalism for medical TSs is getting increasing attention. The most prominent examples of DL-based medical TSs are GALEN⁵ and SNOMED-CT⁶. It seems however, that the merits of DL-based

inference are still not fully understood. This paper presents two methods that aim at applying DL-based inference for auditing medical TSs in order to better understand the potential of DL-based methods.

METHODS

The two methods are based on making a DL-based interpretation of a frame-based system. With minor modification, they can also be applied to audit systems based on an inexpressive DL, but we will not discuss this. One method uses non-primitive definitions to detect concepts with equivalent definitions. This method aims at detecting concepts that have duplicate definitions, and concepts that are under-defined. The other method is characterized by a process in which stringent assumptions are made about concept definitions. This method aims at detecting inconsistencies. The examples given throughout this paper are largely extracted from the Foundational Model of Anatomy (FMA)⁷, a frame-based ontology for anatomical knowledge, containing about 70,000 distinct anatomical concepts. We will first discuss how both representations are generated, and then look at results of DL-inference for concept classification.

Representation for Detection of Equivalent Definitions

The first method aims at detection of concepts with equivalent definitions, which indicate either concepts with duplicate definitions, or under-defined concepts. For example, in the FMA the concepts

Figure 1: Frame-based representation of FMA concepts

Frame: Paraaortic body
 Superframe: Homogeneous organ
 Systemic part of: Endocrine system
Frame: Paraganglion
 Superframe: Homogeneous organ
 Systemic part of: Endocrine system
Frame: Coccygeal body
 Superframe: Homogeneous organ
Frame: Carotid body
 Superframe: Homogeneous organ
Frame: Parathyroid gland
 Superframe: Homogeneous organ
 Arterial supply: Superior thyroid artery,
 Inferior thyroid artery
...

“Paraganglion” and “Paraortic body” are defined in exactly the same way, as shown in Figure 1. These are actually different concepts, but the distinction between them is not represented in FMA. Although it is impossible to represent every characteristic for many concepts⁸, studying concepts with equivalent definitions can help bringing about better distinctions between definitions.

Our method to detect equivalent definitions comprises the following procedure, in which frames are expressed as DL statements.

All frames that contain a reference to exactly one superframe and have no specified slot-fillers, are represented as primitive concept definitions, $B \sqsubseteq A$. All other frames are represented as non-primitive.

Specified superframes and slot-fillers are represented as a logical conjunction, where slot-fillers are interpreted as existential quantifications.

The results of this interpretation are shown in Figure 2a. Primitive concepts are easily recognizable by their definition, and they form the first point of interest for further study. To detect concepts with equivalent definitions, the resulting model is classified using a DL reasoner, as described below.

Representation for Detection of Inconsistencies

This process of DL-based representation of a frame-based system is based on a number of assumptions and modeling decisions.

Frames can be defined using necessary properties, necessary and sufficient properties, or prototypical properties, but in general this is ambiguous. We will assume that definitions contain only necessary properties.

The other assumptions are guided by the aim of the

process: semi-automatic detection of inconsistent concept definitions. In order to be able to detect as many potential inconsistencies as possible, maximally stringent definitions are assumed. These stringent definitions are aimed at restraining the open world assumption, for example by explicitly stating disjointness of siblings, and universal as well as existential quantification. Without such stringent assumptions, no inconsistencies can be detected. For example, to detect inconsistency in role values, disjointness must be made explicit, and role values must be both universally and existentially quantified. Six basic assumptions are made, which are mentioned and discussed below. Two additional assumptions related to representation of anatomy are separately and more extensively discussed.

1. *All concepts are defined as primitive.*
This is based on the assumption that frames are defined using necessary properties. Besides, to infer inconsistency, non-primitive definitions have no additional value.
2. *Multiple superframes are interpreted as conjunction of multiple superclasses.*
This is again based on the assumption that all properties are necessary properties.
3. *All sibling frames are interpreted as mutually disjoint concepts.*
This is based on the assumption that siblings specialize the superframe by further specification of one and the same property.
4. *Slots for which fillers are specified override slot-fillers defined by superclasses.*
This implies that properties are interpreted as refined specification, not as an additional

Figure 2: Results of to methods of interpreting frame-based statements

a) DL-based representation for detection of equivalence:

Paraortic_body \equiv Homogeneous_organ \sqcap \exists syst_part_of Endocrine_system
 Paraganglion \equiv Homogeneous_organ \sqcap \exists syst_part_of Endocrine_system
 Coccygeal_body \sqsubseteq Homogeneous_organ
 Carotid_body \sqsubseteq Homogeneous_organ
 Parathyroid_gland \equiv Homogeneous_organ \sqcap
 \exists arterial_supply Superior_thyroid_artery \sqcap \exists arterial_supply Inferior_thyroid_artery \sqcap ...

b) DL-based representation for detection of inconsistency:

Paraortic_body \sqsubseteq Homogeneous_organ \sqcap \exists syst_part_of Endocrine_system \sqcap \forall syst_part_of Endocrine_system
 Paraganglion \sqsubseteq Homogeneous_organ \sqcap \exists syst_part_of Endocrine_system \sqcap \forall syst_part_of Endocrine_system
 Coccygeal_body \sqsubseteq Homogeneous_organ
 Carotid_body \sqsubseteq Homogeneous_organ
 Parathyroid_gland \sqsubseteq Homogeneous_organ \sqcap
 \exists arterial_supply Superior_thyroid_artery \sqcap \exists arterial_supply Inferior_thyroid_artery \sqcap
 \forall arterial_supply (Superior_thyroid_artery \sqcup Inferior_thyroid_artery) \sqcap ...
 disjoint Parathyroid_gland Paraortic_body Paraganglion Coccygeal_body Carotid_body

property.

5. *Superframes and slots for which fillers have been specified are conjuncted.*

This is in accordance with the default interpretation of frame-based descriptions.

6. *Slot-fillers are interpreted as conjunctions of existential quantifications of the role values and universal quantification of the disjunction of the role values.*

This is shown in the definition of Parathyroid gland in Figure 2b. Both slot-fillers for “Arterial supply”, Superior thyroid artery and Inferior thyroid artery, are specified in an existential quantification, and by means of universal quantification they are defined as the only allowed role values.

Representing anatomy using SEP triplets

Two additional assumptions involve the representation of anatomical knowledge in terminological systems.

7. *The Anatomy taxonomy consists of anatomical structures, which are represented using Structure-Entity-Part (SEP) triplets⁹.*
8. *The Anatomy taxonomy is treated as a partition, i.e. parts are considered to be disjoint.*

In accordance to the “rules for part-whole relationships”¹⁰ we make the assumption that parts are not overlapping within one context, and that each context (a particular viewpoint) requires a different parthood relation. For example, right side and left side of the heart are functional or clinical partitions, whereas a subdivision into walls and cavities is an anatomical partition.

In Figure 3 an example is shown of the chest and the heart, both being part of the thorax, and both having a left side and a right side. The introduction of universal quantifications, combined with disjointness and non-overlap complicates parthood modeling.

Figure 3: DL-based representations of parthood, with and without use of SEP triplets

a) Without SEP triplets

Heart $\sqsubseteq \exists$ part of_D Thorax $\sqcap \forall$ part of_D Thorax
 Left side of heart $\sqsubseteq \exists$ part of_D Heart $\sqcap \forall$ part of_D Heart
 Right side of heart $\sqsubseteq \exists$ part of_D Heart $\sqcap \forall$ part of_D Heart
 Chest $\sqsubseteq \exists$ part of_D Thorax $\sqcap \forall$ part of_D Thorax
 Left side of chest $\sqsubseteq \exists$ part of_D Chest $\sqcap \forall$ part of_D Chest
 Right side of chest $\sqsubseteq \exists$ part of_D Chest $\sqcap \forall$ part of_D Chest
 disjoint Heart Chest
 disjoint (\exists part of Heart) (\exists part of Chest)
 $\langle \text{anat} \rangle_S \equiv \langle \text{anat} \rangle_E \sqcup \langle \text{anat} \rangle_P$
 $\langle \text{anat} \rangle_P \equiv \exists$ part of $\langle \text{anat} \rangle_E$
 disjoint $\langle \text{anat} \rangle_E \langle \text{anat} \rangle_P$
 Heart disease \sqsubseteq disease $\sqcap \exists$ anatomy (Heart $\sqcup \exists$ part of Heart)
 $\sqcap \forall$ anatomy (Heart $\sqcup \exists$ part of Heart)

The modeling solution described in Figure 3a requires distinction of an intransitive “direct” part of role (denoted as part of_D), which is subsumed by a transitive “part of” role. The definition of Heart disease requires use of a construct “Heart $\sqcup \exists$ part of Heart”. This is actually an “anonymous” form of the Heart_S (Heart structure) concept of the SEP triplet, which subsumes Heart_E (Heart entity) and Heart_P (Heart part). Other ways of modeling (e.g. defining “part of” to be subsumed by the “anatomy” role) have been assessed as well, but they all demonstrate these anonymous SEP triplets. We have therefore used the representation using SEP-triplets, as shown in Figure 3b.

PROCESSING THE MODELS

Based on the two methods, two DL-based representations of a terminological system can be generated. The axioms can be represented in KRSS syntax¹¹ or OWL¹² and the model can be classified using a DL reasoner, for example RACER¹³.

Detection of Equivalent Definitions

Classification of the model results in sets of concepts with equivalent definitions. Sets can be evaluated with regard to the lexical similarity of the terms that denote the concepts. For example, a pair, found in FMA, “Left subcostal muscle”, “Right subcostal muscle” indicates that laterality is not specified in the definition. A pair such as “Paraganglion”, “Paraortic body” indicates a more intricate distinction between the concept definitions, which might not be possible to represent. Larger sets may point out concepts that lack specification of various characteristics, as well as multiple siblings that are equivalent to their subsuming class. For example, analysis of the triple “Nerve to right subclavius”, “Nerve to left subclavius”, “Nerve to subclavius” reveals that the definitions of the subsumed concepts (Nerve to left/right subclavius) do not specify nerve supply of

b) With use of SEP triplets

Heart_S \sqsubseteq Thorax_P
 Left side of heart_S \sqsubseteq Heart_P
 Right side of heart_S \sqsubseteq Heart_P
 Chest_S \sqsubseteq Thorax_P
 Left side of chest_S \sqsubseteq Chest_P
 Right side of chest_S \sqsubseteq Chest_P
 disjoint Heart_S Chest_S
 - (inferred from disjointness of structures)

left/right subclavius and are logically equivalent to that of the subsuming concept (Nerve to subclavius, which specifies “nerve supply of” subclavius). In this case, the subsumed concepts redundantly specify “has physical state: Solid”, which is also part of the specification of “Nerve to subclavius”.

In summary, 3 major types of equivalent definitions can be detected:

1. concepts without any specification (which are defined as primitive)
2. equivalent siblings, which lack specification of their distinguishing characteristics
3. concepts equivalent to their stated subsumer, which contain redundant specification of characteristics.

Detection of Inconsistencies

Classification of the model generated for detection of inconsistencies results in a number of unsatisfiable concepts, i.e. concepts that have an inconsistent definition. As the definition is based on an interpretation of the original definition, one needs to determine whether the interpretation or the original definition is incorrect. However, pinpointing the cause of inconsistency is not straightforward¹⁴. If a concept is rendered unsatisfiable, all concepts that it subsumes as well as all concepts that refer to it will become unsatisfiable. Moreover, the characteristic(s) that lead to unsatisfiability are not readily available, but as yet need to be determined by hand. Hence, the number of unsatisfiable concepts is no indication for the number of actual inconsistencies.

AUDITING IN PRACTICE

We have applied the methods described above on the FMA in order to assess the usefulness of these methods for a real-world TS and perform a provisional audit of the FMA.

The local installation of the FMA was used, which can be accessed through Protégé¹⁵. The model is processed using the Protégé API. Those slots that are not part of the concept definition were ignored, i.e. all Protégé system slots and the slots “UWDAID” and “definition”, which specify respectively a unique identifier for a concept and a free-text definition. The resulting two models, for equivalence and consistency detection, were audited separately.

Starting from the top of the hierarchy, all subframes were recursively represented using DL according to both methods described above. Disjointness of siblings could be stated because multiple inheritance, though allowed in FMA, was not encountered. The axioms were represented in KRSS syntax and processed using RACER.

Due to the complexity of the models, it was not possible to classify the DL-based representations of

the FMA as a whole, which contained 68781 concept definitions. The complexity of the model is the result of the large number of concepts and the numerous cyclic definitions, which are caused by the use of relations and their inverses, e.g. “branch” and “branch of”; “part” and “part of”.

Therefore, we have restricted the audit to the “Organ” taxonomy of FMA, which contained 3826 definitions, comprising about 5% of the FMA.

Equivalent Definitions

All of the FMA contained 35425 (=52%) primitive and 33356 (=48%) non-primitive definitions.

The “Organ” taxonomy had 1167 (=31%) primitive and 2659 (=69%) non-primitive definitions.

The model has been classified using RACER, and the output of RACER was processed using simple scripting and text manipulation tools (sed, awk and grep). Classification resulted in 494 concepts having non-unique definitions. There were 157 sets of concepts with equivalent definitions, ranging in size from 2 concepts (106 sets) to 54 concepts (1 set).

28 sets contained concepts that referred to laterality (e.g. Left phrenic nerve, Right phrenic nerve), without explicit reference to laterality in the definition. In general, many of the equivalent concepts contained positional information, e.g. distal/middle/proximal, or posterior/anterior.

109 definitions were found for concepts that were equivalent to their stated subsumer, hence contained redundant specification of characteristics.

Inconsistencies

The “Organ” taxonomy created according to the process described above could not be classified using RACER-1.7.24 (on a 2.4 GHz 1 GB Pentium 4), probably due to the presence of definitions using relations and their inverses, in combination with the use of SEP triplets. Leaving out the SEP triplets rendered the model classifiable, and 307 inconsistent concepts were found. 230 inconsistencies originate from two characteristics of “Organ”, respectively “regional part of” Organ system, and “part of” Organ system. In many cases, fillers of these slots are not an organ system. For example Periodontium has characteristic: part of Tooth. Tooth is an organ, not an organ system, rendering Periodontium inconsistent. Manual review also revealed various concepts (e.g. Coccyx) that were specified as part of both a Male and a Female body part (e.g. Male pelvis and Female pelvis). Preferably, such concepts refer to a gender-neutral body part (e.g. Pelvis).

DISCUSSION AND CONCLUSIONS

A major advantage of the methods described in this paper is that they use readily available reasoning

capabilities of DL reasoners. This makes it possible to find concepts with logically equivalent or inconsistent definitions, with relatively little effort. One drawback of the method used is the lack of support for processing the results of the classification, e.g. lexical methods², and methods to pinpoint sources of inconsistencies¹⁴. Research is ongoing to support this.

Another drawback is the fact that it is not possible to classify a large TS, such as FMA, as a whole. This can be resolved by partitioning a TS and applying the methods to all the resulting parts of a TS. To what extent this complicates the methods or influences the outcomes is yet to be determined.

It must be stressed that the models resulting from our methods are useful for auditing purposes, but the underlying assumptions by which they are generated may not be in correspondence with the actual semantics, hence these models are by no means a replacement for the original TS.

As demonstrated for the FMA, the methods described provide guidance in finding concepts for which the definition can be enhanced, and concepts for which the definition should be revised. In this way, they contribute to the auditing of terminological systems.

Acknowledgements

This work has been partially funded by the NICE foundation and the Netherlands Organization for Scientific Research (NWO) program "Information & Communication Technology in Healthcare" (ICZ) for the project entitled "Terminology and Semantics: Making semantics explicit", number 014-18-014.

References

1. Ceusters W, Smith B, Kumar A, Dhaen C. Ontology-based Error Detection in SNOMED-CT(R). In: Proceedings from Medinfo 2004 p. 482-6.
2. Cimino JJ. Auditing the Unified Medical Language System with semantic methods. *J Am Med Inform Assoc* 1998;5(1):41-51.
3. Bodenreider O, Smith B, Kumar A, Burgun A. Investigating Subsumption in DL-Based Terminologies: A Case Study in SNOMED CT. In: KR 2004 Workshop on Formal Biomedical Knowledge Representation (KR-MED 2004) p. 12-20.
4. Pisanelli DM, Gangemi A, Battaglia M, Catenacci C. Coping with medical polysemy in the semantic web: the role of ontologies. In: Proceedings from Medinfo 2004 p. 416-9.
5. Rector AL, Solomon WD, Nowlan WA, Rush TW, Zanstra PE, Claassen WM. A Terminology Server for medical language and medical

information systems. *Methods Inf Med* 1995;34(1-2):147-57.

6. Spackman KA. SNOMED CT milestones: endorsements are added to already-impressive standards credentials. *Healthc Inform* 2004;21(9):54, 56.
7. Rosse C, Mejino J, Jose L. V. A reference ontology for biomedical informatics: the Foundational Model of Anatomy. *J Biomed Inform* 2003;36(6):478-500.
8. Doyle J, Patil R. Two Theses of Knowledge Representation: Language Restrictions, Taxonomic Classifications, and the Utility of Representation Services. *Artif Intell* 1991;48(3):261--298.
9. Schulz S, Romacker M, Hahn U. Part-whole reasoning in medical ontologies revisited--introducing SEP triplets into classification-based description logics. In: Proceedings of the 1998 AMIA Annual Fall Symposium p. 830-4.
10. Mejino JV, Jr., Agoncillo AV, Rickard KL, Rosse C. Representing complexity in part-whole relationships within the Foundational Model of Anatomy. In: Proceedings of the 2003 AMIA Annual Symposium p. 450-4.
11. Patel-Schneider P, Swartout B. Description-Logic Knowledge Representation System Specification from the KRSS Group of the ARPA Knowledge Sharing Effort: KRSS Group of the ARPA Knowledge Sharing Effort; 1993 1 november 1993.
12. OWL website, <http://www.w3.org/2004/OWL/>, Last Accessed: 2005, March 11th
13. Haarslev V, Möller R. RACER System Description. In: Proceedings of the International Joint Conference on Automated Reasoning p. 701-706.
14. Schlobach S, Cornet R. Non-Standard Reasoning Services for the Debugging of Description Logic Terminologies. In: International Joint Conference on Artificial Intelligence p. 355-360.
15. Gennari JH, Musen MA, Fergerson RW, Grosso WE, Crubézy M, Eriksson H, et al. The Evolution of Protégé: An Environment for Knowledge-Based Systems Development. *Int J Hum-Comput Stud* 2003;58(1):89-123.

Address for correspondence

Ronald Cornet (r.cornet@amc.uva.nl)
Dept. of Medical Informatics, J1B-114
Academic Medical Center
P.O. Box 22700
1100 DE Amsterdam, The Netherlands