Proceedings of 2004 the Pacific Symposium on Biocomputing. World Scientific; 2003. p. 164-165.

INVESTIGATING IMPLICIT KNOWLEDGE IN ONTOLOGIES WITH APPLICATION TO THE ANATOMICAL DOMAIN

S. ZHANG & O. BODENREIDER

U.S. National Library of Medicine
8600 Rockville Pike, MS 43, Bethesda, Maryland 20894, USA
E-mail: {szhang, olivier}@nlm.nih.gov

Knowledge in biomedical ontologies can be explicitly represented (often by means of semantic relations), but may also be implicit, i.e., embedded in the concept names and inferable from various combinations of semantic relations. This paper investigates implicit knowledge in two ontologies of anatomy: the Foundational Model of Anatomy and GALEN. The methods consist of extracting the knowledge explicitly represented, acquiring the implicit knowledge through augmentation and inference techniques, and identifying the origin of each semantic relation. The number of relations (12 million in FMA and 4.6 million in GALEN), broken down by source, is presented. Major findings include: each technique provides specific relations; and many relations can be generated by more than one technique. The application of these findings to ontology auditing, validation, and maintenance is discussed, as well as the application to ontology integration.

1 Introduction

Biomedical ontologies can be developed manually, semi-automatically or automatically, with the support of knowledge acquisition tools, or by knowledge servers reasoning on formal knowledge representation languages [1]. The resulting ontologies generally consist of concepts modeled by hierarchical relationships. Concepts are identified by names or formal definitions, and described by properties and associative relationships with other concepts. The inter-concept relationships, either hierarchical or associative, direct or indirect, constitute the explicit knowledge represented in the ontologies.

Ontologies may also contain knowledge less explicitly represented. The notion of implicit knowledge has been explored in various contexts in AI-related areas including expert systems, knowledge acquisition, and knowledge representation and reasoning [2, 3]. Explicit knowledge generally refers to what is represented through formal models or procedures. Implicit knowledge, on the other hand, is defined differently and may include human experiences, informal representations such as images and visions, and formal implications deduced from the explicit knowledge. In this paper, we investigate the implicit knowledge embedded in the concept names and inferable from various combinations of semantic relations.

In a previous study [4], we proposed several techniques for acquiring implicit knowledge in biomedical ontologies. Our motivation was to facilitate ontology integration by making different ontologies more directly compatible. Additionally, we showed that acquiring implicit knowledge can help reveal latent inconsistencies within ontologies, as well as conflicts between representations of the same domain.

Knowledge may not always need to be represented explicitly. For example, in description logic-based systems [5], reasoners and classifiers rely on metaknowledge expressed through axioms to generate additional knowledge from the explicit representation. Such systems would generally perform similarly to our techniques for inferring new knowledge. However, these systems do not usually take advantage of the knowledge implicitly embedded in concept names as we do.

The contribution of this paper is to study the respective proportions of explicit and implicit knowledge in biomedical ontologies and the relative contribution of various techniques to making implicit knowledge explicit. We show later on that formally representing the origin of the relations is of interest as it may contribute to maintaining consistency in ontologies, to auditing and validating ontologies, and would, more generally, benefit tasks such as ontology merging [6, 7] and alignment [8], and agent communication in the Semantic Web [9, 10].

2 Materials

Our domain of interest for this study is anatomy. We selected two comprehensive ontologies representing anatomical knowledge: the Foundational Model of Anatomy[1] (FMA) [March 4, 2003 version] and the GALEN[2] common reference model [v. 6].

The Foundational Model of Anatomy (FMA) is an evolving ontology that has been under development at the University of Washington since 1994 [11, 12]. Its objective is to conceptualize the physical objects and spaces that constitute the human body. The underlying data model for FMA is a frame-based structure implemented with Protégé-2000. With 66,879 concepts, FMA claims to cover the entire range of gross, canonical anatomy.

The Generalized Architecture for Languages, Encyclopedias and Nomenclatures in medicine (GALEN) has been developed as a European Union AIM project led by the University of Manchester since 1991 [13, 14]. The GALEN common reference model is a clinical terminology represented using GRAIL, a formal language based on description logics. GALEN contains 52,006 concepts and intends to represent the biomedical domain, of which canonical anatomy is only one part.

Both FMA and GALEN are modeled by *IS-A* and *PART-OF* relationships and allow multiple inheritance. Relationships in GALEN are finer-grained than in FMA. For the purpose of this study, we considered as only one *PART-OF* relationship the various kinds of partitive relationships present in FMA (e.g., *part of*, *gen-*

[1] http://sig.biostr.washington.edu/projects/fm/AboutFM.html

[2] http://www.opengalen.org/

eral part of) and in GALEN (e.g., *isStructuralComponentOf*, *isDivisionOf*). *IS-A* and *PART-OF* have inverse relationships, *INVERSE-IS-A* and *HAS-PART*. In canonical anatomy, the inverse relations are essentially always valid, although this may not necessarily be the case in the real world [15].

3 Methods

3.1 Acquiring explicit knowledge

Inter-concept relationships are generally represented by semantic relations <*$concept_1$, relationship, $concept_2$*>, where *relationship* links *$concept_1$* to *$concept_2$*. In this study, we limited our investigation to hierarchical relationships, i.e., *IS-A*, *INVERSE-IS-A*, *PART-OF*, and *HAS-PART*. Acquiring explicit knowledge simply consisted of extracting the semantic relations explicitly represented. In addition, we refined these explicit relations by a series of complementing and cleaning actions. First, in order to make the relations more easily comparable across systems, we added to each ontology the missing inverse relations[3]. Additionally, and only for FMA, we assigned to a more generic concept the *PART-OF* relationships common to all its leaf descendants[4]. Finally, we identified and removed a small number of hierarchical cycles within each ontology. The knowledge resulting from these actions is still considered explicit, either because the tasks are relatively trivial or because this knowledge was expected to be represented in the first place. The resulting relations are called the base semantic relations, to which implicit knowledge will be compared.

3.2 Acquiring implicit knowledge

Augmentation and inference were two main techniques used to acquire implicit knowledge from FMA and GALEN [4]. **Augmentation** attempts to represent with relations knowledge that is otherwise embedded in the concept names through reification or other linguistic phenomena such as nominal modification and prepositional attachment. Augmentation based on reified *PART-OF* relationships consists of creating a relation <*P*, *PART-OF*, *W* > between concepts *P* (the part) and *W* (the whole) from a relation <*P*, *IS-A*, *Part of W*>, where the concept *Part of W* reifies, i.e., embeds in its name, the *PART-OF* relationships to *W*. For example, <*Neck of Femur*, *PART-OF*, *Joint*> was added from the relation <*Neck of Femur*, *IS-A*, *Component of*

[3] For example, <*Hand, HAS-PART, Index finger*>, was added to GALEN, complementing <*Index finger, PART-OF, Hand*>, explicitly represented.

[4] For example, <*Lung, PART-OF, Intrathoracic part of chest*> was added to FMA because all leaf descendants of *Lung*, i.e., *Left lung* and *Right lung*, are in the *PART-OF* relationship with *Intrathoracic part of chest*. Such *PART-OF* relationships should have been assigned to more generic concepts and inherited downwards in the ontology modeling stage.

Joint>, where the concept *Component of Joint* reifies a specialized *PART-OF* relationship. Examples of augmentation based on nominal modification and prepositional attachment include <*Thyroid gland, IS-A, Gland*> (from the concept name *Thyroid gland*) and <*Leaflet of pulmonary valve, PART-OF, Pulmonary valve*> (from the concept name *Leaflet of pulmonary valve*).

Inference generates additional semantic relations by applying inference rules to the existing relations. These inference rules, specific to this study, represent limited reasoning along the *PART-OF* hierarchy, generating a partitive relation between a specialized part and the whole or between a part and a more generic whole. For example, <*Hand, PART-OF, Free limb*> was inferred based on the explicit relations <*Hand, PART-OF, Free upper limb*> and <*Free upper limb, IS-A, Free limb*>.

3.3 Identifying the origin of semantic relations

Semantic relations may be acquired by several methods. They can be explicitly represented, added by complementation, as well as generated by augmentation and by inference. The former two categories constitute explicit knowledge (i.e., the base semantic relations in this study) and the latter two implicit knowledge. In other words, each method produces a set of semantic relations. Augmentation relies solely on concept names and only one set of augmented relations obtains. In contrast, inference can be applied to the base relations only, to the augmented relations only, or to both, resulting in three distinguishable sets of inferred relations. The five sets of semantic relations studied are: *B* (base semantic relations), *A* (augmented semantic relations), I_B (inferred semantic relations based on the base relations alone), I_A (inferred semantic relations based on the augmented relations alone), and $I_{B \cup A}$ (inferred semantic relations based on the base and augmented relations).

Depending on which method (or methods) can generate it, each semantic relation belongs to at least one and at most five sets *B*, *A*, I_B, I_A, and $I_{B \cup A}$. When a relation can be generated by several methods, it is therefore common to the corresponding sets of relations and, thus, belongs to the intersection of these sets. We use the intersection of sets as a unique identifier for the origin of a relation, hereafter referred to as its source. For example, the source $(B \cap A \cap I_{B \cup A} \cap I_A)$ identifies the relations common to the sets *B*, *A*, $I_{B \cup A}$, and I_A, but absent from I_B. More concretely, the semantic relation <*Anterior lobe of prostate, PART-OF, Prostate*> in FMA belongs to the intersection $(B \cap A \cap I_{B \cup A} \cap I_A)$ because the relation: is explicitly represented in FMA (i.e., in *B*); can be augmented from the name of the concept *Anterior lobe of prostate* (i.e., in *A*); can be inferred from two augmented relations <*Anterior lobe of prostate, IS-A, Lobe of prostate*> and <*Lobe of prostate, PART-OF, Prostate*> (i.e., in I_A); can be inferred from a combination of base relation <*Anterior lobe of prostate,* IS-A, *Lobe of prostate*> and augmented relation <*Lobe of prostate,* PART-OF, *Prostate*> (i.e., in $I_{B\cup A}$); and cannot be inferred solely from base relations using our inference rules (i.e., **not** in I_B).

4 Results

4.1 Number of semantic relations acquired

The number of semantic relations acquired from FMA and GALEN is presented in Table 1. The base semantic relations include the relations explicitly represented and those added by complementation, as described earlier. The implicit relations are generated by augmentation and inference. Because semantic relations may be acquired by several methods, the total number of unique semantic relations is slightly less than the sum of the number of relations in the four subcategories listed.

Semantic relations		**FMA**	**GALEN**
Base semantic relations	Explicit	342,238	228,522
	Complemented	305,194	23,268
Implicit semantic relations	Augmented	392,314	32,922
	Inferred	11,896,508	4,356,244
Total (unique semantic relations)		12,388,812	4,584,504

Table 1. Number of semantic relations acquired from FMA and GALEN

4.2 Origin of the semantic relations acquired

From the perspective of the semantic relations, the source of a relation represents the method (or methods) by which this relation can be generated. From the five individual methods studied in this paper (B, A, I_B, I_A, and $I_{B\cup A}$), nineteen sources in FMA and sixteen in GALEN were found to partition the total set of relations into disjoint subsets. To each subset corresponds a combination of methods by which the relations in the subset can be generated. As shown in Figure 1, four sources contribute the vast majority of relations in both FMA (about 95%) and GALEN (nearly 99%). These sources are: $(I_{B\cup A} \cap I_B)$, $(I_{B\cup A})$, (B), and $(B \cap I_{B\cup A} \cap I_B)$. The number and percentage of relations coming from each source for FMA and GALEN are presented in Table 2.

For example, 105,084 relations in FMA can be generated by both A (augmentation) and $I_{B\cup A}$ (inference based on the base and augmented relations), but not by the other three methods. As shown in the table next to the label $(A \cap I_{B\cup A})$, these 105,084 relations are represented by two gray slots in column A and $I_{B\cup A}$ and white slots in the other three columns. Note that row *(A)* represents the relations that can only be generated by augmentation, while a gray slot in column *A* identifies the relations that may be generated by augmentation.

Source of the semantic relations						FMA		GALEN	
	B	A	I_B	I_A	$I_{B\cup A}$	Number	%	Number	%
(B)	■					355,550	2.8699	217,816	4.7511
(A)		■				96,194	0.7765	4,286	0.0935
$(I_{B\cup A})$					■	4,158,676	33.5680	197,608	4.3103
$(I_{B\cup A} \cap I_B)$			■		■	7,052,658	56.9276	4,082,178	89.0430
$(I_{B\cup A} \cap I_A)$				■	■	157,252	1.2693	9,366	0.2043
$(B \cap A)$	■	■				40,560	0.3274	6,158	0.1343
$(B \cap I_{B\cup A})$	■				■	75,218	0.6071	262	0.0057
$(A \cap I_{B\cup A})$		■			■	105,084	0.8482	148	0.0032
$(B \cap A \cap I_{B\cup A})$	■	■			■	1,048	0.0085	0	
$(B \cap I_{B\cup A} \cap I_B)$	■		■		■	170,330	1.3749	22,148	0.4831
$(B \cap I_{B\cup A} \cap I_A)$	■			■	■	1,534	0.0124	0	
$(A \cap I_{B\cup A} \cap I_B)$		■	■		■	27,716	0.2237	15,402	0.3360
$(A \cap I_{B\cup A} \cap I_A)$		■		■	■	82,362	0.6648	136	0.0030
$(I_{B\cup A} \cap I_B \cap I_A)$			■	■	■	24,122	0.1947	17,388	0.3793
$(B \cap A \cap I_{B\cup A} \cap I_B)$	■	■	■		■	1,334	0.0108	466	0.0102
$(B \cap A \cap I_{B\cup A} \cap I_A)$	■	■		■	■	234	0.0019	0	
$(B \cap I_{B\cup A} \cap I_B \cap I_A)$	■		■	■	■	1,158	0.0093	4,816	0.1050
$(A \cap I_{B\cup A} \cap I_B \cap I_A)$		■	■	■	■	37,316	0.3012	6,202	0.1353
$(B \cap A \cap I_{B\cup A} \cap I_B \cap I_A)$	■	■	■	■	■	466	0.0038	124	0.0027
Total						12,388,812	100	4,584,504	100

Table 2. Source of the semantic relations acquired from FMA and GALEN

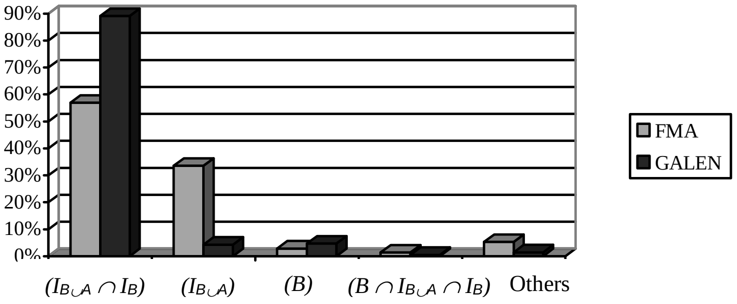

Figure 1. Contribution of the top four sources of relations in FMA and GALEN

4.3 *Base semantic relations*

The base semantic relations come from all sources involving *B*, i.e., not only the row *(B)* in Table 2, but all ten rows marked in grey in column *B*, including, for example, $(B \cap I_{B \cup A})$. While some of these relations are only present in the base, some of them may also be augmentable, be inferable, or both. The proportion of base relations for each of these categories in FMA and GALEN is shown in Table 3.

Base semantic relations	**FMA** (N=647,432)	**GALEN** (N=251,790)
Only present in the base	54.92 %	86.51 %
Also augmentable	6.74 %	2.68 %
Also inferable	38.83 %	11.05 %
(Both augmentable and inferable)	0.48 %	0.24 %

Table 3. The base semantic relations

4.4 *Augmented semantic relations*

The augmented semantic relations come from all sources involving *A*, i.e., not only the row *(A)* in Table 2, but all ten rows marked in grey in column *A*, including, for example, $(A \cap I_{B \cup A})$. While some of these relations can be generated only by augmentation, some of them may also be present in the base, be inferable, or both. The proportion of augmented relations for each of these categories in FMA and GALEN is shown in Table 4.

Augmented semantic relations	**FMA** (N=392,314)	**GALEN** (N=32,922)
Can only be augmented	24.52 %	13.02 %
Also present in the base	11.12 %	20.50 %
Also inferable	65.14 %	68.28 %
(Both in the base and inferable)	0.78 %	1.80 %

Table 4. The augmented semantic relations

4.5 *Inferred semantic relations*

The inferred semantic relations come from all sources involving $I_{B \cup A}$, I_B, or I_A, i.e., not only the rows $(I_{B \cup A})$, (I_B), and (I_A) in Table 2, but all rows except *(B)* , *(A)*, and $(B \cap A)$. These rows are all marked in grey in column $I_{B \cup A}$, I_B, or I_A, and include, for example, $(I_{B \cup A} \cap I_A)$. While some of these relations can be generated only by inference, some of them may also be present in the base, be augmentable, or both. The proportion of inferred relations for each of these categories in FMA and GALEN is shown in Table 5.

Inferred semantic relations	**FMA** (N=11,896,508)	**GALEN** (N=4,356,244)
Can only be inferred	95.77 %	98.86 %
Also present in the base	2.11 %	0.64 %
Also augmentable	2.15 %	0.52 %
(Both in the base and augmentable)	0.03 %	0.02 %

Table 5. The inferred semantic relations

The last row in Tables 3, 4, and 5 corresponds in all three cases to relations which are present in the base and are also augmentable and inferable (3,082 in FMA and 590 in GALEN). These relations correspond to the following four rows in Table 2: $(B \cap A \cap I_{B \cup A})$, $(B \cap A \cap I_{B \cup A} \cap I_B)$, $(B \cap A \cap I_{B \cup A} \cap I_A)$, and $(B \cap A \cap I_{B \cup A} \cap I_B \cap I_A)$.

5 Discussion

5.1 *Specificity and common features of the various methods generating relations*

Each method provides specific relations. With the exception of I_B and I_A, each method contributes specific relations, i.e., relations that could not be generated by other methods. By definition, $I_{B \cup A}$ includes both I_B and I_A, i.e., every relation in I_B or I_A is also in $I_{B \cup A}$. However, as reflected by the two non-empty sets ($I_{B \cup A} \cap I_B$) and ($I_{B \cup A} \cap I_A$), not every relation generated by I_B can also be generated by I_A, and vice-versa. The largest proportion of specific relations is associated with inference (more than 95% of the relations inferred from FMA and GALEN can be generated only by inference). The base relations represent the second pool of specific relations (the proportion of base relations which cannot be generated by augmentation or inference is nearly 55% in FMA and 86% in GALEN).

Many relations can be generated by more than one method. Many relations generated by augmentation (11% in FMA and 20% in GALEN) and, to a lesser extent, by inference (2.1% in FMA and .6% in GALEN) are also present in the base, i.e., explicitly represented in most cases. There is also a significant overlap between the relations generated by augmentation and by inference, especially when examined from the perspective of augmented relations (about two thirds of augmented relations can also be inferred). Finally, a few hundred relations can be generated by all the methods under investigation. These relations, $B \cap A \cap I_{B \cup A} \cap I_B \cap I_A$, are present in the base, augmentable, and inferable from both the base and augmented relations. Examples of such relations include <*Variant muscle of thorax*, PART-OF, *Thorax*> in FMA and <*Deep Vein Of Leg*, PART-OF, *Leg*> in GALEN.

Relative contribution of each method. The source of the relations can be used to study the generative capabilities of the various methods producing these relations. From Figure 1, it is clear that, in both FMA and GALEN, the most important contribution comes from ($I_{B \cup A} \cap I_B$), i.e., inference based on relations present only in the base. This should not be surprising since inference performs similarly to a transitive closure applied to a combination of IS-A and PART-OF relations. In GALEN, relations from ($I_{B \cup A} \cap I_B$) account for nearly 90% of all relations. In FMA, however, this proportion is only 57%, but ($I_{B \cup A} \cap I_B$) and ($I_{B \cup A}$) together account for about 90%. What this illustrates is the role played by augmentation in FMA: while augmentation generally contributes few relations which could not have been generated otherwise, in FMA, augmented relations participate in a significant number of inferred relations.

Some sources do not provide any relations in GALEN. As mentioned earlier, only sixteen sources are found to contribute relations in GALEN, while there are nineteen such combinations in FMA. The three combinations missing in GALEN are ($B \cap A \cap I_{B \cup A} \cap I_A$), ($B \cap A \cap I_{B \cup A}$) and ($B \cap I_{B \cup A} \cap I_A$), which in all account for about 0.02% of relations in FMA. Augmentation plays a role in these three sources – directly or through inference – and it is consistent with earlier findings to see augmentation more strongly associated with FMA than GALEN.

5.2 Applications

5.2.1 Ontology auditing, validation, and maintenance

This study showed that the relations represented in ontologies – explicitly or not – may be redundant. When relations can be acquired from several different methods (e.g., explicitly represented and inferable from a combination of other relations), the relations in the ontology are no longer independent of each other. Redundancy may have beneficial effects for users of the ontology, such as providing direct links between important concepts. However, the dependence among equivalent relations or combination thereof is rarely explicit. Therefore, there is a chance that, over time, one relation be modified without modifying the dependent relations accordingly, leading to inconsistency.

Recognizing redundancy. Using techniques such as augmentation and inference, we showed that it is possible to identify relations which can be generated by more than one method, i.e., redundant relations. The percentage of redundant relations can be used as an indicator for auditing ontologies. A small percentage is likely to be associated with consistency and ease of maintenance, but the ontology may be more difficult to use by humans without the help of an inference engine.

Identifying dependence among relations. An ontology in which dependence among equivalent relations is explicit would be easier to maintain in a consistent state. For example, the following guidelines, inspired by the two ontologies of anatomy under investigation, could be adopted: (1) If a relation to be modified is represented explicitly and augmentable (6.74% in FMA as shown in Table 3), modify the explicit representation (e.g., <*P*, *PART-OF*, *W*>) and the equivalent concepts and relations (e.g., <*P*, *IS-A*, *Part of W*>, where *Part of W* embeds a reified *PART-OF* relationship). (2) If a relation to be modified is specific to the base relations (e.g., 54.92% in FMA as shown in Table 3), find all relations inferable from this relation (or using it for inference) and check their validity. (3) If a relation to be modified is represented explicitly and inferable (e.g., 38.83% in FMA as shown in Table 3), identify all relations from which this relation can be inferred, and check their validity.

Detecting inconsistency. Both FMA and GALEN were found to contain a small number of hierarchical cycles, resulting from either reflexive or circular hierarchical relations. Cycles may be found among the relations explicitly represented (e.g., <*Basal Ganglia*, *HAS-PART*, *Basal Ganglia*> in GALEN). More often, they are revealed while making explicit the implicit relations by augmentation and inference. For example, a *PART-OF* reflexive cycle was identified while augmenting from explicit relation <*Internal spermatic fascia*, *IS-A*, *Organ component of internal spermatic fascia*> in FMA. Additionally, the explicit relation <*Apex of urinary bladder*, *HAS-PART*, *Urinary bladder*> and the relation augmented from <*Apex of urinary bladder*, *IS-A*, *Subdivision of urinary bladder*> composed a direct hierarchical cycle in FMA.

5.2.2 Integration of multiple ontologies

Facilitating comparisons across ontologies. The ontologies to be integrated may use different modeling conventions, resulting not only in different relations being represented, but also in different ways to represent the same relations. In both cases, integration is facilitated by forcing all relations to be explicitly represented. This enables comparisons across systems based on simple matches among <$concept_1$, *relationship*, $concept_2$> relations on each side.

Detecting inconsistencies across ontologies. As mentioned earlier, applying augmentation and inference to the relations represented explicitly helped detect inconsistencies within ontologies. The same techniques are similarly powerful for detecting inconsistencies across ontologies. For example, the relationship between *Shoulder* and *Pectoral girdle* is *PART-OF* in FMA and *HAS-PART* in GALEN. However, while hierarchical cycles within ontologies are generally indicative of wrong relations, inconsistencies across ontologies may reveal either wrong relations (at least one of the two hierarchical relations is wrong) or errors in the alignment (the two concept names, although lexically similar, may stand for distinct objects in the world) [16]. In this case, the two concepts and their relations must be reviewed.

5.3 Advantages and limitations of this approach

Formalism. While other ontology tools (e.g., [6, 7]) require OKBC-compliance, the approach described in this paper is not tied to a particular formalism. FMA is a frame-based system and GALEN is based on description logics (DL). One requirement is to extract hierarchical relations from the system (e.g., superclass-subclass). The other requirement is to augment knowledge using linguistic clues in concept names. This presupposes the existence of concept names and is therefore not applicable to some 3,000 anonymous concepts in GALEN. Of note, the relations resulting from applying inference rules to hierarchical relations would certainly have been generated by a reasoner in a DL-based system. By generating these relations independently of such a system, however, our method is applicable to ontologies represented in other formalisms as well.

Domain. As a method for auditing ontologies (see section 5.2.1), this approach can be used with any ontology, as long as the requirements mentioned above are met. In its application to integrating multiple ontologies (section 5.2.2), this method requires that the ontologies to be integrated be of the same domain or, at least, have a significant overlap, as it is the case with FMA and GALEN. With other alignment methods (e.g., [17]), our method has in common that it intersects the content of several ontologies. However, we take advantage of techniques such as augmentation and inference, described in this paper and quantified for the FMA-GALEN alignment, to maximize the intersection.

Validation. One limitation of this study is that no validation of the relations generated has been performed yet. However, some elements of validation are built in the method. Redundant relations are likely to be valid, as are the relations represented in several ontologies. Finally, relations resulting from inference mechanisms should generally be valid. The evaluation provided by this method is essentially quantitative, resulting from auditing the ontology automatically. For this reason, our method can be seen as complementary of a qualitative analysis of taxonomic relationships (e.g., [18]), which requires extensive manual work.

6 Acknowledgements

The research was supported in part by an appointment to the National Library of Medicine Research Participation Program administrated by the Oak Ridge Institute of Science and Education through an interagency agreement between the U.S. Department of Energy and the National Library of Medicine. Thanks for their support and encouragement to Cornelius Rosse, José Mejino, and Kurt Richards for FMA and Alan Rector, Jeremy Rogers, and Angus Roberts for GALEN. Thanks also to Pieter Zanstra at Kermanog for providing us with an extended license for the GALEN server.

References

1. Corcho O, Fernandez-Lopez M, Gomez-Perez A. Methodologies, tools and languages for building ontologies. Where is their meeting point? Data & Knowledge Engineering 2003;46(1):41-64
2. Duc HN. Resource-bounded reasoning about knowledge [PhD Thesis]: University of Leipzig; 2001
3. Sima J, Cervenka J. Neural knowledge processing in expert systems. In: Cloete I, Zurada JM, editors. Knowledge-based neurocomputing. Cambridge, Mass.: MIT Press; 2000. p. 419-466
4. Zhang S, Bodenreider O. Aligning representations of anatomy using lexical and structural methods. Proc AMIA Symp 2003:(to appear)
5. Baader F, Horrocks I, Sattler U. Description logics as ontology languages for the Semantic Web. In: Hutter D, Stephan W, editors. Festschrift in honor of Jörg Siekmann: Springer; 2003. p. (to appear)
6. Noy NF, Musen MA. PROMPT: algorithm and tool for automated ontology merging and alignment. Proc of AAAI 2000:450-455
7. McGuinness DL, Fikes R, Rice J, Wilder S. The Chimaera ontology environment. Proc of AAAI 2000:1123-1124
8. Reed SL, Lenat D. Mapping Ontologies into Cyc. Proc of AAAI 2002 http://citeseer.nj.nec.com/509238.html.
9. Bailin SC, Truszkowsk W. Ontology negotiation as a basis for opportunistic cooperation between intelligent information agents. In: Cooperative Information Agents V, Proceedings; 2001. p. 223-228
10. Uschold M, Gruninger M. Creating semantically integrated communities on the world wide web. Proc International Workshop on the Semantic Web 2002 http://semanticweb2002.aifb.uni-karlsruhe.de/USCHOLD-Hawaii-InvitedTalk2002.pdf.
11. Rosse C, Mejino JL, Modayur BR, Jakobovits R, Hinshaw KP, Brinkley JF. Motivation and organizational principles for anatomical knowledge representation: the digital anatomist symbolic knowledge base. J Am Med Inform Assoc 1998;5(1):17-40
12. Noy NF, Musen MA, Mejino JL, Rosse C. Pushing the envelope: challenges in a frame-based representation of human anatomy: Technical Report of Stanford Medical Informatics; 2002. Report No.: SMI-2002-0925
13. Rector AL, Bechhofer S, Goble CA, Horrocks I, Nowlan WA, Solomon WD. The GRAIL concept modelling language for medical terminology. Artif Intell Med 1997;9(2):139-71
14. Rogers J, Rector A. GALEN's model of parts and wholes: experience and comparisons. Proc AMIA Symp 2000:714-8
15. Schulz S. Bidirectional mereological reasoning in anatomical knowledge bases. Proc AMIA Symp 2001:607-11
16. Bodenreider O. Circular Hierarchical Relationships in the UMLS: Etiology, Diagnosis, Treatment, Complications and Prevention. Proc AMIA Symp 2001:57-61
17. Wiederhold G. An Algebra for Ontology Composition. Proceedings of the 1994 Monterey Workshop on Formal Methods 1994:56-61
18. Welty C, Guarino N. Supporting ontological analysis of taxonomic relationships. Data & Knowledge Engineering 2001;39(1):51-74