Influence of the Digital Anatomist Foundational Model on Traditional Representations of Anatomical Concepts

Augusto V. Agoncillo, MD, José L.V. Mejino Jr., MD and Cornelius Rosse, MD, DSc
Structural Informatics Group, Department of Biological Structure,
University of Washington, School of Medicine, Seattle, WA 98195

ABSTRACT

A principled and logical representation of the structure of the human body has led to conflicts with traditional representations of the same knowledge by anatomy textbooks. The examples which illustrate resolution of these conflicts suggest that stricter requirements must be met for semantic consistency, expressivity and specificity by knowledge sources intended to support inference than by textbooks and term lists. These next-generation resources should influence traditional concept representation, rather than be constrained by convention.

INTRODUCTION

Textbooks and atlases remain the principal sources of knowledge for the education and training of health care providers. As far as we are aware, essentially all current educational computer programs rely on the same paradigms of knowledge organization; they use narrative text enhanced by interactive images. Such computer-based and hard copy sources, however, do not meet the needs of knowledge modelers engaged in the development of applications for managing biomedical information in clinical medicine and research. These needs have led to the establishment of clinical terminology projects which have generated machine-parseable, structured lists of terms that were culled from hard copy sources. The initial objective of our own work in anatomical knowledge representation[1] was to enhance the anatomy content of UMLS, which interrelates more than 30 vocabularies. Using the high level semantic types and relationships of the UMLS Semantic Network as a starting point, we extended the network and began to populate a hierarchy of anatomical semantic types (classes and subclasses) with concepts culled from time-honored and widely used anatomy textbooks[2,3]. We defined knowledge elements that we judged unique to anatomy among the biomedical sciences and proposed to generate a comprehensive and logical representation of this domain[4,5]. These knowledge elements concern the physical entities that constitute the body and the relationships that account for the integrity of the body as a structured object. We hypothesize, however, that this structural component, though more specific and narrow, is indispensable to anatomical knowledge, and it is this domain which is invoked explicitly or implicitly when reference is made to the body and its parts in any context of biomedical discourse.

For these reasons we call the scheme for this structural knowledge the Foundational Model of anatomy[5]. The model 1. declares a set of principles according to which knowledge should be structured; 2. specifies generalia and differentia as defining attributes, according to which concepts may be grouped together and distinguished from one another; and 3. explicitly defines the relationships between concepts. We are in the process of instantiating components of the model for major body parts. We encounter, and have to resolve, a number of conceptual problems in this process, because the semantic specificity enforced by the Foundational Model conflicts with traditional textbook representations of the information we are entering.

The objective of this report is to describe these conflicts and their resolutions so that knowledge modelers in other biomedical fields may benefit from our experience. We first describe the distinctions that can be made between anatomy textbooks, vocabularies, foundational models and knowledge bases and then illustrate with examples the kinds of conflicts and problems we have encountered, before drawing conclusions from our knowledge modeling experience.

TEXTBOOKS, VOCABULARIES, FOUNDATIONAL MODELS AND KNOWLEDGE BASES

The primary purpose of textbooks is to teach the uninitiated about a specific and narrow concept domain. Most textbooks are initially written for quite specific audiences (e.g., nursing students versus medical students). By contrast, the primary purpose of hard copy or computer-based terminologies (term lists, vocabularies) is to serve as a reference in order to assist in assuring standards and consistency in the use of terms.

Textbooks use examples liberally from concept domains outside their own discipline (e.g., anatomy texts refer to physiologic function and clinical examples). Hard copy term lists, (e.g., *Nomina Ana-*

tomica and its successor, *Terminologia Anatomica*), are focussed on their own domain, but capture very little meaning through the structure of their content. Computer-based vocabularies (e.g., SNOMED), however, tend to encompass much broader and more heterogeneous fields of knowledge than hard copy term lists. Many of their users may lack familiarity with several fields of the vocabulary, yet they rely on it as a knowledge source or reference. Therefore, the need for incorporating more and more meaning in vocabularies has become increasingly evident. Through its semantic network and semantic type definitions, UMLS has superimposed layers of meaning on the vocabularies that it interrelates.

A foundational model, as we have defined it[5], is a hybrid between textbooks and vocabularies. It is more constrained, and aims to be more expressive, than either of these knowledge sources. The purpose of a foundational model is to represent the structure of knowledge explicitly, comprehensively and consistently in a narrow and specific domain. Moreover, like a textbook, the model should be understandable to both novice and expert users. Unlike textbooks, however, the model should generalize to all fields of discourse that engage the model's concepts, and it should meet the requirements of any application for the knowledge that is represented by the model. Therefore a foundational model is primarily a resource for authors and application developers, who may or may not be familiar with the field of knowledge represented by the model. For instance, applications of the Digital Anatomist Foundational Model, even for anatomy education, require that (as in a textbook) authors associate images, physiologic function, radiology and some symptoms and physical signs with the model's own concepts. Such enhancements of an *anatomy foundational model* result in an *anatomy knowledge base*, which is suitable for meeting a set of educational objectives by a particular group of learners.

CONFLICTS IN CONCEPT REPRESENTATION

In previous publications we have described the Digital Anatomist Foundational Model and the methods for instantiating it[4,6]. This report addresses the differences and conflicts in concept representation we have encountered during instantiation of the model, particularly in the areas of semantic expressivity and specificity.

Semantic Expressivity. The constraint principle dictates that the Foundational Model should represent the physical organization of the human body. How well the model can express the structural relationships between the great variety of material objects that constitute the body depends on the class structure of inheritance hierarchies or ontologies, which group together and distinguishes from one another these objects, and on the sets of relationships that specify the associations between them. There are difficulties in deriving a sufficiently expressive classification from anatomy textbooks and term lists. Only in rare instances do the texts imply a classification in terms of structure. For example, bones, muscles and joints are classified according to their shape or constituent parts (e.g., long bones and short bones, unipennate and multipennate muscles, fibrous and synovial joints). However, essentially all other anatomical entities are described sequentially in chapters that are devoted either to organ systems (e.g., cardiovascular system), or to major body parts (so called regions), such as the thorax or upper limb, without reference to a classification system and without explicit definitions. *Nomina Anatomica/Terminologia*, the time-honored reference for classifying anatomical concepts, groups together anatomical entities according to the broad physiological processes in which the entities participate (e.g., circulation and digestion), regardless of their structural similarities or differences. This results, for instance, in a closer implied relationship between the liver and the appendix than between the liver and the spleen**.** Most clinical terminology projects base the organization of their anatomy term lists on *Nomina Anatomica*. Both the Read Codes[7] and GALEN[8] have enhanced this shallow classification by arbitrary subclasses. However, none of these enhancements aid the structural description of the body and its parts.

For these reasons we found it difficult to formulate a logical ontology that would capture the physical organization of the body as long as we limited ourselves to available sources. Therefore, we proposed an inheritance hierarchy, which includes a number of classes that are new in anatomy knowledge representation[4,6,9]. We explicitly defined all classes in accord with the definition principle: "Defining attributes of anatomical entities must be stated in terms of their constituent parts". This *Anatomy Ontology* (**Ao**) is the fundamental component of the Foundational Model. The logical foundation for the **Ao** is provided by the principle that designates *Organ* as the organizational unit of macroscopic anatomy, and classifies other structures according to whether they constitute organs or are constituted by organs. Whenever possible we have integrated in this ontology the classification implied by *Nomina/Terminologia Anatomica* and the textbooks we used for reference. To date, **Ao** has been instantiated for the thorax, abdomen, pelvis and perineum with over 25,000 anatomical concepts, and data entry continues for the remaining body parts.

Together with its class definitions, the **Ao** (established through the *-is a-* relationship) captures a substantial amount of knowledge about the kinds of entities that constitute the body. The result is, that such organs as the liver, lung and spleen are grouped together in the *Organ* subclass *Parenchymatous Viscus*. They all satisfy the definition of this subclass: "a viscus, some organ parts of which constitute lobes, segments, lobules, acini, or cortex and medulla". Stomach, urinary bladder, gall bladder share another set of structural attributes according to which they are assigned to the subclass *Hollow Viscus*: "a viscus, some organ parts of which constitute a wall; the viscus wall surrounds the organ (viscus) cavity". Although these two sets of organs are quite unlike one another, nevertheless they share a number of other attributes: all are organs located in the body cavity and the organ parts of all are embryologically derived from endoderm and splanchnic mesoderm. These are some of the defining attributes of *Viscus*, the parent class of *Parenchymatous Viscus* and *Hollow Viscus.* Although anatomy term lists and texts disregard such aspects of anatomical knowledge, the concepts of viscus, hollow and parenchymatous viscus, are taken for granted by most anatomists. Indeed, they use them in their teaching, although they may not have attempted to define them. Because of their emphasis on concrete leaf concepts and on the differences rather than the similarities between these concepts, traditional sources neglect abstract concepts. We have found, however, that abstract concepts of classes are a requirement for the logical organization of leaf concepts. The instantiation of these classes with many thousands of leaf concepts has provided an initial measure for the formative evaluation of **Ao.**

The amount of knowledge the Foundational Model can express is enhanced by its second component, the *Anatomical Structural Abstraction*, or **ASA**, which consists of a *Spatial ontology* (**So**) and a number of structural networks[5,10]. A companion report in these Proceedings illustrates the cross mapping of **Ao** concepts to the classes of **So**[6], which are defined in terms of spatial dimensions and shape. Conflicts that arise as a consequence of such cross mapping are discussed below in reference to specificity. We have instantiated the *Part-of Network* of **ASA** for concepts that have been entered in **Ao**. A logical representation of *-part of-* relationships is critical for the description of structured objects such as the human body and its parts**.**

For instance, a part of the right ventricle is designated as the 'infundibulum', the synonyms for which include 'outflow part [or tract] of right ventricle'. However, neither *Nomina/Terminologia Anatomica*, nor anatomy textbooks (except one[3]) name the remaining and larger part of the ventricle. A rule for instantiating the *Part-of Network* stipulates that if an entity of a given spatial dimension has one named part, its complementary remaining part or parts must also be named. In order to comprehensively instantiate the *Part-of Network* of the right ventricle, its part that is not the 'infundibulum' must be identified by a name. We have proposed naming this part the 'inflow part of the right ventricle'. The 'inflow part of the right ventricle' can be distinguished from the infundibulum not only by its function (it receives blood from the right atrium*)*, but also by its structure and anatomical features, as well as by its embryological derivation. Having made this decision, the selection of the preferred term for the 'infundibulum' (Latin for funnel) must be reconsidered. Although it conflicts with traditional naming conventions, the logical choice must fall on 'outflow part of right ventricle'; relegating 'infundibulum' to the role of a synonym. Figure 1 illustrates the outcome of this process.

This is but one example of the conflicts that had to be resolved in order to assure that the representation of the heart in the foundational model is logical and comprehensive. Based on these and similar examples, we contend that the Foundational Model is more expressive and logical than traditional sources, an assertion that will have to be empirically evaluated through applications of the Foundational Model in diverse biomedical fields.

Semantic Specificity. Much of the specificity of anatomical terms depends on the context in which they appear. These contexts include images embedded in narrative text, as well as other concepts that the text relates to an anatomical term. In the Foundational Model and in other ontologies, class assignments and relationships provide context. Therefore, in these knowledge sources, there is a stricter requirement for semantic specificity. For this reason, the Foundational Model does not admit homonyms, which leads to conflicts with traditional sources. For instance, anatomy sources and anatomists associate the term 'base of heart' with the posterior surface of the heart constituted by the left atrium. On the other hand, when clinicians use this term, they refer to the part of the chest wall that overlies the outflow tracts of the two ventricles, which are anterior in the heart. The narrative text context, as a rule, minimizes ambiguity, which is not the case in an ontology. Therefore we specify the meaning of the term by an extension, which eliminates ambiguity: 'base of heart (viewed anatomically)' and 'base of heart (viewed clinically)'.

The rule we apply is that each preferred term must be a associated only with one physical anatomical entity, and each physical anatomical en-

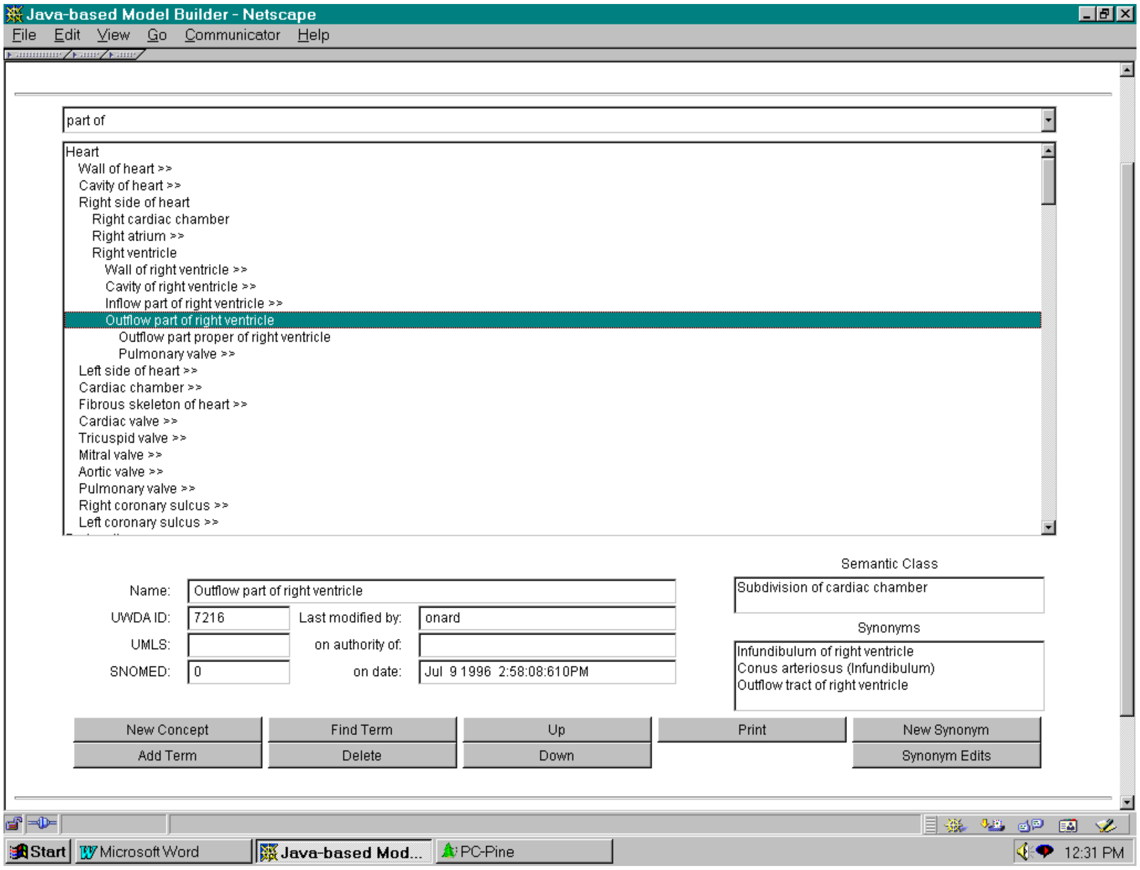

Figure 1. Screen capture from the Java-based Model Builder authoring program, showing a segment of the *Part-of Network* of the heart.

tity must have only one preferred term associated with it. This rule eliminates also homonyms that occur in strictly anatomical contexts. For instance, depending on the context, the term ‘muscle’ may refer to an organ, such as the biceps, or to the tissue that constitutes parts of the biceps. Adding extensions to the term eliminates ambiguity: ‘muscle (organ)’ and ‘muscle (tissue)’. The same rule enforces a qualification of terms for which real physical anatomical entities do not exist within the body. For instance, the term ‘right border of heart’ (and its various synonyms) can only be associated with a PA chest x-ray, but not with the real heart, because on the right, the heart presents a rounded surface, which projects on the x-ray film as a margin or border. We qualify the term by an extension: ‘right border of heart (viewed radiologically)’.

As mentioned earlier, cross mapping concepts of **Ao** to **So** classes of zero to three dimensions enforces specificity of yet another kind and leads to conflicts with traditional sources. This issue relates to the use of the term ‘region’, which we also explore elsewhere in these Proceedings in relation to anatomical spatial concepts[6]. Since it has not been explicitly defined, conventional use associates the term with several distinct classes of anatomical concepts. ‘Thorax’ and ‘neck’ are regarded as regions of the body by *Nomina/Terminologia Anatomica*, as well as by textbooks. In another context, however, the term ‘region’ may imply areas on the body’s surface, which are demarcated by arbitrary lines and other landmarks (e.g., epigastrium, abdominal quadrant). In yet another context, the same terms are used to designate arbitrary volumes within the body that underlie the designated surface areas, as in the phrase ‘The appendix is located in the right lower abdominal quadrant’.

The Foundational Model enforces specificity on the terms that we associate with these diverse concepts. We constrain the use of the term ‘body region’ by defining it as “a two-dimensional anatomical spatial entity that is demarcated by anatomical features or anatomical landmarks on the external or internal surfaces of anatomical structures”. We distinguish ‘right lower abdominal quadrant (surface)’ from ‘right lower abdominal quadrant (volume)’ and assign them to 2-D and 3-D classes of **So**, respectively. The definition of ‘body region’ precludes referring to major body parts such as ‘thorax’ and ‘neck’ by the term ‘body region’, because they all map to the subclass *Volume* in **So**. Therefore, we proposed that the term *Body Part* should specify this subclass of structures, and restricted the meaning of the term by defining this subclass: "Body part is an anatomical structure that consists of members of diverse subclasses of *Organ*, which is surrounded or partially covered by skin; together with all other body parts, a body part constitutes the body."

CONCLUSIONS AND DISCUSSION

By establishing the Digital Anatomist Foundational Model for macroscopic anatomy, we wish to stimulate interest by knowledge modelers and other investigators in order to 1. extend the anatomy model to microscopic anatomy and molecular structure; 2. generate knowledge bases for anatomy education and clinical practice; and 3. develop foundational models in other biomedical fields that rely on anatomy. Such next-generation knowledge sources are intended for supporting inference. They have to satisfy more stringent requirements for concept representation than time-honored textbooks and term lists. They also have to have a more expressive semantic structure than currently available clinical terminologies. Our experience suggests that knowledge modelers should not be bound by the conventions of traditional knowledge sources. Rather, they should look on the development of new knowledge sources as an opportunity for evaluating semantic structure in a domain with the objective of enhancing expressivity and specificity.

The first requirement is to specify a sufficiently narrow field of knowledge that can be consistently modeled. Second, a set of principles should be declared according to which knowledge is to be modeled. Third, sets of abstract concepts should be proposed as classes of an inheritance hierarchy, which should be populated with a preliminary but substantial cohort of instances in accord with explicit definitions that specify the attributes according to which instances should be included in, or excluded from, classes of the ontology. A structured set of foundational relationships should be proposed and associated with concepts of the ontology in order to model a coherent segment of knowledge. These representations should be evaluated and modified before the model is instantiated for the comprehensive domain.

Because of its generalizable concept domain, we contend that the Digital Anatomist Foundational Model provides a prototype for next-generation knowledge sources in the biomedical sciences. Although we do not anticipate that these new resources will replace textbooks, they should exert an influence on them. Conceptualization of the large body of leaf concepts prevalent in these books should be aided by the logical classification, explicit definitions and semantic specificity that the new knowledge sources will provide. These enhancements of the texts should promote reasoning with the domain's concepts and replace rote memorization during the learning process.

Acknowledgment

This work was supported by NLM contract LM43546 and NLM grant LM06316.

References

1 Rosse C, Said MB, Eno K, Brinkley JF. Enhancements of Anatomical Information in UMLS Knowledge Sources. J Am Med Inform Assoc. AMIA'95 Symp. Suppl. 1995;873-877.
2. Williams PL. (editor) Gray's Anatomy 38th ed. New York: Churchill-Livingstone, 1995.
3. Rosse C, Gaddum-Rosse P. Hollinshead's Textbook of Anatomy, 5th ed. Philadelphia: Lippincott-Raaven, 1997.
4. Rosse C, Mejino JL, Modayur BR, Jakobovits R, Hinsaw K, Brinkley JF. Motivation and organization principles for anatomical knowledge representation: The Digital Anatomist Symbolic Knowledge Base, J Am Med Inform Assoc. 1998;5:17-40.
5. Rosse C, Shapiro LG, Brinkley, JF. The Digital Anatomist Foundation Model: Principles for Defining and Structuring its Concept Domain. J Am Med Inform Assoc. 1998;820-824.
6. Mejino JLV, Rosse C. Conceptualizations of Anatomical Spatial Entities in the Digital Anatomist Foundational Model. Proceedings AMIA Fall Symposium 1999. Submitted.
7. Schulz EB, Price C, Brown PJB, Symbolic anatomic knowledge representation in the Read Codes Version3. J Am Med Inform Assoc. 1997;4:38-48.
8. Rector AL, Gangemi A, Galeazzi E, Glowinski AJ, Rossi-Mori A. The GALEN CORE model schemata for anatomy: Proc 12th Internat Congr Eur Fed Med Inform (MIE 94), Lisbon. IOS Press, 1994;229-233.
9. Mejino JLV, Rosse C. The potential of the Digital Anatomist Foundational Model for assuring consistency in UMLS sources. J Am Med Inform Assoc. AMIA'98 Symp Suppl 1998;285-829.
10. Neal P, Shapiro LG, Rosse C. The Digital Anatomist Structural Abstraction: a scheme for the spatial description of anatomical entities. J Am Med Inform Assoc. AMIA'98 Symp Suppl. 1998;423-427.