The Digital Anatomist Foundational Model: Principles for Defining and Structuring its Concept Domain

Cornelius Rosse M.D., D.Sc.,[1]
Linda G. Shapiro Ph.D.,[2] and James F. Brinkley, MD, Ph.D.[1,2]
Structural Informatics Group,
Department of Biological Structure[1] and Department of Computer Science and Engineering[2],
University of Washington, Seattle, WA 98195

ABSTRACT

*We define a foundational model as an abstraction of a body of knowledge that explicitly declares the principles and concepts necessary for coherently and consistently modelling a knowledge domain. Principles for a foundational model of anatomy are defined and used to specify the components of such a model. These components include an anatomy ontology (**Ao**), an anatomical structural abstraction (**ASA**), an anatomical transformation abstraction (**ATA**) and metaknowledge (**Mk**), which comprises the rules for representing relationships in the other three components of the model. The foundational model **Fm** is therefore specified as the four-tuple **Fm=(Ao,ASA,ATA,Mk)**. We hypothesize that this abstraction captures the information that is sufficient and necessary for describing the anatomy of any physical entity that constitutes the body, as well as that of the body itself.*

INTRODUCTION

A requirement for logic-based representation of clinical data is the establishment of foundational models for those concept domains that generalize to diverse fields of clinical medicine[1]. Time and probability are among such concepts, and so are the sciences or knowledge domains that are basic to clinical medicine. Among the basic biomedical sciences, anatomy occupies perhaps the most fundamental position: not only clinical medicine, but also such basic sciences as physiology and pathology depend on anatomical concepts for communicating knowledge in their respective fields. A foundational model of anatomy, therefore, is a prerequisite for logic-based concept representation in both the basic and the clinical sciences.

Current medical terminology projects that are motivated by computerizing the clinical record incorporate a large number of anatomical terms[2-4]; however, they are neither consistent, compatible, nor sufficiently comprehensive in their representation of anatomical concepts to meet the requirements of a foundational model of anatomy. The explanation for these shortcomings may be that the principles have not been articulated for defining and structuring the anatomical concept domain of these knowledge sources. The anatomy section in these large, computer-based, clinical vocabularies is designated as the 'T' (topography) axis by SNOMED[2], the 'Symbolic Anatomic Knowledge Base' by the Read Codes[3] and the 'Anatomy Model' by GALEN[4]. The schemes used for modeling the concepts in these three vocabularies are as varied as the terms that refer to them, suggesting that an underlying foundational model has either not been identified, or that more than one representation can capture it. The GALEN anatomy model[4] has been regarded by some (though not its developers) as a foundational model. However, since its authors, by their own admission, could not identify principles according to which anatomical entities could be classified, the taxonomy is arbitrary.

Our objective is to develop a foundational model as a hypothesis; the model can then be empirically evaluated in different fields of the biomedical sciences to determine whether the model can meet their respective needs for anatomical information. This report describes the processes that have led to the formulation of the model, the principles that guided the first phase of its development, and the plans for its enhancements and evaluation.

METHOD OF APPROACH

We selected anatomy education for medical and dental students as the primary application for driving the development of the foundational model. Only in such educational programs is anatomy explored comprehensively; clinical fields and their problems focus on selected domains of anatomy.

Based on our experience with the writing and editing of anatomy textbooks[5,6], classroom teaching over three decades, and the design of computer-based educational programs[7-9], we have proposed a set of principles for formulating the foundational model. We gauged the scope of the information the model had to represent through three processes: 1. surveying anatomy exams at five universities; 2. experience gained through developing the anatomy component of USMLE Step 1 examinations for the National Board of Medical Examiners; and 3. developing templates for the kinds of information students need to acquire and synthesize for conceptualizing different classes of anatomical entities (e.g., bone, muscle, nerve, viscus). The usefulness of these templates for knowledge organization by many generations of students has been validated, independent of the teaching staff, by the Department of Medical Education, University of Washington. We used the information gathered through these processes to modify the principles and derive the requirements for the foundational model. We prioritized the tasks for formulating the model and used a conjoint graduate course in computer science and biological structure (CS 590BR: Anatomical Knowledge Representation) as a forum for evaluating segments of the evolving model. As portions of the model became instantiated for thoracic anatomy, they were subjected to the same evaluation process.

THE MODEL

First we define a foundational model; then we declare the principles that are guiding the development of the Digital Anatomist foundational model, propose the abstraction for the high level model, and discuss the priorities for its implementation and the strategy for its empirical evaluation.

Definition.

A foundational model is a conceptualization (abstraction) of a coherent body of knowledge about a domain, which declares the constraints (principles) for including concepts and relationships in the model, explicitly defines the concepts, and also makes explicit the foundational assumptions about relationships between concepts that are both necessary and sufficient for coherently and consistently modelling the structure of the specified knowledge domain.

We have previously distinguished between two concepts denoted by the term *anatomy*: anatomy(structure) and anatomy(science), and provided definitions for each[10]. Although anatomical discourse in education and other fields embraces diverse concepts, such as function, dysfunction and biomechanics, it is the information that pertains to the physical organization (structure) of the body that distinguishes anatomy(science) from other biomedical sciences. A foundational model of anatomy, therefore, must strictly concern itself with anatomy(structure). The abstraction that aims to represent this structure should resemble, in its own structure, the physical organizational plan of the body.

Foundational Principles.

Principles are assertions that provide the basis for reasoning and action. The formulation of the Digital Anatomist foundational model is guided by seven principles:

1. Constraint principle: The abstraction should model the physical organization of the human body. Physical anatomical entities are material objects, spaces and substances; the generation of these entities is regulated by the coordinated expression of groups of genes.

2. Root concept principle: All physical anatomical entities should be assigned to one of three top-level classes (root concepts): *Anatomical structure*, the dominant class, and *Anatomical spatial entity* and *Body substance*, which are defined in relation to *Anatomical structure* [10].

3. Definition principle: Defining attributes of anatomical entities must be stated in terms of their constituent parts, and in terms of the entities which they, in turn, constitute.

4. Constitutive principle: Anatomical structures are constituted by anatomical structures; the largest of these is the whole *organism* (in the present instance, the human body), and the smallest - for the first iteration of the model - the *cell. Anatomical spatial entities* are constituted by anatomical spatial entities of the same dimension, the smallest of which is a point, which has zero dimensions and no constituent parts.

5. Organizational unit principle: The unit of macroscopic anatomy is *Organ*; subclasses of *Anatomical structure*, other than *Cell*, either constitute organs or are constituted by organs. (Cells exist in body substances as well as in organs.)

6. Spatial relationship principle: Spatial relationships among anatomical structures and anatomical spatial entities are physical relationships which are established by processes regulated by coordinated gene expression.

7. Representation principle: The model should be formulated as an abstraction of *canonical anatomy* but should support queries and reasoning about *instantiated anatomy* (in accord with the definitions we have provided for these concepts[10]).

The High Level Model.

In addition to restricting the concept domain, the principles also specify four classes of attributes among anatomical concepts: 1. taxonomic or class inclusion relationships (-is a-, -type of-); 2. constitutive or part-whole relationships; 3. spatial relationships; and 4. transformational relationships. The latter include morphological changes resulting from inductive, developmental, growth, and involutional processes. These four classes of relationships are the ones that seem necessary for capturing the structure of the knowledge that deals with physical entities of the body, warranting their designation as *foundational relationships* or *attributes.*

The practical approach we chose to the foundational model was to conceptualize it as having four components: 1. an *Anatomy ontology* (***Ao***), which specifies the taxonomic relationships of anatomical entities and assigns them to classes according to defining attributes which they share with one another and by which they can be distinguished from one another; 2. an *Anatomical structural abstraction* (***ASA)***, which describes the partitive (meronymic) and spatial relationships of the concepts represented in the ontology; 3. an *Anatomical transformation abstraction* (***ATA***), which describes the time-dependent morphological transformations of the concepts represented in the ontology during the human life cycle; and 4. *Metaknowledge* (***Mk***), which comprises the principles and sets of rules, according to which the relationships are represented in the model's other three component abstractions. Thus the foundational model ***Fm*** may be specified as the four-tuple

$$\boldsymbol{Fm = (Ao, ASA, ATA, Mk)} \qquad (1)$$

We hypothesize that this abstraction captures the information that is sufficient and necessary for describing the anatomy of any structure (physical object) or space that constitutes the body, as well as that of the entire human body itself. The foundational model for the anatomy of the entire body ($\boldsymbol{Fm}_{\mathrm{BODY}}$) may, therefore, be conceived as a composite of all the foundational models of physical anatomical entities ($\{\boldsymbol{Fm}_{\mathrm{PHYSICAL\_ANATOMICAL\_ENTITY}}\}$) that constitute the body. Thus,

$$\boldsymbol{Fm}_{\mathrm{BODY}} = \{\boldsymbol{Fm}_{\mathrm{PHYSICAL\_ANATOMICAL\_ENTITY}}\} \qquad (2)$$

The latter assertion has practical implications: it provides for building the foundational model for the entire body, stepwise, from the models of individual physical anatomical entities. It is possible, therefore, to test the validity of the model and refine it as the model is incrementally instantiated for organs, body parts and organ systems.

Implementation Priorities.
Consistency between different components of the foundational model is best assured by an ontology that is formulated in accord with the foundational principles. The ***ASA*** and ***ATA*** represent the concepts of ***Ao*** in different contexts. The ontology is, therefore, a prerequisite for both the ***ASA*** and ***ATA***. Inheritance of spatial and transformational attributes in the respective abstractions must rely on the classification scheme of the ***Ao***. Therefore, an ontology has to be in place for a particular segment of the concept domain before formulation of the structural and transformational models for that segment can begin.

Relying on the UMLS semantic network as a starting point[11], we have formulated an ***Ao***[10]. We implemented the ontology as a semantic network, and gained a measure of the validity of its classes through instantiating the ontology with canonical concepts of thoracic anatomy. Instantiation of the ***Ao*** for other body parts is in progress and will proceed in parallel with the development of the ***ASA***. We have reordered concepts of thoracic anatomy into acyclic directed graphs using relationships such as -part of-, -branch of-, -tributary of-, which provide elementary components of ***ASA***[10].

Currently there is no consistent ontology or formal representation of concepts in the domains of embryology, growth and involution. Physical entities that must be included in a transformation ontology satisfy the defining attributes of *Anatomical structure* and *Anatomical spatial entity*. Therefore, the transformation ontology should be developed as an extension of the ontology for macroscopic anatomy (***Ao***)[10]. However, currently we are focusing on the development and instantiation of ***ASA*** for the entire body, since we perceive that the most urgent need, both in education and clinical medicine, is a model in terms of which spatial relationships among macroscopic anatomical entities can be described.

We hypothesize that an abstraction comprised of an ontology of spatial objects and of three interacting networks, which describe topological, part-whole, and spatial association relationships among anatomical concepts, will meet requirements for the logic-based modeling of the three-dimensional structure of the body and its parts. We specify the ***ASA*** as a critical part of the Foundational Model equation and define it as a 4-tuple:

$$\boldsymbol{ASA = (So, Tn, Pn, SAn)} \qquad (3)$$

where: So = Spatial Object ontology
Tn = Topology network
Pn = Part-of network

SAn = Spatial Association network

In analogy with the composite foundational model, our hypothesis proposes that the ***ASA*** captures the information that is sufficient and necessary for describing the structure of any physical object or space that constitutes the body, as well as that of the entire human body itself. The ***ASA*** of the entire body may, therefore, be conceived as a composite of all ***ASA***s of anatomical structures and anatomical spaces that constitute the body. The Anatomical Structural Abstraction is distinct from other approaches that have been proposed for the symbolic description of anatomical spatial relationships (for examples see Tagare et al.[12] and Robinson et al.[13]), in that it is not limited to object recognition in medical images, it generalizes to all parts of the body and all types of physical anatomical entities, and it accommodates all relationships that are necessary for describing the 3D structure of the body.

Component schemes of ***Fm*** are described by two separate reports in these proceedings[14,15], which illustrate the schemes with anatomical examples.

Operationally, the anatomy ontology and the part-of network of the ***ASA*** (including -branch of- and -tributary of- relationships) have been implemented as semantic networks, consistent with the UMLS Semantic Network and Metathesaurus, and are accessible through UMLS 1998. However, in order to provide a sufficiently expressive representation language and scheme for implementing the multiple relationships of the foundational model, and the ***ASA*** in particular, we have opted for the Protégé system[16]. Protégé promotes the development of domain ontologies and couples them with problem solving methods that can be tailored to the requirements of various applications in biomedical education, research and clinical practice. The design and formal representation of the foundational model of anatomy will be supported by the Protégé suite of knowledge-acquisition tools in collaboration with the developers of the Protégé system.

Formative Evaluation.

We report one measure for evaluating the foundational model in these proceedings[14]: inconsistencies in anatomical concept representation we detected in UMLS source vocabularies could be explained and reconciled by the current version of the foundational model. We will make use of selected problem domains in anatomy education for evaluating successive iterations and enhancements of the model. We will make the model available to medical and dental students through the Web-based Digital Anatomist information system[9]. We report in these proceedings the development of the next generation, knowledge-based Digital Anatomist interface[17], which will display information represented in ***Ao*** and ***ASA***, in response to queries, and will associate the answers with a reference image. We will use the knowledge organization templates for evaluating the answers, and thereby test the hypothesis that the information provided with the aid of the model is not only necessary but also sufficient for medical and dental students to acquire anatomical knowledge about selected entities. Once the model is validated for its consistency and educational applications, we will assess whether it can be generalized to the problem domain of "anatomy intensive" clinical fields, such as radiation treatment planning.

DISCUSSION

Formalized foundational models for concept domains that generalize to diverse fields of the biomedical sciences, seem as yet to be a goal rather than an accomplishment. We believe that formulating such a model for anatomy should prove useful and promote the development of foundational models in other fields of the biomedical sciences. We regard foundational models as distinct from knowledge bases in that the models are restricted to a specified concept domain that generalizes to a number of fields, whereas a knowledge base may combine a number of foundational models relevant to a particular problem domain and embellish the models with other information, including non-foundational relationships. In the case of an anatomy knowledge base, functional relationships would fall into such a category.

The scheme we propose as a foundational model is the first in the history of anatomy which 1. declares a set of principles according to which knowledge should be structured; 2. defines and constrains the concept domain of the field; 3. specifies generalia and differentia as defining attributes, according to which concepts may be grouped together and distinguished from one another; 4. explicitly defines relationships between concepts; and 5. proposes a generalizable and comprehensive scheme for the symbolic representation of the structure and spatial disposition of the material objects and spaces that constitute the body. The knowledge organization schemes implicit in textbooks of anatomy are inadequate for meeting the needs of a new generation of users in medical informatics, and fall short of the requirements for logic-based modelling of anatomy(science). The development of a logic-based scheme for organizing anatomical knowledge calls for a deep understanding of the field and extensive experience with its practice. Such a background can provide the confidence for proposing logic-based schemes that are not only explicit, but also more consistent than those implied in the traditional sources, even if this calls for breaking with traditional ways of organizing anatomical knowledge.

We recognize that the foundational model we are proposing for anatomy is far from complete and its validation requires empirical evaluation. Our justification for describing the model at its current stage of development is to invite such evaluations through its educational and clinical applications.

Acknowledgments

This work was supported in part by contract LM13506 and grant LM06316, National Library of Medicine.

References

1. Campbell KE, Das AK, Musen MA. A logical foundation for representation of clinical data. J Am Med Inform Assoc 1994;1:218-32.
2. Côté R (ed). Systematized Nomenclature of Human and Veterinary Medicine (SNOMED International), version 3.1. Northfield, IL: College of American Pathologists, and Schaumburg, IL, American Veterinary Medical Association, 1995.
3. Schulz EB, Price C, Brown PJB. Symbolic anatomic knowledge representation in the Read Codes Version 3: Structure and application. J Am Med Inform Assoc 1997;4:38-48.
4. Rector AL, Gangemi A, Galeazzi E, Glowinski AJ, Rossi-Mori A. The GALEN CORE model schemata for anatomy: Towards a re-usable application-independent model of medical concepts. In: Barahona P, Veloso M, Bryant J, editors. Proceedings of the 12th International Congress of the European Federation for Medical Informatics (MIE 94), Lisbon. IOS Press, 1994:229-33.
5. Rosse C, Clawson DK. The musculoskeletal system in health and disease. Hagerstown: Harper & Row, 1980.
6. Rosse C, Gaddum-Rosse P. Hollinshead's Textbook of Anatomy. 5th ed. Philadelphia: Lippincott-Raven, 1997.
7. Rosse, C., 1995 The potential of computerized human anatomy in the training of health care providers. Acad Med 70:499-505.
8. Conley DM, Kastella KG, Sundsten JW, Rauschning W, Rosse C. Computer-generated three-dimensional reconstruction of the mediastinum correlated with sectional and radiological anatomy. Clinical Anatomy 1992;5:185-202.
9. Brinkley JF, Bradley SW, Sundsten JW, Rosse C. The Digital Anatomist information system and its use in the generation and delivery of web-based anatomy atlases. Comp Biomed Res 1997;30:472-503.
10. Rosse C, Mejino JL, Modayur BR, Jakobovits R, Hinshaw KP, Brinkley JF. Motivation and organizational principles for the Digital Anatomist Symbolic Knowledge Base: an approach toward standards in anatomical knowledge representation. J Am Med Inform Assoc 1998;5:17-40.
11. McCray AT, Nelson SJ. The representation of meaning in the UMLS. Meth Inform Med 1995;34:193-201.
12. Tagare HD, Vos F, Jaffe CC, Duncan JS. Arrangement: a spatial relation comparing part embeddings and its use in medical image comparisons. In: Proceedings of the 13th International Conference on Information Processing in Medical Imaging (IPMI 93), Flagstaff (AZ), 1993;132-48.
13. Robinson GP, Colchester ACF, Griffin LD. Model-based recognition of anatomical objects from medical images. Image and Vision Computing 1994;8:499-507.
14. Mejino JL, Rosse C. The Potential of the Digital Anatomist Foundational Model for Assuring Consistency in UMLS Sources. Proceedings AMIA Fall Symposium 1998.
15. Neal PJ, Shapiro LG, Rosse C. The Digital Anatomist Spatial Abstraction: a scheme for the spatial description of anatomical entities. Proceedings AMIA Fall Symposium 1998.
16. Musen MA, Gennari JH, Eriksson H, Tu SW, Puerta AR. PROTÉGÉ II:computer support for development of intelligent systems from libraries of components. MEDINFO 95, The eighth World Congress of Medical Informatics, Vancouver, B.C. Canada, pp.766-770.
17. Brinkley, J.F. and Rosse, C. 1998. Requirements for an on-line knowledge-based anatomy information system. Proceedings AMIA Fall Symposium 1998.