Ontology-Based Representation of Simulation Models of Physiology

Daniel L. Rubin, MD, MS[1], David Grossman, PhD[1], Maxwell Neal, BS[2], Daniel L. Cook, PhD[3], James B. Bassingthwaighte, MD, PhD[2], and Mark A. Musen, MD, PhD[1]
[1]Stanford Medical Informatics, Stanford University School of Medicine, Stanford, CA, USA
[2]Departments of Bioengineering and [3]Physiology & Biophysics, University of Washington, Seattle, WA

Abstract

Dynamic simulation models of physiology are often represented as a set of mathematical equations. Such models are very useful for studying and understanding the dynamic behavior of physiological variables. However, the sheer number of equations and variables can make these models unwieldy, difficult to understand, and challenging to maintain. We describe a symbolic, ontologically-guided methodology for representing a physiological model of the circulation. We created an ontology describing the types of equations in the model as well as the anatomic components and how they are connected to form a circulatory loop. The ontology provided an explicit representation of the model, both its mathematical and anatomic content, abstracting and hiding much of the mathematical complexity. The ontology also provided a framework to construct a graphical representation of the model, providing a simpler visualization than the large set of mathematical equations. Our approach may help model builders to maintain, debug, and extend simulation models.

A. Introduction

Mathematical models are needed for quantitative representation of physiological processes, and they can help researchers understand how components of these processes contribute to disease. The approach to building these physiological models is to create detailed mathematical representations of the essential features of the physiological system at multiple scales, including anatomy and mechanics, in an attempt to produce a realistic computational model of physiology [1]. The resulting models usually comprise a large set of ordinary differential equations (ODEs) describing time-varying physiological variables and parameters.

While mathematical models are a useful representation for generating quantitative simulations, they can be unwieldy for the researchers who build them. It is difficult to comprehend, maintain, and extend models containing huge numbers of equations. Model designers often think in terms of graphical or schematic representations of the system being modeled, rather than the long lists of equations entailed by that representation. The equations themselves are compiled by studying the schematic diagram of these models, a process that can be difficult and error-prone when models are large. It can also be difficult to extend these models to add new components or conditions because of the need to relate changes in the schematic representation to the necessary updates to the equations in the mathematical formulation of the model.

Graphical representations of complex biomedical systems are useful to help users comprehend the components and attributes of these systems, but purely graphical models do not directly support analyzing and simulating these systems mathematically. There is a need for a representation paradigm that enables researchers to create models in a visual schematic format while maintaining explicit connections to the mathematical representation of those models in order to simulate and analyze biological systems.

Our hypothesis is that it is possible to represent both the anatomic components and the mathematical elements of a physiological model of hemodynamics using an ontology, and that the ontology can enable a user to work in an environment where anatomical relationships can be represented graphically and where symbolic depiction of physiological constraints drives the generation of appropriate mathematical equations. In this paper, we describe our work in creating this ontology, containing explicit knowledge of the anatomical components of a model of cardiovascular hemodynamics as well as the types of equations needed to describe the quantitative aspects of the model. We also describe our approach of instantiating a particular model as a set of linked instances created from the ontology, of connecting the instances of anatomic components to create a circulatory loop, and of describing attributes of those connections using a graphical view of the ontology instances. Finally, we describe our methodology for deriving the set of mathematical equations for the model given its schematic representation. We believe this methodology may enable a generalizable, extensible approach to represent, maintain, and evolve multi-scale simulation models using the graphical schematic form in which these models are originally conceived.

B. Methods

We undertook this work in stages. First, we studied an existing ODE model of cardiovascular hemodynamics [2, 3] in order to develop abstractions of the physiological and mathematical knowledge needed to describe this and similar models using entities and rela-

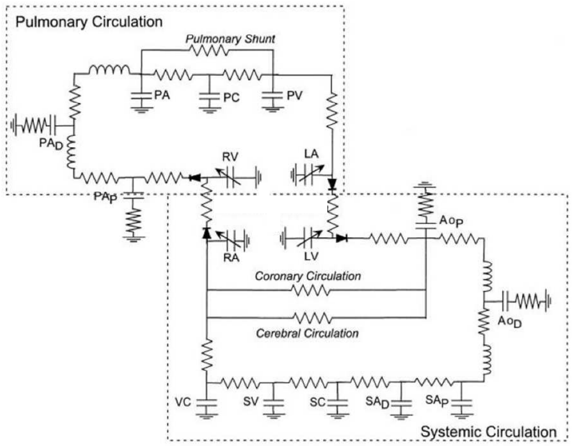

Figure 1. Representation of a closed-loop circulatory model. *The representation is a graphical schematic diagram, consisting of anatomic components, such as Pulmonary Vein (PV) and Left Atrium (LA) that are linked by connections that have properties such as resistance, capacitance, and inductance (modified after [2]).*

tions in an ontology. Next, we built an ontology to describe these entities and relations, and we subsequently created instances using the ontology to represent a particular physiological model. Finally, we assessed the potential benefits of our approach.

B.1. Model of Cardiovascular Hemodynamics

In order to create an ontology of cardiovascular physiology models, we examined an existing mathematical model of cardiovascular hemodynamics [3] with the goal of abstracting its content. In that model, the cardiovascular system is represented as a circulatory loop of connected components (anatomic entities), some of which represent a four-chambered heart, a systemic circulation, and a pulmonary circulation (Figure 1). The model consists of (1) a **schematic diagram** representing the structural relationships in the model, and (2) a **set of mathematical equations** representing the quantitative behavior in the model and that are used to generate a time-varying simulation. In the schematic diagram representation of the model, there are **anatomic entities, connections** between the components, and **physiological attributes** related to the hemodynamics of the components and their connections, such as resistance, capacitance, and inductance (Figure 1). In the mathematical representation of the model, there is a list of ordinary differential equations that express the quantitative relationships among the physiological variables and parameters (155 variables and parameters, 83 equations, and 21 initial conditions).

There were two types of equations in the model: those that occur throughout the model in the same canonical form ("canonical equations"), and specialized equations that occur only once in the model and usually substitute for a canonical equation to describe unique behavior in a specific portion of the model ("custom equations"). There were five types of canonical equations, each describing particular physiological attributes: the dependence of flow on volume, on resistance, on capacitance, and/or on inductance, as well as flow conservation (sum of flows entering a junction equals the sum of flows leaving junction). Thus, particular combinations of physiological attributes determine the mathematical equations appropriate to express those attributes.

Accordingly, we determined that the key *knowledge representation elements* needed for our ontology included **anatomic entities**, **connections** between anatomic entities, and **physiological attributes**. In addition, there was domain knowledge relating the *physiological attributes* that describe connected anatomic entities to the *mathematical equations* used to express the quantitative representation of that portion of the model. In addition to representing the details of the physiological model, w explored the opportunity to use our knowledge representation to relate a graphical schematic diagram view of the model to the mathematical equations that express it in quantitative terms.

B.2. Ontology of physiological models

From the knowledge requirements described in Section B.1, we created an ontology containing two types knowledge needed for models of cardiovascular hemodynamics: (1) representation of the elements used to create schematic physiological models, and (2) explicit description of the types of equations in the model. We used the Protégé ontology development suite (http://protege.stanford.edu) to build the ontology. To represent schematic physiological models, the ontology contains classes for nodes, arcs, and physiological attributes (Figure 2). Nodes correspond to anatomic entities in the model; arcs represent connections among the anatomic entities; and attributes are physiological parameters that are associated with individual anatomic entities (such as pressure) or the connections between pairs of anatomic entities (such as flow and resistance attributes).

The anatomic entities are drawn from those in the Foundational Model of Anatomy (FMA) [4] in order to enable interoperability among models that use common terminology. The FMA is an ontology containing the concepts and relationships that pertain to the structural organization of the human body. The FMA contains rich anatomic knowledge, relating anatomic entities to other entities via relations such as connectivity and adjacency that could be used in checking the structure of physiological models for consistency.

Our ontology of physiological models also contains knowledge about the types of equations in cardiovascular models. The types of equations used in these models depend on the physiological attributes; thus, the physiological attribute classes contain knowledge

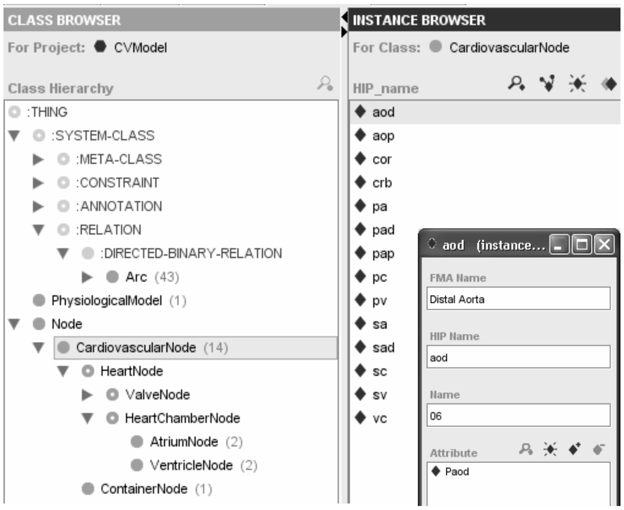

Figure 2. ***Ontology of cardiovascular physiology models.*** *The ontology classes are shown on the left, and instances are shown on the right. The ontology contains classes describing components of physiology models, including anatomic entities (nodes), connections among those entities (arcs), and physiological attributes on nodes and arcs. The ontology also describes the types of mathematical equations needed to describe these attributes quantitatively. The anatomical entity for distal aorta (lower right) has a reference to the corresponding entity in the FMA, and internal name, identifier, and a pressure attribute.*

about the mathematical equations used to express those attributes in the model. There are two types of physiological attribute classes, canonical and custom (Figure 3). Canonical attribute classes contain the information necessary to generate the canonical form of the equation appropriate to the attribute, including the value and units for that attribute. For example, the canonical resistance attribute class would be used to generate the standard form of the equation for resistance, the Ohm's Law equation (P = FR), providing the value and units for that particular equation. Custom attribute classes contain strings representing the particular specialized equations needed for the model and that would not be generated using the canonical form of the equation.

B.3. Using the ontology to build graphical models and generate equations

While the ontology describes the components of the graphical and mathematical aspects of cardiovascular models generically, it does not contain any instance data—the ontology does not specify the details of a particular cardiovascular model. To create a new ontology-based model, instances of the appropriate ontology classes are created and given the necessary data values. Specifically, instances of the Node or Arc classes are created to represent the anatomical entities and their connections in the cardiovascular model (Figure 2). In addition, instances of the physiological attribute classes can be associated with anatomic entities and arcs, to describe the mathematical properties of those elements. Instances of physiological attributes are associated with arcs using attributed relations (Figure 2). The type of physiological attribute a user chooses for a connection (canonical or custom) depends on whether one wants to specify a specialized form of the equation governing that attribute (custom attribute), or whether the canonical form of the equation is appropriate (canonical attribute).

In order to visualize the cardiovascular model as a set of linked instances, we use the Protégé Diagram Widget [5], which provides a graphical paradigm for creating, linking, and visualizing instances created using an ontology. We created different glyphs to represent the different types of components in cardiovascular hemodynamic models to clarify the distinction among different types of anatomic structures (Figure 4).

The information in the ontology-based model can be used to generate a listing of the equations needed to produce a time-varying simulation of that model. We created a script that iterates through instances of the physiological attribute classes, evaluating the nodes and arcs related to those attributes as well as the type of attribute (canonical or custom) to generate an equation file that can be processed by a separate simulation program. The detailed information needed to create the equations resides in the ontology instances; therefore, the ontology model can be altered and extended without needing to update the application code which

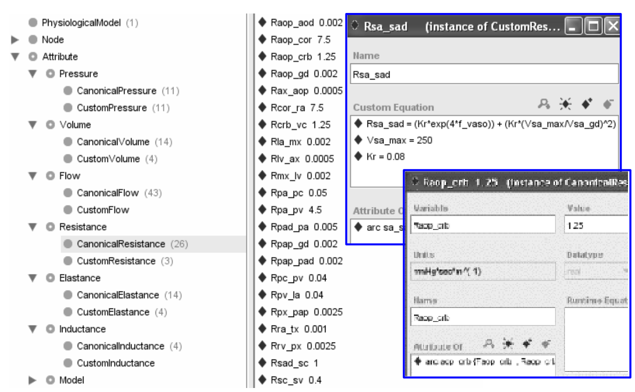

Figure 3. Representing physiological attributes and equations in the ontology. *Equations are represented in terms of physiological attributes, that latter determining the type of general equations (canonical attributes) or hand-crafted custom equations (custom attributes). A custom resistance attribute (upper right) is associated with the connection between systemic arteries and arterioles (arc* sa_sad*), specified by the custom equations shown. A canonical attribute (lower right) is associated with the connection between proximal aorta and cerebral arteries (arc* aop_crb*). This resistance has a value of 1.25 mmHg-sec/ml, and it will be used to create the equations related to the resistance on this connection: Faop_crb = (Paop - Pcrb)/Raop_crb and Raop_crb = 1.25 mmHg-sec/ml.*

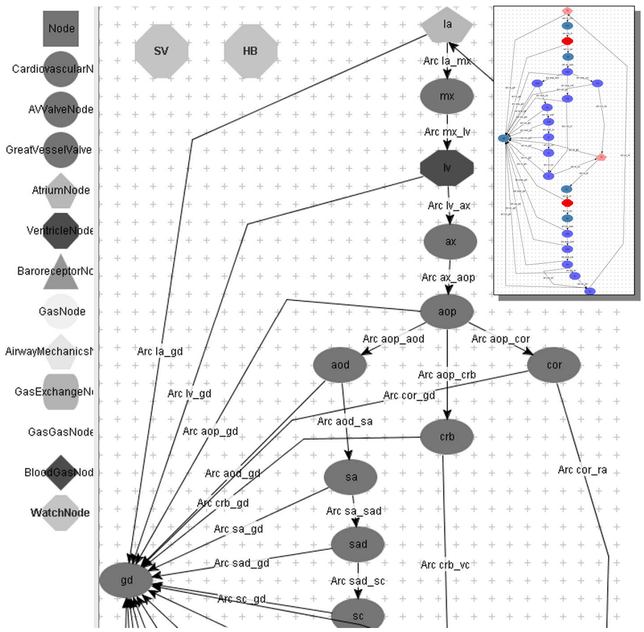

Figure 4. Ontology-based representation of model of cardiovascular hemodynamics. *A portion of the complete model (upper right) is shown. This model, based on that in Figure 1, is represented as a set of linked instances using the Protégé Diagram Widget. The nodes and arcs shown to the user are Protégé instances in the knowledge base. The user can interact with the diagram, adding new anatomic entities, change connections and physiological attributes while simultaneously altering the underlying mathematical model. Physiological attributes are attributed relations on the arcs and can be changed by double-clicking on the arcs.*

creates the mathematical equations from those instances.

To assess the potential benefits of our approach, we interviewed the creator of the original equation-only model and qualitatively compared the development process with that entailed using our methodology.

C. Results

We used our ontology to build a representation of the cardiovascular model shown in Figure 1 by creating ontology instances for the anatomic entities in that model, specifying their connections, and adding physiological attributes to those connections. Anatomic entities were represented in the physiological model as nodes. For example, the Distal Aorta is represented as an instance of CardiovascularNode (Figure 2). Connections between the entities to create physiological continuities were represented by creating instances of the Arc class; for example, the connection between proximal aorta (aop) and cerebral arteries (crb) was established by creating arc aop_crb, and that between the systemic arteries (sa) and arterioles (sad) by creating arc sa_sad. Nodes (anatomic entities) and arcs (connections) were assigned the necessary physiological attributes (resistance, capacitance, etc) as required by the model in Figure 1.

The Protégé Diagram Widget and our ontology enabled us to produce a graphical schematic version of the model similar to that shown in Figure 1 (Figure 4). At the same time, this schematic representation recorded the necessary information required to create an output file of the mathematical equations corresponding to the graphical representation. Our ontology-based model introduced a few additional distinctions among the components of the circulatory system that differed from the original model, such as valves and the structure of the coronary and cerebral circulation; however, except for these differences, the equations among the two models and the simulation results were similar.

While comparing the equations produced by our ontology-based method and those in the original model, we discovered a few discrepancies which subsequently proved to constitute errors in the original model. We also identified two cases where a physiological attribute was missing from the original model. The creator of the original model told us that such errors, while rare, do arise in complex models and that a graphical-based approach such as ours improve the model development process. In addition, the schematic view of the ontology-based model enabled users to interrogate focused pieces of the model and to examine a limited subset of the model equations. Thus, the mathematics could be localized regionally within the model, assisting with comprehension and debugging.

D. Discussion

Complex physiological models generally contain many mathematical equations as well as a graphical schematic representation of the model. The schematic diagram is useful to help users visualize and comprehend the model; however, the schematic representation usually serves only to document the structure of the model (Figure 1). Such graphical representations are usually disconnected from the mathematical equations that provide the quantitative description of the model; if the diagram is changed, the equations need to be updated.

The task of amassing and maintaining the mathematical equations necessary to the schematic form of models is a time-intensive and error-prone effort. Potential errors become more likely as models become large and complex. In addition, extending or changing an existing model is challenging because alterations in the graphical representation do not automatically translate into changes in the mathematical equations.

We represented the information content of a model of cardiovascular hemodynamics in an ontology, providing a structured and explicit framework, enabling simultaneous graphical and mathematical representation of the model. Our implementation in Protégé allows users to create and edit their models in a graphical paradigm, to describe the physiological attributes, and to automatically create the mathematical equations needed to simulate and analyze the model. Our approach permits unification of graphical symbolic representations of physiological models with their mathematical formulation.

Previous work has focused on creating graphical languages for describing biomedical systems [6]. Standard languages for exchanging simulation models have also been created, such as SBML [7]. However, neither of these approaches directly addresses the need to connect the graphical schematic representation to the mathematical formulation of models. Graphical methods have been developed to create kinetic models [8], though in such systems there is no declarative representation of the connection between the graphical and quantitative models. There has also been work in using object-oriented approaches for modeling biological systems; however, we believe the ontological approach is advantageous because ontologies are separate from the implementing software, and they can be browsed by developers and processed by machines. Ontologies are also extensible, enabling interoperability and reuse of information in other ontologies. In fact, our ontology used anatomic knowledge of organ identity contained in the FMA (Figure 2).

Although we have not performed a formal evaluation, we believe our approach has several potential advantages. The ontology can help manage the complexity of physiological models because it localizes the information as instances of particular classes. In our ontology, we could readily identify portions of the model where custom equations were employed, indicating components of the model where there were different mathematical representations. Recognizing and relating particular equations to the corresponding portion of the schematic representation of the model would be difficult without the ontological representation.

Another advantage is that explicit knowledge representation in the ontology makes it possible to detect inconsistencies in models. We identified missing physiological attributes and a few errors in equations in the original physiological model by browsing the ontological representation of the model. Our ontology can link to other ontologies, which could provide knowledge to perform additional validations, such as checking that anatomic entities connected in the model are known to have continuity relations in the FMA. We also believe our approach is extensible and could be applied in representing other types of biomedical models.

There are limitations to our approach. We assume that the knowledge necessary to generate all equations can be localized to the arcs and nodes in the model, but for some models there could be equations needed to describe groups of nodes and arcs or other types of components. We believe the ontology could be extended to handle such cases. Another limitation is that we have not yet developed a debugging facility to ensure that valid combinations of attributes or values are used in the ontology. Debugging knowledge could be added to the ontology in the future.

Our ultimate goal is to expand our ontology-based system to permit researchers to build many different types of biomedical models, giving developers a graphical environment in which to develop and modify their models while dynamically creating and executing the corresponding quantitative simulation equations. This methodology may provide a more robust approach to represent, maintain, and extend large multi-scale simulation models.

E. Acknowledgments

This work was supported by a contract from DARPA, executed by the U.S. Army Medical Research and Materiel Command/TATRC Cooperative Agreement, Contract W81XWH-04-2-0012. This work was also supported by the Protégé resource, under grant LM007885 from the U.S. National Library of Medicine.

F. References

[1] Bassingthwaighte JB. Back to fundamentals: anatomy-based physiological bioengineering. Ann Biomed Eng 2000;28(7):701-3.

[2] Lu K, Clark JW, Jr., Ghorbel FH, Ware DL, Bidani A. A human cardiopulmonary system model applied to the analysis of the Valsalva maneuver. Am J Physiol Heart Circ Physiol 2001;281(6):H2661-79.

[3] Neal M, Bassingthwaighte JB. spheart: A closed loop cardiovascular model composed of a systemic circulation, a pulmonary circulation, and a four-chamber varying-elastance heart model. http://nsr.bioeng.washington.edu/PLN/Members/mneal/integrated_html. Accessed 2/3/05.

[4] Rosse C, Mejino JL, Jr. A reference ontology for biomedical informatics: the Foundational Model of Anatomy. J Biomed Inform 2003;36(6):478-500.

[5] Vendetti J. The Protege Diagram Widget. http://protege.stanford.edu/doc/tutorial/graph_widget/.

[6] Cook DL, Farley JF, Tapscott SJ. A basis for a visual language for describing, archiving and analyzing functional models of complex biological systems. Genome Biol 2001;2(4):RESEARCH0012.

[7] Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics 2003;19(4):524-31.

[8] Cook DL, Atkins WM. Enhanced detoxication due to distributive catalysis and toxic thresholds: a kinetic analysis. Biochemistry 1997;36(36):10801-6.

Address for correspondence

Daniel L. Rubin (dlrubin@stanford.edu), Stanford Medical Informatics, MSOB X-215, Stanford, CA 94305, USA.