Bridging Biological Ontologies and Biosimulation: The Ontology of Physics for Biology

Daniel L. Cook[1,2], Jose L. V. Mejino, Jr.[2], Maxwell L. Neal[3], and John H. Gennari[3]

Departments of [1]Physiology & Biophysics, [2]Biological Structure, and [3]Medical Education & Biomedical Informatics, University of Washington, Seattle, WA

Abstract

We introduce and define the Ontology of Physics for Biology (OPB), a reference ontology of physical principles that bridges the gap between bioinformatics modeling of biological structures and the biosimulation modeling of biological processes. Whereas modeling anatomical entities is relatively well-studied, representing the physics-based semantics of biosimulation and biological processes remains an open research challenge. The OPB bridges this semantic gap—linking the semantics of biosimulation mathematics to structural bio-ontologies. Our design of the OPB is driven both by theory and pragmatics: we have applied systems dynamics theory to build an ontology with pragmatic use for annotating biosimulation models.

The Biophysical Semantics of Biosimulation

The *static* structures of biological organisms have been comprehensively encoded in such bioinformatics resources as the Foundational Model of Anatomy[1], Gene Ontology[2], and ChEBI.[3] The behavior—the *dynamic processes*—of such biological systems have, for the past half-century, been captured in the mathematical language of physics-based biosimulation modeling. Recently, researchers have aimed at building a complete Physiome[4], a flexible integration of component models into large-scale or special-purpose biosimulations for application to clinical and investigatory problems. To date, there have been only rudimentary attempts to bridge the wealth of structural knowledge developed by the bioinformatics community and the process knowledge developed by the biosimulation community.

We aim to bridge this gap, and to link the semantics of biosimulation to the knowledge in structural bio-ontologies. Pragmatically, we have shown in preliminary work that this semantic linkage can help with the task of integrating and linking biosimulation models[5], such as is needed by the Physiome project. The fundamental challenge in carrying out such work is that although biosimulations are based on classical physics and formally expressed in mathematics, the semantics of these models—the meaning of variables and equations—is only implicit in model computational code or, at best, annotated using *ad hoc* in-line code comments. Although current best practices in biosimulation modeling include adherence to some annotation standards[6], integrating models remains a daunting task which is further hampered by conflicting computational languages, differences in implicit assumptions, and pervasive coding errors[6 7].

We show here that the semantics of biosimulation models can be expressed in a formal ontology that describes the entities, the properties, and the physical laws that are encoded in the mathematical equations of a simulation model. Thus, we introduce the Ontology of Physics for Biology (OPB) that is based on systems dynamics and makes explicit the biophysical semantics of physics-based biosimulation models. Capturing these semantics is a critical step toward aligning, integrating, or even de-bugging biosimulation models[5]. Thus, we view the OPB as a pragmatic bridge between the fields of biomedical ontologies and biosimulation: it is a missing link in current biomedical ontologies, and it will be of pragmatic use to the biosimulation research community.

Principles and scope of the OPB

The OPB is based on ontological principles espoused by the Open Biomedical Ontologies (OBO)[8]. The OPB is a reference ontology of biophysics that is orthogonal and complementary to representations of biological structure as exemplified by the Foundational Model of Anatomy Ontology (FMA)[1]. Thus, for example, the OPB represents *volume* and *pressure* (OPB classes will be distinguished by *Arial* font) that are physical properties of anatomical entities but does not represent the anatomical entities themselves nor the structural relations between entities—these are the purview of the anatomy ontologies.

We base the OPB on the principles of physics as described in a variety of resources including textbooks of physics, textbooks of systems dynamics (e.g., [9]), treatises on biological network thermodynamics (e.g., [10]), and existing ontologies of engineering.[11] Within the broad scope of these resources, we have constrained the OPB (at least initially) to representing the biophysics that are pertinent to physics-based biosimulation models such as those included in the Physiome project[4]. Thus, the OPB is multiscale—by

Domain	Example
Fluid	flow of blood or respiratory gasses
Solid	contraction, skeletal movements
Electrical	action potentials, synaptic activity
Chemical	cell metabolism, cell signaling
Diffusion	vesicle diffusion, ion diffusion
Thermal	heat generation and dissipation

Table 1. OPB *Physical domains*

which we mean that it encompasses multiple structural scales such as represented in the FMA (organs, organ parts, cells, molecules, etc.)—and is *multidomain*—the OPB represents physical properties and laws in each of six *Physical domains* (Table 1).

We have implemented an initial version of the OPB within the Protégé ontology editing environment[12], and although we expect the OPB to continue to evolve, we do not expect massive additions. The OPB is a representational framework for the principles and laws of classical physics as they are applied in biosimulation modeling. As such, OPB classes serve as reference concepts by which the computational variables and equations may be annotated (see[5], for example). Thus, although aortic blood pressure may be represented in a biosimulation system by a model variable "Paorta", the OPB will simply contains the class *Fluid pressure*, and this class does not have subclasses representing, for example, the pressures of specific portions of blood (e.g., pressures of the aorta or vena cava). Thus, whereas the goal of the FMA has been to represent all the parts of the canonical human body, the OPB is a much smaller ontology whose sparseness is attributed to the concision of physical theory.

We believe that adherence to the above ontological and physical principles is essential for achieving our pragmatic goal: that the OPB will serve as a utilitarian, reusable resource of biophysical knowledge in service to the biosimulation research community.

OPB representational schema

Our goal is to use the OPB as a reference knowledge resource for annotating variables and equations of models and for deriving computable modeling code. Consequently, we have adopted a simple representational schema (Fig 1) as is used implicitly by practitioners of physics-based biosimulation. In this view, *physical entities* (e.g., a bone, a portion of blood, an electrical field) have (*has_property* relation) *physical properties* (e.g., force, volume, electrical potential). The laws of physics, such as Ohm's law or Newton's law, are represented by *physical dependencies* that have as *players* (via the *has_player* relation) particular *physical properties*. This physics-based view is represented in the OPB class hierarchy with top-level classes *Physical entity*, *Physical property*, and *Physical dependency*.

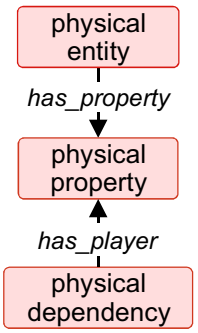

Fig 1. OPB Schema

Because biosimulation models treat biological (both normal and pathological) and non-biological entities in the same physics-based terms, we broadly define *Physical entity* as an "entity in the world that has spatial extent and possesses thermodynamic energy". Thus, OPB represents entities that have mass but are not strictly biological (e.g., air, water, proton) as well as mass-less energetic entities such as gravitational or electrical fields. In view of existing ontologies of human anatomy (e.g., the FMA), genes and gene products[2], and small molecules of biologic interest[3], the OPB will not subsume these ontologies, but rather, provide links to them.

The OPB includes a fourth top-level class, *Physical domain*, whose subclasses correspond to the six domains shown in Table 1, as well as a *Spatial domain* for represent geometrical principles. Subclasses of *Physical property* and of *Physical dependency* take one of these *Physical domain* classes as a slot value according to physical principles (e.g., *Fluid pressure* is assigned to the *Fluid domain*).

Physical properties :: biosimulation variables

In physics-based biosimulation modeling, mathematical variables represent the magnitudes of physical properties and how such property magnitudes vary as a function of time. Various cardiovascular models, for example, calculate the volume flow rate of blood in the aorta yet use different variable names (e.g., "Faorta", "flowA") as well as different units of pressure (e.g., liters/min, ml/sec). Despite such differences in mathematical and code representation each model refers to a single physical concept; the flow of blood in the aorta. Although in-line code comments could point to a common reference such as the FMA class *Blood in aorta*, aside from a few specialized nomenclatures, there is no reference informatics resources to which "volume flow rate" and other physical properties may be mapped. Therefore, a key goal of the OPB is to represent and organize physical properties as a class subsumption hierarchy based on the principles of systems dynamics within the scope of the OPB *Physical domains*[9-11].

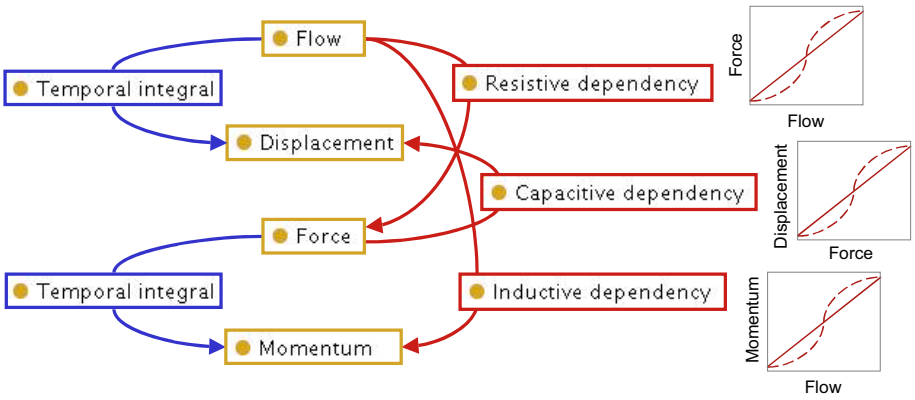

Fig 3. OPB *Physical properties* are related by *Physical dependency* subclasses: *Temporal integral dependency* (left) and three kinds of *Constitutive dependency* (right).

The *Physical property* taxonomy is organized according to mathematical analogies, familiar to physicists and engineers, about how the properties of a given *Physical entity* and *Physical domain* depend on one another. Figure 2 (a Protégé screenshot) partially expands the OPB *Physical property* taxonomy to show several subclasses; here we discuss only *Kinetic physical property* subclasses. For example, *Rate property* and *State property* have, as respective subclasses, *Flow* and *Force*, and *Displacement* and *Momentum*. Furthermore, each of these classes has a subclass for each *Physical domain* as expanded for *Flow* and *Displacement* in Figure 2. Thus, a physical entity such as a portion of blood in the aorta, as an entity of the *Fluid domain*, will have four properties: *Fluid flow, Fluid pressure, Fluid volume,* and *Fluid momentum.*

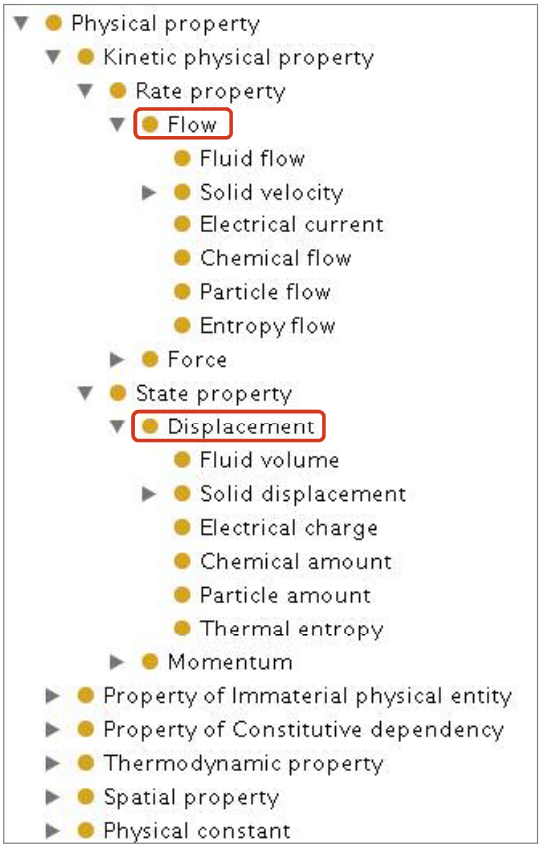

Fig 2. *Physical properties* such as *Flow* and *Displacement* have subclasses for each *Physical domain*.

Physical dependencies :: biosimulation equations

The mathematical relations between model variables are the essence of biosimulation models. Whereas, two models may declare both the pressure and volume as properties of aortic blood, only one model may represent how aortic blood pressure depends on blood volume; i.e., one model represents the distensibility of the aorta while the other treats the aorta as a rigid tube. In terms of systems dynamics, the former model includes a "constitutive capacitive" relation between pressure and volume, whereas the latter model does not. Constitutive and other physical relations are critical model-building choices and are the theoretical basis for the equations in physics-based models.

The great strength of the systems dynamics approach is that it posits only a few fundamental physical relationships between physical properties that are based on widely recognized mathematical analogies that apply irrespective of physical domain. For example, Ohm's Law, originally described for electrical current flow has a direct analogue as Ohm's Law for fluid flow. Thus, the OPB *Physical dependency* taxonomy includes four subclasses: *Temporal integral dependency,* Constitutive dependency, Summation

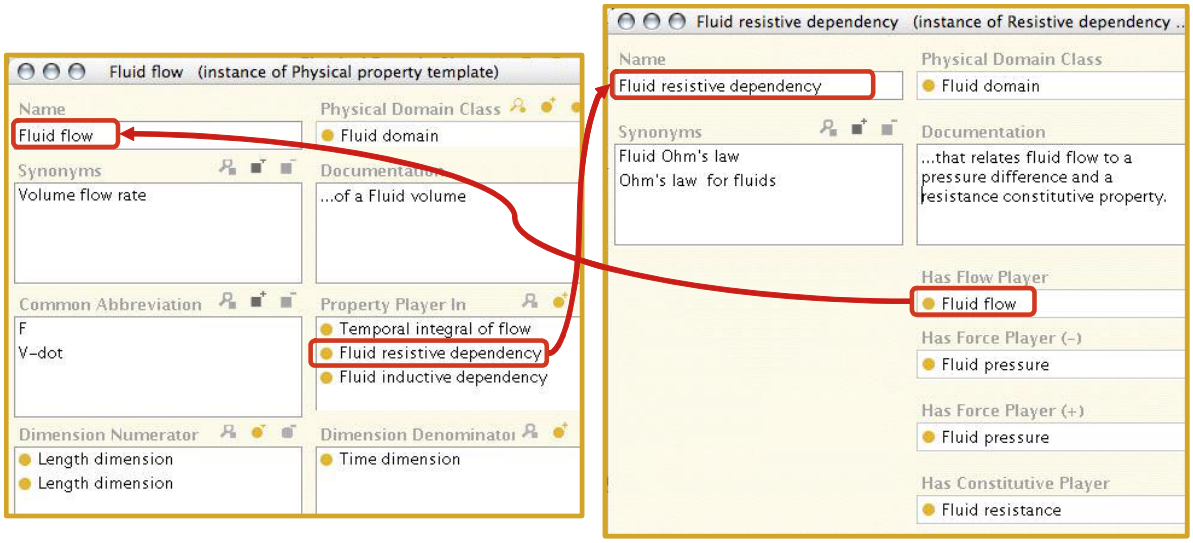

Fig 4. Slot relations between *Fluid flow* and *Fluid resistive dependency*.

dependency, and Thermodynamic dependency of which we will discuss only the first three.

Figure 3 illustrates the systems dynamics schema by which the basic *Physical property* classes are related to *Temporal integral dependency* and *Constitutive dependency* classes within a single *Physical domain*. For example, in the fluid domain, the *Temporal integral dependency* encodes that the net *Fluid flow* of blood over a span of time changes the *Volume* of the portion of blood into which it flows. Less intuitive, but no less true, is the analogous relation by which the temporal integral of a force (e.g., *Fluid pressure*) results in a change of momentum (e.g., *Fluid pressure momentum*). Whereas *Temporal integral dependency* represents a theorem of physics (i.e., is true by definition), constitutive relations such as Ohm's law or Hooke's law are derived empirically and depend on both the material composition (e.g., density, viscosity) and configuration (e.g., shape, size) of the physical entity to which they apply.

By analogy with electrical circuit elements (resistor, capacitor, inductor), OPB *Constitutive dependency* has three subclasses: *Resistive dependency, Capacitive dependency,* and *Inductive dependency*. Figure 3 shows diagrammatically how each of the four *Physical properties* are related by these *Constitutive dependencies*. *In the Fluid domain, Fluid resistive dependency* is an analogue of Ohm's law, *Fluid capacitive dependency* relates *Fluid pressure* and *Fluid volume* according to the elastance of a vessel containing the fluid, and *Fluid inductive dependency* represents how a *Fluid pressure* differences change *Fluid momentum*.

The insets on the right of Figure 3 illustrate that *Constitutive property dependencies* may be linear or nonlinear depending on the specific physical composition and constitution of the physical entities to which they apply. In engineered systems, these are often designed to be linear, and thus mathematically more tractable. Unfortunately, non-linearity is the hallmark of biological systems, providing considerable computational challenge. The OPB encodes constitutive relations as classes irrespective of the specific mathematical functions (e.g., linear, hyperbolic, polynomial) that specific simulations use to approximate their shapes.

In addition to *Constitutive dependencies* that apply to a single entity and *Physical domain* (e.g., blood flow in the *Fluid domain*), the OPB also includes *Transformation dependency* and *Transduction dependency* classes that represent how energy can flow between two entities of the same or different domains (e.g., how ventricular wall contraction increases ventricular blood pressure). Finally, the OBP includes *Summation dependency* classes that represent physical conservation (e.g., conservation of charge) and multiscale summation laws. These laws are required for multiscale models whereby, for example, the mass of a heart is the sum of the masses of all of its parts.

Relations between properties and dependencies

Whereas taxonomies of classes provide unique identifiers for annotating and identifying the biophysical meaning of code variables and equations, additional biophysical knowledge is required to check, correct and encode biosimulation models. Here we describe the OPB's representation of such knowledge as frames for each *Physical property* and *Physical property dependency* and the slots by which relations between frames are encoded.

Each *Physical property* frame is built on a template that includes slots for the property's name, physical domain, synonyms, definition ("documentation"), common abbreviations, specification of the property's physical dimensions (e.g., length^2/time for a

volume flow rate), and a slot that lists each of the *Physical dependencies* in which the property is a role player. For the example shown in Figure 4, *Fluid flow* is a role player in three *Physical dependencies* (as so far encoded). Each *Physical dependency* frame encodes the same knowledge (synonyms, etc.) as a *Physical property* frame but encodes a specific relation to each *Physical property* that is a role player in the dependency. On the right side of Figure 4, the *Fluid resistive dependency* (analogous to Ohm's law for fluids) includes slots for specifying a *Fluid flow*, a positive *Fluid pressure*, a negative *Fluid pressure* and a *Fluid resistance* parameter.

Application evaluation

As a partial test of the utility of the OPB, we created a set of semantic annotations based on the OPB, and stored as a Protégé ontology, of the variables (properties) and equations (property dependencies) of three biosimulation models of cardiovascular regulation[5]. These models are fully described in our prior publication[5], but in brief, they model: 1) the cardiovascular system consisting of a beating heart, and arterial and venous vessels, 2) the baroreceptor reflex by which aortic blood pressure controls heart rate, and 3) arterioles whose resistance depends on calcium uptake of arteriolar smooth muscle cells. We captured the semantics of each model variable as a duple combining an *OPB>Physical property* with the *FMA>Anatomical entity* to which it applies. Thus the variable "Paorta" in one model and the variable "Paor" in another model can be found to apply to the same physical entity represented by the duple {*OPB>Fluid pressure: FMA>Blood of aorta*}. We claim that such semantic annotations will assist with aligning, integrating, and even de-bugging complex biosimulation models[5].

By such identifications and distinctions, we were able to merge the semantic maps of the three source models to produce a merged semantic model that we re-encoded (by hand) into an executable model that is multiscale (subcellular processes to organ system processes), and multidomain (neural reflexes, fluid dynamics, cell signaling). Thus, the merged model was able to reproduce the expected effects of increased arteriolar calcium uptake to increase blood pressure and to decrease heart rate—effects that cannot be reproduced in the separate source models.

Discussion and status

We have introduced the OPB as a reference ontology of the properties and principles of classical physics that are the foundations of biosimulation modeling. By applying a simple representational schema (Figure 1) across multiple biologically relevant physical domains (Table 1), we leverage the concepts of systems dynamics as have been articulated (separately) in the engineering[9, 11] and biological[10] sciences.

To date, we have established a stable ontological framework that is based both on the principles of biophysical systems dynamics and of ontology design. Although the OPB is by no means complete, we have within the *Fluid domain* demonstrated that the OBP can capture semantics in sufficient detail to merge and re-encode models into a composite, multiscale bio-simulation[5]. As the systems dynamics approach has great generality for physics-based systems, we believe that the OBP reference ontology will be of pragmatic use for a wide variety of biosimulation models.

Acknowledgements

This work was partially funded by NIH grant #R01 HL087706-01. We thank Drs. Cornelius Rosse, Fred Bookstein and Adriana Emmi for critical input to the OPB.

References

1. Rosse C, Mejino JLV. A Reference Ontology for Bioinformatics: The Foundational Model of Anatomy. *Journal of Biomedical Informatics* 2003;36:478-500.
2. Gene Ontology, http://www.geneontology.org/, 2005.
3. Chemical Entities of Biological Interest (ChEBI), http://www.ebi.ac.uk/chebi/, 2006.
4. Hunter PJ, Borg TK. Integration from proteins to organs: the Physiome Project. *Nat Rev Mol Cell Biol* 2003;4(3):237-43.
5. Gennari JH, Neal ML, Carlson BE, Cook DL. Integration of multi-scale biosimulation models via light-weight semantics. *Pac Symp Biocomput* 2008;13:414 – 425.
6. Le Novere N, Finney A, Hucka M, Bhalla US, Campagne F, Collado-Vides J, et al. Minimum information requested in the annotation of biochemical models (MIRIAM). *Nat Biotechnol* 2005;23(12):1509-15.
7. Alves R, Antunes F, Salvador A. Tools for kinetic modeling of biochemical networks. *Nat Biotechnol* 2006;24(6):667-72.
8. OBO Foundry: A New Paradigm for Biomedical Ontology Development. http://obofoundry.org/, 2006.
9. Karnopp D, Margolis DL, Rosenberg RC. *System dynamics: a unified approach*. 2nd ed. New York: Wiley, 1990.
10. Oster GF, Perelson AS, Katchalsky A. Network thermodynamics: dynamic modelling of biophysical systems. *Q Rev Biophys* 1973;6(1):1-134.
11. Borst P, Akkermans H, Top J. Engineering ontologies. *Int. J. Human–Computer Studies* 1997;46:365-406.
12. Gennari JH, Musen MA, Ferguson RW, Grosso WE, Crubezy M, Eriksson H, et al. The evolution of Protege: an environment for knowledge-based systems development. *Int. J. Human–Computer Studies* 2003;58:89-123.