What should be the central purpose for Working Group 10?

Jim Bassingthwaighte

Frank T. Bergmann

Herbert M. Sauro

Topics:

Introduction: JBB
Review of Model Archiving Systems: Frank, Herb, Jim
Discussion: All

Reproducibility in Science
(A proposal for a central mission for WG10)

Adherence to principles:

- Data gathering with hypotheses to test
- Platt's (1964) alternative hypotheses in experimentation
- Completeness in reporting

Standardization:

- Ontologies
- Standards for quantitative modeling
- Journal publication standards and peer review
- Full disclosure of thinking, failures, successes
- Unfettered dissemination
- Standards for curation levels

Building the repositories

- Completeness in reporting (scientific literature)
- Databases of biological information (gene,protein, … physiological time series and images (e.g. PhysioNet generalized)
- Databases of working or even untested hypotheses: Models.

These are not new ideas!

"Attempts are being made to establish a standard form of what a model should be like, how it should be described and documented, as well as to find out how efficient different computers are in computing its behavior. This is intended in part to facilitate communication of information about models, which may be difficult owing to their complexity. In building a model one has to keep track of why one made the decisions one did, what mistakes were made, what was learned along the way, etc. It may be desirable in describing a model to also include its unsuccessful predecessors, and the reasons for their failure to fit the data or other inadequacies."

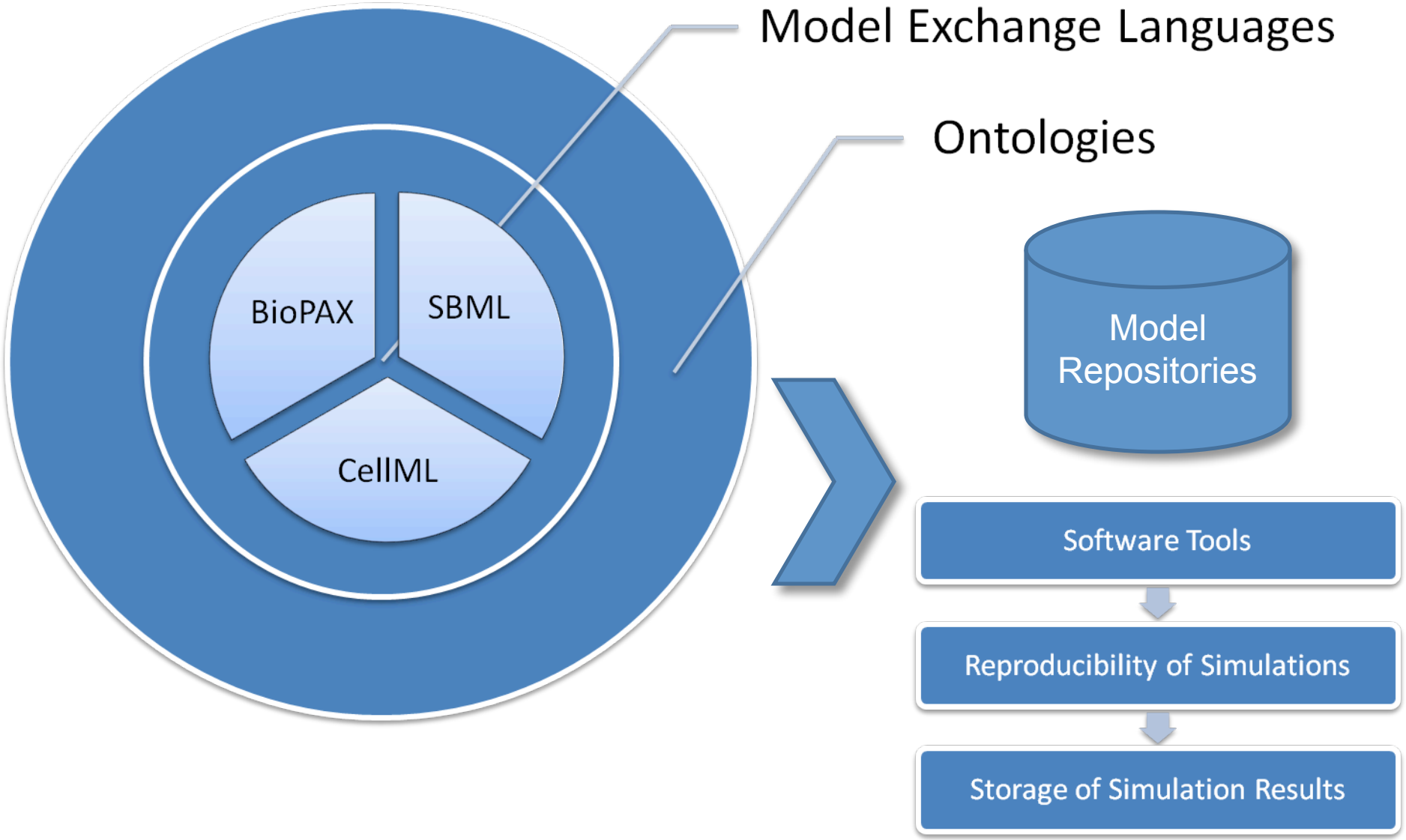
-D. Garfinkel. FEBS Letters 2:S9-S13,1969.

Standards & Ontologies for biochemical / biophysical modeling

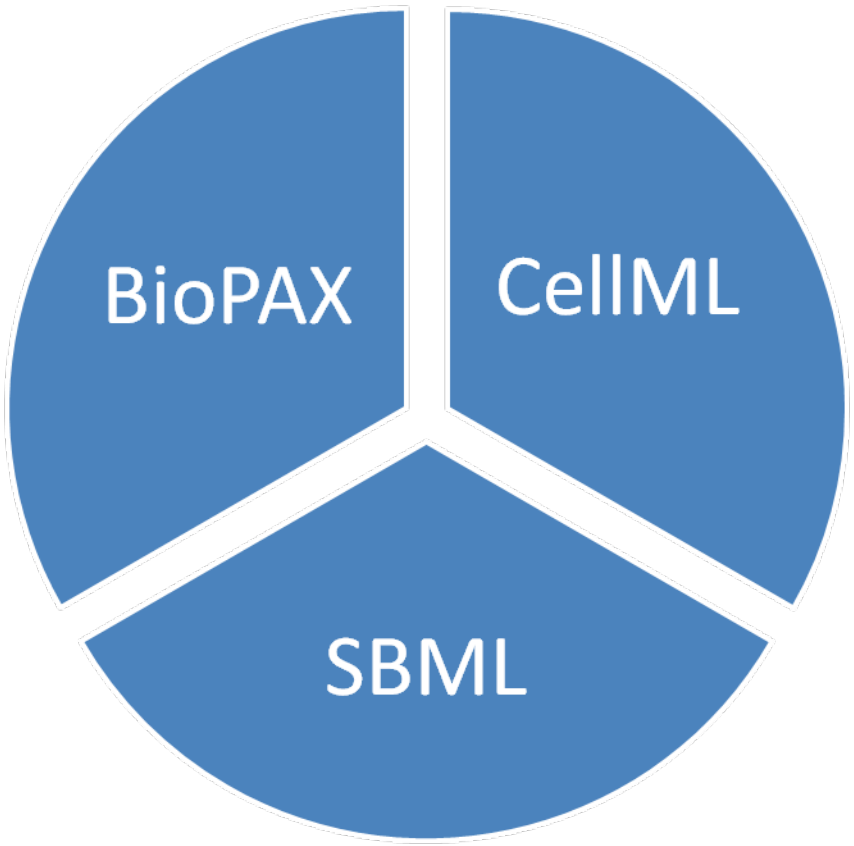
Frank T. Bergmann

Herbert M. Sauro

Overview


Model Exchange Languages
Ontologies
BioPAX
SBML
CellML
Model Repositories
Software Tools
Reproducibility of Simulations
Storage of Simulation Results

Model Exchange Standards



Cell Markup Language

- Designed to support the definition and sharing of models of biological processes.
- Intended to provide consistency in the mathematical representation.
- Encourages model evolution and reuse.
- Started 1999, around same time as SBML.
- CellML and SBML have different scopes:
  - "SBML is designed for representing models of biochemical reaction networks". *(http://www.sbml.org/)*
  - "The purpose of CellML is to store and exchange computer-based mathematical models". *(http://www.cellml.org/)*

[Peter Hunter, *Distributing and maintaining models in CellML,* Life Sciences 2008, Montreal, August, 2008]

Cell Markup Language

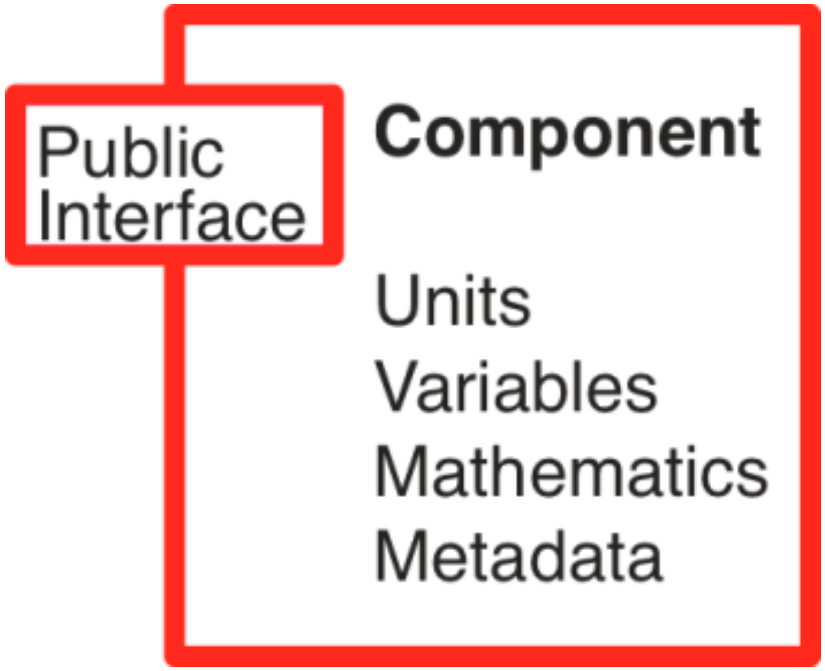
- CellML includes information about:
  - Model structure (how the parts of a model are organizationally related to one another);
  - Mathematics (equations describing the underlying biological processes);
  - Metadata (additional information about the model that allows scientists to search for specific models or model components in a database or other repository).
- CellML includes mathematics and metadata by leveraging existing XML-based languages, such as Content MathML, XML Linking Language (XLink), and Resource Description Framework (RDF).

*(C. M. Lloyd, M. D. B. Halstead, and P. F. Nielsen, "**CellML: its future, present and past**" Progress in Biophysics & Molecular Biology, vol. 85, pp. 433-450, June-July 2004)*

[Peter Hunter, *Distributing and maintaining models in CellML,* Life Sciences 2008, Montreal, August, 2008]

CellML components

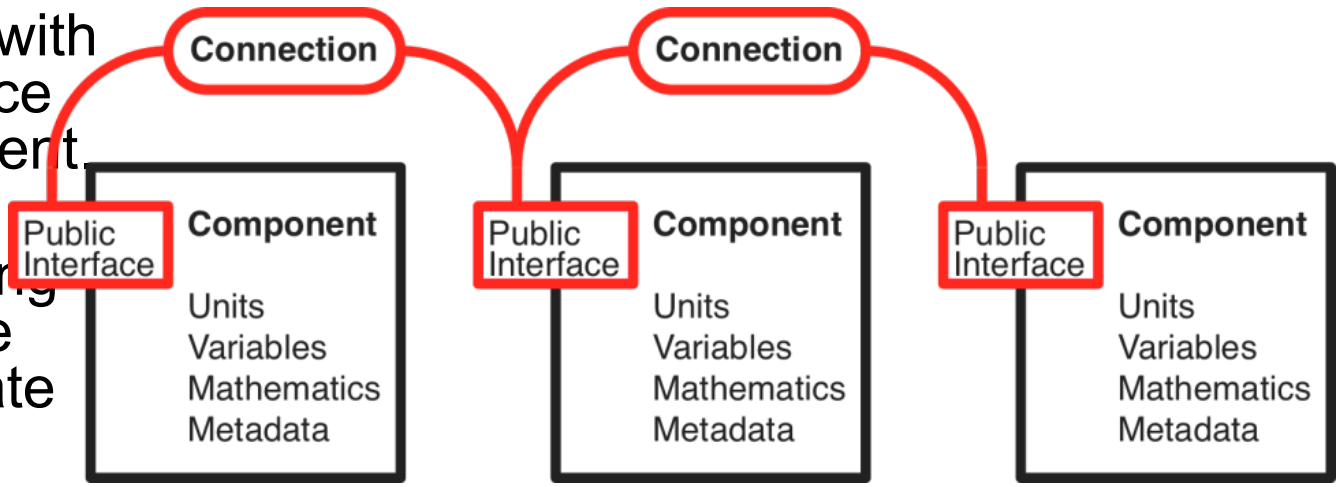
- CellML has a simple structure based upon connected *components*.
- Components abstract concepts by providing well-defined interfaces to other components.
- Components encapsulate concepts by hiding details from other components.



[Peter Hunter, *Distributing and maintaining models in CellML,* Life Sciences 2008, Montreal, August, 2008]

CellML connections

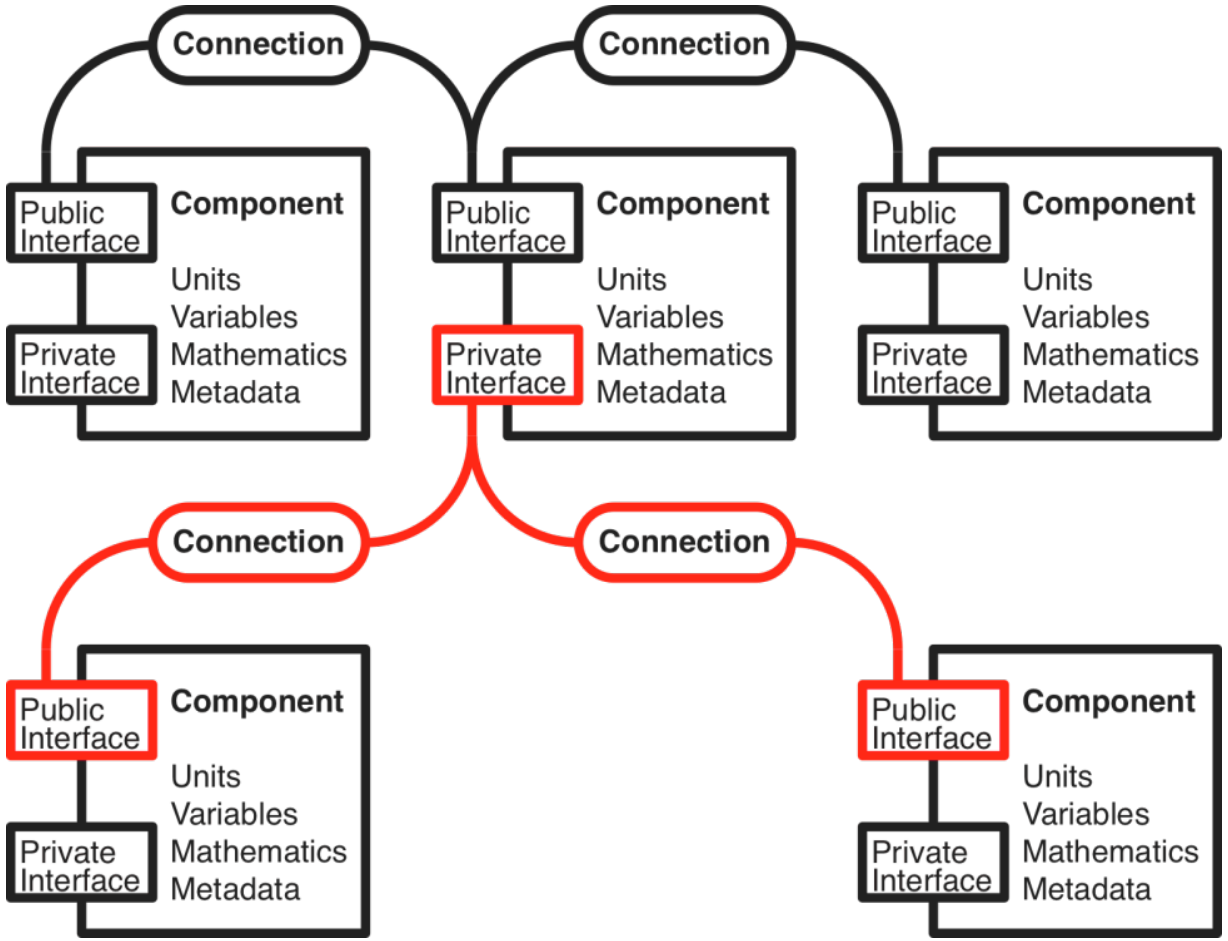
- Connections provide the means for sharing information by associating variables visible in the interface of one component with those in the interface of another component.
- Consistency is enforced by requiring that all variables be assigned appropriate physical units.



[Peter Hunter, *Distributing and maintaining models in CellML,* Life Sciences 2008, Montreal, August, 2008]

CellML encapsulation

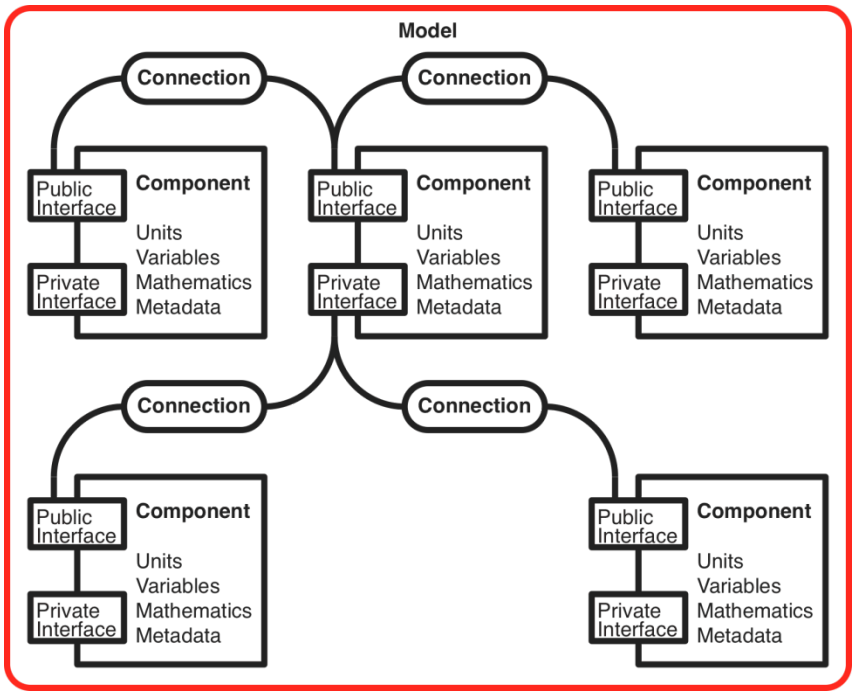
- Encapsulation hierarchies are enabled using *private interfaces*.



[Peter Hunter, *Distributing and maintaining models in CellML,* Life Sciences 2008, Montreal, August, 2008]

CellML model

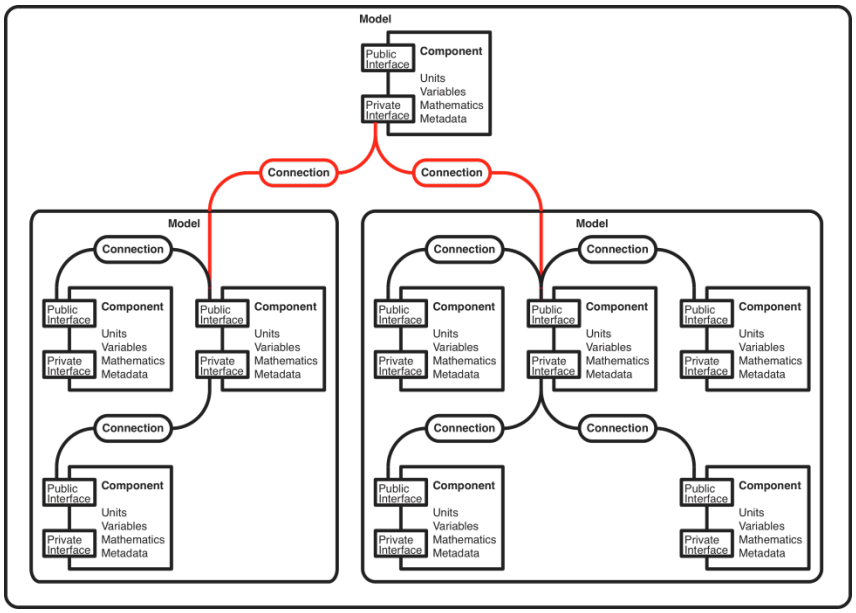
- A *model* is the root element for a CellML document. It is a container for components, connections, units, and metadata.



[Peter Hunter, *Distributing and maintaining models in CellML,* Life Sciences 2008, Montreal, August, 2008]

CellML import

- Model reuse is enabled by the *import* element.
- New models may thus be constructed by combining existing models into model hierarchies.



[Peter Hunter, *Distributing and maintaining models in CellML,* Life Sciences 2008, Montreal, August, 2008]

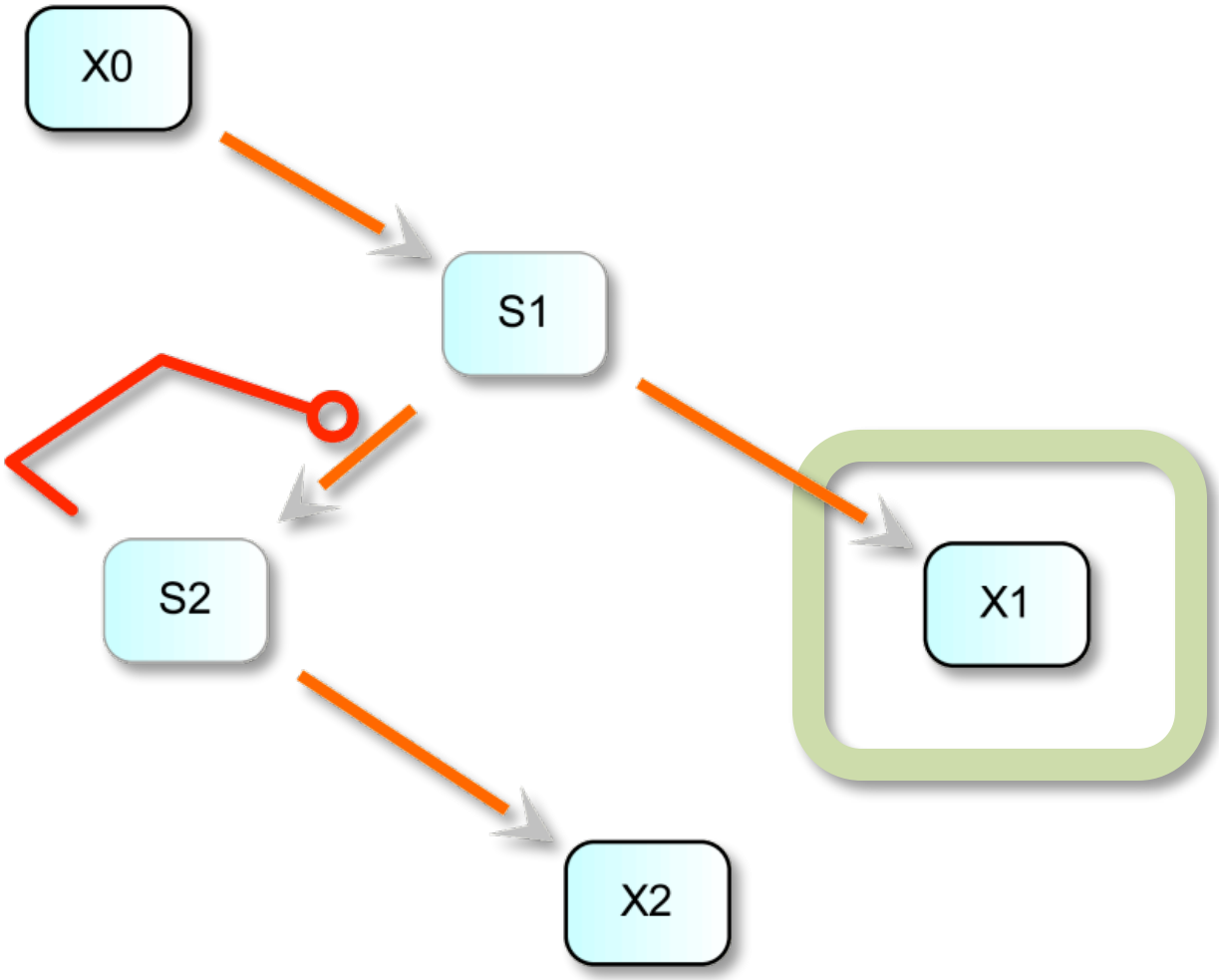
Systems Biology Markup Language

- The **Systems Biology Markup Language** (**SBML**) is a computer-readable format for representing models of biochemical reaction networks in software.
- SBML developed in stages or ***Levels***
  - *Level 1:* mostly basic compartmental modeling
  - *Level 2:* new features (but more complexity), such as:
    - MathML instead of text strings for math expressions
    - Support for user-defined functions
    - Support for conditional events
  - *Level 3:* under development; expect modular support for
    - Multistate species
    - Model composition (submodels)
    - Diagrams
    - Spatial features

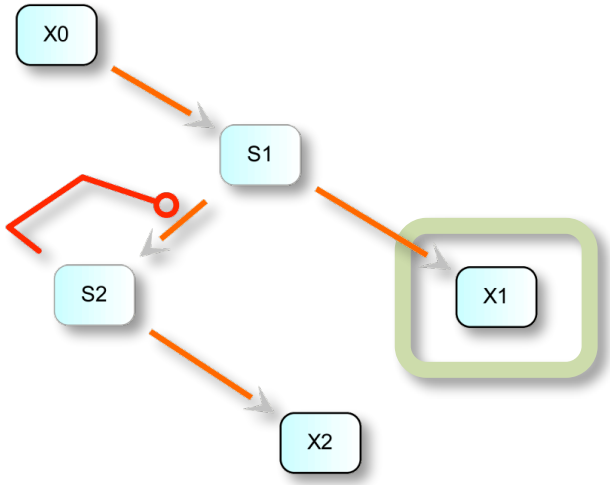
Structure of SBML

- Beginning of SBML model definition
  - List of function definitions
  - List of unit definitions
  - List of compartment types
  - List of molecular species types
  - List of compartments
  - List of species
  - List of parameters
  - List of initial assignments
  - List of rules
  - List of constraints
  - List of reactions
  - List of events
- End of SBML model definition

Systems Biology Markup Language



Systems Biology Markup Language



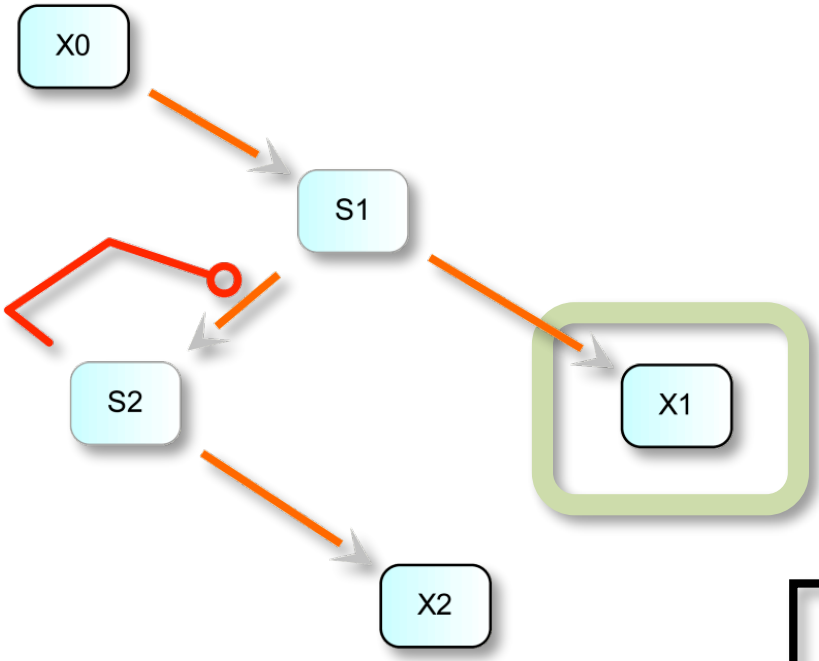
$$\$X0 \rightarrow S1;\ v_0;$$

$$S1 \rightarrow \$X1;\ k3 \cdot S1;$$

$$S1 \rightarrow S2;\ (k_1 \cdot S1 - k_{-1} \cdot S2) \cdot (1 + c \cdot S2^q);$$

$$S2 \rightarrow \$X2;\ k_2 \cdot S2;$$

Systems Biology Markup Language



$$\$X0 \rightarrow S1;\ v_0;$$

$$S1 \rightarrow \$X1;\ k3 \cdot S1;$$

$$S1 \rightarrow S2;\ (k_1 \cdot S1 - k_{-1} \cdot S2) \cdot (1 + c \cdot S2^q);$$

$$S2 \rightarrow \$X2;\ k_2 \cdot S2;$$

```
<?xml version="1.0" encoding="UTF-8"?>
<sbml xmlns = "http://www.sbml.org/sbml/level2" level = "2"
   <model id = "oscli" name = "Oscli (Heinrich model)">
      <listOfCompartments>
         <compartment id = "compartment" size = "1"/>
      </listOfCompartments>
      <listOfSpecies>
         <species id = "S1" initialConcentration = "0" comp
         <species id = "S2" initialConcentration = "1" comp
         <species id = "X0" boundaryCondition = "true" init
         <species id = "X1" boundaryCondition = "true" init
         <species id = "X2" boundaryCondition = "true" init
      </listOfSpecies>
      <listOfParameters>
         <parameter id = "J0_v0" value = "8"/>
         <parameter id = "J1_k3" value = "0"/>
         <parameter id = "J2_k1" value = "1"/>
         <parameter id = "J2_k_1" value = "0"/>
         <parameter id = "J2_c" value = "1"/>
         <parameter id = "J2_q" value = "3"/>
         <parameter id = "J3_k2" value = "5"/>
      </listOfParameters>
      <listOfReactions>
         <reaction id = "J0" reversible = "false">
            <listOfReactants>
               <speciesReference species = "X0" stoichiometry
            </listOfReactants>
            <listOfProducts>
               <speciesReference species = "S1" stoichiometry
            </listOfProducts>
            <kineticLaw>
               <math xmlns = "http://www.w3.org/1998/Math/M
                  <ci> J0_v0 </ci>
               </math>
            </kineticLaw>
         </reaction>
[ . . . ]
```

Systems Biology Markup Language

- Broad Acceptance
  - Supported by over 100 software systems
    - Simulators
    - Databases
    - Analysis tools
    - Editing tools
  - Supported by several alliances

    DARPA Bio-SPICE, IECA, others
  - Supported by journals

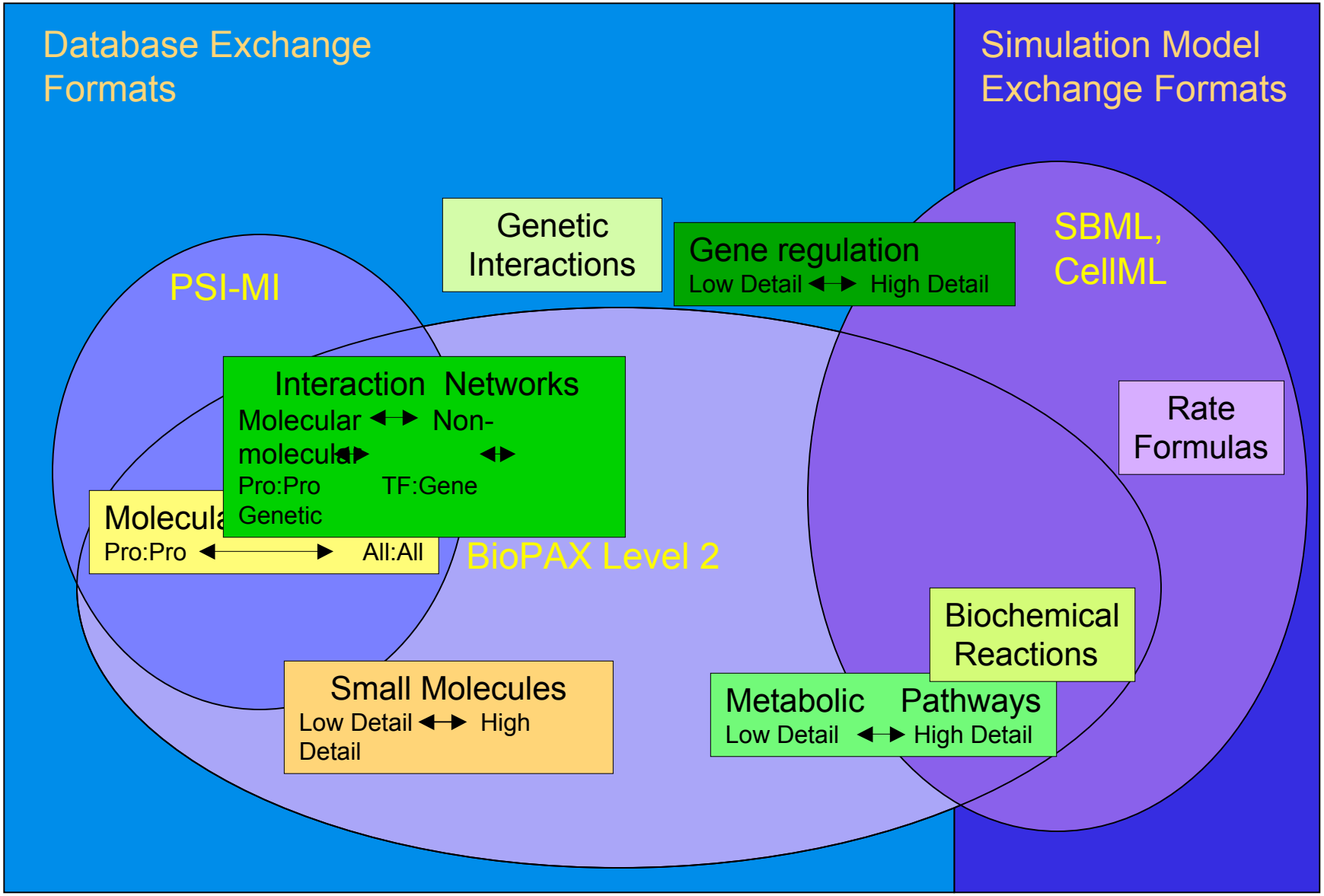
    "*Nature* journals and *Molecular Systems Biology* support submissions involving SBML." [*Nature*, p.1, May 5, 2005]

BioPAX
Biological Pathway Exchange Language

- BioPAX = Biological Pathway Exchange
- Data exchange format for pathway data
- Include support for these pathway types:
  - Metabolic pathways
  - Signaling pathways
  - Protein-protein, molecular interactions
  - Gene regulatory pathways
  - Genetic interactions
- Accommodate representations used in existing databases such as BioCyc, BIND, WIT, aMAZE, KEGG, Reactome, etc.

[Gary Bader, *Proposing a Standard Biological Pathway Exchange Format for caBIG*, BioPAX Wiki,2005]

Exchange Formats in the Pathway Data Space



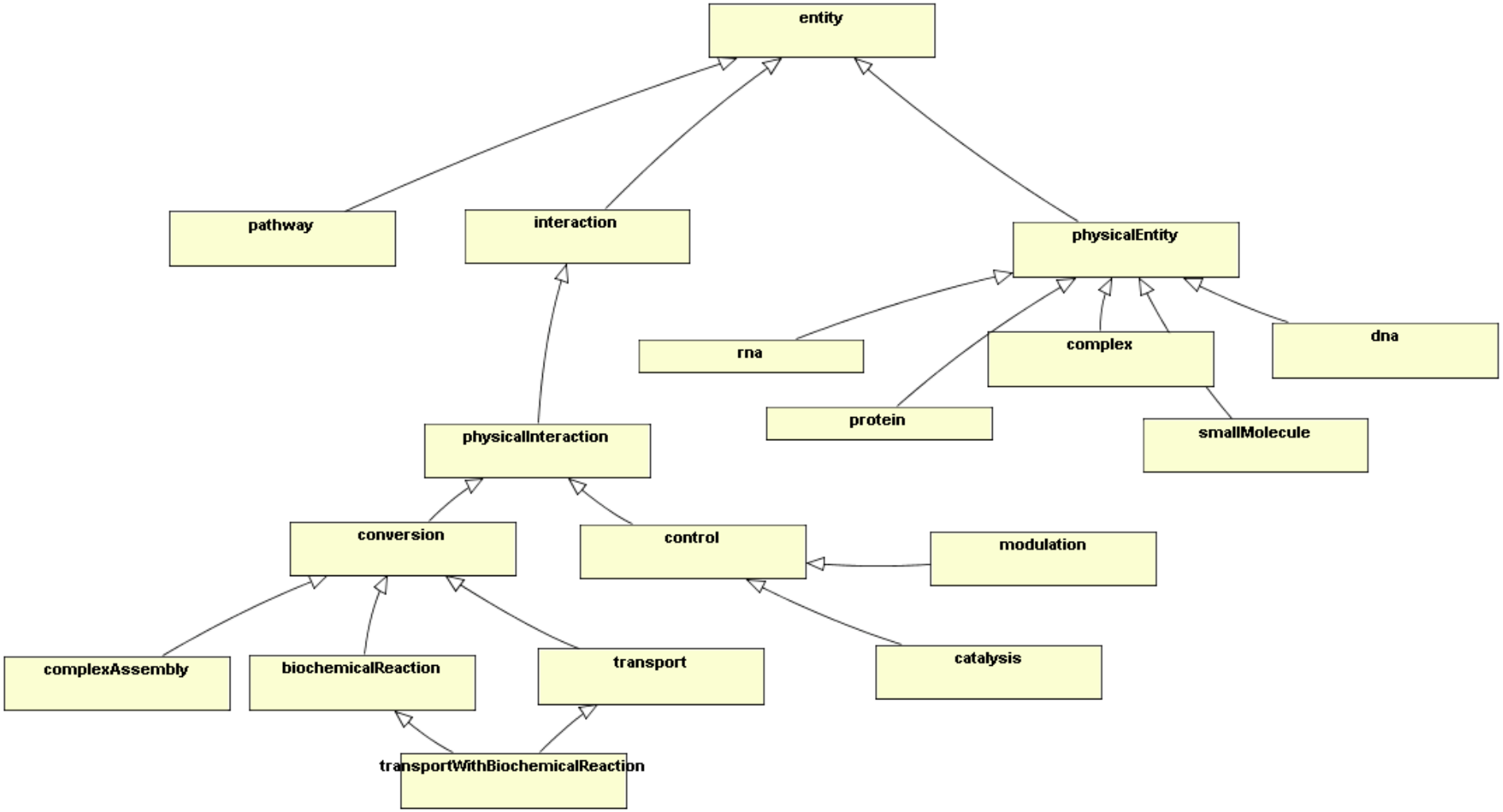
[Gary Bader, *Proposing a Standard Biological Pathway Exchange Format for caBIG*, BioPAX Wiki,2005]

BioPAX

- Conceptual framework based upon existing DB schemas:
  - aMAZE, BIND, BioCyc, WIT, KEGG, Reactome, PATIKA, iNOH, TRANSPATH etc.
  - Allows wide range of detail, multiple levels of abstraction
- BioPAX ontology and format in OWL (XML)
- Ontology built using GKB Editor and Protégé
- Semantic mapping still an issue
- Level 1 represents metabolic pathway data
- Level 2 adds support for molecular interactions, post-translational modifications, experimental description from PSI-MI model (Backwards compatible)

[Gary Bader, *Proposing a Standard Biological Pathway Exchange Format for caBIG*, BioPAX Wiki,2005]

BioPAX Ontology: Overview



Level 2 (May, 2005)

[Gary Bader, *Proposing a Standard Biological Pathway Exchange Format for caBIG*, BioPAX Wiki,2005]

Future Directions

- Level 3
  - Molecular states
  - Gene regulation
- Level 4
  - Genetic interactions
  - Generic entities
- Tool support: VisANT, PATIKA, caBIG

[Gary Bader, *Proposing a Standard Biological Pathway Exchange Format for caBIG*, BioPAX Wiki,2005]

Ontologies

- MIRIAM – Minimal Information Requested In the Annotation of biochemical Models
- SBO – Systems Biology Ontology
- TEDDY - TErminology for the Description of Dynamics

- In Conception / Planning
  - KiSAO - Kinetic Simulation Algorithm Ontology

MIRIAM

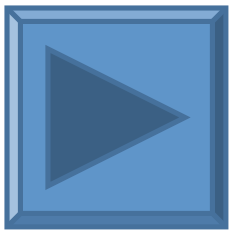
```
[...]
<species

          id="L_EGFR"
          compartment="compartment"
          initialConcentration="0">
</species>
[...]
```

MIRIAM

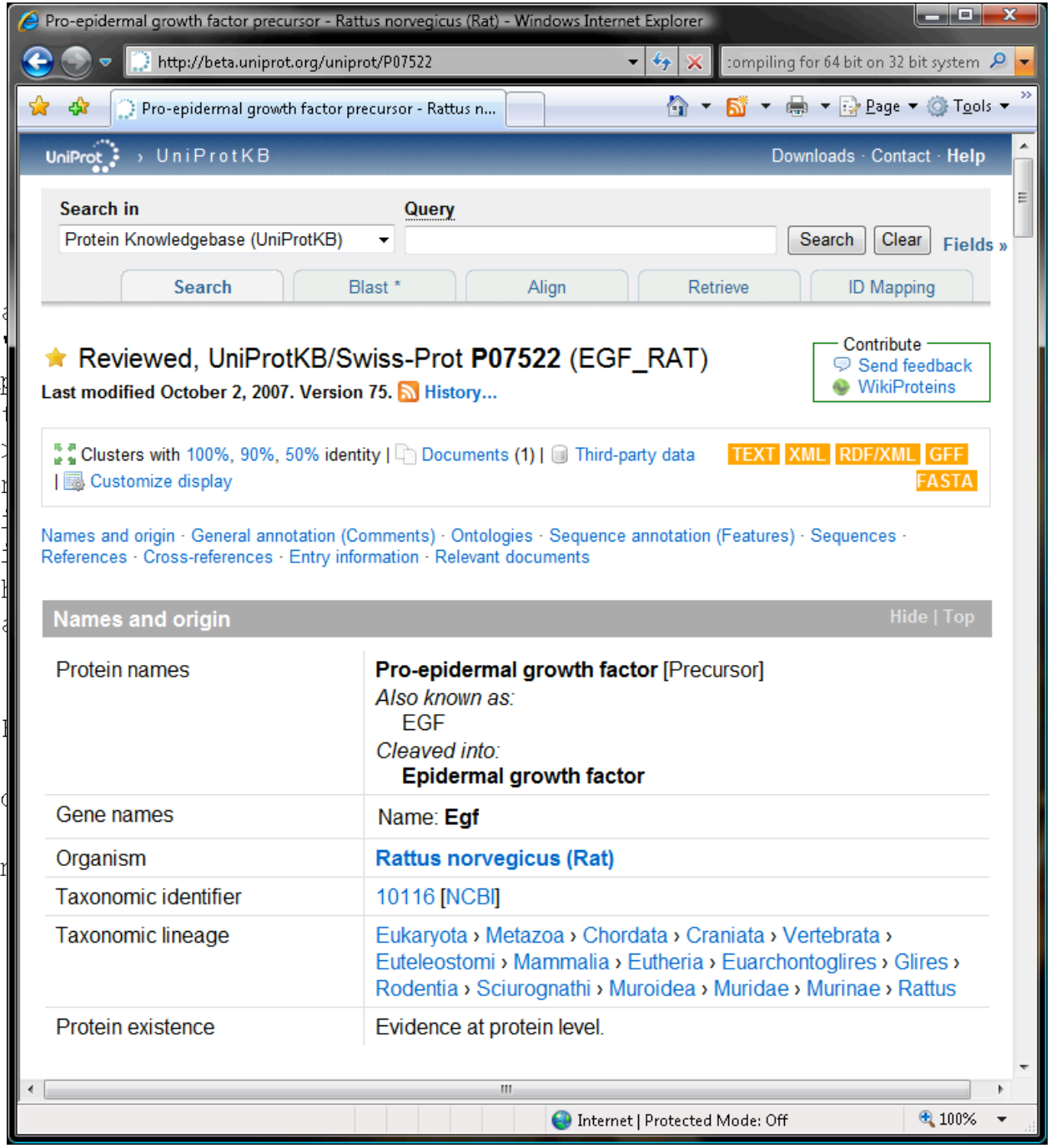
```
[...]
<species metaid="metaid_0000006"
         id="L_EGFR"
         compartment="compartment"
         initialConcentration="0">

 <annotation>
  <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22rdfsyntaxns#"
   xmlns:bqbiol="http://biomodels.net/biologyqualifiers/">
  <rdf:Description rdf:about="#metaid_0000006" >
  </rdf:Description>
  </rdf:RDF>
 </annotation>
</species>

[...]
```

MIRIAM

```
[...]
<species metaid="metaid_0000006"
         id="L_EGFR"
         compartment="compartment"
         initialConcentration="0">
 <annotation>
  <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22rdfsyntaxns#"
   xmlns:bqbiol="http://biomodels.net/biologyqualifiers/">
  <rdf:Description rdf:about="#metaid_0000006">
    <bqbiol:hasPart>
      <rdf:Bag>
        <rdf:li rdf:resource="http://www.uniprot.org/#P07522" />
        <rdf:li rdf:resource="http://www.uniprot.org/#Q9QX70_RAT" />
      </rdf:Bag>
    </bqbiol:hasPart>
   </rdf:Description>
  </rdf:RDF>
 </annotation>
</species>

[...]
```



```
[...]
<species meta
          id="
          comp
          init
 <annotation>
  <rdf:RDF xm
   xmlns:bqbi
  <rdf:Descri
    <bqbiol:h
      <rdf:Ba
        <rdf
        <rdf
      </rdf:l
    </bqbiol
   </rdf:Desc
  </rdf:RDF>
 </annotation
</species>
[...]
```



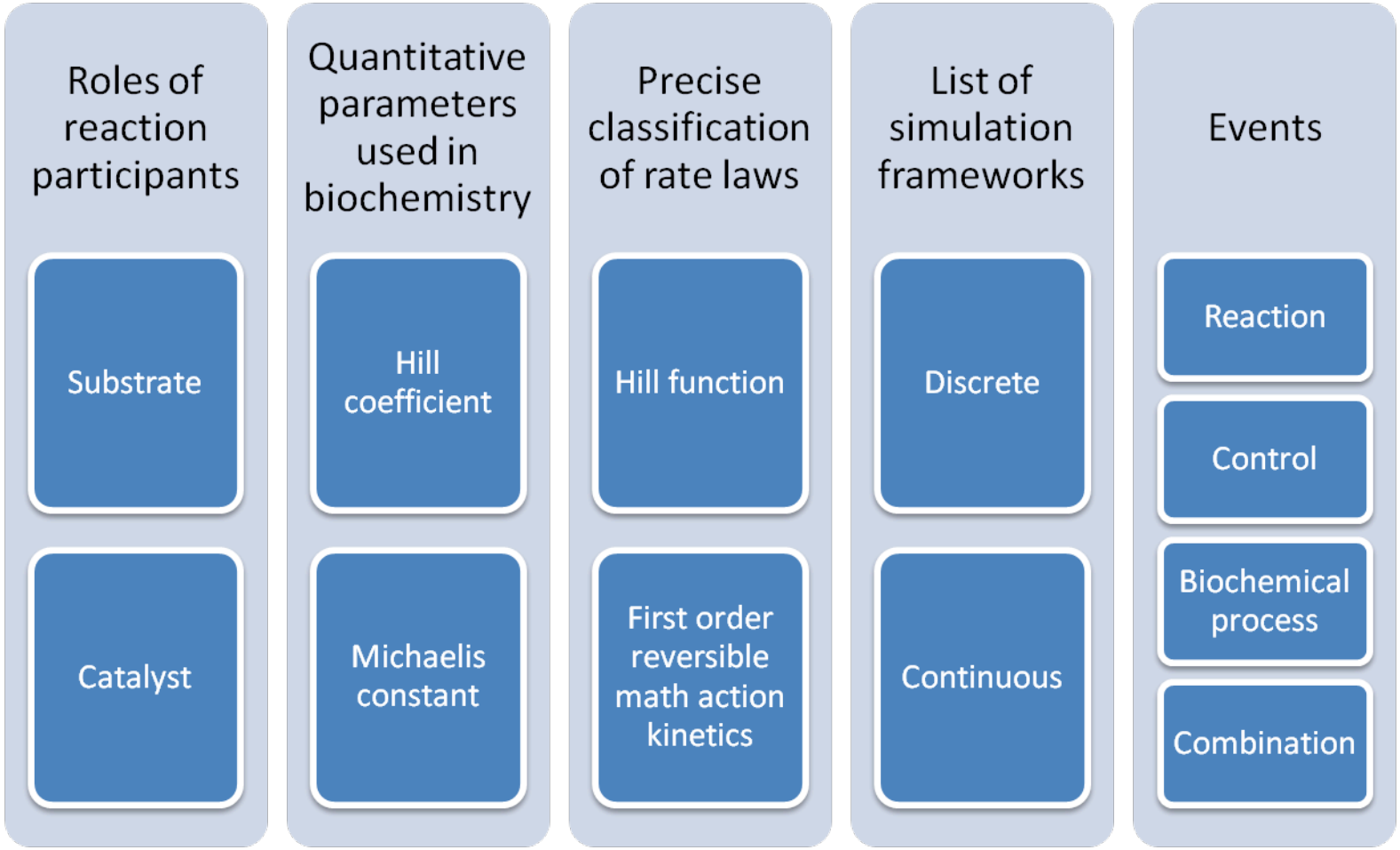
MIRIAM

- Summarizing
  - MIRIAM allows to link a model to:
    - the publication,
    - the author and
    - to uniquely identify model participants.
  - Computational Resources available:
    - http://www.ebi.ac.uk/miriam/
- BUT: What about all the assumptions made during the creation of the model?

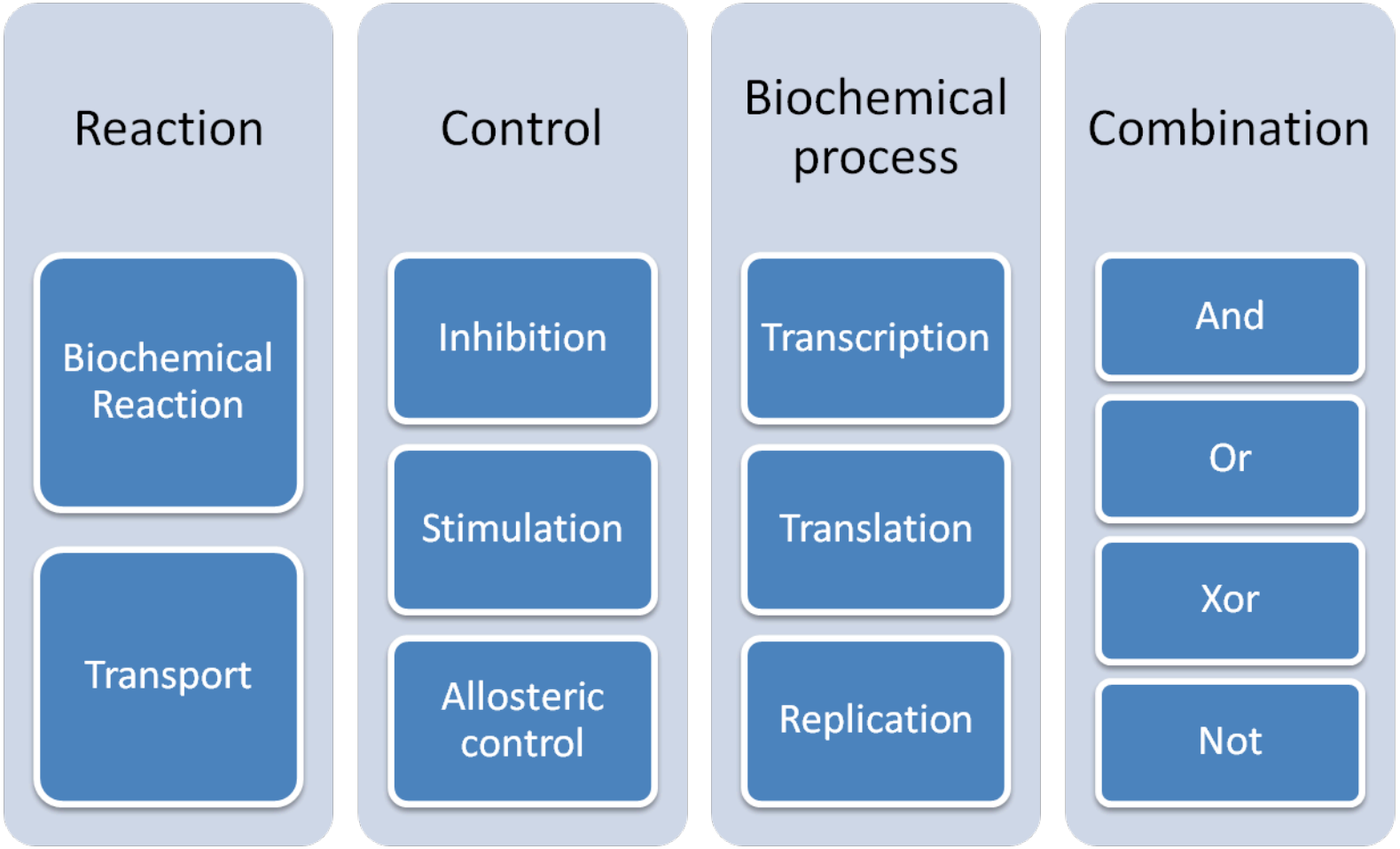
Systems Biology Ontology

- Classification of concepts used in quantitative modeling
- Identify each model component and the relationship between it and the model structure
- “A Simple Proposal for Indicating Modeling Framework Assumptions Underlying an SBML Model”

SBO – Controlled Vocabularies



Events


Reaction
Biochemical Reaction
Transport
Control
Inhibition
Stimulation
Allosteric control
Biochemical process
Transcription
Translation
Replication
Combination
And
Or
Xor
Not

Systems Biology Ontology

- Taxonomy of the roles of reaction participants
- Examples:
  - Catalysis
  - competitive inhibitor
  - Protein
  - enzyme

Systems Biology Ontology

- Controlled vocabulary for parameter roles
- Examples:
  - forward uni-molecular rate constant
  - Michaelis constant
  - total concentration of enzyme
  - pKa

Systems Biology Ontology

- Classification of mathematical expression used in biochemical modeling
- Examples:
  - mass action kinetic
  - Henri-Michaelis-Menten kinetics

Systems Biology Ontology

- List of modeling frameworks (specifies how to interpret a mathematical expression)
- Examples:
  - Continuous
  - Discrete
  - logical framework

Systems Biology Ontology

- Classification of events represented by biochemical models
- Examples:
  - Binding
  - Transport
  - phosphorylation

Why do we need it?

Henri-Michaelis-Menten:

$E+S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_p} E+P$; Rapid equilibrium $\quad K_m = k_{-1}/k_1$

Van Slyke-Cullen:

$E+S \xrightarrow{k_1} ES \xrightarrow{k_p} E+P$; Total duration equal sum of durations

$K_m = k_p/k_1$

Briggs-Haldane:

$E+S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_p} E+P$; Quasi-steady state $\quad K_m = (k_{-1}+k_p)/k_{-1}$

Why do we need it?

```
<reaction>
   <listOfReactants>
     <speciesReference species="S" />
   </listOfReactants>
   <listOfProducts>
     <speciesReference species="P" />
   </listOfProducts>
   <listOfModifiers>
     <speciesReference species="E" />
   </listOfModifiers>
   <kineticLaw>
     <listOfParameters>
       <parameter id="Km"/>
       <parameter id="kp"/>
     </listOfParameters>
     <math xmlns="http://www.w3.org/1998/Math/MathML">
       <apply>
         <divide/><apply>
                    <times/><ci>E</ci>
                            <ci>kp</ci>
                            <ci>S</ci>
                  </apply>
                  <apply>
                    <plus/><ci>Km</ci>
                           <ci>S</ci>
                  </apply>
       </apply>
     </math>
   </kineticLaw>
 </reaction>
```

$E+S \rightarrow E+P$

Import in a discrete simulator

?

$$E+S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_p} E+P \; ; \; k_1 = k_{-1}/K_m$$

$$E+S \xrightarrow{k_1} ES \xrightarrow[k_{-1}]{k_p} E+P \; ; \; k_1 = k_p/K_m$$

$$E+S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_p} E+P \; ; \; k_1 = (k_{-1}+k_p)/K_m$$

SBO – How to use it

```
<reaction>
   <listOfReactants>
     <speciesReference species="S" />
   </listOfReactants>
   <listOfProducts>
     <speciesReference species="P" />
   </listOfProducts>
   <listOfModifiers>
     <speciesReference species="E" />
   </listOfModifiers>
   <kineticLaw sboTerm="SBO:0000031">
     <listOfParameters>
       <parameter id="Km"/>
       <parameter id="kp"/>
     </listOfParameters>
     <math xmlns="http://www.w3.org/1998/Math/MathML">
       <apply>
         <divide/><apply>
                    <times/><ci>E</ci>
                            <ci>kp</ci>
                            <ci>S</ci>
                  </apply>
                  <apply>
                    <plus/><ci>Km</ci>
                           <ci>S</ci>
                  </apply>
       </apply>
     </math>
   </kineticLaw>
 </reaction>
```

Briggs-Haldane equation

SBO – How to use it

```
<reaction sboTerm="SBO:0000062">
   <listOfReactants>
     <speciesReference species="S" sboTerm="SBO:0000015"/>
   </listOfReactants>
   <listOfProducts>
     <speciesReference species="P" sboTerm="SBO:0000011"/>
   </listOfProducts>
   <listOfModifiers>
     <speciesReference species="E" sboTerm="SBO:0000014"/>
   </listOfModifiers>
   <kineticLaw sboTerm="SBO:0000031">
     <listOfParameters>
       <parameter id="Km" sboTerm="SBO:0000008"/>
       <parameter id="kp" sboTerm="SBO:0000025"/>
     </listOfParameters>
     <math xmlns="http://www.w3.org/1998/Math/MathML">
       <apply>
         <divide/><apply>
                    <times/><ci>E</ci>
                            <ci>kp</ci>
                            <ci>S</ci>
                  </apply>
                  <apply>
                    <plus/><ci>Km</ci>
                           <ci>S</ci>
                  </apply>
       </apply>
     </math>
   </kineticLaw>
 </reaction>
```

continuous framework

substrate

product

enzyme

Briggs-Haldane equation

Michaelis constant

catalytic rate constant

SBO – How to use it

```
<reaction sboTerm="SBO:0000062">
  <listOfReactants>
    <speciesReference species="A" sboTerm="SBO:0000015"/>
  </listOfReactants>
  <listOfProducts>
    <speciesReference species="B" sboTerm="SBO:0000011"/>
  </listOfProducts>
  <listOfModifiers>
    <speciesReference species="C" sboTerm="SBO:0000014"/>
  </listOfModifiers>
  <kineticLaw sboTerm="SBO:0000031">
    <listOfParameters>
      <parameter id="U" sboTerm="SBO:0000008"/>
      <parameter id="V" sboTerm="SBO:0000025"/>
    </listOfParameters>
  </kineticLaw>
</reaction>
```

continuous simulator

$$v = [C] \cdot V \cdot \frac{[A]}{(U + [A])}$$

discrete simulator

$$v_1 = \frac{(k_{-1} + V)}{U} \cdot [A] \cdot [C]$$

$$v_2 = k_{-1} \cdot [D]$$

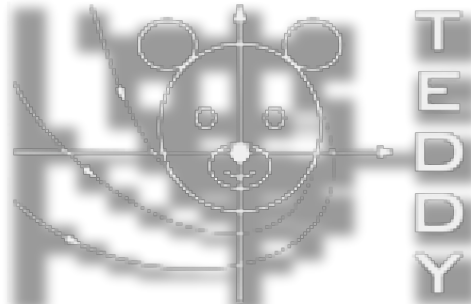
$$v_3 = V \cdot [D]$$

Systems Biology Ontology

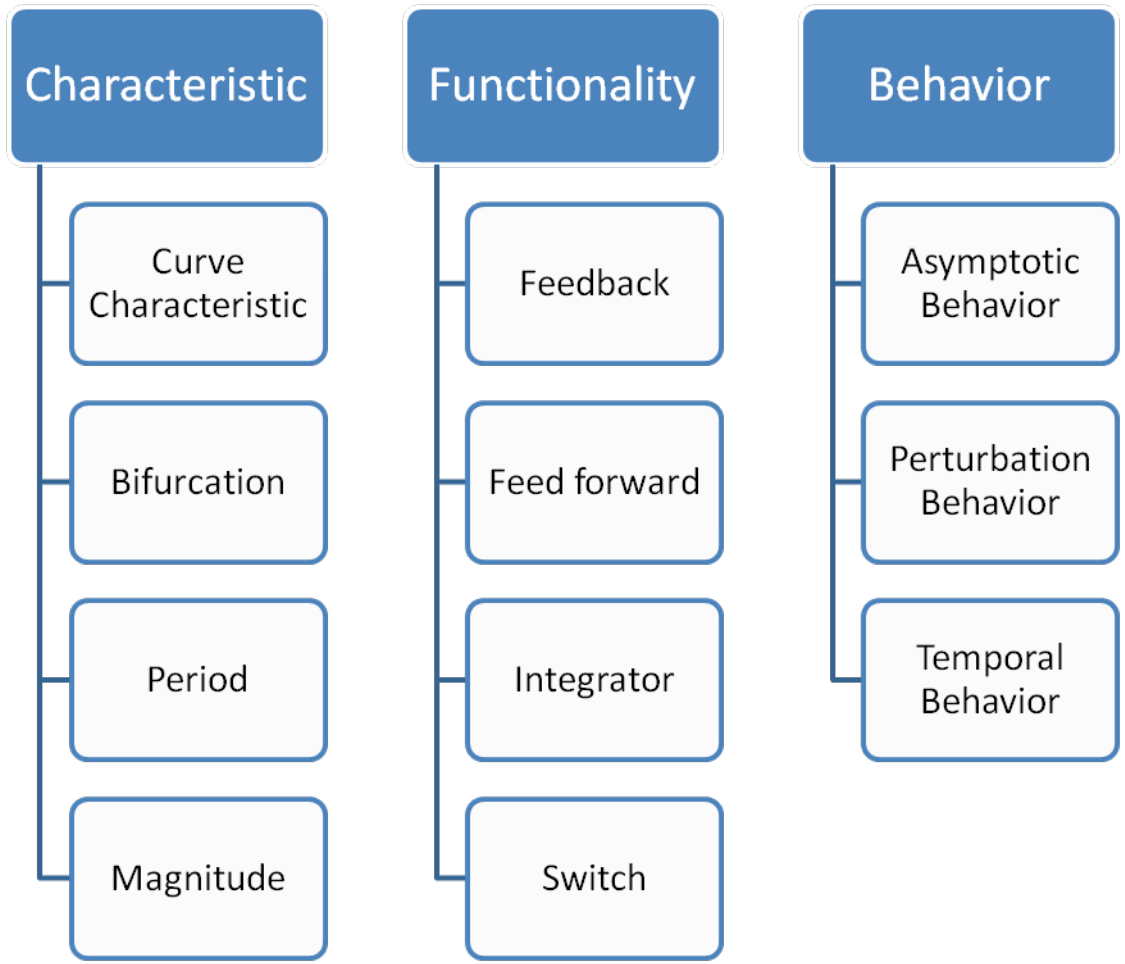
- Semantic layer:
  - link between models encoded in SBML and graphical notations (such as SBGN)
  - conversion to semantically enriched computing formats (such as BioPAX)
  - translation of models between *continuous deterministic frameworks and discrete stochastic framework*
  - merging/integration of models

TEDDY

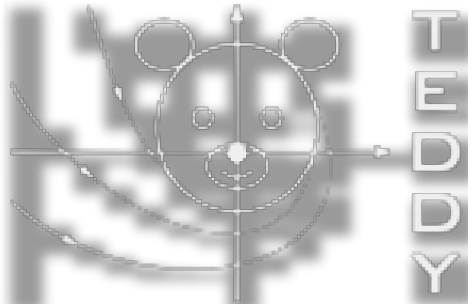
- TErminology for the Description of DYnamics
- an ontology for
  - dynamical behaviors,
  - observable dynamical phenomena,
  - and control elements of bio-models and biological systems in Systems Biology and Synthetic Biology
- > 130 classes
- 6 relationships / 3 vocabularies



TEDDY vocabulary

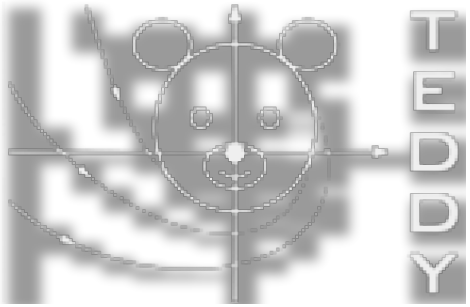


http://teddyontology.sourceforge.net/teddy/current/doc/index.html



TEDDY

- Bifurcation
  - Local Bifurcation
    - Hopf Bifurcation
    - Subcritical Hopf Bifurcation
    - Supercritical Hopf Bifurcation
    - Degenerate Hopf Bifurcation
  - Saddle-Node on Invariant Circle Bifurcation
  - Zero-Eigenvalue Bifurcation
    - Saddle-Node Bifurcation
    - Transcritical Bifurcation
    - Pitchfork Bifurcation
      - Subcritical Pitchfork Bifurcation
      - Supercritical Pitchfork Bifurcation
  - Bifurcation from Limit Cycle
    - Saddle-Node Bifurcation of Limit-Cycles
    - Period-Doubling Bifurcation of Limit-Cycles
    - Neimark-Sacker Bifurcation of Limit-Cycles
  - Saddle-Node Homoclinic Bifurcation



Reproducing / Evaluating Simulation Results

- MIASE - Minimum Information needed About a Simulation Experiment
- SBRML - Systems Biology Results Markup Language

MIASE – Why is it necessary?

- Problem:
  - "The model, when instantiated within a suitable simulation environment, must be able to reproduce all relevant results given in the reference description that can readily be simulated"
- However: MIRIAM does not include guidelines about how a relevant result can be reproduced

MIASE – Motivation

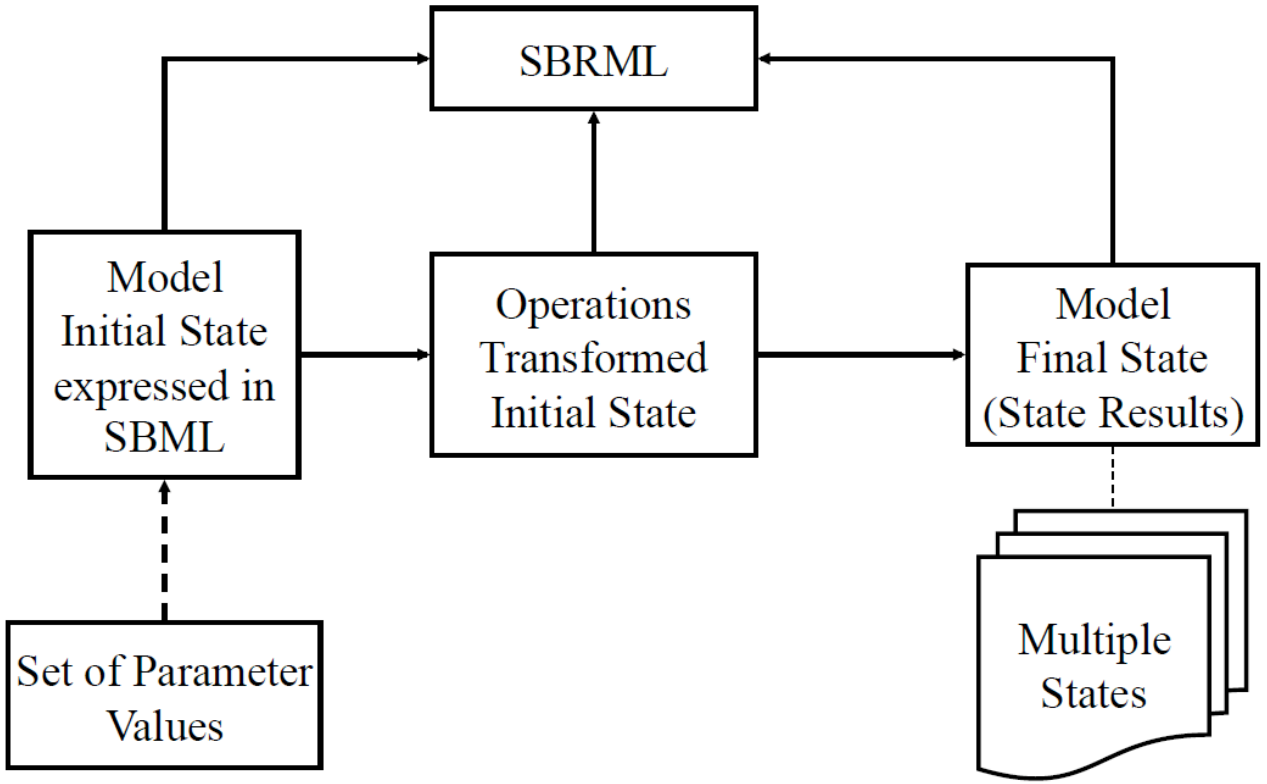
- Aims
  - researchers should be able to exchange and share their results
- Problem
  - most models are supplied as plain SBML/CellML/BioPAX code
  - a model curator does not have any information about simulation settings that led to the validation of the supplied model
  - after curation, the model is made available on biomodels.net, including a screenshot of the validated model
  - the next user has to re-do the whole process of identifying parameter settings before he/she can use the model
- Solution:
  - Provide information about the simulation settings along with the model

MIASE / KiSAO

- MIASE - The description of the Minimum Information needed About a Simulation Experiment in order to repeat it
  - Simulation Algorithm used to simulate a given model (uses KiSAO)
  - Simulation Tool used
  - Parameter settings
- KiSAO - Kinetic Simulation Algorithm Ontology
  - simulation algorithms / simulation methods
  - and their relation to each other

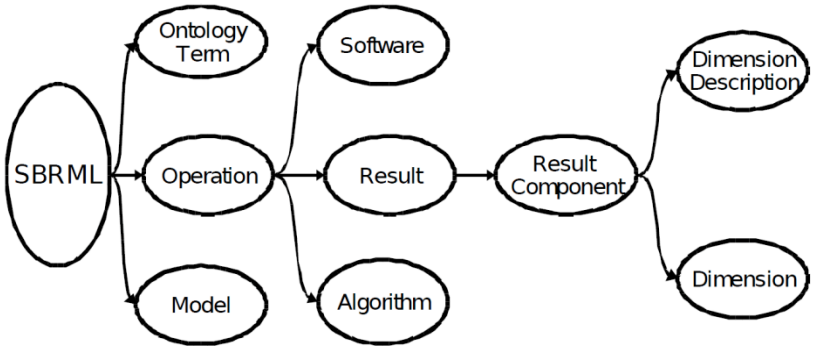
SBRML

- Aim: Linking numerical results to the model that gave rise to them



SBRML

- Ontology Term :
  - Defines all the ontology terms used in SBRML.
- Model :
  - Defines the model used to generate the result.
- Operation :
  - Captures the operation that is performed on the model.
- Software :
  - Defines the software used to perform the operation on the model.
- Algorithm:
  - Defines the algorithm used in the operation.
- Result :
  - Captures all the result of an operation on the model.
- Result Component
  - Captures the individual component of result.
- Dimension Description
  - A header of the result component. It described the data in the component.
- Dimension
  - Captures the actual data of the result component.



SBRML

```
<?xml version="1.0" encoding="UTF-8"?>
<sbrml xmlns="http://www. comp-sys-bio.org/sbrml/level1" version="1.0"
level="1.0">
    <ontologyTerms>
        <ontologyTerm id="term1" term="Steady State"
    sourceTermId="Term00001" ontologyURI="TestOntoWebSite.org" />
    </ontologyTerms>
    <model id="model1" name="Edelstein1996_.." sourceURI="file://BIO..1" />
    <operations>
        <operation id="op1" ontologyTerm="term1">
            ...
        </operation>
    </operations>
</sbrml>
```

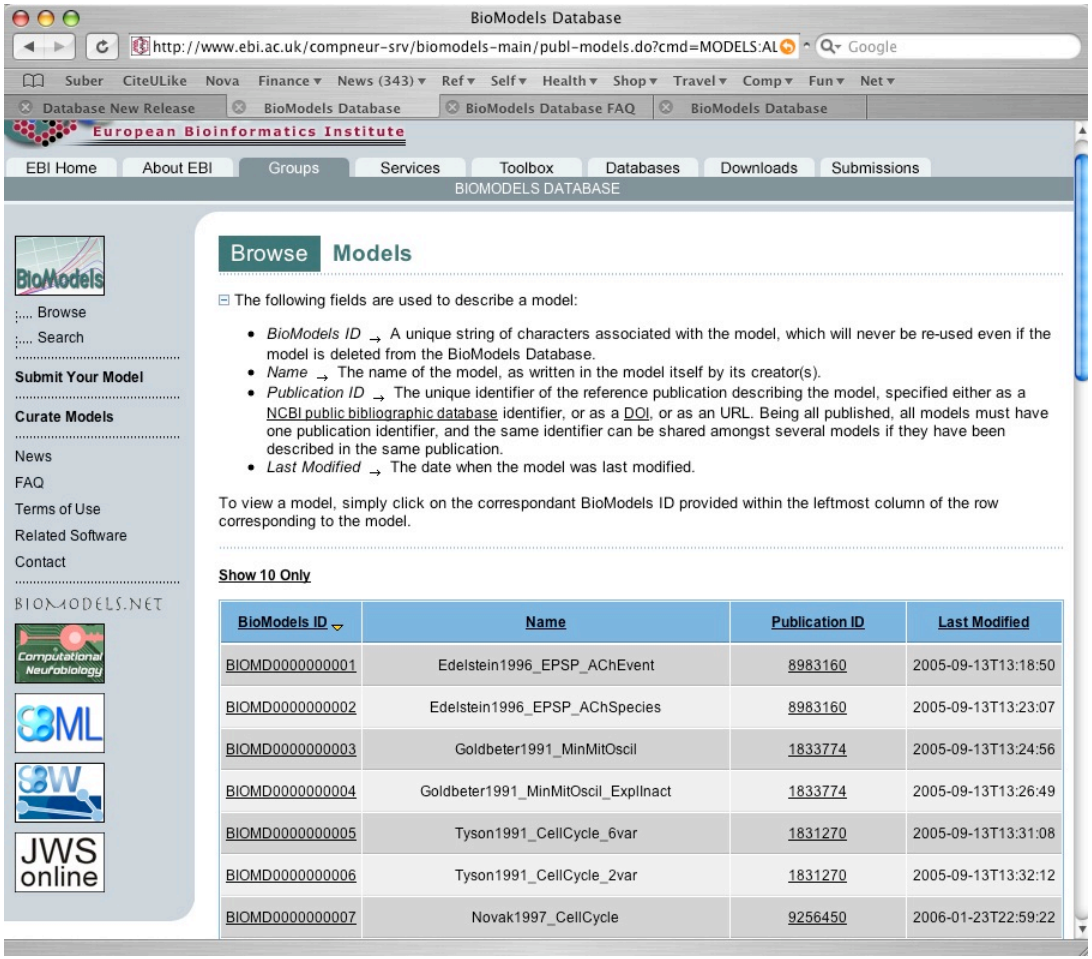
Model Repositories

- BioModels Database
- CellML Repository
- Physiome.org

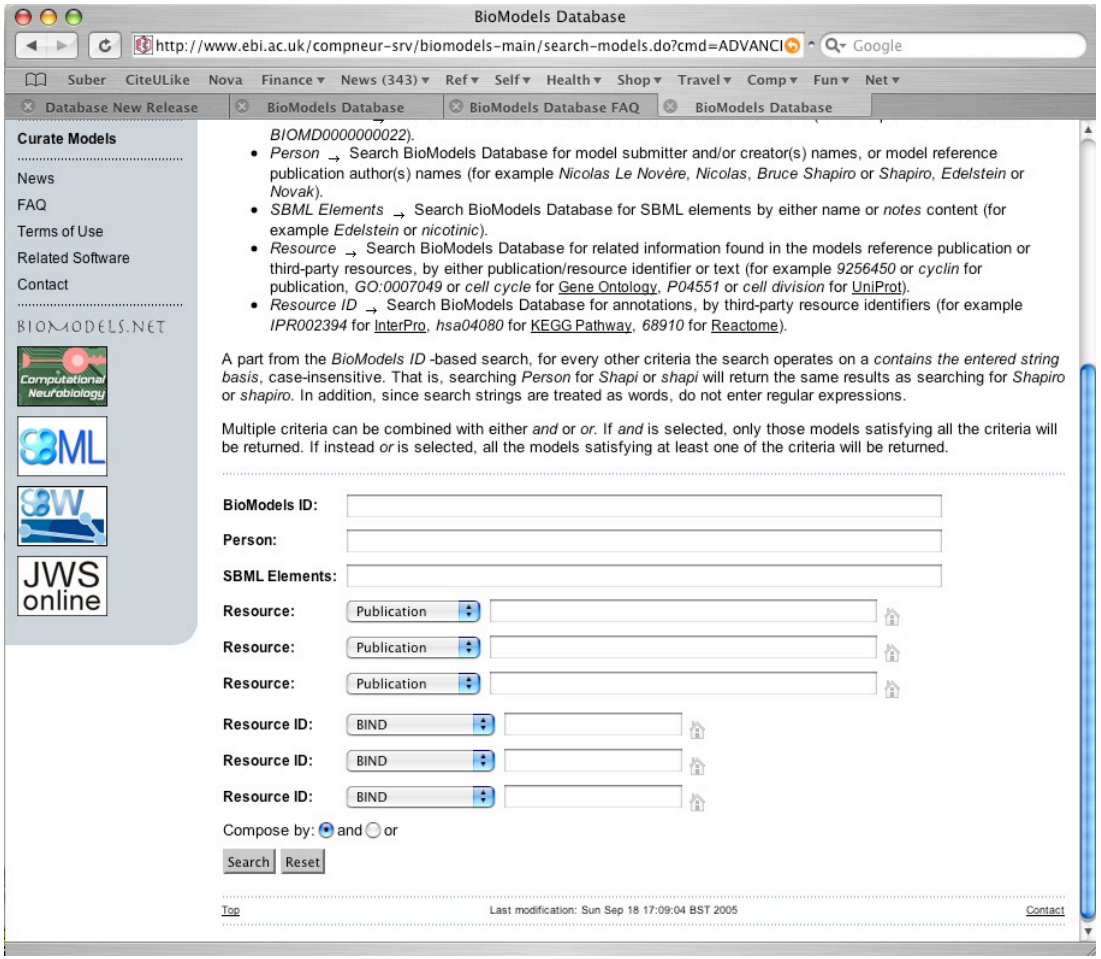
BioModels Database

- Aims to be the Swiss-Prot of quantitative modeling
- Stores & serves quantitative models of biomed. interest
  - Only models described in peer-reviewed scientific literature
- Models are curated by humans: computer software checks syntax, humans check semantics
- Models are simulated to check correspondence to reference
- Model components are annotated to improve identification and retrieval
- Accepted in SBML and CellML formats, served in several (SBML, XPP, CellML, diagram; more coming)

www.ebi.ac.uk/biomodels



Search


BioModels Database
http://www.ebi.ac.uk/compneur-srv/biomodels-main/search-models.do?cmd=ADVANCI
Google
Suber CiteULike Nova Finance News (343) Ref Self Health Shop Travel Comp Fun Net
Database New Release
BioModels Database
BioModels Database FAQ
BioModels Database
Curate Models
News
FAQ
Terms of Use
Related Software
Contact
BIOMODELS.NET
Computational Neurobiology
SBML
SBW
JWS online
BIOMD0000000022).
Person → Search BioModels Database for model submitter and/or creator(s) names, or model reference publication author(s) names (for example Nicolas Le Novère, Nicolas, Bruce Shapiro or Shapiro, Edelstein or Novak).
SBML Elements → Search BioModels Database for SBML elements by either name or notes content (for example Edelstein or nicotinic).
Resource → Search BioModels Database for related information found in the models reference publication or third-party resources, by either publication/resource identifier or text (for example 9256450 or cyclin for publication, GO:0007049 or cell cycle for Gene Ontology, P04551 or cell division for UniProt).
Resource ID → Search BioModels Database for annotations, by third-party resource identifiers (for example IPR002394 for InterPro, hsa04080 for KEGG Pathway, 68910 for Reactome).
A part from the BioModels ID -based search, for every other criteria the search operates on a contains the entered string basis, case-insensitive. That is, searching Person for Shapi or shapi will return the same results as searching for Shapiro or shapiro. In addition, since search strings are treated as words, do not enter regular expressions.
Multiple criteria can be combined with either and or or. If and is selected, only those models satisfying all the criteria will be returned. If instead or is selected, all the models satisfying at least one of the criteria will be returned.
BioModels ID:
Person:
SBML Elements:
Resource: Publication
Resource: Publication
Resource: Publication
Resource ID: BIND
Resource ID: BIND
Resource ID: BIND
Compose by: and or
Search Reset
Top
Last modification: Sun Sep 18 17:09:04 BST 2005
Contact

Annotation Sources

model	Submitter, Creators, publication, Gene Ontology, ICD10, KEGG pathway, OMIM, Reactome, Taxonomy
compartment	Gene Ontology
species	BIND complex, ChEBI, Ensembl, Gene Ontology, InterPro, KEGG compound, OMIM, Taxonomy, UniProt
rule	BIND interaction, EC code, Gene Ontology, IntAct, KEGG pathway, KEGG Reaction, Reactome
reaction	BIND interaction, EC code, Gene Ontology, IntAct, KEGG pathway, KEGG Reaction, Reactome
event	BIND interaction, EC code, Gene Ontology, IntAct, KEGG pathway, KEGG Reaction, Reactome

From Nicolas Le Novère @ EBI

CellML Repository

- CellML model repository has over 350 published models of:
  - Signal transduction pathways;
  - Metabolic pathways;
  - Electrophysiological;
  - Calcium dynamics;
  - Immunology;
  - Cell cycle;
  - Smooth and skeletal muscle models;
  - Mechanical and constitutive relationships.
- **http://www.cellml.org/models/**

  *(C. M. Lloyd, J. R. Lawson, P. J. Hunter, and P. F. Nielsen, "**The CellML model repository**" Bioinformatics, 2008)*

[Peter Hunter, *Distributing and maintaining models in CellML,* Life Sciences 2008, Montreal, August, 2008]

Physiome.org Repository

- The physiome.org model repository has about 300 live models (and is undergoing a big revision):
  - Convection-diffusion reaction and exchange;
  - Tissue and organ models for PET and MRI analysis;
  - Electrophysiology, mostly cardiac;
  - Physico-chemical, osmotic processes, cells, tissues;
  - Enzymatic reactions, metabolic networks;
  - Transport and exchange of respiratory gases, ;
  - Circulatory and respiratory mechanics, coupled with transport and control mechanisms;
  - Membrane transporters and pumps.
  - Pharmacokinetic models for complex systems.
  - Families of tutorials for physiological transport and metabolism.
- **http://www.physiome.org/Models/**

Raymond GM, Butterworth E, and Bassingthwaighte JB. JSIM: Free software package for teaching physiological modeling and research. Exper Biol 2003 280.5, p102, 2003.

The Physiome.org Repository

- JSim is a powerful modeling interface system:
  - Designed for analyzing experimental data, providing parameter estimation and confidence limits. Sensitivity analysis built in. Multiple optimizers for weighted non-linear, least squares minimization.
  - Simple human readable coding of ODEs and PDEs and algebraic differential equations, MML. Has multiple solvers for ODEs and PDEs
  - Project files (preserved as mathML) provide for multiple models and model comparisons, multiple data sets, varied and multiple graphs, extensive notes, enough for all figures in a paper.
  - Downloads and runs models from CellML, SBML, Biomodels, JWS.
  - Runs Wikis over the web. Allows free editing to all, and download.
  - Code at the speed of thought (or your typing, whichever is slower).
  - Runs on Linux, MAC, Windows. Free and open.
  - Automated unit balance checking and unit scaling conversion.
  - Can runs models written in C, FORTRAN and Matlab.
  - Can be used to drive Finite Element/ Finite Volume solvers (Nike, Open Foam) to provide dynamic boundary conditions for 3D models.
  - Easily learned. Grad students start changing the programs in 10 min.
- **http://www.physiome.org/JSim/**

  Raymond GM, Butterworth E, and Bassingthwaighte JB. JSim: Free software package for teaching physiological modeling and research. Exper Biol 2003 280.5, p102, 2003.

Physiome.org Repository

- JSim is incomplete:
  - Events in SBML not handled until next release.
  - Requires explicit coding to check mass, charge, osmotic, and thermodynamic balances.
  - Needs translators from MML to SBML or CellML for the ODE level models. Not able to link to Beard's enzyme database.
  - Does not have specific domain notation for compartmental regions but does for spatially distributed regions.
  - Does handle linear programming problems without special calls, but does solve complex and time-variable boundary conditions for PDEs.
  - Does not use an ontology specification, yet.
  - Can be stretched to 2-D PDE problems but not 3D or Greens functions.
  - The automated model construction from modules works, but is limited.
- Virtues:
  - Great generality of problems that can be solved.
  - Handles models as big as over 100,000 ODEs.
  - Uses multiprocessing (tested on 8); uses re-entrant solvers.
  - Fast: 300 times faster than Matlab/Simulink, about 1/2 times C. Runs a 200+ ODE closed loop, regulated circulatory/respiratory exchange model in real time.

Discussion:

Define the central mission(s) for WG10

How do we promote adherence to principles:

- Data gathering guided by hypotheses testing strategies?
- Can we figure out alternative hypotheses to define efficient experimentation?
- How do we persuade our community (societies, journals) to review to high standards?

Standardization, will these topics suffice or deter?

- Ontologies
- Standards for quantitative modeling
- Journal publication standards and peer review
- Full disclosure of thinking, failures, successes
- Unfettered dissemination
- Standards for curation levels: Uniform descriptions covering the several repositories?

How should one design and build the repositories?

- Completeness in reporting (scientific literature)
- We have good databases of digitizable biological information (gene, proteins). How can we develop real physiological databases for time series (like PhysioNet generalized), images and image sequences, experimental methods of all sorts?
- Model databases of working models are coming along. How about for untested hypotheses: models to be tested against difficult to get data?.
- How do we get stable federal support for essential basic structures like databases?
- How do we compete with Europe and Japan in funding basic resources, repositories?