Merging of systems biology models with semanticSBML

Wolfram Liebermeister, Falko Krause, Edda Klipp
Max Planck Institute for Molecular Genetics, Berlin
I. Mathematical models in systems biology
Biochemical pathway models

Budding yeast

Biochemical entities
Metabolites, reactions, ...

Biochemical quantities
Concentrations, velocities, ...

Mathematical statements
Values, equations, ...

Full-scale model of glycolysis in Saccharomyces cerevisiae, F. Hynne et al., 2001, Biophysical Chemistry (94), 121-163.
Mathematical frameworks used in systems biology

- Ordinary differential equations
- Partial differential equations
- Stochastic processes
- Particle-based models
- Constraint-based models
- Optimality-based models

A common feature: the **semantics**
- List of mathematical objects
- Mapping to biochemical objects / quantities
- List of mathematical statements

Statements can be seen as **facts** or **rules for computation**
### Kinetic models with formal semantics

#### Network scheme

- **F6P**
- **ATP**
- **FBP**
- **FBA**

#### Model elements

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Variable</th>
<th>Mathematical statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>F6P conc. [mM]</td>
<td>(c_{F6P})</td>
<td>(C_{F6P} = 0.1)</td>
</tr>
<tr>
<td>ATP conc. [mM]</td>
<td>(c_{ATP})</td>
<td>(C_{ATP} = 0.5)</td>
</tr>
<tr>
<td>PFK vel. [mM/s]</td>
<td>(v_{PFK})</td>
<td>(v_{PFK} = v_{PFK}(c_{F6P}, c_{ATP}))</td>
</tr>
<tr>
<td>FBP conc. [mM]</td>
<td>(c_{FBP})</td>
<td>(dc_{FBP}/dt = v_{PFK} - v_{FBA}) (c_{FBP}(0) = c_{F6P})</td>
</tr>
<tr>
<td>FBA vel. [mM/s]</td>
<td>(v_{FBA})</td>
<td>(v_{FBA} = v_{FBA}(c_{FBP}))</td>
</tr>
</tbody>
</table>
Species called “ATP” represents KEGG C06262 (ATP)

"A simple scheme for annotating SBML with references to controlled vocabularies and database entries"
Le Novere and Finney, 2005
II. Merging of biochemical models
Playing with biochemical models?
Playing with biochemical models?

Model composition

Model merging
Merging of graphical models

Elements are characterised by:

Name: "How is the element called?"
(need not be compatible between models)

Semantics ("annotation"): "What does the element represent?"
(used for matching the elements between models)

Statements: "What is stated about the element?"
(can lead to conflicts between models / have to be chosen)

Additional information
(e.g., icon, position in graphics etc)
Matching elements between models


Can yeast glycolysis be understood in terms of in vitro kinetics of t, Bas Teusink, et al., 2000 European Journal of Biochemistry, 267, 5313-5329.
Matching elements between models


1. **Independent**: nothing in common, no constraints, no conflict
Semantic relations between quantities

1. **Independent**: nothing in common, no constraints, no conflict

   - **Model 1**: ATP concentration
   - **Model 2**: ADP concentration

2. **Identical**: *statement conflict possible*; choose between statements

   - **ATP concentration**

---

**Note:**
- ATP: Adenosine Triphosphate
- ADP: Adenosine Diphosphate
Semantic relations between quantities

1. **Independent**: nothing in common, no constraints, no conflict

   - **Model 1**: ATP concentration
   - **Model 2**: ADP concentration

2. **Identical**: *statement conflict possible*; choose between statements

   - ATP concentration

3. **Interconvertible**: need to be converted in advance (afterwards, see “identical”)

   - ATP concentration [mM]
   - ATP concentration [M]
   - ATP concentration
   - ATP amount
Semantic relations between quantities

1. **Independent**: nothing in common, no constraints, no conflict
   - Model 1: ATP concentration
   - Model 2: ADP concentration

2. **Identical**: *statement conflict possible*; choose between statements
   - ATP concentration ↔ ATP concentration

3. **Interconvertible**: need to be converted in advance (afterwards, see “identical”)
   - ATP concentration [mM] ↔ ATP concentration [M]
   - ATP concentration ↔ ATP amount

4. **Semantic overlap**: *uncontrollable conflicts*; models cannot be merged automatically
   - lumped reaction ↔ individual reaction steps
   - ribosome concentration ↔ total RNA concentration
A merging algorithm for semantic models with explicit equations

1. Convert all quantities to standard units
A merging algorithm for semantic models with explicit equations

1. Convert all quantities to standard units
2. Match redundant / conflicting elements

Conflicts between elements? merging is impossible
A merging algorithm for semantic models with explicit equations

1. Convert all quantities to standard units
2. Match redundant / conflicting elements

3. Choose between statements
# Validity criteria for systems biology models

<table>
<thead>
<tr>
<th>Aspects of validity</th>
<th>Example validity criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Syntax</strong></td>
<td>• correct file format ...</td>
</tr>
<tr>
<td>Model can be read and processed</td>
<td></td>
</tr>
<tr>
<td><strong>2. Computation</strong></td>
<td>• statements unique + complete</td>
</tr>
<tr>
<td>Model can be used for predictive simulations</td>
<td>• sequential evaluation possible ...</td>
</tr>
<tr>
<td><strong>3. Semantic correctness</strong></td>
<td>• valid statements</td>
</tr>
<tr>
<td>Model statements agree with the model semantics</td>
<td>• no semantic dependencies ...</td>
</tr>
<tr>
<td><strong>4. Empirical correctness</strong></td>
<td>• realistic numerical values</td>
</tr>
<tr>
<td>Model agrees with physical and biochemical facts</td>
<td>• correct thermodynamics</td>
</tr>
<tr>
<td></td>
<td>• correct reaction balances, ...</td>
</tr>
<tr>
<td><strong>5. Context</strong></td>
<td>• agreement with data</td>
</tr>
<tr>
<td>Model performs well and suits its purpose</td>
<td>• plausible assumptions</td>
</tr>
<tr>
<td></td>
<td>• no irrelevant parts, ...</td>
</tr>
</tbody>
</table>

III. Model annotation and merging with semanticSBML
SemanticSBML: an interactive tool for model annotation and merging

Functions in semanticSBML
- Annotate
- Check
- Merge
- Build
- Display

Versions of semanticSBML
- Stand-alone version with GUI
- Web version (merging still under development)
- API (programming interface)

http://semanticSBML.sf.net
http://www.semanticsbml.org
Annotation GUI
Annotation GUI

- **Compartments**
  - [cytosol] [ID:cytosol]
    - is Gene Ontology: GO:0005829 cytosol
  - [extracellular region] [ID:extracellular]
- **Species**
  - Glycerol [ID:GLY]
    - is ChEBI: CHEBI:17754 glycerol
    - is KEGG Compound: C00116 glycerol

Search by Name

**glycerol**

Add

- **Database**
  - **Identifier**
    - CAS: S6-40-6
    - PubChem: 3339
    - KEGG: C00037
    - ChEBI: CHEBI:15428
Annotation GUI

Current Annotations

- Bis(is (change)
- is ChEBI C00167 by glycerol

Search by Name

- GLY

Add

- Database: Identifier
- 3DMET (add)
- CAS: S340-6
- glycine (add)
- PubChem: 3339 (add)
- glycine
- KEGG: C00037 (add)
- glycine
- ChEBI: C00167 (add)
Annotation GUI
Annotation GUI
Annotation GUI
Annotation GUI

- Model BIOMD0000000064
  - ✅ Models
  - ✅ Compartments
    + ✅ [cytosol] [ID:cytosol]
      - is Gene Ontology: GO:0005829 cytosol

Compartment

Automatic annotation: Annotate all elements in the Compartments with the first search hit found

annotate now

- Extracellular Glucose [ID:OL.C0]
- Succinate [ID:SUC.C]
- CO2 [ID:CO2]
- Trehalose [ID:Thr]
- Glycogen [ID:Glyo]
- NADH [ID:NADH]
- NAD [ID:NAD]
- High energy phosphates [ID:ATP]
- Acetate/acetate [ID:ACET]
- Pyruvate [ID:PYR]
- Phosphoenolpyruvate [ID:PEP]
- 2-phosphoglycerate [ID:F2P]
- 3-phosphoglycerate [ID:F3G]
- 1,3-bisphosphoglycerate [ID:BPG]
- Triose phosphate [ID:TRIO]
- Fructose-1,6-bisphosphate [ID:F16P]
- Fructose 6 Phosphate [ID:F6P]
- Glucose 6 Phosphate [ID:G6P]
- Glucose in Cytosol [ID:GLC]
Merge GUI

Overview

Merged Model

Models 1...n

Merge GUI

Merged Model

Teusink2000 Glycolysis

Hynne2001 Glycolysis

Cytoplasm

Annotations

Gene Ontology: GO:0005520

cytosol

Quantity:

Type: volume
Unit: litre

Location:

no outside specified

Statement:

Constant True

Size: 1.0
Volume: 1.0

Dependencies

This Element is Depending On

Elements Depending On This Element

Resolve

Resolve All Conflicts With Model Priority:

0 - Teusink2000 Glycolysis

- Hynne2001 Glycolysis

Resolve Conflicts
Element Cytosolic glucose has conflicting values. Element Glycerol has conflicting values. Element NADH has conflicting values. Element 1,3-Bisphosphoglycerate has conflicting values. Element Fructose 1,6-bisphosphate has conflicting values. Element EtOH has conflicting values. Element Pyruvate has conflicting values. Element Mixed flow glucose61 has conflicting values. Element NAD has conflicting values. Element Phosphoenolpyruvate has conflicting values. Element Glyceraldehyde 3-phosphate has conflicting values. Element Acetalddehyde has conflicting values. Element ATP has conflicting values. Element Hexokinase has conflicting values. Element Phosphoglucoisomerase has conflicting values. Element PhosphofrUCTokinase has conflicting values. Element Glyceraldehyde 3-phosphate dehydrogenase has conflicting values. Element Aldolase has conflicting values. Element Pyruvate kinase has conflicting values. Element Pyruvate decarboxylase has conflicting values. Element Glucose uptake has conflicting values. Element Phosphoenolpyruvate synthesis has conflicting values.
Merge GUI

**Annotations**

- bio:hasPart
  - ChEBI CHEBI:15422
    - atp
    - delete
  - bio:is
    - Reactome 243599
    - delete
  - bio:is
    - Reactome 70102
    - delete

**Merged Model**

- cytosol
  - cytosol_merge
  - Annotations
    - bio:is
      - Gene Ontology GO:0005520
      - cytosol
  - Quantity
    - Type: volume
    - Unit: litre
  - Location
    - no outside specified
  - Statement
    - Constant True
  - Size 1.0
  - Volume 1.0

**Resolve All Conflicts With Model Priority:**

0 - Teusink2000_Glycolysis

Resolve

**Resolve Conflicts**
Merge GUI

Merging of Annotation Not Possible

You are trying to merge two MIRIAM annotations with qualifier 'is' pointing to the database EC code. Merging is not possible.

bio:is

Reactome 70102

Resolve Conflicts
Merge GUI

Glycerol

GLY0

Merge with Elements from

Similar: Extracellular glycerol

Other Model(s): Extracellular glucose 6

Do Not Merge
Future issues in merging of SBML models

- **How to annotate models**
  - Be explicit
  - Avoid lumped substances and reactions or annotate them carefully

- **Our "to do" list for semanticSBML**
  - Support for SBO terms
  - Improved internal database for element matching
  - Automatic recognition and annotation of kinetic rate laws
  - Smarter ways to handle biological qualifiers

- **Community efforts needed**
  - Agreement about annotation of protein complexes and modifications
  - Standard format for graphical representation (support by software)
Acknowledgements

Computational Systems Biology
MPI for Molecular Genetics

Edda Klipp
Falko Krause
Jannis Uhlendorf
Timo Lubitz
Marvin Schulz
Anselm Helbig
Firdevs Topcu-Alici
Anna Krivchanskaya
Hatice Celik