SemanticSBML: a tool for annotating, checking, and merging of biochemical models in SBML format

Wolfram Liebermeister, Falko Krause, Jannis Uhlendorf, Timo Lubitz, and Edda Klipp

Humboldt-Universität Berlin, Theoretical Biophysics, Invalidenstr. 42, 10115 Berlin

lieberme@molgen.mpg.de http://www.molgen.mpg.de/~lieberme

Introduction

Merging of cell models is one of the main future challenges in systems biology. Many structural and dynamic models are available in the XML-derived exchange format SBML[1] and stored in model databases like BioModels.net[2]. Computer-assisted merging of such models would considerably facilitate building of large-scale models. A central task in model merging is the matching of individual model elements: to allow for automatic comparisons, the biological meaning of model elements has to be specified by computer-readable annotations.

Our tool semanticSBML[7,8] assists the user in annotating, checking, and merging of SBML models; it also allows to create SBML models directly from lists of biochemical reactions (specified, e.g., by KEGG[3] IDs). For comparison of model elements, the program relies on semantic annotations in a standard format (MIRIAM[4]-compliant RDF annotations with bioqualifiers[5]).
SemanticSBML is developed under the GNU public license and can be run and downloaded at www.semanticsbml.org.

MIRIAM-compliant annotations

The MIRIAM standard[4] ("Minimum information requested in the annotation of biochemical models") is a set of rules to ensure the quality of systems biology models.

Among other things, MIRIAM describes rules for machine-readable, unique descriptions of biological objects. These rules are realised in the SBML format in the form of RDF (Resource description framework) annotations[5].
An SBML/MIRIAM annotation consists of three parts: the biological object to be described, a reference to an object in an external database and a qualifier that specifies a relation between the described object and the referenced object.

```
<species name="ATP" compartment="c70101" id="s113592"
metaid="metaid_s113592">
<annotation>
        <rdf:RDF ... iers/" >
        <rdf:Description rdf:about="#metaid_s113592">
        <bqbiol:isVersionOf>
        <rdf:Bag>
        <rdf:li rdf:resource="urn:miriam:reactome:REACT_15422"/>
        </rdf:Bag>
        </bqbiol:isVersionOf>
        </rdf:Description>
        </rdf:RDF>
</annotation>
...
</species>
```

Annotated SBML tag for molecule species ATP

Object	Qualifier	Reference
ATP	isVersionOf	urn:miriam:reactome:REACT_15422

Workflow for model annotation and merging

Biochemical models in SBML format

SBMLannotate

Helps the user to annotate model elements (compartments, metabolites, and reactions) with SBO (systems biology ontology) terms and unique identifiers from various databases.

SBMLcheck

Tests if SBML models are consistent with their annotations.

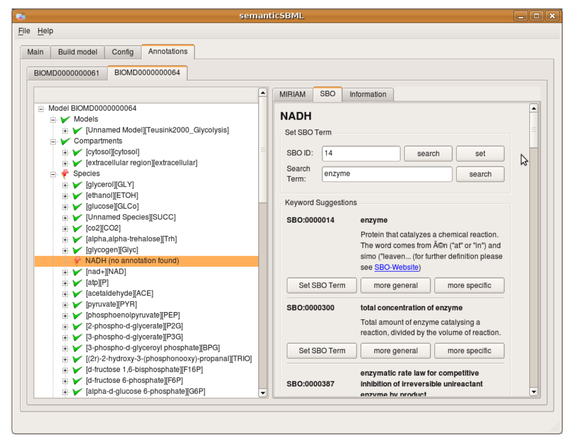
SBMLmerge

Helps the user to combine SBML models and to detect possible conflicts between them.

Merged SBML file

Model annotation

SBMLannotate allows the user to add and modify SBO terms and SBML/MIRIAM annotations: model elements are linked to unique identifiers (IDs) from a variety of databases. For a number of databases (currently GO, KEGG, ChEBI, PubChem, 3DMET, and CAS), IDs are stored internally, which allows users to find annotations by a string search. The program can also suggest annotations, e.g., based on the name attribute of a species element.

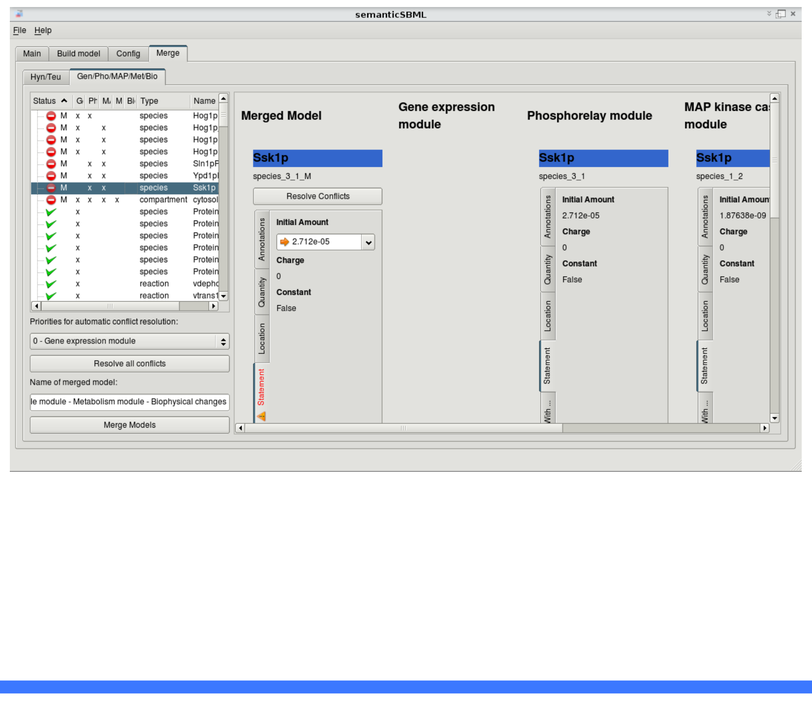


Model checking

SBMLcheck verifies whether the model contains readable and consistent annotations. Elements without annotations are reported to the user. Furthermore, SBMLcheck detects mistakes in the model structure, such as circular dependencies among algebraic equations ("assignment rules") or violations of the mass conservation constraints in chemical reactions.

Model merging

SBMLmerge allows you to merge several models: during merging, it detects and resolves various kinds of syntactic and semantic conflicts, such as conflicting variable names, elements that appear in several input models, or mathematical problems arising from the combination of equations.



For an initial automatic matching, the program searches the input models for matching elements.
SBMLmerge relies on the annotations in the model; however, the user can also match elements differently.
If two models make contradicting statements about the same biochemical quantity, the user has to choose between the different possibilities.
Afterwards, the program merges the models and resolves conflicts that have occurred during the merging process. The result is a new, valid SBML model, which can then be visualised as a reaction network graph[6].

Financial support by

[1] Finney *et al.* (2003) **Systems Biology Markup Language (SBML) Level 2: Structures and Facilities for Model Definitions**

[2] **http://biomodels.org/**

[3] Kanehisa *et al.* (2006) **From Genomics To Chemical Genomics: New Developments KEGG (Nucleic Acid Res.) 34, D354 - 357**

[4] LeNovere *et al.* (2005) **Minimum information requested in the annotation of biochemical models (MIRIAM) (Nat. Biotechnol.)**

[5] LeNovere *and Finney* (2005) **A simple scheme for annotating SBML with references to controlled vocabularies and database entries**

[6] **http://www.graphviz.org**

[7] Liebermeister W., Krause F., Klipp E. (2008), **Merging of systems biology models with semanticSBML**, 5th Workshop on Computation of Biochemical Pathways and Genetic Networks

[8] Krause F., Uhlendorf J., Lubitz T., Klipp E., Liebermeister W., **SemanticSBML, a program for the annotation and merging of SBML models**, in preparation