

Model merging with semanticSBML

Wolfram Liebermeister, MPI-MG Berlin
Computational Systems Biology

SBML composition meeting
UCHC, September 9-10 2007

What is semanticSBML?

Tool for handling SBML models;
focus on “semantic” aspects (annotations!!)

SBMLannotate

SBMLannotate

check, add, and modify annotations
(“MIRIAM” style annotation tags)

SBMLcheck

SBMLcheck

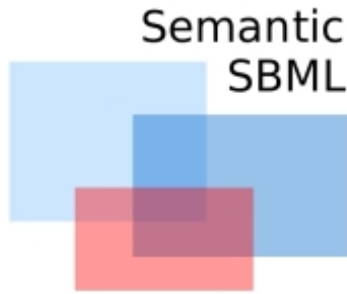
apply model checks based on annotations
(e.g., balances of atom numbers)

SBMLmerge

SBMLmerge

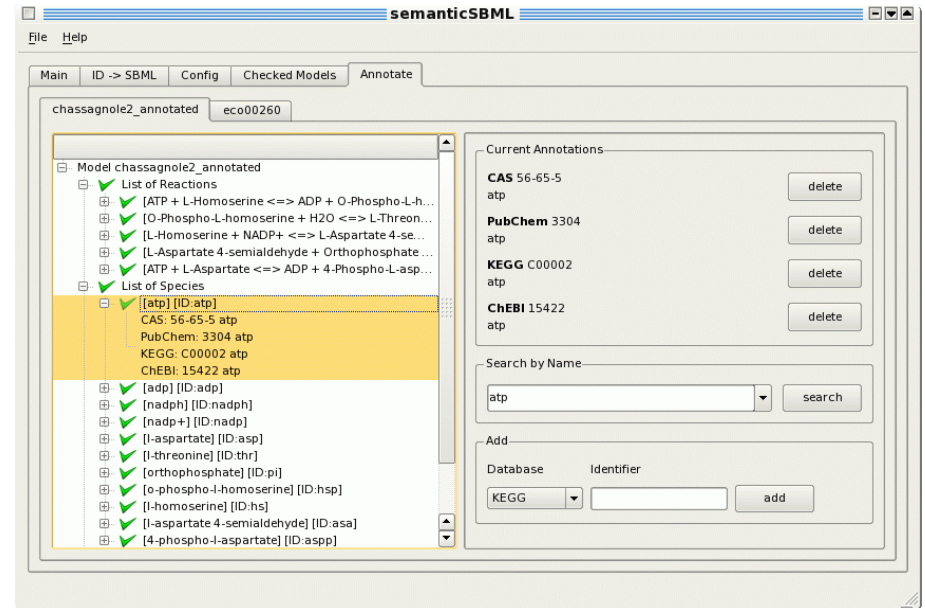
- combine (=fuse) two SBML models
- detect conflicts between them
- resolve conflicts or produce a warning

Current state of semanticSBML



semanticSBML

- written in python, based on libSBML
- GUI + command line tool



latest version

- semanticSBML-0.9.3 at <http://sysbio.molgen.mpg.de/semanticsbml/>
- requires python2.4, some python packages, libSBML 2.3.4, QT4, graphviz

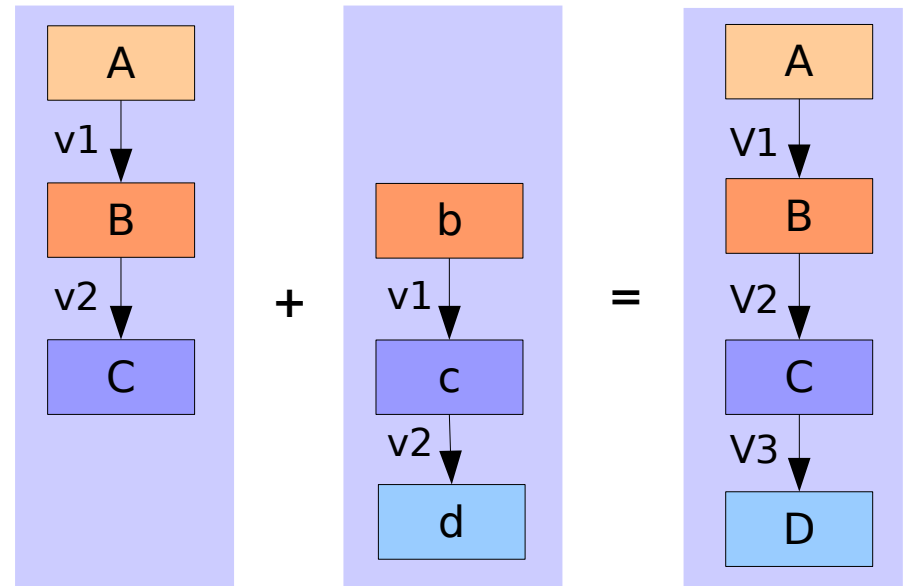
next version (November??)

- uses libSBML 3.0
- new internal data structures
- support qualifiers in annotations ("version of", ...)

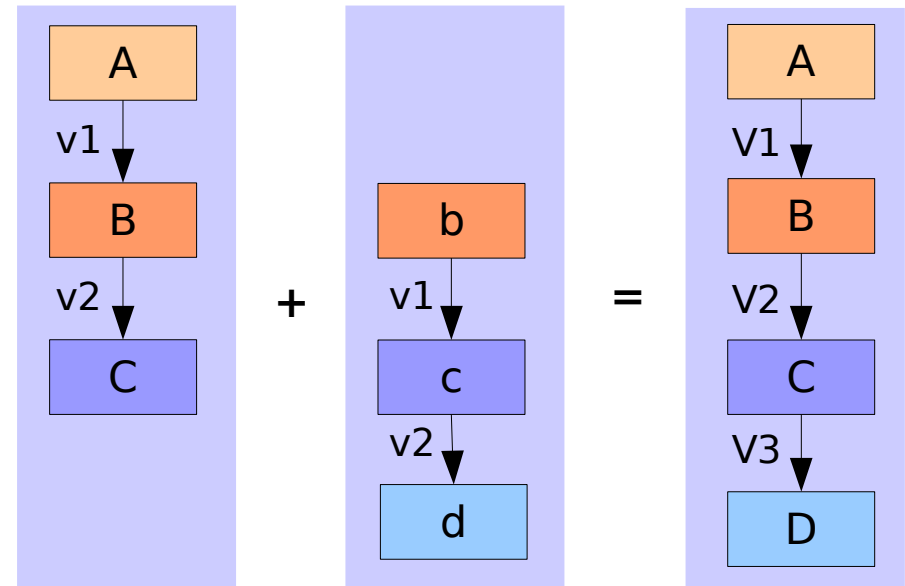
online semanticSBML

update to new version -> access currently blocked

Our view of model merging



Our view of model merging

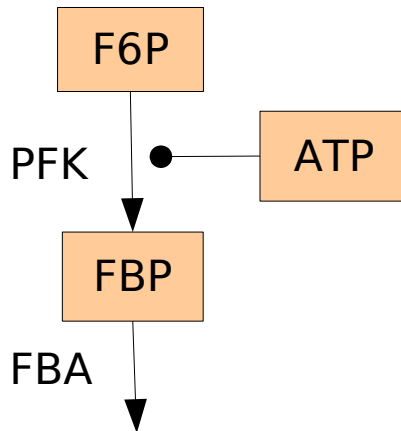


DIFFICULTIES and solutions

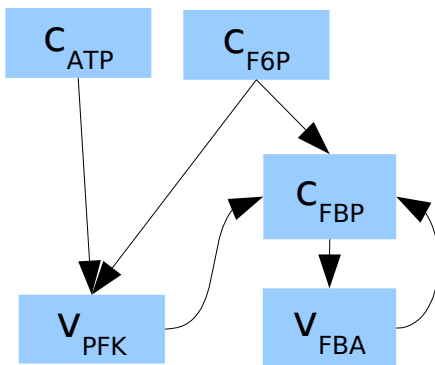
- Incompatible names
compare elements
by annotations, not by model IDs
- Comparing the elements
use database for comparing annotations;
find duplicates and conflicting elements
- Conflicting elements
must not appear together in a model -> abandon merging
- Incompatible statements for duplicates
user chooses between conflicting statements
- Computational cycles
do not allow algebraic rules;
constrain the choice between statements

Semantic models: statements and semantics

Biological entities



Mathematical elements



Model		
ATP conc. [mM]	c_{ATP}	$C_{ATP} = 0.5$
F6P conc. [mM]	c_{F6P}	$C_{F6P} = 0.1$
FBP conc. [mM]	c_{FBP}	$dc_{FBP}/dt = v_{PFK} - v_{FBA}$ $c_{FBP}(0) = c_{F6P}$
PFK vel. [mM/s]	v_{PFK}	$v_{PFK} = v_{PFK}(c_{F6P}, c_{ATP})$
FBA vel. [mM/s]	v_{FBA}	$v_{FBA} = v_{FBA}(c_{FBP})$

Biochemical quantity
= element semantics
determined by "MIRIAM
style" annotations

Mathematical statement
"algebraic rules"
 $0=f(a,b,c,...)$ not allowed!!

Mathematical variable
represents one or several SBML tags

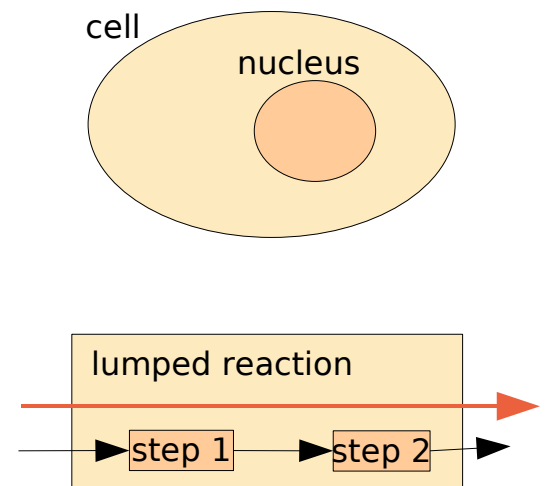
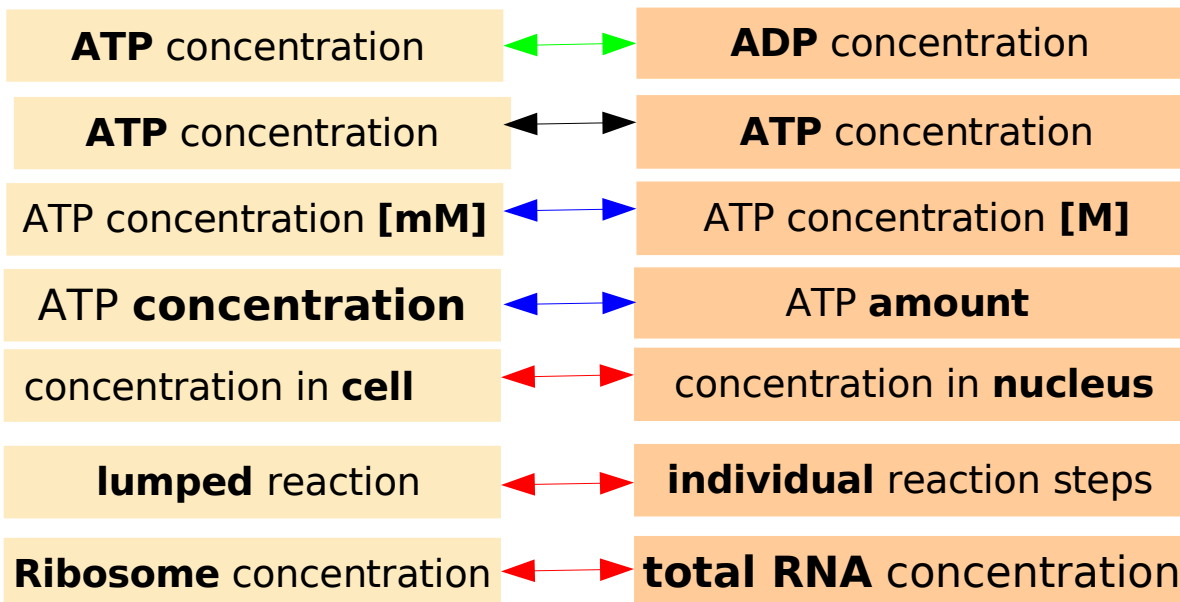
Semantic comparison of elements

Relations between element pairs:

- independent -> **no conflict**
- identical -> **conflict**; choose between statements
- interconvertible -> **conflict**; need to be converted in advance
- semantic overlap -> **severe conflict**; models cannot be merged

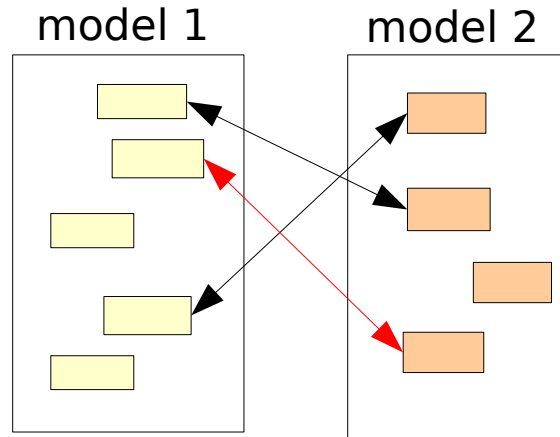
Elements from model 1

Elements from model 2



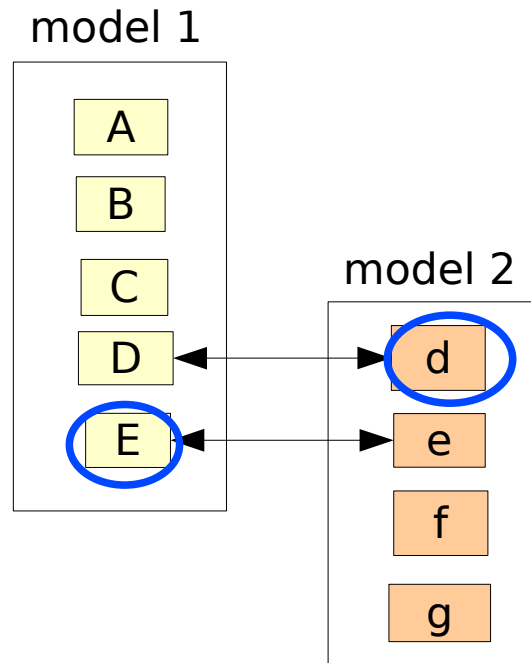
Main steps in SBMLmerge

1. Semantic comparison



Semantic overlap!!
failed after comparison
-> no merging possible

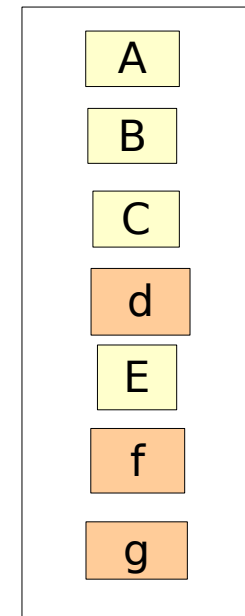
2. Choice of statements



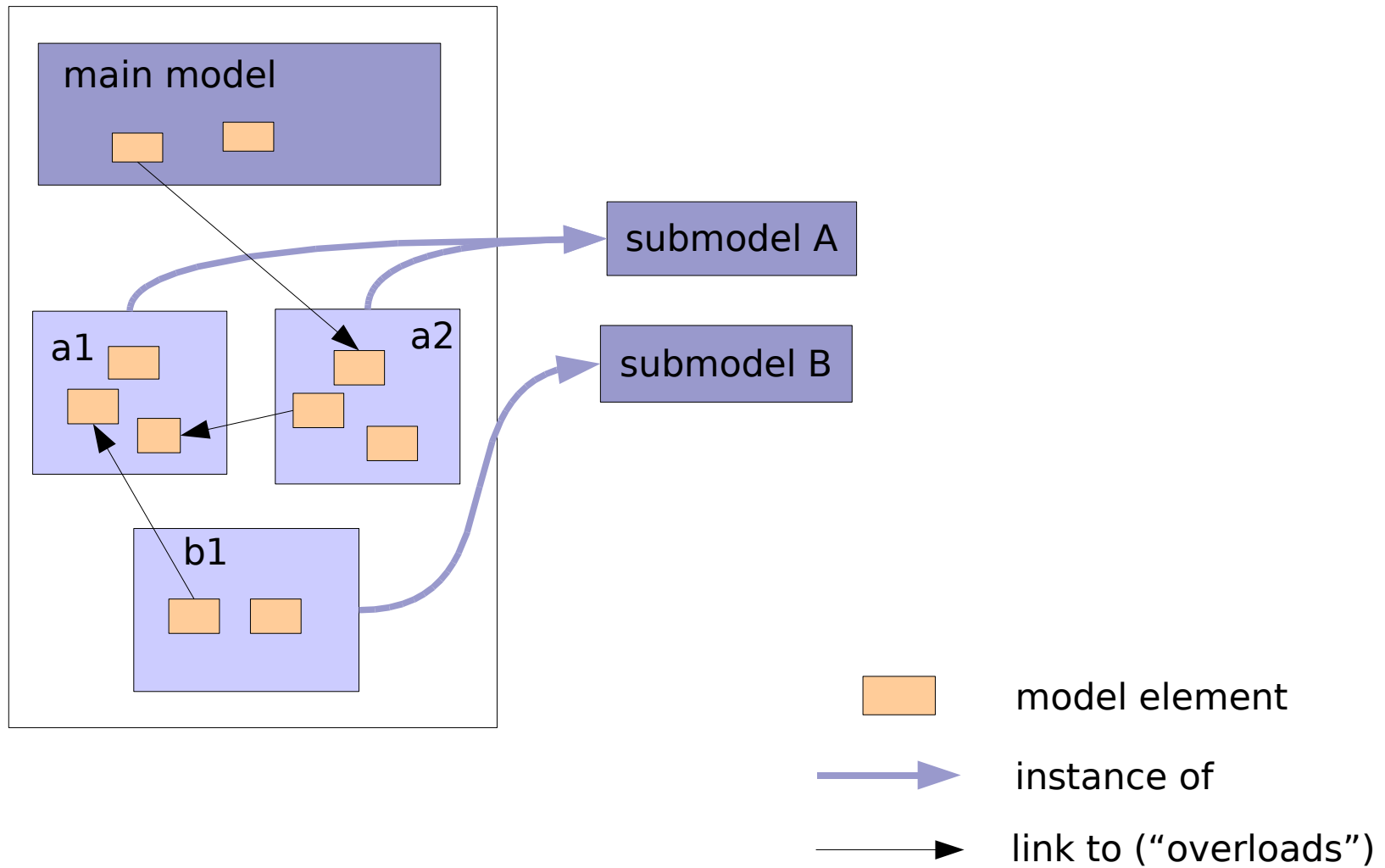
**choose
statements**



merged model



Model composition (Andrew's proposal)



What's the semantics of a composed model?

1. Multiple submodel instances represent “different things” (??)

Requires annotations for the submodel instances

- an annotation for spatial regions
- attributes or convention saying “no semantic overlap between instances”

-> if there is dependence, links must be set!!!

2. Links represent “identical things” (!!)

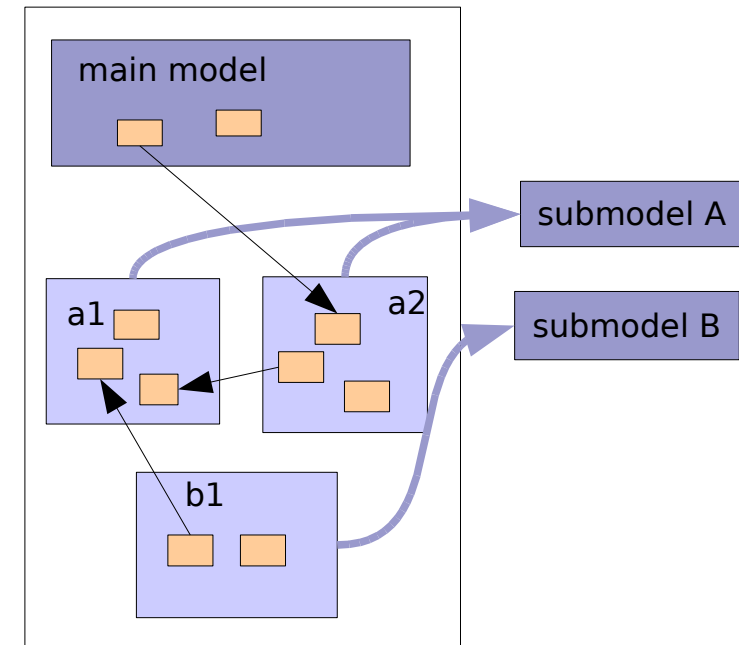
Sensible constraints for links

- No cycles of links
- No converging links

Linked elements should have compatible semantics (... by annotations...)

Overloading by links should affect ...

- ... MIRIAM annotations
- ... other semantics-related annotations (CellDesigner???)
- ... assignment rules, rate rules, etc. belonging to species



A future SBMLmerge could use links and model aggregation

Core SemanticSBML:

Semantic comparison of input models

- overlap -> stop merging
- identity -> count as duplicate

User chooses between duplicates
(avoid computational cycles)

Current flat merging

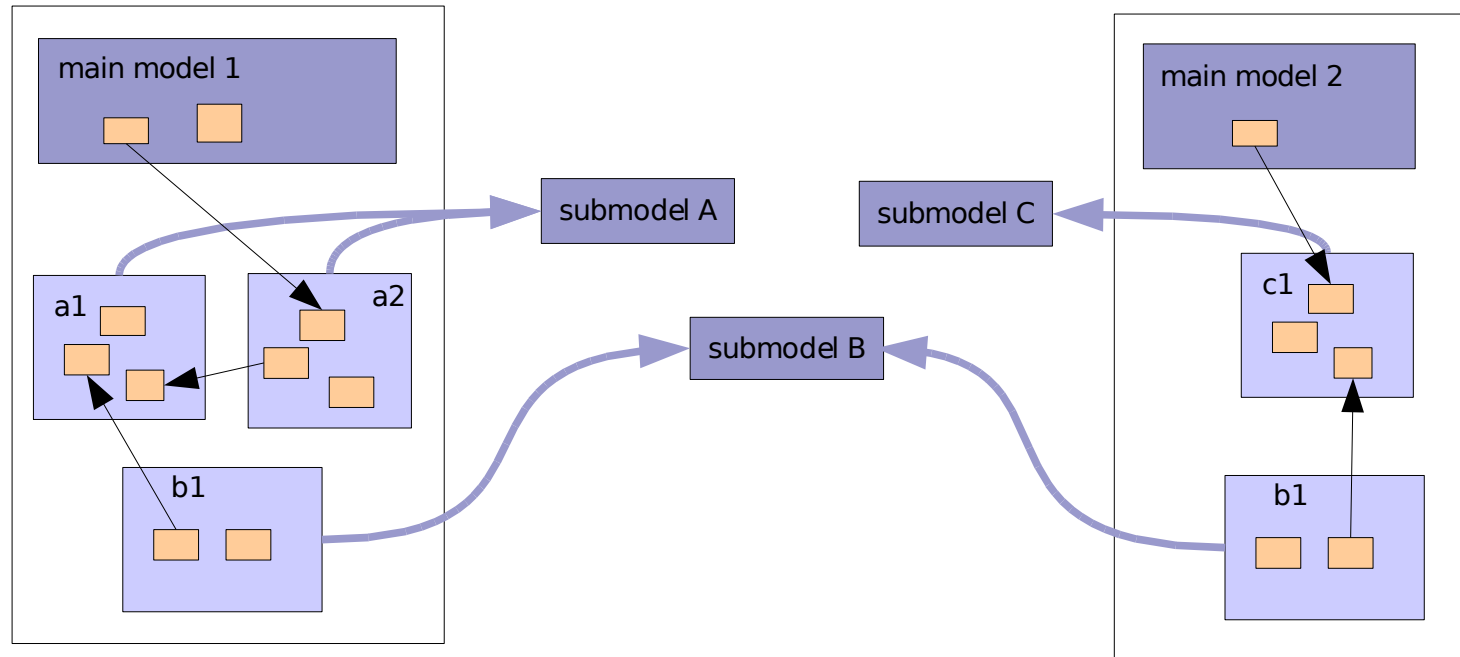


- Build **flat SBML file** with all elements
- **Remove** discarded elements from duplicate pairs.

Alternative: aggregation

- Include input models as **instances of submodels** into SBML file
- For each duplicate pair, **overload** discarded element by chosen element

A future SBMLmerge for composed models ??



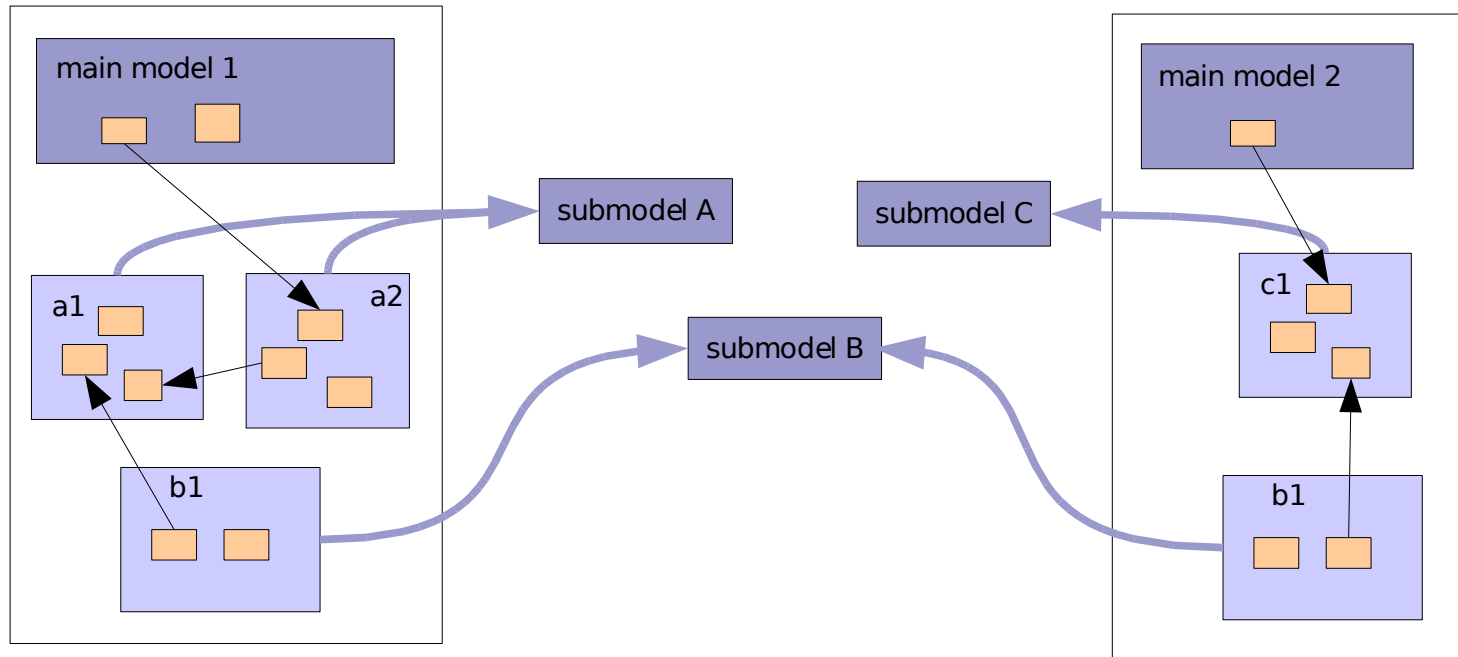
Modularity can save time in semantic comparison

- element comparison for main/submodels, **not for instances**
- **no comparison** necessary for common submodels

To be specified by annotations:

- are instances of submodel B identical/overlapping or not??
(... depends on the intended location ...)
- If overlaps are found between submodel A and C:
are their instances a1, a2 \leftrightarrow c1 overlapping or not?

How could a future SBMLmerge handle links ??



Disregard overloaded objects

- in semantic comparison
- in statement choice

Will link constraints be violated during merging?

- If cycles appear -> flatten some of the linked elements
- convergent links -> user has to choose

Acknowledgements

MPI für molekulare Genetik

Edda Klipp
Falko Krause
Jannis Uhlendorf
Marvin Schulz
Anselm Helbig

EML Heidelberg

Isabel Rojas, Martin
Golebiewski, Ralph Gauges

Humboldt-Universität Berlin

Ulf Leser

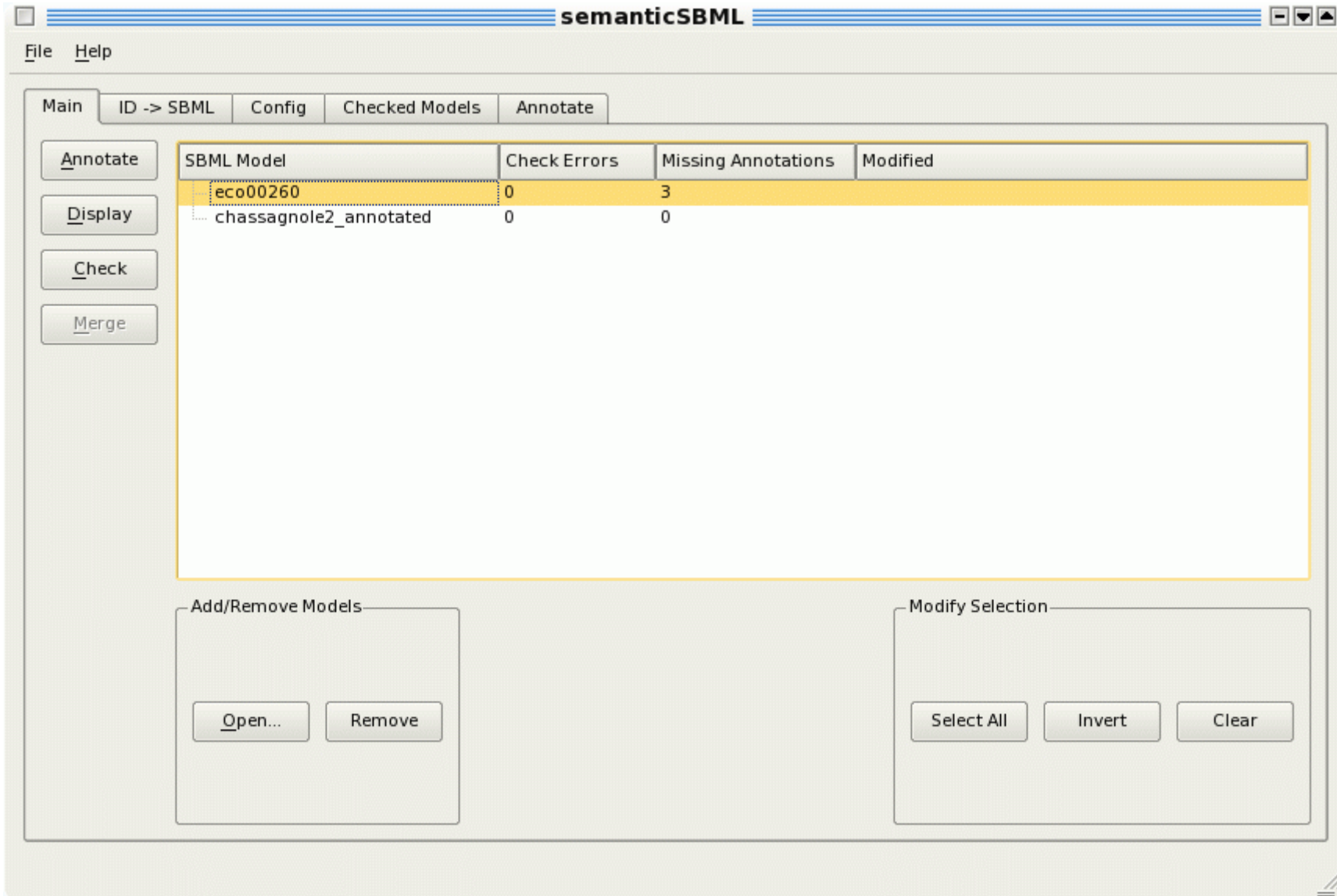
Universität Heidelberg

Ursula Kummer

Funding by European integrated project BaSysBio



semanticSBML



semanticSBML

The screenshot displays the semanticSBML application interface. The main window is titled "semanticSBML" and contains a menu bar with "File" and "Help". Below the menu bar are several tabs: "Main", "ID -> SBML", "Config", "Checked Models", and "Annotate". The "Annotate" tab is active, showing a model named "chassagnole2_annotated" with a selected identifier "eco00260".

The central panel is divided into two main sections. On the left is a tree view of the model's components:

- Model chassagnole2_annotated
 - List of Reactions
 - [ATP + L-Homoserine <=> ADP + O-Phospho-L-h...
 - [O-Phospho-L-homoserine + H2O <=> L-Threon...
 - [L-Homoserine + NADP+ <=> L-Aspartate 4-se...
 - [L-Aspartate 4-semialdehyde + Orthophosphate...
 - [ATP + L-Aspartate <=> ADP + 4-Phospho-L-asp...
 - List of Species
 - [atp] [ID:atp] (highlighted)
 - CAS: 56-65-5 atp
 - PubChem: 3304 atp
 - KEGG: C00002 atp
 - ChEBI: 15422 atp
 - [adp] [ID:adp]
 - [nadph] [ID:nadph]
 - [nadp+] [ID:nadp]
 - [l-aspartate] [ID:asp]
 - [l-threonine] [ID:thr]
 - [orthophosphate] [ID:pi]
 - [o-phospho-l-homoserine] [ID:hsp]
 - [l-homoserine] [ID:hs]
 - [l-aspartate 4-semialdehyde] [ID:asa]
 - [4-phospho-l-aspartate] [ID:aspp]

Current Annotations

CAS 56-65-5
atp

delete

PubChem 3304
atp

delete

KEGG C00002
atp

delete

ChEBI 15422
atp

delete

Search by Name

atp

search

Add

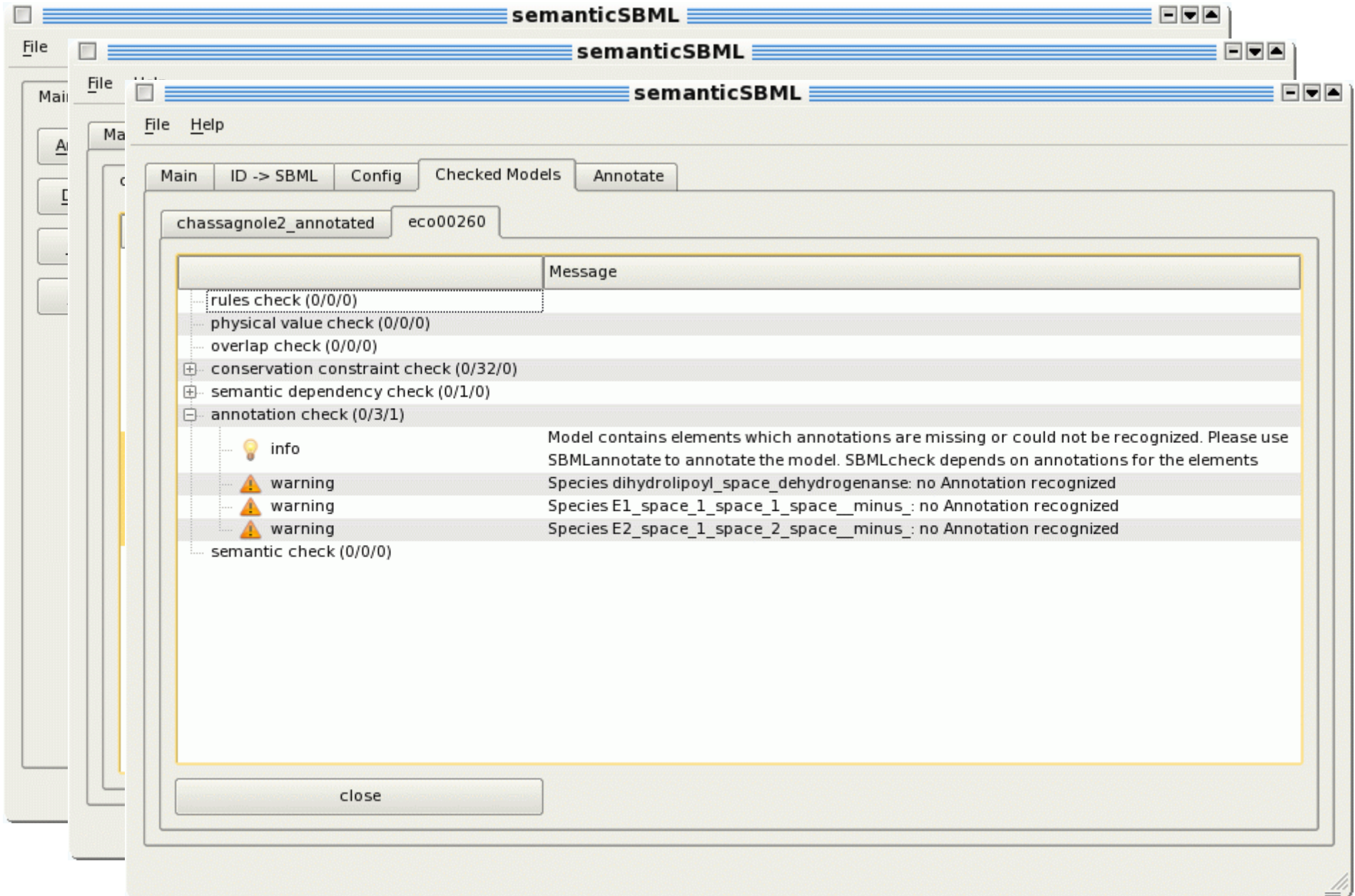
Database

Identifier

KEGG

add

semanticSBML



Harmless conflicts during model merging

INCOMPATIBLE STATEMENTS (about the same quantity)

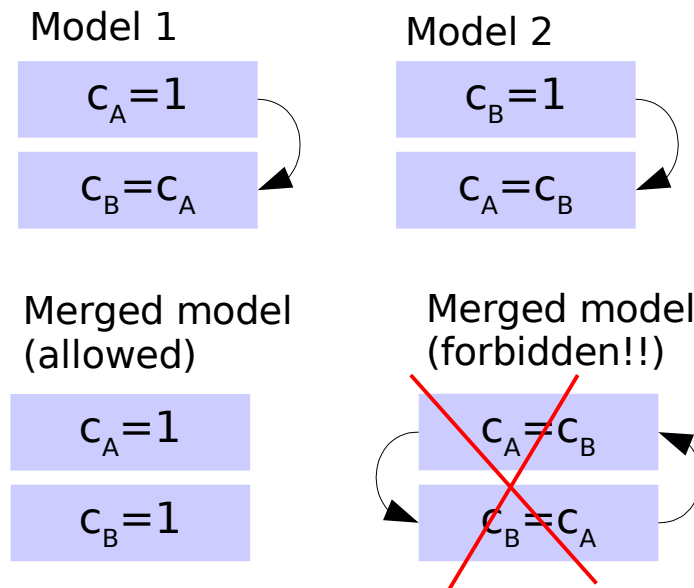
Model 1 says	FBP conc. [mM]	c_{FBP}	$\frac{dc_{\text{FBP}}}{dt} = v_{\text{PFK}} - v_{\text{FBA}}$ $c_{\text{FBP}}(0) = c_{\text{F6P}}$
Model 2 says	FBP conc. [mM]	c_{FBP}	$c_{\text{FBP}} = 0.5$

Choose one of them (or choose a third, alternative one)

COMPUTATIONAL CYCLES

Problem: algebraic equations refer to variables that have not been evaluated before; -> model is not computable

Computational cycles are forbidden!
 Can they always be avoided?
 Yes; choose all statements from model 1



SBMLmerge: current features

- the output model describes all elements of the input models
- for each element, a statement is chosen from an input model
- no conflicting statements are made
- the solution is not unique;
choices are made based on rules or by the user
- the output model is computable (no cycles)
- several models can be merged subsequently

SBMLmerge: future features

To be supported in coming versions:

- events, species types, SBO terms
- more database identifiers
- qualifiers in annotations (“version of” etc) -> detection of overlap
- annotations for things implicitly described by the model
- SBML standards for model composition (???)

Not supported (...or with big warning signs ...)

- algebraic rules, such as “ $f(a,b,c,d)=0$ ”
- constraints
- incompletely annotated models
- semantically overlapping elements