

Minutes from the 9th SBML Forum, Heidelberg, Oct 14-15, 2004

Recorded by Joanne Matthews

14th October 2004

Introduction and Overview (Mike Hucka)

Gave a brief overview of the days agenda. Delegates introduced themselves.

Quick overview of the state of SBML (Mike Hucka)

All specifications, FAQs, wiki, discussion list, papers can be found online at SBML.org. Current acceptance level of L2V1 is that 60–70 tools support SBML. Details are on the website. Previous versions of SBML = L1V2, L1V1.

The SBML process continues to be informal. AF and MH organise and manage activities. SBML forum and the main channels for discussion was described. SBML Team members introduced. Funding agencies acknowledged.

Software for stochastic reaction diffusion simulation (MesoRD)

Presentation by Johan Elf. More information can be found here:

<http://mesord.sourceforge.net>

Python simulator for cellular systems (PySCeS)

Presentation by Johann Rohwer. Tool for modelling cellular systems. Coded in Python. Extended using the Python scipy package.

Can find the steady state using a few different solvers like FINTSLV. Limited to ODE simulations (solved using LSODA). More information can be found here:

<http://pysces.sourceforge.net>.

CellDesigner and hardware accelerated simulator

Akira Funahashi Presented new release 2.0 of CellDesigner as well as a new project to develop a hardware accelerator for simulations, based on FPGA boards.

CellDesigner download available from <http://systems-biology.org>. The FPGA-based accelerator allows much faster simulation of stochastic models.

BioUML

Fedor Kolpakov presented BioUML. Written in Java and consists of a number of tools to enable modelling of biological systems. Also has a tool for formally describing system, tools for visualisation and simulation, and two databases:

Cyclonet database www.cyclonet.biouml.org a database on cell cycle regulation in eukaryotes, and BioPath, <http://www.biopath.biouml.org>

Can use SBML models using SBML plug in in BioUML workbench.

SBMLToolbox

Sarah Keating presented the SBMLToolbox. Download from <http://SBML.org>. SBMLToolbox is intended to let Matlab users interact with SBML models more directly.

SBML support in E-Cell

Takeshi Sakurada presented an update on recent work in E-Cell. E-Cell is a tool for modelling, simulation and analysis. More information can be found here www.e-

cell.org/software . SBML models are supported but requires libSBML. All 3 levels and versions of SBML are supported. They also have a new web site for the E-Cell community: www.ecdn.e-cell.org
E-Cell models in version 2 will be converted to version 3.

SBML arrays in MathSBML and the Computable Plant Project

Bruce presented a proposed implementation of SBML arrays in MathSBML. Arrays are one of the possible modules being considered for inclusion in SBML Level 3. See <http://www.SBML.org/wiki/arrays> for more on arrays. See <http://SBML.org/software/mathSBML> for more on MathSBML, a tool built for Mathematica. Bruce also described the computable plant project; see www.computableplant.org

From Logical Regulatory Networks to Petri Nets to SBML

Claudine Chaouiya presented her current work.

The Petri Net community has developed many tools here:

www.daimi.au.dk/tools/Petrinets/tools

PNML is the Petri Net Markup Language.

Andrew Finney asked: why not use something like GML to SBML?

Claudine answered: PNML is more intuitive and tailored for Petri Nets.

Model integration into SBML using BioPAX Ontology

Jeremy Zucker presented BioPAX = Biological Pathway Exchange Ontology, for representing pathway knowledge.

BioPAX hackathons, forums, conference are planned. See www.biopax.org and Biopax-discuss@biopax.org

BioNetGen

Michael Blinov presented BioNetGen, a rule-based reaction networks generation tool. See

<http://cellsignaling.lanl.gov/bionetgen>

Layout proposal

Ralph Gauges summarized the current version of the layout proposal.

Significant changes are:

Removed semantic overload for bounding box. (No way to determine what the reaction glyph would look like.)

MH asked: Has the issue re bounding box been resolved with Herbert Sauro?

RG said: Some people didn't want to provide dimensions but it is not possible to automatically generate it, so have agreed that it stays in.

Current definition of layout proposal is available at

<http://projects.eml.org/bcb/SBML>.

LibSBML

Ben Bornstein presented LibSBML. Current version is 2.2. It has had more than 100 downloads of the latest version from <http://SBML.org/software/libSBML>.

Among its new features is consistency checking of models. The current checks are a subset of the ones suggested on the SBML.org wiki. Users of libSBML are encouraged to subscribe to SBML mailing list SBML-discuss.

Discussion: Someone recommended a consistency check to prevent a species existing that wasn't involved in a reaction.

MH: What kinds of APIs are required for handling annotations?

Ben: Annotations are a vexing problem. One way to do it would be to set up SAX events handled by the calling application. What about a DOM-based solution? Well, the problem there is that Expat doesn't provide a DOM. Also, using a DOM-based solution would potentially run into issues of memory footprint. One possibility is that libSBML could implement just enough DOM to allow annotations support.

Ben asked: Should the model checking/validation be part of the library or a separate tool? Responses were mixed.

Discussion: Should the SBML standard be relaxed to allow reordering of annotations?

Announcements

Rick Paxson from MathWorks: The Matlab Systems Biology toolbox will include tools for mathematical modelling and analysis. Release date expected in the first or second quarter of 2005.

Timo Hannay from Nature Publishing Group: NPG and EMBO are launching a new journal, Molecular Systems Biology. Looking for ideas to provide added value. E.g., one might be to provide support for SBML. Also providing search facilities on published models. Timo would appreciate feedback regarding what readers want. MSB will be an online-only publication and free to access, but will use the author-pay model.

Someone else announced: The European Conference on Mathematical and Theoretical Biology - ECMTB05 Germany, <http://www.ecmtb05.org/>.

SBML Semantic Test Suite

Andrew Finney presented the SBML Semantic Test suite. The tests are designed to check the interpretation of SBML, to ensure that different tools interpret SBML in the same way. The test suite consists of models and simulation output

<http://www.cds.caltech.edu/~afinney/semantic-test-suite.tar.gz>

Documentation can be found on the SBML wiki here:

http://sbml.org/wiki/Semantic_Test_Suite

AF and MH suggested a future enhancement would be to use the test suite online by providing a way for user to submit results for their tool. The server side could plot results, display a chart detailing which tools pass each test, etc.

Discussion:

Have to be careful to test the tool's interpretation of SBML and not its simulation performance

The validation suite is needed because there needs to be some sort of benchmark to validate tools. But the results have to be compared to some accepted "gold standard". Currently this is MathSBML. Is there enough confidence in MathSBML as the gold standard? (This is a question that can't be answered yet.)

The validation suite currently uses wrappers to drive tools. Does there need to be a way (maybe a web site) for people to register wrappers for their tools?

The test suite can be useful for developers. Testing if numerical outputs are reasonably accurate can be useful for developers as they're developing the tools. The reference data/results would need to be curated.

How to encourage building the suite? One way might be to do reviews of other people's simulator by pairing up with someone else.
There needs to be a review process in any case.
Another point: some differentiation might be appropriate between different simulators.
Jeremy Zucker offered to provide code of example basic test features.

15th October 2004

Overview of the Day's Agenda

Mike Hucka outlined the day's agenda.

Merging and Visualisation of SBML Models

Karthik Raman presented his proposal for merging KEGG and SBML models

<http://www.systems-biology.org/001>

Scripts available for cygwin at

<http://rishi.serc.iisc.ernet.in/~kathik/sbml/sbmlmerge.htm>

Jeremy Zucker suggested reactions in KEGG are actually not balanced reactions—need to watch out for that when formulating models from KEGG.

Coherent organisation of the model would be good.

Suggested central bug reporting place on KEGG.

Model composition proposal could be used to address this use case of merging models?

Another issue: exact name matching on specie name based on id. Using EC numbers (Enzyme commission numbers) would help but they're not always available. Article by Rob Pinter addresses this problem.

Pathway information can be available for every genome. BioPAX aims to provide this ability. BioCyc and BioPAX provides for EC number even if the enzyme is not associated with it. Thermodynamic section on NIST provides database of EC number.

Modularization of SBML

Jörg Weimar presented his proposal for the modularization of SBML. Motivation: no tool supports all the features provided by SBML; this proposal addresses a mechanism to specify partial support. The proposal suggests producing XML sub-schemas that is restricted to a subset of features, then letting individual tools create “metamodules” that group some of the features together to describe what the tool understands. Using UML modelling techniques, JW identified > 1009 dependencies for L1 and >5657 dependencies for L2. Certain modules always use the same subset of features so could be possible to modularize schema. Metamodules could be employed to compose schema bits that you can break up.

Discussion:

Must be careful to support those features with required dependencies.

Would it be possible to handle extensions to schema for Level 3, i.e., by making packages be a set of modules?

Must be specific about terminology. Ben Bornstein suggested “aspects” as a replacement name for modules.

Another issue: deeper semantic stuff like what sort of stoichiometric units are in use. The approach offers the possibility that every tool could define an associated schema indicating what it is capable of interpreting.

Need to consider advantages of many schemas versus just a list of supported features. Subsets would be useful for consistency checking, allowing common subset validation. But would this lead to proliferation of xsd files and problems managing them? Version problems with xsd's. Ben Bornstein suggested putting xsd's into code. Jörg suggests that the xsd's be automatically generated, maybe with some hand tuning afterwards. Jeremy Zucker suggested using the Jaxb package with a model could be used to generate code to define the classes.

Nicholas Le Novère suggested having a tree of modules. The top level set of modules would be the current schema. Lower branches would be optional, so that the Xsd file can be like a package.

Pedro discussed round-tripping models. No tool can guarantee it. Should the user be prevented from saving a model as SBML if some of the features have been lost? People agreed that a tool that reads something in doesn't have to guarantee to write it out without losing any features.

Semantic and syntactic dependencies are great but must find a way to minimize the (currently very high) number of dependencies.

Another issue: how to store this information – stored vocabularies?

Multicomponent Species: A Proposal for SBML Level 3

Andrew Finney presented his proposal for support of multicomponent species in SBML Level 3. Motivation: current SBML has limitations in support for species with states and components. Proposal for multicomponent species would extend SBML to handle these. Some features of the proposal include: Constraints do not have to be specified somewhere else. Let reaction system define what possible states a reaction can get into. Only specify the binding sites you care about if not relevant to your system don't specify. Rather than list of rules, is there general consensus on how to handle dependant binding and wild cards. Multiple array of binding sites.

Conformational change is defined here as a property of a complex, to avoid problems with hierarchical structure

Discussion: AF feels state variables need to be decoupled from binding sites.

Need also to allow for grouping reactions where some reactions are exclusive to each other, for example like in Molecuizer.

Agree good terminology for different bits.

Clarify semantics, ie more in chemical terms.

Represent species spanning compartments.

Allow writing raw ODEs based on semantic state.

Question: Are groups general over species or molecules? AF: only unit compartment. 4 things that do molecular action multi state activity unaddressed question.

Pedro: Need to know what domains available and where.

AF: Where do you put the unit for a species that is in 3 domains?

Could transitions be specified by rules rather than ODE's?

AF: If properly written, the kinetics constrain you to what really happens.

JZ: BioPAX 2 deals with the issues of states. Encouraged people to join the BioPAX work groups.

G and S question: Consider looking at more graphical stuff.

Jim Shaff: recommends looking at GNS notation for graphical diagrams.

Blinov: Cannot generate complete set of species in advance using rules because some of the information would be lost.

People would like to see a standard graphical notation, but one is not available yet.

There have been proposals such as Kitano's proposal.

Group Discussion: Proposal for L2V2—what to put into v2, what to leave for L3?

We examined the L2V2 proposal in somewhat random order. Start with Section 8.

Dependent variable attribute on AlgebraicRule

Martin Ginkel and Jörg Weimar: we shouldn't have information about dependent variable on algebraic rule—this is information about the numerics. SBML did not specify information about numerical solutions until now.

Pedro: the dependent variable info is not necessary.

JZ: Is it needed for describing metadata of stiff solvers?

Pedro: says that's actually irrelevant, a technicality of the simulator.

Conclusion of group: voted 'no' on dependent variable attribute.

Unitkind and nested unit definitions

Jim Schaff: VCell has extensive list of preferred units that the tool provides.

Convenience, make the model easier to read.

Conclusions of group: voted 'yes' to allow nested unit definitions.

SpeciesType

Without this, can't name a molecule in 2 different places without having subscript in place. Currently can't parse species without knowing the molecule.

Question from Jörg Weimar: does this interact with the multicomponent species proposal?

Another question: what about multiple levels of species types?

Another idea: Could package it in with the modularity proposal.

Conclusion: defer. First need to make sure there will not be unexpected interactions with the (future) multicomponent species proposal and also consider the multiple levels of species type. AF will review the issues and mail sbml-discuss his conclusions.

Diagram layout

Should it be incorporated into L2 directly, rather than waiting to make it a package of L3? AF suggests another reason to keep it as an SBML package is so that it can be developed independently of SBML.

Conclusion: make this an SBML package for L3. In the meantime, for use in L2, people will put layout into annotations (everyone can agree to recognize the appropriate annotations).

Will need to modify libSBML parser to allow diagram layouts to be patched in.

Controlled vocabularies

General point: they need to be developed independently of SBML.

Require generic mechanism to relate controlled vocabulary terms to SBML object

Should tools know about controlled vocabularies in advance.

What approach to take?

JZ suggests OWL like.

But how to relate, e.g., proteins that have many synonyms?

BioPAX doesn't do reaction equations. Limitations due to OWL implementation.

GO has advantages. Could get round the limited tool support as all open source.

GO edits facility is good.

Pedro: Ontology is a set of relationships between named objects. GO is not this.

GO does not contain the set of vocabularies required by systems biology, e.g.,

Michaelis-Menten.

NL: Several vocabularies are needed, e.g., one for kinetic laws.

Conclusion: Significant disagreement on how to define stuff.

Pedro concerned that we risk carving mistakes in stone.
GO resolves this issue by having different versions. Terms don't get deleted.
Keep ontology as snapshots and specify which version you are using.
Biospice is taking approach of defining annotations that define equation types.
SBML contains references to particular terms.
User might want description for term. SBML tells you where that might be. But doesn't specify version information. What happens when terms change? AF suggested that audit trail is used.

Regarding the proposal to extend SBase in support of CVs:

Don't introduce new name space but have a predicate that refers to a resource.

Conclusions: CV deferred. Jörg summarized the sentiments: people need to be more confident in suggested implementation before we can agree on using a particular technology. (OWL or Go?)

Assertions

Discussion: Do assertions belong to model or analyses methods?

MH: could we just use events to implement this? Maybe with small mods to events to extend them in the right way, rather than defining a whole new thing?

Martin Ginkel agrees that assertions belong to models, not analyses.

AF: Crosses analysis boundaries, so is part of the model.

Modellers may want to convey to users what's going on in rate laws.

MH: Maybe events are the wrong thing. OK, what about using rules? This is like identifying rules that, if violated, take some action. Is there really a need for a new concept? What if Rules had messages associated with them? Or maybe could do it using events, by setting a new flag. If it notifies the user without action, then it would be considered an event, otherwise it would be considered a rule.

Someone else: Pre and post conditions for the whole model?

Jörg: Continuous models don't have pre and post conditions; they're just always active. Pre and post conditions thus don't make sense.

Someone else: Would like to express optional rule conditions and message pairs.

Jörg: agrees that assertions are rules, in the sense of constraint rules.

Conclusion: create a new kind of rule, "ConstraintRule", and use the <notes> element of the rule to provide messages to the user.

Flux and variable bounds

Jörg: What should be done if flux goes beyond the stated bounds?

JZ: uses conventions in annotations to solve this problem.

Question: Is it possible to use local parameter to define bounds?

Jörg: With flux it is. Parameters are in kinetic law.

Suggestion: Contemplate introducing some bounds.

Pedro proposes a 4th kind of rule: Rule that add inequality constraints. Constraint rule evaluates to Boolean expression.

Allow expression of probability equations.

Allow reaction ids.

Sign of the equality is determined

Tentative mechanism for implementing flux bounds and how to interpret in general case. No general agreement whether to incorporate. MH proposed to formulate some ideas on this and circulate on SBML-discuss.

Id's on SimpleSpeciesReference

Conclusion: people voted 'yes' to add them as per the proposal.

Id's on SBase

Conclusion: agreed with proposal that id's should not be on SBase.

Hackathon

Options:

April/may timeframe.

Tokyo one option.

Caltech another option.

The upcoming CCB meeting in March another option.

RECOMB at MIT 14–18 may

San Diego FASEB 31st March – 6th April

Announcements

Jim Schaff announced the third CCB computational biology meeting

<http://www.nrcam.uchc.edu>

March 19-23 2005 in Massachusetts.

Jim Schaff announced that Virtual Cell version 4 beta will be released end of this month.

Pedro announced availability of beta version of COPASI at <http://www.copasi.org>