

Recent progress on SBGN in **Arcadia**

Alice Villéger

alice.villegier@manchester.ac.uk

Bioanalytical Sciences Group / Advanced Interface Group
Faculty of Engineering and Physical Sciences
School of Chemistry / School of Computer Science
The University of Manchester

CellML/etc. Meeting, SBGN discussion, Waiheke Island, 2009/04/07

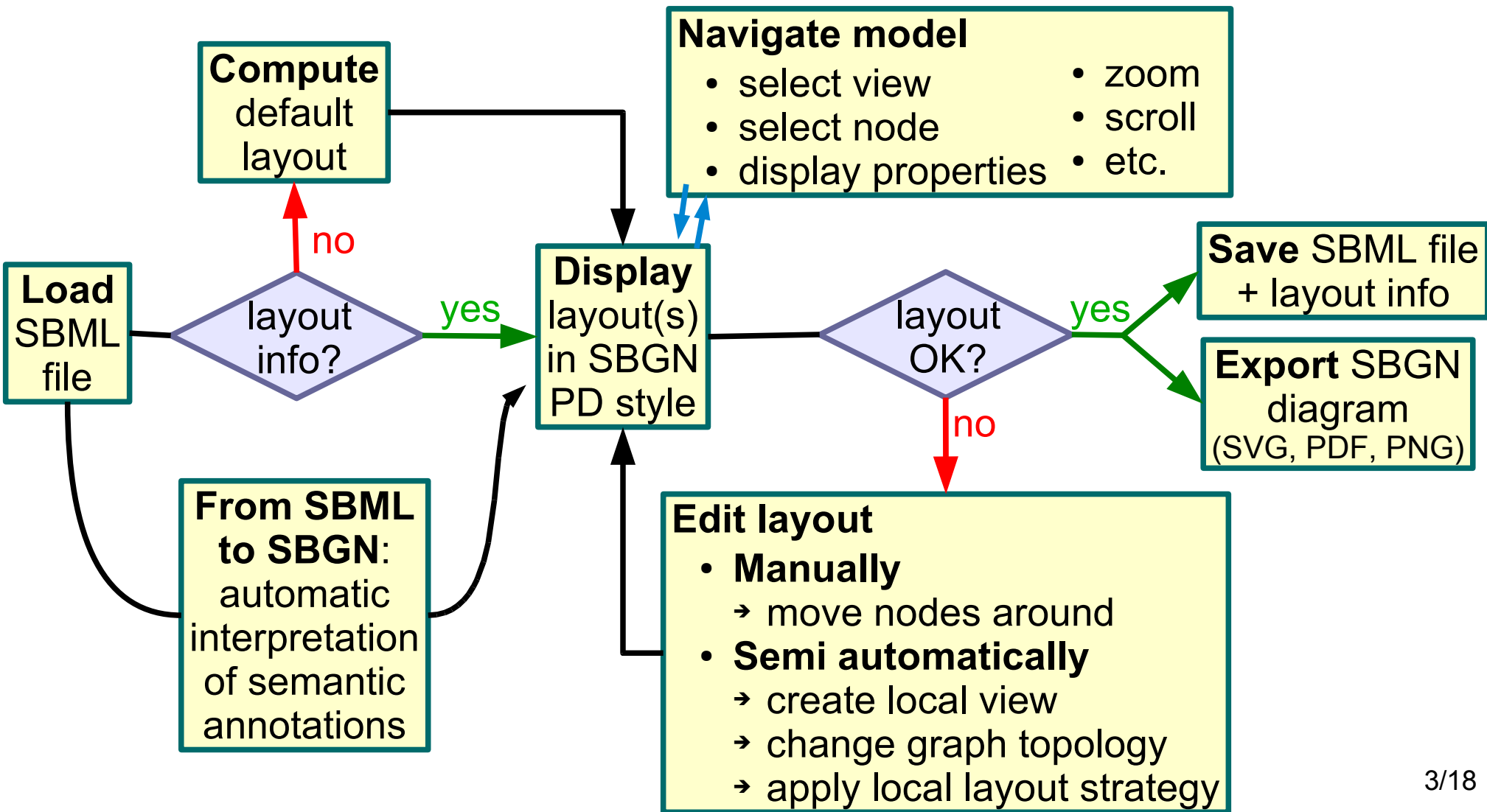
Arcadia in a nutshell

- **What?** systems biology model **viewer**/browser (≠ editor)
- **Why?** to visualise **metabolic pathways** (from small to **HUGE**)
 - **How?**
 - simple **navigation** interface between **multiple views**
 - light-weight, standalone, and **interoperable**
 - supports and promotes **standards** (SBML, SBGN, MIRIAM, SBO, ...)
 - **flexible** Model/View/Controller software architecture
 - **integration** of new file formats, **interaction** with external tools
 - **cross-platform, open-source** implementation (C++, Qt)
 - powered by existing libraries (libsbml, boost, graphviz, libavoid, ...)

most biomodels.net models

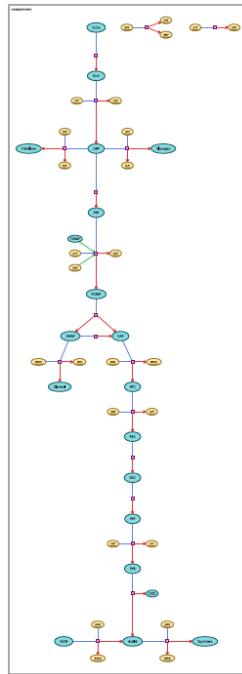
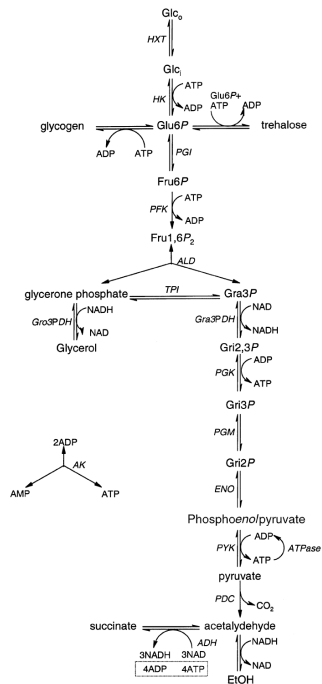
!! Yeast and Human Jamboree

Main features: typical workflow



Current results

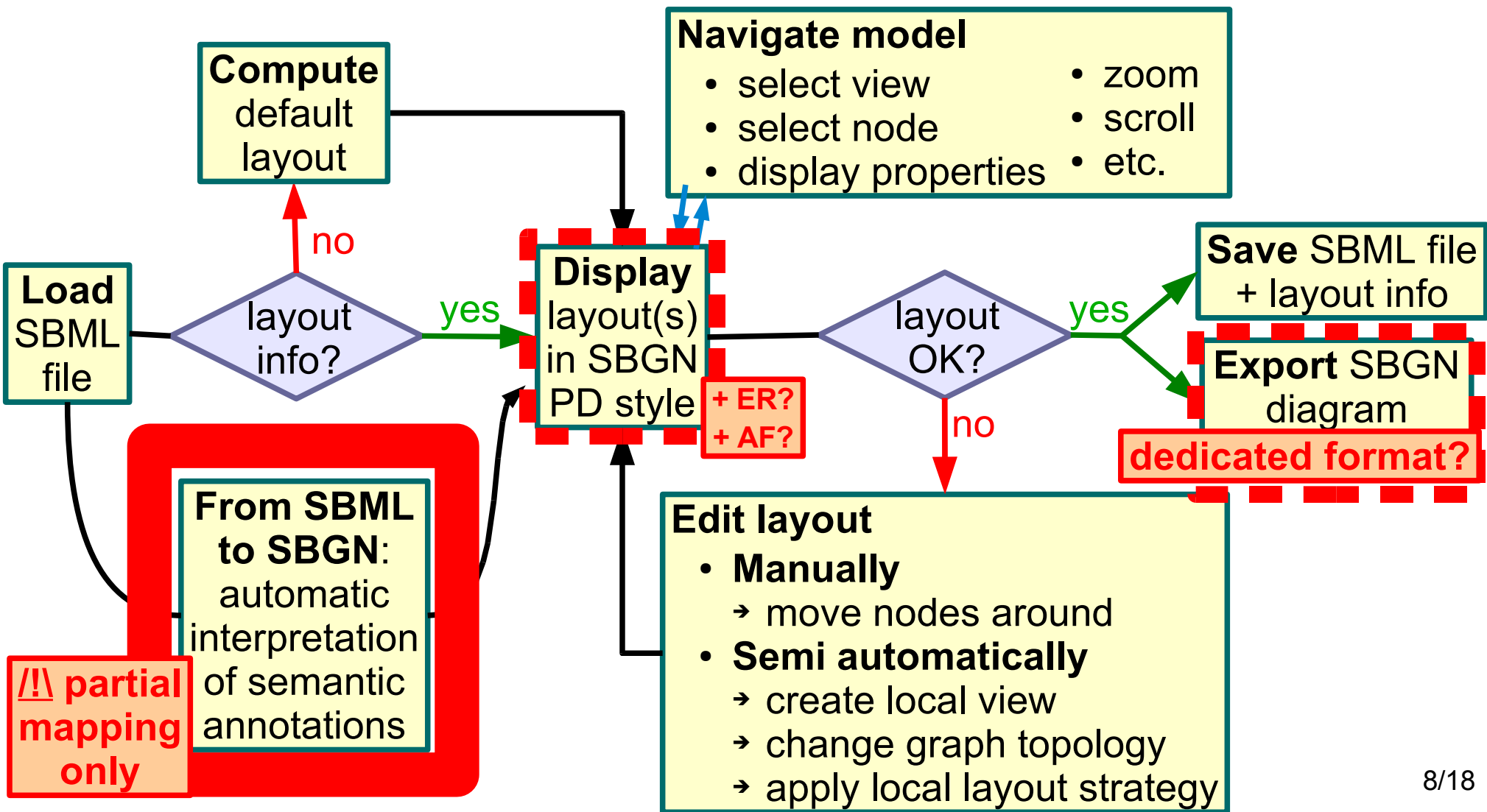
semi automated layout
similar to
hand-drawn diagram



- Tested on a few real use cases
 - illustrate models (communication)
 - interpret experimental results
 - assist curation process
- Mostly SBGN compatible
 - ✓ fixed **transition** nodes problem
 - ✓ all the **glyphs** are supported
 - x but **incomplete mapping** between SBML and SBGN
- Imminent first public release!

arcadiapathways.sourceforge.net

What I will be talking about now



From SBML to SBGN through SBO

- General idea behind the conversion

(annotated, structured content + stylesheet)

x rendering engine = visual representation

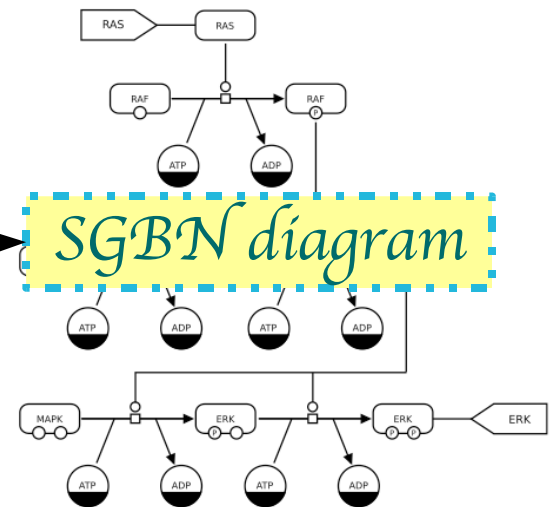
SBML

with SBO and MIRIAM annotations

stylesheet???

model browser?

External
resources?
(via URI)



SBO = missing link?

- SBML x SBO (SBML L2v4r1)

SBML Component	SBO Branch	Branch Identifier
Model	interaction	SBO:0000231
FunctionDefinition	mathematical expression	SBO:0000064
CompartmentType	material entity	SBO:0000240
SpeciesType	material entity	SBO:0000240
Compartment	material entity	SBO:0000240
Species	material entity	SBO:0000240
Reaction	interaction	SBO:0000231
Parameter	quantitative parameter	SBO:0000002
SpeciesReference	participant role	SBO:0000003
ModifierSpeciesReference	participant role	SBO:0000003
KineticLaw	rate law	SBO:0000001
InitialAssignment	mathematical expression	SBO:0000064
AlgebraicRule	mathematical expression	SBO:0000064
AssignmentRule	mathematical expression	SBO:0000064
RateRule	mathematical expression	SBO:0000064
Constraint	mathematical expression	SBO:0000064
Event	interaction	SBO:0000231
Trigger	mathematical expression	SBO:0000064
Delay	mathematical expression	SBO:0000064
EventAssignment	mathematical expression	SBO:0000064

Table 6, page 90, SBML L2v4r1

SBO:0000000 sbo term
SBO:0000002 quantitative parameter [Parameter]
SBO:0000003 participant role [SpeciesReference] [ModifierSpeciesReference]
SBO:0000064 mathematical expression [FunctionDefinition] [InitialAssignment] [AlgebraicRule] [AssignmentRule] [RateRule] [Constraint] [Trigger] [Delay] [EventAssignment]
SBO:0000001 rate law [KineticLaw]
SBO:0000231 interaction [Model] [Reaction] [Event]
SBO:0000236 entity
SBO:0000240 material entity [CompartmentType] [SpeciesType] [Compartment] [Species]

$SBO \cap SBML$

SBGN x SBO (PD L1r1)

SYSTEMS BIOLOGY GRAPHICAL NOTATION REFERENCE CARD

<h3>Entity Pool Nodes</h3> <p>SBO 285 unspecified entity</p> <p>SBO 247 simple chemical</p> <p>SBO 245 macromolecule</p> <p>SBO 354 nucleic acid feature</p> <p>SBO 286 multimer</p> <p>SBO 291 source/sink</p> <p>SBO 358 observable</p> <p>SBO 357 perturbation</p>	<h3>Auxiliary units</h3> <p>SBO / unit of information</p> <p>SBO / state variable</p> <p>SBO / clone marker</p>	<h3>Process Nodes</h3> <p>SBO 167 transition</p> <p>SBO 397 omitted process</p> <p>SBO 396 uncertain process</p> <p>SBO 177 association</p> <p>SBO 180 dissociation</p>	<h3>Connecting Arcs</h3> <p>SBO 394 consumption</p> <p>SBO 393 production</p> <p>SBO 168 modulation</p> <p>SBO 170 stimulation</p> <p>SBO 172 catalysis</p> <p>SBO 169 inhibition</p> <p>SBO 171 trigger</p> <p>SBO 398 logic arc</p> <p>SBO 392 equivalence arc</p>	<p>SBO:000000 sbo term</p> <p>SBO:000003 participant role</p> <p>SBO:000289 functional compartment [Compartment]</p> <p>SBO:000231 interaction</p> <p>SBO:000374 relationship</p> <p>SBO:000168 control [Modulation]</p> <p>SBO:000169 inhibition [Inhibition]</p> <p>SBO:000170 stimulation [Stimulation]</p> <p>SBO:000171 necessary stimulation [Trigger]</p> <p>SBO:000172 catalysis [Catalysis]</p> <p>SBO:000393 production [Production]</p> <p>SBO:000394 consumption [Consumption]</p> <p>SBO:000237 logical combination</p> <p>SBO:000173 and [And]</p> <p>SBO:000174 or [Or]</p> <p>SBO:000238 not [Not]</p> <p>SBO:000392 equivalence [Equivalence arc]</p> <p>SBO:000398 logical relationship [Logic arc]</p> <p>SBO:000375 process</p> <p>SBO:000167 biochemical or transport reaction [Transition]</p> <p>SBO:000176 biochemical reaction</p> <p>SBO:000177 non-covalent binding [Association]</p> <p>SBO:000180 dissociation [Dissociation]</p> <p>SBO:000357 biological effect of a perturbation [Perturbation]</p> <p>SBO:000358 process that affects an observable [Observable]</p> <p>SBO:000395 encapsulating process [Submap]</p> <p>SBO:000396 uncertain process [Uncertain process]</p> <p>SBO:000397 omitted process [Omitted process]</p> <p>SBO:000236 entity</p> <p>SBO:000240 material entity</p> <p>SBO:000245 macromolecule [Macromolecule]</p> <p>SBO:000247 simple chemical [Simple chemical]</p> <p>SBO:000253 non-covalent complex [Complex]</p> <p>SBO:000286 multimer [Multimer]</p> <p>SBO:000285 material entity of unspecified nature [Unspecified entity]</p> <p>SBO:000291 empty set [Source Sink]</p> <p>SBO:000354 informational molecule segment [Nucleic Acid Feature]</p>
	<h3>Container Nodes</h3> <p>SBO 253 complex</p> <p>SBO 289 compartment</p> <p>SBO 395 submap</p>		<h3>Logical Operators</h3> <p>SBO 173 and operator</p> <p>SBO 174 or operator</p> <p>SBO 238 not operator</p>	<h1>SBO ∩ SBGN</h1>

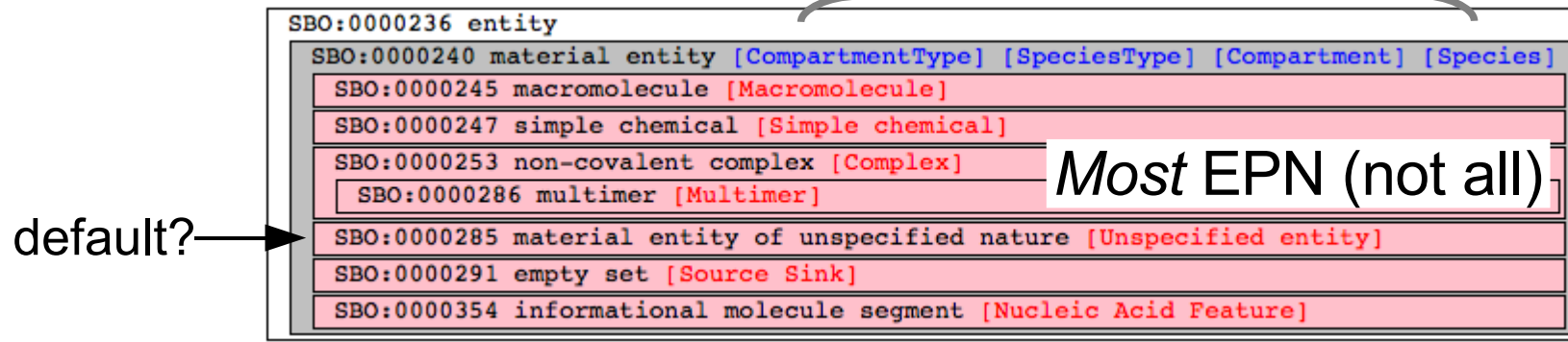
Converting SBML to SBGN

- Relevant SBML elements → Look at SBO term
 - Obvious ones
 - Species[Type]
 - Compartment[Type]
 - Reaction (reactants, products)
 - [Modifier]SpeciesReference
 - Less obvious ones?
 - Parameter
 - Event
 - ...?
- While no equivalence with SBGN, go up in the SBO tree
- If glyph found, display it
- If no match found (or no SBO term), pick up a default SBGN glyph, e.g.:
 - Compartment (SBML) = Compartment (SBGN)
 - Reaction = Transition
 - Modifier = Modulation
 - etc.

SBO \cap (SBML \cup SBGN)

- SBO:0000236 entity = SBML Species and Compartment
 - ✓ Species/SpeciesType maps well to *most* EPN
 - ✗ But SBGN Compartment are not here! \Rightarrow change specs?

Used to span all entities in SBML L2v3



Q: is it ok to assume any unqualified Species is an unspecified EPN by default?
Should the unspecified EPN be associated to SBO:0000240 material entity?

SBO:0000231 interaction

= SBML Model, Reaction, Event

↙ ? ↘
? Maps quite well ?

SBO:0000231 interaction [Model] [Reaction] [Event]	
SBO:0000374 relationship	
SBO:0000168 control [Modulation]	
SBO:0000169 inhibition [Inhibition]	
SBO:0000170 stimulation [Stimulation]	
SBO:0000171 catalysis [Catalysis]	
SBO:0000172 catalysis [Catalysis]	
SBO:0000393 production [Production]	
SBO:0000394 consumption [Consumption]	
SBO:0000237 logical combination	
SBO:0000173 and [And]	
SBO:0000174 or [Or]	
SBO:0000238 not [Not]	
SBO:0000392 equivalence [Equivalence arc]	
SBO:0000398 logical relationship [Logic arc]	
SBO:0000375 process	
SBO:0000167 biological process reaction [Transition]	
SBO:0000176 biological process reaction [Transition]	
SBO:0000177 non-covalent binding [Association]	
SBO:0000180 dissociation [Dissociation]	
SBO:0000357 biological effect of a perturbation on a process that affects an observable [Observable]	
SBO:0000395 encapsulating process [Submap]	
SBO:0000396 process [Uncertain process]	
SBO:0000397 process [Omitted process]	

Modulating arcs

Logical arcs

Transitions

Perturbation/Observable

Process

Problem! In SBML,
Species Reference is SBO:0000003
⇒ can be solved by changing
current SBGN specs (next slide)

Not in SBML?

Equivalence arc and Submap
= Layout information only?

Reaction ⇒ transition

Reaction/Event ⇒ EPN??

Reaction ⇒ [?] [//] process?

(could be layout info?)

Other SBO terms

= Parameter, **SpeciesReference**, Rules, Law, Assignments, Delay, etc.

SBML and SBGN Compartments don't map??

SBO:0000002	quantitative parameter	[Parameter]
SBO:0000003	participant role	[SpeciesReference] [ModifierSpeciesReference]
SBO:0000289	functional compartment	[Compartment]
SBO:0000064	mathematical expression	[FunctionDefinition] [AlgebraicEquation] [Constraint] [EventAssignment] [KineticLaw]

SBGN Unit of information
(also found in entity and process, but not dealt with in that talk)



Constraint, Function, Trigger, etc.

Math expression \Rightarrow not in SBGN (?)

SBO:0000003	participant role	[SpeciesReference] [ModifierSpeciesReference]
SBO:0000010	reactant	
SBO:0000015	substrate	
SBO:0000336	interactor	
SBO:0000000	modulating arc	
SBO:0000000	modulating arc	
SBO:0000000	modulating arc	
SBO:0000000	modulating arc	
SBO:0000206	competitive inhibitor	
SBO:0000207	non-competitive inhibitor	
SBO:0000021	potentiator	
SBO:0000289	functional compartment	[Compartment]
SBO:0000408	biological activity	

Modulating arcs should be here

To be mapped to appropriate connecting arcs in the next SBGN PD release?

Tentative summary

✓ in Arcadia
 ✗ not yet
 / does not apply?

KineticLaw, FunctionDefinition, InitialAssignment,
 AlgebraicRule, AssignmentRule, RateRule,
 Constraint, Trigger, Delay, EventAssignment

no SBGN equivalent? /

Model no SBGN equivalent? / (maybe Submap in SBML L3?) ✗

Compartment **Compartment** (but SBO mismatch) ✓
 CompartmentType could be interpreted as Submap? ✗

Species, SpeciesType **Most EPN:** ✓ **Unspecified entity**, Simple chemical, Macromolecule,
 Nucleic Acid Feature, Multimer, Source/Sink, + Complex (?) ~

Product/Reactant SpeciesReference **Production, Consumption**
 ModifierSpeciesReference **Modulation**, Stimulation, Catalysis, Inhibition, Trigger (after patch) ✓

Reaction **Transition, Association, Dissociation** ✓
 Event? Uncertain process?, Omitted process? (could be layout info?) ✗
 Perturbation, observable? (should be PN, not EPN?) ✗

Layout info? **clone marker**, ✓ equivalence arc, ✗ tag, ✗ submap ✗

Does not apply? and, or, not, logic arc /

Parameter (among others) Unit of information (not fully studied yet) ✗

??? State variable ✗

Problems to solve

- Need machine readable expression of:
 - SBML x SBO
 - SBGN x SBO } XML file, other format, web service?
- Few models have appropriate SBO annotations
 - Generate a test case file covering all SBGN glyphs?
 - MIRIAM to SBO conversion? (e.g. on biomodels.net DB)
- Ad-hoc conversion in Arcadia ⇒ not reusable
 - Define a SBML to SBGN “stylesheet”? (PD, + ER, AF?)
 - Open question: SBGN file format...?

Thank you!

More information:

<http://arcadiapathways.sourceforge.net>

<http://personalpages.manchester.ac.uk/staff/Alice.Villeger>

Questions, comments, suggestions, etc.:

alice.villeger@manchester.ac.uk

Also feel free to ask me for a demo during the meeting!