

Models For All

*Standards for describing the whole life-cycle of modeling
in the life sciences*

Nicolas Le Novère, EMBL-EBI

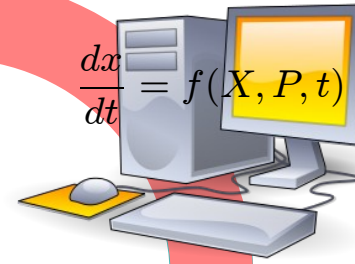
New way of doing biomedical research

Needs for interplay between models and reality tests

Experiment

model

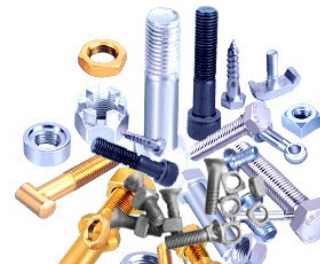
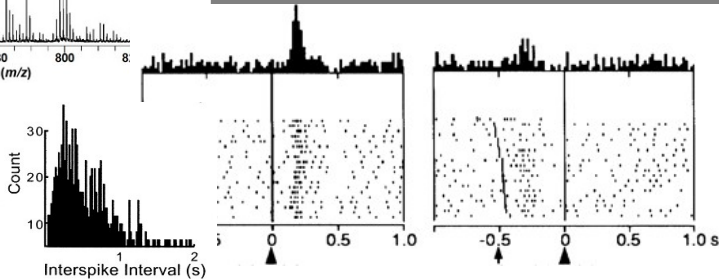
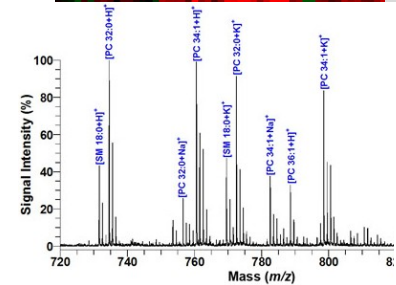
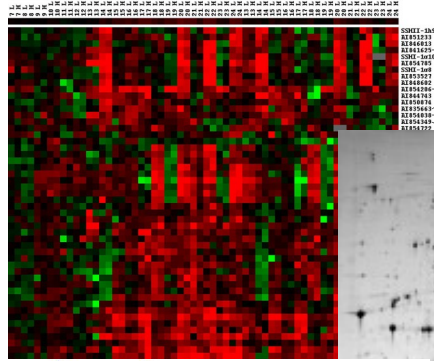
$$\frac{dx}{dt} = f(X, P, t)$$



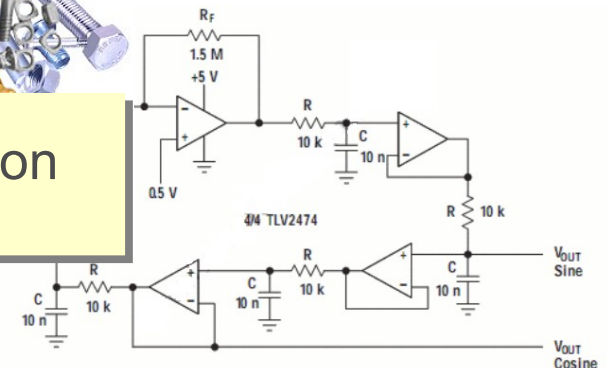
hypothesis



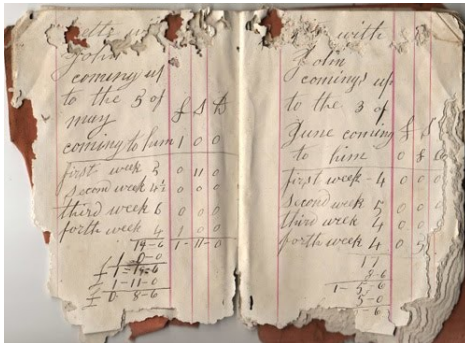
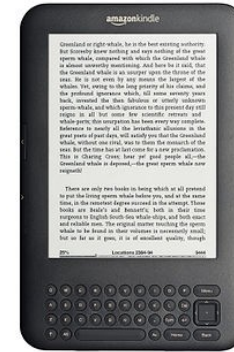
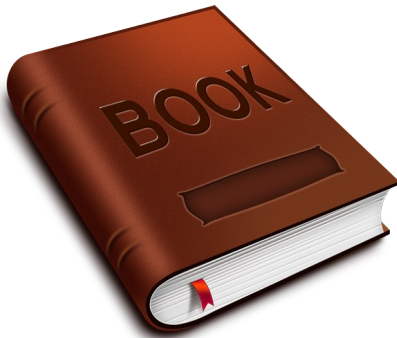
Needs for systems thinking and integration of heterogeneous knowledge



Needs for cooperation and standardisation



What are standards good for?



Accounts.gruntic : Gruntic

File Edit View Insert Format Tools Data Help

Helvetica

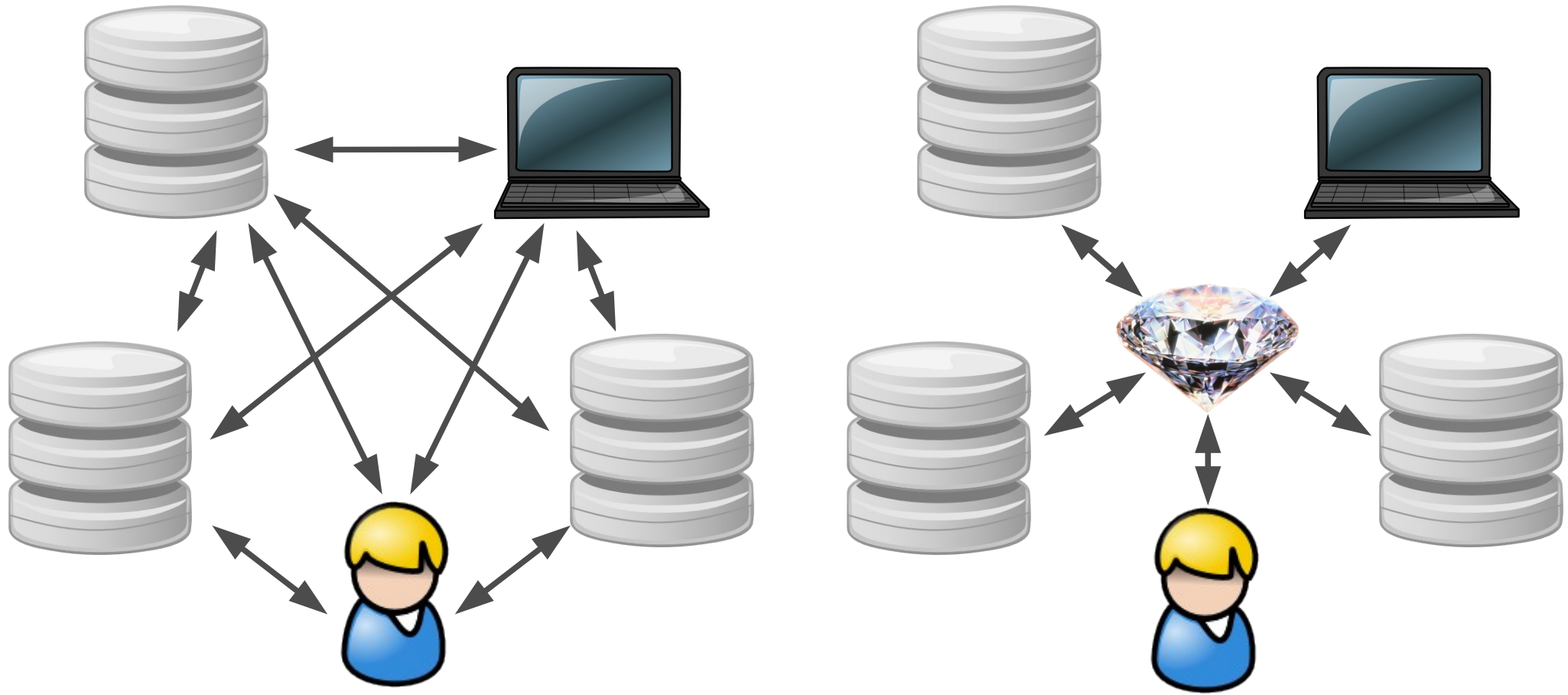
	A	B	C	D	E	F	G	H
1	Anita's Account							
2								
3	Task	Date	Start	End	Income	Total	Adjustments	
4	Carry Forward					\$ 40.00		
5	Clothes + Wash Floor	21-Sep-2000	4:30 PM	5:00 PM	\$ 2.00	\$ 42.00		
6	Cooking		5:35 PM	6:21 PM	\$ 3.07	\$ 45.07		
7	Fold Clothes		6:25 PM	6:35 PM	\$ 0.67	\$ 45.75		
8	Fold Clothes		6:45 PM	7:10 PM	\$ 1.67	\$ 47.40		
9	Clothes Away	22-Sep-2000	6:30 AM	6:38 AM	\$ 0.40	\$ 47.80		
10	Wash Car		10:03 AM	11:23 AM	\$ 5.33	\$ 53.13		
11	Lolli Pop				\$ 0.00	\$ 53.83	\$ -0.10	
12	Cooking		0:55 PM	1:04 PM	\$ 0.60	\$ 53.83		
13								
14								
15								
16								
17								
18								
19								
20								
	Anita [Scan]					Sum=0		

$$\int_1^t \frac{dx}{x}$$

```
<int/>
  <bvar><ci>x</ci></bvar>
  <lowlimit><cn>1</cn></lowlimit>
  <uplimit><ci>t</ci></uplimit>
  <apply>
    <divide/>
    <cn>1</cn>
    <ci>x</ci>
  </apply>
</int>
```

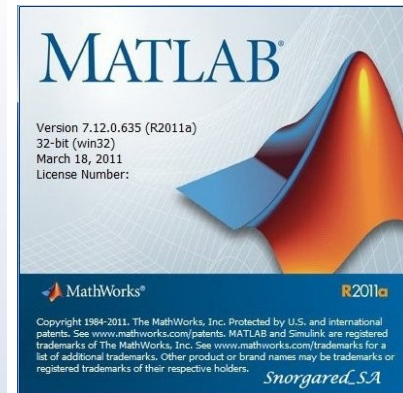
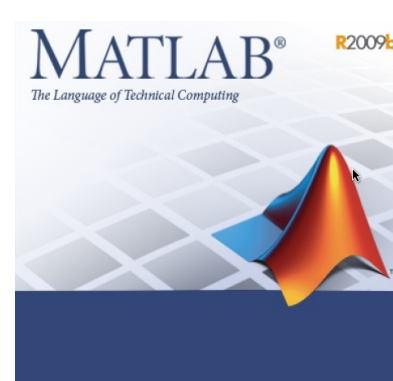
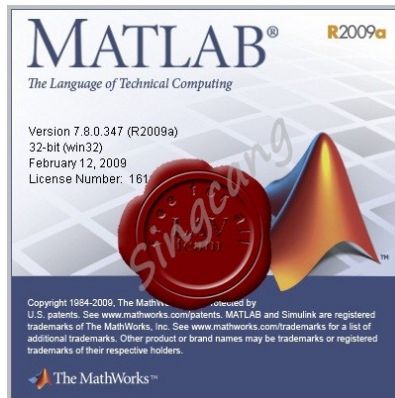
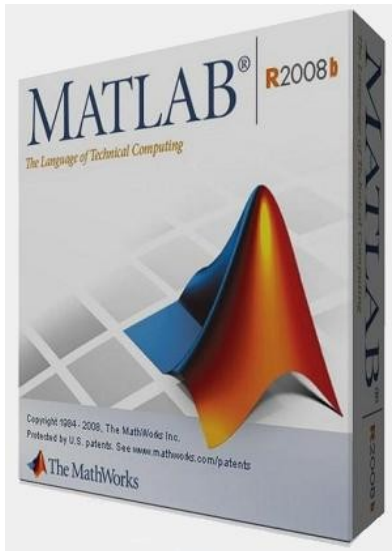
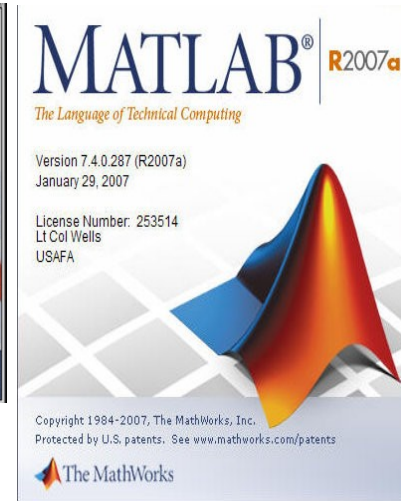
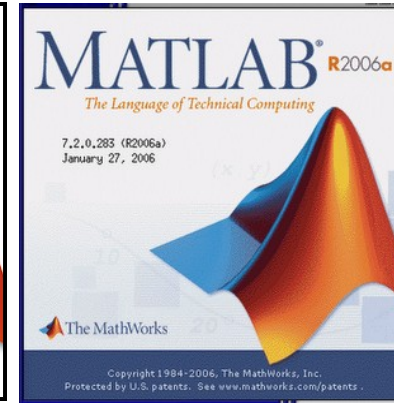
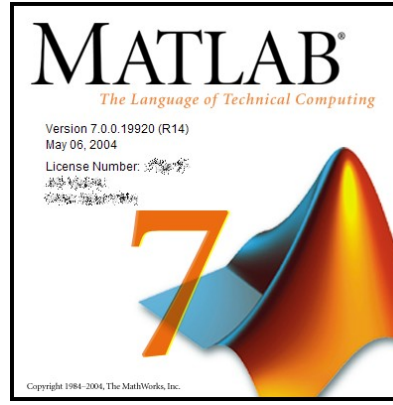
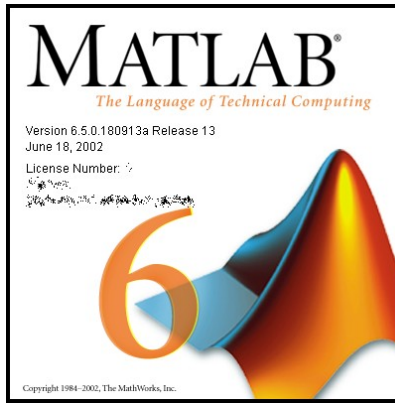
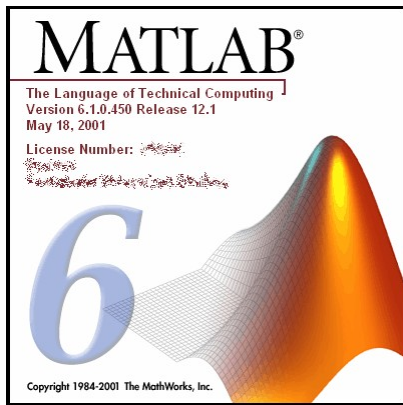
Formal languages can be read and written by computers

What are standards good for?



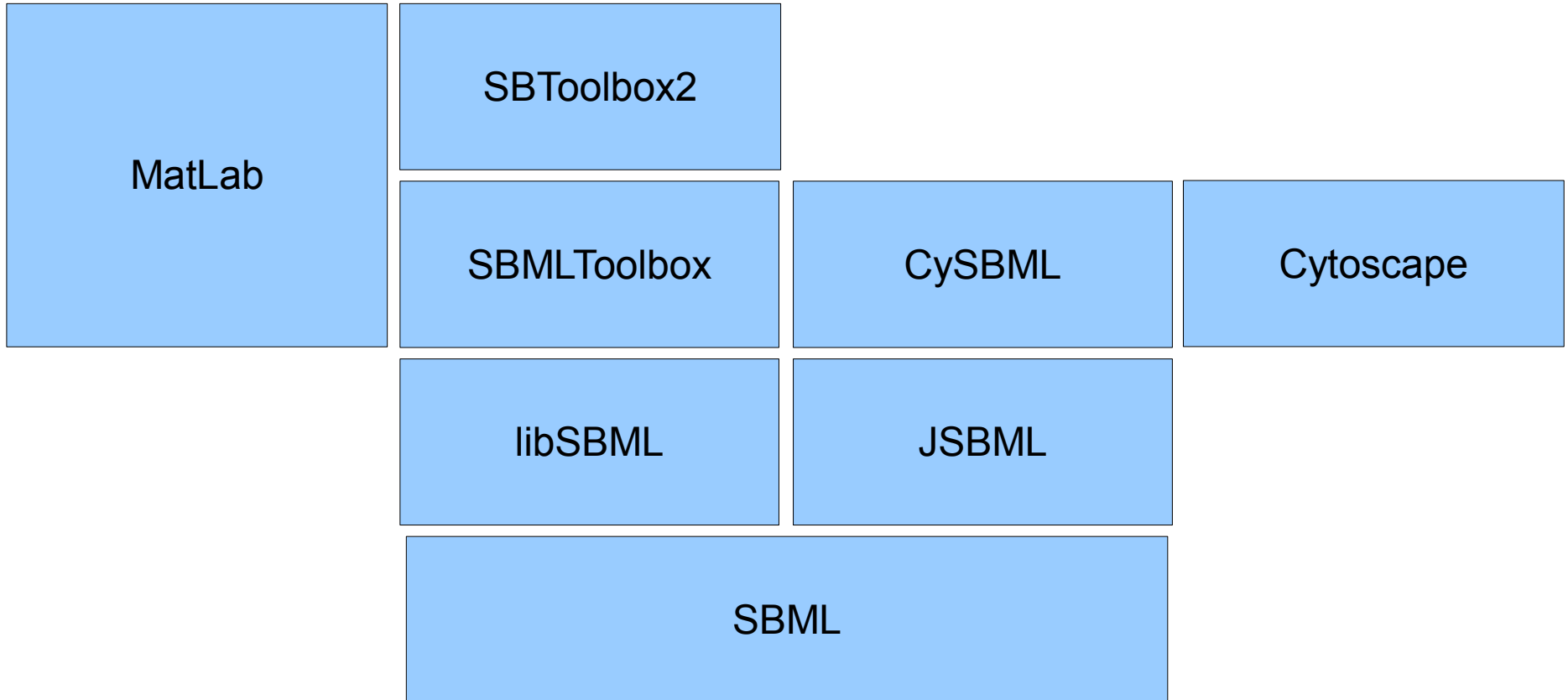
N tools require N conversions for exchange and not $N(N-1)/2$

What are standards good for?



Open standards are more stable than proprietary formats

What are standards good for?



Open standards can be built on. They generate new science

Three types of standards

Minimal requirements	What to encode in order to share experiments and understand results 
Data-models	How to encode the information defined above in a computer-readable manner 
Terminologies	Structured representation of knowledge, with concept definitions and their relationships 

Three types of standards

Minimal requirements	What to encode in order to share experiments and understand results  MIBBI
Data-models	How to encode the information defined above in a computer-readable manner  combine
Terminologies	Structured representation of knowledge, with concept definitions and their relationships  OBO foundry

A language to describe computational models in biology

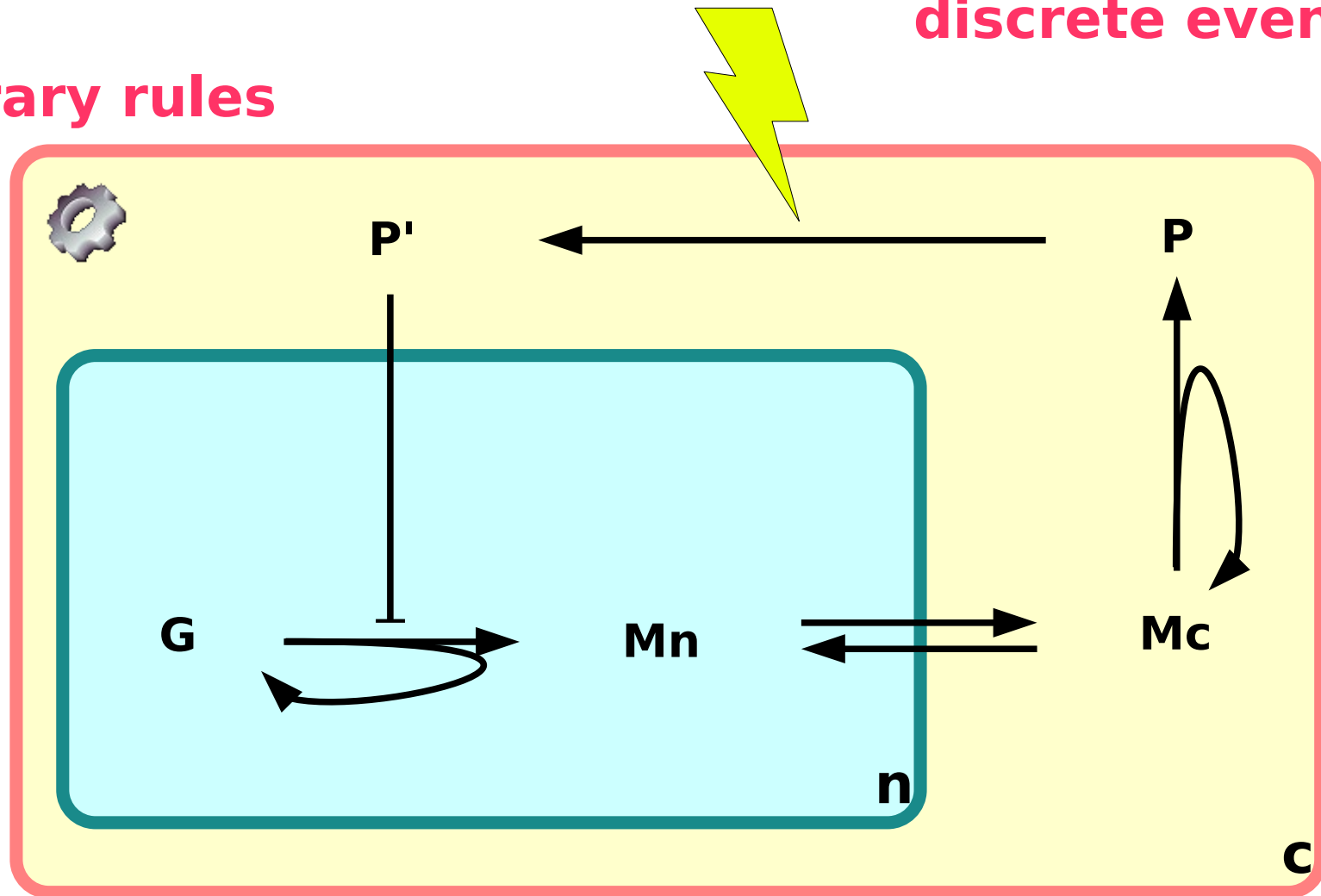
	Model descriptions
Data-models	

Born in Caltech 2000

What can we encode in SBML (core)?

arbitrary rules

discrete events



Why the Extensible Markup Language (XML)?

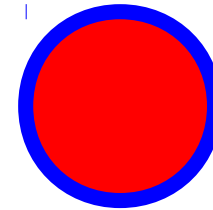
- **HTML**

A **strong word** and an [hyperlink](http://www.w3.org/)

A strong word and an [hyperlink](#)

- **SVG**

```
<circle r="100" fill="red"
stroke="blue" stroke-width="10" />
```



- **MathML**

```
<apply>
  <int/>
  <bvar><ci> x </ci></bvar>
  <lowlimit><cn> 0 </cn></lowlimit>
  <uplimit><ci> a </ci></uplimit>
  <apply><ci> f </ci><ci> x </ci></apply>
</apply>
```

$$\int_0^a f(x) \, dx$$

- **Excel**

```
<row r="1">
  <c r="A1"><v>1</v></c>
  <c r="B1">
    <f>C11*E11</f>
    <v>102</v>
  </c>
</row>
```

B1		▼		Σ	=	=C11*E11
	A	B	C			
1		102				
2						

Why the Extensible Markup Language (XML)?

- Easy to define and validate
 - Rapid prototyping, processing tools can be generated and thrown away
- Existence of a very large toolkit
 - Libraries in every programming languages
 - A very large number of description formats in life sciences are in XML
- Associated technologies
 - Definition: XML Schema, Schematron (themselves XML)
 - Conversion: XSLT (using XSL in XML)
 - Linking: XPath and XQuery

Global structure of a SBML file

```
<?xml version="1.0" encoding="UTF-8"?>
<sbml level="3" version="1".
  xmlns="http://www.sbml.org/sbml/level3/version1/core">
  <model>
    <listOfFunctionDefinitions> <!-- --> </listOfFunctionDefinitions>
    <listOfUnitDefinitions> <!-- --> </listOfUnitDefinitions>
    <listOfCompartments> <!-- --> </listOfCompartments>
    <listOfSpecies> <!-- --> </listOfSpecies>
    <listOfParameters> <!-- --> </listOfParameters>
    <listOfInitialAssignments> <!-- --> </listOfInitialAssignments>
    <listOfRules> <!-- --> </listOfRules>
    <listOfConstraints> <!-- --> </listOfConstraints>
    <listOfReactions> <!-- --> </listOfReactions>
    <listOfEvents> <!-- --> </listOfEvents>
  </model>
</sbml>
```

variables

relationships

A very simple SBML file (A $\rightarrow$ B)

```
<?xml version="1.0" encoding="UTF-8"?>
<sbml xmlns="http://www.sbml.org/sbml/level2/version4" level="2" version="4">
  <model name="Simple Model">
    <listOfCompartments>
      <compartment id="cell" size="1" />
    </listOfCompartments>
    <listOfSpecies>
      <species id="A" compartment="cell" initialConcentration="1"/>
      <species id="B" compartment="cell" initialConcentration="1"/>
    </listOfSpecies>
    <listOfParameters>
      <parameter id="k1" value="0.1"/>
    </listOfParameters>
    <listOfReactions>
      <reaction id="r1" reversible="false">
        <listOfReactants>
          <speciesReference species="A"/>
        </listOfReactants>
        <listOfProducts>
          <speciesReference species="B"/>
        </listOfProducts>
        <kineticLaw>
          <math xmlns="http://www.w3.org/1998/Math/MathML">
            <apply>
              <times/>
              <ci> cell </ci>
              <ci> k1 </ci>
              <ci> A </ci>
            </apply>
          </math>
        </kineticLaw>
      </reaction>
    </listOfReactions>
  </model>
</sbml>
```

A very simple SBML file (A $\rightarrow$ B)

```
<?xml version="1.0" encoding="UTF-8"?>
<sbml xmlns="http://www.sbml.org/sbml/level2/version4" level="2" version="4">
  <model name="Simple Model">
    <listOfCompartments>
      <compartment id="cell" size="1" />
    </listOfCompartments>
    <listOfSpecies>
      <species id="A" compartment="cell" initialConcentration="1"/>
      <species id="B" compartment="cell" initialConcentration="1"/>
    </listOfSpecies>
    <listOfParameters>
      <parameter id="k1" value="0.1"/>
    </listOfParameters>
    <listOfReactions>
      <reaction id="r1" reversible="false">
        <listOfReactants>
          <speciesReference species="A"/>
        </listOfReactants>
        <listOfProducts>
          <speciesReference species="B"/>
        </listOfProducts>
        <kineticLaw>
          <math xmlns="http://www.w3.org/1998/Math/MathML">
            <apply>
              <times/>
              <ci> cell </ci>
              <ci> k1 </ci>
              <ci> A </ci>
            </apply>
          </math>
        </kineticLaw>
      </reaction>
    </listOfReactions>
  </model>
</sbml>
```

MathML



SimpleSBML - COPASI 4.6 (Build 32) /home/.../2011-03-ModelsIndustry/SimpleSBML.cps

File Tools Help

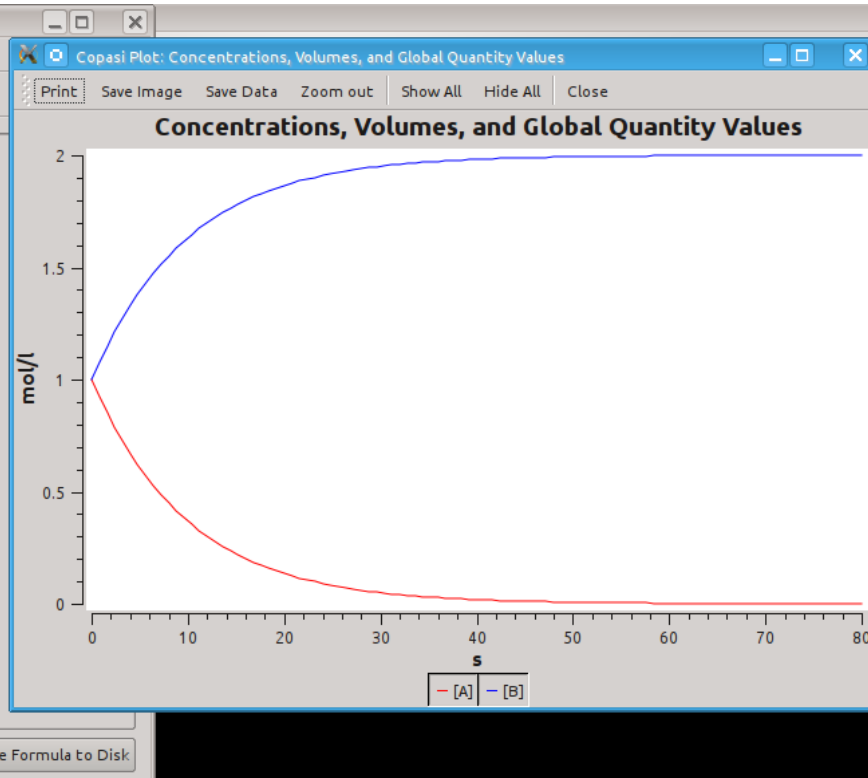
Concentrations

Copasi

- Model
 - Biochemical
 - Compartments
 - cell
 - Species
 - A
 - B
 - Reactions
 - r1
 - Global Quantities
 - Events
 - Parameter Overview
 - Mathematical
 - Differential Equations
 - Matrices
 - Diagrams
 - Tasks
 - Steady-State
 - Stoichiometric Analysis
 - Time Course
 - Metabolic Control Analysis
 - Lyapunov Exponents
 - Time Scale Separation Analysis
 - Parameter Scan
 - Optimization
 - Parameter Estimation
 - Sensitivities

$$\frac{d([A] \cdot V_{cell})}{dt} = -V_{cell} \cdot (k1 \cdot [A])$$
$$\frac{d([B] \cdot V_{cell})}{dt} = +V_{cell} \cdot (k1 \cdot [A])$$

local parameters display numerical value Save Formula to Disk



CellDesigner

File Edit Component View Database Layout Simulation Plugin Window SBW Preference Help

Model

- Compartments
- Species
- Reactions

Layer base

SimpleSBML.xml *

Species Proteins Genes RNAs asRNAs Reactions Compartments Par

Edit Export

class	id	name	speciesType	compar...	positi...	included
PROTEIN	A	A		cell	inside	
PROTEIN	B	B		cell	inside	

ControlPanel SimpleSBML.xml

File Edit Data Simulation

Time span: End Time 80, Num. o... 100

Error tolerance: Exp. -6

Solver: ☒ SLSlib, ☐ COPASI

Species Parameters Change amount Parameter Scan

Id	Name	Compart...	Quantity ...	Initial
A	A	cell	Concentra...	
B	B	cell	Concentra...	

Graph Table

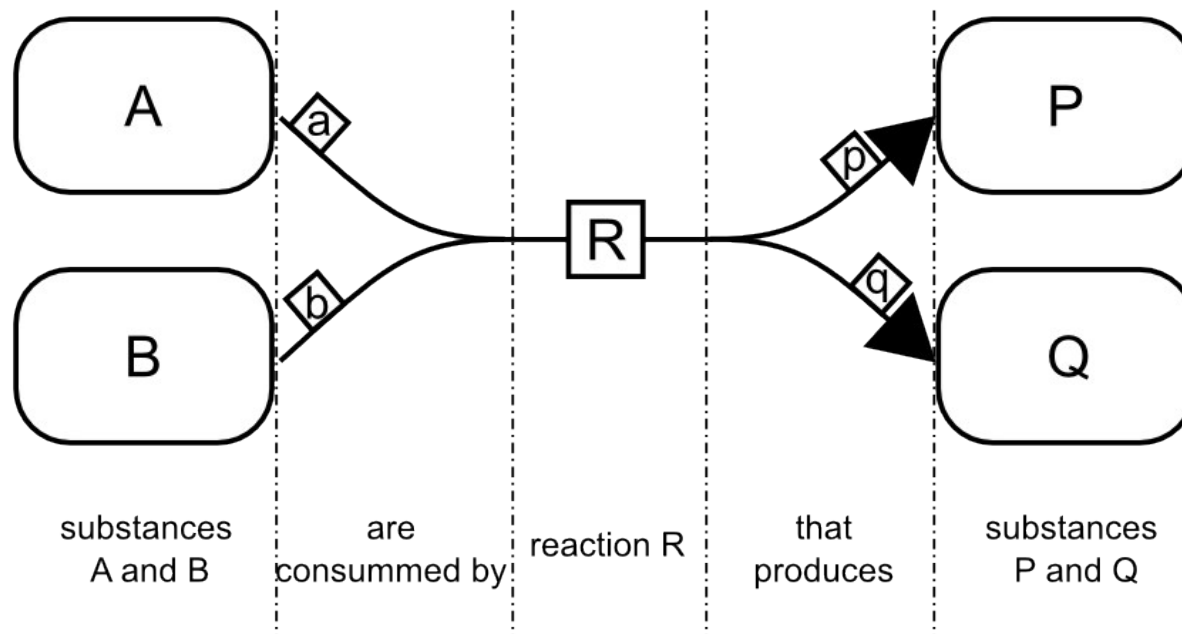
Concentration raw ☒ Species ☐ Fluxes ☐ Parameters ☐ Comp...

Time

80.00

Initialize Save As Execute Close show scatter plot

Unselect all species search



Species

Species
Reference

Kinetic
Law

Species
Reference

Species

A more realistic example ...

```

<species.
  id="A".
  name="a-tubulin"
  compartment="cell"
  initialAmount="1000"
  substanceUnits="item"
  hasOnlySubstanceUnits="true"
  boundaryCondition="true"
  constant="false"
  charge="0"
  metaid="PX"
  sboTerm="SBO:0000245" >
<notes>
  <body xmlns="http://www.w3.org/1999/xhtml">
    <p>One of the components of a microtubule</p>
  </body>
</notes>
<annotation>
  <rdf:RDF.
    xmlns:bqbiol="http://biomodels.net/biology-qualifiers/"
    xmlns:bqmodel="http://biomodels.net/model-qualifiers/"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
    <rdf:Description rdf:about="#PX">
      <bqbiol:is>
        <rdf:Bag>
          <rdf:li rdf:resource="urn:miriam:uniprot:P68370"/>
          <rdf:li rdf:resource="urn:miriam:obo.go:GO%3A0045298"/>
        </rdf:Bag>
      </bqbiol:is>
    </rdf:Description>
  </rdf:RDF>
</annotation>
</species>

```

A more realistic example ...

```
<species.
  id="A".
  name="a-tubulin"
  compartment="cell"
  initialAmount="1000"
  substanceUnits="item"
  hasOnlySubstanceUnits="true"
  boundaryCondition="true"
  constant="false"
  charge="0"
  metaid="PX"
```

```
sboTerm="SBO:0000245" >
```

biological semantics: macromolecule

```
<notes>
```

XHTML

```
<body xmlns="http://www.w3.org/1999/xhtml">
  <p>One of the components of a microtubule</p>
</body>
```

```
</notes>
```

```
<annotation>
```

RDF

```
<rdf:RDF.
  xmlns:bqbiol="http://biomodels.net/biology-qualifiers/"
  xmlns:bqmodel="http://biomodels.net/model-qualifiers/"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:Description rdf:about="#PX">
    <bqbiol:is>
      <rdf:Bag>
        <rdf:li rdf:resource="urn:miriam:uniprot:P68370"/>
        <rdf:li rdf:resource="urn:miriam:obo.go:GO%3A0045298"/>
      </rdf:Bag>
    </bqbiol:is>
  </rdf:Description>
</rdf:RDF>
```

```
</annotation>
```

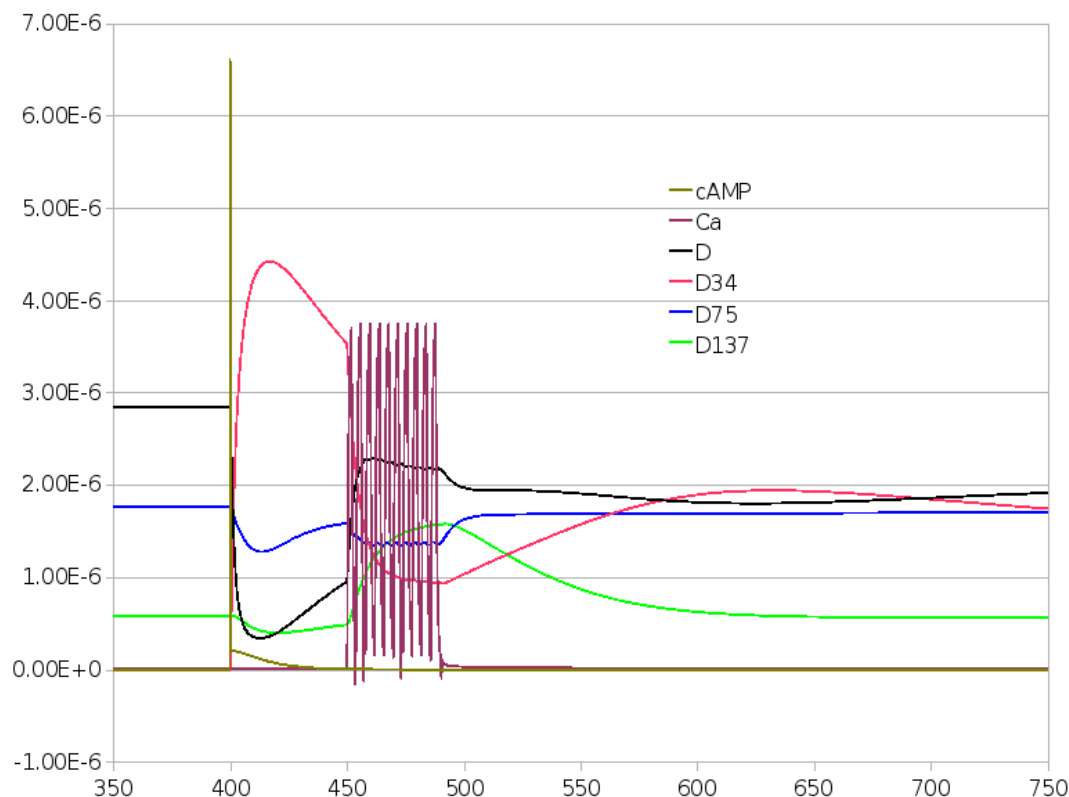
```
</species>
```

SBML is not limited to biochemistry!

- A **species** is a pool of entities participating to a reaction, **not always** a **chemical** entity
 - It can be a pool of molecules
 - It can be a pool of cells
 - It can be a pool of organs
 - It can be a population of organism
- **Rate Rules** can describe the temporal evolution of any quantitative parameter, e.g. transmembrane voltage, tumour size etc.
- **Events** can describe any discontinuous change, e.g. neurotransmitter release, repolarisation, cell division etc.

→ SBML is about process descriptions

Biochemical models



reaction:

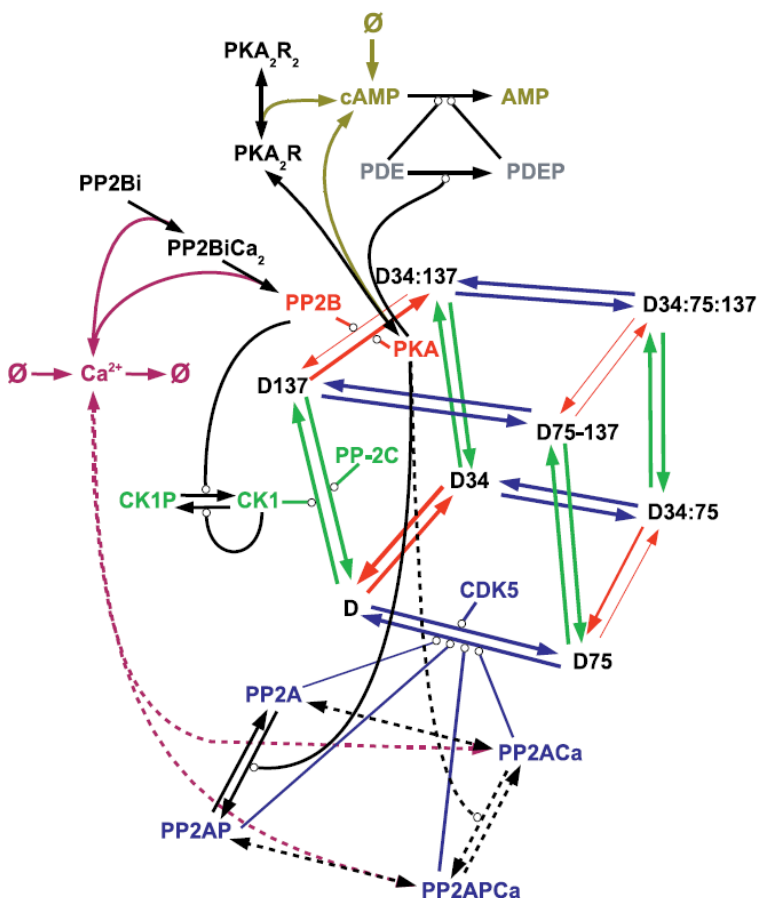
$$v_{on1} = k_{on1} \times [D] \times [CDK5] \times Vol$$

→ Vijayalakshmi Chelliah, Saturday

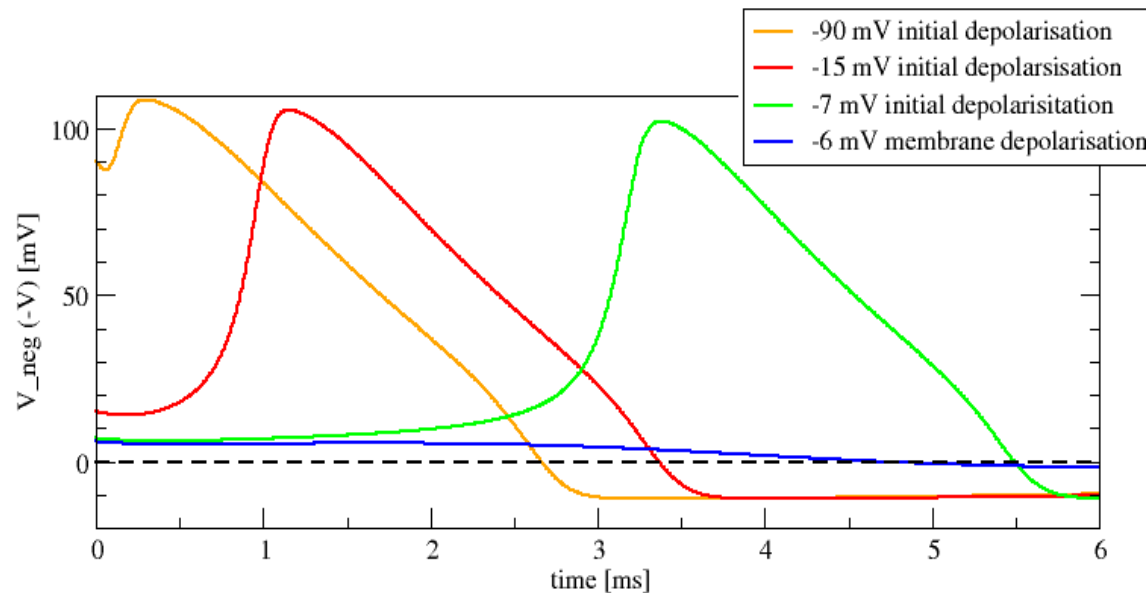
Fernandez et al. DARPP-32 is a robust integrator of dopamine and glutamate signals
PLoS Comput Biol (2006) 2: e176.



BIOMD0000000153



Conductance-based model



Hodgkin AL, Huxley AF.
A quantitative description of
membrane current and its
application to conduction
and excitation in nerve.
J Physiol (1952) 117:500-544.



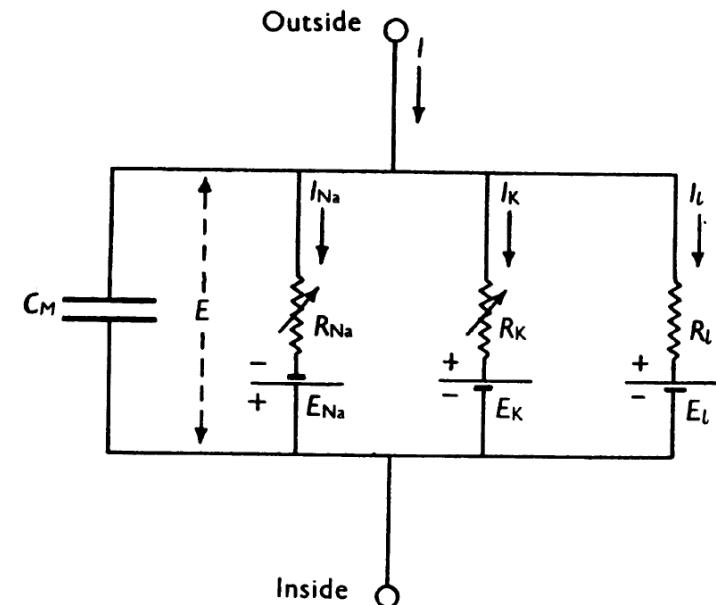
BIOMD0000000020

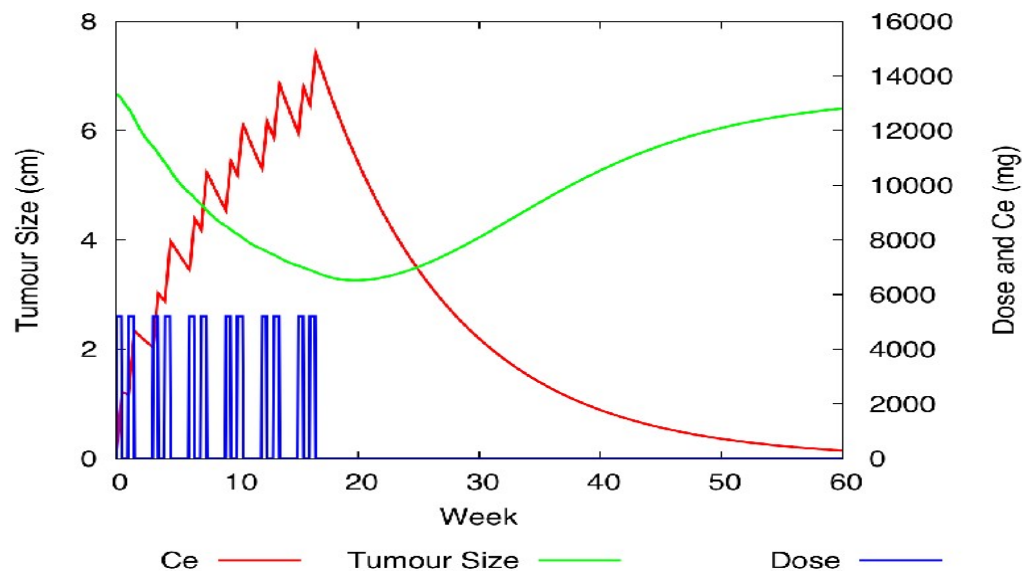
rate rule:

$$\frac{dv}{dt} = \frac{I - (i_{Na} + i_K + i_L)}{C_m}$$

assignment rule:

$$i_{Na} = g_{Na} \times m^3 \times h \times (V - E_{Na})$$





Tham et al (2008) A pharmacodynamic model for the time course of tumor shrinkage by gemcitabine + carboplatin in non-small cell lung cancer patients. *Clin Cancer Res.* 2008 14(13): 4213-8.

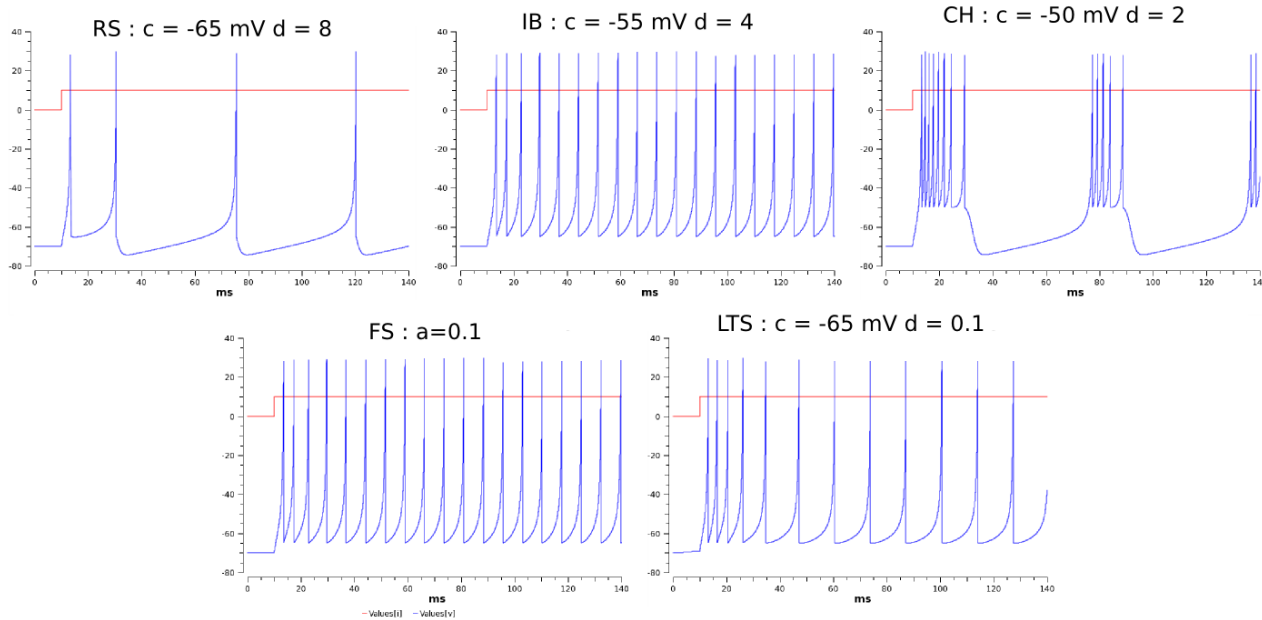


rate rule:

$$\frac{dSize}{dt} = (Rate_{in} \times Effect - K_{over} \times Size) \times Size$$

assignment rule:

$$Effect = 1 - \frac{E_{max} \times Ce}{Amt_{50} + Ce}$$



Izhikevich EM. Simple model of spiking neurons. *IEEE Trans Neural Netw* (2003) 14(6):1569-1572.



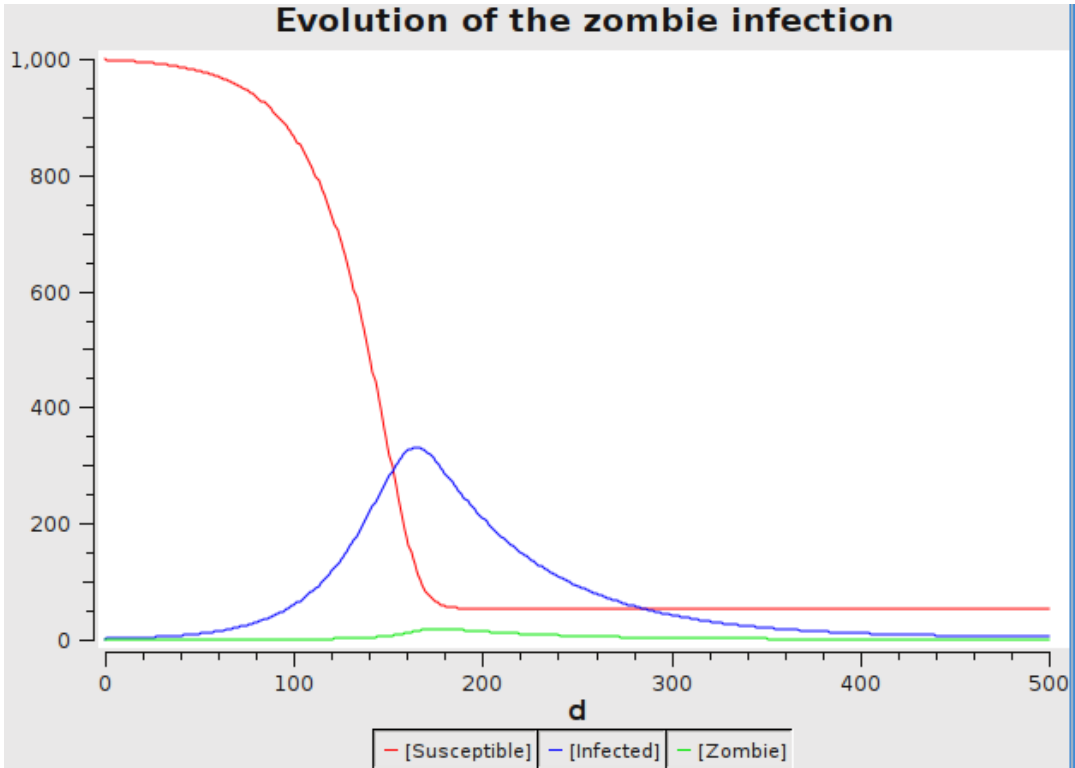
rate rule:

$$\frac{dv}{dt} = 0.04v^2 + 5 \times V + 140 - U + i$$

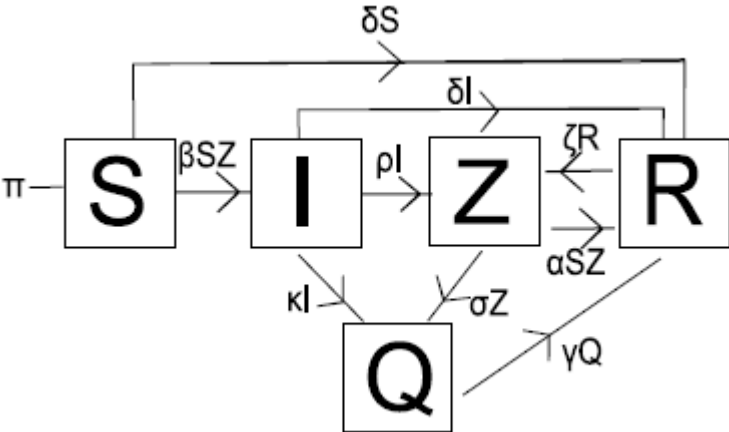
event:

$$\text{when } v > V_{thresh} \begin{cases} v = c \\ U = U + d \end{cases}$$

Spread of infection diseases ...



Munz P et al. When zombies attack!: Mathematical modelling of an outbreak of zombie infection. in "Infectious Disease Modelling Research Progress", (2009)133-150



SBML Level 3 packages

- Core package – public specification
- Model composition – public specification
- Qualitative models – public specification
- Flux balance constraint – public specification
- Complex species – specification finalised
- Graph Layout – specification finalised
- Groups - specification finalised
- Distributions and ranges - specification under discussion
- Spatial diffusion – specification under discussion
- Enhanced metadata – specification under discussion
- Graph rendering – specification proposed
- Arrays and sets – specification proposed
- Dynamic structures - needed

???

SBML definition and API

- SBML syntax and semantics are very precisely defined
 - SBML specification document: Level 3 Version 1 = 167 pages, small margins
 - XML schema (L1 and L2) and Schematron (forthcoming for L3)
 - Hundreds of validation rules to check compliance
- A standard Application Programming Interface with two implementations
 - LibSBML in C and C++, with binding to C#, Java, Python, Perl, MatLab, Octave, Ruby
 - JSBML, native Java version
- Test suite of 5514 models either testing a feature or a documented error

→ Sarah Keating, Nicolas Rodriguez, Saturday

SBML Software Guide

The following pages describe SBML-compatible software packages known to us. We offer different ways of viewing the information, all drawn from the same underlying data collected from the systems' developers via our [software survey](#). The *Matrix* provides a table listing all known software and a variety of their features; the *Summary* provides general descriptions of most of the software; and the *Showcase* provides a sequential slideshow of a subset of the software.

Number of software packages listed in the matrix today: **232**.

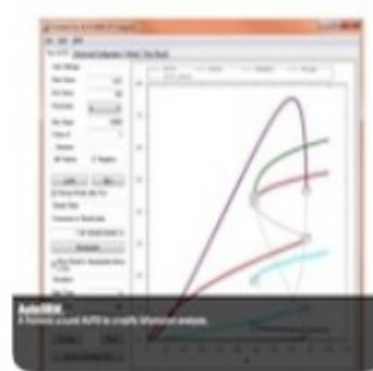
Go to the SBML
Software Matrix

The screenshot shows the SBML Software Matrix, a table with columns for software name, version, language, platform, and various features. The table lists numerous software packages and their compatibility with different SBML standards and features.

Go to the SBML
Software Summary

The screenshot shows the SBML Software Summary page, which provides a detailed description of the software packages listed in the matrix. It includes sections for software categories, features, and a list of software packages with their descriptions.

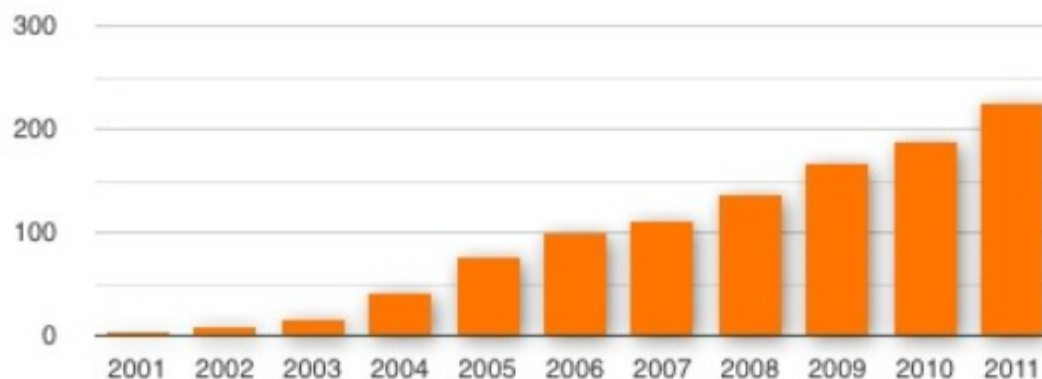
Go to the SBML
Software Showcase



Please [use the survey form](#) to notify us about additions and suggestions.

Historical trend

The following graph shows the total number of known SBML-compatible software packages each year, as counted by the SBML Team. The counts shown are for approximately the middle of each year.



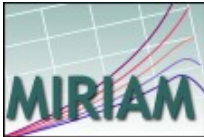
SBML supporting tools

- Simulators
 - Discrete stochastic (25)
 - Continuous deterministic (42)
 - Spatial (4)
- Modelling and simulation environments (29)
 - Based on Mathematica (3)
 - Based on Matlab (12)
 - Based on Python/SciPy (9)
 - Based on R (3)
- Flux/metabolic analysis (16)
- Integrated framework (3)
- Libraries (3)
- Model Management, Data Integration, and Analysis (12)
- Model development tools (18)
- Model visualisation (7)
- Model Repositories, Test Suites, and Databases (16)
- Converters (7)
- Analysis and utility (12)

Adding the semantics to the syntax

	Model descriptions
Minimal requirements	 The MIRIAM logo features the word "MIRIAM" in a bold, green, sans-serif font. It is set against a light blue background with a grid pattern and several colorful, curved lines in shades of red, orange, and purple.
Data-models	 The S8ML logo consists of the letters "S8ML" in a blue, sans-serif font. The "8" is stylized with two circular shapes. A small "TM" trademark symbol is located to the upper right of the "L". The logo is on a light blue background with a grid and curved lines.
Terminologies	 The S8O logo features the letters "S8O" in a green, sans-serif font. The "8" is stylized with two circular shapes. It is set against a light blue background with a grid pattern and several colorful, curved lines in shades of red, orange, and purple.

Born in Heidelberg 2004



SBML and MIRIAM cross-references

```
<species id="ca_calmodulin" metaid="cacam">
  <annotation>
    <rdf:RDF
      »   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
          xmlns:bqbiol="http://biomodels.net/biology-qualifiers/">
        <rdf:Description rdf:about="#cacam">
          »   <bqbiol:hasPart>
              <rdf:Bag>
                »   <rdf:li rdf:resource="http://identifiers.org/uniprot/62158"/>
                »   <rdf:li rdf:resource="http://identifiers.org/chebi/CHEBI:29108"/>
              </rdf:Bag>
            </bqbiol:hasPart>
          </rdf:Description>
        </rdf:RDF>
      </annotation>
    </species>
```

<http://co.mbine.org/standards/miriam>
<http://identifiers.org>
<http://www.ebi.ac.uk/miriam>

→ Camille Laibe



Systems Biology Ontology

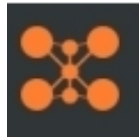
- Download
- Recent changes

- Term request
- Edit tree

Web Services

- FAQ
- News
- Project on SourceForge
- Contact

BIOMODELS.NET



SBO:0000000 - systems biology representation

SBO:0000064 - mathematical expression

- SBO:0000355 - conservation law
- SBO:0000474 - convenience function
- SBO:0000001 - rate law
- SBO:0000391 - steady state expression

SBO:0000544 - metadata representation

SBO:0000004 - modelling framework

SBO:0000231 - occurring entity representation

SBO:0000003 - participant role

SBO:0000236 - physical entity representation

SBO:0000545 - systems description parameter

SBO:0000546 - qualitative systems description parameter

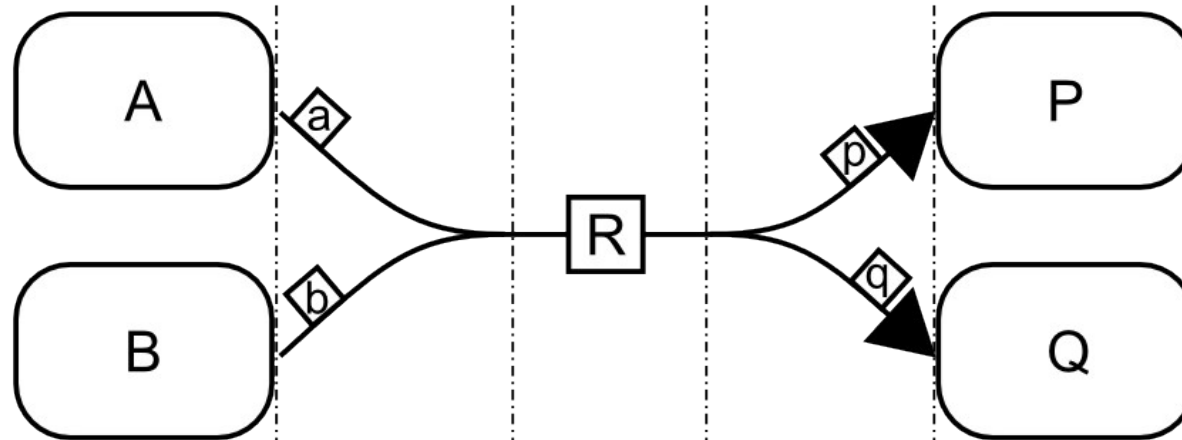
SBO:0000002 - quantitative systems description parameter

- SBO:0000492 - amplitude
- SBO:0000542 - basic reproductive ratio
- SBO:0000380 - biochemical coefficient
- SBO:0000258 - capacitance
- SBO:0000257 - conductance
- SBO:0000254 - electrical resistance
- SBO:0000308 - equilibrium or steady-state characteristic

Essential activator

```
<listOfModifiers>  
  <modifierSpeciesReference  
    sboTerm="SBO:0000461"  
    species="Y"/>  
</listOfModifiers>
```

<http://www.ebi.ac.uk/sbo/>



substances
A and B

are
consumed by

reaction R

that
produces

substances
P and Q



SBO:0000240
material entity

SBO:0000394
consumption

SBO:0000375
process

SBO:0000393
production

SBO:0000240
material entity



Species

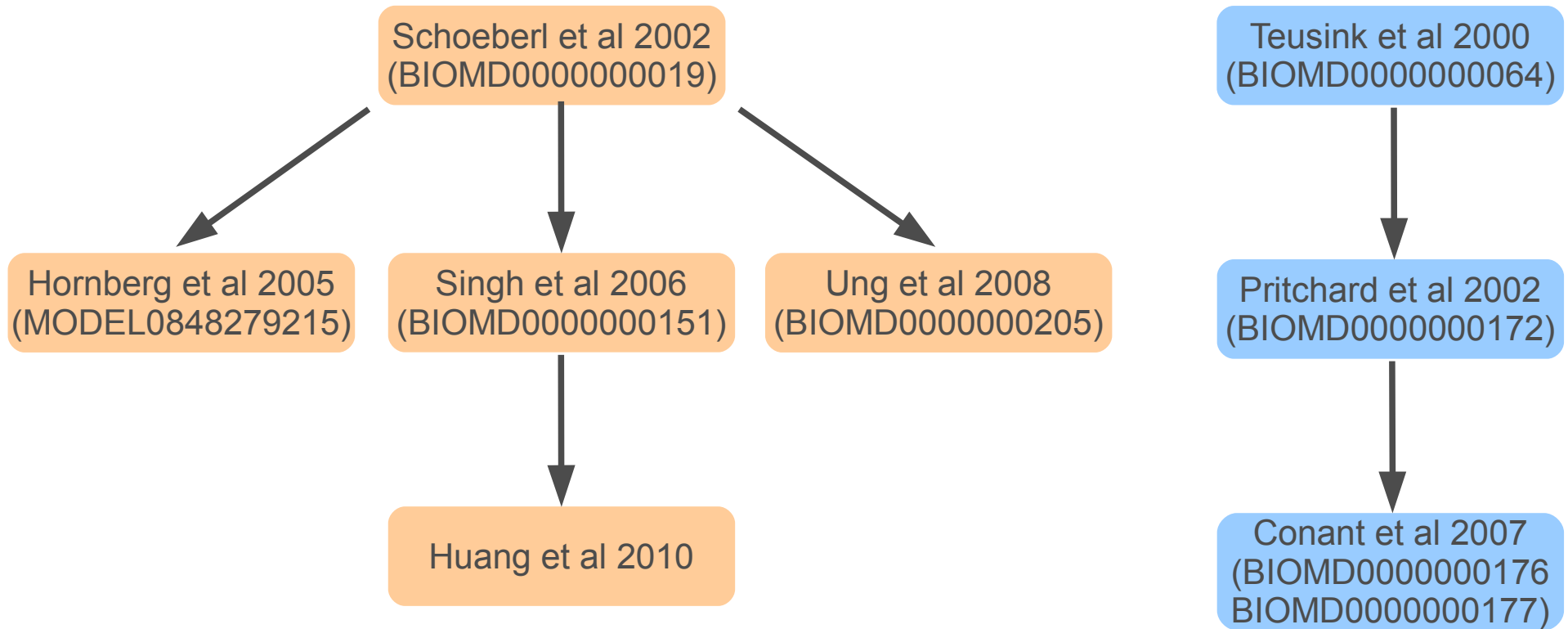
Species
Reference

Kinetic
Law

Species
Reference

Species

Direct model re-use: e.g. EGFR signalling and glycolysis



Standard formats generate new research

- Herrgård et al (2008) A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nature Biotechnol*, 26: 1155-1160



MODEL0072364382: 2152 species, 1857 reactions

- stoichiometric map, no concentrations, no kinetics
- Smallbone et al (2010) Towards a genome-scale kinetic model of cellular metabolism. *BMC Syst Biol*, 4:6



MODEL1001200000: 1748 species, 1059 reactions

- Concentrations and flux from BioModels Database
- Constraint-based model and simplified linlog kinetics
- Dobson et al (2010) Further developments towards a genome-scale metabolic model of yeast. *BMC Syst Biol*, 4:145



MODEL1012110000: 2657 species, 1865 reactions

- Li et al (2010) Systematic integration of experimental data and models in systems biology. *BMC Bioinfo*, 11: 582



MODEL1012110001

- Workflows using experimental kinetic information database (SABIO-RK) plus metabolomics and proteomics database
- Full quantitative chemical kinetics descriptions

Clustering models (and data) based on metadata

BIOMD00000000010_0
 BIOMD00000000009_Huang1996_MAPK_ultrasens)
 BIOMD00000000011_Levchenko2000_MAPK_noScaffold)
 BIOMD00000000014_Levchenko2000_MAPK_Scaffold)
 BIOMD00000000026_Markevich2004_MAPK_orderedElementary)
 BIOMD00000000028_Markevich2004_MAPK_phosphoRandomElementary)
 BIOMD00000000030_Markevich2004_MAPK_AllRandomElementary)
 BIOMD00000000029_Markevich2004_MAPK_phosphoRandomMM)
 BIOMD00000000027_Markevich2004_MAPK_orderedMM)
 BIOMD00000000031_Markevich2004_MAPK_orderedMM2kinases)
 BIOMD00000000032_Kofahl2004_pheromone)
 BIOMD00000000116_McClean2007_CrossTalk)
 BIOMD00000000084_Hornberg2005_ERKcascade)
 BIOMD00000000149_Kim2007_Wnt_ERK_Crosstalk)
 BIOMD00000000033_0
 BIOMD00000000048_Kholodenko1999_EGFRsignaling)
 BIOMD00000000049_Sasagawa2005_MAPK)
 BIOMD00000000205_Ung2008_EGFR_Endocytosis)
 BIOMD00000000019_0
 BIOMD00000000175_Birtwistle2007_ErbB_Signaling)

ATP:protein_phosphotransferase_(non-specific)
 RAF_proto-oncogene_serine/threonine-protein_kinase
 inactivation_of_MAPKKK_activity
 inactivation_of_MAPKK_activity
 protein_amino_acid_dephosphorylation
 protein_amino_acid_phosphorylation
 MAP_kinase_kinase_kinase_kinase_activity
 MAP_kinase_kinase_kinase_activity
 activation_of_MAPKKK_activity
 activation_of_MAPKK_activity
 Ras_small_GTPase,_Ras_type
 mitogen-activated_protein_kinase_kinase_kinase_binding
 urn:miriam:reactome:REACT_143
 urn:miriam:reactome:REACT_996
 urn:miriam:reactome:REACT_614
 Serine/threonine-protein_kinase_mos
 urn:miriam:reactome:REACT_525
 Mitogen-activated_protein_kinase_1
 ATP:protein_phosphotransferase_(MAPKKK-activated)
 MAP_kinase_kinase_activity
 activation_of_MAPK_activity
 inactivation_of_MAPK_activity
 Dual_specificity_mitogen-activated_protein_kinase_kinase_1
 urn:miriam:reactome:REACT_136
 urn:miriam:reactome:REACT_2247
 urn:miriam:uniprot:Q90W58
 phosphoprotein_phosphatase_activity
 mitogen-activated_protein_kinase_binding
 mitogen-activated_protein_kinase_kinase_binding
 urn:miriam:reactome:REACT_1780
 urn:miriam:reactome:REACT_495
 peptidyl-threonine_phosphorylation
 peptidyl-tyrosine_phosphorylation

Ranking and retrieval of models

Model	BioModel	Score	p-val num	
Huang1996_MAPK_ultrasens	BIOMD0000000009	1.000	< 1/1000	
Levchenko2000_MAPK_noScaffold	BIOMD0000000011	0.925	< 1/1000	
Levchenko2000_MAPK_Scaffold	BIOMD0000000014	0.865	< 1/1000	
Kholodenko2000_MAPK_feedback	BIOMD0000000010	0.816	< 1/1000	
Markevich2004_MAPK_orderedElementary	BIOMD0000000026	0.737	< 1/1000	
Markevich2004_MAPK_phosphoRandomElementary	BIOMD0000000028	0.690	< 1/1000	
Markevich2004_MAPK_AllRandomElementary	BIOMD0000000030	0.690	< 1/1000	
Markevich2004_MAPK_orderedMM	BIOMD0000000027	0.673	< 1/1000	
Markevich2004_MAPK_orderedMM2kinases	BIOMD0000000031	0.673	< 1/1000	
Markevich2004_MAPK_phosphoRandomMM	BIOMD0000000029	0.617	< 1/1000	
Hornberg2005_ERKcascade	BIOMD0000000084	0.468	< 1/1000	
McClean2007_CrossTalk	BIOMD0000000116	0.397	< 1/1000	
Kofahl2004_pheromone	BIOMD0000000032	0.348	< 1/1000	
Kim2007_Wnt_ERK_Crosstalk	BIOMD0000000149	0.323	< 1/1000	
Brown2004_NGF_EGF_signaling	BIOMD0000000033	0.260	< 1/1000	
Goldbeter1995_CircClock	BIOMD0000000016	0.244	< 1/1000	
Sasagawa2005_MAPK	BIOMD0000000049	0.240	< 1/1000	
Ung2008_EGFR_Endocytosis	BIOMD0000000205	0.234	< 1/1000	
Leloup1999_CircClock	BIOMD0000000021	0.229	< 1/1000	
Goldbeter1991_MinMitOscil_ExpInact	BIOMD0000000004	0.222	< 1/1000	
Goldbeter1991_MinMitOscil	BIOMD0000000003	0.214	< 1/1000	
Leloup1998_CircClock_LD	BIOMD0000000171	0.201	< 1/1000	
Tyson1991_CellCycle_6var	BIOMD0000000005	0.199	< 1/1000	
Veening2008_DegU_Regulation	BIOMD0000000240	0.180	< 1/1000	
Neves2008_Cell_Shape	BIOMD0000000182	0.168	4.0e-03	
Fisher2006_Ca_Oscillation_dpdnt_NFAT_dynamics	BIOMD0000000122	0.164	5.0e-03	
Leloup2003_CircClock_DD	BIOMD0000000073	0.163	5.0e-03	

Ranking and retrieval of models

Model	BioModel	Score	p-val num	
Huang1996_MAPK_ultrasens	BIOMD0000000009	1.000	< 1/1000	
Levchenko2000_MAPK_noScaffold	BIOMD0000000011	0.925	< 1/1000	
Levchenko2000_MAPK_Scaffold	BIOMD0000000014	0.865	< 1/1000	
Kholodenko2000_MAPK_feedback	BIOMD0000000010	0.816	< 1/1000	
Markevich2004_MAPK_orderedElementary	BIOMD0000000026	0.737	< 1/1000	
Markevich2004_MAPK_phosphoRandomElementary	BIOMD0000000028	0.690	< 1/1000	
Markevich2004_MAPK_AllRandomElementary	BIOMD0000000030	0.690	< 1/1000	
Markevich2004_MAPK_orderedMM	BIOMD0000000027	0.673	< 1/1000	
Markevich2004_MAPK_orderedMM2kinases	BIOMD0000000031	0.673	< 1/1000	
Markevich2004_MAPK_phosphoRandomMM	BIOMD0000000029	0.617	< 1/1000	
Hornberg2005_ERKcascade	BIOMD0000000084	0.468	< 1/1000	
McClean2007_CrossTalk	BIOMD0000000116	0.397	< 1/1000	
Kofahl2004_pheromone	BIOMD0000000032	0.348	< 1/1000	
Kim2007_Wnt_ERK_Crosstalk	BIOMD0000000149	0.323	< 1/1000	
Brown2004_NGF_EGF_signaling	BIOMD0000000033	0.260	< 1/1000	
Goldbeter1995_CircClock	BIOMD0000000016	0.244	< 1/1000	
Sasagawa2005_MAPK	BIOMD0000000049	0.240	< 1/1000	
Ung2008_EGFR_Endocytosis	BIOMD0000000205	0.234	< 1/1000	
Leloup1999_CircClock	BIOMD0000000021	0.229	< 1/1000	
Goldbeter1991_MinMitOscil_ExpInact	BIOMD0000000004	0.222	< 1/1000	
Goldbeter1991_MinMitOscil	BIOMD0000000003	0.214	< 1/1000	
Leloup1998_CircClock_LD	BIOMD0000000171	0.201	< 1/1000	
Tyson1991_CellCycle_6var	BIOMD0000000005	0.199	< 1/1000	
Veening2008_DegU_Regulation	BIOMD0000000240	0.180	< 1/1000	
Neves2008_Cell_Shape	BIOMD0000000182	0.168	4.0e-03	
Fisher2006_Ca_Oscillation_dpdnt_NFAT_dynamics	BIOMD0000000122	0.164	5.0e-03	
Leloup2003_CircClock_DD	BIOMD0000000073	0.163	5.0e-03	

MAPK

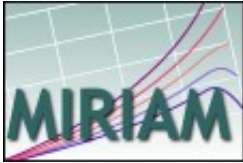
Ranking and retrieval of models

Model	BioModel	Score	p-val num	
Huang1996_MAPK_ultrasens	BIOMD0000000009	1.000	< 1/1000	
Levchenko2000_MAPK_noScaffold	BIOMD0000000011	0.925	< 1/1000	
Levchenko2000_MAPK_Scaffold	BIOMD0000000014	0.865	< 1/1000	
Kholodenko2000_MAPK_feedback	BIOMD0000000010	0.816	< 1/1000	
Markevich2004_MAPK_orderedElementary	BIOMD0000000026	0.737	< 1/1000	
Markevich2004_MAPK_phosphoRandomElementary	BIOMD0000000028	0.690	< 1/1000	
Markevich2004_MAPK_AllRandomElementary	BIOMD0000000030	0.690	< 1/1000	
Markevich2004_MAPK_orderedMM	BIOMD0000000027	0.673	< 1/1000	
Markevich2004_MAPK_orderedMM2kinases	BIOMD0000000031	0.673	< 1/1000	
Markevich2004_MAPK_phosphoRandomMM	BIOMD0000000029	0.617	< 1/1000	
Hornberg2005_ERKcascade	BIOMD0000000084	0.468	< 1/1000	
McClearn2007_CrossTalk	BIOMD0000000116	0.397	< 1/1000	
Kofahl2004_pheromone	BIOMD0000000032	0.348	< 1/1000	
Kim2007_Wnt_ERK_Crosstalk	BIOMD0000000149	0.323	< 1/1000	
Brown2004_NGF_EGF_signaling	BIOMD0000000033	0.260	< 1/1000	
Goldbeter1995_CircClock	BIOMD0000000016	0.244	< 1/1000	
Sasagawa2005_MAPK	BIOMD0000000049	0.240	< 1/1000	
Ung2008_EGFR_Endocytosis	BIOMD0000000205	0.234	< 1/1000	
Leloup1999_CircClock	BIOMD0000000021	0.229	< 1/1000	
Goldbeter1991_MinMitOscil_ExpInact	BIOMD0000000004	0.222	< 1/1000	
Goldbeter1991_MinMitOscil	BIOMD0000000003	0.214	< 1/1000	
Leloup1998_CircClock_LD	BIOMD0000000171	0.201	< 1/1000	
Tyson1991_CellCycle_6var	BIOMD0000000005	0.199	< 1/1000	
Veening2008_DegU_Regulation	BIOMD0000000240	0.180	< 1/1000	
Neves2008_Cell_Shape	BIOMD0000000182	0.168	4.0e-03	
Fisher2006_Ca_Oscillation_dpdnt_NFAT_dynamics	BIOMD0000000122	0.164	5.0e-03	
Leloup2003_CircClock_DD	BIOMD0000000073	0.163	5.0e-03	

MAPK

Other kinase cascades

The interface with all biologists

	Model descriptions
Minimal requirements	
Data-models	  BioPAX
Terminologies	

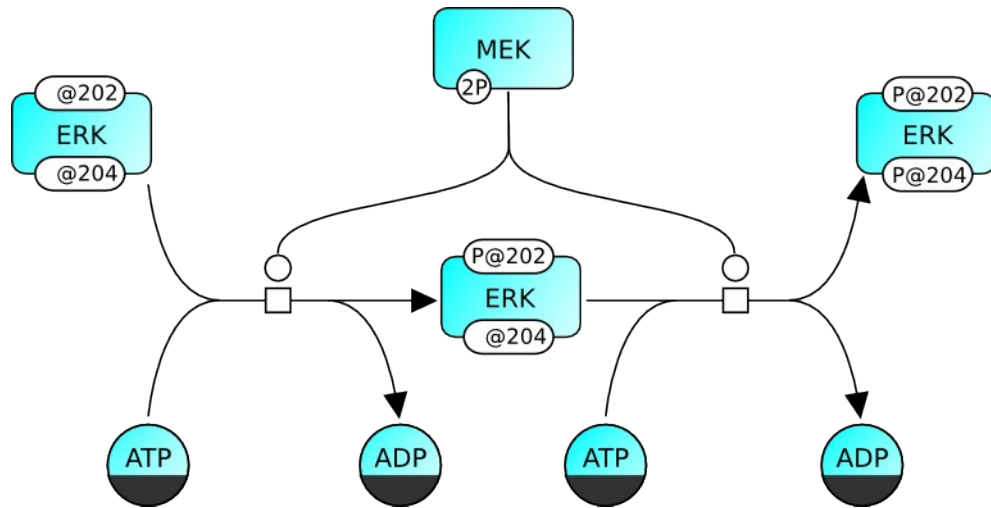
Born in Tokyo 2005

What is SBGN?

- An unambiguous way of graphically describing and interpreting biochemical and cellular events
- Limited amount of symbols
Re-use existing symbols
-  Smooth learning curve
- Can represent logical or mechanistic models, biochemical pathways, at different levels of granularity
- Detailed technical specification, precise data-models and growing software support
- Developed over four years by a diverse community, including biologists, modellers, computer scientists etc.

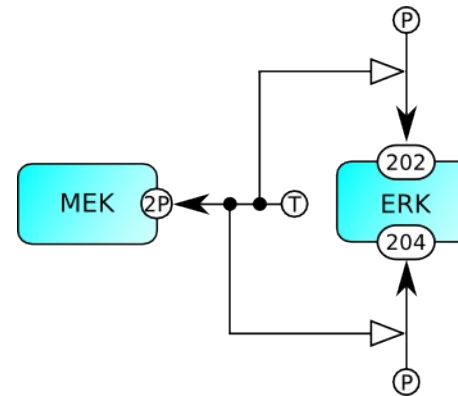
Graph trinity: three languages in one notation

Process Descriptions



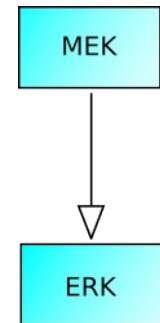
- Unambiguous
- Mechanistic
- Sequential
- Combinatorial explosion

Entity Relationships



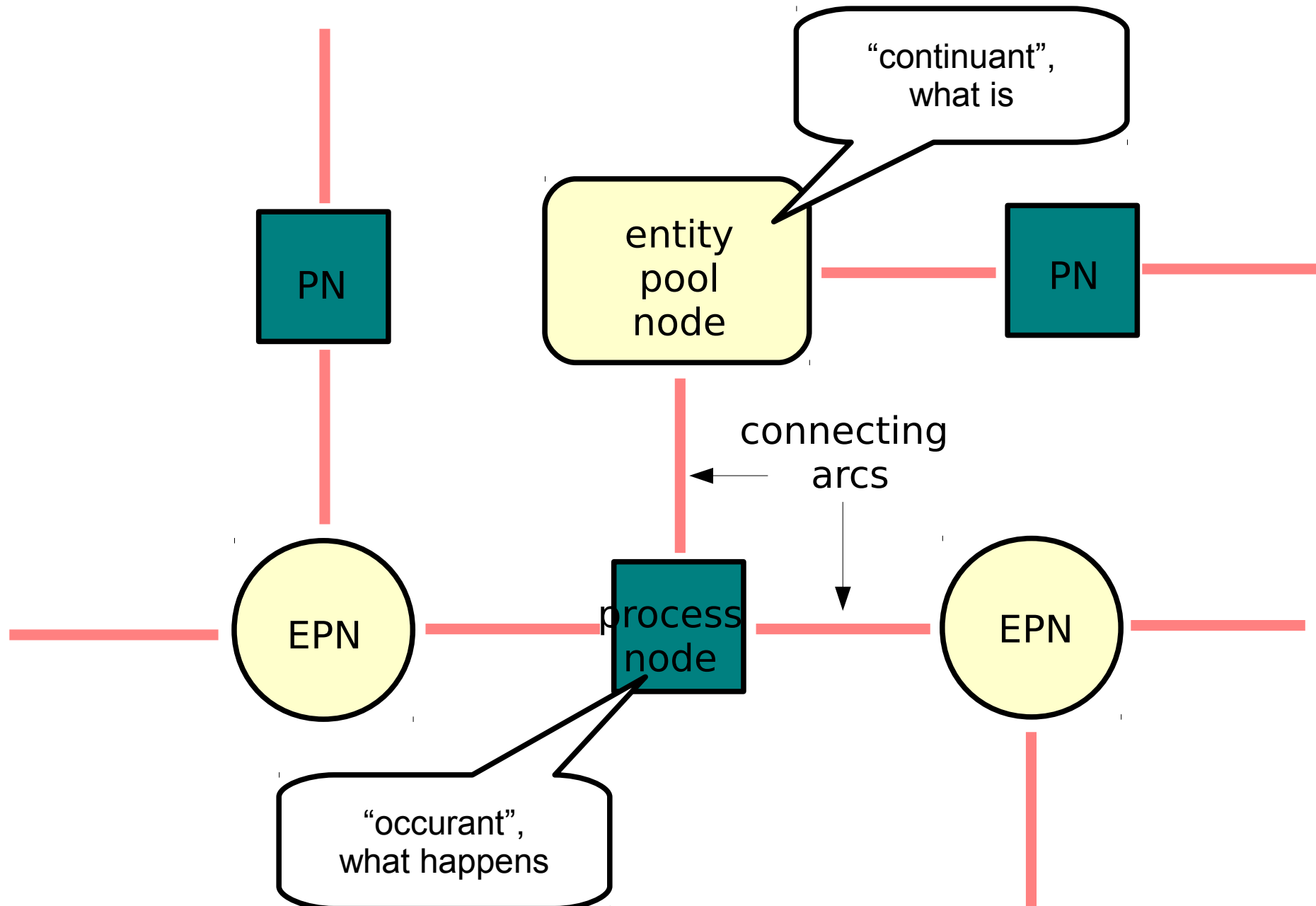
- Unambiguous
- Mechanistic
- Non-sequential
- Independence of relationships

Activity Flows

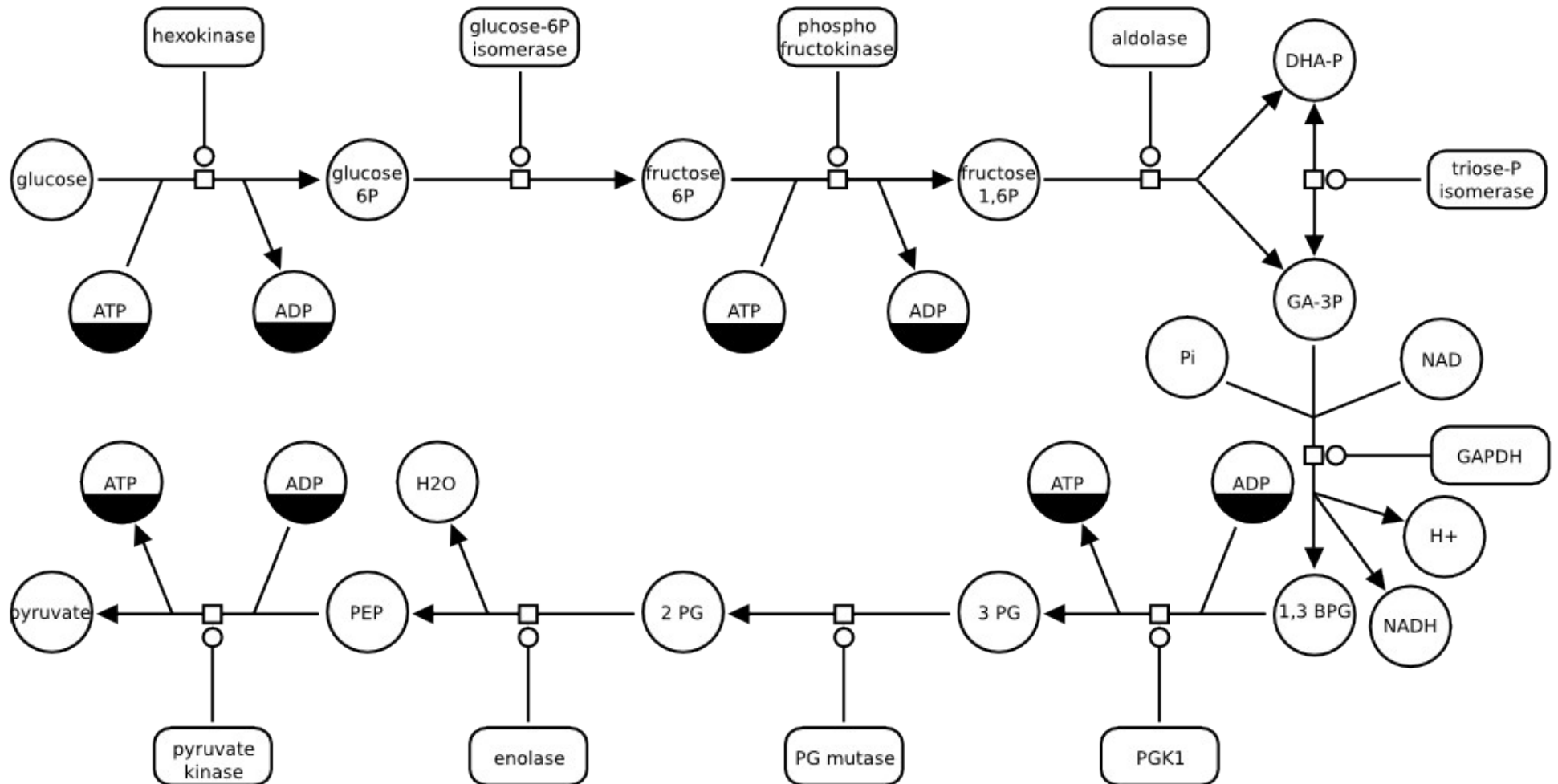


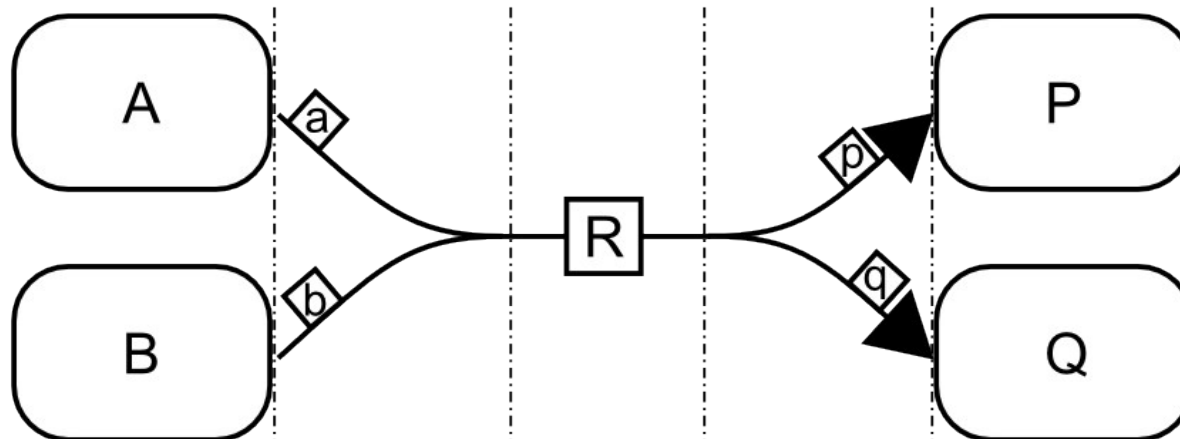
- Ambiguous
- Conceptual
- Sequential

Process Descriptions are bipartite graphs



Metabolic network in Process Description Language





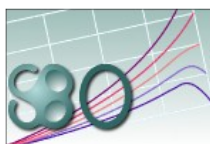
substances
A and B

are
consumed by

reaction R

that
produces

substances
P and Q



SBO:0000240
material entity

SBO:0000394
consumption

SBO:0000375
process

SBO:0000393
production

SBO:0000240
material entity



Species

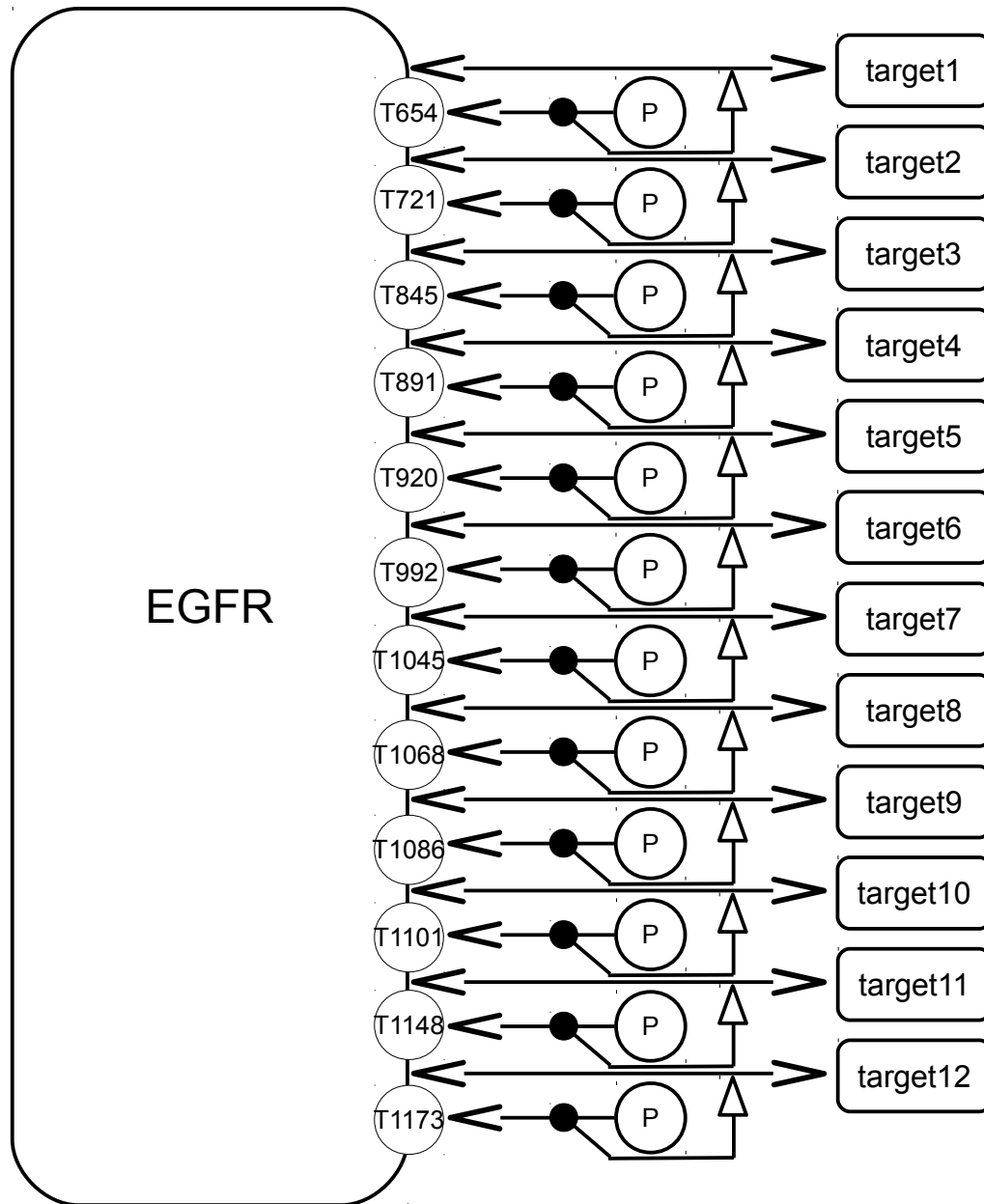
Species
Reference

Kinetic
Law

Species
Reference

Species

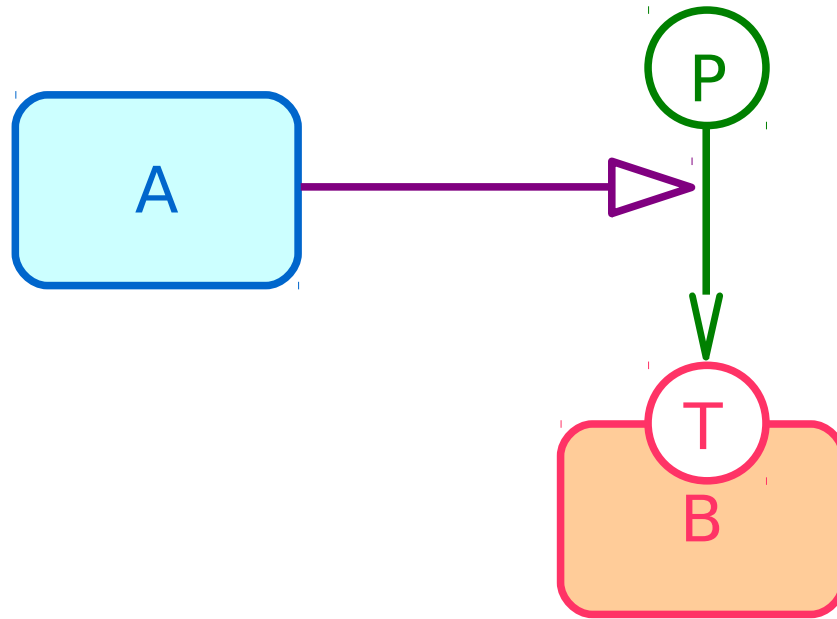
Multi-state and combinatorial explosion



Process Descriptions:
“once a state variable value,
always a state variable value”

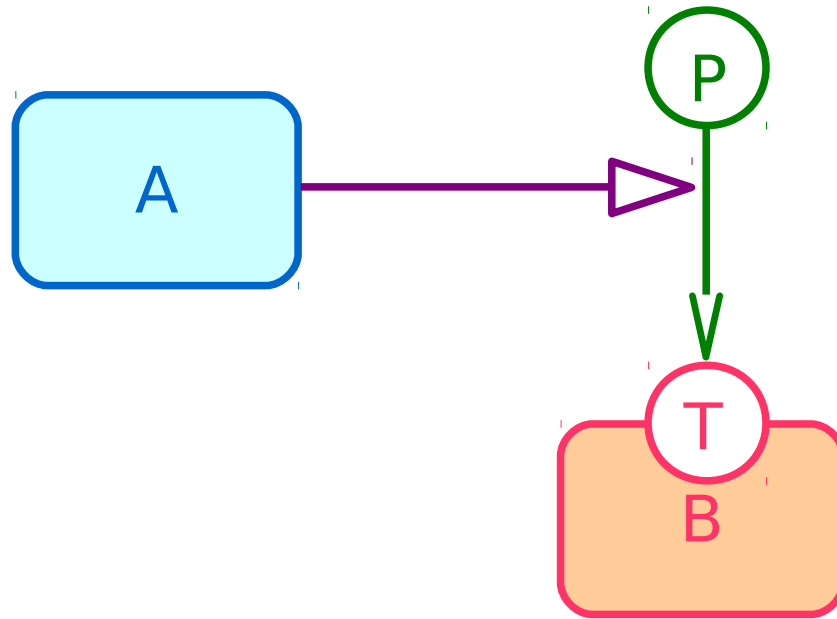
$2^{12} = 4096$ states
(i.e. EPN glyphs) for EGFR
and 4096 complexes between
EGFR and targets

Entity Relationships can be viewed as rules



If **A exists**, the **assignment of the value P** to the **state variable T of B** is **increased**

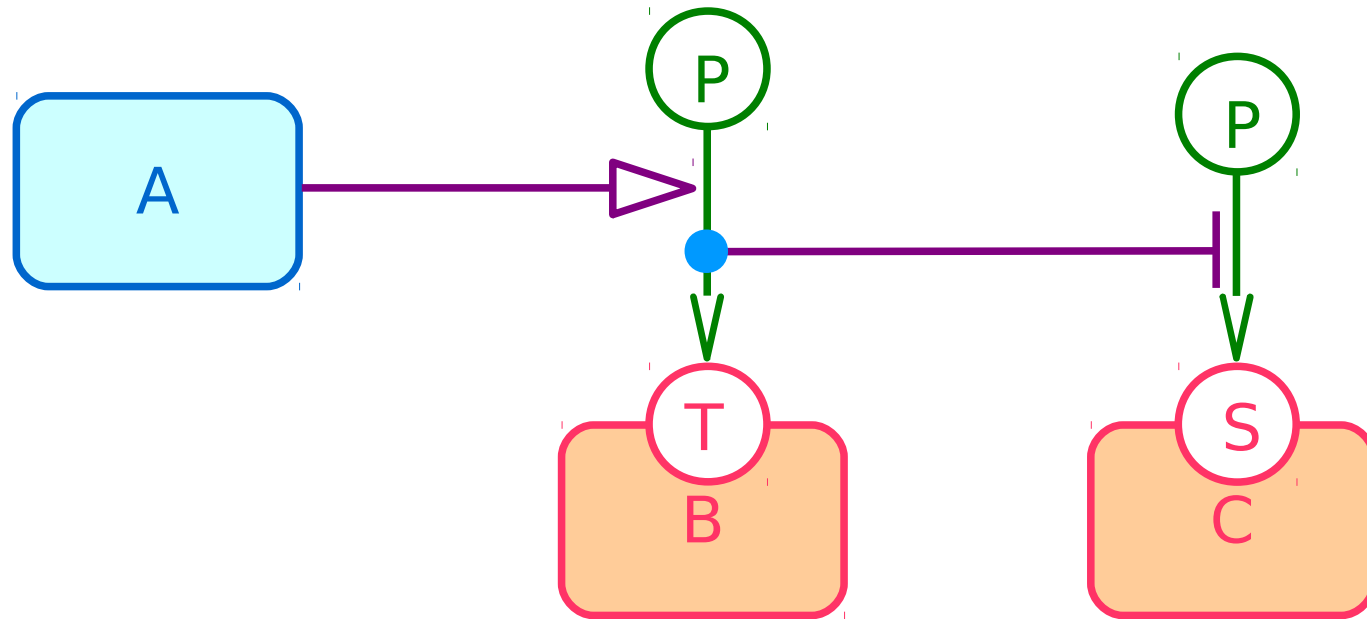
Entity Relationships can be viewed as rules



If **A exists**, the **assignment of the value P** to the **state variable T of B** is **increased**

(**A stimulates** the **phosphorylation** of **B on the threonine**)

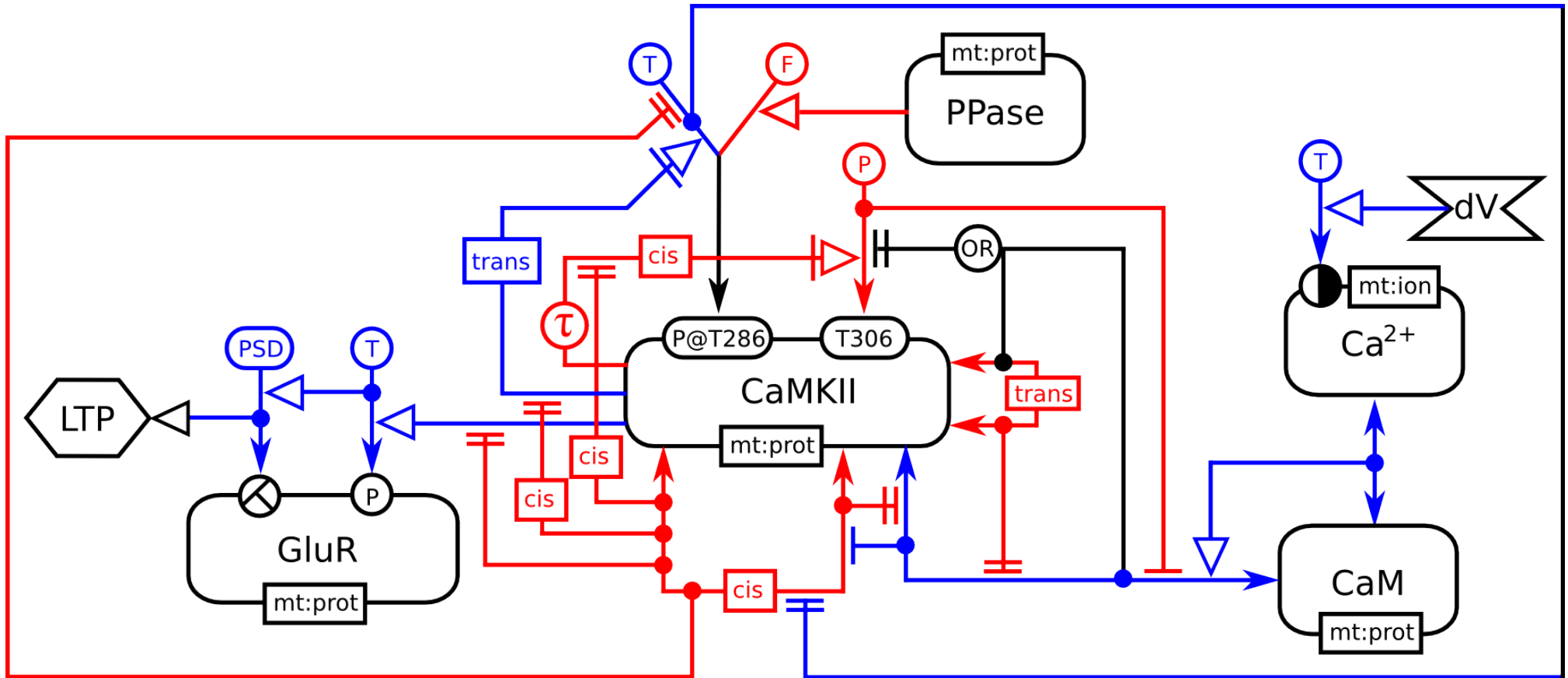
Entity Relationships can be viewed as rules



If **A exists**, the **assignment of the value P** to the **state variable T of B** is **increased**

If **P** is assigned to the **state variable T of B**, the **assignment of the value P** to the **state variable S of B** is **decreased**

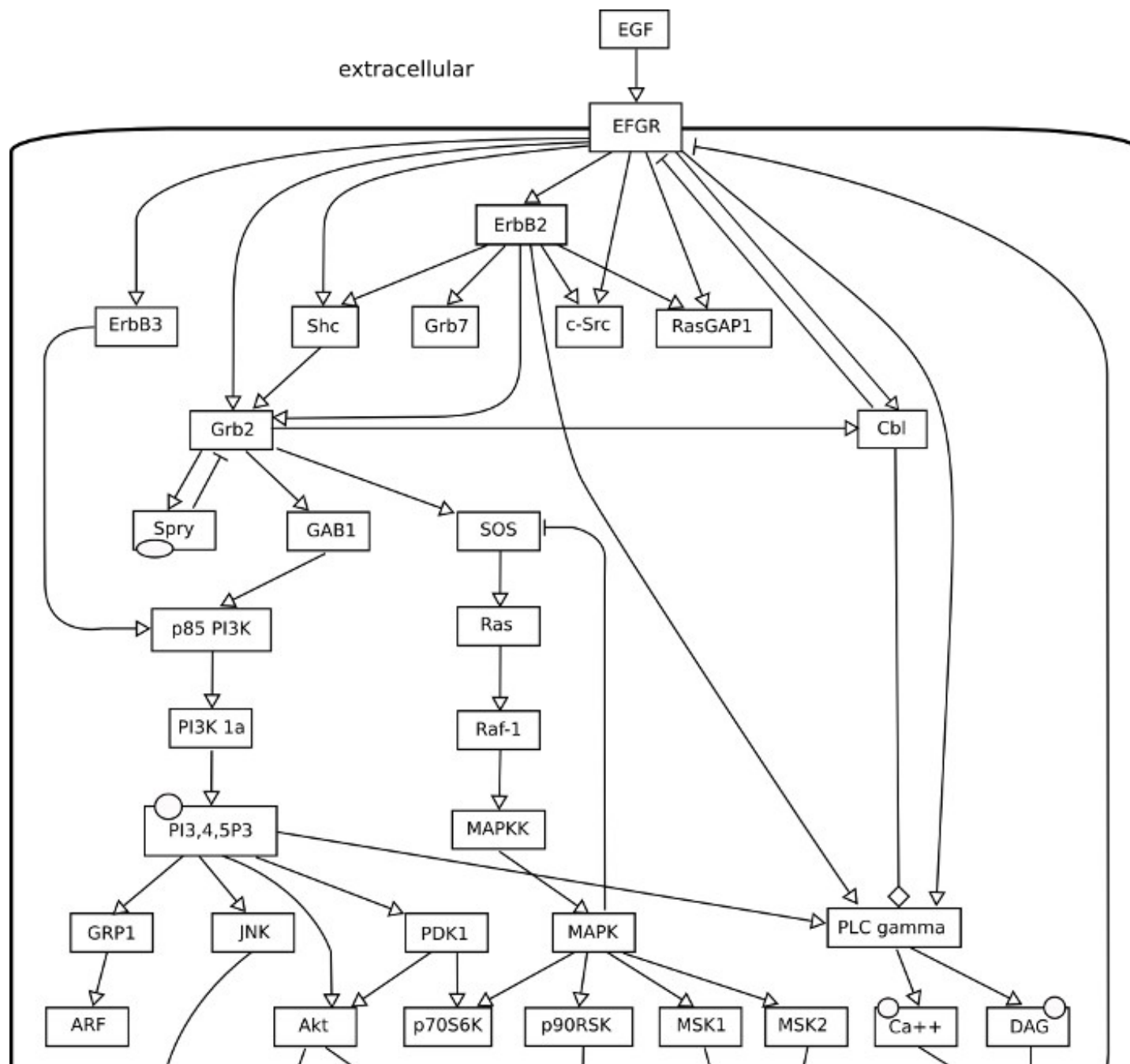
GN ER map of calcium-regulated synaptic plasticity



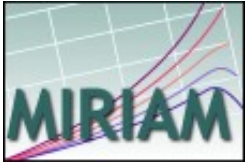
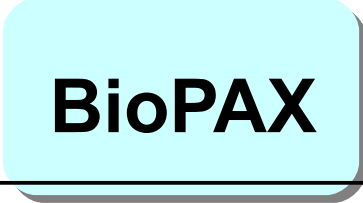
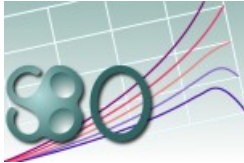
increases synaptic weight

decreases synaptic weight

Example of Activity Flow map



Merging biologist semantics for database

	Model descriptions
Minimal requirements	
Data-models	  
Terminologies	

Biological Pathway Exchange format - BioPAX

- **RDF/OWL**-based language to represent biological pathways at the molecular and cellular level
- Covers molecular and genetic interactions, metabolic network, signalling pathways, gene regulatory networks
- **Interactions**: Control (Catalysis, Modulation, TemplateReactionRegulation), Conversion (BiochemicalReaction, ComplexAssembly, Degradation, Transport, TransportWithBiochemicalReaction), GeneticInteraction, MolecularInteraction, TemplateReaction
- **Physical entities**: Complex, DNA, DNARegion, Protein, RNA, RNARegion, SmallMolecule
- **Utility classes**: BioSource, ChemicalStructure, ControlledVocabulary, DeltaG, EntityFeature, EntityReference, Evidence, ExperimentalForm, kPrime, PathwayStep, Provenance, Score, SequenceLocation, Stoichiometry, Xref

Resource Description Framework - RDF

- Language of the semantic web
- Semantic entirely embedded in the language (does not require external schemas and specifications)
- Series of statements

Subject

Predicate

Object

EGFR

located_in

plasma membrane

P00533

OBO_REL:0000008

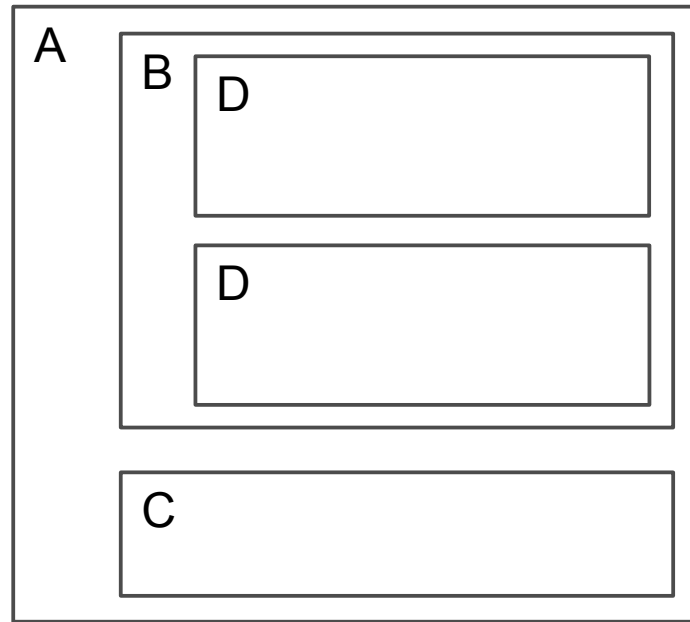
GO:0005886

(UniProt)

(OBO Relation ontology)

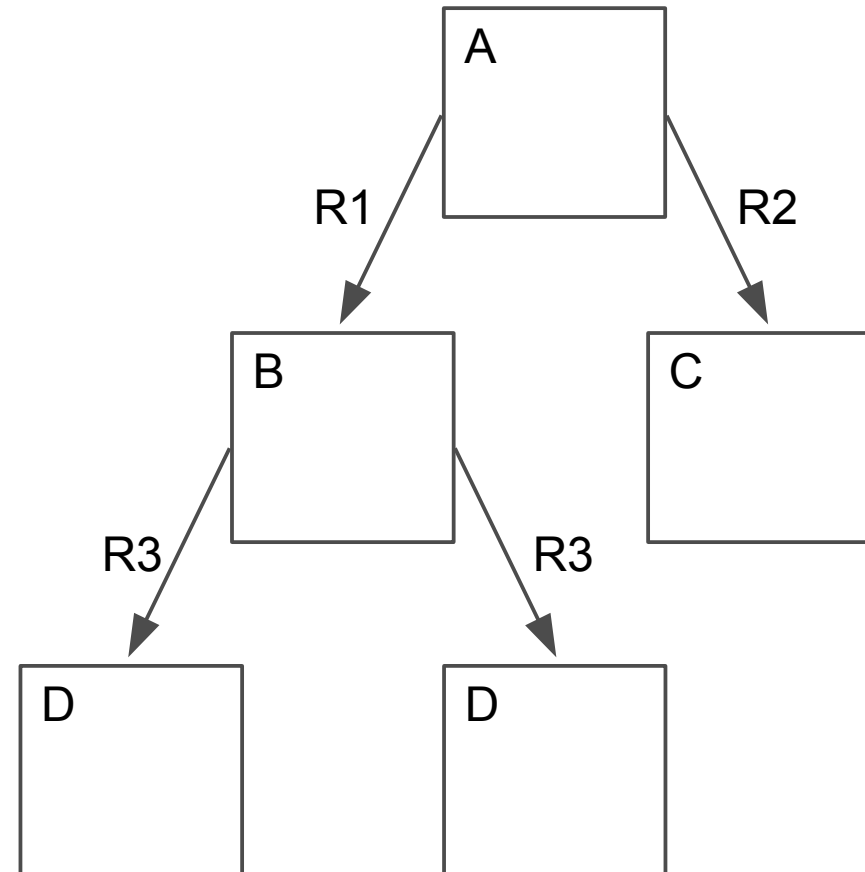
(Gene Ontology)

“XML”



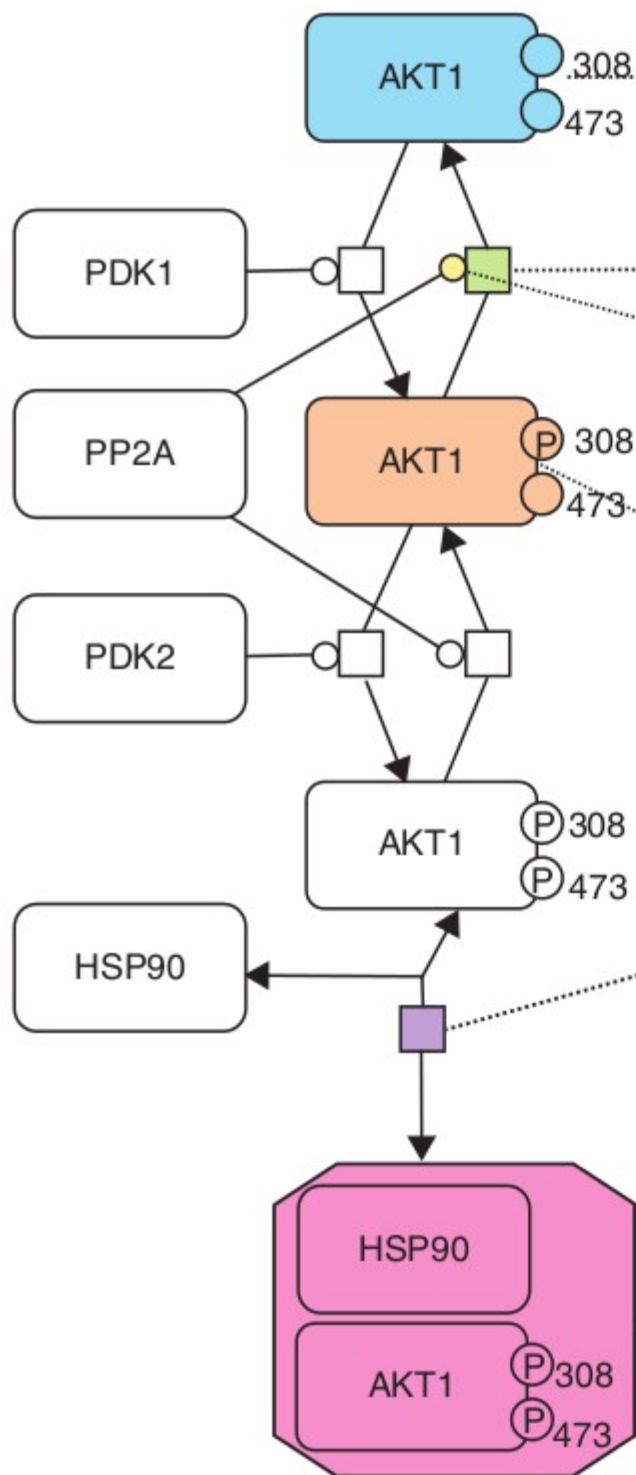
Vs

RDF



Web Ontology Language - OWL

- BioPAX contains classes, relations and properties, constraints, objects, values.
- Classes are arranged in a specialisation hierarchy. E.g. a “protein” is a “physical entity”
- Classes may have properties of specific types. The types are linked to other classes by relations. E.g. “reaction X” has *participant* “protein P”. An attribute is a property that has a simple type.
- Constraints define allowable values and connexions. E.g. “MOLECULAR_WT” must be a positive real number
- Objects are instances of classes.



AKT1.1 is a *Protein*
 has *proteinReference* rAKT1
 has *notFeature* p@308
 has *notFeature* p@473

reaction1 is a *BiochemicalReaction*
 has *left* AKT1.2
 has *right* AKT1.1
 is *left-to-right*.

catalysis1 is a *Catalysis*
 has *controller* PP2A.1
 has *controlled* reaction1
 has *direction* irr-left-to-right

AKT1.2 is a *Protein*
 has *proteinReference* rAKT1
 has *feature* p@308
 has *notFeature* p@473

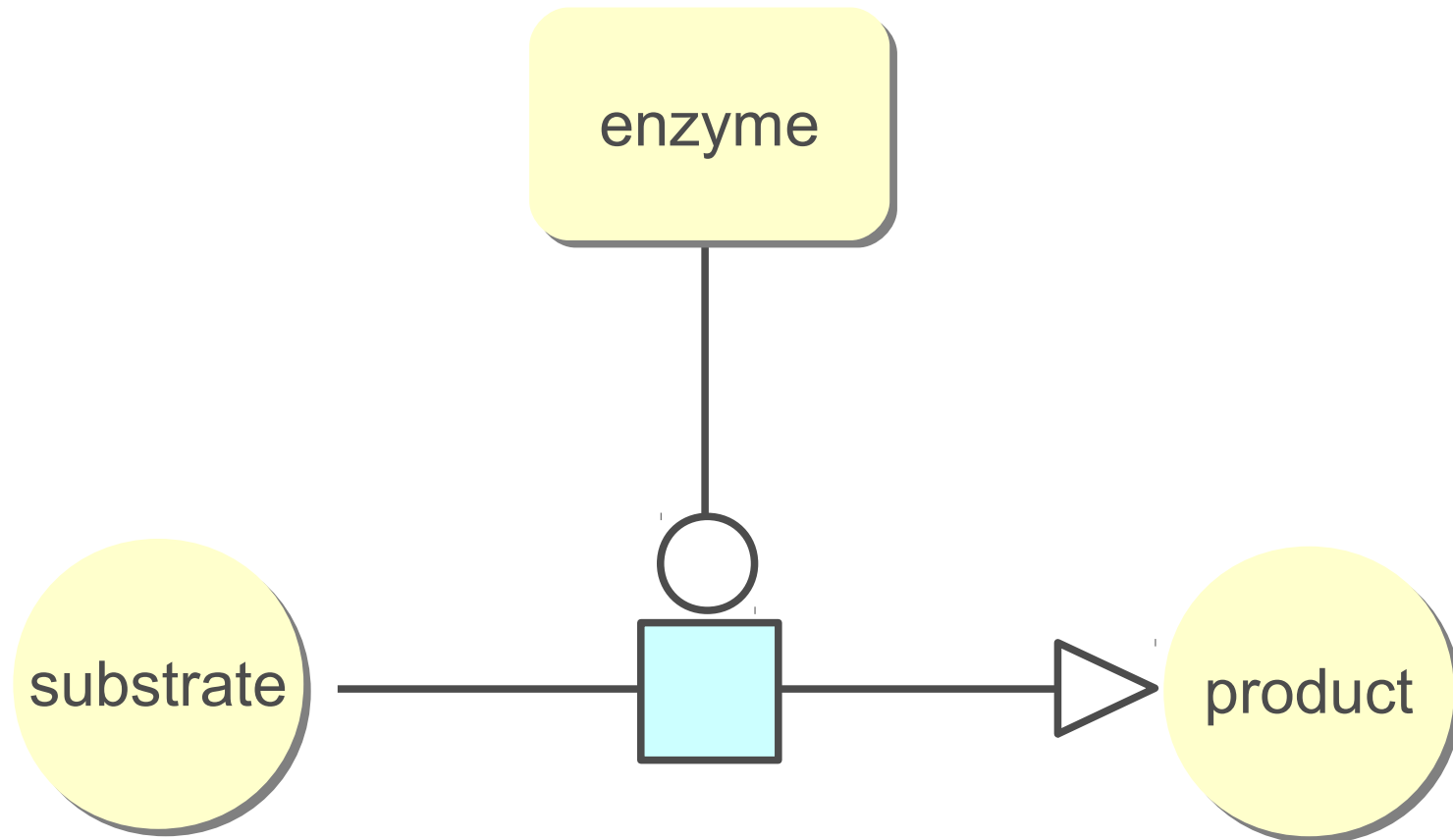
assembly1 is a *ComplexAssembly*
 has *left* HSP90.1
 has *left* AKT1.3
 has *right* complex1
 is *reversible*

complex1 is a *Complex*
 has *component* AKT1.4
 has *component* HSP90.2

HSP90.2 is a *Protein*
 has *proteinReference* rHSP90
 is *boundTo* AKT1.4

AKT1.4 is a *Protein*
 has *proteinReference* rAKT1
 has *feature* p@308
 has *feature* p@473
 is *boundTo* HSP90.2

A catalysis in BioPAX



A catalysis in BioPAX

```
<!-- physicalEntities -->
```

```
<bp:smallMolecule rdf:ID="smallMolecule1" />  
<bp:smallMolecule rdf:ID="smallMolecule2" />  
<bp:protein rdf:ID="protein1">
```

```
<!-- physicalEntityParticipants -->
```

```
<bp:physicalEntityParticipant rdf:ID="substrate">  
  <bp:STOICHIOMETRIC-COEFFICIENT>1.0</bp:STOICHIOMETRIC-COEFFICIENT>  
  <bp:PHYSICAL-ENTITY rdf:resource="#smallMolecule1" />  
</bp:physicalEntityParticipant>  
<bp:physicalEntityParticipant rdf:ID="product">  
  <bp:STOICHIOMETRIC-COEFFICIENT>1.0</bp:STOICHIOMETRIC-COEFFICIENT>  
  <bp:PHYSICAL-ENTITY rdf:resource="#smallMolecule2" />  
</bp:physicalEntityParticipant>  
<bp:physicalEntityParticipant rdf:ID="enzyme">  
  <bp:PHYSICAL-ENTITY rdf:resource="#protein1" />  
</bp:physicalEntityParticipant>
```

```
<!-- physicalInteractions -->
```

```
<bp:biochemicalReaction rdf:ID="biochemicalReaction1">  
  <bp:LEFT rdf:resource="#substrate">  
  <bp:RIGHT rdf:resource="#product">  
</bp:biochemicalReaction>  
<bp:catalysis rdf:ID="catalysis1">  
  <bp:CONTROLLER rdf:resource="#enzyme" />  
  <bp:CONTROLLED rdf:resource="#biochemicalReaction1">  
</bp:catalysis>
```

A catalysis in BioPAX

```
<!-- physicalEntityParticipants -->
```

```
<bp:physicalEntityParticipant rdf:ID="substrate">
  <bp:STOICHIOMETRIC-COEFFICIENT>1.0</bp:STOICHIOMETRIC-COEFFICIENT>
  <bp:PHYSICAL-ENTITY>
    <bp:smallMolecule rdf:ID="smallMolecule1" />
  </bp:PHYSICAL-ENTITY>
</bp:physicalEntityParticipant>
<bp:physicalEntityParticipant rdf:ID="product">
  <bp:STOICHIOMETRIC-COEFFICIENT>1.0</bp:STOICHIOMETRIC-COEFFICIENT>
  <bp:PHYSICAL-ENTITY>
    <bp:smallMolecule rdf:ID="smallMolecule2" />
  </bp:PHYSICAL-ENTITY>
</bp:physicalEntityParticipant>
<bp:physicalEntityParticipant rdf:ID="enzyme">
  <bp:PHYSICAL-ENTITY>
    <bp:protein rdf:ID="protein1">
    </bp:PHYSICAL-ENTITY>
  </bp:PHYSICAL-ENTITY>
</bp:physicalEntityParticipant>
```

```
<!-- physicalInteractions -->
```

```
<bp:biochemicalReaction rdf:ID="biochemicalReaction1">
  <bp:LEFT rdf:resource="#substrate">
  <bp:RIGHT rdf:resource="#product">
</bp:biochemicalReaction>
<bp:catalysis rdf:ID="catalysis1">
  <bp:CONTROLLER rdf:resource="#enzyme" />
  <bp:CONTROLLED rdf:resource="#biochemicalReaction1">
</bp:catalysis>
```

A catalysis in BioPAX

```
<!-- physicalInteractions -->
```

```
<bp:biochemicalReaction rdf:ID="biochemicalReaction1">
  <bp:LEFT>
    <bp:physicalEntityParticipant>
      <bp:STOICHIOMETRIC-COEFFICIENT>1.0</bp:STOICHIOMETRIC-COEFFICIENT>
      <bp:PHYSICAL-ENTITY>
        <bp:smallMolecule rdf:ID="smallMolecule1" />
      </bp:PHYSICAL-ENTITY>
    </bp:physicalEntityParticipant>
  </bp:LEFT>
  <bp:RIGHT>
    <bp:physicalEntityParticipant rdf:ID="product">
      <bp:STOICHIOMETRIC-COEFFICIENT>1.0</bp:STOICHIOMETRIC-COEFFICIENT>
      <bp:PHYSICAL-ENTITY>
        <bp:smallMolecule rdf:ID="smallMolecule2" />
      </bp:PHYSICAL-ENTITY>
    </bp:physicalEntityParticipant>
  </bp:RIGHT>
</bp:biochemicalReaction>
```

```
<bp:catalysis rdf:ID="catalysis1">
  <bp:CONTROLLER/>
  <bp:physicalEntityParticipant rdf:ID="enzyme">
    <bp:PHYSICAL-ENTITY>
      <bp:protein rdf:ID="protein1">
      </bp:PHYSICAL-ENTITY>
    </bp:physicalEntityParticipant>
  </bp:CONTROLLER>
  <bp:CONTROLLED rdf:resource="#biochemicalReaction1"
</bp:catalysis>
```

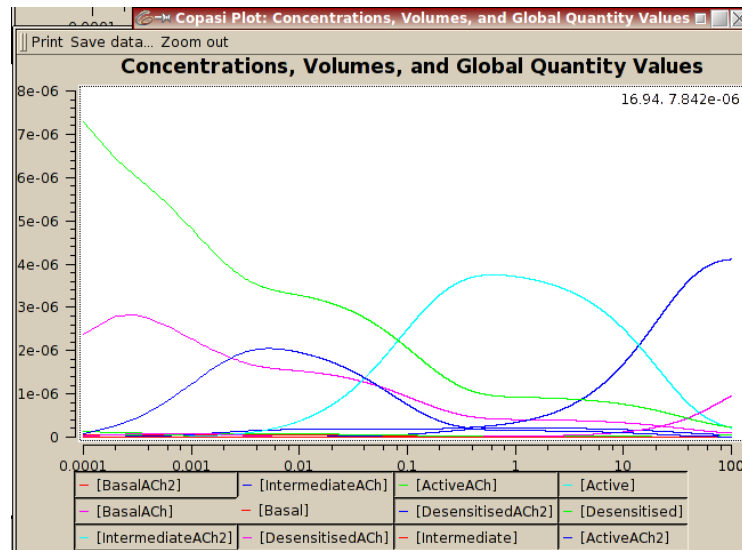
A catalysis in BioPAX

```
<!-- physicalInteractions -->
```

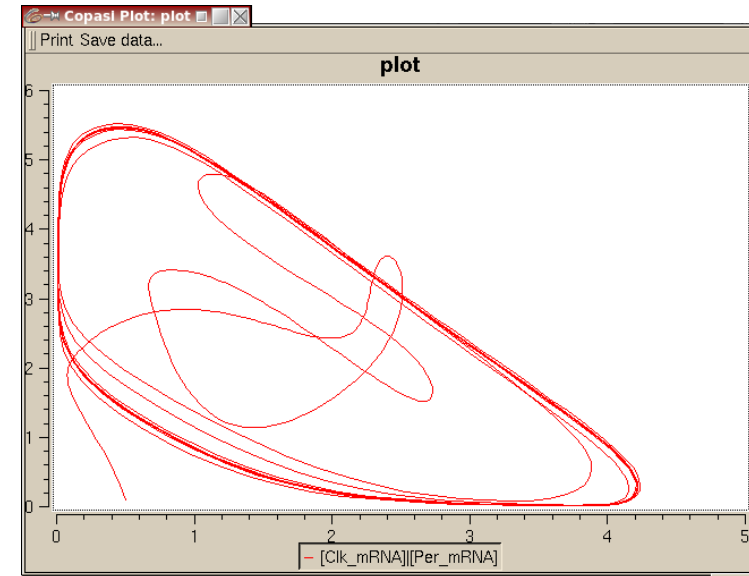
```
<bp:catalysis rdf:ID="catalysis1">
  <bp:CONTROLLER/>
    <bp:physicalEntityParticipant rdf:ID="enzyme">
      <bp:PHYSICAL-ENTITY>
        <bp:protein rdf:ID="protein1">
        </bp:PHYSICAL-ENTITY>
      </bp:physicalEntityParticipant>
    </bp:CONTROLLER>
    <bp:CONTROLLED>
      <bp:biochemicalReaction rdf:ID="biochemicalReaction1">
        <bp:LEFT>
          <bp:physicalEntityParticipant>
            <bp:STOICHIOMETRIC-COEFFICIENT>1.0</bp:STOICHIOMETRIC-COEFFICIENT>
            <bp:PHYSICAL-ENTITY>
              <bp:smallMolecule rdf:ID="smallMolecule1" />
            </bp:PHYSICAL-ENTITY>
          </bp:physicalEntityParticipant>
        </bp:LEFT>
        <bp:RIGHT>
          <bp:physicalEntityParticipant rdf:ID="product">
            <bp:STOICHIOMETRIC-COEFFICIENT>1.0</bp:STOICHIOMETRIC-COEFFICIENT>
            <bp:PHYSICAL-ENTITY>
              <bp:smallMolecule rdf:ID="smallMolecule2" />
            </bp:PHYSICAL-ENTITY>
          </bp:physicalEntityParticipant>
        </bp:RIGHT>
      </bp:biochemicalReaction>
    </bp:CONTROLLED>
  </bp:catalysis>
```

Reproduction of published simulation results

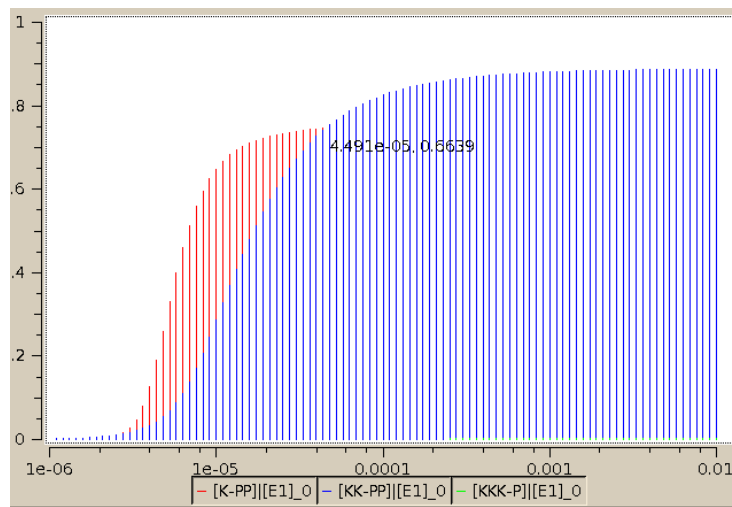
Edelstein et al 1996 (BIOMD0000000002)



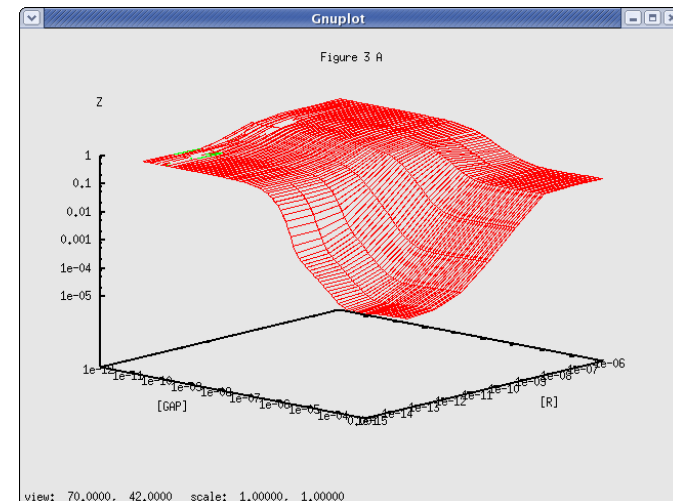
Ueda, Hagiwara, Kitano 2001 (BIOMD0000000022)



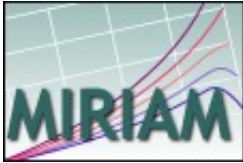
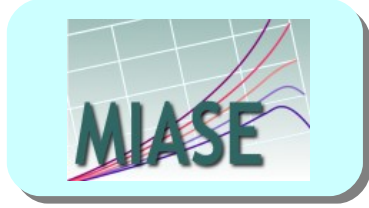
Huang & Ferrell (BIOMD0000000009)



Bornheimer et al 2004 (BIOMD0000000086)



Description of simulations and analyses

	Model descriptions	Simulations and analysis
Minimal requirements		
Data-models	  BioPAX	
Terminologies		

Born in Hinxton 2007

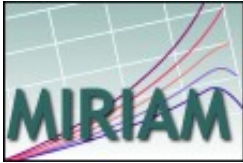
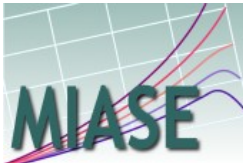
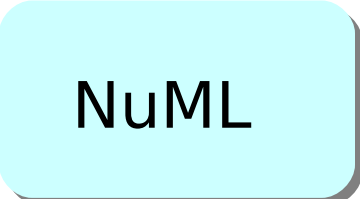
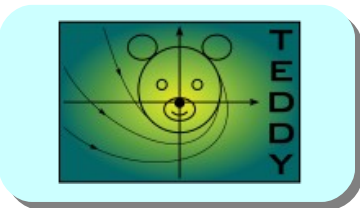
Description of model simulation and analysis

Minimum Information About a Simulation Experiment (MIASE) common set of information a modeller needs to provide in order to enable the execution and reproduction of a numerical simulation experiment, derived from a given set of quantitative models

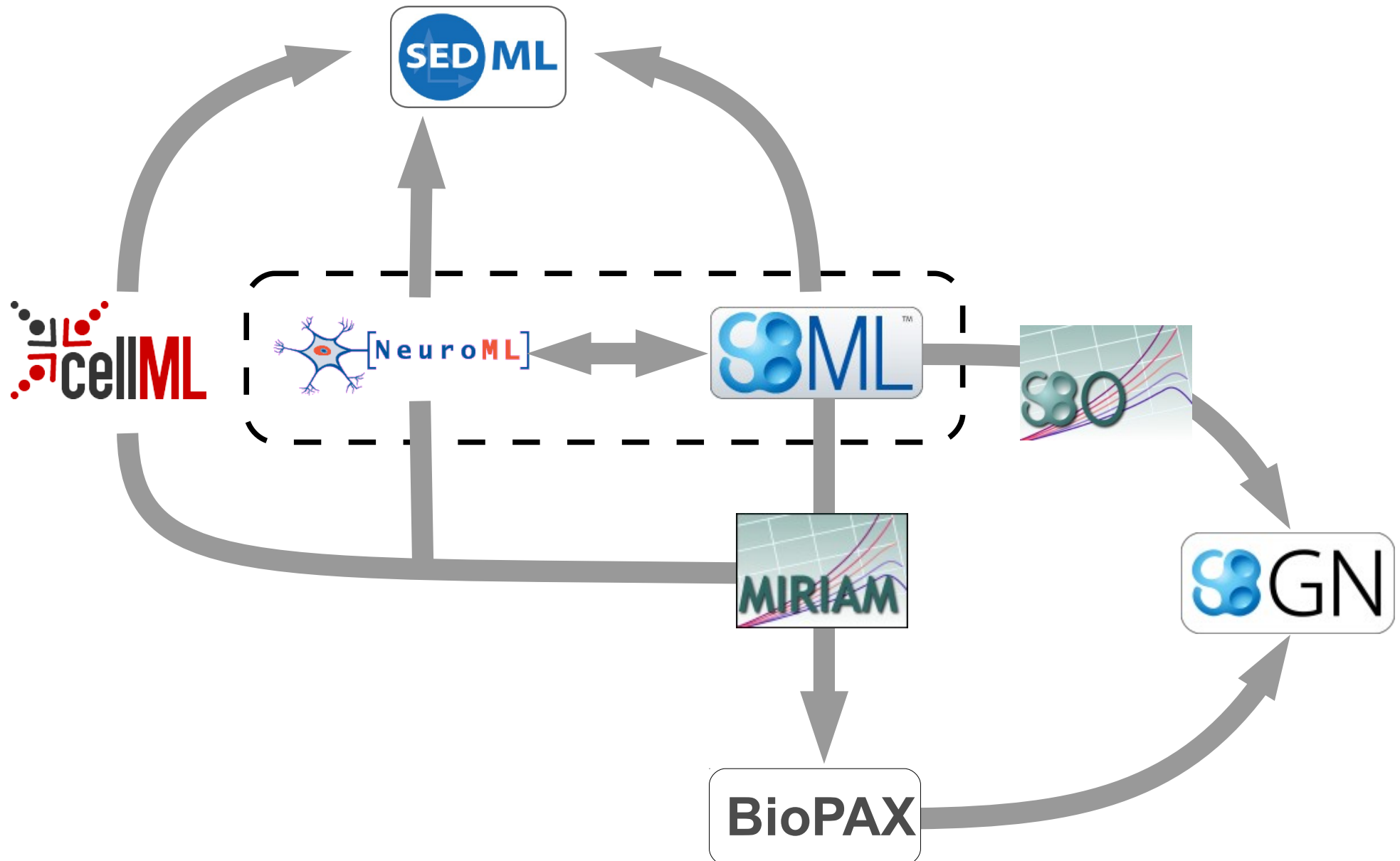
Simulation Experiment Description Markup Language (SED-ML) XML-based format for encoding simulation experiments, following the requirements defined in the MIASE guidelines

Kinetic Simulation Algorithm Ontology (KiSAO) covers the most important simulation algorithms and simulation methods used to simulate biological kinetic models and puts those algorithms and methods into relation

Characterising dynamical behaviours

	Model descriptions	Simulations and analysis	results
Minimal requirements			
Data-models	  BioPAX		
Terminologies			

Existing standards interoperability



Overarching standardisation structure



<http://co.mbine.org/>

- ISCB 2013 end-user workshop (presentations and tutorials)
 - 4 September 2013, Copenhagen
 - http://co.mbine.org/events/ISCB_2013
- COMBINE 2013 (standard further development)
 - 16 to 20 September 2013, Paris
 - http://co.mbine.org/events/COMBINE_2013
- HARMONY 2014 (development of software support)
 - 22 to 25 April 2014, Manchester
 - http://co.mbine.org/events/HARMONY_2014

Where to find more information?

Communities

Semantics

Coordination

<http://co.mbine.org/>

<http://biopax.org/>

<http://sbgn.org/>

<http://sbml.org/>

<http://sed-ml.org/>

<http://biomodels.net/>

<http://biomodels.net/biomodels/>

<http://biomodels.net/kisao>

<http://biomodels.net/sbo>

<http://biomodels.net/teddy>

<http://biomodels.net/miase>

<http://biomodels.net/miriam>

