

BioUML – open source extensible workbench for systems biology

<http://www.biouml.org>

Biosoft.Ru/DevelopmentOnTheEdge.com

Laboratory of Bioinformatics,
Design Technological Institute of Digital Techniques

Novosibirsk, Russia.

Main BioUML concepts and ideas

- Visual modeling
- Plug-in based architecture (Eclipse platform runtime from IBM).
- Meta model – problem domain neutral level of abstraction that describes system as compartmentalized graph.
- Diagram type concept – formally defines graphical notation and provides its incorporation into BioUML workbench.
- Module concept - allows developer to incorporate databases on biological pathways into BioUML framework taking into account database peculiarities.

Visual Modeling

The problem of modeling and simulating of complex systems can be significantly simplified for customers by using computer systems providing visual modeling.

These visual depictions offer alternative syntax to completely and formally specify models.

A number of visual syntaxes were developed and implemented for formal description of biological systems, but there is no common standard.

Some graphical notations for biological pathways

Kohn K.W. (1999). Molecular Interaction Map of the Mammalian Cell Cycle Control and DNA Repair Systems. *Mol. Biol. Cell.* 10, 2703-2734.

Kitano H. (2003). A graphical notation for biochemical networks. *BIOSILICO Vol. 1. No. 5.*

R. Maimon and S. Browning (2001). Diagrammatic Notation and Computational Grammar for Gene Networks. *Proceedings of the International Conference on Systems Biology.* 2001.


Cook D.L. et al. (2001). A basis for a visual language for describing, archiving and analyzing functional models of complex biological systems. *Genome Biol.* 2. RESEARCH 0012.


Database specific notations:

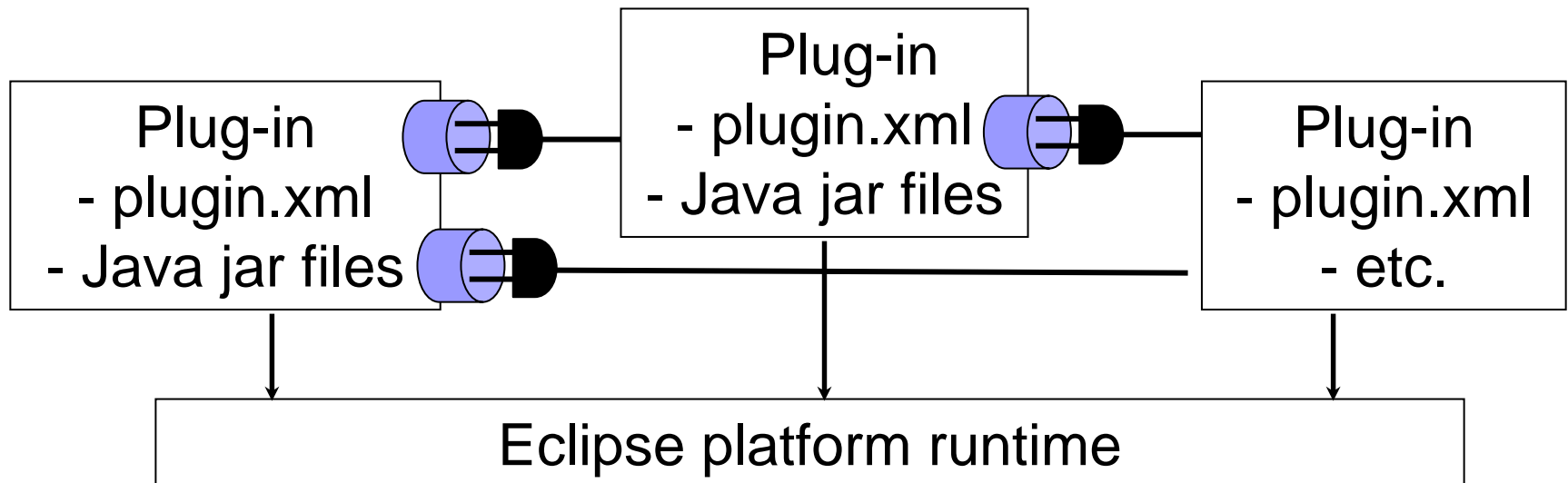
- KEGG/Metabolic pathways; GeneNet system; TRANSPATH
- ...

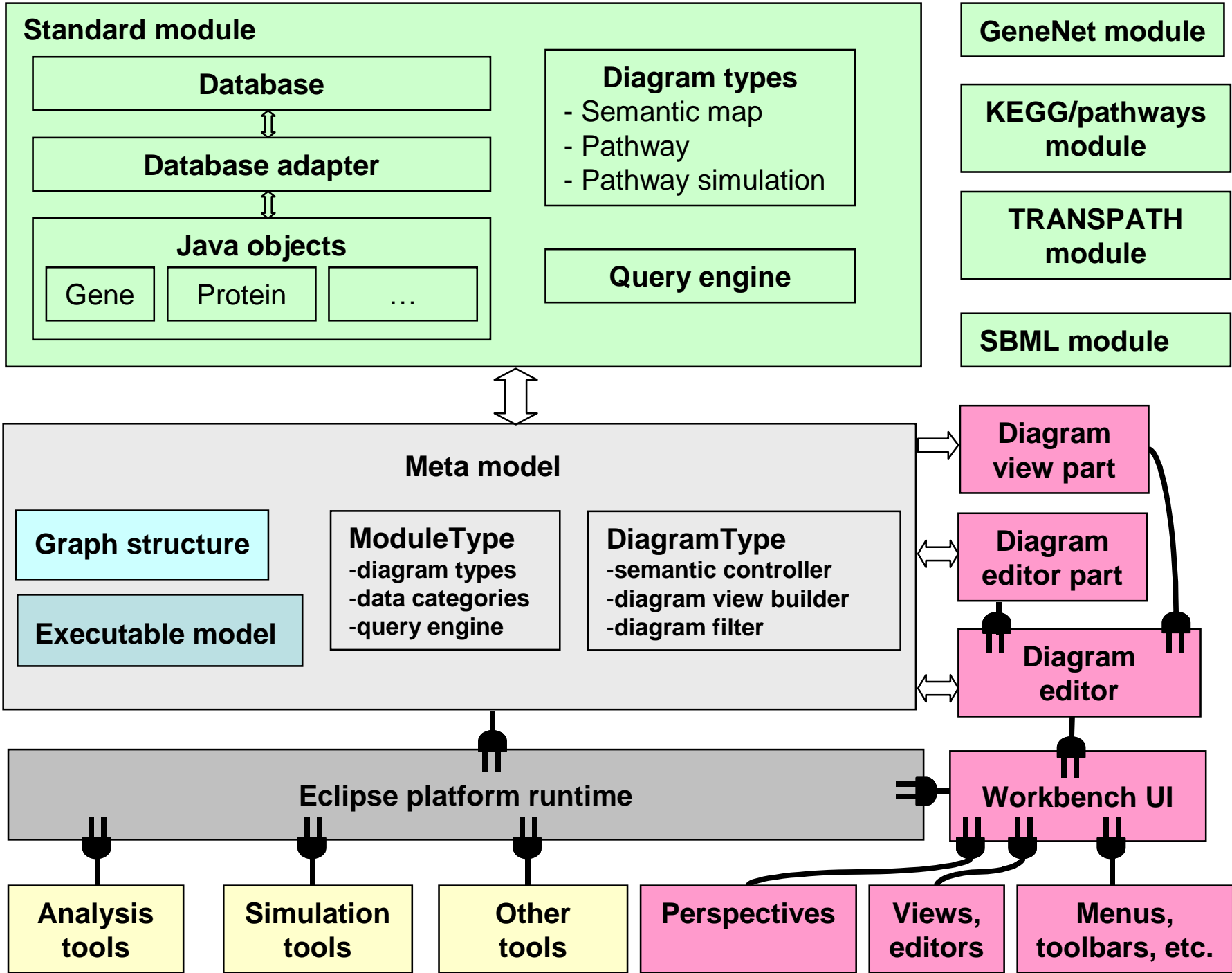
Plug-in based architecture

A **plug-in** is the smallest unit of BioUML workbench function that can be developed and delivered separately into BioUML workbench. A plug-in is described in an XML manifest file, called plugin.xml. The parsed contents of plug-in manifest files are made available programmatically through a **plug-in registry** API provided by Eclipse runtime.

 - **extension points** are well-defined function points in the system where other plug-ins can contribute functionality.

 - **extension** is a specific contribution to an extension point. Plug-ins can define their own extension points, so that other plug-ins can integrate tightly with them.





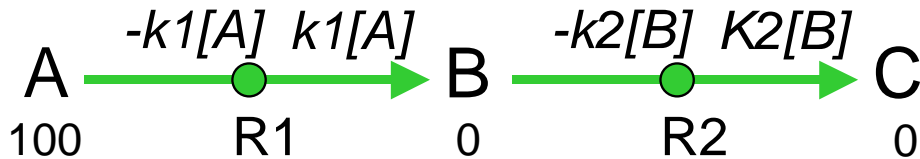
BioUML meta model

The core of BioUML workbench is meta model.

Unlike UML meta mode BioUML meta model is problem domain neutral and provides an abstract layer for comprehensive formal description of wide range of biological and other complex systems.

Content of databases on biological pathways or SBML models are expressed in terms of meta model and then can be used by other workbench plug-ins.

Example of formalized description of system from two chemical reactions



ID	A
CC	..
...	
//	

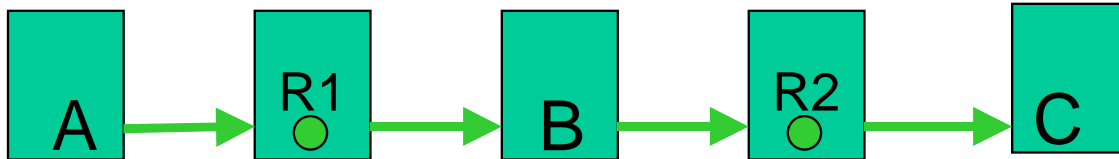
ID	R1
	A->B
...	
//	

ID	B
CC	..
...	
//	

ID	R2
	B->C
...	
//	

ID	C
CC	..
...	
//	

Description of system components in the database



System structure is described as a graph

100	$-k_1[A]$	$k_1[A]$	0	$-k_2[B]$	$K_2[B]$	0
-----	-----------	----------	---	-----------	----------	---

Mathematical model of the system

Mapping of SBML models into BioUML meta model

http://www.biouml.org/plugin_sbml.shtml?overview.

SBML

Model
Compartment
Species
Reaction
SpeciesReference
FunctionDefinition
Rule
Event

RDF
notes

DATA

Description of system
components

graph

System structure is
described as a graph

Math

Mathematical model
of the system

math
variable
volume
initialAmount
trigger, delay
eventAssignments

Diagram type concept

Diagram type defines:

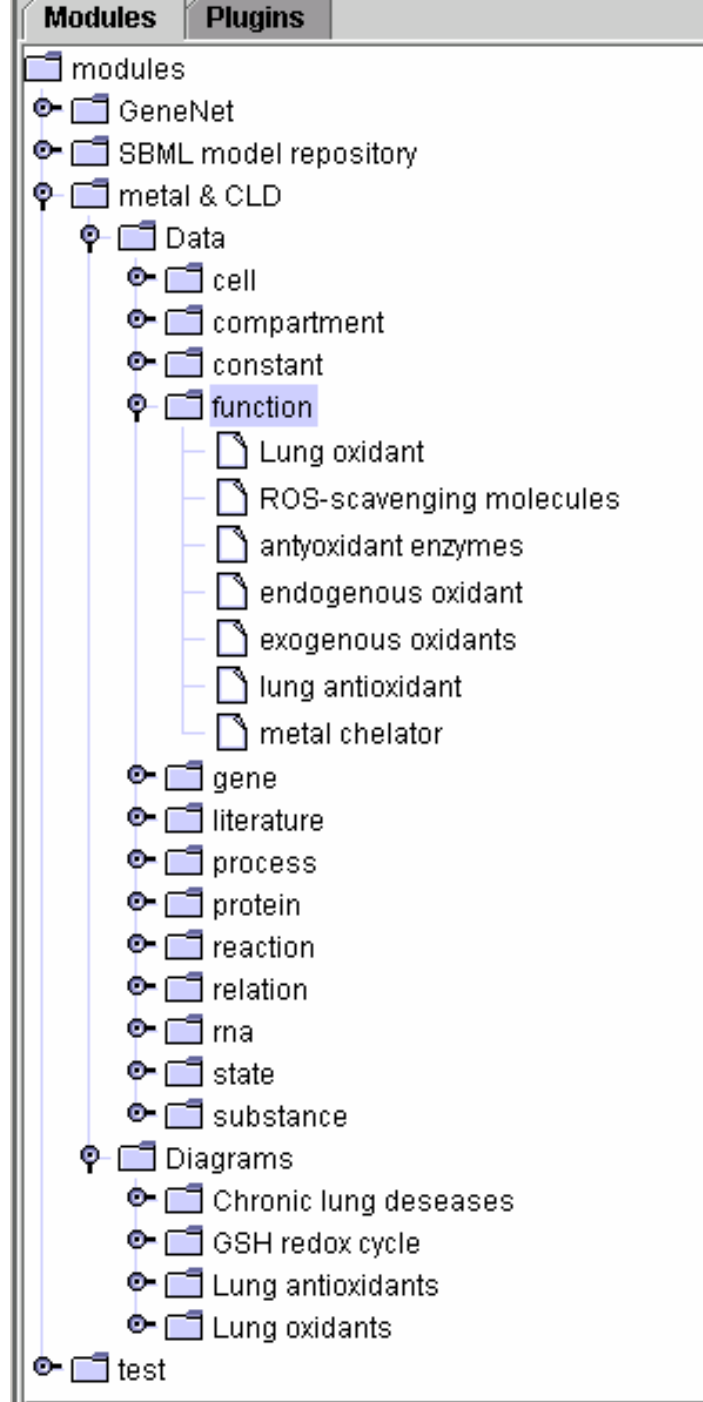
- what system components can be shown in the diagram;
- diagram view builder - it is used to generate view for each diagram element taking into account problem domain peculiarities;
- semantic controller - provides semantic integrity of the diagram during its editing;
- filters – hide or highlight diagram elements according to some selection criteria.

Module concept

The module concept allows to developer define new diagram types and incorporate other databases on biological pathways into BioUML framework.

The module defines mapping of database content into diagram elements and diagram types that can be used with the database.

Module also provides query engine that can be used by BioUML workbench to find interacting components of the system.



Graph search engine

SEARCH_FORM_TITLE

Search options

Property	Value
SearchOptions	
Diagram type	GeneNet diagram
Diagram name	At_seed_1
Type	Protein
Identifier	Hs:HSF1
Direction	Both
Depth	1
Layout	Layered

Start Save Cancel

Node

Title: Hs:HSF1-p
Data: DataElement[Hs:HSF1-p] class=class biouml.modules.genenet.type.Protein
ID: Hs:HSF1-p
OS: Homo sapiens (human).
NM: heat shock factor 1
SN: HSF1
GN: Hs:HSF1
FN: active
MM: multimer
MD: phosphorylated
SO: Hs:K562

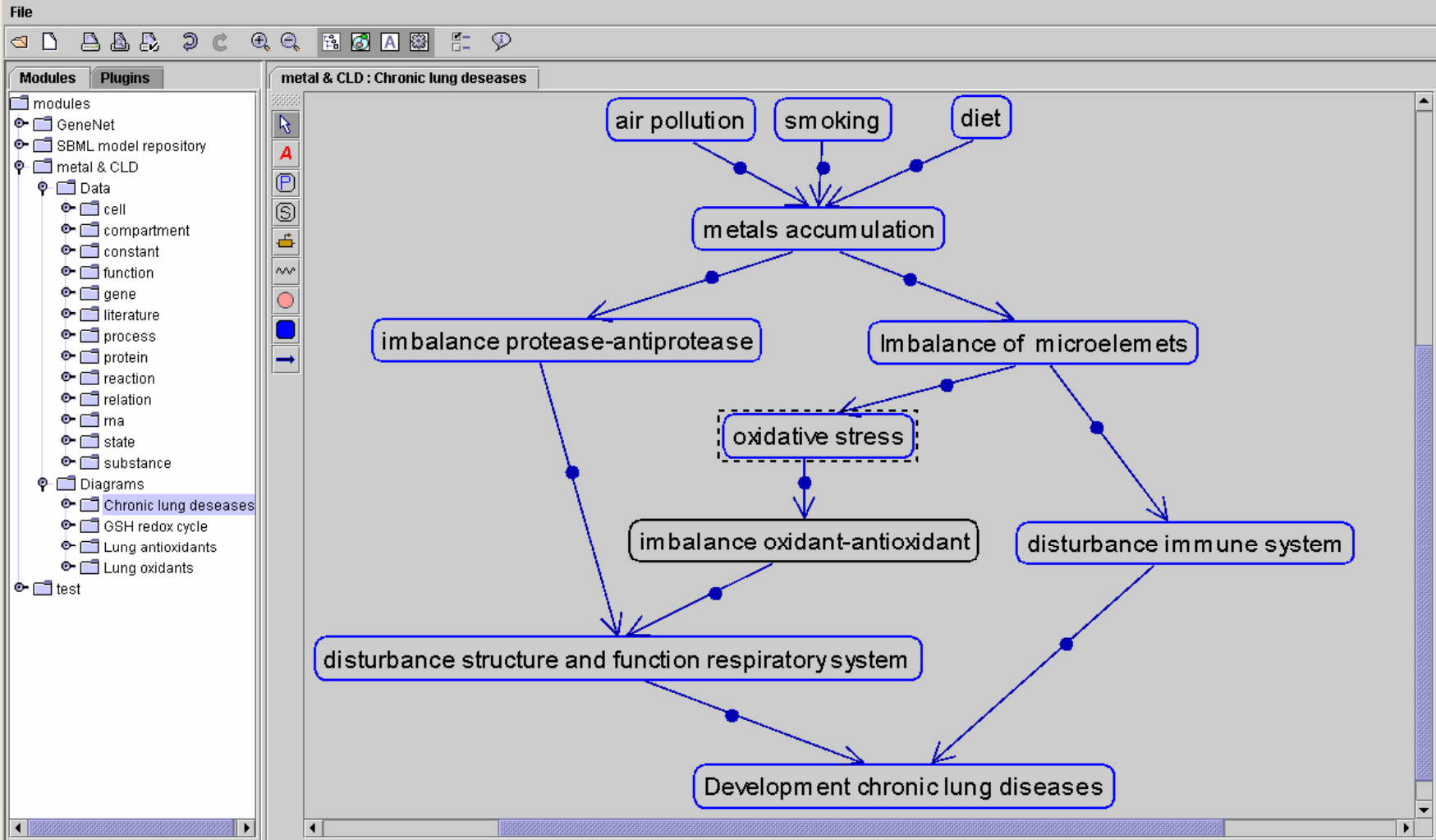
Modules

- standard BioUML module for biological pathways;
- module for models in SBML format;
- module for models in CellML format (only biochemical models);
- GeneNet database module;
- module for KEGG/Pathways database (draft);
- module for TRANSPATH database (draft);
- GeneOntology (under development).

Standard BioUML module for biological pathways

The module defines most common biological data types (gene, protein, RNA, substance, reaction, etc.), they mapping into simple text database and three diagram types for description of biological pathways on several semantic levels:

1. Semantic network (ontology) - this diagram type is used to describe semantic relationships between system components, system states, and related problem domain concepts.
2. Pathway diagram type is used for formalized description of biological pathway structure. This diagram type uses GeneNet graphical notation.
3. Pathway simulation diagram type is extension of pathway structure diagram, where variables are associated with graph nodes and differential equations with graph edges. This allows to BioUML workbench automatically generate mathematical model of the system and simulate its dynamics.



Title: oxidative stress
Data:
 ID: oxidative stress
 RF: [Ljava.lang.String;@14d4313
 [0]: Bagchi et al., 2001
 [1]: Buzard and Kasprzak2000
 [2]: Cavallo et al., 2002

Normal
HTML
Preview

Conceptual model of heavy metals influence on pathogenesis of chronic lung diseases

Description Legend JavaScript

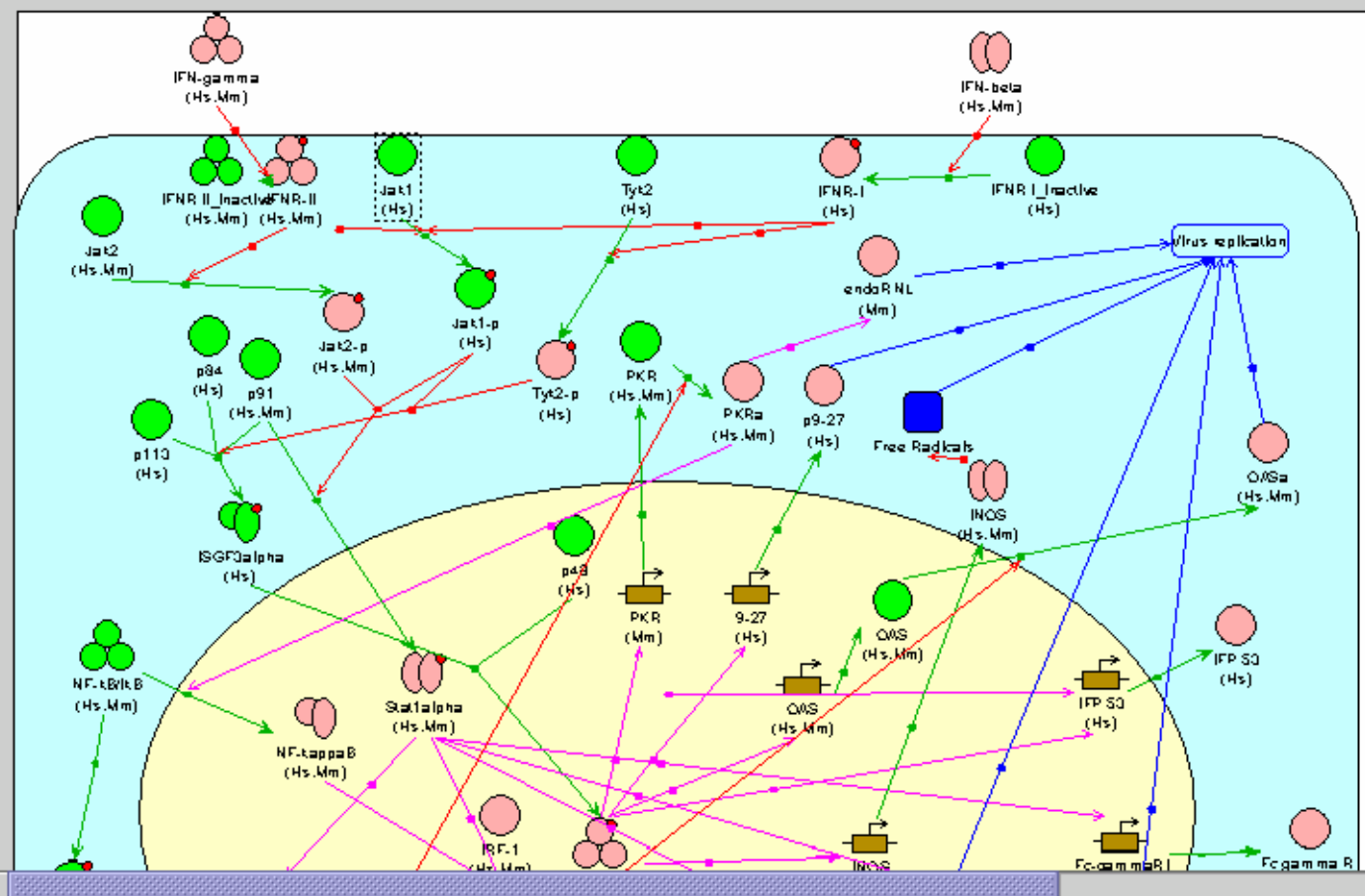


Plugins **Services**

Modules

- Biopath
- Biopath (local)
- CellML models
- Cyclonet
- Cyclonet (local)
- GeneNet
- Data
- Diagrams
 - Adipocyte
 - Antiviral response
 - Apoptosis (ge)
 - Cell Cycle (G)
 - Cholesterol
 - Cholesterol r
 - Cholesterol_
 - Erythroid diffe
 - Flower Devel
 - Germination
 - HSP70 autort
 - Heat Shock F
 - LEA program
 - Leptin (organ
 - Lipid metabo
 - Lipid metabo

GeneNet : Antiviral response



Normal **Html** Preview

Description Legend Filter ApplicationLog Clipboard JavaScript SearchLinked

IFN signal transduction pathways. form. </p><p>LFR activity

File

Modules Plugins

modules

- SBML model repository
 - Diagrams
 - CellCycle-1991Gol.xml
 - CellCycle-1991Tys-2.xml
 - CellCycle-1991Tys.xml
 - CellCycle-1997Nov.xml
 - CellCycle-1998Gar.xml
 - CircClock-1999LeI_periodic.xml
 - CircClock-2002Vil.xml
 - Genetic-2000Elo.xml
 - MAPKcasc-2000Kho.xml
 - MAPKcasc-2000Lev-2.xml
 - MAPKcasc-2000Lev.xml
 - Metabolism-2000Teu.xml
 - Metabolism-2001Hel.xml
 - Metabolism-2002Hoe.xml
 - Metabolism-2002Lam.xml
 - Miscellaneous-1963Lor.xml
 - Miscellaneous-1974Fie.xml
 - Receptor-1992DeY.xml

SBML model repository: CellCycle-1991Gol.xml

Figure No. 1

File Edit View Insert Tools Window Help

Solving Goldbeter problem

Property	Value
Node	
Title	Reaction5
Comment	
Data	
Identifier	Reaction5
Name	Reaction5
Reversible	<input type="checkbox"/>
Fast	<input type="checkbox"/>
Kinetic law	
Formula	(Reaction5_VM3...
Time units	
Substance u	

Normal
Html
Preview

Minimal Mitotic Oscillator

Citation
Goldbeter, A. A minimal cascade model for the mitotic oscillator involving cyclin and cdc2 kinase. Proc. Natl. Acad. Sci. USA 88: 9107-1101 (1991). <http://www.pnas.org/cgi/content/abstract/88/20/9107>

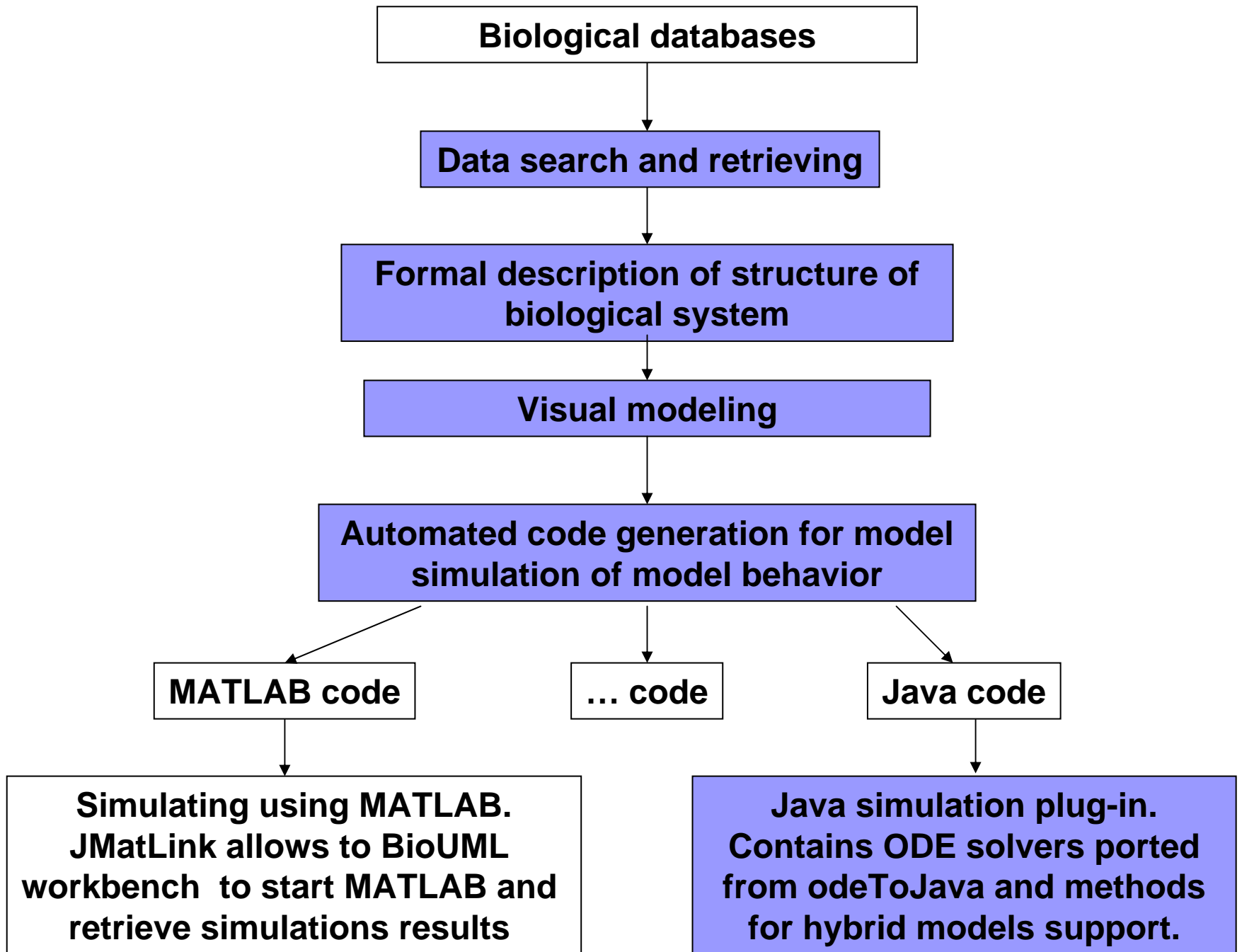
Description
This is a widely cited minimal (3-variable) model for a mitotic oscillator. The three variables represent Cyclin (C), inactive cdc-2 Kinase (M) and an active cdc-2 Kinase (X).
<http://www.pnas.org/cgi/content/abstract/88/20/9107>

Rate constant	Reaction
K1 = 0.1 (hill khalf)	Comp[M] -> M

View Edit Editors

Description Legend Constants Variables MATLAB JavaScript

- Suggested approach can be applied for modeling biological systems using:
 - Systems of ordinary differential equations
 - Systems of algebra-differential equations
 - State and transition diagrams
 - Hybrid models
 - Boolean and logical networks
 - Petri nets
 - Markov chains
 - Stochastic models
 - ...
- Limitations (to be overcome)
 - Spatial models
 - PDE
 - ...

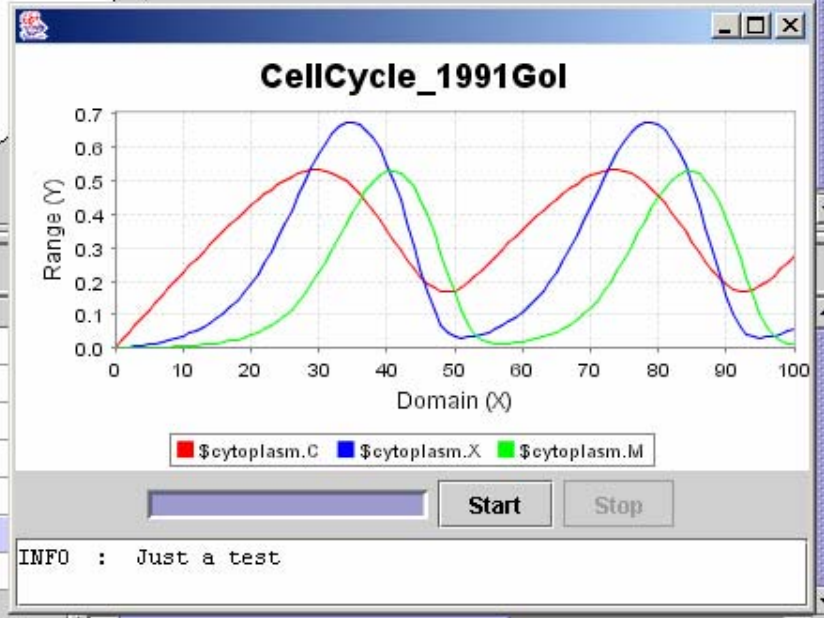
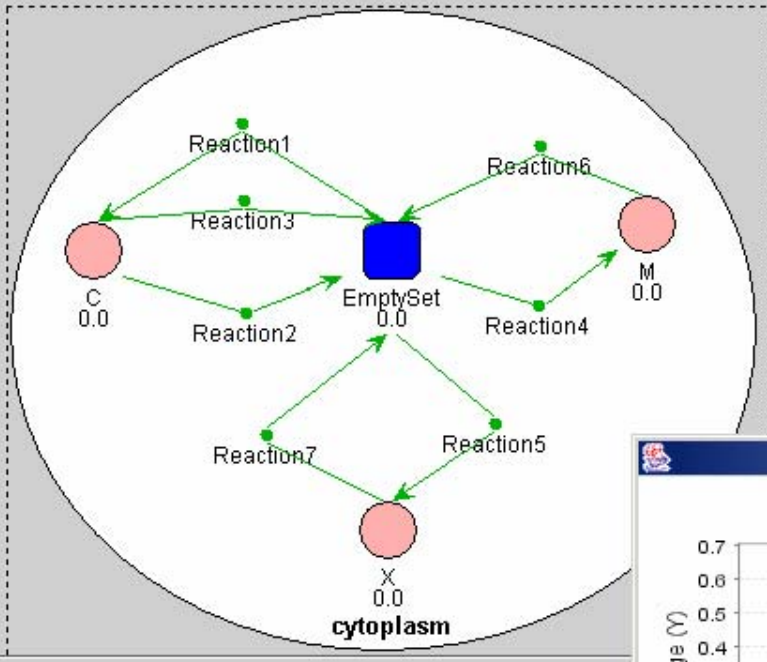


Plugins **Services**

Modules

- KEGG pathways
- SBML model repository
 - Diagrams
 - CellCycle-1991G
 - CellCycle-1991T
 - CellCycle-1991T
 - CellCycle-1997M
 - CellCycle-1998C
 - CircClock-1999L
 - CircClock-2002V
 - Genetic-2000Eld
 - MAPKcasc-2000
 - MAPKcasc-2000
 - MAPKcasc-2000
 - Metabolism-200
 - Metabolism-200
 - Metabolism-200
 - Metabolism-200

GeneNet : Antiviral response SBML model repository : CellCycle-1991Gol.xml



Diagram

Title: Goldbeter

Role:

Initial time: 0.0

Completion time: 100.0

Data:

ID: CellCycle-1991Gol.xml

TI: CellCycle-1991Gol.xml

Property	Value
Java simulation engine	
ODE solver name	
Output dir	.out
Initial time	0
Completion time	100
Time increment	1
Needed Plot Simulation	<input checked="" type="checkbox"/>

View Edit Editors Description Legend ApplicationLog Clipboard Parameters Variables JAVA MATLAB JavaScript

Modules Plugins Services

- modules
 - Biopath
 - Biopath (local)
 - CellML models
 - Cyclonet
 - Cyclonet (local)
 - GeneNet
 - KEGG pathways
 - SBML model repository
 - Diagrams
 - Simulation
 - plot
 - gold
 - result
 - res 1
 - res 2
 - Transpath
 - test

Plots

Plot parameters Plot

Plot properties

Property	Value
Module name	SBML model repository
Open plot	new_plot
Save as...	new_plot
Plot	
Name	new_plot
Title	
Description	
X axis title	
X: from	0
X: to	1
Y axis title	
Y: from	0
Y: to	1

Add series

Simulation Result: **res 1**

Variable X: **\$cytoplasm.C**

Variable Y: **\$cytoplasm.X**

Add **Remove**

Series properties

	Simulation result	X Variable	Y Variable	Legend	Specification
0	res 1	\$cytoplasm.C	\$cytoplasm.X	\$cytoplasm.X	

Ok **Cancel**

Simulation result

ID: res 1
 SN: ode23
 FROM: 0.0
 TO: 100.0

View Edit Editors

Description Legend ApplicationLog Clipboard Parameters Variables JAVA MATLAB JavaScript

11Cycle-1
 11Cycle-1
 11Cycle-1
 11Cycle-1
 11Cycle-1
 11Cycle-1
 11Cycle-1
 11Cycle-1



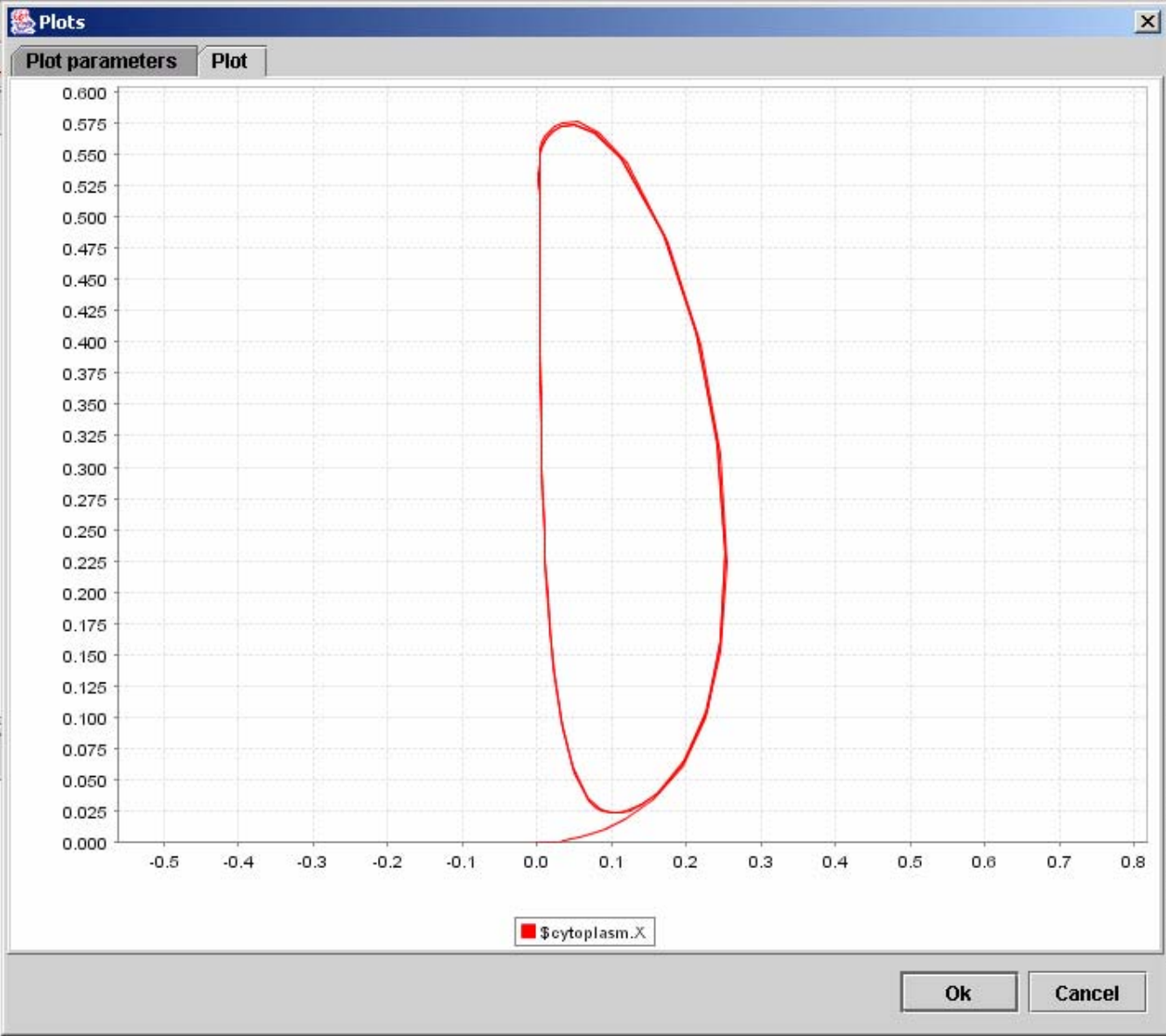
Modules Plugins Services

- modules
 - Biopath
 - Biopath (local)
 - CellML models
 - Cyclonet
 - Cyclonet (local)
 - GeneNet
 - KEGG pathways
 - SBML model repository
 - Diagrams
 - Simulation
 - plot
 - gold
 - result
 - res 1
 - res 2
 - Transpath
 - test



Simulation result

ID: res 1
SN: ode23
FROM: 0.0
TO: 100.0



View Edit Editors

Description Legend ApplicationLog Clipboard Parameters Variables JAVA MATLAB JavaScript

Cyclonet database

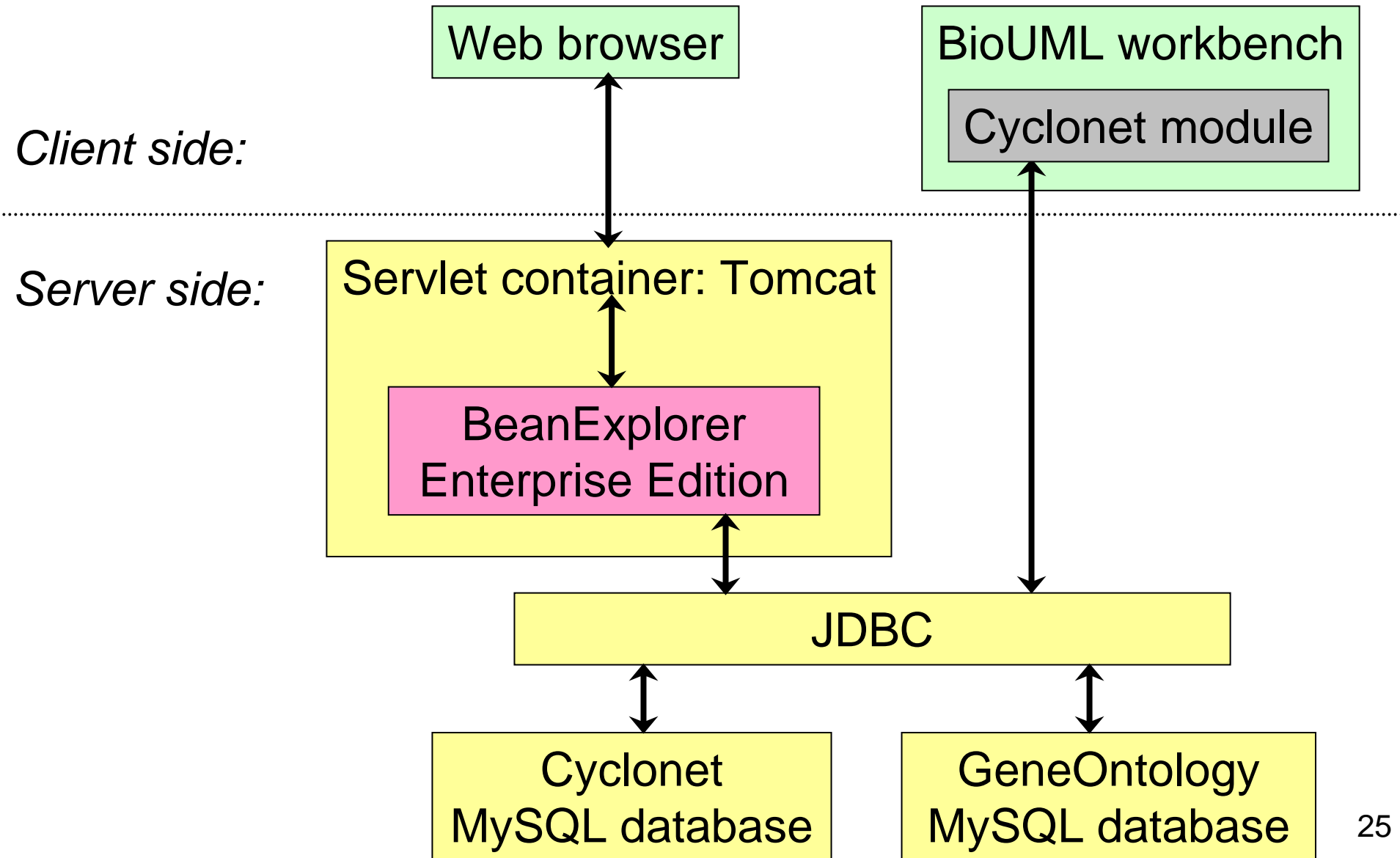
Using BioUML workbench we are developing Cyclonet database – a database on cell cycle regulation in eukaryotes.

Using BeanExplorer Enterprise Edition database content is available through the Internet.

<http://cyclonet.biouml.org>

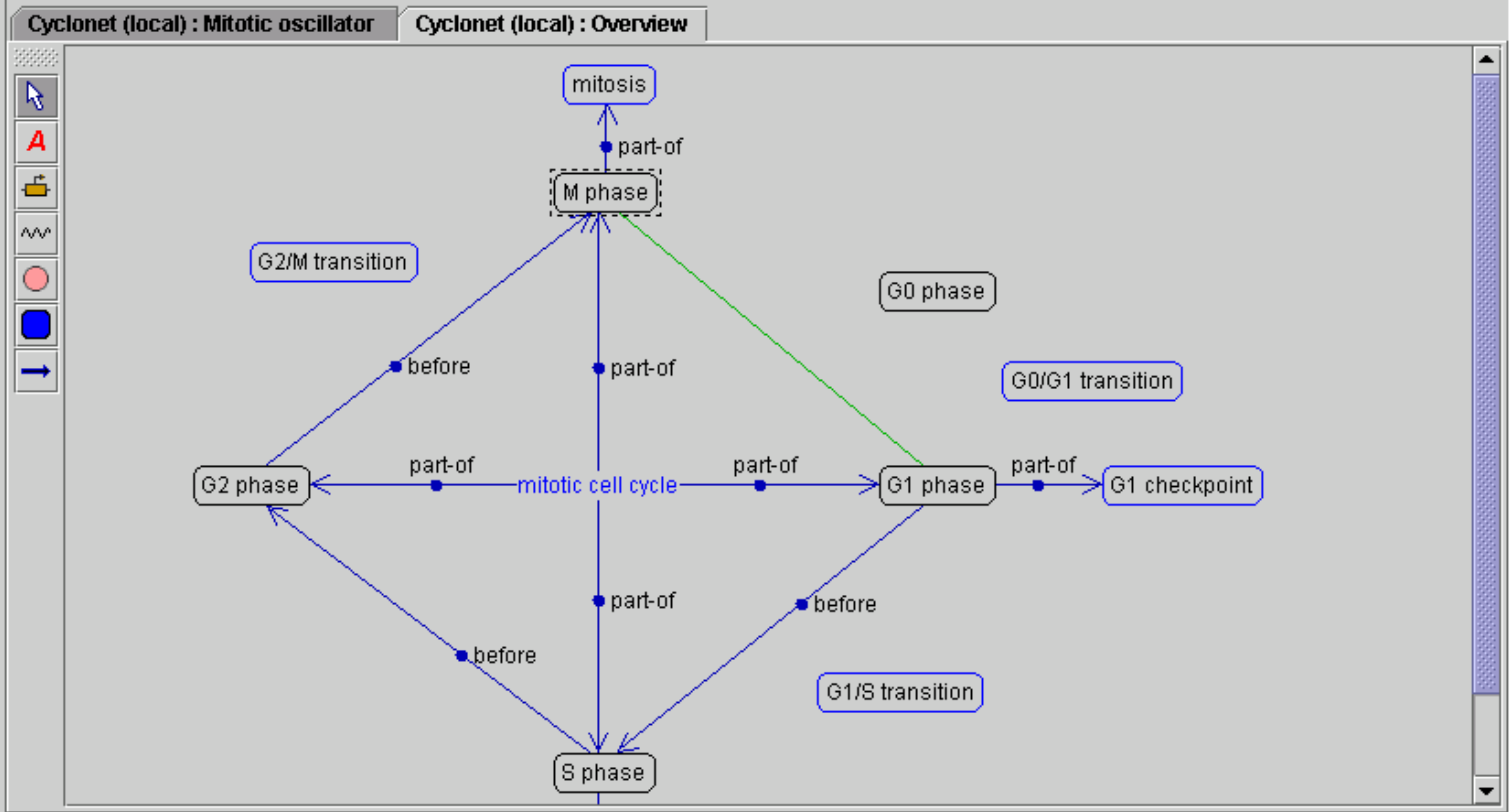
The database contains information about cell cycle specific genes, proteins, protein complexes and their interactions, diagrams of cell cycle regulation for vertebrates, models of cell cycle and results of their analyses, microarray data, literature references and other related resources.

Cyclonet system architecture



Modules **Plugins**

- modules
 - Cyclonet
 - Cyclonet (local)
 - Data
 - cell
 - compartment
 - concept
 - gene
 - protein
 - reaction
 - relation
 - ma
 - substance
 - Diagrams
 - Mitotic oscillator
 - Overview
 - Regulation
 - Dictionaries
 - Transpath



Property	Value
Node	
Title	M phase
Comment	
Data	
Identifier	CPT0000005
Type	semantic-concept-state

View Edit Editors

Ontology diagram type

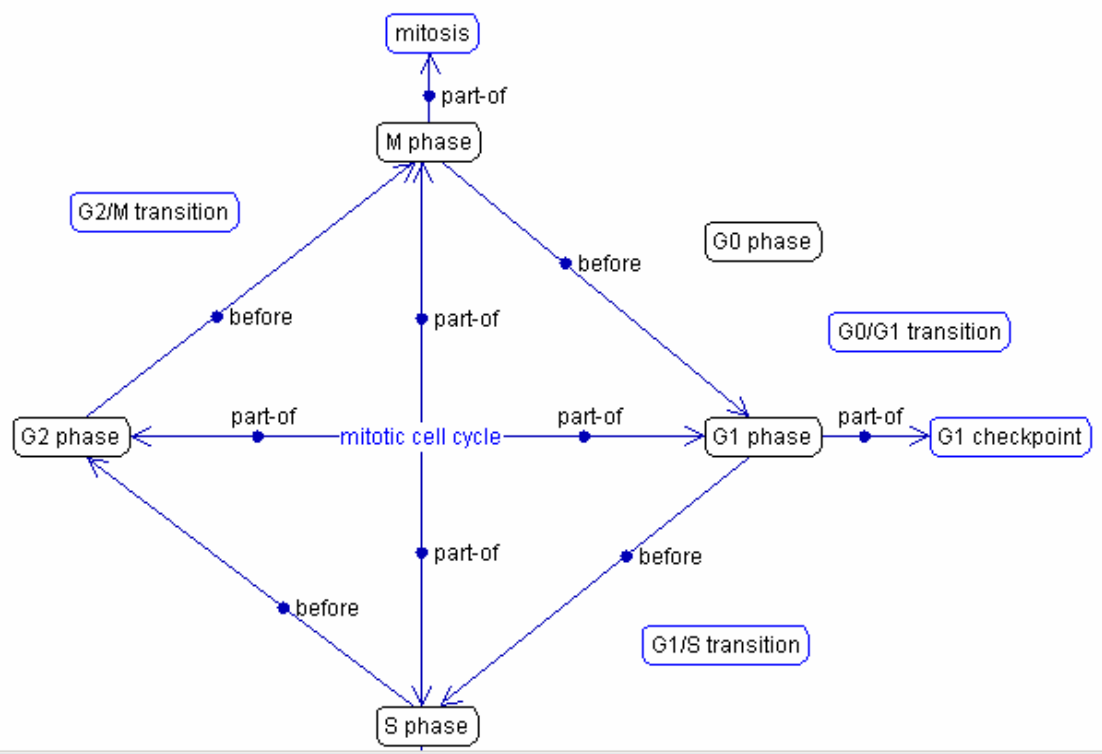
... the description should be here ...

Description Legend JavaScript

- Cyclonet project
 - [Home](#)
 - [About](#)
- Diagrams
- Data
 - [Cells](#)
 - [Compartments](#)
 - [Concepts](#)
 - [Genes](#)
 - [Proteins](#)
 - [Reactions](#)
 - [Relations](#)
 - [RNAs](#)
 - [Substances](#)
- Publications
 - [Compact view](#)
 - [Detailed view](#)
- Dictionaries
 - [Databases](#)
 - [Relation types](#)
 - [Species](#)
 - [Units](#)
- Registered users
 - [Login](#)
 - [Forgot Password?](#)



Overview



Concepts

Filter

#	ID	Name	Description	Synonyms	DB references	References
1	CPT000005	M phase, M phase of mitotic cell cycle	The part of a standard cell cycle during which mitosis and cytokinesis take place.	0	1	0

Biopath (codename) database

<http://biopath.biouml.org>

Purpose of this project is create formal description of pathogeneses of chronic respiratory diseases and optimization of their treatment.

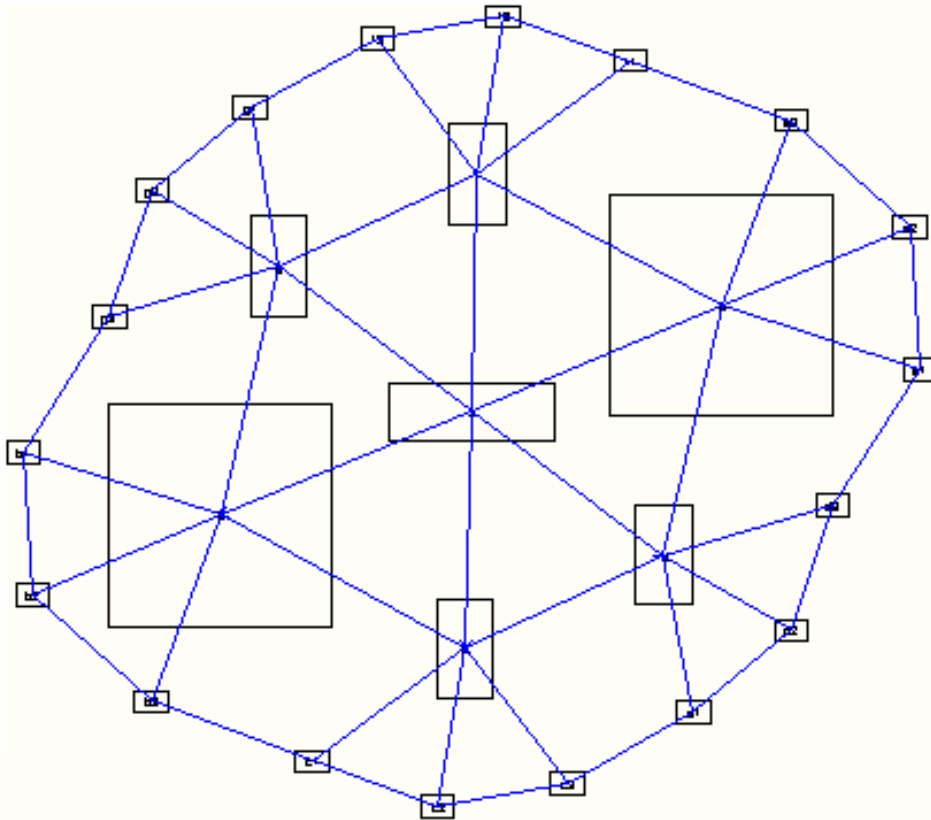
The work is doing jointly with Institute of Medical Problems of North (Krasnoyarsk) and Institute of Pulmonology (Moscow).

Current works

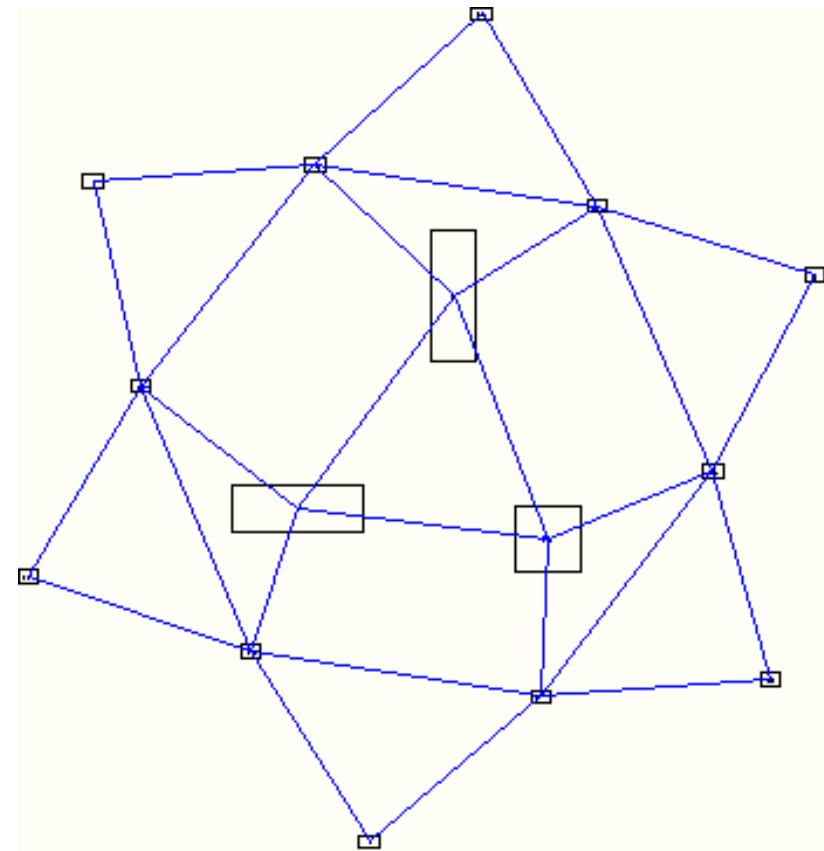
- Graph search engine
 - Graph search engine completion and testing
 - adaptation of new graph layout library
- Complete support of SBML level 2
 - Time delay
 - Algebraic rules
 - Units
 - SBML semantic tests
- Formula editor
- State and transition diagrams.

Improved version of force directed graph layout

(takes into account nodes size)



(lay-out time ~50ms, 300 iterations)



$$\text{justAtest}(x_1) = \begin{cases} \sin(-x_1) & \text{if } x < 0 \\ \frac{0}{1 + \frac{y_1 + y_2}{2}} & \text{if } x == 0 \\ x_1 + e^{x+5} & \text{otherwise} \end{cases}$$

Text **MathML**

```
function justAtest(x_1) =
  if(x<0)    sin(-x_1);
  if(x==0)   0/(1+(y_1+y_2)/2);
  otherwise  x_1+e^(x+5)
```

Validate

Dump

Messages:

```
AST tree dump:
  Start
    Function declaration: justAtest
      Variable: "x_1"
      Piecewise: 3
        Piece
          Function "<"
            Variable: "x"
            Constant: 0
          Function "sin"
```

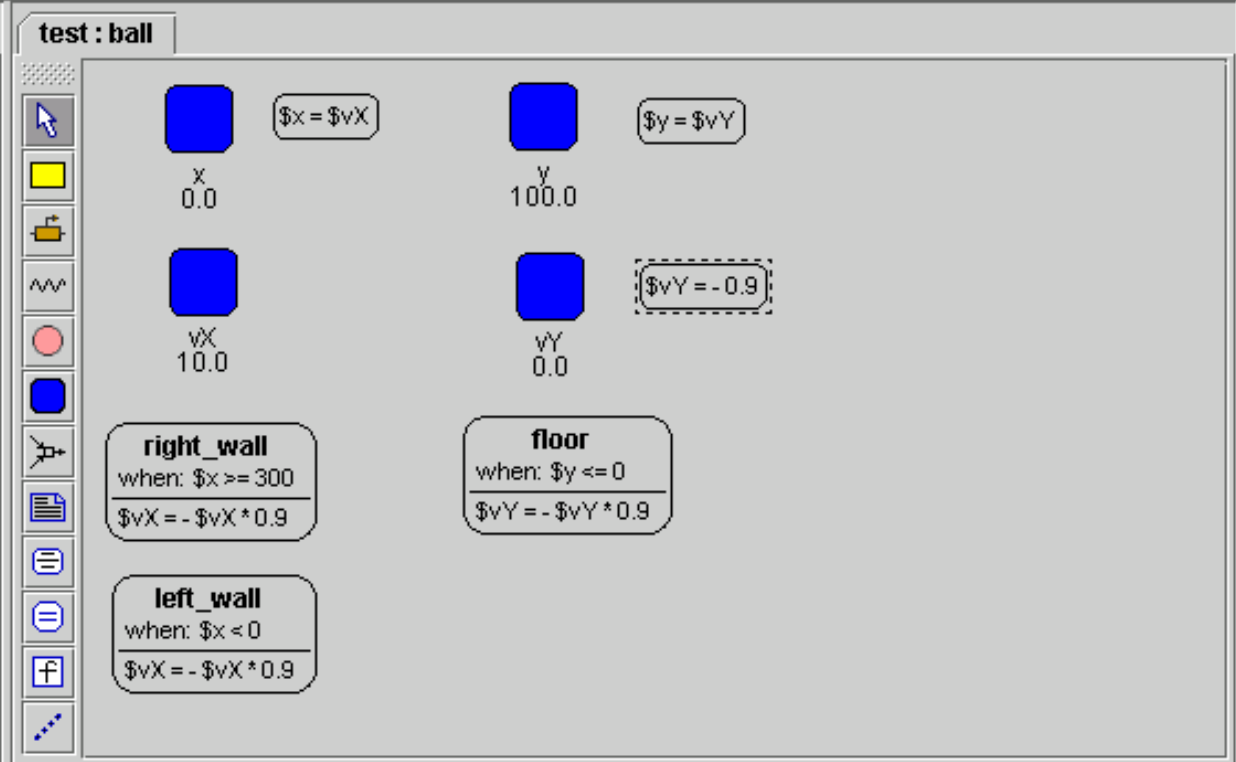
Ok

Cancel



Modules **Plugins** **Services**

- Some models
- Cyclonet
- Cyclonet (local)
- GeneNet
- KEGG pathways
- SBML model repository
- Transpath
- test
 - Data
 - Diagrams
 - ball
 - equation_vY
 - equation_x
 - equation_y
 - floor
 - left_wall
 - right_wall
 - vX



Property	Value
Node	
Title	equation_vY
Comment	
Size	57.0 20.0
Role	
Equation	-0.9
Comment	

Property	Value
Matlab simulation engine	
ODE solver name	ode23
Output dir	C:/matlabR12/work
Initial time	0
Completion time	100
Time increment	0

test : ball

$x = vx$ $y = vy$

$x = 0.0$ $y = 100.0$

$vx = 10.0$ $vy = 0.0$ $vy = -0.9$

right_wall
when: $x \geq 300$
 $vx = -vx * 0.9$

floor
when: $y \leq 0$
 $vy = -vy * 0.9$

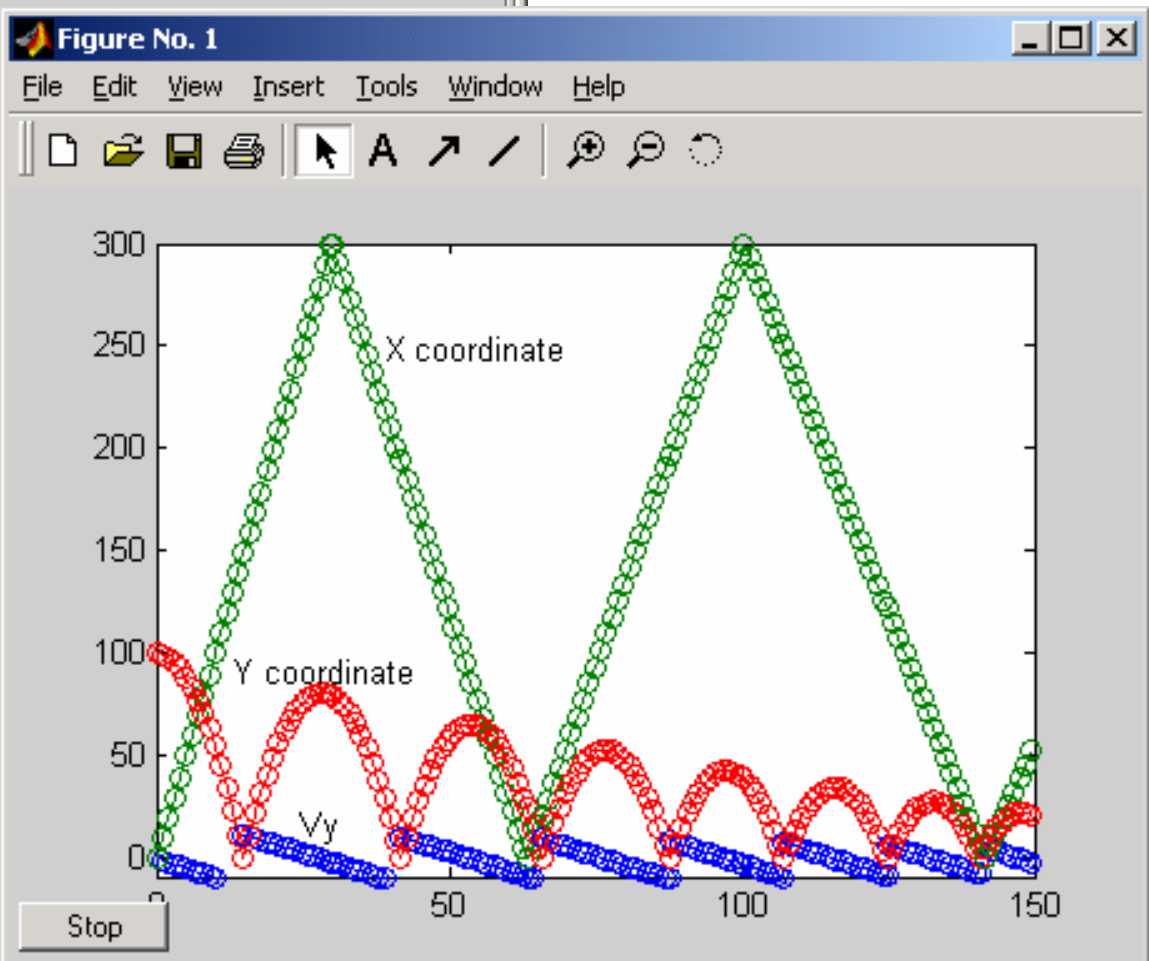
left_wall
when: $x < 0$
 $vx = -vx * 0.9$

Property

Matlab simulation engine	
ODE solver name	ode23
Output dir	C:/matlabR12A
Initial time	0
Completion time	100
Time increment	0

Description Legend Filter

Parameters Variables JAVA MATLAB Ja



Further works

- Library of predefined kinetic laws
- Stochastic modeling
- New diagram type for gene regulatory regions
- ...
- BioUML can be a test platform where we can try to implement proposals for SBML level 3, so people will get a possibility to evaluate them in practice.
- We will be glad to help third party developers to integrate their software with BioUML workbench.
- We are open for any collaboration, joint projects, etc.

Availability

BioUML workbench (including source code) is freely available at <http://www.biouml.org>

There is special forum dedicated to BioUML workbench where you can post your questions and suggestions:

<http://groups.yahoo.com/group/biouml/>

Cyclonet database – <http://cyclonet.biouml.org>

Biopath database – <http://biopath.biouml.org>

Acknowledgments

Part of this work was partially supported by following grants:
Volkswagen-Stiftung (I/75941), INTAS Nr. 03-51-5218 and
RFBR Nr. 04-04-49826-a .

Author is grateful to for useful comments, discussions and
technical support

Alexander Kel
Sergey Zhatchenko

Software developers

Mikhail Puzanov

Igor Tyazhev

Vasiliy Hudyakov

Cheremushkina Alexandr Koshukov

Artem Shaidukov

Annotators

Ruslan Sharipov

Elena

Oleg Onegov