系统生物学

(Systems Biology)

马彬广

专用建模工具

（第十五讲）

梗概（Synopsis）

- 通用建模工具（数学计算软件）
- 专用建模工具（细胞生化体系建模）

专用建模工具

- 模型描述语言与数据格式
- 模型及建模相关的数据库
- 细胞生化体系建模工具

模型描述语言

模型描述语言大多是基于XML (Extensible Markup Language) 的，并用到了MathML。

- SBML (Systems Biology Markup Language)、SBGN (Systems Biology Graphical Notation)
- BioPAX （Biological Pathway Exchange）
- CellML、FieldML
- CMDL （Chemical Model Definition Language, of Dizzy）

XML

- Extensible Markup Language (XML) is a set of rules for encoding documents in machine-readable form. It is defined in the XML 1.0 Specification produced by the W3C, and several other related specifications, all gratis open standards.

- XML's design goals emphasize simplicity, generality, and usability over the Internet. It is a textual data format with strong support via Unicode for the languages of the world. Although the design of XML focuses on documents, it is widely used for the representation of arbitrary data structures, for example in web services.

- Many application programming interfaces (APIs) have been developed that software developers use to process XML data, and several schema systems exist to aid in the definition of XML-based languages.

- To now, hundreds of XML-based languages have been developed, including RSS, Atom, SOAP, and XHTML. XML-based formats have become the default for most office-productivity tools, including Microsoft Office (Office Open XML), OpenOffice.org (OpenDocument), and Apple's iWork.

（From WiKi, check the WiKi page for more information）

MathML

- Mathematical Markup Language (MathML) is an application of XML for describing mathematical notations and capturing both its structure and content. It aims at integrating mathematical formulae into World Wide Web pages and other documents. It is a recommendation of the W3C math working group. (from WiKi)

- Site: http://www.w3.org/Math/

- Check WiKi for more details.

SBML- Systems Biology Markup Language

- SBML is a machine-readable language, based on XML, for representing models of biological processes. SBML can represent metabolic networks, cell-signaling pathways, regulatory networks, and many other kinds of systems. (from WiKi)

- Simply put, SBML is a machine-readable format for representing models.

- Version: SBML Level 3 Version 1 Core, Release 1, 6 October 2010

- Website: sbml.org; (show the main page)

- Show the simple introduction; Show more details;

- Check the original paper; Check the specification;

SBGN - Systems Biology Graphical Notation

- The Systems Biology Graphical Notation (SBGN) project, an effort to standardize the graphical notation used in maps of biochemical and cellular processes studied in systems biology.

- Standardizing the visual representation is crucial for more efficient and accurate transmission of biological knowledge between different communities in research, education, publishing, and more. When biologists are as familiar with the notation as electronics engineers are familiar with the notation of circuit schematics, they can save the time and effort required to familiarize themselves with different notations, and instead spend more time thinking about the biology being depicted.

- SBGN is made up of three orthogonal languages, representing different visions of biological systems. Each language defines a comprehensive set of symbols with precise semantics, together with detailed syntactic rules how maps are to be interpreted.

- Ref: Nat Biotechnol. 2009 27(8):735-41. Site: www.sbgn.org

Three Languages

- Process Description language (formerly Process Diagram language)

The SBGN Process Description (PD) language shows the temporal courses of biochemical interactions in a network. It can be used to show all the molecular interactions taking place in a network of biochemical entities, with the same entity appearing multiple times in the same diagram. (Level 1 Version 1.2)

- Entity Relationship language

The SBGN Entity Relationship (ER) language allows you to see all the relationships in which a given entity participates, regardless of the temporal aspects. Relationships can be seen as rules describing the influences of entities nodes on other relationships. (Level 1 Version 1.1)

- Activity Flow language

The SBGN Activity Flow (AF) language depicts the flow of information between biochemical entities in a network. It omits information about the state transitions of entities and is particularly convenient for representing the effects of perturbations, whether genetic or environmental in nature. (Level 1)

(Specifications can be found at Nature Proceedings)

BioPAX - Biological Pathway Exchange

- BioPAX is a standard language that aims to enable integration, exchange, visualization and analysis of biological pathway data.

- Specifically, BioPAX supports data exchange between pathway data groups and thus reduces the complexity of interchange between data formats by providing an accepted standard format for pathway data.

- By offering a standard, with well-defined semantics for pathway representation, BioPAX allows pathway databases and software to interact more efficiently. In addition, BioPAX enables the development of pathway visualization from databases and facilitates analysis of experimentally generated data through combination with prior knowledge.

- The BioPAX effort is coordinated closely with that of other pathway related standards initiatives namely; PSI-MI, SBML, CellML, and SBGN in order to deliver a compatible standard in the areas where they overlap.

- Website: www.biopax.org (link & demo)

CellML & FieldML

- The CellML language is an open standard based on the XML markup language. CellML is being developed by the Auckland Bioengineering Institute at the University of Auckland and affiliated research groups. The purpose of CellML is to store and exchange computer-based mathematical models. CellML allows scientists to share models even if they are using different modelling tools. It also enables them to reuse components from one model in another, thus accelerating model development. (see http://www.cellml.org/ for more details)

- FieldML's (Field Modelling/Markup Language) goal is to be a declarative language for building hierarchical models represented by generalized mathematical fields. Its primary use will be to represent the dynamic geometry and solution fields from computational models of cells, tissues and organs. It is intended to provide a framework for modelling software development and model interchange for bioengineering and general engineering analysis communities. (see http://www.fieldml.org/ for more details)

模型及建模相关数据库

模型数据库收集文献中报道的生命活动模型，而建模相关的数据库，则收录了相应生物学过程的实验信息或已有知识。

- BioModels (mathematical models of biological interests)
- KEGG, Reactome, ConsensusPathDB, BioCyc
- Sabio-RK (System for the Analysis of Biochemical Pathways - Reaction Kinetics)
- PANTHER (Protein ANalysis THrough Evolutionary Relationships)

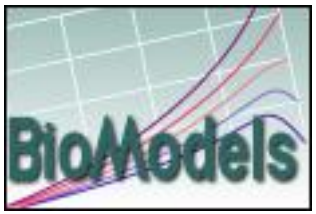
BioModels

- BioModels Database is a data resource that allows biologists to store, search and retrieve published mathematical models of biological interests. Models present in BioModels Database are annotated and linked to relevant data resources, such as publications, databases of compounds and controlled vocabularies.

- BioModels Database also allows users to generate sub-models, provides access to online simulation tools and features programmatic access via Web Services.

- URL: http://www.ebi.ac.uk/biomodels-main/

- Check the above website and give some explanation.



SABIO-RK

- The SABIO-RK (System for the Analysis of Biochemical Pathways - Reaction Kinetics) is a web-based application based on the SABIO relational database that contains information about biochemical reactions, their kinetic equations with their parameters, and the experimental conditions under which these parameters were measured.

- It aims to support modellers in the setting-up of models of biochemical networks, but it is also useful for experimentalists or researchers with interest in biochemical reactions and their kinetics.

- Information about reactions and their kinetics can be exported in SBML (Systems Biology Mark-Up Language) format.

- URL: http://sabio.villa-bosch.de/

- Check the above site

PANTHER

- The PANTHER (**P**rotein **AN**alysis **TH**rough **E**volutionary **R**elationships) Classification System is a unique resource that classifies genes by their functions, using published scientific experimental evidence and evolutionary relationships to predict function even in the absence of direct experimental evidence. Proteins are classified by expert biologists according to:

  - Gene families and subfamilies, including annotated phylogenetic trees
  - Gene Ontology classes: molecular function, biological process, cellular component
  - PANTHER Protein Classes
  - Pathways, including diagrams

- PANTHER is part of the Gene Ontology Reference Genome Project.

- URL: http://www.pantherdb.org/

- Check the above website

专用建模工具

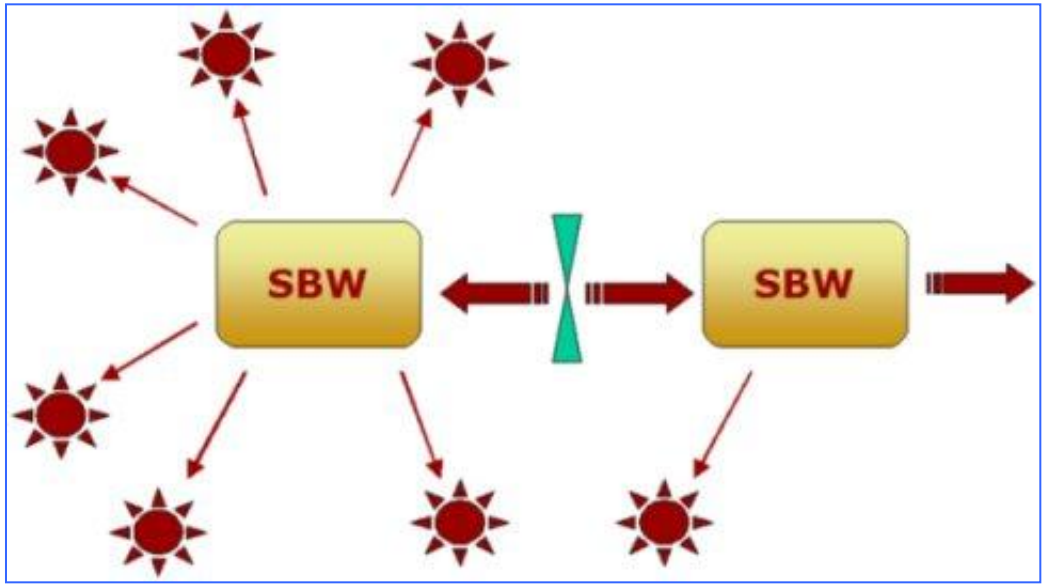
生化细胞体系生物过程建模的工具，有stand alone的程序，也有基于通用(数学)软件的扩展包，还有网络服务。

- SBW, CellDesigner 、COPASI
- Dizzy （Chemical Model Definition Language）
- MatLab-based: SimBiology, SBMLtoolbox, SBtoolbox, SBtoolbox2, CellNetAnalyzer
- Mathematica-based: MathSBML
- Python API: PySCeS; R-interface: R-SBML / SBMLR
- WebService: PyBioS, JWS online

SBW- Systems Biology Workbench

What is SBW: The Systems Biology Workbench (SBW), is a software framework that allows heterogeneous application components-written in diverse programming languages and running on different platforms-to communicate and use each others' capabilities via a fast binary encoded-message system. SBW enables applications (potentially running on separate, distributed computers) to communicate via a simple network protocol. The interfaces to the system are encapsulated in client-side libraries that are provided for different programming languages.

How does it work?



CellDesigner

- CellDesigner is a structured diagram editor for drawing gene-regulatory and biochemical networks.

- Networks are drawn based on the process diagram, with graphical notation system proposed by Kitano, and are stored using the Systems Biology Markup Language (SBML), a standard for representing models of biochemical and gene-regulatory networks.

- Networks are able to link with simulation and other analysis packages through Systems Biology Workbench (SBW).

- CellDesigner supports simulation and parameter scan by an integration with SBML ODE Solver and Copasi.

- By using CellDesigner, you can browse and modify existing SBML models with references to existing databases, simulate and view the dynamics through an intuitive graphical interface.

- CellDesigner (Current version 4.2, SBML level 2) is free to use.

COPASI

- COPASI is a software application for simulation and analysis of biochemical networks and their dynamics. COPASI is a stand-alone program that supports models in the SBML standard and can simulate their behavior using ODEs or Gillespie's stochastic simulation algorithm; arbitrary discrete events can be included in such simulations.

- COPASI carries out several analyses of the network and its dynamics and has extensive support for parameter estimation and optimization. COPASI provides means to visualize data in customizable plots, histograms and animations of network diagrams.

- Team: The COPASI project is an international collaboration between three groups at the Virginia Bioinformatics Institute, the University of Heidelberg, and the University of Manchester (the project has also been developed at the EML Research in the past). All aspects of coding, documentation, and testing are distributed between the three groups.

- Site: http://www.copasi.org; Current Version: Stable: COPASI 4.8 (Build 35)

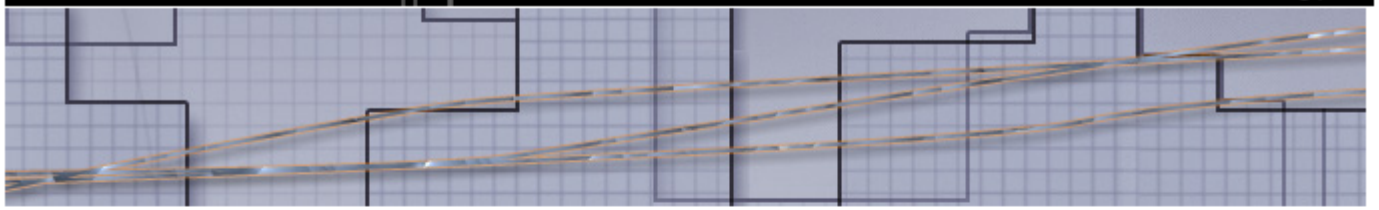
COPASI: feature list

- Model:
  - Chemical reaction network.
  - Arbitrary kinetic functions.
  - ODEs for compartments, species, and global quantities.
  - Assignments for compartments, species, and global quantities.
  - Initial assignments for compartments, species, and global quantities.
  - SBML import (L1V1+2, L2V1-4) and export (L1V2, L2V1-4).
- Analysis:
  - Stochastic and deterministic time course simulation
  - Steady state analysis (including stability).
  - Metabolic control analysis/sensitivity analysis.
  - Elementary mode analysis .
  - Mass conservation analysis.
  - Time scale separation analysis
  - Calculation of Lyapunov exponents.
  - Parameter scans.
  - Optimization of arbitrary objective functions.
  - Parameter estimation using data from time course and/or steady state experiments simultaneously.
- Graphical User Interface (CopasiUI)
  - Sliders for interactive parameter changes.
  - Color-coded tables
  - 3D bar charts
  - Plots and Histograms.
  - Network diagram visualization of results.
- Command Line (CopasiSE) for batch processing.
- Versions for MS Windows, Linux, Mac OS X, and Solaris SPARC.
- Loading of legacy Gepasi files.
- Export to Berkeley Madonna, XPPAUT, and C source code of the ODE system generated from the model.
- Saving of mathematical formulas and ODEs in MathML or LaTeX

Dizzy

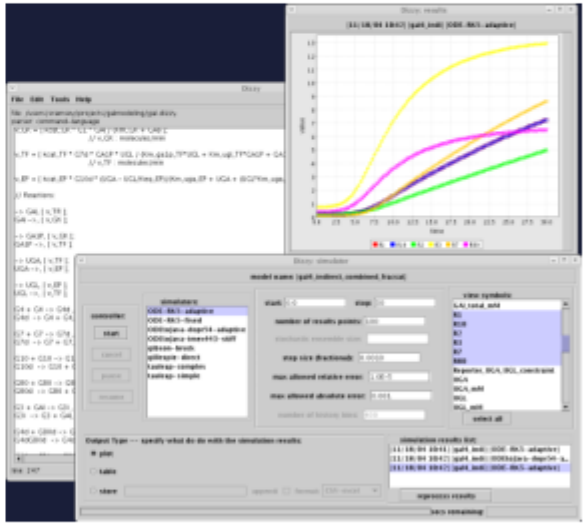
Dizzy Home Page

INSTITUTE FOR Systems Biology



Download ____ Overview ____ User Manual ____ Versions ____ License Agreement ____ API ____ Publication

Dizzy is a chemical kinetics stochastic simulation software package written in Java. It provides a model definition environment and an implementation of the Gillespie, Gibson-Bruck, and Tau-Leap stochastic algorithms. Dizzy is capable of importing and exporting the SBML model definition language, as well as displaying models graphically using the Cytoscape software system. Dizzy was written by Stephen Ramsey in the laboratory of Hamid Bolouri at ISB. Dizzy is based on the ISBJava library. The software is available for download in a self-contained archive format. It is distributed under the GNU Lesser General Public License (LGPL), which is a standard "free software" and "open source" license. Dizzy requires the Java 2 Runtime Environment (JRE) version 1.4 or newer, or an equivalent JRE. For more information about the Dizzy system, please consult the Dizzy User Manual.



Current version: 1.11.4, released on 2006/09/28

The development of Dizzy was supported by a grant from the National Institute of Allergy and Infectious Disease (NIAID), a division of the National Institutes of Health (NIH).

MatLab-based: SimBiology

SimBiology software provides an integrated environment for modeling biological processes, simulating the dynamic behavior of these processes, and analyzing the model with simulation and experimental data. Biological processes include metabolic, genetic, and signaling pathways with transform, binding, and transport reactions.

- Model — Design and build models by entering reactions, species, parameters, kinetic laws, rules, and events with a graphical user interface, a block diagram editor, or using the MATLAB® Command Window. Verify that the model can be simulated, and use the verification results to fix any incompatibilities in the model. Support SBML.

- Simulate — Observe changes in species amounts and parameter values over time. Convert your model to a system of differential equations and simulate the model numerically with various differential equation solvers. The deterministic solvers include stiff and nonstiff ordinary differential equation (ODE) solvers. The stochastic solvers include a stochastic simulation algorithm with implicit and explicit tau variations. Perform multiple stochastic ensemble runs.

- Analyze — Save data from a simulation, compare simulation and experimental data, perform sensitivity analysis, species or parameter scans, parameter estimation, and search for conserved moieties.

MatLab-based: SBMLToolbox

SBMLToolbox

Version 3.1.2 released 27 April 2010
Sarah M. Keating

SBMLToolbox is built on top of **libSBML** and provides a set of basic functions allowing SBML models to be used in **MATLAB**. SBMLToolbox provides functions for reading, writing, and validating SBML models; viewing model structures in a simple GUI; converting models into a symbolic form suitable for use with MATLAB's **Symbolic Math Toolbox**; and simulate models using MATLAB's ordinary differential equation solvers.

SBMLToolbox works by translating SBML models to/from a MATLAB **structure** called `MATLAB_SBML`. It provides facilities for manipulating this and its substructures within MATLAB, as well as functions for saving and loading `MATLAB_SBML` structures to/from MATLAB data files.

The toolbox is not intended to be a complete Systems Biology toolbox for MATLAB, but rather a platform facilitating getting SBML in and out of MATLAB and serving as a starting point from which users can develop their own functionality.

- **Download**
- **User's Manual**
- **Screenshots**
- **Release notes**
- **Report/view bugs**
- **License**
- **SBMLToolbox paper**

Please cite the paper if you use SBMLToolbox

Article citations are crucial to our academic careers. If you use SBMLToolbox and you publish papers about work that was performed using SBMLToolbox, we ask that you please cite the **SBMLToolbox paper**:

> Keating, S. M., Bornstein, B. J., Finney, A., and Hucka M. (2006) SBMLToolbox: an SBML toolbox for MATLAB users. *Bioinformatics*, 22(10):1275–1277.

MatLab-based: SBtoolbox

Fraunhofer ITWM
CHALMERS Research Centre Industrial Mathematics

Systems Biology Toolbox for MATLAB Version 1.8

- Start
- Toolbox Documentation
- Download Toolbox
- Installation Guide
- Contact

Downloads: 2100
(since April 2005)

SBTOOLBOX Addons

- SBaddon Package

Other Links

- PottersWheel

The Systems Biology Toolbox for MATLAB offers systems biologists an open and user extensible environment, in which to explore ideas, prototype and share new algorithms, and build applications for the analysis and simulation of biological systems. It features a wide range of functions:

- Experiment description functionality **(NEW in Ver. 1.8)**
- Modeling (based on ODEs and chemical reaction equations)
- Handling of measurement data
- Import of SBML models
- Simulation (deterministic and stochastic)
- Steady-state and stability analysis
- Determination of the stoichiometric matrix
- Parameter sensitivity analysis
- Network identification
- Parameter estimation
- Determination of moiety conservations
- Simple model reduction
- Bifurcation analysis
- Localization of mechanisms leading to complex behaviors
- Optimization algorithms
- and some specialized functions ...

Reference:
H. Schmidt, M. Jirstrand: Systems Biology Toolbox for MATLAB: A computational platform for research in Systems Biology, *Bioinformatics* Advance Access, November 2005, DOI 10.1093/bioinformatics/bti799 [PDF]

http://www.sbtoolbox.org

MatLab-based: SBtoolbox2

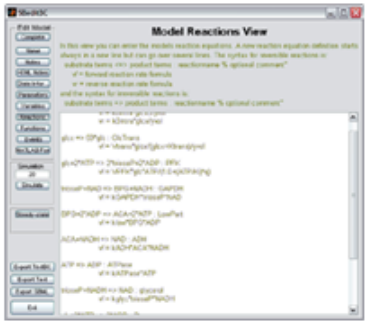
sbtoolbox2

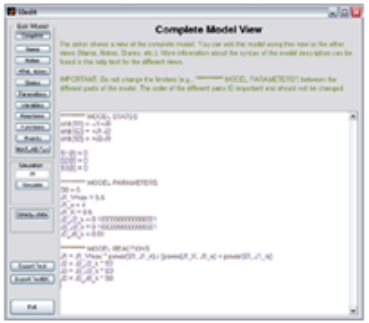
Version 2.1

Overview and Features of the SBT2

Model Representation

Define your models using a simple syntax based on differential equations or based on biochemical reaction equations. To see examples you can click on the figures to the right.





Additionally it is possible to define differential equations, initial conditions, variables, and reactions using a vectorized syntax. An example for which this syntax is useful is:

```
d/dt x[n] = kon*x[1]*x[n-1] + koff*x[n+1] - (kon*x[1]+koff)*x[n]
```

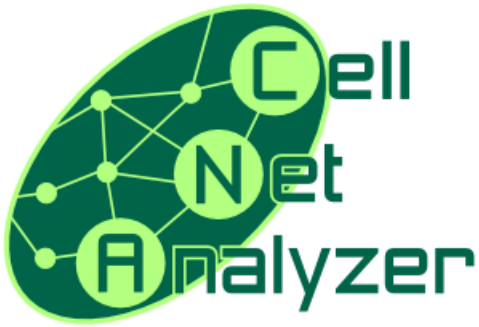
Further examples for models and their syntax (including the vectorized syntax) can be found in the model repository.

SBML Enabled

The toolbox supports the import and export of SBML models. On this page you can find information about its SBML compatibility.

http://www.sbtoolbox2.org/

MatLab-based: CellNetAnalyzer



Home
General Information
Download
Online Manual
Detailed description
Model Repository
Screenshots
Related Links
Publications
Contact

latest version: 9.6.

CellNetAnalyzer

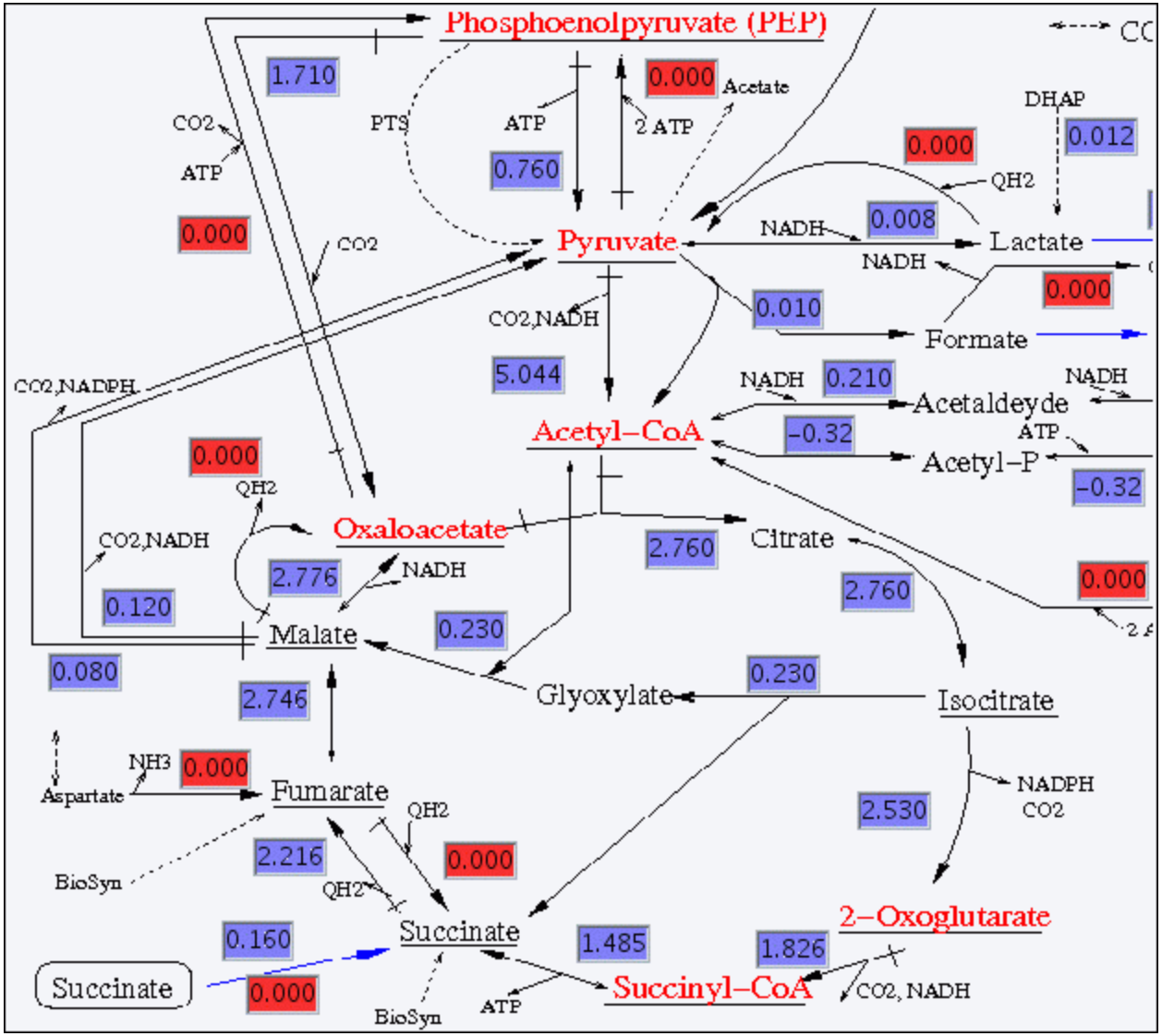
Structural and Functional Analysis of Cellular Networks

CellNetAnalyzer is a package for MATLAB and provides a comprehensive and user-friendly environment for structural and functional analysis of biochemical networks. CNA facilitates the analysis of metabolic (stoichiometric) as well as signaling and regulatory networks solely on their network topology, i.e. independent of kinetic mechanisms and parameters. CNA provides a powerful collection of tools and algorithms for structural network analysis which can be started in a menu-controlled manner within *interactive network maps*. Recently, API functionalities have been added to enable interested users to call algorithms of CNA from external programs. Applications of CNA can be found in systems biology, biotechnology, metabolic engineering, pharmacology, microbiology, chemical engineering.

CellNetAnalyzer is the successor and further development of *FluxAnalyzer*. The latter was originally developed for structural analysis of mass-flow (metabolic) networks; CNA extends *FluxAnalyzer* with a framework for signal-flow networks.

http://www.mpi-magdeburg.mpg.de/projects/cna/cna.html

MatLab-based: CellNetAnalyzer



Mathematica-based: MathSBML

- MathSBML is an open-source package for working with SBML models in Mathematica.

- It provides facilities for reading SBML models, converting them to systems of ordinary differential equations for simulation and plotting in Mathematica, and translating the models to other formats.

- Check the Introduction page (http://sbml.org/Software/MathSBML) for more information.

PySCeS

PySCeS: The Python Simulator for Cellular Systems

home download documentation contact

PySCeS provides a variety of tools for the analysis of cellular systems

- Input via a text based model description language.
- A structural analysis module.
- Integrators for the simulation of a cellular system's evolution over time (LSODA).
- Non-linear solvers that can be used to calculate a systems steady state (HYBRD, NLEQ2).
- A module for performing Metabolic Control Analysis (i.e. elasticities, flux and concentration control coeffcients).
- A bifurcation module for the study of systems which exhibit multiple (stable and unstable) steady states (PITCON)
- A variety of extra utilites such for performing parameter scans, data output and plotting.
- SBML import and export capability.
- PySCeS is developed as Open Source software distributed under a BSD style licence.

PySCeS is developed by the Triple-J Group for Molecular Cell Physiology at Stellenbosch University and the Molecular Cell Physiology Group at the VU University Amsterdam

2010-07-27: PySCeS has been succesfully build on Mac OS X using the Enthought Python Distribution
2010-06-27: PySCeS 0.7.4 test binaries for 2.5 and 2.6 are available for download from SF.net (see all files)
2010-06-21: PySCeS 0.7.4 has been repackaged and now is available as a Ubuntu package

http://pysces.sourceforge.net/index.html

PySCeS

PySCeS download area (add-on modules and related software)

PySCeS software

File (all)	Linux	Mac OS X	Win32 binaries		Ubuntu 10
PySCeS 0.7.0(all files)	src		2.5		
PySCeS 0.7.5 (coming soon)	*src*	*src*	*2.5*	*2.6*	*2.6*

PySCeS dependencies (required for running/building)

It is highly recommended to consider using the Enthought Python Distribution (available for Windows & Macintosh) which contains all the dependencies (with the exception of libSBML) required to run or build PySCeS.
Python: PySCeS has been tested with Python versions 2.5 and 2.6
NumPy: compatible with versions newer than 1.2.1
SciPy: compatible with versions newer than 0.7
Matplotlib: the default plotting interface requires versions 0.98 or newer (requires the **TKagg** backend)

Optional addon packages (required for extended functionality)

libSBML: PySCeS requres libSBML 3 and 4 for SBML support (please ensure that the Python bindings are installed)
Gnuplot: PySCeS supports GnuPlot version 4 as an alternative plotting interface (includes 3D plotting)
IPython interactive shell (under Windows install the pyreadline and pywin32 packages first)
Pyreadline (tab completion for IPython, requires ctypes for Python 2.4)
Pywin32 (needed for Windows IPython install)

SBML-R interface

- R package: SBMLR;
  url: http://rss.acs.unt.edu/Rdoc/library/SBMLR/html/00Index.html

- R package: R-SBML, a package to enable use of the SBML data format in the R computing environment using the libSBML shared library.

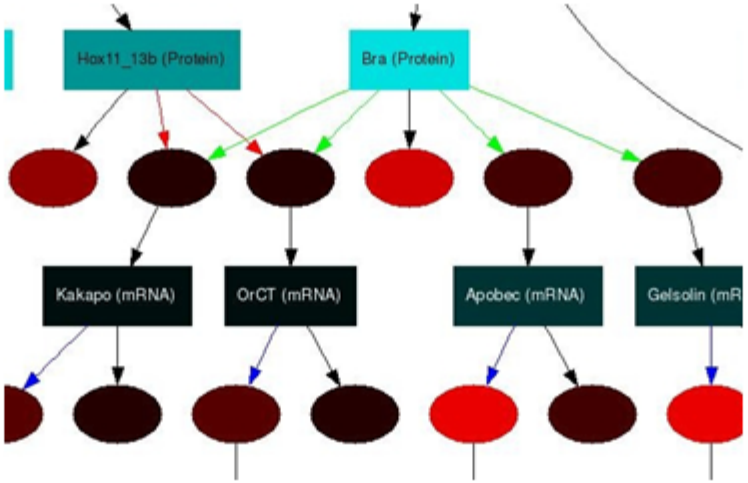
- None of them is available in CRAN, yet.

PyBioS

Max Planck Institute
for Molecular Genetics

MAX-PLANCK-GESELLSCHAFT

a tool for modeling and simulation of cellular systems



HUAZHONG AGRICULTURAL UNIVERSITY

http://pybios.molgen.mpg.de/

PyBioS

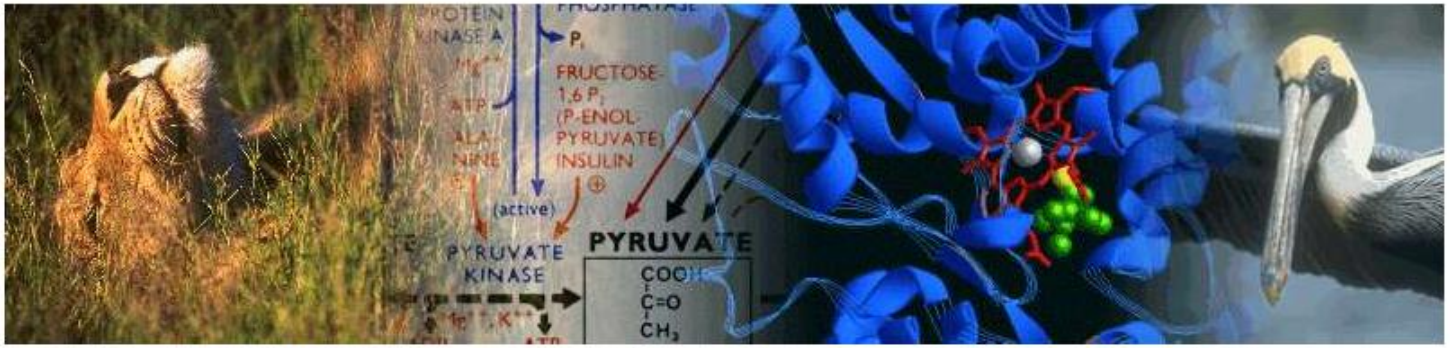
- PyBioS is a system for the modeling and simulation of cellular processes. It is developed at the Max-Planck-Institute for Molecular Genetics in the department of Prof. Lehrach by Christoph Wierling, Elisabeth Maschke-Dutz, Ralf Herwig, and Hans Lehrach.

- PyBioS acts as a model repository and supports the generation of large models based on publicly available information like data of the Reactome database. An ODE-system of a model can be generated automatically based on pre- or user-defined kinetic laws and used for subsequent simulation of time course series and further analysis of the dynamic behavior of the underlying system. The forward-modeling approach supports the formulation of hypotheses, e.g. for in silico knock-out experiments or time series.

注册帐号，在线使用

Mathematica-based: JWS online

JWS Online



Home | Model Database | Project Info | News | Upload | Help | Online servers

The JWS Online Project

JWS Online aims to provide a service to the Systems Biology community by 1) giving access to a database of curated models of biological systems, and 2) allowing the users to run these models in a web browser via an easy to use interface, and 3) helping in reviewing of manuscripts containing kinetic models. In addition to this service role, JWS Online is an important component of an ambitious research initiative: The Silicon Cell. Whereas as a service JWS Online is a repository of published models, in the research activity models are not stored as published but are changed to reflect standardized notation of metabolites and enzymes, and a direct link between model and experimental data is provided.

The JWS Online team

JWS Online was started in 2000 by Jacky Snoep and Brett Olivier, just to try and run simulations in a browser and to make curated models available for everyone. Currently Brett works on his stand-alone simulation tool (PySCeS) while Jacky continued working on JWS Online, now together with Carel van Gend as full-time programmer and Riaan Conradie, Franco du Preez, DuToit Schabbort, Gerald Penkler and Francois du Toit helping out with coding models and Kora Holm making the metabolic schemes. In Amsterdam and Manchester mirror sites of JWS Online are set up and maintained by Cor Stoof and Steve O'Hagan respectively. Cor works in the team as Java programmer on the applet code. Hans Westerhoff is one of the drivers behind the Silicon Cell project which is closely connected to JWS Online more..

Mathematica-based: JWS online

The Inner Workings

more..

Collaborations

The JWS Online team works together with other model databases (Biomodels, SABIO-RK), Systems Biology initiatives (The Silicon Cell, SBML, YSBN, HepatoSys, ESCEC), simulation engines (PySCeS, Copasi), and scientific journals (Microbiology, FEBS J, IET Systems Biology and Metabolomics). more..

Funding

The South African National Bioinformatics Network funds the JWS Online project, providing with infrastructure, student bursaries (Riaan, Franco and Gerald) and a postdoc salary (Carel). more..

Publications

JWS Online and the Silicon Cell project have been described in a number of publications, see more.. When citing JWS Online, please refer to: B.G. Olivier and J.L. Snoep (2004), Web-based kinetic modelling using JWS Online, Bioinformatics 20, 2143-2144.

Visit http://jjj.mib.ac.uk/index.html