Taverna services for systems biology

Paweł Banasik, Anna Gambin and Mikołaj Rybiński
p.banasik@students.mimuw.edu.pl, {aniag, trybik}@mimuw.edu.pl
Institute of Informatics, Warsaw University, Poland

Wydział Matematyki, Informatyki i Mechaniki
Uniwersytet Warszawski

Introduction

Taverna Workbench [1]

What is it for?

- Computational biology tool for facilitating the design and execution of *in silico* experiments.
- Provides easy access to various Web Services available transparently via the Internet.
- Allows to model experiments as workflows.
- Integrates variety of tools (different programming languages, different platforms).

Taverna workflow: set of services, which are to be used, together with connections between their inputs and outputs.

Figure 1: Taverna workbench splash screen.

What are the benefits?

- Experiments are easily repeatable.
- Each run is fully documented.
- The provenance of data is recorded.

The vast majority of Web Services available for use with Taverna are related to sequence analysis (bioinformatics).

Mathematical modelling of biochemical reactions

Example: **simple enzymatic reaction**.

1. Model structure

$$E + S \underset{k_2}{\overset{k_1}{\longleftrightarrow}} ES \xrightarrow{k_3} E + P.$$

2. Kinetics (deterministic semantics; ODEs)
   Law of mass action:

$$\begin{bmatrix} d[S](t)/dt \\ d[E](t)/dt \\ d[ES](t)/dt \\ d[P](t)/dt \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 1 & 1 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} k_1 \cdot [S](t) \cdot [E](t) \\ k_2 \cdot [ES](t) \\ k_3 \cdot [ES](t) \end{bmatrix}$$

   Michaelis–Menten approximation:

$$\frac{d[P](t)}{dt} \approx k_3[E](0)\frac{[S](t)}{[S](t) + K_M} \text{ where } K_M = \frac{k_2 + k_3}{k_1}$$

3. Model parametrization
   Determining values of parameters requires conducting a series of costly and difficult biological experiments.

Sensitivity analysis can assess relative importance of parameters narrowing set of experiments to perform.

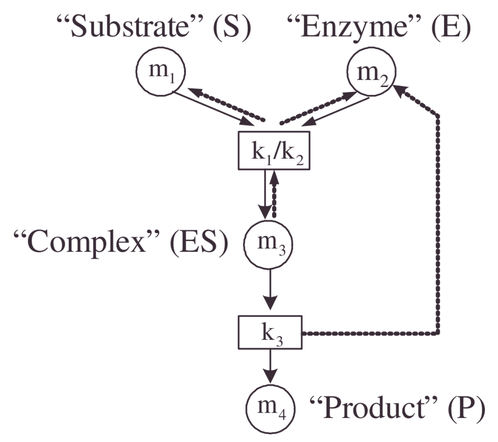

Figure 2: Graphical representation of simple enzymatic reaction model structure (fig. from [2]).

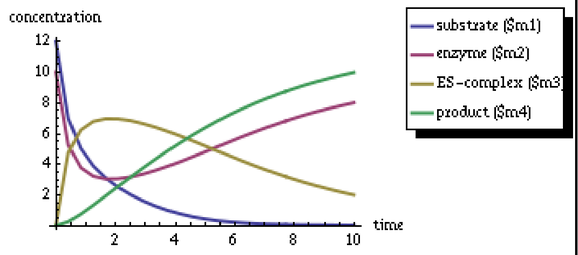

Figure 3: Numerical simulation of the enzymatic reaction mass action law ODEs.

Framework technology

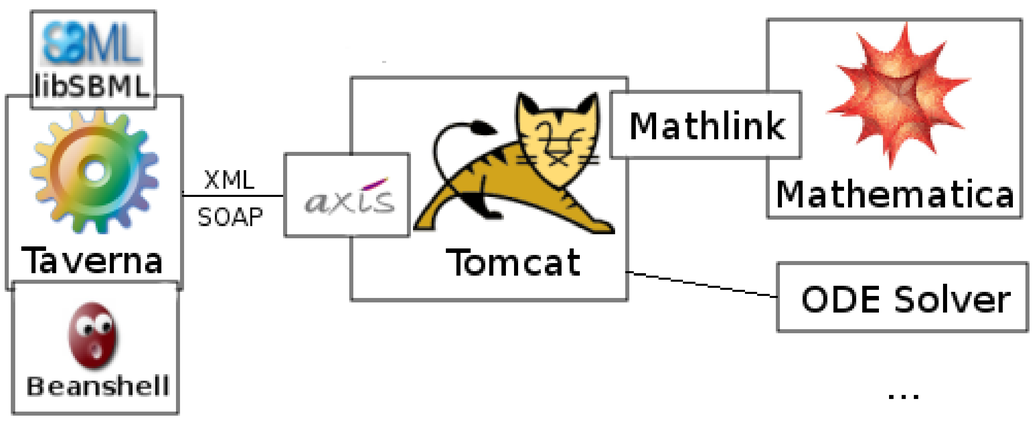

Figure 4: Architecture behind the services. In the left the client side, in the middle and right — the server side.

Client side

- Main element: Taverna workbench.
- Embedded Beanshell scripting host allows for writing custom (simplified Java) code, e.g. for data conversion or dialog boxes.
- libSBML library allows for SBML–based models, in particular for extracting parameters and manipulate their values.
- Client communicates with server via Web Services (using XML messages that follow the SOAP standard).

Server side

- Tomcat web server exposes Web Services to client through Apache Axis2 tool.
- Calculations on the server side are performed by Mathematica and SBML ODE Solver Library.
- Tomcat communicates with Mathematica via J/Link, which enables the usage of Mathematica kernel from a Java program.
- ODE Solver is used via command-line to simulate ODEs derived from SBML models.

Experiments: workflows and results

Multi–Parameter Sensitivity Analysis (MPSA)

MPSA procedure [2, 3]

Step 1. Select parameters to assess.
Step 2. Set parameters range.
Step 3. Uniformly generate independent samples.
Step 4. Calculate samples errors.
Step 5. Classify samples as acceptable or unacceptable.
Step 6. Statistically evaluate sensitivity.

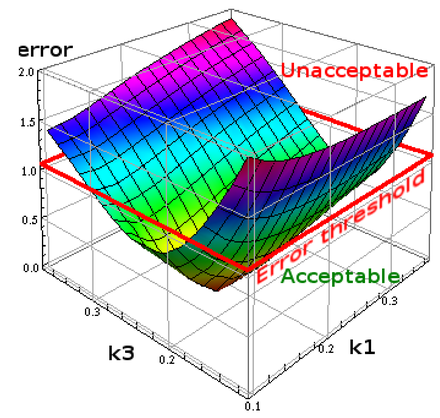

Figure 5: Mean sum squared error surface for the enzymatic reaction $k1$ and $k3$ parameters samples.

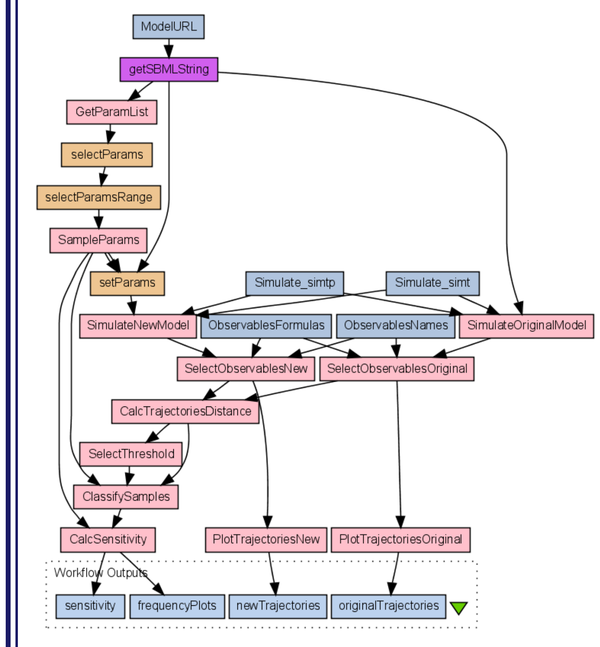

Figure 6: Taverna MPSA workflow.

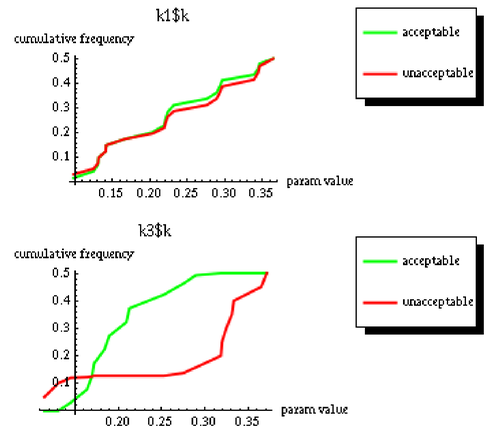

Figure 7: Empirical cumulative distribution functions of acceptable and unacceptable samples projections on $k1$ and $k3$ coordinates.

Pearson correlation coefficient:

- 0.998 for $k1$,
- 0.675 for $k3$.

Local parameter sensitivity 3D plots

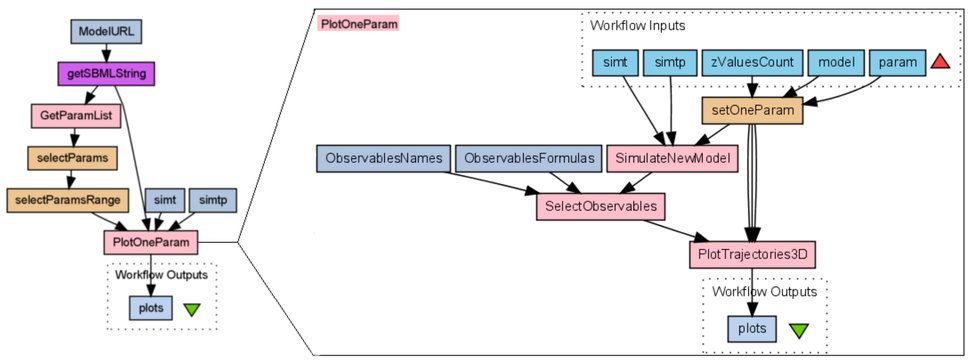

Figure 8: Taverna local sensitivity 3D plot workflow (loop implemented by nested workflow).

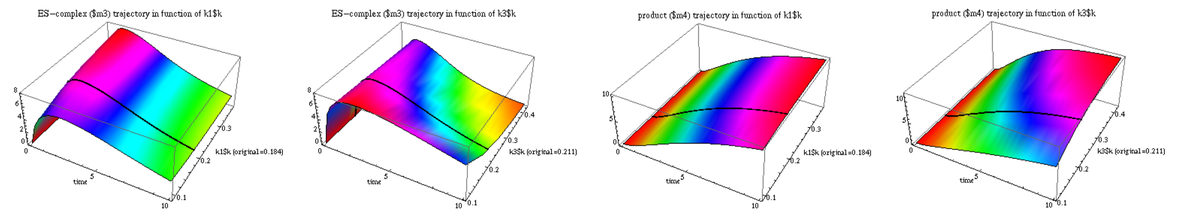

Figure 9: ES complex and P species trajectories sensitivity to local (single) $k1$ and $k3$ parameters variations.

Conclusions

1. We created a set of Taverna services and workflows for SBML MPSA and local sensitivity 3D plots.
2. Application to simple enzymatic reaction model indicated larger influence of $k3$ parameter than $k1$ on system's behavior.
3. Taverna provides data standardization, tools integration, services accesibility and computations transparency.

Future work

- Stochastic systems sensitivity analysis (exploiting gridMathematica).
- Sensitivity analysis with arbitrarily defined error functions (robustness analysis; CSL formulas verification in PRISM tool for stochastic systems).
- Simplification of adding new or modifying existing Taverna services in presented framework architecture.

References

[1] T Oinn, M Addis, J Ferris, D Marvin, M Senger, M Greenwood, T Carver, K Glover, M R Pocock, A Wipat, and P Li. Taverna: a tool for the composition and enactment of bioinformatics workflows. *Bioinformatics*, 20(17):3045–3054, Nov 2004.

[2] K-H Cho, S-Y Shin, W Kolch, and O Wolkenhauer. Experimental Design in Systems Biology, Based on Parameter Sensitivity Analysis Using a Monte Carlo Method: A Case Study for the TNFalpha-Mediated NF-kappaB Signal Transduction Pathway. *SIMULATION*, 79:726–739, Dec 2003.

[3] Zhike Zi, Kwang-Hyun Cho, Myong-Hee Sung, Xuefeng Xia, Jiashun Zheng, and Zhirong Sun. In silico identification of the key components and steps in IFN-gamma induced JAK-STAT signaling pathway. *FEBS letters*, 579:1101–8, Feb 2005.

Acknowledgments: These studies are supported by Polish Ministry of Science and Higher Education (PBZ-MIN 014/P05/2004) and CoE BioExploratorium, University of Warsaw.