Development of a systems biology infrastructure for mathematical modelling and parameter estimation

ID - 197

Mosca Ettore[1,2], Alfieri Roberta[2,3], Merelli Ivan[2], Milanesi Luciano[2,3]
[1]Department of Biotechnology and Biosciences, University of Milano-Bicocca, Milano
[2]Institute for Biomedical Technologies, National Research Council, Segrate (Milano)
[3]CILEA, Consorzio Intrauniversitario per l'Elaborazione Automatica, Segrate (Milano)

Motivation

Nowadays it is important to describe integrated biological processes such as the cell cycle or signal transduction in quantitative functional terms [1]. Since these pathways are not just an assembly of genes and proteins, some properties cannot be fully understood merely by drawing diagrams of their interconnections [2]. In order to model the cell cycle mechanism it is appropriate to use ordinary differential equations (ODEs), if the two basic assumptions of this approach, the well-stirred chemical reactor and sufficient great concentrations, are satisfied [3,4]. According to this approach, each component of the modelled system is associated with a differential equation which describes its temporal rate of change. One of the main problems in developing new mathematical models, using the ODEs approach, lies in the experimental outline of the specific kinetic parameters, such as rate and association constants, needed in the kinetic laws. Moreover, although standards for codification of biological models in computer-readable formats (SBML and CellML) exist, many published models are not included in models repositories, like BioModels, CellML models repositories and JWS on-line. To overcome these limitations we have developed an infrastructure with the aim to improve the creation, the sharing and the simulation of models based on ODEs in the context of the systems biology.

Methods

The technology for the simulation consists in five modules: a model repository, an XML based parser, the MathML to HTML converter, the ODEs simulation engine, and the parameter estimation tool. PHP code links all the modules and generates web interfaces which allow users to interact with the system. Most of the models collected in the repository comes from the BioModels database at EBI, some other models derive from the model authors web sites, or have been manually implemented using the JigCell software [5]. The models are encoded in Systems Biology Markup Language (SBML), an XML-based language specific for biological models. The parser has been written in object oriented PHP and its structure follows the SBML data structure specifications. The implemented system allows users to translate formulas, which in SBML files are encoded with MathML specifications, to HTML using an XSLT library [6] and TTH, a tool to translate TEX to HTML [7]. The simulation engine relies on XPPAUT [8] and the output results are show using GNUPLOT [9]. The parameter estimation is performed using a evolution strategy algorithm, named Stochastic Ranking Evolution Strategy [10] coupled to sundials library [11] for numerical calculations of the differential equations. Parameter estimation execution is parallelized thanks to a C based Message Passing Interface implementation of SRES algorithm [12] on high performance computing device.

Results

The web interface allows the user to acquire information about models in two sections. The first section regards the model publication paper information, including model wiring-diagram. The second allows the users to explore the SBML components of the selected model, including its complete mathematical definition, which consists in species initial concentration and parameters values, differential equations, algebraic equations, assignment rules and events. Our technology allows the real-time simulation of the models: this feature is essential to fully understand the complex behaviours of the pathway, such as oscillations. Moreover, it is possible to observe model response against change of initial conditions, i.e. proteins species concentrations and kinetic parameter values. Results are instantaneously plotted on 2D graphs which show time courses and phase diagrams. In the parameter estimation section it is possible to submit a novel model and optimize its kinetics parameter values according to experimental data provided. The algorithm we used, SRES, has demonstrated good performance in the estimation of continuous variables [13]. Since the computational load of this optimization problem scales rapidly both with the cost of a single model simulation and with the number of its kinetic parameters, the parameter estimation job is executed using high performance computing.

Availability: http://www.itb.cnr.it/cellcycle/models.php

Email: ettore.mosca@studenti.unimib.it