Meeting report
Computational models in systems biology
Laurence Loewe* and Jane Hillston†

Addresses: *Centre for Systems Biology at Edinburgh, School of Biological Sciences, The University of Edinburgh, Kings Buildings, Mayfield Road, Edinburgh EH9 3JU, Scotland, UK. †Laboratory for the Foundations of Computer Science, School of Informatics, The University of Edinburgh, Informatics Forum, 10 Crichton Street, Edinburgh EH8 9AB, Scotland, UK.

Correspondence: Laurence Loewe. Email: Laurence.Loewe@ed.ac.uk

---

A report of the 6th International Conference on Computational Methods in Systems Biology, Rostock, Germany, 12-15 October 2008.

One of the chief goals of systems biology is to build mechanistic mathematical models of biological systems to further the understanding of biological detail. Such models often aim at predicting the outcome of potentially interesting biological experiments, and if such predictions are confirmed by wet-lab observations, an important step forward is made. How exactly such models are constructed and how predictions are computed were at the core of a recent conference on Computational Methods in Systems Biology that brought 80 participants to Rostock, Germany (for conference proceedings see volume 5307 of *Lecture Notes in Bioinformatics* [http://dx.doi.org/10.1007/978-3-540-88562-7]).

A simplistic approach to model construction might be to capture everything that is known about a system and simulate it in supercomputers. While this is appropriate for some systems, it is impossible or highly impracticable for many others. This is mostly due to the complexity of biological systems, which demand simplification to make them amenable to modeling. Such simplifications have to capture the essence of the processes of interest, while neglecting as many of the less important details as possible. Thus, one can consider model building in systems biology as the art of building caricatures of life: capture the essence, ignore the rest.

Constructing systems biology models
Two formalisms called process algebras and Petri nets offer alternative ways of constructing computational systems biology models. Both are concerned with how to specify (mostly quantitative) models of molecular reaction networks in an abstract way that is independent of particular mathematical techniques of analysis. Access to techniques such as differential equations or stochastic simulations is then facilitated automatically by software tools that translate the abstract model into a concrete model that is ready for computation. The advantage of this approach is that one needs to describe the model only once in order to access a variety of analytical techniques, which reduces implementation time and errors.

Corrado Priami (COSBI, Trento University, Italy) gave an overview of the development of the new BlenX programming language, which is based on a process algebraic core, and can be used for modeling biochemical reaction networks that change their topology over evolutionary time. It aims to be more accessible to biologists than process algebras and is supported by a range of software tools. Process algebras rely on abstract formal (textual) notations that describe a biochemical network as a concurrent system, using approaches that were first developed to describe communicating processes in computer science (a concurrent system consists of a number of entities all of which can act simultaneously).

The stochastic pi calculus is another process algebra widely used for modeling cellular systems. Andrew Phillips (Microsoft Research, Cambridge, UK), described a visualizer he has developed to facilitate the understanding of stochastic pi calculus models. He has applied this to the modeling of actin polymerization. Matthias John (University of Rostock, Germany) presented the attributed pi calculus, an extension of the pi calculus that is parameterized with a special language for the definition of the characteristics of interest. In this approach, developed by Uhrmacher's group in Rostock, 'attributes' capture additional information about the entities being represented and allow the simulation models derived to be more faithful to biological reality. This approach can be used to model the phototaxis of *Euglena*. Marek Kwiatkowski

(University of Edinburgh, UK) introduced the continuous pi calculus, which he has been developing. It provides a means of describing a biological system as a set of ordinary differential equations and has been applied to a circadian clock.

One of us (JH) gave an overview of Bio-PEPA, a stochastic process algebra newly developed by Hillston's group that is designed for analyzing biochemical networks using a wide range of mathematical techniques. It has been applied to a number of biochemical pathways, including cyclin regulation in the cell cycle.

Like process algebras, Petri nets also describe a biochemical network as a concurrent system, but primarily use graphical diagrams ('nets') to represent it. They support the abstract theory behind applied methods such as flux-balance analysis, and feature particularly well developed structural analysis techniques, which give insight into how the topology of the network influences the behavior.

Michael Pedersen (University of Edinburgh, UK) introduced compositional definitions of minimal flows in Petri nets and applied them to examples in the Petri-net-based 'Calculus of Biochemical Systems'. Being able to carry out analyses in a compositional way allows much larger models to be handled efficiently. Annegret Wagler (University of Magdeburg, Germany) presented a combinatorial approach to reconstruct all possible Petri nets that could explain a given set of experimental data. Such work is important as it underlies the development of methods for understanding flows through metabolic networks.

Nonlinear kinetic models are often excessively complex to analyze and Hidde de Jong (INRIA, Grenoble, France) presented one approach to make their qualitative analysis easier. He described how on some occasions they can be simplified to piecewise linear differential equation models. This makes it much faster to compute an overview of all possible outcomes, as linear differential equations are easy to solve and a corresponding tool, the Genetic Network Analyzer (GNA) [http://www-helix.inrialpes.fr/gna], is available online.

A method widely used in computational systems biology is the stochastic simulation of molecular systems. Matthias Jeschke (University of Rostock, Germany) has compared the performance of various simulation algorithms and found that the underlying data structures used to implement an algorithm can have a substantial impact on computing time.

Model checking
Model checking is an approach that evaluates a logical expression against a model that describes the behavior of the system. It is used extensively to check the properties of computer systems and is now being applied to systems

biology. In a biological context, model checking might seek to answer the question: 'In the model, will the concentration of a given metabolite exceed a certain value within a certain time?' With the aim of exploring how model checking could be used to estimate parameters in simulation models, Aurélien Rizk (INRIA, Rocquencourt, France) presented a logic language that replaces the usual yes/no satisfaction of a condition with a measure of how close the model is to satisfying the condition. Rizk then relies on optimization algorithms to find sets of parameters that minimize the distance of the model to the formal condition on the basis of observed data and assumed algorithms. It performed well on a MAP kinase (MAPK) signal transduction model with 22 unknowns (examples of unknowns include reaction rates).

In an independent study with the same objective, David Gilbert (Brunel University, London, UK) also estimated parameters in a MAPK signal transduction model. The genetic algorithm Gilbert uses for optimization is not as fast as the algorithms employed by Rizk, but could have other benefits, such as being less affected by local optima. It is too early to assess how these two approaches compare, or to advise on which should be adopted for a particular problem. Future work will have to show how they will compare with established statistical methods for parameter estimation such as maximum likelihood or Bayesian approaches.

Modeling in three dimensions
Substantial advances were reported in building spatial models of intracellular processes. Robert Murphy (Carnegie Mellon University, Pittsburgh, USA) reviewed the work of his group in developing automated methods for determining the intracellular location of proteins tagged with green fluorescent protein. Traditionally, this could only be done by expert microscopic examination. Using image-recognition and machine-learning techniques, it is now possible to automate the task, while retaining a degree of certainty for each identification. The accuracy of the automated system can now match that of human experts if sufficient certainty is demanded for a valid match. It is hoped that such work will help with the automated construction of spatially more explicit models of intracellular processes.

Such models might eventually be passed on to the impressive E-Cell simulation machinery of Koichi Takahashi (RIKEN, Yokohama, Japan), who presented a new, highly efficient method for simulating Brownian dynamics in cells and reviewed the spatial modeling approaches that are supported by the E-Cell Project [http://www.e-cell.org].

Practical applications
Good quantitative models have practical applications. Dieter Oesterhelt (Max Plank Institute of Biochemistry, Martinsried, Germany) demonstrated the down-to-earth potential of

models in systems biology using examples from halophilic archaea, including a simple mechanism for regulating the flagellar motor in a way that guarantees an optimal response of the cell to light. He also presented a flux-balance analysis model with about 700 reactions that is capable of predicting growth rates in response to the addition of 16 amino acids to the growth medium. Nicolas Le Novère (European Bioinformatics Institute, Cambridge, UK) presented models of neurological signaling at multiple scales, where lower scales are used to build helpful abstractions at higher scales. Such models are designed to contribute towards understanding dopamine biochemistry and can generate up to 10 terabytes of data for simulating 1 second in the life of a synapse.

At a simpler level, Carolyn Talcott (SRI International, Menlo Park, USA) presented an abstract model of the central pattern generator located in the buccal ganglia of the marine mollusc *Aplysia*, a long-time model organism for neurobiology. Her simple model considered only a small number of discrete levels of excitation for each neuron but was able to faithfully reproduce the predicted behavior of a more detailed continuous, ordinary differential equation model. Thus, instead of the continuous-state-space approach of differential equations, a logic-based discrete-state-space description of the system is used. This is conceptually much simpler.

The future is likely to see many more mechanistically understood models of actual biological systems that rely on computational methods in systems biology for their construction. The conference showed that there is the potential to greatly enhance our understanding of life if theoretical computer scientists, practical programmers, theoretical biologists and experimental biologists work closely together in order to develop the tools that are needed for constructing and analyzing systems biological models.