Submitted to the Technical Communications of the International Conference on Logic Programming (ICLP'10)
http://www.floc-conference.org/ICLP-home.html

A LOGICAL PARADIGM FOR SYSTEMS BIOLOGY (INVITED TALK)

FRANÇOIS FAGES

EPI Contraintes, INRIA Paris-Rocquencourt,
Domaine de Voluceau, 78150 Rocquencourt, France
E-mail address: Francois.Fages@inria.fr
URL: http://contraintes.inria.fr/

Biologists use diagrams to represent complex systems of interaction between molecular species. These graphical notations encompass two types of information: interactions (e.g. protein complexation, modification, binding to a gene, etc.) and regulations (of an interaction or a transcription). Based on these structures, mathematical models can be developed by equipping such molecular interaction networks with kinetic expressions leading to quantitative models of mainly two kinds: ordinary differential equations for a continuous interpretation of the kinetics, and continuous-time Markov chains for a stochastic interpretation of the kinetics.

The Systems Biology Markup Language (SBML) [7] uses a syntax of reaction rules with kinetic expressions to define such reaction models in a precise way. Nowadays, an increasing collection of models of various biological processes is available in this format in model repositories, such as for instance www.biomodels.net [8].

Since 2002, we investigate the transposition of programming concepts and tools to the analysis of living processes at the cellular level. Our approach relies on a logical paradigm for systems biology which consists in making the following identifications:

biological model = quantitative state transition system
biological properties = temporal logic formulae
biological validation = model-checking
model inference = constraint solving

Our modelling software platform BioCham [6] (implemented in Prolog) is founded on this paradigm. An SBML model can be interpreted in BioCham at three abstraction levels:

- the Boolean semantics (asynchronuous Boolean state transitions on the presence/absence of molecules),
- the continuous semantics (ODE on molecular concentration),
- the stochastic semantics (CTMC on numbers of molecules).

These semantics have been related in the framework of abstract interpretation in [5], showing for instance that the Boolean semantics is an abstraction of the stochastic semantics, i.e. that the possible stochastic behaviors can be checked in the Boolean semantics, and that if a Boolean behavior is not possible, it cannot be achieved in the quantitative semantics for any kinetics. The temporal logics used to formalize the properties of the behavior of the system are respectively the Computation Tree Logic (CTL) for the Boolean semantics,

1998 ACM Subject Classification: algorithm, theory,verification.
Key words and phrases: temporal logic, model-checking, systems biology, hybrid systems.

and a quantifier-free Linear Time Logic with constraints over the reals (LTL($\mathbb{R}$)) for the quantitative semantics. BioCham has been used for querying large Boolean models of the cell cycle by symbolic model-checking [1], formalizing phenotypes in temporal logic [3], searching parameter values from temporal specification [9], measuring the robustness of a system w.r.t. temporal properties [10], and developping in this way quantitative models of cell signalling and cell cycle for cancer therapies [2].

For some time, an important limitation of this approach was due to the logical nature of temporal logic specifications and their Boolean interpretation by true or false. By generalizing model-checking techniques to temporal logic constraint solving [3, 4], a continuous degree of satisfaction could be defined for temporal logic formulae, opening the field of model-checking to optimization.

We believe that this mixing of discrete logical and continuous dynamics, pioneered by constraint logic programming and hybrid systems, and illustrated here in systems biology, is a deep trend for the future in programming and verification.

References

[1] Nathalie Chabrier and François Fages. Symbolic model checking of biochemical networks. In Corrado Priami, editor, *CMSB'03: Proceedings of the first workshop on Computational Methods in Systems Biology*, volume 2602 of *Lecture Notes in Computer Science*, pages 149–162, Rovereto, Italy, March 2003. Springer-Verlag.

[2] Elisabetta De Maria, François Fages, and Sylvain Soliman. On coupling models using model-checking: Effects of irinotecan injections on the mammalian cell cycle. In *CMSB'09: Proceedings of the seventh international conference on Computational Methods in Systems Biology*, volume 5688 of *Lecture Notes in BioInformatics*, pages 142–157. Springer-Verlag, 2009.

[3] François Fages and Aurélien Rizk. On temporal logic constraint solving for the analysis of numerical data time series. *Theoretical Computer Science*, 408(1):55–65, November 2008.

[4] François Fages and Aurélien Rizk. From model-checking to temporal logic constraint solving. In *Proceedings of CP'2009, 15th International Conference on Principles and Practice of Constraint Programming*, number 5732 in Lecture Notes in Computer Science, pages 319–334. Springer-Verlag, September 2009.

[5] François Fages and Sylvain Soliman. Abstract interpretation and types for systems biology. *Theoretical Computer Science*, 403(1):52–70, 2008.

[6] François Fages, Sylvain Soliman, and Aurélien Rizk. *BIOCHAM v2.8 user's manual*. INRIA, 2009. `http://contraintes.inria.fr/BIOCHAM`.

[7] Michael Hucka et al. The systems biology markup language (SBML): A medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531, 2003.

[8] Nicolas le Novère, Benjamin Bornstein, Alexander Broicher, Mélanie Courtot, Marco Donizelli, Harish Dharuri, Lu Li, Herbert Sauro, Maria Schilstra, Bruce Shapiro, Jacky L. Snoep, and Michael Hucka. BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acid Research*, 1(34):D689–D691, January 2006.

[9] Aurélien Rizk, Grégory Batt, François Fages, and Sylvain Soliman. On a continuous degree of satisfaction of temporal logic formulae with applications to systems biology. In Monika Heiner and Adeline Uhrmacher, editors, *CMSB'08: Proceedings of the fourth international conference on Computational Methods in Systems Biology*, volume 5307 of *Lecture Notes in Computer Science*, pages 251–268. Springer-Verlag, October 2008.

[10] Aurélien Rizk, Grégory Batt, François Fages, and Sylvain Soliman. A general computational method for robustness analysis with applications to synthetic gene networks. *Bioinformatics*, 12(25):il69–il78, June 2009.