Theory of Dynamical Model Reductions in Systems Biology and Implementation in Constraint Logic Programming

PhD Thesis Subject, Advisors:

François Fages, Inria Paris-Rocquencourt, http://contraintes.inria.fr

Ovidiu Radulescu, Université de Montpellier II, http://www.umr5235.univ-montp2.fr/spip.php?rubrique21

At the end of the 90s, research in Bioinformatics evolved, passing from the analysis of the genomic sequence to the analysis of post-genomic interaction networks (expression of RNA and proteins, protein-protein interactions, transport, etc.). Systems Biology is the name given to a multidisciplinary research field involving biology, computer science, mathematics, physics, to illustrate this change of focus towards system-level understanding of high-level functions of living organisms from their biochemical bases at the molecular level.

Dynamical models of biochemical reaction networks are composed of a reaction graph (bipartite graph with vertices for species and reactions) in which the reaction vertices are labelled with kinetic expressions (mass action law, Michaelis-Menten, Hill, etc.). Most of the work of our labs consists in analysing the *interplay between the structure* (reaction graphs) *and the dynamics* (Ordinary Differential Equations ODE, Continuous-Time Markov Chains CTMC, or hybrid discrete-continuous interpretations derived from the kinetic expressions) of these models by applying Mathematical and Computer Science methods in Systems Biology.

Model reduction is a key concept in Systems Biology for modelling a biological process at different levels of abstraction and for developing multi-scale models. Model reduction approaches may be used to identify critical components or interactions that should be described with precision, while non-critical parts can be simplified. There exist several classical approaches for model reduction in chemical kinetics (quasi-steady state, QSSA, quasi-equilibrium, QE approximations, lumping, averaging, finding a limiting rate reaction step). In recent years, our labs have been investigating scalable reduction methods for both qualitative [Gay *et al*, 2010, 2011, Fages *et al.*, 2012] and quantitative models [Crudu *et al.*, 2009, Gorban *et al*, 2004, 2008 & 2010; Radulescu *et al*, 2008, 2012; Noel *et al* 2011].

The subject of the Thesis is to merge these techniques to develop a general theory of dynamical model reduction encompassing QSSA, QE etc., together with a general purpose implementation using Constraint Logic Programming to master the computational complexity of several NP-hard problems involved in the computation of reduced dynamical models [Soliman, 2012].

We will combine tropical geometry ideas and constraint logic programming with the aim of developing new tools for model reduction of biochemical reaction network models. Model reduction of multiscale dynamical reaction networks can be based on the idea of dominance. When reaction rates are polynomial or rational functions of the species concentrations, the natural mathematical framework for studying dominance relations is tropical geometry, a new field of algebraic geometry. For instance, fast dynamics of excitable species can be described by keeping only the largest dominant term among the reactions acting on them. Slow dynamics on low dimensional invariant manifolds, preserving species concentration boundedness and leading to reduced models, involve some form of species equilibration. In this case, at least two reaction terms, positive and negative, are equal and dominate the others. We call this problem *tropical equilibration*. Thus, given the structure of the model and the orders of the kinetic parameters, a first step in the proposed model reduction algorithm is to compute possible

tropical equilibrations. It consists in determining a priori unknown orders of species concentrations that satisfy a set of constraints (equality of orders of at least two monomials per species and a set of inequalities with respect to the remaining monomials). This is a NP hard problem.

Once the set of equilibrations is found, the model can be reduced by graph rewritting. The graph rewritting operations consist in species and reactions pooling. The pools of reactions and species correspond to convex bases of the left and right kernels of the stoichiometric matrices of the fast subsystem. The computation of these convex bases (elementary modes) is also NP hard but can be efficiently computed in practice as a constraint satisfaction problem [Soliman 2012].

The Thesis should provide solutions for the two computational problems identified in these reduction methods and a complete implementation allowing for their evaluation on a large scale in model repositories like e.g. *biomodels.net*.

Required skills: strong background in mathematics and computer science, knowledge of constraint logic programming, interest for biology.

Bibliography:

Chabrier-Rivier N, Fages F, Soliman S (2004b). The Biochemical Abstract Machine BIOCHAM. *Lecture Notes in Computer Science* **3082**: 172-91

Crudu A., Debussche A., Radulescu O., Hybrid stochastic simplifications for multiscale gene networks, BMC Sys tems Biology (2009) 3:89.

Fages F., Gay S., Soliman S. An automatic method for curating SBML models based on their ODE semantics. *Submitted*, 2012.

Fages F, Soliman S (2008a). Formal Cell Biology in Biocham. *Lecture Notes in Computer Science* **5016**: 54–80.

Fages F, Soliman S (2008c). From reaction models to influence graphs and back: a theorem. *Lecture Notes in Computer Science* **5054**: 90-102.

Gay S, Soliman S, Fages F (2010). A graphical method for reducing and relating models in systems biology. *Bioinformatics* **26**: i575-81.

Gay S., Fages F., Martinez T., Soliman S., and Solnon C. On the subgraph epimorphism problem. *Submitted*, 2011.

Gorban A.N.and Radulescu O. Dynamic and static limitation in reaction networks, revisited. Advances in Ch emical Engineering (2008) 34:103-173.

Gorban AN, Radulescu O, Zinovyev A (2010). Asymptotology of chemical reaction networks. *Chemical Engineering Science* **65**: 2310-24.

Noel V, Grigoriev D, Vakulenko S, Radulescu O (2011). Tropical geometries and dynamics of biochemical networks. Application to hybrid cell cycle models. Venice, SASB 2011, in press *Electronic Notes in Theoretical Computer Science*, arxiv:1109.4085.

Radulescu O, Gorban A, Zinovyev A, Lilienbaum A (2008). Robust simplifications of multiscale biochemical networks. *BMC Systems Biology* **2**: 86.

Noel V, Grigoriev D, Vakulenko S, Radulescu O (2011). Tropical geometries and dynamics of biochemical networks. Application to hybrid cell cycle models. Venice, SASB 2011, in press *Electronic Notes in Theoretical Computer Science*, arxiv:1109.4085.

Radulescu O, Gorban A, Zinovyev A, Lilienbaum A (2008). Robust simplifications of multiscale biochemical networks. *BMC Systems Biology* **2**: 86.

Radulescu, O. Gorban A.N., Zinovyev A., Noel V., Reduction of dynamical biochemical reaction networks in computational biology. Frontiers in Genetics, 3, 2012.

Soliman S.. Invariants and other structural properties of biochemical models as a constraint satisfaction problem. *Algorithms for Molecular Biology*, 7(15), May 2012.