Journal of Theoretical Biology 459 (2018) 79–89

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/jtb

Graphical requirements for multistationarity in reaction networks and their verification in BioModels☆

Adrien Baudier, François Fages*, Sylvain Soliman

Inria Saclay Île-de-France, Palaiseau, France

ARTICLE INFO

Article history:
Received 4 June 2018
Revised 22 August 2018
Accepted 23 September 2018
Available online 26 September 2018

Keywords:
Multistability
Reaction networks
Influence networks
Positive circuits
Systems biology

ABSTRACT

Thomas' necessary conditions for the existence of multiple steady states in gene networks have been proved by Soulé with high generality for dynamical systems defined by differential equations. When applied to (protein) reaction networks however, those conditions do not provide information since they are trivially satisfied as soon as there is a bimolecular or a reversible reaction. Refined graphical requirements have been proposed to deal with such cases. In this paper, we present for the first time a graph rewriting algorithm for checking the refined conditions given by Soliman, and evaluate its practical performance by applying it systematically to the curated branch of the BioModels repository. This algorithm analyzes all reaction networks (of size up to 430 species) in less than 0.05 second per network, and permits to conclude to the absence of multistationarity in 160 networks over 506. The short computation times obtained in this graphical approach are in sharp contrast to the Jacobian-based symbolic computation approach. We also discuss the case of one extra graphical condition by arc rewiring that allows us to conclude on 20 more networks of this benchmark but with a high computational cost. Finally, we study with some details the case of phosphorylation cycles and MAPK signalling models which show the importance of modelling the intermediate complexations with the enzymes in order to correctly analyze the multistationarity capabilities of such biochemical reaction networks.

1. Introduction

The wide variety of cells in a multicellular organism show that cells with identical copies of DNA may differentiate in different cell types. In the late 40's, Max Delbruck at Caltech suggested that each type of cell could correspond to a distinct steady state in the dynamics of their shared gene expression network. In order to analyze such large networks, René Thomas conjectured in 1980 that the existence of a positive (resp. negative) feedback loop was a necessary condition for multistationarity (resp. sustained oscillations) Thomas (1981). Those conjectures were later proved in various formalisms (Boolean or discrete transition systems, differential equations) with various degrees of generality. In 2003, Christophe Soulé finally proved Thomas' necessary condition for multistationarity with full generality for dynamical systems defined by differential equations (Soulé, 2003).

In his mathematical formalization of the conjecture, Soulé considers a differentiable mapping F from a finite dimensional real vector space to itself, and for each point a, the directed graph $G(a)$ where the arcs are the non-zero entries of the Jacobian matrix of F, labeled by their sign. He shows that if F has at least two non-degenerate zeroes, there exists a such that $G(a)$ has a positive circuit.

When applied to (protein) reaction networks however, Thomas' necessary condition for multistationarity fails short since it is trivially satisfied as soon as there exists either a bimolecular or a reversible reaction. Indeed, a bimolecular reaction such as a complexation reaction immediately creates a mutual inhibition between the two reactants, i.e. a positive circuit, and a reversible reaction produces a mutual activation, i.e. again a positive circuit, making Thomas' necessary condition always true in those networks.

Nevertheless, reaction models are widespread in computational systems biology and it would be very desirable to be able to predict the absence of multistationarity by systematically checking such conditions with efficient algorithms. For instance, the BioModels database[1]~(Chelliah et al., 2013) is a repository of more than 600 hand-curated models written in the Systems Biology

☆ Article submitted to a special issue *in memoriam* of René Thomas.

* Corresponding author.
E-mail addresses: francois.fages@inria.fr (F. Fages), sylvain.soliman@inria.fr (S. Soliman).

[1] http://biomodels.net/.

https://doi.org/10.1016/j.jtbi.2018.09.024

Markup Language (SBML) (Hucka et al., 2008) mostly with reaction rules, over several tenths or hundred of molecular species. There are hundreds more models in the non-curated branch, and thousands of models imported from metabolic networks databases with even larger numbers of reactions and species.

Soulé's proof, as most preceding and following proofs, uses the fact that the existence of multiple steady states implies a non-injectivity property which is shown to be equivalent to a determinant being zero for some values of reaction rate constants. One approach, called the Jacobian approach, is thus to use symbolic computation methods to directly compute the roots of that determinant. If it is non-zero, one can conclude to the absence of multistationarity. This is the approach taken by Feliu and Wiuf (2013a). Interestingly, they evaluated their algorithm, implemented in Maple 16, on the curated branch of BioModels (323 networks in their case), showing that 31,6% were injective and that only 8,3% of the networks of this benchmark caused memory overflow by that method. On the sequences of r phosphorylation cycles of Wang and Sontag (2008), they could check non-injectivity up to $r = 17$ cycles in 1200 s.

In this paper, we follow the alternative graphical approach to multistationarity analyses. We describe a graph rewriting algorithm which deals with sequences of $r = 1000$ phosphorylation cycles in a second, and analyzes the curated branch of BioModels (506 networks in our case) with a maximum computation time of 50 milliseconds per network (including large networks of size up to 430 species), while concluding to the non existence of multiple steady states in 160 networks of size up to 54 species in that benchmark, i.e. with a similar ratio of 31.6% of results concluding to non-multistationarity.

This algorithm is based on a refinement of the graphical requirements of Soulé (2003) given by the third author in Soliman (2013) as a necessary condition for the existence of multiple steady states in (biochemical) reaction networks. Similar graphical requirements have also been given in Banaji and Craciun (2010) without restriction to mass-action law kinetics, but to our knowledge, it is the first time that they are implemented and evaluated systematically in model repositories. For instance, we are not aware of similar evaluations obtained with the Chemical Reaction Network Toolbox[2] for systematically checking the graphical conditions for multistationarity of Feinberg's Chemical Reaction Network Theory (CRNT) (Craciun and Feinberg, 2006; Feinberg, 1977).

More specifically, we present a series of graph rewriting algorithms for checking the different graphical requirements of Soliman (2013), and analyze their practical performance in the curated models of BioModels, in order to:

- evaluate when the original condition of Thomas allows us to rule out multistationarity;
- evaluate when the following three extra conditions given in Soliman (2013) become conclusive, namely:
  1. the positive circuit must not come from twice the same reaction;
  2. the positive circuit must not come from a reaction and its reverse reaction;
  3. the positive circuit must not involve all species of a conservation law;
- evaluate when even stronger conditions based on the rewirings detailed in Soulé (2003) and Soliman (2013) are necessary to conclude, namely
  1. by sign change of incoming arcs on a set of species,
  2. or by permuting the arcs to a set of target species.

[2] https://crnt.osu.edu/toolbox-history-and-explanation.

For this study, we used our software modelling environment BIOCHAM[3](Calzone et al., 2006; Fages et al., 2017) to load all models from the curated branch of BioModels, improve their writing in SBML with well-formed reactions using the algorithm described in Fages et al. (2015), compute the conservation laws (Soliman, 2012), compute their influence multigraph labelled by the reactions (Fages et al., 2018; Fages and Soliman, 2008b) and export the labelled multigraph in the Lemon library format[4]. Then we used an implementation in C++ of the algorithm presented in this paper to search for positive circuits with the different refined conditions on the labelled influence multigraph, and evaluate their respective contributions for the analysis of multistationarity in BioModels. All the computation times obtained with this algorithm given in this paper were obtained on a standalone desktop Linux machine with an Intel Xeon 3.6 GHz processor[5].

The rest of this article is organized as follows. The next section presents the refined necessary conditions for multistationarity in reaction networks described in Soliman (2013) and detailed here with five levels of conditions. The following section presents a graph rewriting algorithm for checking those conditions, and evaluates its computational complexity. Section 4 shows the remarkable performance of this algorithm by applying it systematically to the curated part of the model repository BioModels, including models out of reach of Jacobian-based symbolic computation methods, and details the effect of the five levels of refined conditions in this benchmark. Section 5.1 considers the models of double phosphorylation cycles of Wang and Sontag (2008) and shows a very low quadratic empirical complexity of the graphical algorithm, again in sharp contrast to symbolic computation methods. Section 5.2 focuses on model 270 of ERK signalling that contains 33 species and 42 reactions resulting in an influence multigraph of 126 arcs with many positive and negative feedback loops, yet for which our graphical algorithm demonstrates the absence of multistationarity. These examples illustrate the importance of modelling the intermediate complexes in enzymatic reactions to obtain multiple steady states, and show the sensitivity of both the dynamical properties of the models and of our graphical conditions to the writing of enzymatic reactions with or without intermediate complexes. We conclude on the remarkable performance of the graphical approach to analyze multistationarity in reaction models of large size, and on some perspectives to further improve our algorithm and generalize this approach.

2. Necessary condition for multistationarity in reaction networks

Let us consider a biochemical reaction system with n species $S_1, \ldots, S_n$ and m reactions $R_1, \ldots, R_m$. Using notations from Kaltenbach (2012) we write:

$$R_j = \sum_{i=1}^{n} y_{ij} S_i \longrightarrow \sum_{i=1}^{n} y'_{ij} S_i$$

The y and y' represent the stoichiometric coefficients of the reactants and products of the reaction. The rate law associated with reaction R_j will be written v_j. This defines a dynamical system, in the form of an Ordinary Differential Equation (ODE): $\dot{x} = F(x)$ where x_i is the concentration of species S_i and

$$f_i(x) = \sum_j v_j(x) \cdot (y'_{ij} - y_{ij})$$

[3] http://lifeware.inria.fr/biocham4.
[4] http://lemon.cs.elte.hu/.
[5] For the sake of reproducibility, our programs and data are available at https://lifeware.inria.fr/wiki/Main/Software#JTB18.