BioSystems 149 (2016) 113–124

Contents lists available at ScienceDirect

BioSystems

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

Predicting changes of reaction networks with partial kinetic information

Joachim Niehren[b,c], Cristian Versari[a,c,*], Mathias John[a,c], François Coutte[a,d], Philippe Jacques[a,d]

[a] Université de Lille, France
[b] Inria, Lille, France
[c] BioComputing Team of CRIStAL Lab (CNRS UMR 9189), Lille, France
[d] Research Institute Charles Viollette, EA-7394-ICV, Lille, France

ARTICLE INFO

Article history:
Received 6 January 2016
Received in revised form 18 July 2016
Accepted 1 September 2016

Keywords:
Reaction networks
Model based prediction
Abstract interpretation
Constraint solving
Metabolic engineering
Genetic engineering

ABSTRACT

We wish to predict changes of reaction networks with partial kinetic information that lead to target changes of their steady states. The changes may be either increases or decreases of influxes, reaction knockouts, or multiple changes of these two kinds. Our prime applications are knockout prediction tasks for metabolic and regulation networks.

In a first step, we propose a formal modeling language for reaction networks with partial kinetic information. The modeling language has a graphical syntax reminiscent to Petri nets. Each reaction in a model comes with a partial description of its kinetics, based on a similarity relation on kinetic functions that we introduce. Such partial descriptions are able to model the regulation of existing metabolic networks for which precise kinetic knowledge is usually not available.

In a second step, we develop prediction algorithms that can be applied to any reaction network modeled in our language. These algorithms perform qualitative reasoning based on abstract interpretation, by which the kinetic unknowns are abstracted away. Given a reaction network, abstract interpretation produces a finite domain constraint in a novel class. We show how to solve these finite domain constraints with an existing finite domain constraint solver, and how to interpret the solution sets as predictions of multiple reaction knockouts that lead to a desired change of the steady states. We have implemented the prediction algorithm and integrated it into a prediction tool.

This journal article extends the two conference papers John et al. (2013) and Niehren et al. (2015) while adding a new prediction algorithm for multiple gene knockouts. An application to single gene knockout prediction for surfactin overproduction was presented in Coutte et al. (2015). It illustrates the adequacy of the model-based predictions made by our algorithm in the wet lab.

1. Introduction

Mathematical methods for analysing reaction networks (Feinberg, 1987; Calzone et al., 2006; Hoops et al., 2006; Fages et al., 2015) often require full kinetic information for all reactions, while in systems biology practice only partial information is usually available. Therefore, we study the problem of how to model reaction networks with partial kinetic information, and how to reason qualitatively about such models with methods from computer science. In particular, we wish to predict changes of reaction networks with partial kinetic information that may or must lead to a target change of their steady state.

When full kinetic information is not available, the existing model-based reasoning methods tend to ignore the kinetic information all over. Most typically, this holds for flux balance analysis (Orth et al., 2010; Papin et al., 2004) when applied to metabolic networks (Otero and Nielsen, 2016; Sohn et al., 2010). The missing information is then compensated heuristically by the adoption of ad hoc optimization criteria. Alternatively, pathway analysis approaches (Papin et al., 2004) rely on the structure of reactions networks, but the combinatorial nature of the problem makes difficult their application to densely interconnected networks. Both above methods have extensions that deal with partial kinetic information about inhibitors. This is done by adding Boolean constraints

* Corresponding author at: Université de Lille, France.
E-mail address: cristian.versari@univ-lille1.fr (C. Versari).

http://dx.doi.org/10.1016/j.biosystems.2016.09.003
0303-2647/

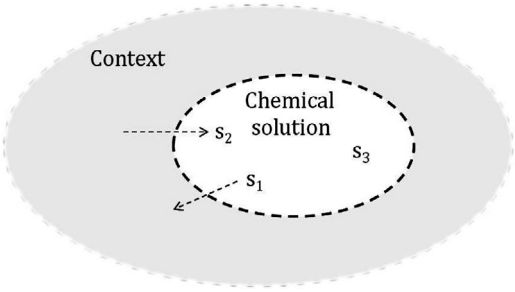

Fig. 1. A reaction network describes a system of chemical reactions that acts on a chemical solution, the inflow of molecules from the context, and the outflow of molecules into the context.

that state the conditions on which an inhibitor blocks a reaction Jungreuthmayer and Zanghellini (2012). However, blocking inhibitors are not really appropriate in deterministic semantics, where the average over blocked and unblocked situations is to be considered. Therefore, it remains open how to deal with nonblocking inhibitors, which only slow down a reaction in average.

In the first part of this article, we propose a new modeling language for reaction networks with partial kinetic information, a short version of which was presented at the CMSB'2015 conference Niehren et al. (2015). Our language is parameterized by a similarity relation on kinetic functions, so that the rate laws of chemical reactions need only to be specified up to similarity. For instance, two kinetic functions could be considered similar if they have the same monotonic behavior. Then, the kinetic function mapping x to $42x$ or to $5x/(7+x)$ would be similar, since whenever x increases then the values of both terms increase, and whenever x decreases then they both decrease.

Reaction networks are applied to a chemical solution, which is typically placed within a context. This may be another chemical solution that is subject to an adjacent reaction network or to an experimental environment. The situation is illustrated in Fig. 1. Since we do not want to model the context, we equip reaction networks with an interface to possible contexts only. The interface specifies which species can inflow from the context and which species may outflow into the context. It should be noticed that the inflows are controlled by the context, while the outflows are controlled by the network.

We also assume that some of the reactions of the network may be candidates for knockout. The choice of whether a reaction is knocked out or not remains external to the network. For the prediction task we are interested in network changes possibly combining multiple reaction knockouts and influx changes.

The models of reaction networks in our language have a graphical syntax that is reminiscent of Petri nets, and also an equivalent XML syntax. They are given a steady state semantics in terms of arithmetic equations, which define the relation between influxes and outfluxes of the network in steady states, depending on which knockout candidates were knocked out. The steady state semantics subsumes the usual flux balance equations, but enhanced with equations on the rates of reactions based on the partial kinetic information. This information is expressed by variables for kinetic functions that are subject to similarity constraints. In this way, the inhibitors of a reaction slow its rate down rather than shutting it down completely, while the activators of a reaction speed its rate up.

Our language can be used to model metabolic networks with complex regulation such as for *subtilis* in Mäder et al. (2011). An example is the regulation network of the PIlv-Leu promoter of *subtilis*, which regulates the metabolism of the branched-chain amino acids Valine, Leucine, and Isoleucine. Previous models of these metabolic networks in the Subtiwiki were not given any formal semantics, so that they could not be used directly for qualitative prediction algorithms. A detailed model of this network in the language presented here was published in Coutte et al. (2015) recently.

In the second part of this article, we develop a new prediction algorithm that can be applied to any reaction network modeled in our language. This algorithm does qualitative reasoning (Forbus, 1997) based on abstract interpretation (Cousot and Cousot, 1979), by which the partial kinetic knowledge is discretized while the unknowns are abstracted away. A first version of the algorithm was presented at the VMCAI'2013 conference John et al. (2013), but restricted to networks of reactions with mass action kinetics. The second version, presented at the CMSB'2015 conference Niehren et al. (2015) and extended here, removes this limitation.

Given a reaction network, abstract interpretation can be applied to infer a difference constraint that relates network changes to changes in the steady states. This can be used to predict which network changes may or must lead to an expected change of the outfluxes. As already stated above, the network changes that we are interested in are reaction knockouts, influx changes or multiple combinations thereof. This becomes possible since the models of reaction networks in our language do describe how network changes affect the steady state semantics.

Difference constraints are finite domain constraints from a novel class. The second step of our prediction algorithm consists in computing the set of all solutions of a difference constraint. Since difference constraints are finite domain constraints, their solution set is always finite. We built two constraint solvers for difference constraints based on finite domain constraint programming. The solver reported earlier in the conference paper preceding this article was developed from scratch in the programming language Scala Odersky (2014). Since then we developed a new solver by reduction to the MiniZinc solver (Nethercote et al., 2007) for finite domain constraints. While our Scala solver could enumerate only the n-best solutions where $n \leq 3000$ while consuming considerable time (more than 10 min), the MiniZinc solver can indeed return the complete set of all solutions sets in all our applications, while enumerating more than 5000 solutions in less then a second.

Solutions of difference constraints can be interpreted as predictions of network changes that may or must lead to an overproduction target. When only seeing the n-best solutions, one can find solutions with few network changes that are compatible with the overproduction target. But since we are now having access to the complete solutions sets of difference constraints obtained from reaction networks in practice, we can distinguish the solutions that are merely compatible with the overproduction target from those that safely entail it. As we will argue, safe solutions correspond to predictions that must satisfy the overproduction target, while compatible solutions yield predictions that may satisfy the overproduction target or not. It also turns out that multiple network changes may be necessary to obtain safe solutions, while single knockouts may not be enough.

We have implemented the prediction algorithm and integrated it into a prediction tool. We illustrate this tool and our prediction algorithm at the example of two simplified models of leucine production of *B. subtilis*, that focus on the regulation of the PIlv-Leu promoter. The target here is leucine overproduction. For the simpler of the two models, the prediction based on the graphical model can be done manually by humans. This illustrates that our algorithm formalizes a natural kind of qualitative reasoning.

In a follow up work Coutte et al. (2015), our algorithms were applied to single knockout prediction for leucine overproduction in a larger and more realistic model. The predictions obtained there were not safe, but still 6 out of 14 predictions could be verified successfully by gene knockouts in the wet lab.

Compared to the previous two conference papers John et al. (2013) and Niehren et al. (2015), we extended the prediction