BIOINFORMATICS

Vol. 00 no. 00 2012
Pages 1–2

CRNreals: a toolbox for distinguishability and identifiability analysis of biochemical reaction networks

Gábor Szederkényi [1,2]*, Julio R. Banga [1] and Antonio A. Alonso [1]

[1](Bio)Process Engineering Group, IIM-CSIC, Spanish National Research Council, C/Eduardo Cabello, 6, 36208 Vigo, Spain
[2]Process Control Research Group, Systems and Control Laboratory, Computer and Automation Research Institute of the Hungarian Academy of Sciences (MTA SZTAKI), Kende u. 13-17, H-1111, Budapest, Hungary

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Associate Editor: XXXXXXX

ABSTRACT

Summary: Chemical reaction network theory (CRNT) is widely used in modeling and analysing complex biochemical systems such as metabolic networks and cell signalling pathways. Being able to produce all the biologically and chemically important qualitative dynamical features, CRNs have attracted significant attention in the systems biology community. It is well-known that the reliable inference of CRN models generally requires thorough identifiability and distinguishability analysis together with carefully selected prior modeling assumptions. Here we present a software toolbox CRNreals that supports the distinguishability and identifiability analysis of CRN models using recently published optimization-based procedures.

Availability and Implementation: The CRNreals toolbox and the associated documentation are available at `http://www.iim.csic.es/~gingproc/CRNreals/`. The toolbox runs under the popular MATLAB computational environment and supports several free and commercial linear programming (LP) and mixed integer linear programming (MILP) solvers.

Contact: szeder@scl.sztaki.hu

1 INTRODUCTION

The key importance of dynamics in the explanation of complex phenomena occurring in living systems is now a commonly accepted view (Alon, 2007) that underscores the importance of mathematical model building and model parameter estimation (Banga and Balsa-Canto, 2008). An important family of nonnegative nonlinear dynamical systems is the class of deterministic chemical reaction network (CRN) models obeying the mass-action law (Horn and Jackson, 1972). Such networks can be used to describe pure chemical reactions, but they are also widely used to model the dynamics of intracellular processes, metabolic or cell signalling pathways.

Chemical reaction network theory (CRNT) is originated in the 1970's by the pioneering works of Horn, Jackson and Feinberg (Horn and Jackson, 1972; Feinberg, 1987). Since then, CRNT have gained an increasing attention, and many strong results have been published in the field on the relation between network structure and qualitative dynamical properties (Shinar and Feinberg, 2010; Conradi and Flockerzi, 2011).

It is known from the early literature that different CRNs can produce exactly the same kinetic differential equations (Horn and Jackson, 1972). This phenomenon is called *macro-equivalence* or *dynamical equivalence*. However, effective optimization-based procedures for the analysis and synthesis of dynamically equivalent CRNs have been developed only recently, using the possibility of transforming propositional logic statements encoding structural properties into MILP problems (Szederkényi, 2009; Szederkényi *et al.*, 2011b,a). The primary purpose of our toolbox is to make these numerical methods accessible in an easily usable way.

As a related software tool, first we have to mention the Chemical Reaction Network Toolbox (Ellison and Feinberg, 2000) that can produce detailed reports about the basic network properties, multiple and degenerate steady states, injectivity and (strong) concordance. Additionally, the open source ERNEST Toolbox (Soranzo and Altafini, 2009) performs a detailed model analysis of the input CRN by determining the basic system features and by using the Deficiency Zero or Deficiency One Theorems. The toolbox is also capable of running the Deficiency One Algorithm where applicable. However, both of the above-mentioned toolboxes assume that the structure of the analyzed network is a priori known, therefore they have no functionality for examining dynamical equivalence.

2 METHODS AND IMPLEMENTATION

The basic model form of CRN dynamics describing the time-evolution of species is the following: $\dot{x} = Y \cdot A_k \cdot \psi(x)$, where $x \in \mathbb{R}^n$ is the concentration vector of the species, $Y \in \mathbb{R}^{n \times m}$, called the *complex composition matrix*, stores the stoichiometric coefficients of the complexes, $A_k \in \mathbb{R}^{m \times m}$ contains the information corresponding to the weighted directed graph of the reaction network, and $\psi_j(x) = \prod_{i=1}^{n} x_i^{Y_{ij}}$, for $j = 1, \ldots, m$ (see, e.g. (Horn and Jackson, 1972) for more details). Therefore the basic data structure used by the CRNreals toolbox for uniquely representing a CRN is the matrix pair (Y, A_k). Two reaction networks with distinct specific values of the reaction rate

*to whom correspondence should be addressed

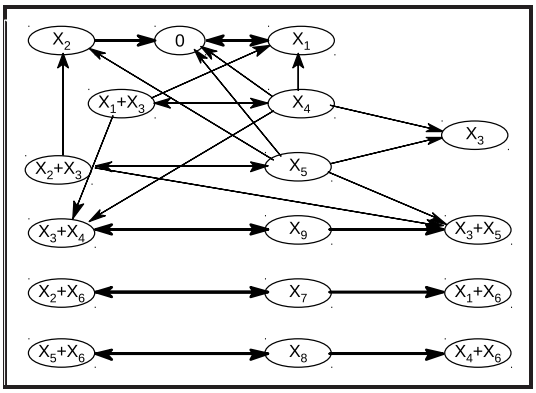

Fig. 1. Dense realization of a biochemical switch analysed in (Szederkényi *et al.*, 2011a). Core reactions are indicated by thick directed edges.

coefficients are called *dynamically equivalent realizations, if they give rise to the same set of differential equations.*

Besides several utility functions for basic model analysis and manipulation, the main high-level functionality of the toolbox includes the following.

1. Algorithmically building the so-called canonical CRN mechanism (Érdi and Tóth, 1989) from kinetic polynomial ordinary differential equations.
2. Finding dense and sparse realizations containing the maximal and minimal nonzero reaction rate coefficients (see Fig. 1). We note that dense realizations give a unique super-structure with a fixed complex set, and that the CRN structure is unique if and only if the structures of the dense and sparse realizations are identical (Szederkényi *et al.*, 2011b).
3. Finding reversible and weakly reversible realizations. The existence of such realizations has a key effect on the boundedness of solutions and on the robust stability of the system (depending also on the deficiency).
4. Finding detailed balanced and complex balanced realizations if they exist. These properties have also important stability implications.
5. Computing dynamically equivalent realizations with the minimal and maximal number of chemical complexes from a previously defined complex set.
6. Finding the so-called 'core' and 'non-core' reactions of the CRN. Core reactions are present in any dynamically equivalent CRN realization and thus they are indispensable components of the system, while non-core reactions can (at least mathematically) be substituted by others (Szederkényi *et al.*, 2011a).

The toolbox was implemented in the MATLAB computational environment becuase of the easy and straightforward manipulation of matrices and the flexible setup of LP and MILP solvers. The toolbox provides conversion functions for importing and exporting mass-action model structures using the libSBML API. Therefore, computed structures can be easily transferred for further analysis e.g. to the ERNEST toolbox.

Fig. 1 shows the structure of the dense realization of a biochemical switch model analyzed in (Szederkényi *et al.*, 2011a). The dense realization and the core reactions were determined using the `CRN_FindSDRealization` and `CRN_FindCoreReactions` functions, respectively (the detailed computations can be found in Example 4 of the toolbox).

3 CONCLUSION

A software toolbox called CRNreals was presented in this note for the analysis and synthesis of dynamically equivalent CRNs. The tools provided by our software aid modeling and dynamical analysis in the following areas: (i) they support the distinguishability and identifiability analysis of CRN models by deciding whether a given CRN has a unique structure or not, (ii) they clearly extend the application scope of CRNT results by searching for dynamically equivalent network structures with preferred properties such as density/sparsity, (weak) reversibility, detailed/complex balance, etc., (iii) they allow the CRNT-based analysis of nonnegative polynomial models possibly coming from application areas other than (bio)chemistry. Further work in future versions will be focused on making the applied numerical methods more effective, and on including the recently developed algorithms related to linear conjugacy thus extending the notion of dynamical equivalence.

Funding: Spanish MICINN project 'MultiScales' (DPI2011-28112-C04-03) and CSIC intramural project 'BioREDES' (ref. PIE-201170E018), Hungarian Scientific Research Fund (grant no. 83440).

Conflict of Interest: None declared.

REFERENCES

Alon, U. (2007). *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman & Hall / CRC Mathematical & Computational Biology.

Banga, J. R. and Balsa-Canto, E. (2008). Parameter estimation and optimal experimental design. *Essays in Biochemistry*, **45**, 195–209.

Conradi, C. and Flockerzi, D. (2011). Multistationarity in mass action networks with applications to ERK activation. *Journal of Mathematical Biology*, **available online**, DOI: 10.1007/s00285-011-0453-1.

Ellison, P. and Feinberg, M. (2000). How catalytic mechanisms reveal themselves in multiple steady-state data: I. Basic principles. *Journal of Molecular Catalysis A: Chemical*, **154**, 155–167.

Érdi, P. and Tóth, J. (1989). *Mathematical Models of Chemical Reactions. Theory and Applications of Deterministic and Stochastic Models*. Manchester University Press, Princeton University Press, Manchester, Princeton.

Feinberg, M. (1987). Chemical reaction network structure and the stability of complex isothermal reactors - I. the deficiency zero and deficiency one theorems. *Chemical Engineering Science*, **42**, 2229–2268.

Horn, F. and Jackson, R. (1972). General mass action kinetics. *Archive for Rational Mechanics and Analysis*, **47**, 81–116.

Shinar, G. and Feinberg, M. (2010). Structural sources of robustness in biochemical reaction networks. *Science*, **327**, 1389–1391.

Soranzo, N. and Altafini, C. (2009). ERNEST: a toolbox for chemical reaction network theory. *Bioinformatics*, **25**, 2853–2854.

Szederkényi, G. (2009). Computing sparse and dense realizations of reaction kinetic systems. *Journal of Mathematical Chemistry*, **47**, 551–568.

Szederkényi, G., Banga, J. R., and Alonso, A. A. (2011a). Inference of complex biological networks: distinguishability issues and optimization-based solutions. *BMC Systems Biology*, **5**, 177.

Szederkényi, G., Hangos, K. M., and Péni, T. (2011b). Maximal and minimal realizations of reaction kinetic systems: computation and properties. *MATCH Communications in Mathematical and in Computer Chemistry*, **65**, 309–332.