Available online at www.sciencedirect.com
ScienceDirect
Automatica 44 (2008) 982–989
automatica
www.elsevier.com/locate/automatica

Brief paper

Symbolic reachability analysis of genetic regulatory networks using discrete abstractions ☆

Grégory Batt[a,b,c,1], Hidde de Jong[a,*], Michel Page[a,e], Johannes Geiselmann[a,d]

[a] *INRIA Grenoble - Rhône-Alpes, 655 avenue de l'Europe, Montbonnot, 38334 Saint Ismier Cedex, France*
[b] *Center for Information and Systems Engineering, Boston University, USA*
[c] *Verimag, Grenoble, France*
[d] *Laboratoire Adaptation et Pathogénie des Microorganismes (CNRS UMR 5163), Université Joseph Fourier, Grenoble, France*
[e] *Ecole Supérieure des Affaires, Université Pierre Mendès France, Grenoble, France*

Received 3 August 2006; received in revised form 23 April 2007; accepted 2 August 2007
Available online 21 December 2007

Abstract

We use hybrid-systems techniques for the analysis of reachability properties of a class of piecewise-affine (PA) differential equations that are particularly suitable for the modeling of genetic regulatory networks. More specifically, we introduce a hyperrectangular partition of the state space that forms the basis for a discrete abstraction preserving the sign of the derivatives of the state variables. The resulting discrete transition system provides a qualitative description of the network dynamics that is well-adapted to available experimental data and that can be efficiently computed in a symbolic manner from inequality constraints on the parameters.

Keywords: Piecewise-affine differential equations; Qualitative analysis; Discrete abstraction; Hybrid systems; Genetic regulatory networks; Systems biology

1. Introduction

A class of piecewise-affine (PA) differential equations introduced by Glass and Kauffman (1973) in the seventies has been shown particularly suitable for modeling so-called *genetic regulatory networks*, networks of genes, proteins, small molecules, and their mutual interactions that are involved in the control of intracellular processes. The dynamics of these networks is hybrid in nature, in the sense that the continuous evolution of the concentration of proteins and other molecules is punctuated by discrete changes in the activity of genes coding for the proteins. This switch-like character of gene regulation is well-captured by the PA models, which have the additional advantage that the qualitative dynamics of the systems is relatively simple to analyze, even in higher dimensions, without the use of numerical values for the kinetic parameters. Given that such information is usually absent in molecular biology, the PA models have been found to be a valuable tool for the practical analysis of complex genetic regulatory networks, which would be difficult to handle with more conventional nonlinear models.

The dynamical properties of the class of PA models considered here have been the subject of active research for more than three decades (e.g., Glass & Kauffman, 1973; Belta, Esposito, Kim, & Kumar, 2005; Edwards, 2000; Ghosh & Tomlin, 2004; Gouzé & Sari, 2002; Mestl, Plahte, & Omholt, 1995, see Batt, Ropers, de Jong, Page, & Geiselmann, 2007 for further references). In our previous work (de Jong et al., 2004), we have made a contribution to the analysis of these PA models by showing how to use differential inclusions to deal with discontinuities in the righthand side of the equations. Moreover, we have proposed algorithms and tools to compute a discrete representation of the state space dynamics in the form of a state transition graph.

In this note, we carry the analysis of the PA models further on a number of points, borrowing concepts and techniques from the field of hybrid systems. First, we partition the state space

☆ This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Michael Henson under the direction of Editor Frank Allgöwer.

* Corresponding author.
E-mail addresses: Gregory.Batt@inria.fr (G. Batt), Hidde.de-Jong@inrialpes.fr (H. de Jong), Michel.Page@iae-grenoble.fr (M. Page), Hans.Geiselmann@ujf-grenoble.fr (J. Geiselmann).
[1] Present Address: INRIA Paris-Rocquencourt, 78153 Le Chesnay, France.

0005-1098/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2007.08.004

into hyperrectangular regions in which the time derivatives of the solutions have a unique sign pattern. In a second step, this partition motivates the definition of a discrete abstraction (Alur, Henzinger, Lafferriere, & Pappas, 2000) that leads to a discrete transition system providing a finer-grained description of the qualitative dynamics of the system than was hitherto possible and which is better adapted to currently available experimental data. Third, we give rules for the symbolic computation of the discrete state transition system from inequality constraints on the parameters. The implementation of these rules has been shown to scale up to large and complex PA models of genetic regulatory networks.

A long version of this note, containing the proofs of all lemmas and propositions, as well as examples of the application of the method to an actual biological network, is available as supplemental material on the INRIA web site (Batt et al., 2007). This work extends a short and preliminary version of the paper presented at the HSCC conference (Batt et al., 2005).

2. PA systems

The dynamics of genetic regulatory networks can be described by PA differential equation models using step functions to account for regulatory interactions (Glass & Kauffman, 1973; Mestl et al., 1995). Fig. 1 gives an example of the PA model of a simple two-gene network. Below we define the models and review some mathematical properties.

We denote by $x = (x_1, \ldots, x_n)' \in \Omega$ a vector of cellular protein concentrations, where $\Omega = \Omega_1 \times \cdots \times \Omega_n \subset \mathbb{R}^n_{\geqslant 0}$ is a bounded n-dimensional hyperrectangular state space region. For each protein concentration x_i, $i \in \{1, \ldots, n\}$, we distinguish a set of constant, strictly positive threshold concentrations $\{\theta_i^1, \ldots, \theta_i^{p_i}\}$, $p_i > 0$. At its threshold concentrations a protein may affect the expression of genes encoding other proteins or the expression of its own gene, thus changing the regulatory mode of the system. The threshold concentrations induce a natural partition of Ω into hyperrectangular regions (de Jong et al., 2004).

Definition 1 (*Mode domain partition*). $\mathscr{M}$ is the hyperrectangular partition (Batt et al., 2007) of Ω induced by $\{\theta_i^1, \ldots, \theta_i^{p_i}\}$. The sets $M \in \mathscr{M}$ are called *mode domains*.

Fig. 2(a) shows the mode domain partition of the state space of the example network. We distinguish between mode domains like M^2 and M^7, which are located on (intersections of) threshold hyperplanes, and mode domains like M^1, which are not. The former are called *singular* and the latter *regular* mode domains. We denote by $\mathscr{M}_\mathrm{r}$ and $\mathscr{M}_\mathrm{s}$ the sets of regular and singular mode domains, respectively. Note that $\mathscr{M} = \mathscr{M}_\mathrm{r} \cup \mathscr{M}_\mathrm{s}$.

The PA models with step functions can be defined on the mode partition as follows (de Jong et al., 2004; Glass & Kauffman, 1973):

$$\dot{x} = h(x) = \mu^M - \nu^M x, \quad x \in M \in \mathscr{M}_\mathrm{r}, \tag{1}$$

where μ^M is a vector of positive constants and $\nu^M = \mathrm{diag}(\nu_1^M, \ldots, \nu_n^M)$ a diagonal matrix of strictly positive constants. That is, in each mode domain the rate of change of

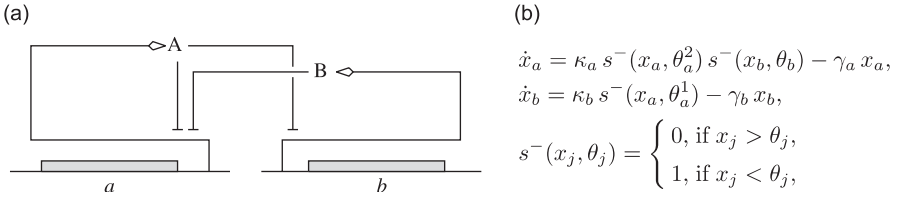

Fig. 1. (a) Example of a genetic regulatory network of two genes (a and b), each coding for a regulatory protein (A and B). Protein B inhibits the expression of gene a, while protein A inhibits the expression of gene b and its own gene. (b) PA model with step functions corresponding to the network in (a). Protein A is synthesized at a rate κ_a, if and only if the concentration of protein A is below its threshold θ_a^2 ($x_a < \theta_a^2$) and the concentration of protein B below its threshold θ_b ($x_b < \theta_b$). The degradation of protein A occurs at a rate proportional to the concentration of the protein itself ($\gamma_a x_a$).

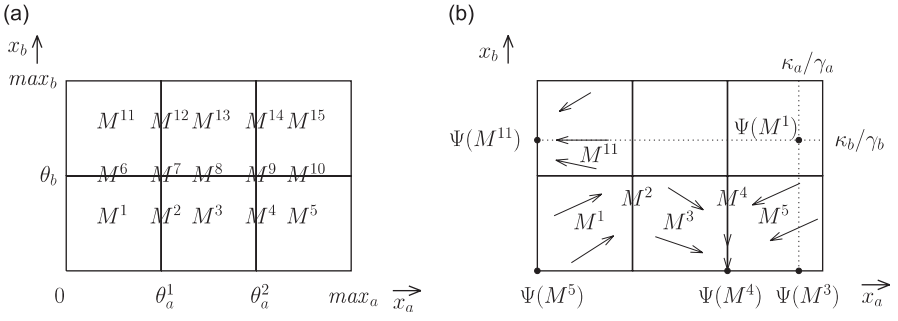

Fig. 2. (a) Mode domain partition of the state space for the model of Fig. 1(b). (b) Focal sets and vector fields associated with the mode domains $M^1, \ldots, M^5, M^{11}$. The use of differential inclusions gives rise to sliding mode solutions in M^4 (de Jong et al., 2004; Gouzé & Sari, 2002).