Gene 518 (2013) 70–77

Contents lists available at SciVerse ScienceDirect

Gene

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

HSM — a hybrid system based approach for modelling intracellular networks

Alvis Brazma [a,1], Karlis Cerans [b,1], Dace Ruklisa [c,1], Thomas Schlitt [d,1], Juris Viksna [b,*,1]

[a] European Bioinformatics Institute, Hinxton, Cambridge CB10 1SD, UK
[b] Institute of Mathematics and Computer Science, Rainis Blvd 29, Riga LV-1459, Latvia
[c] UCL Genetics Institute, Darwin building, Gower Street, London WC1E 6BT, UK
[d] King's College London, Institute for Mathematical and Molecular Biomedicine, Hodgkin Building, Guy's Campus, London SE1 1UL, UK

ARTICLE INFO

Available online 21 December 2012

Keywords:
Gene regulatory networks
Qualitative modelling of biological systems
Analysis of dynamics of biological systems
Hybrid system based modelling

ABSTRACT

The paper proposes a hybrid system based approach for modelling of intracellular networks and introduces a restricted subclass of hybrid systems – HSM – with an objective of still being able to provide sufficient power for the modelling of biological systems, while imposing some restrictions that facilitate analysis of systems described by such models.

The use of hybrid system based models has become increasingly popular, likely due to the facts that: 1) they provide sufficiently powerful mathematical formalism to describe biological processes of interest and do it in a 'natural way' from the biological perspective; 2) there are well established mathematical techniques as well as supporting software tools for analysing such models. However often these models are very dependent on the quantitative parameters of the system (concentrations of proteins, their growth functions etc.) that are seldom exactly known, instead of more limited information of the system that can be observed in practice (directions of change in concentrations, but not the exact values etc.). As a result these models may work well for simulation of the system (prediction of its state starting from some initial conditions), but are too complicated for prediction of all possible qualitatively different behaviours a modelled system might have.

With HSM we try to propose a hybrid system based formalism that is still sufficiently powerful for description of biological systems, while being as restricted as possible to facilitate the analysis of the systems described. We separate between the quantitative system parameters and their qualitative values that can be observed in practice. For HSM we provide an algorithm that analyses the system without the need to know the exact parameter values.

We apply our model and analysis methods to a well-studied gene network of λ-phage. The phage has two well-known qualitatively different behaviours – *lysis* and *lysogeny*. We show that our model has an *attractor structure* that corresponds well to these two behaviours and that these are the *only* stable behaviours that can be exhibited by the system. The algorithm also generates (in principle biologically verifiable) hypotheses about the mutations of λ-phage that should change its observable behaviour.

1. Introduction

Study of the intracellular and gene regulatory network (GRN) behaviour in particular is one of the central topics of the systems biology. No comprehensive definition of a gene network can be given easily, but roughly it can be described along these terms – DNA encodes proteins, a complex cellular machinery 'reads' the DNA and makes (expresses) these proteins, the proteins then interact with other proteins and with the DNA, and these interactions sometimes change the rate of expression. This dynamics can lead to a complex biological behaviour such as cell growth, division, differentiation, death, and other types of qualitative behaviour. When describing gene networks we can distinguish between (1) *network structure* – a diagram showing which proteins interact with what other elements of the network, and (2) *numerical parameters* characterizing the strength of such an interaction and its quantitative effects on expression. Note that experimentally it is often difficult to measure the quantitative parameters of the model accurately. Assuming that it is possible to separate the structure and quantitative parameters of the model, we can study to what extent the qualitative behaviour of the system depends on the structure of the network alone, and to what extent the exact quantitative values (relative or absolute) of the parameters are crucial.

Motivated by these assumptions we propose a *Hybrid System Model* (HSM) for the description of intracellular networks and GRNs

Abbreviations: GRN, gene regulatory network; HS, hybrid system; HSM, Hybrid System Model; *cI*, repressor protein CI gene; *cII*, repressor protein CII gene; *cIII*, repressor protein CII gene; *cro*, control of repressor operator protein gene; *int*, integrase protein gene; *N*, antiterminator protein N gene; *O*, initiation for replication protein O gene; *P*, initiation of replication protein P gene; *Q*, antitermination protein Q gene; *xis*, excisionase protein gene; *Struc*, artificial gene for structural proteins.

* Corresponding author. Tel.: +371 67213716, fax: +371 67820153.
E-mail address: juris.viksna@lumii.lv (J. Viksna).

[1] The authors are listed in alphabetical order and have equally contributed to the paper.

0378-1119/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2012.11.084

in particular. HSM can be viewed as a restricted version of 'typical' hybrid system based models that still provide sufficient power for modelling of biological systems, but the restrictions imposed upon HSM facilitate the analysis of the models. HSM provides a direct way of modelling both discrete events, such as protein binding, and continuous behaviour, such as protein concentration changes.

The HSM approach allows separating the description of the network structure and quantitative model parameters. The dynamical behaviour of such a system can be described as a path through potentially infinite *state space*, where each state is characterized by a discrete mode of the system (e.g., which proteins have bound DNA) and the values of the continuous variables (e.g., protein concentrations) at a certain point in time. We formally define a *qualitative observable behaviour* of a hybrid system and show that it is related to the model's *attractor structure*. An attractor is a region in the state space to which the system may converge and never leave it. It has been hypothesized that the attractors in a biological system correspond to the possible phenotypes, such as cell types (Kauffman, 1969). Our method enables us to study the attractor structure of HSM and its dependence on the system's structure as well as the concrete (absolute or relative) parameter values.

We apply our modelling and analysis methods to a well-studied gene network of λ-phage (McAdams and Shapiro, 1995). λ-phage has two known and qualitatively different behaviours – *lysis* and *lysogeny*. We show that our model has an attractor structure that corresponds to these two behaviours and is largely independent of the concrete parameter values. More precisely, the model exhibits four possible behaviours, one of which corresponds to lysis, one to lysogeny, and the remaining ones to altered 'lysis' and 'lysogeny'-like behaviours. Whether the system will exhibit the normal or altered lysis/lysogeny behaviour depends on the relative ordering of some of the parameter values. Constraints on the relative order of parameter values have to be imposed in order to achieve a correspondence between our model and real world observations.

The hybrid system model proposed here is a continuation of the authors' previous work on *Finite State Linear Model* (FSLM) (Brazma and Schlitt, 2003; Ruklisa et al., 2005; Schlitt and Brazma, 2006). The *Hybrid System Model* we are proposing can be regarded as both simplification and extension of FSLM (by keeping with legacy we will refer to this particular model as HSM and use the abbreviation HS for hybrid systems in general). It drops a number of assumptions, e.g. linear changes of the concentrations of biological substances – a biologically questionable assumption that is not essential for the analysis of a model. It also allows to model some biological processes directly even if they cannot be described simply by the activities of the genes or their binding sites. (Although FSLM has turned out to be suitable for describing λ-phage, it still required artificial 'encoding' of some biological processes.)

Hybrid system based models have become increasingly popular, most likely because: 1) they provide a powerful mathematical formalism for describing biological processes of interest and exploit it in a 'natural way' from the biological point of view; 2) well established mathematical techniques as well as supporting software tools exist for analysing such models. We omit a detailed survey and instead only give references to some of the most known approaches here. It should be emphasized that as far as we know all of these models can be used for the simulation of the system and in some cases also for the formal checking whether the model behaves as the modelled system should. However, typically they are too dependent on the knowledge of parameters when trying to predict all possible observational behaviours of the system.

The first explicit application of a HS based approach to the modelling of biological networks is probably due to (Alur et al., 2001), who considers a very general class of HS models. Their research is based on their earlier comprehensive work on HS (Alur et al., 1995), and uses a special language for defining models in HS framework (Alur et al., 2000). The authors show that HS models are very adequate for description and simulation of biological networks, however they use extremely general HS formalisms and rely heavily on specific differential equations that define the changes of continuous variables.

Another approach was developed by (de Jong et al., 2003; de Jong et al., 2004). Their modelling framework is based on piecewise linear (PL) differential equations describing continuous changes and does not involve discrete states. However, discrete states are introduced implicitly by the domains in which a particular set of equations apply. The authors have developed a tool for simulation and have applied it for modelling of a specific behavioural cycle of *Bacillus subtilis*. A very similar model based on PL differential equations has been used by (Batt et al., 2005) to model nutritional stress response in *E. coli*.

The first HS model specifically tailored for modelling biological systems was likely proposed by (Ghosh and Tomlin, 2004), who introduced piecewise affine hybrid automata. This is a formalism based on restricted HS where the changes of continuous variables are limited by the changes induced by their interaction with an affine vector field (this still allows for quite complicated behaviour). These restrictions allow to design a practical algorithm for computing backward reachable sets from equilibrium states. The method is applied to the model of Delta-Notch signalling, however the authors note that the setting is too simplistic for judging the method's biological relevance.

GRN models that are based on Timed Automata (TA) are related to both hybrid system and FSLM based approaches and have been proposed by (Batt et al., 2007) and (Siebert and Bockmayr, 2008). The models do not include continuous variables explicitly describing concentrations of proteins. Nevertheless continuous variables are used for the time delays and these in turn allow to describe linear changes of concentrations. The approach is demonstrated on a number of examples, including models of λ-phage, although the λ-phage model is rather simplified as it includes just 2 genes: *cro* and *cI*. Noting the difficulties incorporating the gene *cII* into their model Siebert and Bockmayr (2008) suggest that TA models might be too weak for a more realistic description of λ-phage.

A Linear Hybrid Automaton (LHA) that is specifically tailored for GRNs is proposed by (Ahmad et al., 2007). Based on PL differential equations, it is one of a few models which attempt to distinguish between quantitative and qualitative aspects of the system. The authors also study the stability (cyclicity) of the system and provide an algorithm for checking whether the proposed conditions are either sufficient or necessary for such a behaviour to occur. The method has been applied to a study of mucus production system in *P. aeruginosa*. The state space of this model is very small however, consisting of just 6 states.

One of the most recent studies by (Fromentin et al., 2010) proposes a HS based Temporal Evolution Model. The method is applied to a model of *Drosophila* circadian cycle. Model checking software is used to identify the constraints that guarantee the cyclic behaviour of the system. Like in most HS based approaches (but not HSM) the model allows instant resetting of the values of continuous variables to 0. It is not clear whether such a functionality of the model has a biological interpretation or just artificially enforces the desired cyclic behaviour.

2. Materials and methods

2.1. HSM modelling framework

A general hybrid system can be described by a transition diagram (oriented graph) between a finite number of 'modes', a set of continuous variables, a set of continuous functions according to which the variables change in a given mode, and a set of conditions (predicates) which determine when the system switches from one mode to another. When such a system is used to model a gene regulatory network, the modes represent different combinations of transcription factor binding site states (the binding site states are determined by proteins