BioSystems 152 (2017) 24–43

Contents lists available at ScienceDirect

BioSystems

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

Biosystem models, generated from a complex rule/reaction/influence network and from two functionality prototypes

M. Varga [a,*], A. Prokop [b], B. Csukas [a]

[a] Research Group on Process Network Engineering, Kaposvar University, 40 Guba S, 7400, Kaposvar, Hungary
[b] Department of Chemical & Biomolecular Engineering, Vanderbilt University, Nashville, TN, USA

ARTICLE INFO

Article history:
Received 28 July 2016
Accepted 23 December 2016
Available online 3 January 2017

Keywords:
Reaction network
Influence network
Rule network
State prototype
Transition prototype
Model generation
Editable graphical model
Quantitative model
Qualitative model
Multiscale model
p53 signalling at cancer

ABSTRACT

In this work we have further developed the Direct Computer Mapping (DCM) based modelling and simulation methodology. A unified, transition-based representation of complex rule, reaction and influence networks has been introduced and two prototypes (one general state- and another general transition-prototype) have been developed for the unified functional modelling of the state and transition nodes. Starting from the network and from the functional prototypes, an automatic generation method of the graphically editable and extensible GraphML description of biosystem models has been elaborated. The new developments have been implemented in the improved kernel of DCM models. The applied knowledge representation makes possible the unified generation and execution of the balance-based quantitative and influence- or rule-based qualitative, as well as optionally time-driven, multiscale biosystem models. Application of the developed methodology has been illustrated by the improved implementation of the formerly studied and upgraded example biosystem model for combining the detailed, quantitative p53/miR34a signalling system with the pathological model through an extended rule-based coupling model.

1. Introduction

Systems Biology deals with complex, multiscale and partially known process systems, while the knowledge about the structures and functions evolves rapidly. It is a great challenge for the parallel developing holistic and functionality driven modelling approaches.

The purely ***holistic approach*** tries to understand the system's behaviour from the topological properties of the underlying network structure. Networks are simplified mappings of real systems, in which nodes represent the elements (e.g. compounds) and edges describe some form of relationship (interaction, transport, communication, etc.), that exist between nodes. Edges can be non-directed or directed and they can be associated with a weight that differentiates the importance, cost, speed, etc. of the respective relationships. The topological characteristics of the networks offer useful insights into the structure of the studied system (e.g. Strogatz, 2001; Newman, 2003; Newman et al., 2006). The most often used properties of the networks (e.g. Caldarelli, 2007) are the degree, the assortativity, the closeness, the betweenness, the density of links, the average path length, the largest minimum path and the modularity. Average values of these properties describe the global characteristics of the network, while statistical distributions characterize the specific sets of the single nodes.

Network analysis is an important element of modelling in Systems Biology (e.g. Barabasi and Oltvai, 2004; Alberghina and Westerhoff, 2005; Böde et al., 2007). It tries to understand the behaviour of complex systems from the topological properties of the underlying network in a quite formal, but easily computable way.

In contrary, the purely ***functional approach*** of process modelling is based on the functionalities of the systems, without any explicit consideration of the structural features. The foundation of these models had developed before rapid development of IT started. Probably, the first general definition of process models was introduced by (Kalman et al., 1969) in the State Space Model. The state and output functions of the State Space Model for continuous processes were analogously used for the abstract automaton representation of the discrete processes.

Marquardt (1996) and Yang et al. (2013) analysed the state-of-art of process modelling and simulation. The conventional approaches focus on the balance equations and on the event based representation of process systems, while the models are described

* Corresponding author.
E-mail address: varga.monika@ke.hu (M. Varga).

http://dx.doi.org/10.1016/j.biosystems.2016.12.005
0303-2647/

Notations and abbreviations

It is to be noted, that considering the syntax of the model description, the identifiers of components, participating in the models are written by lower case letters.

bcl2	B-cell lymphoma 2, controlling apoptosis
cdkcyc	Refers to CDK4/6 and CylinE2 together
cdk4/6	Cyclin-dependent kinase 4/6
cyclinE2	Associated with CDK2 in a functional kinase complex. exhibits catalytic control over cell cycle
DAE	Differential algebraic equations
DEVS	Discrete event systems
dbc1	Deleted gene in Breast cancer 1
dbc1_like	a hypothetic protein that plays an analogous role to dbc1
DCM	Direct Computer Mapping, a modelling and simulation methodology for a wide range of process systems
mdm2	Murine double minute
mdmx	Structurally homologous protein to mdm2, but it lacks ubiquitin ligase activity (e.g. Kim et al., 2010)
miR34a	A specific miR, participating in p53 controlled signalling processes (Unify this spelling over the MS)
miRNAs	Micro RNAs, small, non-coding RNA molecules
MTBP	Protein, its expression indicates cancer cell types (e.g. Alam et al., 2012)
network	A simplified mapping of real systems, composed by nodes (N) and by (directed or indirected) edges between them, while $E \subset N \times N$
net	Nets composed by two kinds of nodes (S states and T transitions), as well as directed edges (E) between them, while $E \subset (S \times T) \cup (T \times S)$
ODEs	Ordinary differential equations
OpenMI	Open modelling interface, www.openmi.org
p53	Tumour suppressor deacetylated (inactive) protein, in humans it is encoded by the TP53 gene
p53Ac	Acetylated (active) p53
p53mut	Non-functioning, mutant p53
PDEs	Partial differential equations
SBML	Systems Biology Markup Language, http://sbml.org/Main_Page
sirt1	Silent information regulator 1
STN	State transition net

by various mathematical formalisms (e.g. ODEs, PDEs and DAEs, as well as DEVS and STN systems, respectively).

The various process modelling methods and tools, listed comprehensively at http://sbml.org/, were applied for the solution of several specific biosystem models (e.g. Mol et al., 2014; Ryll et al., 2014; Heredia et al., 2015).

Combining of reaction network based and rule-based approaches seems to be a difficult task (e.g. Blinov and Moraru, 2012), but there are effective rule-based solutions (e.g. Chylek et al., 2015). Agent-based approach (e.g. An et al., 2009) was also successfully applied for biosystem modelling.

There are general purpose frameworks (e.g. OpenMI) that offer languages and toolkits for the combined use of various kinds of computer models.

Considering the random nature of functioning of low-copy number species in biological systems, the importance of stochastic reaction networks has also been recognized and used (e.g. Goutsias, 2007). Recently there are interesting efforts to combine stochastic modelling with control theoretical considerations in model-based biosystem engineering (e.g. Briat and Khammash, 2013).

Regarding to the ***combination of structural and functional thinking***, the essential structure of process models can be represented by nets. In the past decades, interesting efforts have been made for the implementation of quantitative functionalities in higher order ***Quantitative Petri Nets*** (e.g. Chen and Hofestädt, 2003; Peleg et al., 2005).

Several authors have emphasized that dynamic characteristics of network properties play a crucial role in the functioning of the system under investigation (e.g. Liu et al., 2011; Csermely et al., 2013). These works ***combined network analysis with the linear state space model***s. However, linearity seems to be a quite arbitrary assumption and there might be obviously more complex, multiscale, event-driven, non-linear, hybrid processes, with considerable time delays behind. Bayesian approach for a non-linear case was analysed by (Oates et al., 2014).

An interesting combination of structural and functional thinking is published by Cardelli (2014), that described dynamics with the so called ***influence networks***, determined by the promoting and inhibiting influences between the components (e.g. proteins). It is worth mentioning that in the real world processes rather the transition (i.e. synthesis, degradation, activation, and inactivation) than the components are promoted or inhibited. Regardless, the influence network helps to discover the ***essential motifs***, as well as useful morphisms between these motifs.

Nowadays, hierarchical modelling of complex systems is being replaced by the paradigm of ***multiscale modelling***. Hierarchy, as an excellent methodology, was developed originally for the recognition. Nevertheless, in the lack of other methodologies, engineering applied hierarchical architectures also for control and design. The basic problem with this approach, however, is that the space of hierarchical problem solving is determined from outside, in a top-down way. In contrary, multiscale methodology tends to link together quite different models, used for the description of the studied sub-systems bottom-up. The multiscale process models are built from various sub-models having different spatial and/or temporal scales. In addition these sub-models can belong to various disciplines. Multiscale modelling seems to be a keynote approach also in computational system's biology (Meier-Schellersheim et al., 2009; Dallon, 2010; Bernard, 2013). According to our understanding, multiscale process models are determined by the event-driven sub-system, describing both the bottom-up and the top-down influences. In this paper we discuss, how the recent improvements of our non-conventional modelling methodology might contribute to develop an intermediate solution, combining holistic and functionality based elements.

2. Applied methodology and investigated biosystem example

2.1. Methodology: Direct Computer Mapping (DCM) of process models

The various (e.g. biosystem) models contain more complex elements and structures, than the theoretically established, specific mathematical apparatuses. Moreover, the execution of the hybrid multiscale models is difficult, because the usual differential equation solvers are not prepared for the discrete events, while the continuous processes cannot be embedded into the discrete models, easily. In addition the continuous and discrete, as well as quantitative and qualitative parts overlap with each other in the low, elementary level. However, these complex structures and functionalities may be mapped directly onto the uniform elements of dynamic databases, associated with locally executable (e.g. declarative) programs, while the simulation may be organized by a general-purpose kernel algorithm.