Computational and Structural Biotechnology Journal 15 (2017) 396–402

ELSEVIER

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

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

Mini Review

Multi-level and hybrid modelling approaches for systems biology

R. Bardini, G. Politano, A. Benso, S. Di Carlo*

Politecnico di Torino, Department of Control and Computer Engineering, 10129 Torino, Italy

ARTICLE INFO

Article history:
Received 15 March 2017
Received in revised form 28 June 2017
Accepted 31 July 2017
Available online 10 August 2017

Keywords:
Systems biology
Computational biology
Multi-level models
Hybrid models
-omics
Multi-scale systems

ABSTRACT

During the last decades, high-throughput techniques allowed for the extraction of a huge amount of data from biological systems, unveiling more of their underling complexity. Biological systems encompass a wide range of space and time scales, functioning according to flexible hierarchies of mechanisms making an intertwined and dynamic interplay of regulations. This becomes particularly evident in processes such as ontogenesis, where regulative assets change according to process context and timing, making structural phenotype and architectural complexities emerge from a single cell, through local interactions. The information collected from biological systems are naturally organized according to the functional levels composing the system itself. In systems biology, biological information often comes from overlapping but different scientific domains, each one having its own way of representing phenomena under study. That is, the different parts of the system to be modelled may be described with different formalisms. For a model to have improved accuracy and capability for making a good knowledge base, it is good to comprise different system levels, suitably handling the relative formalisms. Models which are both multi-level and hybrid satisfy both these requirements, making a very useful tool in computational systems biology. This paper reviews some of the main contributions in this field.

Contents

1. Introduction

Systems biology considers biological entities as complex holistic structures whose behaviour cannot be reduced to the linear sum of the functions of their parts [1]. With the aim of gaining a deeper insight over biological complexity, computational modelling and simulation can support the understanding of experimental data, as well as the capability of generating and testing hypotheses about them [2]. However, given the huge complexity and peculiar features of these systems, it is necessary to carefully understand the specific modelling requirements they pose, in order to define what a good model for systems biology should look like.

In a complex biological structure, overall features emerge from local interactions among its sub-parts [3]. These interactions are in general favoured by the spatial proximity of the sub-parts. *Spatiality* is therefore one of the biological characteristics that must be taken into account when modelling biological systems [4]. More specifically, the probability of two elements to interact is a function of their spatial proximity and the stochasticity guiding such events must be explicitly taken into account in the modelling task [5].

Biological systems evolved different strategies to control the probability of interaction between biological components. One of them is called *compartmentalization* [6,7]. Biological systems are organized in compartments, and boundaries between compartments

http://dx.doi.org/10.1016/j.csbj.2017.07.005

* Corresponding author.

selectively regulate the passage of molecules, thus altering the probability density over space of molecular encounters. In a model, this must translate into the capability of expressing encapsulation and selective communication of each sub-part [8].

Spatial proximity between molecules not always translates into functional activations. The activation of selected functions, in fact, may require biochemical interactions between the molecules leading to structural changes able to alter their functional state. Structural features of biomolecules are encoded in the genome. Thus, the way such information is used determines the quality and quantity of actors and their interactions. The usage of genomic information is regulated at different levels and by different mechanisms, which are in flexible hierarchical relations. Such dynamic interplay of regulations is made of hierarchic relative relations that change according to the process context. This corresponds to the definition of epigenetic regulation in its broader sense: everything acting between a genotype and the corresponding possible phenotypes [9]. Biological models therefore require efficient ways to represent context-dependent and flexible hierarchies.

The modelling of biological systems should also comprise their quantitative aspects. Nevertheless, the way this is taken into account depends on the context. Some biological phenomena fit better with qualitative and discrete information. In other cases, biological quantities need to be represented with continuous quantities, for example referring to molecular concentrations. Therefore, a good model must be able to handle discrete and continuous variables as well as qualitative and quantitative information.

In the large variety of problems to be tackled with a systems biology modelling approach, ontogenetic processes are an example of how the presented modelling requirements are pushed to an extreme. Ontogeny takes the individual organism from the stage of fertilized egg to its fully developed form [10]. This involves a finely tuned and context-dependent processing of the spatiotemporal regulation of the genomic information. In fact, (almost) all cells in an organism share the same genome, yet they have different functional specializations and the overall system exhibits architectural and phenotype diversity. During development, cells undergo differentiation processes guided by their internal states as well as by extrinsic signals. Such signals come from other cells, which are in turn undergoing the same kind of regulations. These inter-cellular interactions can be mediated by concentration gradients over space: different relative positions between the sender and the receiver correspond to different concentration levels determining different results for the same signal. Depending on the context of the process (cellular micro-environment, developmental phase, cell types under analysis, specific regulative state of the cell, etc.) the different regulatory mechanisms involved in ontogenesis change their relative hierarchical relations. In turn, this means that sometimes the genetic regulation determines the future epigenetic state of the cell, other times it is the epigenetic state that determines the availability of the genomic information required to trigger the genetic regulations.

2. An introduction to hybrid and multi-level models

As discussed in the introduction, systems biology models in general must be able to handle different scales of representation, to model the system and its sub-parts into a complex hierarchical structure and to handle various types of information represented with different formalisms.

This review focuses on a particular class of models usually referred to as multi-level and hybrid models. Multi-level models describe a system at least at two different levels. Interactions are taking place within and between those levels [11]. Multi-level models allow for the explicit representation of “upward” and “downward” relations. Upward relations model the fact that the system is somehow constrained by the behaviour of its parts, but at the same time downward relations model the fact that the behaviour of each part is influenced by the behaviour of the system as a whole.

When considering multi-level models it is important to make an explicit distinction between the concept of scale and the concept of level [12]. More specifically, the concept of scale refers to a measurable dimension of the analysis of the considered phenomenon. This dimension can be spatial, temporal, and quantitative. The spatial dimension refers to the size of the entities involved in the phenomenon whereas the temporal dimension is related to the timing associated with the behaviours of these entities and their interactions. The quantitative dimension instead refers to the amount of entities involved in the phenomenon. Differently, the concept of level provides a way to locate the studied phenomenon and/or the entities involved in a phenomenon along the considered dimension of the analysis. A level usually corresponds to all the entities whose size and/or characteristic evolution time have the same or comparable orders of magnitude. For example, a system could be represented at the atomic, molecular, cell, organ, population level.

The concept of multi-level models can be coupled with the concept of hybrid models. According to Stephanou et al., “in its most general definition, a hybrid model corresponds to any interaction or coupling between two or more models that are not based on the same formalism” [13].

Based on this definition, we define models which are both multi-level and hybrid as representations supporting different formalisms and organized in levels encompassing multiple systems scales.

When building up a multi-level and hybrid model, besides choosing the interesting organizational levels, it is necessary to choose the formalisms to describe the different components in the overall model structure. In this sense, it can be useful to briefly revise the formalisms more often employed in modelling biological systems, so that their strengths and limitations can be taken into account when selecting hybrid combinations for the different organizational levels to be modelled. Fig. 1 summarizes the set of considered formalisms and their main characteristics. For a more detailed review of the modelling formalisms used in systems biology, see [14].

In general, biological systems models can be distinguished into mathematical and computational ones. “A computational model is a formal model whose primary semantics is operational; that is, the model prescribes a sequence of steps or instructions that can be executed by an abstract machine, which can be implemented on a real computer. A mathematical model is a formal model whose primary semantics is denotational; that is, the model describes by equations a relationship between quantities and how they change over time.” [15] However, this separation is not strict. Mathematical models can be simulated as well, with the only difference that the computational effort lies into the algorithm chosen to solve the model. One can get insights from a computational model by executing it, or by analyzing it by means of tools for model checking. Mathematical models can instead provide information through formal analysis, but they can be also simulated and solved.

Both mathematical and computational formalisms can be then categorized according to similar opposite features: they can be either qualitative or quantitative, discrete or continuous, deterministic or stochastic.

Usually, mathematical models are based on systems of equations. Difference equations are one of the preferred formalisms when modelling the system using discrete terms [16]. Instead, differential equations are among the preferred formalisms if the model is based on a representation of continuous biological quantities. Ordinary Differential Equations (ODE) are in general used whenever only the temporal aspects of the system are taken into account. Partial