

Chapter 14

Complexity of the BN and the PBN Models of GRNs and Mappings for Complexity Reduction

Ivan V. Ivanov
Texas A&M University, USA

ABSTRACT

Constructing computational models of genomic regulation faces several major challenges. While the advances in technology can help in obtaining more and better quality gene expression data, the complexity of the models that can be inferred from data is often high. This high complexity impedes the practical applications of such models, especially when one is interested in developing intervention strategies for disease control, for example, preventing tumor cells from entering a proliferative state. Thus, estimating the complexity of a model and designing strategies for complexity reduction become crucial in problems such as model selection, construction of tractable sub-network models, and control of the dynamical behavior of the model. In this chapter we discuss these issues in the setting of Boolean networks and probabilistic Boolean networks – two important classes of network models for genomic regulatory networks.

INTRODUCTION

One can think of a *Gene Regulatory Network (GRN)* as a network of relations among strands of DNA (genes) and the regulatory activities associated with those genes (Dougherty and Braga-Neto, 2006). This general definition allows for many mathematical (usually dynamical) systems to be called GRNs. The goodness of each such model is evaluated using several important criteria: the level of description of the biochemical reactions involved, complexity of the model, model parameter estimation, and the predictive power of the model. There have been many attempts to model the structure and dynamical behavior of GRNs, ranging from deterministic with discrete time space to fully stochastic with continuous time

DOI: 10.4018/978-1-60566-685-3.ch014

space. One can find a good review of such attempts in (de Jong, 2002). The so called central ‘dogma’ of molecular biology (Crick, 1970) implies that genes communicate via the proteins they encode. Both stages of protein production, transcription and translation, are controlled by a multitude of biochemical reactions, and are influenced by both internal and external to the cell factors. This perspective suggests that the expression of a given gene i , i.e. the quantity of either protein or messenger RNA, should be considered as a random function $X_i(t)$ of the cell’s internal and external environments. Thus, if one wants to study the dynamical behavior of a GRN, one must design a mathematical model for the gene-expression vector $\mathbf{X}(t) = (X_1(t), X_2(t), \dots, X_n(t))$ for the n genes that form the network. The stochastic differential equation model appears to provide the most detailed description of the dynamics of $\mathbf{X}(t)$. In principle, it could include all of the information about the biochemical processes involved in gene regulation. At the same time, the estimation of its parameters cannot be done without large amount of reliable time-series data. Thus, one is forced to take a more pragmatic approach and look for simpler models for the dynamics of the gene-expression vector. One of the most extreme simplifications is the *Boolean network* (BN) model, originally proposed by Kauffman (1969a). The BN model is based on the observation that during the regulation of its functional states the cell often exhibits switch-like behavior. Recent work using the NCI 60 Anti-Cancer Drug Screen has demonstrated that Boolean logic type interactions can be detected in gene expression data (Pal et al., 2005). While there are instances in gene regulation where the Boolean logic is the appropriate level of description of the interactions – for instance, when transcription factors have to form a complex that binds to the *cis*-regulatory DNA to activate transcription, one should keep in mind that discrete models cannot capture the details of the biochemical reactions involved in those processes. It is not the binary nature of the BN model that is its greatest weakness, one even more important deficiency is its determinism. Deterministic models, such as the BN, cannot represent the consequential perturbations due to external latent variables. In addition, the BN model cannot be used to represent biologically meaningful events, such as gene mutations. The stochastic extension of the BN model - *probabilistic Boolean network* (PBN), was introduced by Shmulevich et al. (2002b) in an attempt to account for those latent variables and gene perturbations while keeping the Boolean logic as the model for the gene-gene interactions. As a collection of BNs with a probability structure, the PBN model could be viewed as a minimal extension of the BN which allows for modeling of the stochastic nature of complex systems with lots of latent variables and random experimental effects. However, even such a minimal extension of the deterministic model exhibits high complexity which impedes its practical applications to model GRN of more than 20 genes. Hence, there is a need for constructing size reducing mappings that produce new and more tractable models that share some of the biologically meaningful properties of the larger-scale models. In addition, one needs to develop methods for complexity estimation of both the model and the mapping used to reduce its size.

In this chapter we use the BN and PBN models to illustrate how one can approach both problems: the complexity estimation of a model and the construction of size-reducing mappings for it. The term *complexity* is overloaded with many different meanings depending on the field of study. GRNs are composed of many parts that interact with each other and those interactions involve feedback loops and are stochastic from the observer’s point of view. Thus, GRN satisfy the definition of a complex system, and one of the aspects of complexity-reduction strategies in this setting is to find subsystems with less number of interacting genes and with simpler rules of interactions between them which preserve, to a degree, important properties of the whole system. This kind of complexity is the main focus of the discussion in the chapter. A different approach to the complexity problem is based on the observation that every GRN can be viewed as an information processing machine which opens the door for applications

16 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/complexity-pbn-models-grns-mappings/38242

Related Content

A Pragmatic Assessment of Approaches and Paradigms in Software Risk Management Frameworks

Alankrita Aggarwal, Kanwalvir Singh Dhindsa and P.K. Suri (2020). *International Journal of Natural Computing Research* (pp. 13-26).

www.irma-international.org/article/a-pragmatic-assessment-of-approaches-and-paradigms-in-software-risk-management-frameworks/241939

A Biologically Inspired Autonomous Robot Control Based on Behavioural Coordination in Evolutionary Robotics

José A. Fernández-León, Gerardo G. Acosta, Miguel A. Mayosky and Oscar C. Ibáñez (2009). *Advancing Artificial Intelligence through Biological Process Applications* (pp. 107-129).

www.irma-international.org/chapter/biologically-inspired-autonomous-robot-control/4975

A Characterization of Convex Fuzzy Mappings

Chih-Yuan Chen, Cheng-Pin Wang and Tetz C. Huang (2010). *International Journal of Artificial Life Research* (pp. 40-44).

www.irma-international.org/article/characterization-convex-fuzzy-mappings/46027

Emergent Structures

Eleonora Bilotta and Pietro Pantano (2010). *Cellular Automata and Complex Systems: Methods for Modeling Biological Phenomena* (pp. 83-113).

www.irma-international.org/chapter/emergent-structures/43218

Evolutionary Turing Machines: The Quest for Busy Beavers

Penousal Machado, Francisco B. Pereira, Jorge Tavares, Ernesto Costa and Amílcar Cardoso (2005). *Recent Developments in Biologically Inspired Computing* (pp. 9-40).

www.irma-international.org/chapter/evolutionary-turing-machines/28322