

Chapter 10

Modelling Gene Regulatory Networks Using Computational Intelligence Techniques

Ramesh Ram

Monash University, Australia

Madhu Chetty

Monash University, Australia

ABSTRACT

This chapter presents modelling gene regulatory networks (GRNs) using probabilistic causal model and the guided genetic algorithm. The problem of modelling is explained from both a biological and computational perspective. Further, a comprehensive methodology for developing a GRN model is presented where the application of computation intelligence (CI) techniques can be seen to be significantly important in each phase of modelling. An illustrative example of the causal model for GRN modelling is also included and applied to model the yeast cell cycle dataset. The results obtained are compared for providing biological relevance to the findings which thereby underpins the CI based modelling techniques.

INTRODUCTION

Biological processes and systems can be abstracted as multi-layered networks interacting with each other to create a complete biological system. Understanding the interactions of genes plays a vital role in the analysis of complex biological systems. The system level view of gene functions provided by gene regulatory networks (GRNs) is of tremendous importance in uncovering the underlying biological process of living organisms, providing new ideas for treating complex diseases, and for designing of new drugs. This chapter presents computational intelligence applications generally to bioinformatics problems and specifically to model networks of genetic regulation or gene regulatory networks (de Jong, 2002) (Someren et al, 2002) (Brazhnik et al, 2002) using gene expression data.

DOI: 10.4018/978-1-60566-685-3.ch010

Living beings are endowed with highly complex information storage and processing systems that are regulated in many different ways. The control of the body is carried out by large networks of regulatory genes, otherwise known as Gene Regulatory Networks (GRN). GRNs are collections of gene-gene regulatory relations in a genome that display relationships between gene activities. Increases in complexity of organisms do not bring an increase in the number of genes in the genome. For example, humans are believed to have about 20,000-25,000 genes (considerably lower than the original estimate), which is not dissimilar to the gene content for less complex organisms, such as the worm *Caenorhabditis elegans*, while a simple organism such as *Drosophila melanogaster*, also known as the fruit fly has about 14,000 genes. Therefore, the complexity may be due to a phenomena such as regulation of expression of genes in both temporal and spatial manners. A prerequisite for cellular behaviour is that the correct genes are expressed in the correct cell over correct time intervals and at correct expression levels. Regulatory networks specify how this gene expression or cellular behaviour is controlled. Over the past two decades, advances in molecular biology, DNA sequencing, and other high-throughput methods have resulted in a vast amount of bioinformatics data as shown in Fig1, including pathway information such as metabolic and regulatory pathways for various organisms. The rapid increase of this data for various organisms offers the possibility to perform analyses for single organisms (intra-species) as well as across different organisms (inter-species). However, the sheer quantity of data generated has exceeded the capacity of a researcher to extract useful information using traditional data analysis techniques. Since the high-throughput data acquisition technology for gene expression measurement known as biological Microarray technology emerged in the late 1990s, application of data mining, machine learning, and computational intelligence techniques to microarray data analysis has drawn attention of the bioinformatics community. Along with this, a significant amount of attention has been focused on modelling genetic regulatory networks from gene expression data. Microarrays allow the monitoring of expression levels of thousands of genes simultaneously and the data provide the basis to discover gene regulation networks, life evolution, and other important bio-problems. However, gene expression microarray data is characterized as massive, heterogeneous (with high dimensions), net character in nature, irregular sampling rate, having measurement errors (leading to noisy data) Hence, its analysis is beyond the ability of traditional analysis methods and decision supporting technologies. Due to the nature of data, previous deterministic models are not able to capture the time-varying dependencies between the different genes in the gene regulatory network. Researchers dealing with gene microarray data are faced with daunting quantities of data in which lies important hidden information, including transcription factor activity profiles.

There are two main objectives for modelling genetic regulatory networks (GRNs). The first is to be able to infer the regulatory network from data, in order to be able to understand the mechanisms behind them. The second is to be able to use the inferred networks in order to be able to predict the behaviour of actual networks for the purpose of diagnosing diseases and the development of drug targets and treatments. Needless to say, the system level view provided by gene networks of gene functions is of tremendous importance in uncovering the underlying biological process of living organisms, providing new ideas for treating complex diseases, and the design of new drugs. Inevitably, its research has become important in the recent times and the extensive research on this topic is timely.

Computational Intelligence (CI) is an area of computer science and engineering that deals with mimicking the intelligence observed from the natural behaviour in, for example, biology, insect societies (ants), neural sciences, immune systems. System Biology (Kitano, 2001), a sub group of computational biology, is a new and developing field of research in bioinformatics which deals with analysing large scale biological systems. It is an interdisciplinary research area combining the fields of (cell) biology and

20 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/modelling-gene-regulatory-networks-using/38238

Related Content

Some Words About Nature-Inspired Computing

Mohamed Arezki Mellal (2022). *Applications of Nature-Inspired Computing in Renewable Energy Systems* (pp. 1-9).

www.irma-international.org/chapter/some-words-about-nature-inspired-computing/294384

Data Analysis Tools for Neural Data Cognitive Computing: Neural-Inspired Algorithms

Shiddarth Srivastava and Rashmi Sharma (2023). *Exploring Future Opportunities of Brain-Inspired Artificial Intelligence* (pp. 139-156).

www.irma-international.org/chapter/data-analysis-tools-for-neural-data-cognitive-computing/320614

Evolutionary Computing to Examine Variation in Proteins with Evolution

Sujay Ray (2017). *Nature-Inspired Computing: Concepts, Methodologies, Tools, and Applications* (pp. 187-202).

www.irma-international.org/chapter/evolutionary-computing-to-examine-variation-in-proteins-with-evolution/161028

An Improved Nondominated Sorting Algorithm

André R. da Cruz (2012). *International Journal of Natural Computing Research* (pp. 20-42).

www.irma-international.org/article/an-improved-nondominated-sorting-algorithm/93622

Folding Theory for Fantastic Filters in BL-Algebras

Celestin Lele (2011). *International Journal of Artificial Life Research* (pp. 32-42).

www.irma-international.org/article/folding-theory-fantastic-filters-algebras/62071