A Phenomic Algorithm for Inference of Gene Networks Using S-Systems and Memetic Search

Rio G.L. D'Souza[1], K. Chandra Sekaran[2], and A. Kandasamy[2]

[1] St Joseph Engineering College, Mangalore, India
[2] National Institute of Technology Karnataka, Surathkal, Mangalore, India
{Rio,kchnitk}@ieee.org, kandy@nitk.ac.in

Abstract. In recent years, evolutionary methods have seen unprecedented success in elucidation of gene networks, especially from microarray data. We have implemented the Phenomic Algorithm which is an evolutionary method for inference of gene networks based on population dynamics. We have used S-systems to model gene interactions and applied memetic search to fine tune the parameters of the inferred networks. We have tested the novel algorithm on artificial gene expression datasets obtained from simulated gene networks. We have also compared the results to those obtained from two other similar algorithms. Results showed that the new method, which we call as Phenomic Algorithm with Memetic Search (PAMS), is an effective method for inference of gene networks.

Keywords: Microarray data analysis, Gene networks, Evolutionary algorithms, S-systems, Memetic search, Phenomic algorithms.

1 Introduction

Ever since the advent microarray technology scientists have been able to study thousands of genes at a time, and this has helped them to analyze the relationships between them. Most microarray experiments result in large datasets which need to be analyzed in order to understand the underlying relationships. There is vast potential for methods that can yield useful patterns from such large datasets without compromising the dimensionality [1]. Gene networks represent relationships between genes, based on observations of how the expression level of each gene affects the expression levels of the others [2]. Several researchers have used evolutionary methods [3] to analyze the relationships between thousands of genes. The Phenomic Algorithm [4], [5] is an approach based on population dynamics. We have implemented the proposed algorithm and validated it on artificial gene network datasets.

The rest of this paper is organized as follows: In Section 2, we provide a review of similar work done by others. We introduce the models and also the basis of the methods that we employ in Section 3; and in Section 4 we discuss the results of our experiments. This is followed by Section 5 which concludes the paper.

J. Suzuki and T. Nakano (Eds.): BIONETICS 2010, LNICST 87, pp. 229–237, 2012.
© Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2012

2 Related Work

The Inference of gene networks from the ever-growing mass of microarray data has become an important research activity in Systems Biology. Among the initial attempts, Somogyi et al. [6] developed a simple method which inferred Boolean networks. Several gene network reconstruction algorithms have been studied by Akutsu et al. [7] and D'haeseleer et al. [8]. While some of these methods infer only qualitative relationships between genes, others which infer quantitative relationships are limited by the scale of networks that they can deduce.

Reliable inference of gene networks is dependent on how closely the chosen model represents the real gene networks. One such model, which is nonlinear and dynamic is the S-System proposed by Savageau [9]. Several researchers [10] have used this model to reverse engineer gene networks. Recently, Ahmed, Song and Xing [11] have used a variant of S-System to construct graphical models for inferring time-varying gene regulatory networks. Most problems in this field can be viewed as some type of optimization and multiobjective evolutionary algorithms (MOEAs) have found remarkable success in reconstructing gene networks from expression data [12].

3 Models and Methods

3.1 S-System Model of Gene Networks

To establish that a change in the expression of gene B was caused by a change in the expression of gene A, it is necessary to show that a dependency exists between the two genes, whereby gene B is dependent on gene A. The Power-law formalism called S-System, which was proposed by Savageau [9] is a nonlinear and dynamic model which we used to capture the relationships between genes.

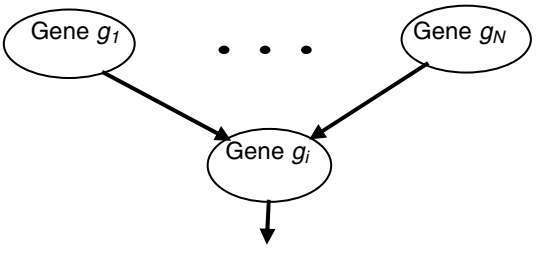

Fig. 1. Generalized gene network model

The behaviour of a cell can be abstracted by a gene regulatory network of N genes and other intermediate gene-products. Each gene g_i produces a certain amount of RNA x_i whenever it expresses. This causes a change in the concentration of this RNA over a time-period. This situation, shown in Fig. 1, can be represented in equation (1):

$$x(t+1)=h(x(t)), \qquad where\ x(t)=(x_1, x_2, \ldots, x_N)\ . \tag{1}$$