Artificial Intelligence formulated this projection for compatibility purposes from the original article published at Global Journals. However, this technology is currently in beta. *Therefore, kindly ignore odd layouts, missed formulae, text, tables, or figures.*

1	Generation of Genetic Networks from a Small Number of Gene
2	Expression Patterns under the Boolean Network Model
3	Jakir Hossain ¹ and Momtaz Begum^2
4	1 Dhaka University of Engineering and technology
5	Received: 29 October 2011 Accepted: 21 November 2011 Published: 3 December 2011
6	

7 Abstract

- ⁸ There are lots of work for inferring genetic network architectures from state transition tables
- ⁹ which correspond to time series of gene expression patterns, using the Boolean network model.
- ¹⁰ Results of those computational experiments suggested that a small number of state transition
- 11 (INPUT/OUTPUT) pairs are sufficient in order to infer the original Boolean network
- ¹² correctly. Tatsuya AKUTSU, Satoru MIYANO and Satoru KUHARA gave a mathematical
- ¹³ proof for this. So there is possibility to devise an algorithm to generate all consistent genetic
- ¹⁴ networks from a small number of gene expression patterns under the Boolean network model.
- 15

16 Index terms— Generation, consistent network.

17 **1 2011**

18 December nference of gene regulation mechanism from time series of gene expression patterns are getting more 19 important especially due to the invention of DNA microarray technology. Expression profiles of several thousands 20 of genes are now being produced for further analyses. Some methods have been proposed for the inference of

21 gene regulation mechanism from time series of gene expression patterns.

A statistical method is proposed by Arkin, Shen and Ross. They used correlation matrices to infer chemical reaction networks from time series of measured concentration of species. Though they treated chemical reaction networks, they also suggested that their method might be applied to genetic networks. However, it seems difficult to apply their method to the inference of large scale networks.

A metabolic pathway is suggested to be inferred by DeRisi, Iyer and Brown from gene expression patterns of Saccharomyces cerevisiae obtained by using DNA microarrays. A network model similar to the Boolean network model is constructed by Yuh, Bolouri and Davidson from time series of expression patterns I General Terms : Gene network, gene regulatory network, Genetic network, gene expression pattern, Boolean network, microarray data, Sum of product, Boolean algebra. Keywords : Generation, consistent network. relating to a sea urchin

- gene. But their inference methods are not systematic or automatic. Besides, some studies have been done on the
- inference of genetic networks from state transition data using the Boolean network .
 On the other hannd, Liang, Fuhrman and Somogyi proposed an algorithm named REVEAL for inference of

Boolean networks (corresponding to genetic networks) from state transition tables (corresponding to time series

of gene expression patterns). REVEAL used information theoretic principles in order to reduce the search space.

They made some computational experiments on REVEAL. The results suggested that only a small number of state transition pairs (100 pairs from 1015) were sufficient for inferring Boolean networks with 50 nodes (genes)

³⁸ whose in degree (the number of input nodes to a node) was bounded by 3.

³⁹ Tatsuya AKUTSU, Satoru MIYANO and Satoru KUHARA gave a mathematical proof for their observation.

- 40 We will extend their algorithm to identify genetic networks from gene expression patterns derived by gene 41 disruptions and gene over expressions using a Boolean network-like model. They proved mathematically a lower
- bound and an upper bound of the number of expression patterns required to identify the network correctly. In
- 43 this paper we will try to extend try to provide an algorithm to generate all the consistent genetic networks
- 44 from a small number of gene expression patterns under the Boolean network model. a) Genetic network A gene

- 45 regulatory network or genetic regulatory network (GRN) is a collection of DNA segments in a cell which interact
- 46 with each other (indirectly through their RNA and protein expression products) and with other substances in
- 47 the cell, thereby governing the rates at which genes in the network are transcribed into mRNA.

⁴⁸ 2 b) Boolean Network

- A Boolean network G (V;F) consists of a set V = {v 1 ,v 2 ,??,v n } of nodes representing genes and a list F =
- (f 1, f 2, ??, f n) of Boolean functions, where a Boolean function f i (v i1, v i2, ?, v ik) with inputs from specified
- nodes v i1 ,v i2 ?,v ik is assigned to each node v i . For a subset U ? ? E-mail : E-mail : E-mail : E-mail : Md. Rokon Uddin Υ c) Definition of Problem
- Let (I j ;O j) be a pair of expression patterns of {v 1 ,v 2 ,?..,v n }, where I j corresponds to the INPUT and O
- j corresponds to the OUTPUT. We call the pair (Ij;Oj) an example. We say that a node vi in a Boolean network
- ⁵⁵ G(V;F) is consistent with an example (Ij;Oj), if Oj(vi) = f i (I j (v i1 ,?I j (v ik)) holds. We say that a Boolean ⁵⁶ network G(V;F) is consistent with (I j ;O j) if all nodes are consistent with (Ij;Oj). There are 2 n minterms of ⁵⁷ n variables, since a variable in the minterm expression can be in either its direct or its complemented form-two
- 58 choices per n variables.
- It is apparent that minterm n gives a true value (i.e., 1) for just one combination of the input variables. If one is given a truth table of a logical function, it is possible to write the function as a "sum of products". This is a special form of disjunctive normal form. For example, if given the truth table for the arithmetic sum bit u of one bit position's logic of an adder circuit, as a function of x and y from the addends and the carry in, ci: of V, an expression pattern of U is a function from U to $\{0,1\}$. An expression pattern of V is also called a state of a Boolean network. That is, represents the states of nodes (genes), where each node is assumed to take either 0 (not-express) or 1 (express) as its state value.©
- Figure 1(a) : A genetic network represented by Boolean network Suppose we have n number of nodes (v 1, v 2,??..v n) and a pair of expression pattern (I,O), where I correspond to input pattern and O corresponds to output pattern. V 1 to v n are assumed to take either 0 (not-express) or 1 (express) as its state value. First we generate a Boolean network consistent with (I,O) and then eliminate nodes which aren't essential to generate more networks. The idea is as follows:-1. Repeat step 2 for each node v i \hat{N} ?" V.1 v 2 v 3 v ' 1 v ' 2 v ' 3 v I 1 I 2 I 3 1 0 0 0 1 0 0 1 1 0 0 1 0 1 1 1 0 0 O 1 O 2 O 3 G I 2 ' 1 v v = , 2 ' 2 v v = AND (NOT 3 v) = ' 3 v NOT
- 72 3 v d)

⁷³ 3 Generate Boolean equation and assign in ?? ??

- ? using Sum of Product (SOP) notation based on input and output expression patterns (I and O). As here are 3 genes we have got 3 Boolean expressions for them. They can be formulated according to previous discussion. They are listed below:-?? 1 ? = ?(3,5), ?? 2 ? = ?(0,5), ?? 3 ? = ?(0,3,6,7).
- ⁷⁷ Using the above example first we'll calculate the running time of the propose 4d algorithm.
- To generate an Boolean equation for any state first we have to check how many 1s it has in output pattern. It can have maximum 2 n 1s. So O(n*2 n) is the running time for generation of each equation. As there are n nodes and one variable is associated with each node so running time of step two of the algorithm is O(n 2 * 2 n).
- Again for each node, there can be maximum 2 n minterms and step 4 eliminates one minterm in each iteration. Now two minterms can be chosen in ? ?? 2 ? ways (let m = 2 n). If we reduce one minterm at a time then complexity of step 4 can be calculated as follows:-
- In first iteration there are m minterms. So we can choose two of them to reduce into one in ? ?? 2 ? ways. 85 In second iteration there are m-1 minterms. So we can choose two of them to reduce into one in ? ?? ?1 2 ? 86 ways and so on. Similarly at the end we should have two minterms and we can choose them in one way. So the 87 complexity will be On the other hand if we assign serial numbers to each row of the input expression patterns 88 from 0 to 2 n -1 then some of products can be expressed with summation notation (?). For example, for the 89 above table as in output column the row 1,2,4 and 7 contain 1, so the output ex[ression can be formulated as 90 91 ?? ?(?? ?2) 2 ? = ? ?? +1 3 ? = (?? +1)! 3!(?? ?2)! = ?? 3 ??? 6 = (2)92
- It is apparent that this notation can be simplified using the help of computer more easily and efficiently.

Observing that the rows that have an output of 1 are the 2nd, 3rd, 5th, and 8th, we can write u as a sum of minterms m 1, m 2, m 4, and m 7. If we wish to verify this: u(ci, x, y) = m 1 + m 2 + m 4 + m 7 = (ci' x' y) + (ci' x y') + (ci x' y') + (ci x y) evaluated for all 8 combinations of the three variables will match the table. network take binary values which are updated synchronously, whereas quantities of gene expressions in real cells are not binary and are changing continuously in time. However, owing to its simplicity, the proposed algorithm can be extended in various way. This algorithm can be extended to identify, count and enumerate all the consistent networks also. Finally, we believe that our theoretical results, along with the mathematical computations encourage the

101 attempts to discover the gene regulation mechanism from time series of gene expression patterns.

 $^{^{1}}$ © 2011 Global Journals Inc. (US)



Figure 1:

 $_1x_1,\ldots,x_n$

	Figure 2: Table 1 (
208 ₁₁	Figure 3:
²⁰⁰ ×2	Figure 4:
208 ₃₈	Figure 5:
3 ¹⁰⁰¹ 1	Figure 6: 3 .
3686	Figure 7: ??) 3 ?2 ?? $6 = 8$?? ?2 ?? 6

1

(a) : Minterms formulation from a Boolean network

Figure 8: Table 1

¹⁰² .1 EXPERIMENTAL OUTCOMES a) Performance Analysis

103 This paper improves the previous woks done by the authors worldwide.

- The improvements are listed below:-1) It generates only the consistent network. So overhead of checking inconsistent networks is eliminated. 2) Huge improvement is achieved in running time.
- 3) As all consistent networks are generated, so counting problem is solved simultaneously. 4) The in degree/out
 degree have no limit. A node (gene) can be expressed by any number of nodes(genes).
- 108

IV.

109 .2 CONCLUSION

- 110 We have proved mathematically that to generate all consistent Boolean networks it requires O(n*8 n) times.
- For that purpose, we proposed a simple algorithm. Of course, real biological systems are different from Boolean networks nodes in a Boolean
- 113 [REFERENCES REFERENCES REFERENCIAS], REFERENCES REFERENCIAS 17.
- [Arkin et al. ()] 'A test case of correlation metric construction of a reaction pathway from measurements'. A
 Arkin , P Shen , J Ross . Science 1997. 277 p. 1275.
- [Kearns and Vazirani ()] An Introduction to Computational Learning Theory, M J Kearns, U V Vazirani. 1994.
 The MIT Press.
- [Mcadams and Shapiro ()] 'Circuit simulation of genetic networks'. H H Mcadams , L Shapiro . Science 1995.
 269 p. 650.
- [Thomas et al. ()] 'Dynamical behaviour of biological regulatory networks -I. Biological role of feedback loops
 and practical use of the concept of the loop-characteristic state'. R Thomas , D Thie_Ry , M Kaufman .
 Bulletin of Mathematical Biology 1995. 57 p. 247.
- [Derisi et al. ()] 'Exploring the metabolic and genetic control of gene expression on a genomic scale'. J L Derisi
 , V R Lyer , P O Brown . *Science* 1997. 278 p. 680.
- 125 [Wu et al.] GeneNetwork: an interactive tool for reconstruction of genetic networks using microarray data, Chia-
- 126 Chin Wu , H
suan-Cheng Huang , H
sueh-Fen Juan , Shui-Tein Chen .
- [Bolouri and Davidson ()] 'Genomic Cis-regulatory logic: experimental and computational analysis of a sea
 urchin gene'. E H Bolouri , Davidson . *Science* 1998. 279 p. 1896.
- [Wuensche ()] 'Genomic regulation modeled as a network with basins of attraction'. A Wuensche . Paci_c
 Symposium on Biocomputing 1998. 3 p. 89.
- [Akutsu et al. ()] 'Identification of gene regulatory networks by strategic gene disruptions and gene over
 expressions'. T Akutsu, S Kuhara, O Maruyama, S Miyano. Proc. 9th ACM-IAM Symp. Discrete Algorithms,
 (9th ACM-IAM Symp. Discrete Algorithms) 1998. p. 695.
- [Tatsuya et al.] Identification Of Genetic networks From A Small Number Of Gene Expression Patterns Under
 The Boolean Network Model, Akutsu Tatsuya , Miyano Satoru , Kuhara Satoru .
- [Xu et al.] Inference of Genetic Regulatory Networks with Recurrent Neural Network Models Using Particle Swarm
 Optimization, Rui Xu, Donald C Wunsch, II, Ronald L Frank.
- [Kightley et al. ()] 'Inferring Gene Regulatory Networks from Raw Data: A Molecular Epistemics Approach'. D
 A Kightley , N Chandra , K Elliston . *Pacific Symposium on Biocomputing*, 2004. 9 p. .
- 140 [Knuth] Donald E Knuth . Concrete Mathematics, p. 14.
- ¹⁴¹ [Wen ()] 'Large-scale temporal gene expression mapping of central nervous system development'. X Wen . *Proc.* ¹⁴² Natl. Acad. Sci. USA, (Natl. Acad. Sci. USA) 1998. 95 p. 334.
- [Somogyi and Sniegoski ()] 'Modeling the complexity of genetic networks: Understanding multigene and
 pleiotropic regulation'. R Somogyi , C A Sniegoski . *Complexity* 1996. 1 p. 45.
- 145 [Liang et al. ()] 'REVEAL, a general reverse engineering algorithm for inference of genetic network architectures'.
- 146 S Liang , S Fuhrman , R Somogyi . *Paci_c Symposium on Biocomputing* 1998. 3 p. 18.
- 147 [Wikipedia] http://en.wikipedia.org/wiki/Genetic_network *Wikipedia*, (the free encyclopedia)