# CISC 889 Bioinformatics (Spring 2004)

Genetic networks and gene expression data

CISC889, S04, Lec20, Liao

#### Gene Networks

- **Definition:** A gene network is a set of molecular components, such as genes and proteins, and interactions between them that collectively carry out some cellular function. A genetic regulatory network refers to the network of controls that turn on/off gene transcription.
- **Motivation**: Using a known structure of such networks, it is sometimes possible to describe behavior of cellular processes, reveal their function and the role of specific genes and proteins

#### • Experiments

- DNA microarray : observe the expression of many genes simultaneously and monitor gene expression at the level of mRNA abundance.
- Protein chips: the rapid identification of proteins and their abundance is becoming possible through methods such as 2D polyacrylamide gel electrophoresis.
- 2-hybrid systems: identify protein-protein interactions
  - (Stan Fields' lab http://depts.washington.edu/sfields/)

CISC889, S04, Lec20, Liao

2

1





## **Genetic Network Models**

- **Linear Model**: expression level of a node in a network depends on linear combination of the expression levels of its neighbors.
- Boolean Model: The most promising technique to date is based on the view of gene systems as a logical network of nodes that influence each other's expression levels. It assumes only two distinct levels of expression - 0 and 1. According to this model a value of a node at the next step is boolean function of the values of its neighbors.
- Bayesian Model: attempts to give a more accurate model of network behavior, based on Bayesian probabilities for expression levels.

CISC889, S04, Lec20, Liao

5

















Step 1. Find expressio was itsel	l all pa on leve f force	airs of el of $x$ ed to a	rows " diffe high o	( <i>i, j</i> ) i rs, exc or low	n <b>E</b> in w cluding p value.	which the rows in which $a_n$				
E = Figu generate	<i>x<sub>0</sub></i> 1 - 1 1 1 1 <b>tre 2</b> : E: ed from	x, 1 - 1 1 the gene	$x_2$ 1 0 0 - 1 expressi etic netv	$x_{j}$ 0 1 0 1 +	$\begin{vmatrix} p_{\theta} \\ p_{1} \\ p_{2} \\ p_{3} \\ p_{4} \end{vmatrix}$ rix fig. 1.	For x3, we find: (p0,p1), (p0, p3), (p1,p2), (p2,p3)				
CISC889, S04, Lec20, Liao 14										





## Chooser

For L hypothetical equiprobable networks generated by the predictor, choose perturbation p that would best discriminate between the L networks, by maximizing entropy  $H_p$  as defined below.

$$H_p = -\sum_{s=1}^{S} (l_s/L) \log_2 (l_s/L)$$

where  $l_s$  is the number of networks giving the state s Note:  $(1 \le s \le S)$ , and  $(1 \le S \le L)$ 

CISC889, S04, Lec20, Liao

17

### Result and Evaluation

- Evaluation of Predictor
- construct a target network *T*: size = *N*, and maximum in-degree = *k* (where the in-degree of a node is its number of incoming edges)
- *sensitivity* is defined as the percentage of edges in the target network that were also present in the inferred network, and *specificity* is defined as the percentage of edges in the inferred network that were also present in the target network.

A	B	С	D	Ε	F	G	H	1	J
N	k	Total Sim. Edges	Num. Inferred Networks	Total Inferred Edges	Num. Shared Edges	Sens- itivity	Spec- ificity	Num. Nodes w/ 1 Soln.	CPU Time (sec)
5	2	4 (0.1)	1 (.02)	3 (0.1)	3 (0.1)	77%	99%	5 (0.0)	0.1 (0.0)
10	2	12 (0.1)	60 (50)	9 (0.1)	9 (0.1)	71%	95%	9 (0.1)	0.1 (0.0)
20	2	27 (0.2)	3×10 <sup>7</sup> (10 <sup>7</sup> )	21 (0.2)	19 (0.1)	71%	92%	18 (0.1)	0.2 (0.0)
50	2	72 (0.2)	$1 \times 10^{12} (10^{12})$	57 (0.3)	51 (0.3)	71%	90%	45 (0.2)	0.8 (0.0)
100	2	146 (0.7)	3×10 <sup>26</sup> (10 <sup>26</sup> )	119 (0.9)	104 (0.7)	70%	88%	89 (0.5)	6.6 (0.3)
20	4	44 (0.3)	2×10 <sup>6</sup> (10 <sup>6</sup> )	28 (0.3)	23 (0.2)	51%	84%	16 (0.1)	0.2 (0.0)
20	6	57 (0.5)	2×10 <sup>7</sup> (10 <sup>7</sup> )	33 (0.3)	27 (0.2)	42%	82%	14 (0.2)	0.2 (0.0)
20	8	69 (0.7)	$9 \times 10^{7} (10^{8})$	38 (0.4)	31 (0.3)	35%	82%	13(0.2)	0.2 (0.0)

CISC889, S04, Lec20, Liao

18







