Modeling the Dynamics of Gene Networks

Michelle Girvan

with:
Andrew Pomerance
Ed Ott
Wolfgang Losert
Lou Staudt (NCI/NIH)
Overview

• **The goal:** To gain insights into the complex process of gene regulation.

• **The approach:** Considering a simple model of genetic control, we explore the effects of network topology.

• **The application:** We hypothesize that a dynamical instability in the gene network may be a causal mechanism contributing to the occurrence of some cancers.
The process of gene regulation

**Transcriptional regulation:** Proteins called transcription factors bind to specific sequences of the DNA to help or hinder the transcription of individual genes.
The Result:
A complex web of interactions

Figure taken from http://rsif.royalsocietypublishing.org/content/5/Suppl_1/S85.full
Input/output regulatory relationships between genes are observed to be strongly sigmoidal and well approximated by step functions.
Modeling Gene Networks: The Boolean Approach

Kauffman’s \( N-K \) model:
- \( N \) Genes on or off
- Each gene has exactly \( K \) inputs, which are randomly chosen
- Discrete updates
- Evolves by a random update function at each node

Our work:
- Focuses on stability of these systems in response to small perturbations
- Explores the effect of network topology on stability
Local update rules: An example

- **Table:**
<table>
<thead>
<tr>
<th>Gene 1</th>
<th>Gene 2</th>
<th>State of gene 3 at t+1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

- **Diagram:**
  - Node with 2 inputs

Gene 1:
- State 0:
  - Inputs: 0, 0
  - Output: 0
- State 1:
  - Inputs: 0, 1
  - Output: 0
- State 0:
  - Inputs: 0, 0
  - Output: 0
- State 1:
  - Inputs: 1, 0
  - Output: 1
- State 1:
  - Inputs: 1, 1
  - Output: 0
Describing the Boolean Network

• Network topology:

\[ A_{ij} = \begin{cases} 
1 & \text{if link from } j \rightarrow i \\
0 & \text{otherwise} 
\end{cases} \]

• Update functions:
  ‣ Output rows randomly filled in
  ‣ Bias - probability of a 0 appearing in the output row
Properties of the Boolean Model

- **Finiteness**: Eventually the system must return to a previously visited state.

- **Determinism**: Upon this return, the subsequent dynamics will be the same as for the previous visit.

- **Attractors**: Every initial condition produces a trajectory that eventually goes to a periodic orbit, called the “attractor” of that initial condition, and different initial conditions can go to different periodic orbit attractors.
Significance of the attractors

- The attractors may represent a specific pattern of protein expression that define a cell’s character.

- In single celled organisms this could be different cell states: growing, dividing, starving, etc.

- In multicellular organisms these could correspond to different cell types.
Thinking about stability

- The state of this system is described by the vector $\underline{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_N)$.

- We look at the Hamming distance, $H(\underline{\sigma}, \tilde{\underline{\sigma}})$ (defined as the number of nodes that differ between the two states).

- Instability: If at $t = 0$ $H \ll N$, then $H$ initially grows in time.
“Chaotic” and stable dynamics for different networks
Why study stability?

• Kauffman hypothesized that gene networks exist at the critical transition between stable and chaotic regimes.

• We want to understand if small perturbations like chemical fluctuations cause a cell to change state.

• We hypothesize that cancer may result from a dynamical instability in the gene network.
Previous work on stability

• Derrida & Pomeau (1986):
  ‣ Uniform number of inputs to every node
  ‣ Uniform probability $p$ of 0 entries in the truth tables for all nodes
  ‣ Annealed network and annealed truth table approximation used to find stability criteria

• Aldana and Cluzel (2003)
  ‣ Stability criteria determined for networks with in degree distribution $p(k_{\text{in}})$ and otherwise random structures
Motivation for our work

- Real networks are far from the idealized models studied previously
- We would like to be able to analyze any fixed network, and we are interested in the effects of:
  - Assortativity
  - Community structure
  - Heterogeneous $p_i$
  - Network motifs
Finding the stability criterion for arbitrary network topologies and sensitivity distributions

• We consider a semi-annealed approximation in which the network is fixed and the output entries of the truth tables are randomized at every step, subject to the $p_i$.

• We perform numerical tests with frozen truth tables to test the validity of this approximation.
Semi-annealed analysis

• Consider two state vectors, $\sigma(t)$ and $\tilde{\sigma}(t)$, that have evolved from slightly different initial conditions

• Let $y_i(t) = \text{the probability that } \sigma_i(t) \text{ and } \tilde{\sigma}_i(t) \text{ differ}$

• Let $q_i = \text{the probability that } \sigma_i(t) \text{ and } \tilde{\sigma}_i(t) \text{ differ, given a difference in the states of the inputs to } i \text{ at time } t - 1$

\[ q_i = 1 - [p_i^2 + (1 - p_i)^2] = 2p_i(1 - p_i) \]
Update equation for $y_i(t)$

\[ y_i(t) = q_i \left\{ 1 - \prod_{j, A_{ij}=1} \left[ 1 - y_j(t-1) \right] \right\} \]

Probability that the inputs at $t-1$ to $i$ are not all the same

Probability that the input from node $j$ is the same

Perturb around $\sigma = \tilde{\sigma} \ (y_i \ll 1)$, linearization gives:

\[ y_i(t) \approx q_i \sum_{j=1}^{N} A_{ij} y_j(t-1) = \sum_{j=1}^{N} Q_{ij} y_j(t-1) \]

where the $Q_{ij} = q_i A_{ij}$ are the elements of a modified adjacency matrix
Stability Criterion

\[ y(t) = Q y(t - 1) \]

\( \lambda_Q \) is the largest eigenvalue of \( Q \), which, according to the Perron-Frobenius theorem is real and positive (\( Q_{ij} \geq 0 \)).

Stability Conditions:

- If \( \lambda_Q < 1 \): stable
- If \( \lambda_Q > 1 \): unstable
- If \( \lambda_Q = 1 \): "edge of chaos"
Numerical tests

We numerically test the predictions of

- $\lambda_Q$ stability criterion

- Saturated normalized Hamming distance between $\sigma$ and $\tilde{\sigma}$:

$$\bar{y} = \lim_{t \to \infty} \frac{1}{N} \sum_{i} y_i(t)$$
Nodes have sensitivity drawn from distribution centered around $q_0$

$y$ is the average saturated distance between two initially close states that have been evolved
**Assortativity**

**Assortativity:** highly connected nodes tend to connect preferentially to other highly connected nodes, tends to increase eigenvalue
Stability and Cancer

- Data from tumor dissections show that nearby cells have vastly different gene expression profiles.
- Could these fluctuations imply a breakdown of genetic control due to dynamical instability?
- What kind of data do we need to answer these questions?
Elucidating the network and the sensitivities from data

- **Network**: Undirected links can be inferred from data by looking at co-expression patterns across a range of perturbation experiments.

- **Sensitivities** can be determined from clinical expression data.
A possible strategy for cancer treatment

Our cancer/stability hypothesis combined with our stability analysis suggests a possible cancer therapy strategy: design drugs that target those genes or links whose disabling would most reduce $\lambda_Q$. 
Summary

- We studied the stability of simple Boolean models of genetic control
  - Considered the case of non-uniform sensitivity
  - Allowed for arbitrary network topology, and considered how specific topological features influence the stability.
- We hypothesize that this stability may be important in cancer.
- We have ideas about how this hypothesis could be tested with real data.

Reference
Future work

- Confine the truth tables to logic functions that better reflect true gene interactions
- Consider three-state instead of Boolean systems
- Explore the relationship between attractors and differentiated cells