Qualitative modeling of biological systems

• The functional form of regulatory relationships and kinetic parameters are often unknown
• Increasing evidence for
  • robustness to changes in kinetic parameters.
  • bistability
Hypothesis: the kinetic details of individual interactions are less important than the organization of the regulatory network
Network topology: identity of nodes, edge attributes (direction, sign)

Topological features such as path redundancy and feedback loops constrain network dynamics.
First step - pseudo-dynamics: propagation of reactions in chemical (interaction) space, starting from a source (signal)

Qualitative models are most closely based on the network topology.
Boolean modeling of biological systems

Main assumption: components have two main states:
Expressed or not expressed, active or inactive, open or closed (ion channel), high or low level. Denote these states by ON (1) or OFF (0).

The changes in state are given by discrete (logical) rules. The future state of a regulated node (the output) depends on the current state of its regulators (inputs), which may or may not include its own current state.

E.g., if transcription factor is active, gene will be transcribed, gene will be expressed in the next time step.

Boole logic: based on the operators NOT, AND, OR
Can be defined based on set intersection and union, or input-output relations (gates, truth tables)
Truth tables for Boolean operators

<table>
<thead>
<tr>
<th>NOT</th>
<th>AND</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>In</strong></td>
<td><strong>Out</strong></td>
<td><strong>In1</strong></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Out = NOT In

Out = In1 AND In2

Out = In1 OR In2
Out = NOT In

Out = In1 AND In2

Out = In1 OR In2

Ex. 1 Give examples for the realization of these Boolean rules in a gene regulatory network.

Ex. 2 Consider a transcription event activated by a transcription factor. Compare the continuous and Boolean description of this process.
**Dose-response curves and Boolean switches**

- $X$ – mRNA
- $Y$ – transcriptional activator

If $\nu$ is large, the dose-response curve becomes a switch

If $Y > K_Y$ \( \frac{dX}{dt} > 0 \)
If $Y < K_Y$ \( \frac{dX}{dt} < 0 \)

The activation threshold is $K_Y$
If activation is weak, protein can decay.

Boolean simplification:
$X^* = Y$
Activation:
If $Y=\text{ON}$ $X^* = \text{ON}$
Decay:
If $Y=\text{OFF}$ $X^* = \text{OFF}$
Boolean models of regulatory networks

• The directed edges in the network indicate regulator – target pairs.

• Assume that the state of each node can be 0 or 1.

• The rule giving the new state of each node is determined by a Boolean function of the states of the nodes that regulate it.

• Synchronous model: the state of each node is updated at multiples of a common timestep.
  • Asynchronous model: the state of each node is updated individually.

• The state of the whole network changes in time.
Ex. 3

Construct a network of three nodes, such that their in-degree is one or two. Associate a Boolean rule to each node. Assume that each node’s state changes at the same time (synchronous update)
Start with an initial state, and update the state of the nodes 10 times.

What is happening to the state of the network?

Start from a different initial state. Will the final behavior be the same?

How many different final states/behaviors can the network have?
Example

A B C D
0 0 0 0
0 0 0 1
.... ... ... ...
1 1 1 1

States

Transitions

<table>
<thead>
<tr>
<th>l1</th>
<th>l2</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 1 1 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 0 1 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 1 0 0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Concepts in Boolean network dynamics

**Attractor** – a set of states that repeats itself in a fixed sequence can be **periodic** or a **fixed point**

Fixed point: Future State = Current State = Previous State

All states lead to or are part of an attractor.

**Basin of attraction** – all states leading to a given attractor

In a network of N nodes the maximum possible length of a periodic attractor is the total number of states, $2^N$

In practice, the period (length) of the attractor is much shorter than this maximum. Cause: many nodes become **frozen**, due partly to **canalizing functions**.
Andy Wuenche, www.ddlab.com
**Canalizing** (forcing) functions:
At least one of the inputs has the property that the output is fixed if this input has one particular value.

e.g. a AND b is canalizing because a=0 implies a AND b = 0

Ex. 4

How many two-input Boolean functions are there?

How many of them are canalizing?

Ex. 5

How would you determine the fixed points of a Boolean network?
Boolean modeling of gene regulatory networks

Cell differentiation is based on differential gene expression

Genes regulate each other’s expression

Stuart Kauffman ~ 1965

Ideas:
- genes can be modeled by on-off switches
- the structure of the gene regulatory networks is unknown
- the regulatory functions are unknown
- network states correspond to cell types
The Kauffman (NK) model

- Construct a network where each node’s in-degree is $K$.
- Assume that the state of each node can be 0 or 1.
- The state of each node is updated at each timestep.
- The rule giving the new state of each node is determined by a random Boolean function of the states of its regulators.
- Find the attractors of the network states. The number of attractors corresponds to the number of possible cell types.

How does the number and type of attractors change with $N$ and $K$?
Attractors in Kauffman networks

- For $K=1$, networks are frozen
  - median number of attractors is close to $2^N$
  - median cycle length close to 1

- For $K>5$, networks are chaotic
  - few attractors
  - median cycle length close to $2^N$

- For $K=2$, interesting level of order
  - median number and length of attractors both scale as $N^\alpha$

This is fairly similar with the number of cell types in different organisms.
Stability of Kauffman networks

What is the effect of a “mutation”, changing the state of a randomly selected node?

If the final number of changed nodes is small – frozen network

Percolating changes – chaotic network

The threshold between order and chaos is K=2

One can bias the Boolean functions so there are more of 0s or 1s.

Then the threshold varies with the bias Q as

\[ K_c = \frac{1}{2Q(1-Q)} \]

Ordered behavior for k<K_c
Does the threshold behavior apply to non-regular networks?

Order: \[ 2Q(1 - Q)K < 1 \]

This relation is maintained if the underlying network is ER with \( <k_{\text{in}}> = K \)

How does this compare with the threshold of a large connected component?
Does the threshold behavior apply to non-regular networks?

For scale-free networks with

\[ P(k) = \left[ Z(\gamma)k^\gamma \right]^{-1} \]

the condition becomes

\[ 2Q(1-Q) \frac{Z(\gamma-1)}{Z(\gamma)} < 1 \]

Scale-free networks with \( \gamma > 2.5 \) are robust to random perturbations.

M. Aldana, P. Cluzel, PNAS 100, 8711 (2003)
As we find out more about gene regulatory networks, it is not necessary to assume random topologies and regulatory functions anymore.

It is still interesting to see how successful an ON/OFF framework and Boolean logic can be (as opposed to chemical kinetics-based models).

Example: Boolean modeling of the segment polarity gene network


Drosophila segmentation is governed by a cascade of genes.
The role of the segment polarity genes

- The segment polarity genes are initiated by the pair-rule genes.
- Several segment polarity genes are expressed (active) in stripes that are repeated in every fourth cell.
- These genes interact via a complex regulatory network.
- The expression pattern of the segment polarity genes is maintained for 3 hours.
- The parasegment borders appear between the cells expressing the two most important segment polarity genes, *engrailed* and *wingless*. 
Segment polarity genes

<table>
<thead>
<tr>
<th>Genes</th>
<th>Proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>wingless (wg)</em></td>
<td>Wingless protein (WG) - secreted</td>
</tr>
<tr>
<td><em>hedgehog (hh)</em></td>
<td>Hedgehog protein (HH) - secreted</td>
</tr>
<tr>
<td><em>engrailed (en)</em></td>
<td>Engrailed protein (EN) - transcription factor</td>
</tr>
<tr>
<td><em>patched (ptc)</em></td>
<td>Patched protein (PTC) - receptor</td>
</tr>
<tr>
<td><em>smoothened (smo)</em></td>
<td>Smoothened protein (SMO) - receptor</td>
</tr>
<tr>
<td><em>sloppy paired (slp)</em></td>
<td>Sloppy paired protein (SLP) - transcription factor</td>
</tr>
<tr>
<td><em>cubitus interruptus (ci)</em></td>
<td>Cubitus interruptus protein (CI)</td>
</tr>
<tr>
<td></td>
<td>Cubitus activator (CIA) - transcription factor</td>
</tr>
<tr>
<td></td>
<td>Cubitus repressor (CIR) - transcription factor</td>
</tr>
</tbody>
</table>

Gene products form a network that maintains a gene expression pattern initiated in an earlier stage.
Evolution of gene expression patterns

**en**  
early stages  
2:50 h

**hh**  
pre-pattern  
3:00-3:30 h

**wg**  
stable pattern  
4:20-7:20 h

**ci**  
3:30 h

**en**  

**ptc**
**Wild type, stable gene patterns**

- *en* is expressed in the anterior part of the parasegment.
- *wg* is expressed in the posterior part of the parasegment.
- Parasegmental grooves form between the *wg* and *en* stripes.

- Two *ptc* stripes in each parasegment.
- *ci* pattern is complementary to that of *en*. 
Reconstructing the topology of the segment polarity network

mRNA

PROTEIN

PROT COMPL

repression

translation, activation, modification

cell

neighbor cell
Synchronous Boolean model

- Transcripts and proteins are either **ON** (1) or **OFF** (0).
- The expression of a node at timestep $t$ is given by a logical rule of the expression of its effectors at time $t-1$.
- Transcription depends on transcription factors; repressors are dominant.
- Translation depends on the presence of the transcript.
- Transcripts and most proteins decay if not produced.
- **Transcription, translation, mRNA/protein decay on the same timescale, protein binding faster.**
Rules for transcription and translation

\[ e_{n_i}^{t+1} = (W_{G_i}^{t-1} \text{ or } W_{G_i}^{t+1}) \text{ and not } S_{LP_i}^t \]

\[ h_{h_i}^{t+1} = E_{N_i}^t \text{ and not } C_{IR_i}^t \]

\[ p_{tc_i}^{t+1} = C_{IA_i}^t \text{ and not } E_{N_i}^t \text{ and not } C_{IR_i}^t \]

\[ c_{i_i}^{t+1} = \text{not } E_{N_i}^t \]

\[ E_{N_i}^{t+1} = e_{n_i}^t \]

\[ W_{G_i}^{t+1} = w_{g_i}^t \]

\[ C_{I_i}^{t+1} = c_{i_i}^t \]

\[ H_{H_i}^{t+1} = h_{h_i}^t \]
wg, PTC and SLP are more stable than other proteins

\[
wg_{i+1}^t = (CIA_i^t \text{ and } SLP_i^t \text{ and not } CIR_i^t) \text{ or } [wg_i^t \text{ and } (CIA_i^t \text{ or } SLP_i^t) \text{ and not } CIR_i^t]
\]

Either of the activators can counter mRNA decay.

\[
PTC_{i+1}^t = ptc_i^t \text{ or } (PTC_i^t \text{ and not } HH_{i-1}^t \text{ and not } HH_{i+1}^t)
\]

Free PTC does not decay.

\[
SLP_{i+1}^t = SLP_i^t
\]

SLP is a source in the segment polarity network.
Rules for post-translational processes

\[ PH_i^t = PTC_i^t \text{ and } (HH_{i-1}^t \text{ or } HH_{i+1}^t) \]

\[ SMO_i^t = \neg PTC_i^t \text{ or } HH_{i-1}^t \text{ or } HH_{i+1}^t \]

\[ SMO_i^{t+\varepsilon} \]

\[ CIA_i^{t+1} = CI_i^t \text{ and } (SMO_i^t \text{ or } hh_{i-1}^t \text{ or } hh_{i+1}^t) \]

\[ CIR_i^{t+1} = CI_i^t \text{ and } \neg SMO_i^t \text{ and } \neg hh_{i-1}^t \text{ and } \neg hh_{i+1}^t \]
**Ingredients of the model**

**Components:** mRNAs and proteins  
**State:** expressed or not in a certain cell  
**Expression pattern:**

Initial state - updating rules - steady state  
The updating rules are determined by the network of interactions.
Start the model from an initial state giving the prepattern of all nodes

Wild type initial state: \( wg \) in the last cell of the parasegment, \( en/hh \) in the first cell of the parasegment, \( ptc \) and \( ci \) complementary to \( en \), no proteins.
The model reproduces the wild type steady state

The net effect of the interactions is enough to capture the functioning of the network!
What happens if the network is perturbed?

The most severe perturbation is caused by gene mutations. To model a null mutation, we assume that the mRNA is kept OFF, thus the protein cannot be translated. The effects of the mutation propagate throughout the network.

The model correctly captures all null mutant experiments.
**wg, en or hh mutations are lethal**

No *wg*, *en* and *hh* stripes, no segmentation, regardless of initial state.
\textit{ptc} mutation broadens the stripes

The \textit{wg}, \textit{en} and \textit{hh} stripes broaden, regardless of initial state.
ci mutation can preserve the prepattern

The effect of ci mutation depends on the initial state. For wild type prepattern, the wg, en, hh stripes remain.
Model correctly reproduces experimental results on knock-out mutants

Tabata, Eaton, Kornberg, Genes & Development 6, 2635 (1992)
Gallet et al., Development 127, 5509 (2000)
There are two other frequently occurring steady states

Broader initiation of \( wg, \ en \) or \( hh \) leads to broad stripes. Absence of \( wg \) leads to a state with no segmentation.

Gallet et al., Development 127, 5509 (2000)
How many initial states lead to the wild type steady state?

<table>
<thead>
<tr>
<th>Minimal prepattern</th>
<th>Steady state</th>
</tr>
</thead>
<tbody>
<tr>
<td>$wg$</td>
<td>$wg$</td>
</tr>
<tr>
<td>$WG$</td>
<td>$WG$</td>
</tr>
<tr>
<td>$en$</td>
<td>$en$</td>
</tr>
<tr>
<td>$EN$</td>
<td>$EN$</td>
</tr>
<tr>
<td>$hh$</td>
<td>$hh$</td>
</tr>
<tr>
<td>$HH$</td>
<td>$HH$</td>
</tr>
<tr>
<td>$ptc$</td>
<td>$ptc$</td>
</tr>
<tr>
<td>$PTC$</td>
<td>$PTC$</td>
</tr>
<tr>
<td>$PH$</td>
<td>$PH$</td>
</tr>
<tr>
<td>$SMO$</td>
<td>$SMO$</td>
</tr>
<tr>
<td>$ci$</td>
<td>$ci$</td>
</tr>
<tr>
<td>$CI$</td>
<td>$CI$</td>
</tr>
<tr>
<td>$CIA$</td>
<td>$CIA$</td>
</tr>
<tr>
<td>$CIR$</td>
<td>$CIR$</td>
</tr>
</tbody>
</table>

3 3 4 4 4 4 8 8 16 16 16 16 8 8 16

Total number of wild-type inducing prepatterns: $6 \times 10^{11} = 8 \times 10^{-6} N_i$
The steady states can be determined analytically

- In the stable state \( x_{i}^{t+1} = x_{i}^{t} \)
- Use the fact that \( SLP_1 = SLP_2 = 0 \) and \( SLP_3 = SLP_4 = 1 \)
- The set of equations reduces to:

\[
\begin{align*}
wg_1 &= wg_1 \text{ and not } wg_2 \text{ and not } wg_4 \\
wg_2 &= wg_2 \text{ and not } wg_1 \text{ and not } wg_3 \\
wg_3 &= wg_1 \text{ or } wg_3 \\
wg_4 &= wg_2 \text{ or } wg_4
\end{align*}
\]

\[
\begin{align*}
PTC_1 &= (\text{not } wg_2 \text{ and not } wg_4) \text{ or } (PTC_1 \text{ and not } wg_1 \text{ and not } wg_3) \\
PTC_2 &= (\text{not } wg_1 \text{ and not } wg_3) \text{ or } (PTC_2 \text{ and not } wg_2 \text{ and not } wg_4) \\
PTC_3 &= PTC_4 = 1
\end{align*}
\]

There are no limit cycles in this network.
Possible stable patterns

The latter states have very small probability.