Programming today is a race between software engineers striving to build bigger and better idiot-proof programs, and the Universe trying to produce bigger and better idiots. So far, the Universe is winning. - Unknown
Biological Networks

Questions:
• Which are the most common motifs among biological networks?
• What are the information processing functions of these motifs?
• What is the difference between a sub graph and a motif?
• How can we detect motifs?
Revision: Solving Differential Equations

- How do we solve $dY / dX = AY$?
- How about $dY / dX = B - AY$?
- Can you plot $Y$ Vs $X$ in each of these cases?
Sub graph Patterns

- Let us consider all possible patterns which can appear in directed networks with
  - size = 3
  - size = 4
  - Etc.

- Such combinations are called ‘sub graph patterns’ - all of them may not occur in a given network.

- Number of sub graph patterns increases if we consider the ‘activating / inhibiting’ nature of a directed link.
Three Node sub graph patterns
Four Node Sub graph patterns
Network Motifs

What are “Network Motifs”? Network Motifs are defined as patterns of interconnections that recur in many different parts of a network at frequencies much higher than those found in randomized networks.

Why do we need them? To help us understand how biological networks work.

Exact forecasting of operation and reaction in the network under given situations.
The concept of “network motifs” was first proposed by Uri Alon’s group:
Schematic View of Network Motif Detection

(A) Real network

(B) Randomized networks

Motif:
Detecting Network Motifs: 
Z scores

- Let us consider N number of random networks with the same size etc.
- What is the average number of a given sub graph pattern in the random networks?
- What is the standard deviation?
- The Z score of a sub graph pattern can be calculated as:

\[ Z_i = \frac{N_i^{real} - < N_i^{rand} >}{\text{std}(N_i^{rand})} \]

- If |z| > 3, then the sub graph pattern can be considered a motif.
Detecting Network Motifs

- Edges easily lost/added
- Compare real networks to randomized networks
- Patterns that occur more often in real networks = Network motifs
Detecting Network Motifs (Cont.)

- N nodes
- possible pairs of nodes: \([N(N-1)] + N = N^2\)
- edge position is occupied: \(p = E / N^2\)
Examples for motifs

- **FeedForward Loop**
  Found in neural networks.
  It seems to be used to neutralize “Biological Noise”. That is, it controls pulses.

- **Single-Input Module**
  Implemented in gene control networks
Examples for motifs

- Parallel paths

Found in neural networks, food webs etc. (and not so much in gene networks)
Biological Network motifs

- Autoregulation (AR)
- Feed Forward Loops (FFL)
- Regulating and Regulated Feedback Loops (RFL)
- Cascade
- Single Input Model (SIM)
- Dense Overlapping Regulon (DOR)
- BiFan
- Diamond

July 2013
N. H. N. D. de Silva
Autoregulation

AS A NETWORK MOTIF
Autoregulation

- Regulation of a gene by its own gene product
- How does it look in the graph?

- E. coli network:
  - 40 self edges
  - 34 repressors
  - 6 activators
Cont.) (Autoregulation)

- **Probability for self edge:**  \( P_{self} = 1/N \)

- **Expected number of self edges:**  
  \[ <N_{self}^{rand}> \sim E \times P_{self} \sim E/N \]

- **Standard deviation:**  \( \sigma_{rand} \sim \sqrt{E/N} \)
Cont.) (Autoregulation)

- Number of self edges:
  - Random network: $\langle N_{self} \rangle_{\text{rand}} \sim 1.2, \quad \sigma_{\text{rand}} \sim \sqrt{1.2} \sim 1.1$
  - E. coli network: 40

- Conclusion: Self edges are a network motif

- But... why?
THE FEED FORWARD LOOP (FFL) AS A NETWORK MOTIF
The Coherent (a) and Incoherent (b) Feedback Loops

- With the Coherent FFL, Y activates Z
- With the Incoherent FFL, Y inhibits Z
Three nodes subgraphs

- 13 possible three-nodes patterns
- Which ones are motifs?
Cont.) (Three nodes subgraphs

- Sub graph G with n nodes and g edges
- \(N^2\) possibilities to place an edge
- Probability of an edge in a given direction between a given pair of nodes: \(p = \frac{E}{N^2}\)
Cont.)(Three nodes subgraphs

- Mean number of appearances:

\[ <N_G> \approx a^{-1}N^{np^g} \]

- Mean connectivity: \( \lambda = \frac{E}{N} \rightarrow p = \frac{\lambda}{N} \)

\[ <N_G> \approx a^{-1}\lambda^g N^{n-g} \]
(Cont.) (Three nodes subgraphs)

- How $<N_G>$ scales with the network size?

$<N_G> \sim N^{n-g}$

- Triangle-shaped patterns (3 nodes and 3 edges):

$<N_{FFL}> \sim \lambda^3 N^0$  
$<N_{3loop}> \sim \frac{1}{3} \lambda^3 N^0$
Cont.) (Three nodes subgraphs

- FFL is the only motif of the 13 three-node patterns

<table>
<thead>
<tr>
<th></th>
<th>FFL</th>
<th>3LOOP</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli</td>
<td>42</td>
<td>0</td>
</tr>
<tr>
<td>Random net</td>
<td>1.7</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Feed-forward loop
3-node feedback loop (cycle)
FFL- Structure

- E. coli example:
FFL- Structure (Cont.)

**COHERENT FFL**

Coherent type 1

Coherent type 2

Coherent type 3

Coherent type 4

**INCOHERENT FFL**

Incoherent type 1

Incoherent type 2

Incoherent type 3

Incoherent type 4
Relative abundance of FLL types in yeast and E. coli:
FFL- Structure (Cont.)

- Logic function
  - AND logic
  - OR logic

- X and Y respond to external stimuli
Coherent Type-1 FFL – AND logic

- Sx appear, X rapidly changes to X*
- X* binds to gene Z, but cannot activate it
- X* binds to gene Y, and begins to transcript it
- Z begins to be expressed after $T_{on}$ time, when $Y^*$ crosses the activation threshold $K_{yz}$
Feed forward Loop: Example

- Coherent Type Feed forward Loop
- Production Rate of $Y = B_y \Theta(X > K_{xy})$
  - Where $U$ is a step Function
- $\frac{dY}{dt} = B_y \Theta(X > K_{xy}) - Ay$
- $\frac{dZ}{dt} = B_y \Theta(x > K_{xy}) B_y \Theta(Y > K_{yz}) - A_z$
- How does the production graphs of $X$, $Y$, $Z$ look?
Coherent Type-1 FFL – AND logic (Cont.)

- Definition:
  - **ON step** - Sx moves from absent to saturated state
  - **OFF step** - Sx moves from saturated to absent state

- Sy is present continuously
The Time Delay in Z

We can notice that after the signal $S_x$, there is a delay $T_{ON}$ before $Z$ increases.
No Time Delay when switching OFF

However, there is no such time delay when $S_x$ switches off. For this reason, this motif is called a
The Coherent FFL as a filter of pulses

We can note that Coherent FFL shows no response to ‘slim’ pulses, but lets through ‘fat’ pulses. It is
Coherent Type-1 FFL – OR logic

- Delay for OFF Steps of Sx
- Flagella system of E. coli:
  - $T_{OFF} = 1$ hour
The Incoherent FFL as a pulse generator

The Incoherent FFL, on the other hand, can work as a ‘pulse generator’. Notice that the signal in $S_x$ results in a ‘pulse’ in $Z$. 

![Graph showing $S_x$, $Y^+$, and $Z$ over time.](image-url)
Incoherent Type-1 FFL-Dynamics (Cont.)

- Dynamic equation of Z:
  - \( Y^* < K_{yz} \)
    - \( \frac{dZ}{dt} = \beta_z - \alpha_z Z \)
    - \( Z_m = \frac{\beta_z}{\alpha_z} \)
    - \( Z(t) = Z_m (1 - e^{-\alpha z t}) \)
  - \( Y^* > K_{yz} \)
    - \( \frac{dZ}{dt} = \beta'_z - \alpha_z Z \)
    - \( Z_{st} = \frac{\beta'_z}{\alpha_z} \)
    - \( Z(t) = Z_{st} + (Z(T_{rep}) - Z_{st}) e^{-\alpha (1 - T_{rep})} \)
  - \( Y^*(T_{rep}) = Y_{ST}(1 - e^{-\alpha y T_{rep}}) \Rightarrow T_{rep} = \frac{1}{\alpha_y} \ln \left[ \frac{1}{1 - K_{yz} / Y_{st}} \right] \)
The ‘Height’ of the Pulse depends on the threshold value
Incoherent Type-1 FFL-Example (Galactose)
Summary of Coherent and Incoherent FFLs.

- Note that the Coherent FFL introduces a delay in Z.
- Coherent FFL - allows the signals through only if they have a certain width.
- Now consider the case where the And function is replaced by the Or function.
- This also introduces a delay, but now on the ‘OFF’ signal rather than the ‘On’ signal.
What do we mean by coding standards and best practices?

- Good coding standards and practices are necessary to ensure software quality.
- Coding – Aesthetic issues
  - Naming the variables
  - Capitalization
  - Modularity
  - Language specific practices
  - How important are these? Important enough to make your project collapse...
- Formal methods to ensure software Quality
- Software Engineering Process Models
Properties

- Coherent FFL - Sign sensitive delay element
- And a filter
- Incoherent FFL - Pulse generator
The SIMs are common in sensory transcription networks:

- Genes from a same Pathway (Arginine synthesis).
- Genes responding to stress (DNA repair).
- Genes that assemble a same biological machine (ribosomal genes).
The SIMs can generate temporal programs of expression Eg: Last In First Out Order:

Last-In First-Out (LIFO) Program
Topological Generalization of network Motifs

- Three node Sub graphs - 13
- Four Node Sub graphs - 199
- Five Node Sub graphs – Over 9000
- Seven Node Subgraphs – Million

Fortunately, Most Biological Network shows families of dominant motifs

- Eg: Transcription Networks Show
  - Feed Forward Loops (FFL)
  - Dense Overlapping Regulons (DOR)
  - Single Input Modules (SIM)
Generalizing motifs – Role replication
First-In First-Out (FIFO) Program

\[ K_{xz1} > K_{xz2} > K_{xz3} \quad K'_{xz1} < K'_{xz2} < K'_{xz3} \]
FIFO program

- The FFL works as an FIFO program here because of the OR functions at $Z_i$
- Notice that threshold levels have to be reversed for $Y$ and $Z$
Aside: E Coli Flagella – Technological wonder

Single cell, 1 micron length
Contains only ~1000 protein types at any given moment

still:

Amazing technology

sensors

computer

engine

Communication bus
FIFO program is governed by a FFL
Multi-input FFL in Neuronal Networks

Nose Touch

FLP

ASH

AVD

AVA

Noxious Chemicals
Nose Touch

Backward movement

July 2013
N. H. N. D. de Silva
DORs

- The Genes in each DOR have a shared Global function
  - Such as stress response
  - Nutrient metabolism
  - Biosynthesis of key classes of cellular components
- The DOR forms the back bone of networks global structure
- So Can we use motifs to simplify networks?
How do Network Motifs Integrate?

The E.coli Transcription Network (partial)

A single DOR Layer

FFLs and SIMs are integrated within DORs

legends:

- transcription factor (TF)
- dense overlapping regulons (DOR)
- single input module (SIM)
- coherent feedforward loop
- incoherent feedforward loop
- single operon

A Master Regulators Layer (lots of Auto-Reg.)

Where are the X→Y→Z?
Compare with…

\[
gene \, x \quad \text{gene} \, y
\]

\[
X \quad Y
\]
DORs in transcription networks

- Form single layers - do not form cascades
- No ‘Single line’ cascades
- Rate limited networks tend not to employ such cascades
- Cascades are found in networks with interactions that are rapid compared to the timescale in which the network needs to function.
Network Motifs in developmental transcription networks

- These are not rate limited
Developmental Transcription Networks

The TF expression profile in a developing Drosophila embryo
Developmental Transcription Networks

Two-node Feedback Loops - Locking

- Both X AND Y are ON at the same time.
- Genes regulated by X and Y belong to the same tissue (or strip).

- X OR Y is ON at a given time.
- Genes regulated by X and Y belong to different tissues (strips).
Developmental Transcription Networks

Regulating Feedback Loops

- Double Positive Loops
- Double Negative Loops

X \rightarrow Y
Y \rightarrow X
Z \rightarrow Y
Y \rightarrow Z
X \rightarrow Y
Y \rightarrow X
Z \rightarrow Y
Y \rightarrow Z

June 2013
N. H. N. D. de Silva
Developmental Transcription Networks
Regulated Feedback Loops as a Memory Element

![Diagram showing regulatory feedback loops]
Developmental Transcription Networks

Cascades

\[ \text{X} \rightarrow \text{Y} \rightarrow \text{Z} \]

\[ \text{X} \rightarrow \text{Y} \rightarrow \text{Z} \]

Time

\[ \text{[X]} \]

\[ \text{[Y]} \]

\[ \text{[Z]} \]

\[ \text{[X]} \]

\[ \text{[Y]} \]

\[ \text{[Z]} \]

July 2013

N. H. N. D. de Silva

71
Feed Back Loops in developmental transcription networks

- X transcriptionally activates X.
- Y inhibits X.

- Z transcriptionally activates X.
- X forms a complex with Y.
- X phosphorylates Y.

Power → Heater → Temperature → Thermostat

(Fast) Protein-Protein Interactions
(Slow) Transcriptional Interactions
Feed Back Loops Produce Oscillation (remember from Control theory?)

Cdc20 oscillator controls Cell Cycle

Mutation of the Drosophila CWO gene
The critically damped response of the oscillator is described by the equation:

$$x = e^{-\gamma t} [x_0 + (v_0 + \gamma x_0)t]$$

which is a combination of an exponential and a linear term.

Oscillator with resonant frequency 10 rad/s started from rest.
After Barger & Olsson
Cell Signaling networks

- What are cell signaling networks?
Signal Transduction Cascades

**Diagram:**

- **Exterior**
  - Signal molecule
  - Activated adenyl cyclase

- **Cytosol**
  - Activated G protein
  - Ste20, Serine/threonine kinase
  - Ste11, Dual-specificity kinase (MEK)
  - Ste7, Ste5 scaffold protein
  - Fus3, Serine/threonine kinase (MAPK)

- **Ste 12 — transcription factor**
  - Mating

- **Cytosol**
  - Activated PKA
  - Activated, phosphorylated CREB
  - CREB-binding protein (CBP)
  - CREB-binding element
  - Activated target gene
  - Transcription
  - Translation
  - New protein

- **Nucleus**
  - Nuclear pore
  - Activated PKA
  - Inactive CREB

- **CYTOSOL**
  - Activated PKA

- **Signal Transduction**
  - X, X, P
  - Y, Y, P
  - Z, Z, P
  - T, T*
Popular Motifs in Signal Transduction Cascades

BiFan

Diamond

Generalization of DOR

Multi-layer Perceptrons (multi-DORs)
Multi Layer Perceptrons in Signal Transduction Cascades
Dynamics of Signal Transduction Cascades

\[
\frac{dY}{dt} = v_1 X_1 Y_o + v_2 X_2 Y_o - \alpha Y_p
\]

\[
\frac{dY}{dt} = 0
\]

At Steady State

\[
\frac{Y_p}{Y} = f(w_1 X_1 + w_2 X_2)
\]

\[
w_2 = \frac{v_2}{\alpha}, \quad w_1 = \frac{v_1}{\alpha}
\]

\[
w_1 X_1 + w_2 X_2 > 1
\]

Activation Threshold

\[
w_2 = 0.7 \quad w_1 = 0.7
\]
Dynamics of Signal Transduction Cascades

“AND” gate

\[ w_2 = 0.7 \quad \text{and} \quad w_1 = 0.7 \]

“OR” gate

\[ w_2 = 2 \quad \text{and} \quad w_1 = 2 \]
Dynamics of Signal Transduction Cascades

\[ Z_p/Z = f(w_{z1}Y_1 + w_zY_2) \]

"OR" gate

"AND" gate
Dynamics of Signal Transduction

Cascades

\[
\begin{align*}
X_1 & \quad 0.7 \quad 1.5 \quad X_2 \\
Y_1 & \quad 2 \quad -3 \\
Z & \quad 2 \quad -3
\end{align*}
\]

\[
\begin{align*}
Y_1 & \quad 0.7 \quad 1.5 \quad Y_2 \\
Z & \quad 2 \quad -3
\end{align*}
\]

\[
\begin{align*}
Y_1 & \quad 1.7 \quad 1.7 \quad Y_2 \\
Z & \quad 2 \quad -3
\end{align*}
\]

\[
\begin{align*}
Y_1 & \quad 1.7 \quad 1.7 \quad Y_2 \\
Z & \quad 2 \quad -3
\end{align*}
\]
Network motifs can function in several biological processes (sensory systems, development).

- different time scales (milliseconds, cell generations).

Network motifs can produce temporal programs (LIFO, FIFO, oscillation).

- Motifs within a network may be arranged in organized structures (perceptrons, interlocking FFL).

- It is possible to understand network topology better by reducing the network to motif-topology.
References

- Slides prepared by Dr. Mahendra Piraveenan for this class in previous years.
- “An Introduction to Systems Biology: Design Principles of Biological Circuits’ - Uri Alon
- “Network properties” by Dr. Nataša Pržulj, Department of Computing, Imperial College London
  (http://www.doc.ic.ac.uk/~natasha/course/docs/network_properties_and_models.ppt)
- “An Introduction To System Biology” - Uri Alon - Presented by Nitsan Chrizman