# Networks in systems biology

Matthew Macauley

Department of Mathematical Sciences
Clemson University
http://www.math.clemson.edu/~macaule/

Math 4500, Spring 2017

# Molecular networks

Many complex systems or processes in molecular biology can be represented as networks.

Typically, the nodes represent biomolecules (proteins, enzymes, etc.) and the edges represent interactions (activation, repression, etc.).

Edges are often signed. The edge

- $X \longrightarrow Y$ means "$X$ *activates* $Y$"
- $X \longrightarrow\!\!\!\!\!\!\mid\ \ Y$ means "$X$ *represses* $Y$" (or "$X$ inhibits $Y$")

Examples of such networks include:

- Protein–protein interaction networks
- Gene regulatory networks
- Signaling networks
- Metabolic networks.

These so-called molecular networks model how cellular processes communicate and react with their surroundings, maintain cellular homeostatis, and carry out necessary cell behavior.

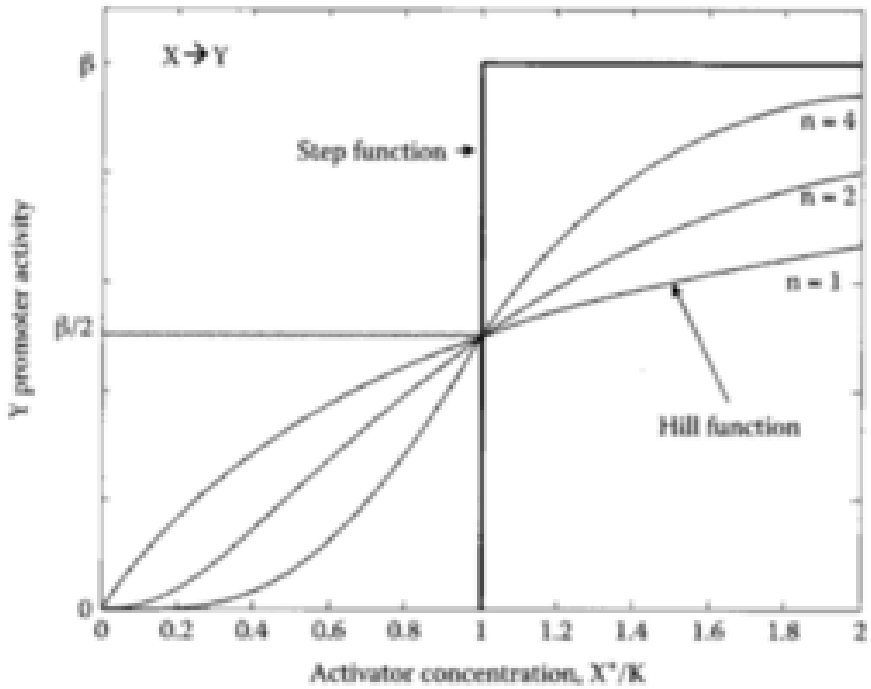
Think of these networks like big natural Rube Goldberg machines.

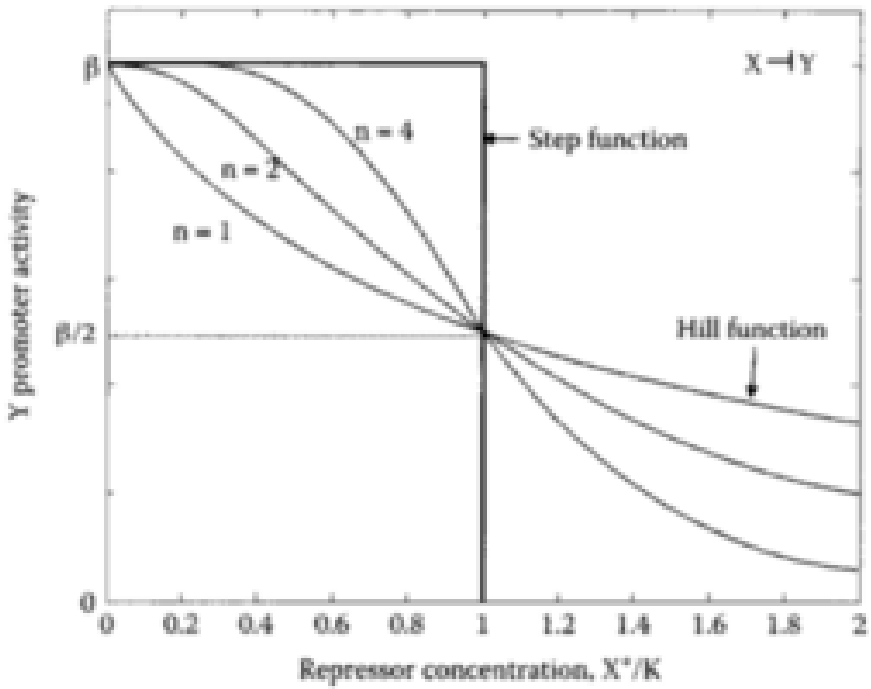
Many disease processes or disorders such as cancer, lactose intolerance, diabetes, vascular diseases, and autoimmunity, arise from problems in signal tranductions or gene regulation.

# Molecular networks

Interactions are usually highly nonlinear. They are often well-modeled by Hill functions. (Due to Michaelis–Menten models of enzyme kinetics.)

- The Hill function for an activator $X \longrightarrow Y$ is $f(X) = \dfrac{\beta X^n}{K^n + X^n}$.
- The Hill function for a repressor $X \longrightarrow\!\!\!\!\mid\ Y$ is $g(X) = \dfrac{\beta}{1 + (X/K)^n} = f(X)X^{-n}$.




"$S$-shaped" functions are generally called sigmoidal.

In the limit, as $n \to \infty$, Hill functions become step functions.

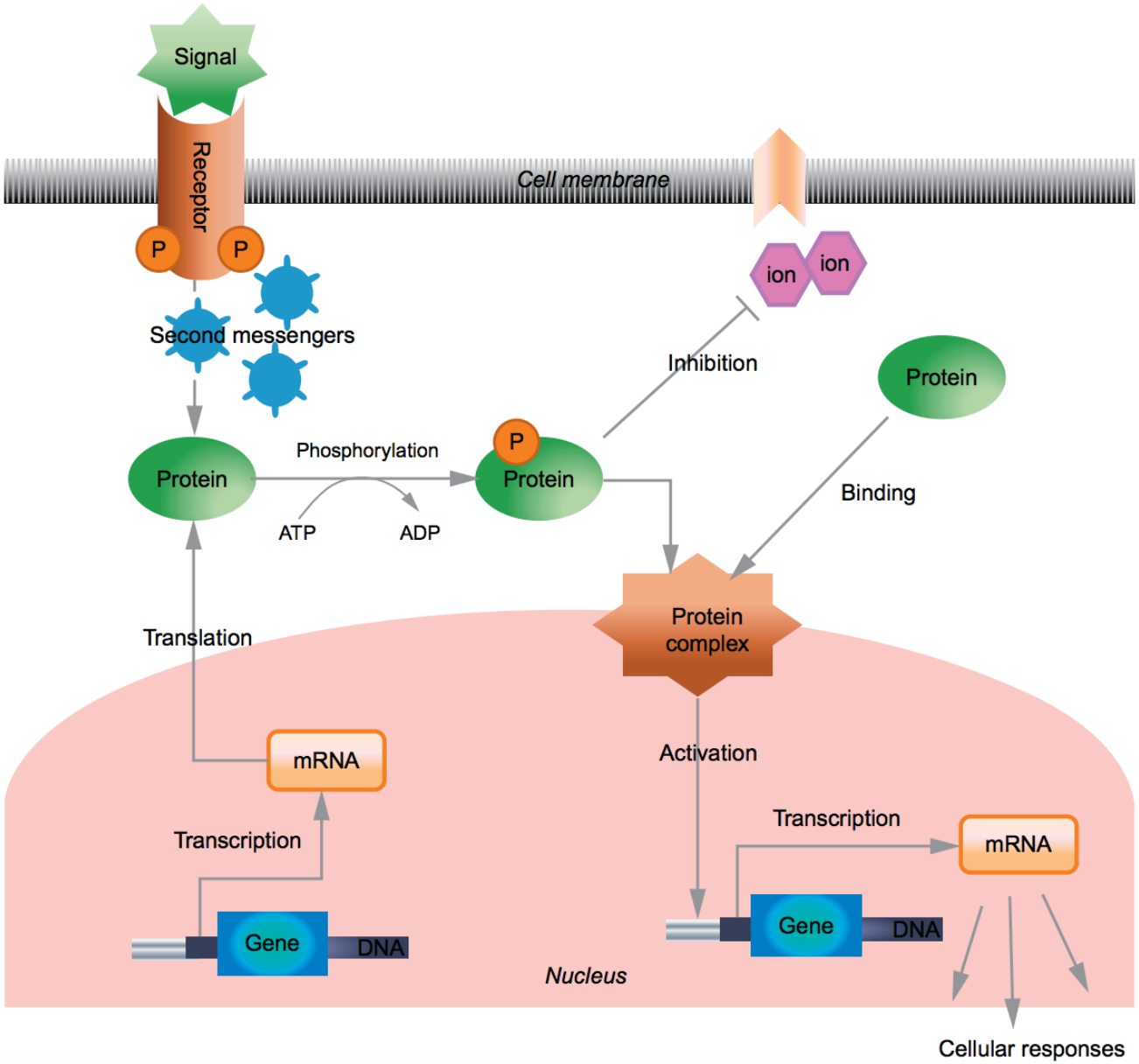



Figure: Scheme of a hypothetical signaling and gene regulatory network.

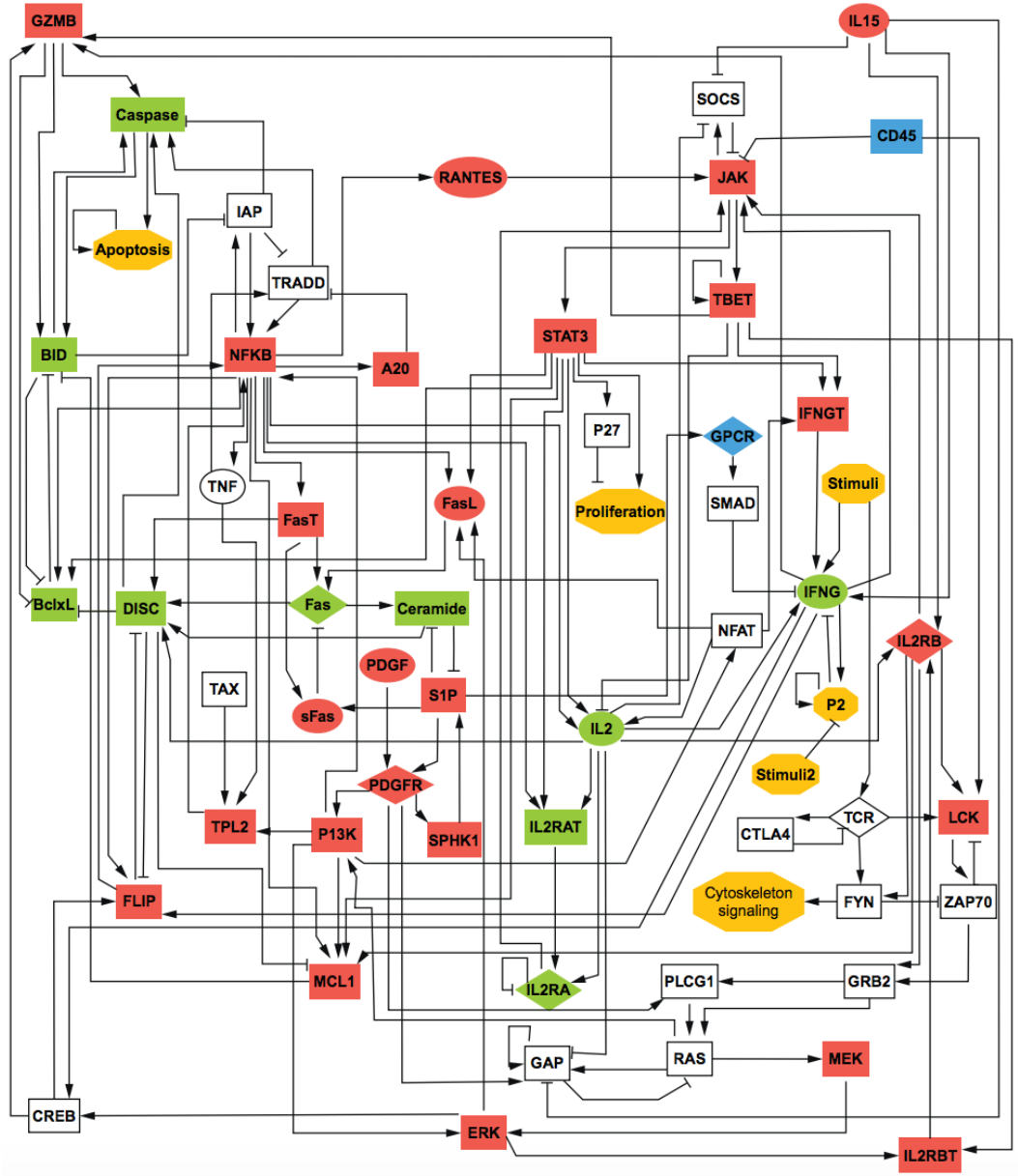



Figure: Signaling network involved in activation-induced cell death of killer T-cells. T-LGL leukemia disrupts this process, causing certain activated T-cells to survive, which later attack healthy cells.

# Network topology

Analysis of the network topology of molecular networks includes graph-theretic measures such as *centrality*, *network motifs*, and *shortest paths*.

Nodes can be categorized as *sources* (signals or parameters), *sinks* (outcomes), or neither.

*Centrality measures* describe the importance of individual nodes in the network. Examples include:

- degree (or in-degree, or out-degree),
- clustering coefficient,
- betweenness.

*Network motifs* are recurring patterns (subgraphs) with well-defined topologies. Common examples include:

- Feed-forward loops (coherent and incoherent)
- Feedback loops (positive and negative)

Feed forward loops tend to arise with greater frequency than in random networks.

## Rule of thumb

Positive feedback loops tend to support multistability while negative feedback loops lead to oscillations.

# Feed-forward loops

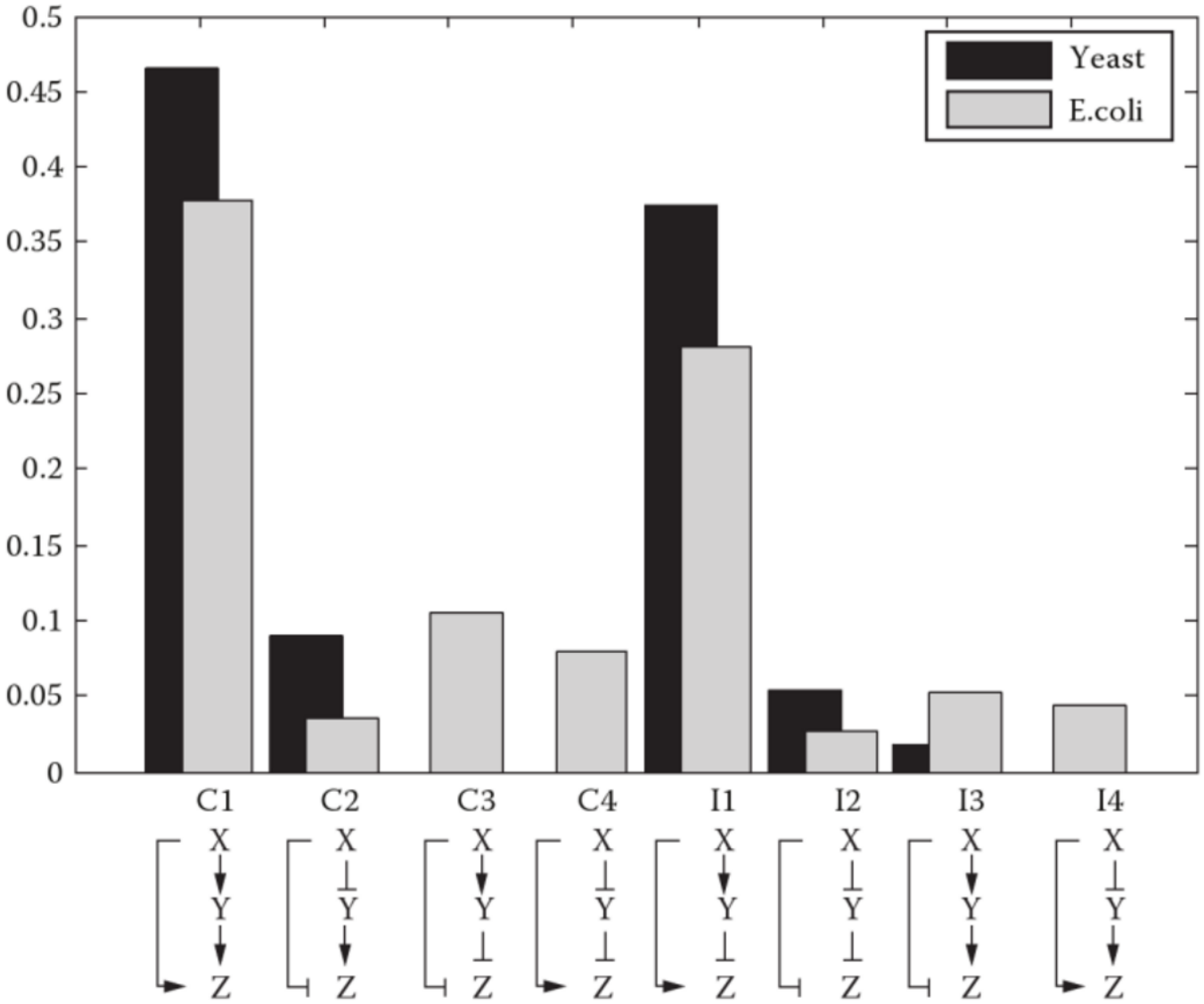



Figure: Relative abundance of the eight types of feed-forward loops in transcription networks (from U. Alon, 2007).

# Strongly connected components

## Definition

A directed graph is strongly connected if for every two nodes $u$ and $v$, there is a (directed) path from $u$ to $v$.

In any directed graph, the strongly connected components form a equivalence relation.

Moreover, there strongly connected components form a directed acyclic graph (i.e., are partially ordered): *add an edge from $C_i$ to $C_j$ if there is a directed path from some $x \in C_i$ to $y \in C_j$ in the original graph*.

Nodes in a strongly connected component tend to have a common task.

Signaling networks tend to have a large strongly connected "core". For example, the previous T-cell network has a core of 44 nodes (75% total).

# A simple Boolean network

Consider a 3-node network with a signal $A$ that activates $B$, which in turn, activates $C$.

For simplicity, suppose that each node takes a state from $\{0, 1\}$. (OFF or ON).

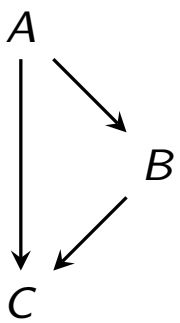
Suppose that $C$ is activated as long as both $A$ and $B$ are.

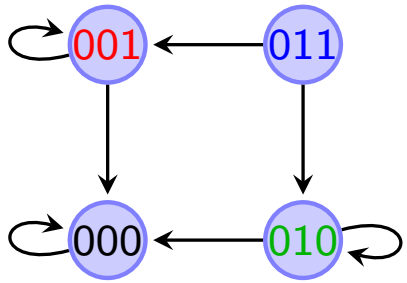
Here's what might happen biologically (not necessarily synchronously):

- $A$ turns on. This activates $B$ and then $C$, and the system settles in the ON steady-state, $(A, B, C) = (1, 1, 1)$.
- Eventually, $A$ turns off. This de-activates $B$ and then $C$, and the system flips to the OFF steady-state, $(A, B, C) = (0, 0, 0)$.

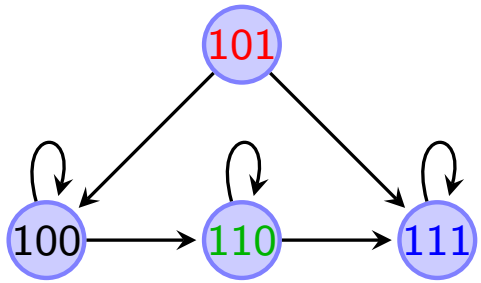
This can be visualized by the following Boolean network:

$$f_A = x_A$$
$$f_B = x_A$$
$$f_C = x_A \wedge x_B$$




$A$ is OFF




$A$ is ON

# Differenial equations vs. Boolean networks

Classically, molecular networks have been modeled using systems of ordinary differential equations (ODEs).

Nodes are represented by real-valued functions, often which represent concentrations.

Mathematicians and scientists have studied how the network topology affect the system dynamics.

## Rule of thumb

ODEs can exhibit complex dynamic behavior such as:

- filtering of noisy input signals (coherent feed-forward loops)
- excitation–adaptation (incoherent feed-forward loops, or negative feedback loops)
- multistability (positive feedback loops)

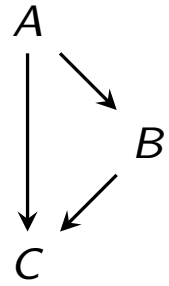
## Question

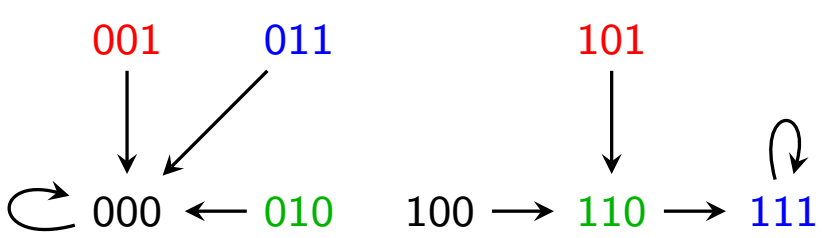
Can Boolean models be used as realistic qualitative approximations of molecular networks in biology?

# Synchronous vs. general asynchronous update

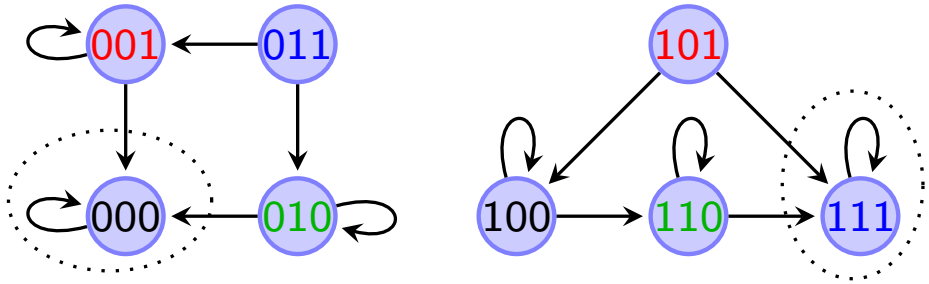
Let's compare the previous example under synchronous vs. asynchronous function update.

$$f_A = x_A$$
$$f_B = x_A$$
$$f_C = x_A \wedge x_B$$




Synchronous Boolean network




Asynchronous Boolean network

In actual biological networks, events and updates might occur randomly and unexpectedly.

Thus, one can make the case that the asynchronous update is "more natural".

Fixed-points correspond to the steady activation states of components (e.g., ON or OFF) or to cellular phenotypes (e.g., cancerous, non-cancerous) in signaling networks.

## Proposition

The set of fixed points of a Boolean network is independent of update scheme (synchronous, asynchronous, stochastic, etc.)

# Excitation–adaptation behavior

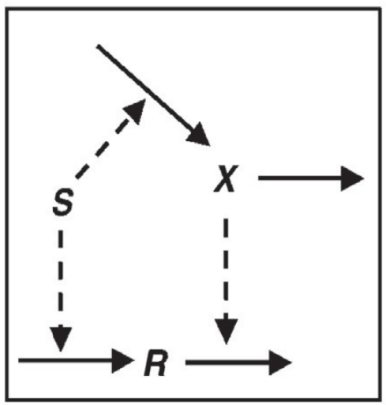
Chemotaxis is the movement of a cell in response to a chemical stimulus (the *signal*).

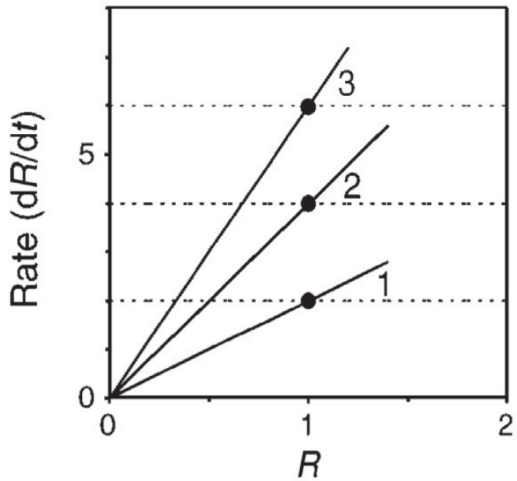
Consider the following system of ODEs, where $X$ and $R$ be concentrations of proteins, $k_i$ ($i = 1, \ldots, 4$) are rate constants, and $S$ be the value of the signal (a parameter):

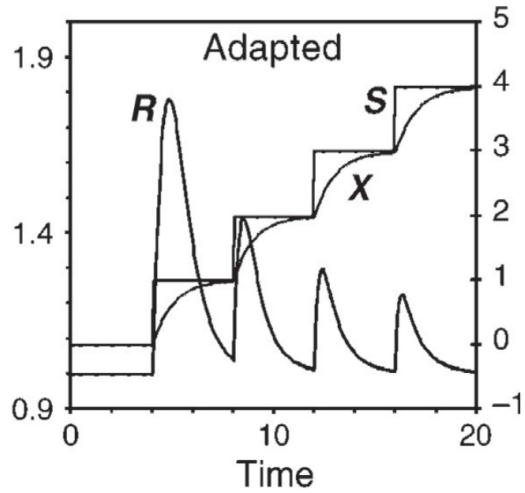
$$\frac{dR}{dt} = k_1 S - k_2 X R$$
$$\frac{dX}{dt} = k_3 S - k_4 X$$

## Analytical results

The (steady-state) concentration $R^*$ does not depend on $S$.



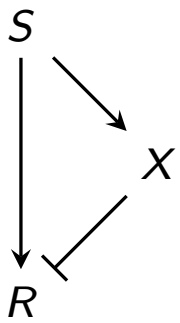



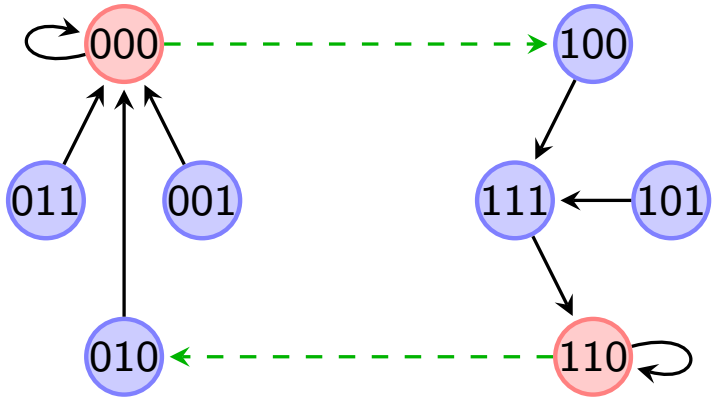
# Excitation–adaptation behavior

Let's create a Boolean model of this. The nodes will be $S$, $X$, $R$. Assume $X$ and $R$ have similar timescales and use synchronous update.

Here are the Boolean functions, wiring diagram, and state space:

$$f_S = x_S$$
$$f_X = x_S$$
$$f_R = x_S \wedge \overline{x_X}$$




The dashed lines describe a step-wise increase in the signal $S$ (i.e., $0 \to 1$ or $1 \to 0$).

## Analysis

(i) Start with $x_S = 0$. The system goes into 000 in one step.

(ii) Now set $x_S = 1$, which leads to 100.

(iii) The system transitions $100 \to 111$ excitation for $R$.

(iv) In the next step $111 \to 110$ adaptation for $R$.

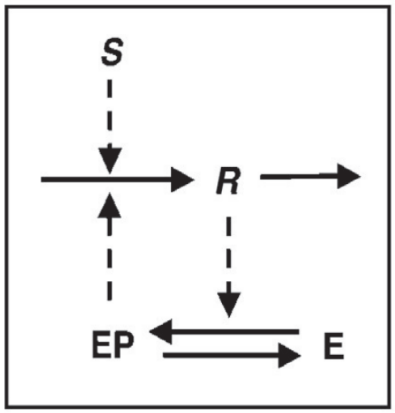
In summary, the change in $x_S$ drove a transient excitation of $x_R : 0 \mapsto 1$ but the steady-state adapted to its original value of $x_R = 0$.

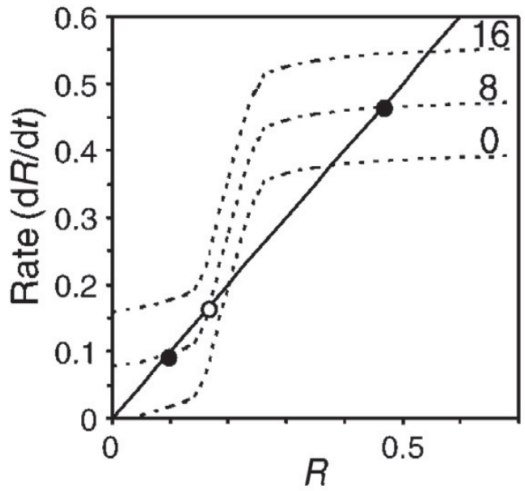
# Multistability and hysteresis

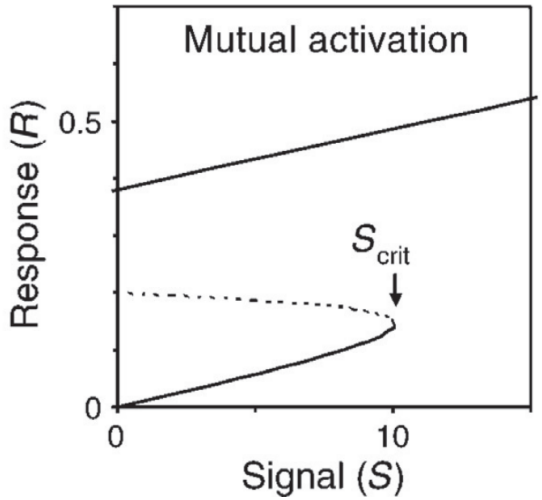
The phenomenon of multistability arises in physics, biology, and chemistry. It is the ability of a system to achieve multipe steady-states under the same external conditions.

Consider the following ODE, where $S$ and $P$ are concentrations of proteins, $k_i$ $(i = 0, 1, 2)$ are rate constants, and $f_E$ is a sigmoidal ("*Hill-like*") function:

$$\frac{dR}{dt} = k_0 f_E(R(t)) + k_1 S(t) - k_2 P(t)$$




Phosphorylation of a protein (adding of a phosphoryl group ($PO_3^{2-}$) changes its function, e.g., like an ON/OFF switch. The $EP \leftrightarrow E$ represents a phosphorylation–dephosphorylation cycle in which concentration of $P$ is constant.

This ODE exhibits irreversible bistability.

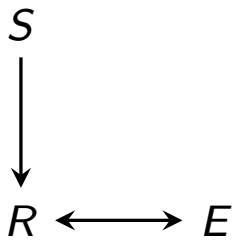
# Multistability and hysteresis

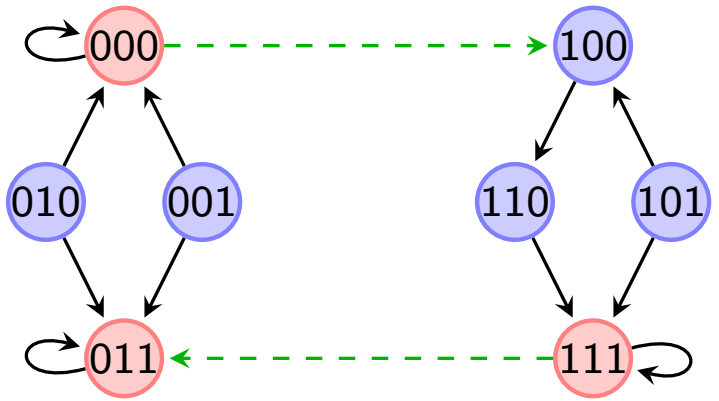
Let's create a Boolean model of this. The nodes will be $S$, $R$, $E$, where $E = 0$ and $E = 1$ are the Boolean approximation of the sigmoidal function $f_E(R)$.

In $R' = k_0 f_E(R(t)) + k_1 S(t) - k_2 P(t)$, synthesis of $R$ is catalyzed *independently* by $E$ and $S$.

Use an asynchronous update.

$$f_S = x_S$$
$$f_R = x_S \vee x_E$$
$$f_E = x_R$$




## Analysis

(i) Start at 000 (OFF). Increase $x_S$ to 1, which leads to 100.

(ii) The system settles to the ON steady-state 111.

(iii) Now, decrease $x_S$ to 0, which leads to the steady-state 011. However, $R$ is still 1.

*Exercise*. Show that the same behavior occurs under synchronous update.