Bistability in ODE and Boolean network models

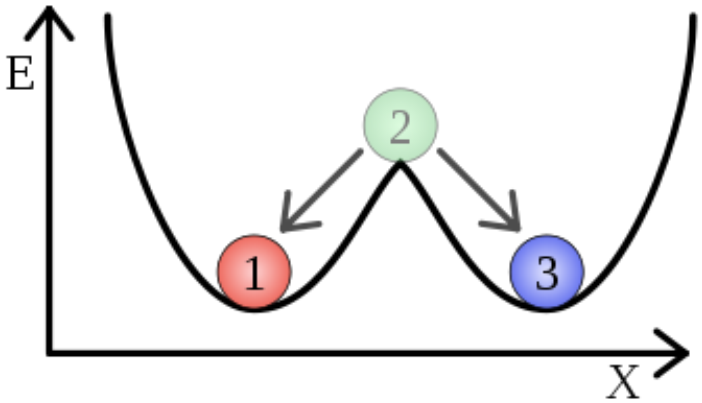
Matthew Macauley

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

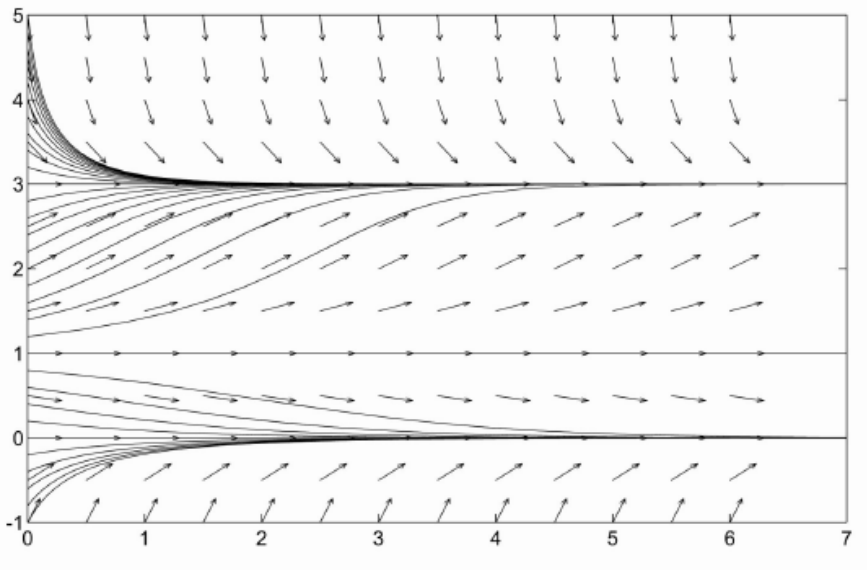
Math 4500, Spring 2017

Bistability

A system is bistable if it has two stable steady-states separated by an unstable state.



From Wikipedia.



The *threshold ODE*: $y' = -ry\left(1 - \frac{y}{M}\right)\left(1 - \frac{y}{T}\right)$.

In the threshold model for population growth, there are three steady-states, $0 < T < M$:

- M = carrying capacity (stable),
- T = extinction threshold (unstable),
- 0 = extinct (stable).

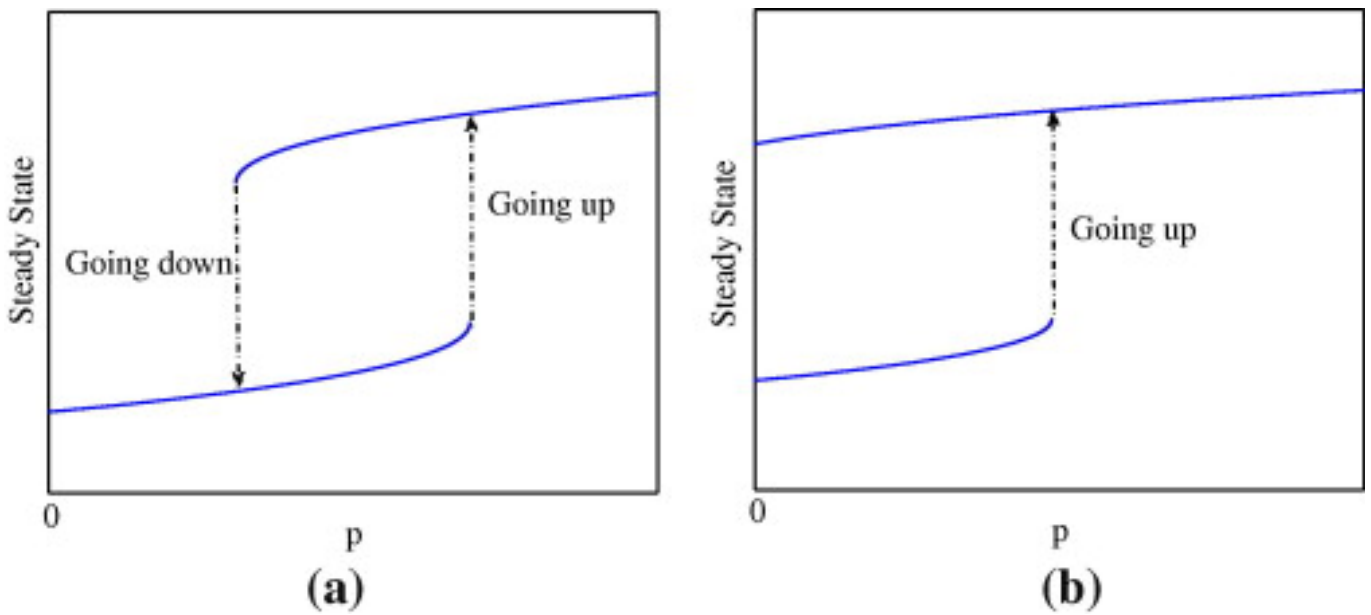
Types of bistability

The *lac* operon exhibits bistability.

The expression level of the *lac* operon genes are either almost zero ("basal levels"), or very high (thousands of times higher). There's no "inbetween" state.

The exact level depends on the concentration of intracellular lactose. *Let's denote this parameter by p.*

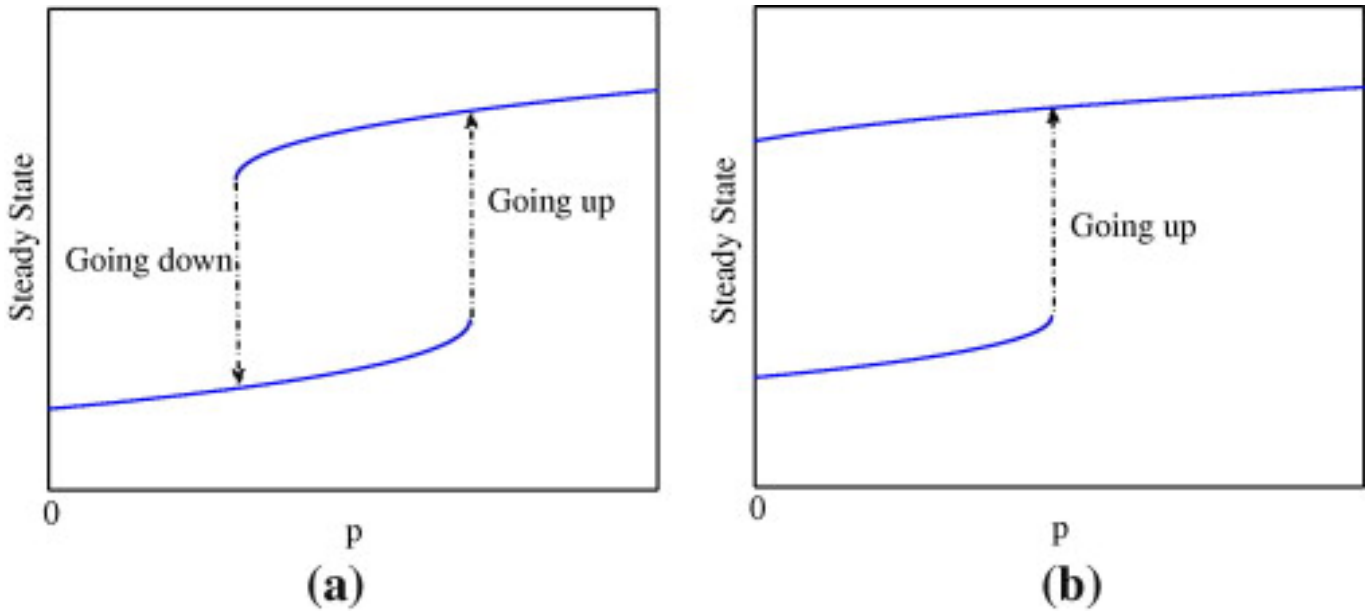
Now, let's "tune" this parameter. The result might look like the graph on the left.



This is reversible bistability. In other situations, it may be irreversible (at right).

Hysteresis

For reversible bistability, the *up-threshold* L_2 *of* p is higher than the *down-threshold* L_1 *of* p.



This is hysteresis: a dependence of a state on its current state *and* past state.

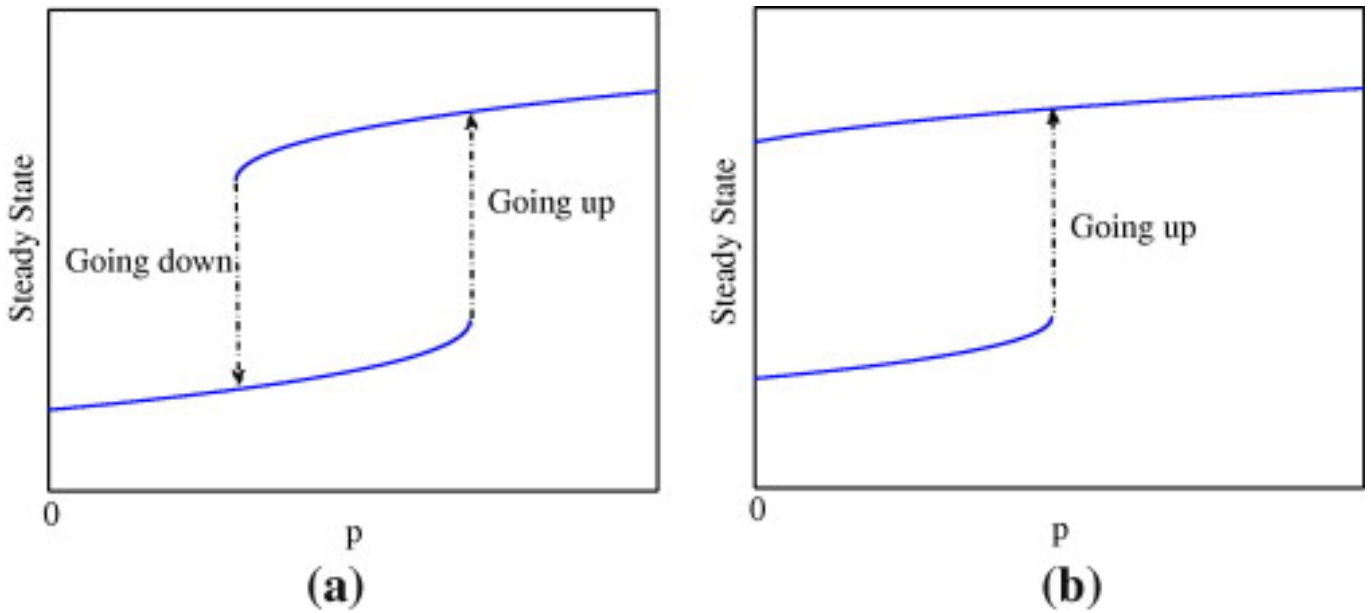
Thermostat example

Consider a home thermostat.

- If the temperature is $T < 66$ (e.g., in winter), the heat is on.
- If the temperature is $T > 76$ (e.g., in summer), the AC is on.
- If $66 < T < 77$, then we don't know whether it's spring or fall.

Hysteresis and the *lac* operon

If lactose levels are medium, then the state of the operon depends on whether or not a cell was grown in a lactose-rich environment.



Lac operon example

Let $[L]$ = concentration of intracellular lactose.

- If $[L] < L_1$, the operon is OFF.
- If $[L] > L_2$, the operon is ON.
- If $L_1 < [L] < L_2$, the operon might be ON or OFF.

In the region of bistability (L_1, L_2), one can find both induced and un-induced cells.

An ODE model of the *lac* operon

The Boolean models we've seen are too simple to capture bistability.

We will derive two different ODE models of the *lac* operon that exhibit bistability: one with 3 variables, and another with 5 variables.

These ODE models were designed using Michaelis–Menten equations from mass-action kinetics which we learned about earlier.

They will also incorporate other features, such as:

- dilution of protein concentration due to bacterial growth
- degredation (decay) of protein concentration
- time delays

After that, we'll see how bistability can indeed be captured by a Boolean model.

In general, bistable systems tend to have positive feedback loops (in their "wiring diagrams") or double-negative feedback loops (=positive feedback).

Modeling dilution in protein concentration due to bacterial growth

E. coli grows fast! It can double in 20 minutes. Thus, ODE models involving concentration can't assume volume is constant.

Let's define:

- V = average volume of an *E. coli* cell.
- x = number of molecules of protein X in that cell.

Assumptions:

- cell volume increases exponentially in time: $\frac{dV}{dt} = \mu V$.
- degradation of X is exponential: $\frac{dx}{dt} = -\beta x$.

The concentration of x is $[x] = \frac{x}{V}$. The derivative of this is (by the quotient rule):

$$\frac{d[x]}{dt} = \left(x'V - V'x\right)\frac{1}{V^2} = \left(-\beta xV - \mu Vx\right)\frac{1}{V^2} = -\left(\beta + \mu\right)\frac{x}{V} = -(\beta + \mu)[x].$$

Modeling of lactose repressor dynamics

Assumptions

- *Lac* repressor protein is produced at a constant rate.
- Laws of mass-action kinetics.

- *Repressor binds to allolactose*:

$$R + nA \underset{1}{\overset{K_1}{\rightleftharpoons}} RA_n \qquad \frac{d[RA_n]}{dt} = K_1[R][A]^n - [RA_n]$$

Assume the reaction is at equilibrium: $\frac{d[RA_n]}{dt} = 0$, and so $K_1 = \frac{[RA_n]}{[R][A]^n}$.

- *The repressor protein binds to the operator region if there is no allolactose*:

$$O + R \underset{1}{\overset{K_2}{\rightleftharpoons}} OR \qquad \frac{d[OR]}{dt} = K_2[O][R] - [OR].$$

Assume the reaction is at equilibrium: $\frac{d[OR]}{dt} = 0$, and so $K_2 = \frac{[OR]}{[O][R]}$.

Modeling of lactose repressor dynamics

Let O_{tot} = total operator concentration (a constant). Then, using $K_2 = \frac{[OR]}{[O][R]}$,

$$O_{tot} = [O] + [OR] = [O] + K_2[O][R] = [O](1 + K_2[R]).$$

Therefore, $\frac{[O]}{O_{tot}} = \frac{1}{1+K_2[R]}$. "*Proportion of free (unbounded) operator sites.*"

Let R_{tot} be total concentration of the repressor protein (constant):

$$R_{tot} = [R] + [OR] + [RA_n]$$

Assume only a few molecules of operator sites per cell: $[OR] \ll \max\{[R], [RA_n]\}$:

$$R_{tot} \approx [R] + [RA_n] = [R] + K_1[R][A]^n$$

Eliminating $[RA_n]$, we get $[R] = \dfrac{R_{tot}}{1 + K_1[A]^n}$.

Now, the proportion of free operator sites is:

$$\frac{[O]}{O_{tot}} = \frac{1}{1 + K_2[R]} = \frac{1}{1 + K_2(\frac{R_{tot}}{1+K_1[A]^n})} \cdot \frac{1 + K_1[A]^n}{1 + K_1[A]^n} = \underbrace{\frac{1 + K_1[A]^n}{K + K_1[A]^n}}_{:=f([A])},$$

where $K = 1 + K_2 R_{tot}$.

Modeling of lactose repressor dynamics

Summary

The proportion of free operator sites is

$$\frac{[O]}{O_{tot}} = \underbrace{\frac{1 + K_1[A]^n}{K + K_1[A]^n}}_{:=f([A])}, \qquad \text{where } K = 1 + K_2 R_{tot}.$$

Remarks

- The function $f([A])$ is (almost) a Hill function of coefficient n.
- $f([A] = 0) = \frac{1}{K} > 0$ "basal level of gene expression."
- f is *increasing* in $[A]$, when $[A] \geqslant 0$.
- $\lim_{[A]\to\infty} f([A]) = 1$ "with lots of allolactose, gene expression level is max'ed."

Modeling time-delays

The production of mRNA from DNA via transcription is not instantaneous; suppose it takes time $\tau > 0$.

Thus, the production rate of mRNA is not a function of allolactose at time t, but rather at time $t - \tau$.

Suppose protein P decays exponentially, and its concentration is $p(t)$.

$$\frac{dp}{dt} = -\mu p \quad \Longrightarrow \quad \int_{t-\tau}^{t} \frac{dp}{p} = -\mu \int_{t-\tau}^{t} dt\,.$$

Integrating yields

$$\ln p(t)\Big|_{t-\tau}^{t} = -\mu t\Big|_{t-\tau}^{t} dt = \ln \frac{p(t)}{p(t-\tau)} = -\mu[t - (t-\tau)] = -\mu\tau.$$

Exponentiating both sides yields $\frac{p(t)}{p(t-\tau)} = e^{-\mu\tau}$, and so

$$p(t) = e^{-\mu\tau} p(t-\tau).$$

A 3-variable ODE model of the *lac* operon

Consider the following 3 quantities, which represent *concentrations* of:

- $M(t) =$ mRNA,
- $B(t) = \beta$-galactosidase,
- $A(t) =$ allolactose.

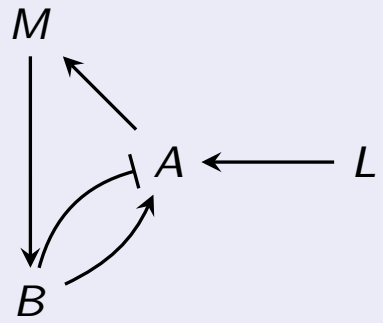
Assumption: Internal lactose (L) is available and is a parameter.

The model (Yildirim and Mackey, 2004)

$$\frac{dM}{dt} = \alpha_M \frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} - \tilde{\gamma}_M M$$

$$\frac{dB}{dt} = \alpha_B e^{-\mu\tau_B} M_{\tau_B} - \tilde{\gamma}_B B$$

$$\frac{dA}{dt} = \alpha_A B \frac{L}{K_L + L} - \beta_A B \frac{A}{K_A + A} - \tilde{\gamma}_A A$$



These are *delay differential equations*, with discrete time delays due to the transcription and translation processes.

There should be a self-loop  at M, B, and A, but we're omitting them for clarity.

A 3-variable ODE model of the *lac* operon

ODE for β-galactosidase (B)

$$\frac{dB}{dt} = \alpha_B e^{-\mu\tau_B} M_{\tau_B} - \tilde{\gamma}_B B,$$

Justification:

- $\tilde{\gamma}_B B = \gamma_B B + \mu B$ represents loss due to β-galactosidase degredation and dilution from bacterial growth.
- Production rate of β-galactosidase, is proportional to mRNA concentration.
- τ_B = time required for translation of β-galactosidase from mRNA, and $M_{\tau_B} := M(t - \tau_B)$.
- $e^{-\mu\tau_B} M_{\tau_B}$ accounts for the time-delay due to translation.

A 3-variable ODE model of the *lac* operon

ODE for mRNA (M)

$$\frac{dM}{dt} = \alpha_M \frac{1 + K_1(e^{\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} - \tilde{\gamma}_M M$$

Justification:

- $\tilde{\gamma}_M M = \gamma_M M + \mu M$ represents loss due to mRNA degredation and dilution from bacterial growth.
- Production rate of mRNA [=expression level!] is proportional to the fraction of free operator sites,
$$\frac{[O]}{O_{tot}} = \frac{1 + K_1 A^n}{1 + K_1 A^n} = f(A).$$
- τ_M = time required for transcription of mRNA from DNA, and $M_{\tau_M} := M(t - \tau_M)$.
- The term $e^{-\mu\tau_M} M_{\tau_M}$ accounts for the time-delay due to transcription.

A 3-variable ODE model of the *lac* operon

ODE for allolactose (A)

$$\frac{dA}{dt} = \alpha_A B \frac{L}{K_L + L} - \beta_A B \frac{A}{K_A + A} - \tilde{\gamma}_A A$$

Justification:

- $\tilde{\gamma}_A A = \gamma_A A + \mu A$ represents loss due to allolactose degredation and dilution from bacterial growth.
- The first two terms models the chemical reaction catalyzed by the enzyme β-galactosidase:

$$L \xrightarrow{\alpha_A} A \xrightarrow{\beta_A} \textit{glucose} + \textit{galactose}.$$

A 3-variable ODE model of the *lac* operon

Steady-state analysis

To find the steady states, we must solve the nonlinear system of equations:

$$0 = \alpha_M \frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} - \tilde{\gamma}_M M$$

$$0 = \alpha_B e^{-\mu\tau_B} M_{\tau_B} - \tilde{\gamma}_B B$$

$$0 = \alpha_A B \frac{L}{K_L + L} - \beta_A B \frac{A}{K_A + A} - \tilde{\gamma}_A A$$

This was done by Yildirim et al. (2004). They set $L = 50 \times 10^{-3}$ mM, which was in the "bistable range."

They estimated the parameters through an extensive literature search.

Finally, they estimated $\mu = 3.03 \times 10^{-2}$ min^{-1} by fitting ODE models to experimental data.

Steady states	A^* (mM)	M^* (mM)	B^* (mM)	
I.	4.27×10^{-3}	4.57×10^{-7}	2.29×10^{-7}	basal (stable)
II.	1.16×10^{-2}	1.38×10^{-6}	6.94×10^{-7}	medium (unstable)
III.	6.47×10^{-2}	3.28×10^{-5}	1.65×10^{-5}	high (stable)

3-variable ODE model

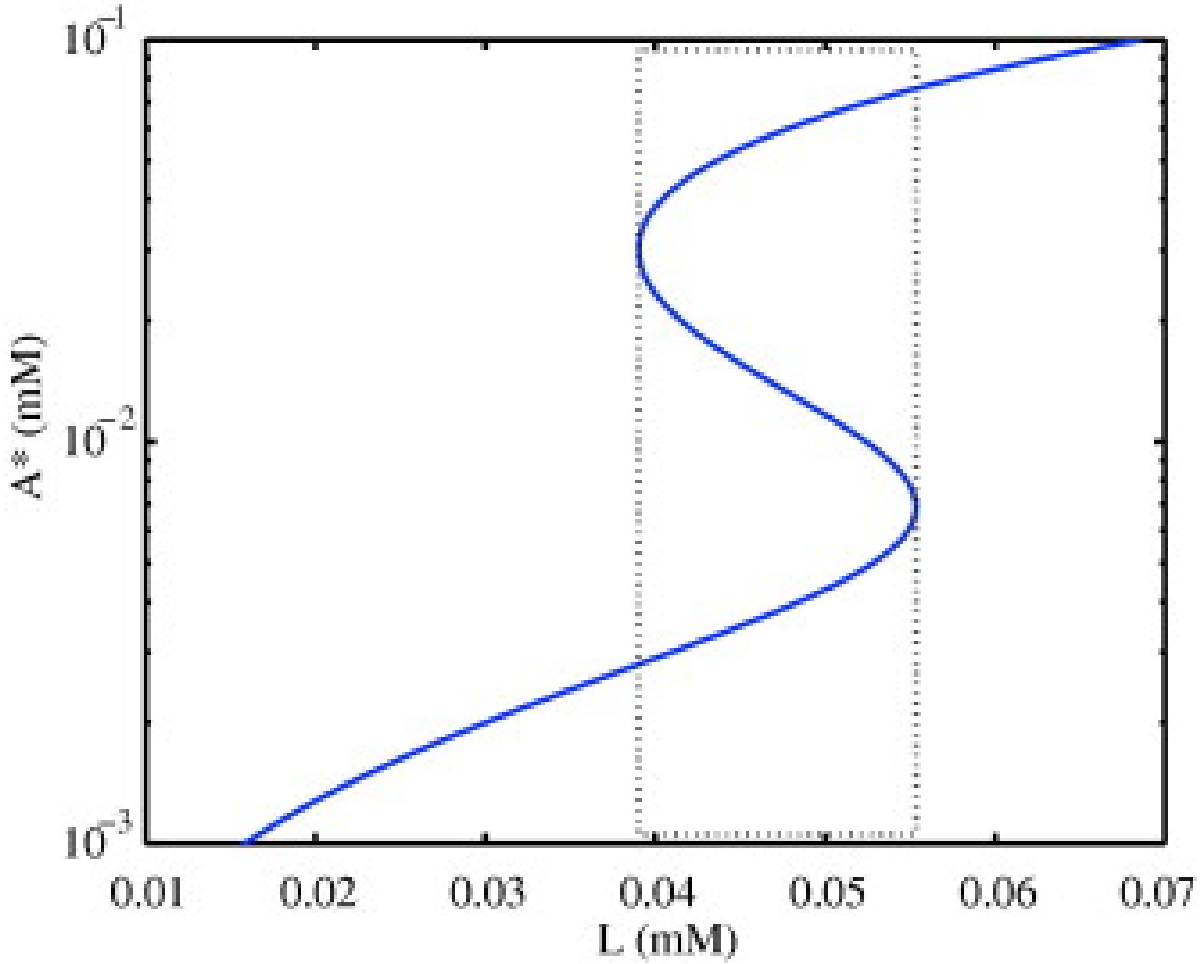


Figure: The fixed points of the allolactose concentration A^* in ODE model as a function of the parameter L (lactose). For a range of L concentrations there are 3 coexisting steady states, which is the phenomenon of bistability.

3-variable ODE model

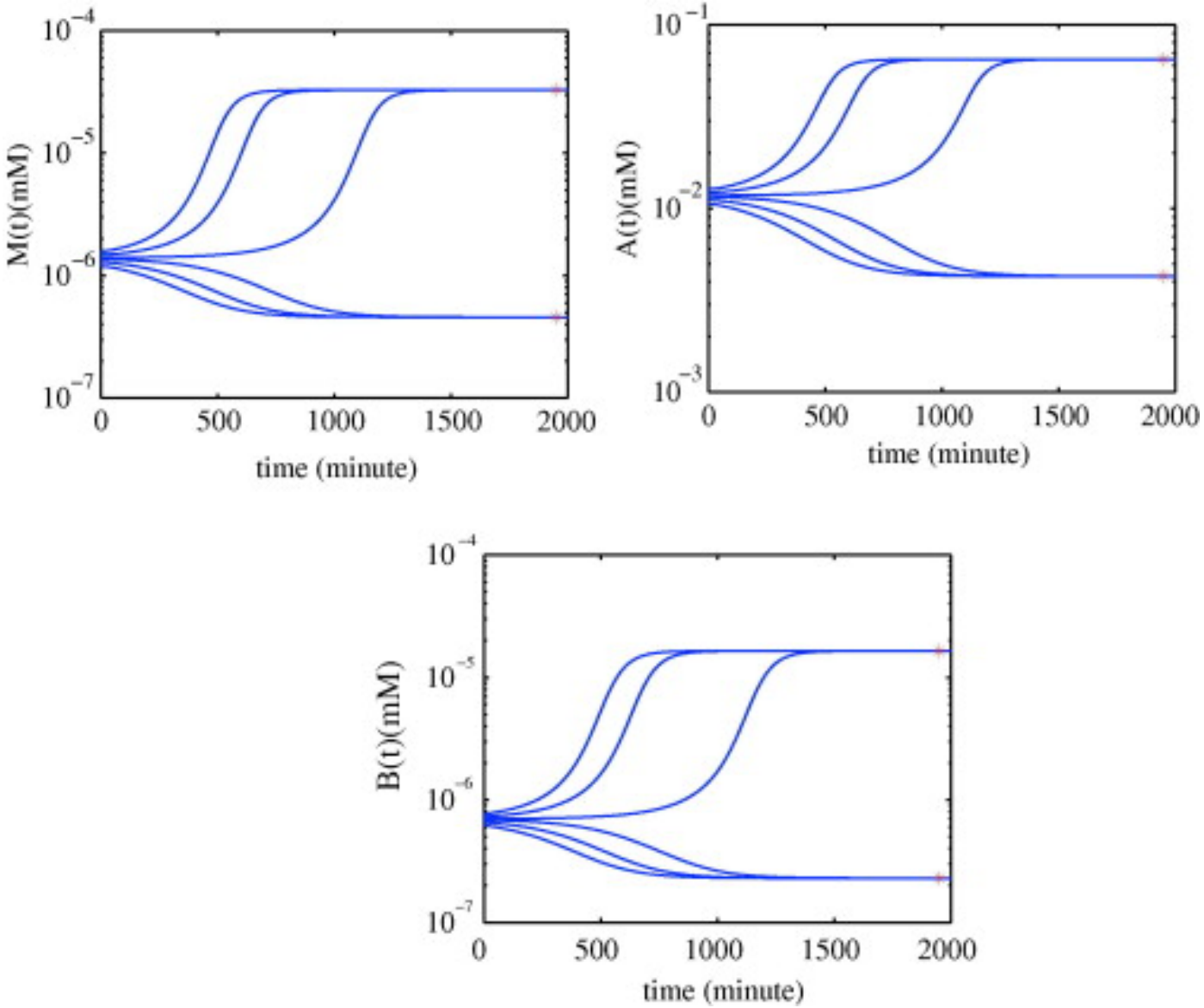


Figure: Numerical solutions of $M(t)$ (mRNA), $B(t)$ (β-galactosidase), and $A(t)$ (allolactose), using $L = 50 \times 10^{-3}$.

5-variable ODE model

Consider the following 5 variables, which represent *concentrations* of:

- $M(t)$ = mRNA,
- $B(t)$ = β-galactosidase,
- $A(t)$ = allolactose.
- $P(t)$ = lac permease.
- $L(t)$ = intracellular lactose.

The model (Yildirim and Mackey, 2004)

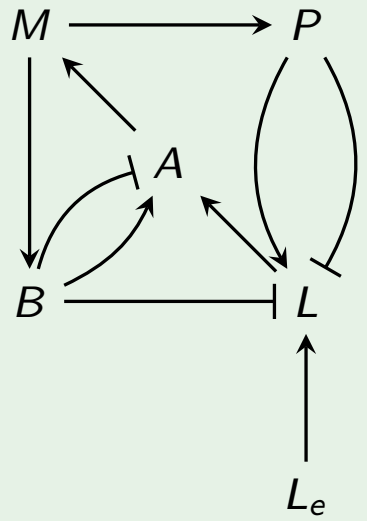
$$\frac{dM}{dt} = \alpha_M \frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} + \Gamma_0 - \tilde{\gamma}_M M$$

$$\frac{dB}{dt} = \alpha_B e^{-\mu\tau_B} M_{\tau_B} - \tilde{\gamma}_B B$$

$$\frac{dA}{dt} = \alpha_A B \frac{L}{K_L + L} - \beta_A B \frac{A}{K_A + A} - \tilde{\gamma}_A A$$

$$\frac{dP}{dt} = \alpha_P e^{-\mu(\tau_B + \tau_P)} M_{\tau_B + \tau_P} - \tilde{\gamma}_P P$$

$$\frac{dL}{dt} = \alpha_L P \frac{L_e}{K_{L_e} + L_e} - \beta_{L_e} P \frac{L}{K_{L_e} + L} - \alpha_A B \frac{L}{K_L + L} - \tilde{\gamma}_L L$$



5-variable ODE model

Remarks

- The only difference in the ODE for M is the extra term Γ_0 which describes the basal transcription rate ($L_e = 0$).
- The ODEs for B and A are the same as in the 3-variable model.
- The ODE for P is very similar to the one for B:
  - production rate of *lac* permease $\propto$ mRNA concentration, with a time-delay.
  - the 2nd term accounts for loss due to degredation and dilution.
- The ODE for lactose,

$$\frac{dL}{dt} = \alpha_L P \frac{L_e}{K_{L_e} + L_e} - \beta_{L_e} P \frac{L}{K_{L_1} + L} - \alpha_A B \frac{L}{K_L + L} - \tilde{\gamma}_L L,$$

  is justified by:
  - The first two terms model the transport lactose by *lac* permease:

$$L_e \underset{\beta_{L_e}}{\overset{\alpha_L}{\rightleftharpoons}} L$$

  - The 3rd term describes the reaction catalyzed by β-galactosidase: $L \xrightarrow{\alpha_A} A$.
  - the 4th term accounts for loss due to degredation and dilution.

A 5-variable ODE model

To find the steady states, we set $M' = A' = B' = L' = P' = 0$ and solve the resulting nonlinear system of equations.

This was done by Yildirim et al. (2004). They set $L_e = 50 \times 10^{-3}$ mM, in the "bistable range."

They also estimated the parameters through an extensive literature search.

Finally, they estimated $\mu = 2.26 \times 10^{-2}$ min^{-1} by fitting the ODE models to experimental data.

SS's	A^* (nM)	M^* (mM)	B^* (mM)	L^* (mM)	P^* (mM)
I.	7.85×10^{-3}	2.48×10^{-6}	1.68×10^{-6}	1.69×10^{-1}	3.46×10^{-5}
II.	2.64×10^{-2}	7.58×10^{-6}	5.13×10^{-6}	2.06×10^{-1}	1.05×10^{-4}
III.	3.10×10^{-1}	5.80×10^{-4}	3.92×10^{-4}	2.30×10^{-1}	8.09×10^{-3}

5-variable ODE model

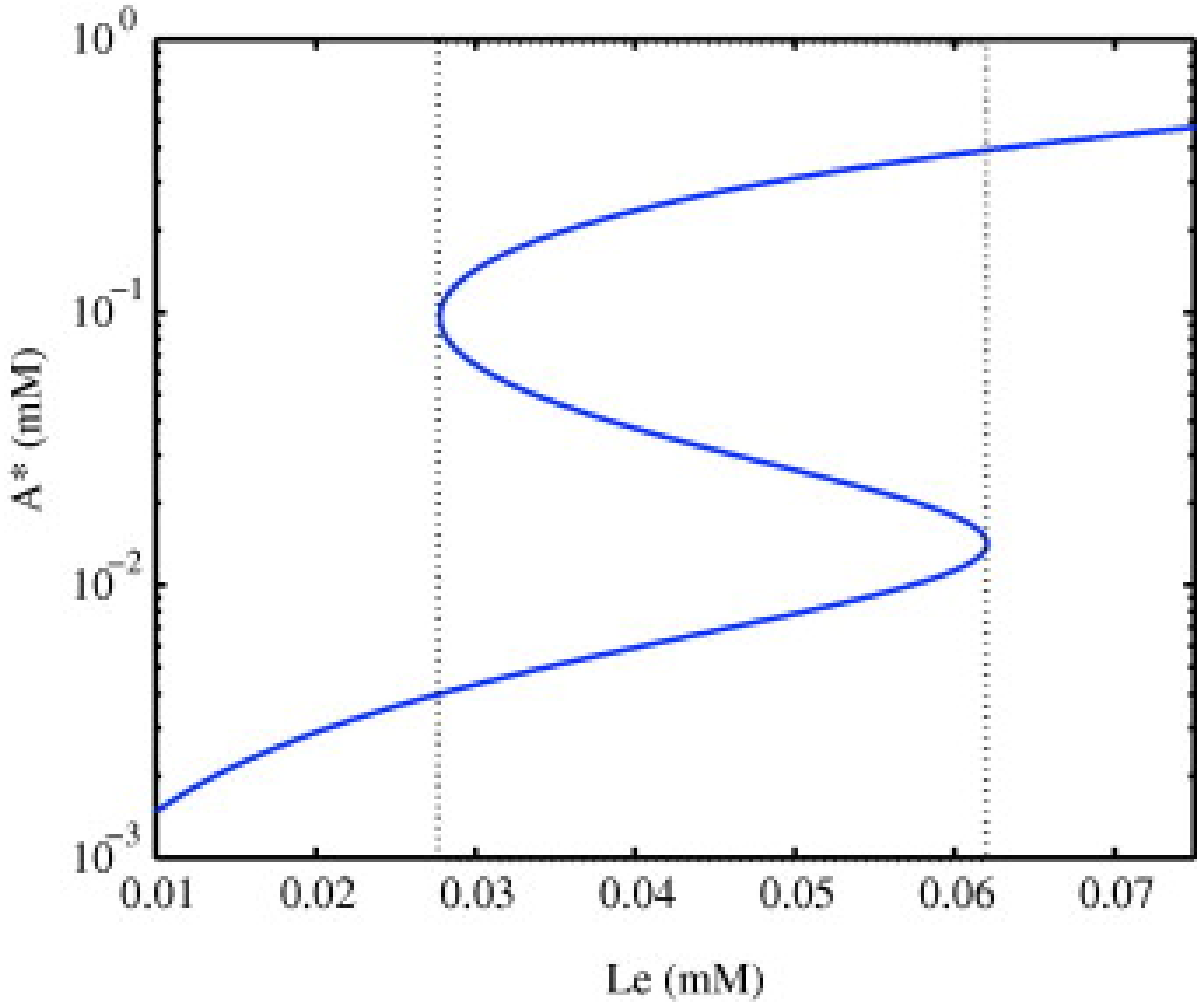


Figure: The fixed points of the allolactose concentration A^* in ODE model as a function of the parameter L_e (external lactose). For a range of L_e concentrations there are 3 coexisting steady states, which is the phenomenon of bistability.

5-variable ODE model

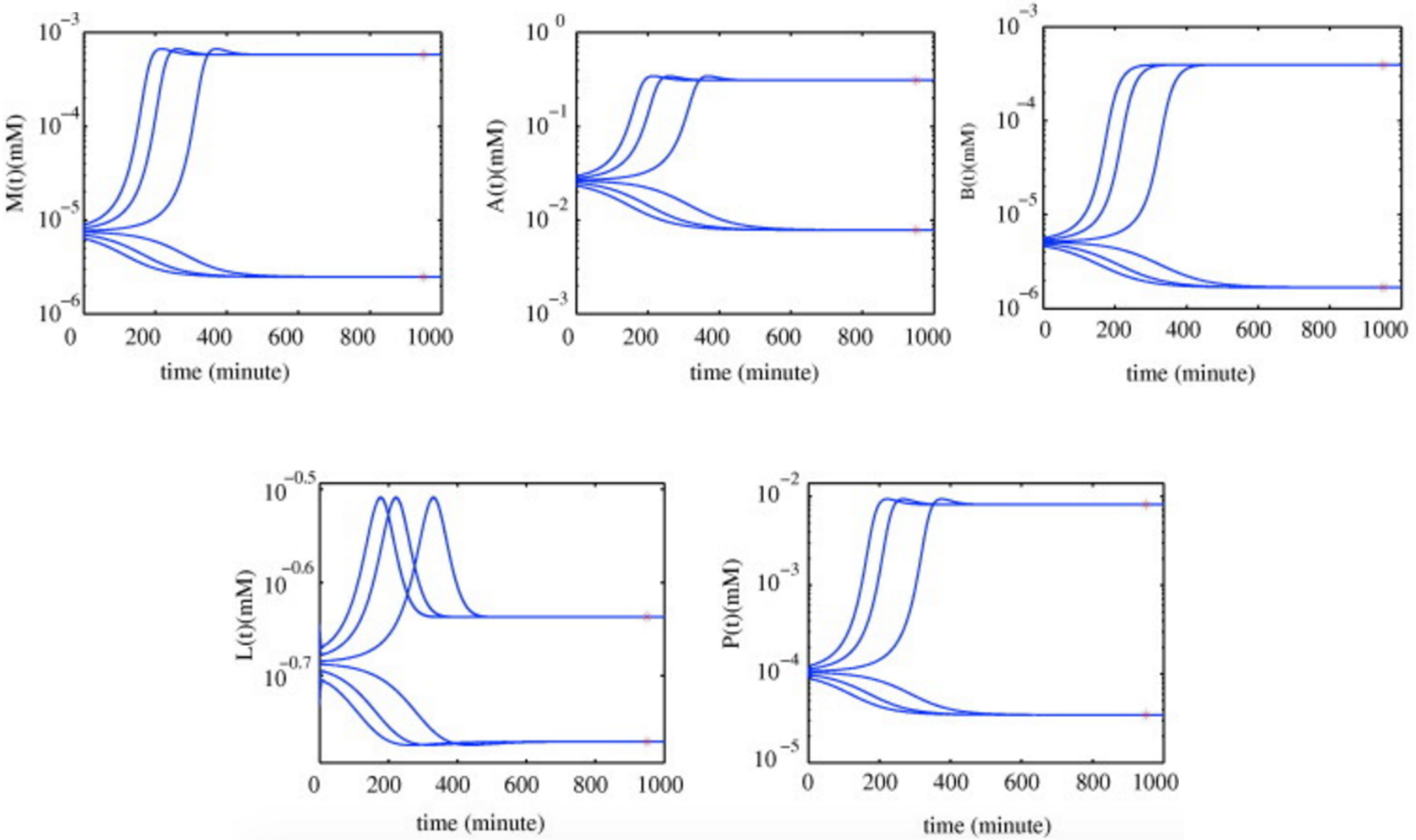


Figure: Numerical solutions of mRNA, β-galactosidase, allolactose, *lac* permease, and lactose concentrations, using $L_e = 50 \times 10^{-3}$.

Bistability in Boolean networks

For bistability to exist, we need to be able to describe three levels of lactose: high, medium, and low.

In a Boolean network framework, one way to do this is to add variable(s):

Medium levels of lactose

Introduce a new variable L_m meaning "at least medium levels" of lactose. Clearly, $L = 1$ implies $L_m = 1$.

- High lactose: $L = 1$, $L_m = 1$.
- Medium lactose: $L = 0$, $L_m = 1$.
- Basal lactose levels: $L = 0$, $L_m = 0$.

We can ignore any state for which $L = 1$, $L_m = 0$.

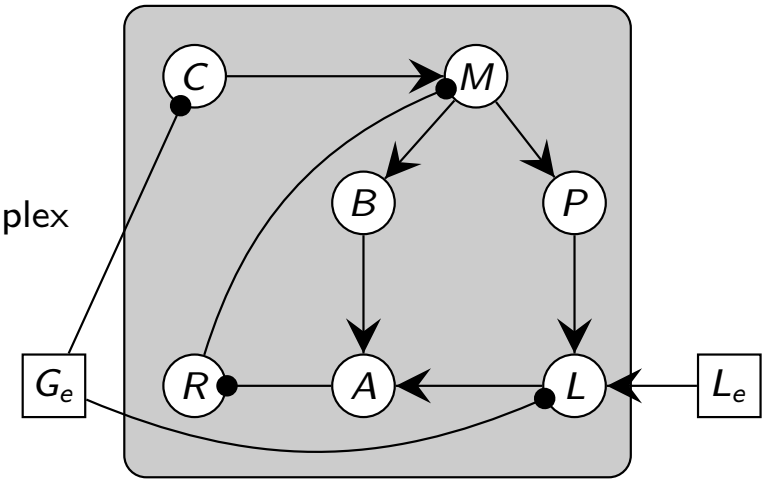
Since β-galactosidase converts lactose into allolactose, it makes sense to add a variable A_m to differentiate between high, medium, and low levels of allolactose.

It's not necessary, but we will also introduce R_m so we can speak of medium levels of the repressor protein.

A Boolean network model of the *lac* operon

Consider the following Boolean network model, which was published in Veliz-Cuba / Stigler (2011).

$M =$ mRNA
$P =$ *lac* permease
$B =$ β-galactosidase
$C =$ cAMP-CAP complex
$R =$ repressor protein
$L =$ lactose
$A =$ allolactose
$G =$ glucose



$$
\begin{cases}
f_M = \overline{R} \wedge \overline{R_m} \wedge C \\
f_P = M \\
f_B = M \\
f_C = \overline{G_e} \\
f_R = \overline{A} \wedge \overline{A_m} \\
f_{R_m} = (\overline{A} \wedge \overline{A_m}) \vee R \\
f_A = L \wedge B \\
f_{A_m} = L \vee L_m \\
f_L = \overline{G_e} \wedge P \wedge L_e \\
f_{L_m} = \overline{G_e} \wedge ((L_{em} \wedge P) \vee L_e)
\end{cases}
$$

Comments

- Circles denote variables, and squares denote parameters.
- The subscript e denotes extracellular concentrations.
- The subscript m denotes medium concentration.

A Boolean network model of the *lac* operon

Here is that model as a polynomial dynamical system:

$x_1 =$ *lac* mRNA (M)	$f_1 = x_4(x_5+1)(x_6+1)$
$x_2 =$ *lac* permease (P)	$f_2 = x_1$
$x_3 = \beta$-galactosidase (B)	$f_3 = x_1$
$x_4 =$ cAMP-CAP complex (C)	$f_4 = G_e + 1$
$x_5 =$ high repressor protein (R)	$f_5 = (x_7+1)(x_8+1)$
$x_6 =$ med. repressor protein (R_m)	$f_6 = (x_7+1)(x_8+1) + x_5 + (x_7+1)(x_8+1)x_5$
$x_7 =$ high allolactose (A)	$f_7 = x_3x_9$
$x_8 =$ med. allolactose (A_m)	$f_8 = x_9 + x_{10} + x_9x_{10}$
$x_9 =$ high intracellular lactose (L)	$f_9 = x_2(G_e+1)L_e$
$x_{10} =$ med. intracellular lactose (L_m)	$f_{10} = (x_2L_{em} + L_e + x_2L_{em}L_e)(G_e+1)$

To find the fixed points, we need to solve the following system of nonlinear equations over $\mathbb{F}_2$, for six choices of initial conditions, (L_e, L_{em}, G_e):

$$\{f_i + x_i = 0, \quad i = 1, 2, \ldots, 10\}.$$

This is an easy task in Sage.

The bistable case

Let's compute the fixed points with medium lactose ($L_e = 0, L_{em} = 1$) and no glucose ($G_e = 0$), which is the case where we hope to observe bistability.

```
P.<x1,x2,x3,x4,x5,x6,x7,x8,x9,x10> = PolynomialRing(GF(2), 10, order='lex'); P
```

```
Multivariate Polynomial Ring in x1, x2, x3, x4, x5, x6, x7, x8, x9, x10 over Finite Field of size 2
```

```
Le=0; Lem=1; Ge=0; print "L_e =", Le; print "L_em =", Lem; print "G_e =", Ge;
```

```
L_e = 0
L_em = 1
G_e = 0
```

```
I = ideal(x1+x4*(x5+1)*(x6+1), x2+x1, x3+x1, x4+Ge+1, x5+(x7+1)*(x8+1), x6+(x7+1)*(x8+1)+x5+(x7+1)*
(x8+1)*x5, x7+x3*x9, x8+x9+x10+x9*x10, x9+x2*(Ge+1)*Le, x10+(x2*Lem+Le+x2*Lem*Le)*(Ge+1)); I
```

```
Ideal (x1 + x4*x5*x6 + x4*x5 + x4*x6 + x4, x1 + x2, x1 + x3, x4 + 1, x5 + x7*x8 + x7 + x8 + 1, x5*x7
```

```
B = I.groebner_basis(); B
```

```
[x1 + x10, x2 + x10, x3 + x10, x4 + 1, x5 + x10 + 1, x6 + x10^2 + 1, x7, x8 + x10, x9, x10^3 + x10]
```

We see immediately that $x_7 = x_9 = 0$ and $x_4 = 1$.

Recall that $x_{10}^k = x_{10}$ for all $k \in \mathbb{N}$. Thus, the last equation, $x_{10}^3 + x_{10} = 0$ doesn't give any information about x_{10}.

The variables x_1, x_2, x_3, and x_8 must equal x_{10}.

The variables x_5 and x_6 must be the opposite of x_{10}. We get two fixed points:

$$(M, P, B, C, R, R_m, A, A_m, L, L_m) = (0,0,0,1,1,1,0,0,0,0) \quad \text{and} \quad (1,1,1,1,0,0,0,1,0,1).$$

Fixed point analysis and bistability

Computing the fixed point(s) for the other 5 initial conditions is an easy task in Sage:

(L_e, L_{em}, G_e)	M	P	B	C	R	R_m	A	A_m	L	L_m	operon
$(0,0,1)$ $(0,1,1)$ $(1,1,1)$	0	0	0	0	1	1	0	0	0	0	OFF
$(0,0,0)$	0	0	0	1	1	1	0	0	0	0	OFF
$(1,1,0)$	1	1	1	1	0	0	0	1	0	1	ON
$(0,1,0)$	0	0	0	1	1	1	0	0	0	0	OFF
	1	1	1	1	0	0	0	1	0	1	ON

Suppose glucose or lactose are both absent ($L_e = L_{em} = G_e = 0$), so the operon is OFF:

$$(M, P, B, C, R, R_m, A, A_m, L, L_m) = (0,0,0,1,1,1,0,0,0,0).$$

Now, let's change L_{em} from 0 to 1, increasing lactose to medium. We are now in the next-to-last fixed point above, so the operon remains OFF.

Conversely, suppose lactose concentration is high ($L_e = L_{em} = 1$), and so the operon is ON:

$$(M, P, B, C, R, R_m, A, A_m, L, L_m) = (1,1,1,1,0,0,0,1,0,1).$$

Now, let's change L_e from 1 to 0, reducing lactose levels to medium. This takes us to the last fixed point above, so the operon remains ON.