Errors using inadequate data are much less than those using no data at all. Charles Babbage.

Artificial Biochemistry

Gene Networks

Luca Cardelli

Microsoft Research

The Microsoft Research - University of Trento
Centre for Computational and Systems Biology

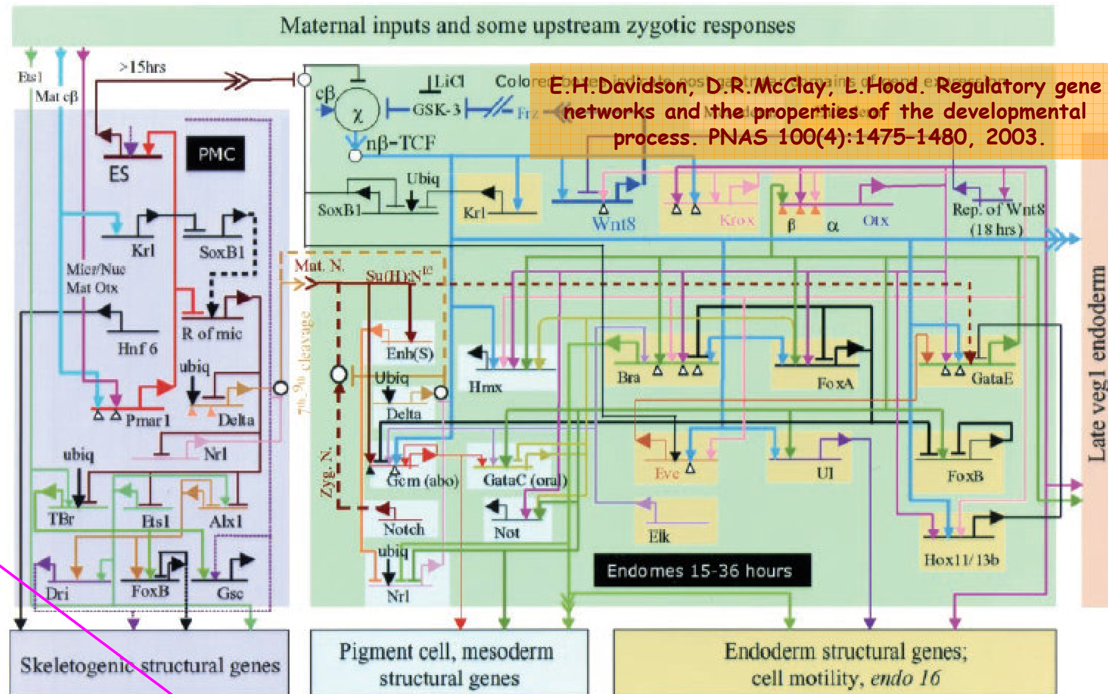
Trento, 2006-05-22..26

www.luca.demon.co.uk/ArtificialBiochemistry.htm

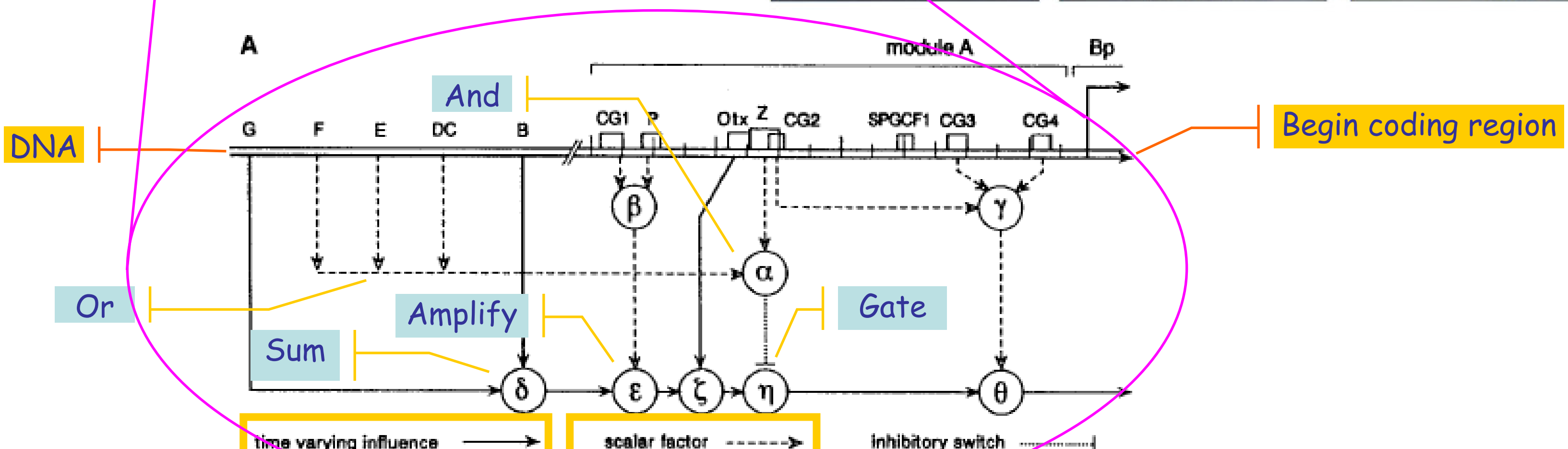
Gene Regulatory Networks

http://strc.herts.ac.uk/bio/maria/NetBuilder/

NetBuilder



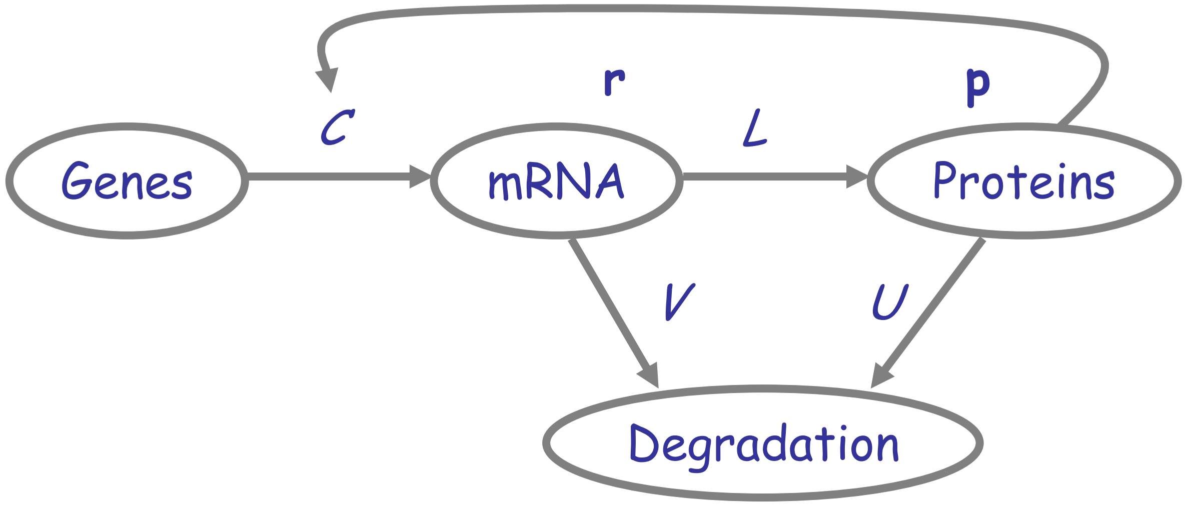
E.H.Davidson, D.R.McClay, L.Hood. Regulatory gene networks and the properties of the developmental process. PNAS 100(4):1475-1480, 2003.



C-H.Yuh, H.Bolouri, E.H.Davidson. Genomic Cis-Regulatory Logic: Experimental and Computational Analysis of a Sea Urchin Gene. Science 279:1896-1902, 1998

The Classical ODE Approach

[Chen, He, Church]



I.e.: to model an operating system, write a set of differential equations relating the concentrations in memory of data structures and stack frames over time. (Duh!)

$$\frac{d\mathbf{r}}{dt} = f(\mathbf{p}) - V\mathbf{r}$$

$$\frac{d\mathbf{p}}{dt} = L\mathbf{r} - U\mathbf{r}$$

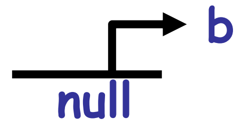
n: number of genes
r mRNA concentrations (n-dim vector)
p protein concentrations (n-dim vector)

$f(\mathbf{p})$ transcription functions:
(n-dim vector polynomials on **p**)

Gene Gates

Nullary Gate

no input



spontaneous ("constitutive") output

interaction site of output protein

$$\mathrm{null}(b) \triangleq \tau_{\varepsilon};\ (\mathrm{tr}(b) \mid \mathrm{null}(b))$$

(recursive, parametric) process definition

and repeat

stochastic delay (τ) with rate ε of constitutive transcription

output protein (transcription factor), spawn out

A stochastic rate r is always associated with each channel a_r (at channel creation time) and delay τ_r, but is often omitted when unambiguous.

Production and Degradation

Degradation is extremely important and often deliberate; it changes unbounded growth into (roughly) stable signals.

$$\mathrm{tr}(p) \triangleq (!p_r;\ \mathrm{tr}(p)) + \tau_\delta$$

transcription factor

interaction site of transcription factor

stochastic choice (race between r and δ)

and repeat

degradation

degradation rate δ

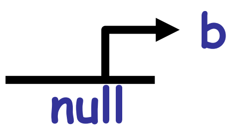
(output, !) interaction with rate r (input, ?, is on the target gene)

A transcription factor is a *process* (not a message or a channel): it has behavior such as interaction on p and degradation.

---

combined effect of production and degradation (without any interaction on b)

product

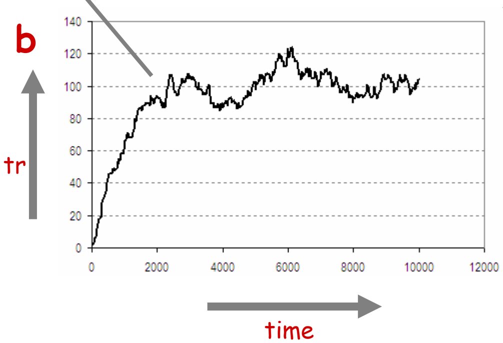


$$\mathrm{null}(b) \triangleq \tau_\varepsilon;\ (\mathrm{tr}(b)\ |\ \mathrm{null}(b))$$

null(b) ε=0.1, δ=0.001

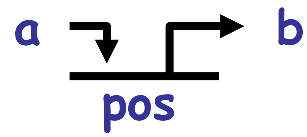
interaction offers on b

(= number of tr processes)



Unary Pos Gate

input (excitatory)



output (stimulated or constitutive)

$pos(a,b) \triangleq$
$?a_r; \tau_\eta; (tr(b) \mid pos(a,b)) +$
$\tau_\varepsilon; (tr(b) \mid pos(a,b))$

transcription delay with rate η

(input, ?) interaction with rate r

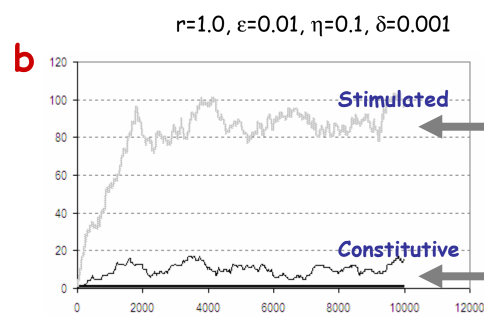
race between r and ε

or constitutive transcription to always get things started

output protein

parallel, not sequence, to handle self-loops without deadlock

unlimited amount of



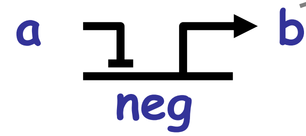
$*tr(a_r) \mid pos(a_r,b)$

$pos(a,b)$

Unary Neg Gate

input (inhibitory)

output (constitutive when not inhibited)



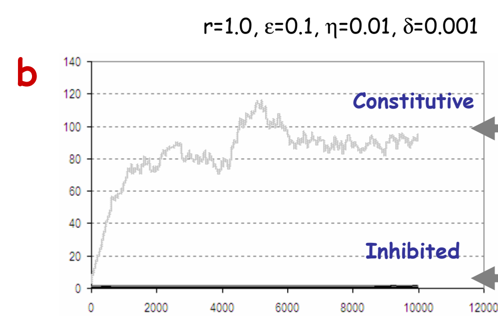
(input, ?) interaction with rate r

or constitutive transcription to always get things started

inhibition delay with rate η

race between r and ε

$$\begin{aligned} &neg(a,b) \triangleq \\ &\quad ?a_r;\ \tau_\eta;\ neg(a,b)\ + \\ &\quad \tau_\varepsilon;\ (tr(b) \mid neg(a,b)) \end{aligned}$$

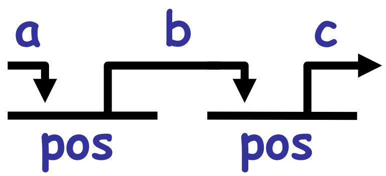


$neg(a_r,b)$

$*tr(a_r) \mid neg(a_r,b)$

Signal Amplification

pos(a,b) | pos(b,c)

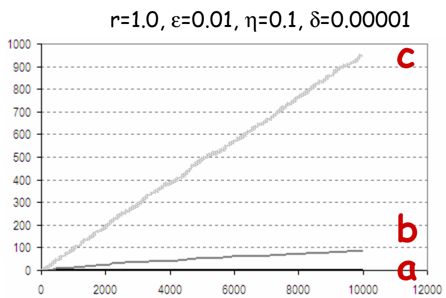


$\mathrm{pos}(a,b) \triangleq$
$\quad ?a_r;\ \tau_\eta;\ (\mathrm{tr}(b) \mid \mathrm{pos}(a,b)) +$
$\quad \tau_\varepsilon;\ (\mathrm{tr}(b) \mid \mathrm{pos}(a,b))$

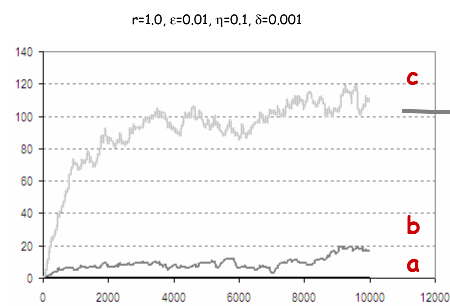
$\mathrm{tr}(p) \triangleq (!p_r;\ \mathrm{tr}(p)) + \tau_\delta$

E.g. 1 a that interacts twice before decay can produces 2 b that each interact twice before decay, which produce 4 c…

With little degradation



pos(a,b) | pos(b,c)

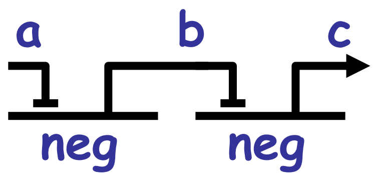


pos(a,b) | pos(b,c)

even with no a input, constitutive production of b gets amplified to a high c signal

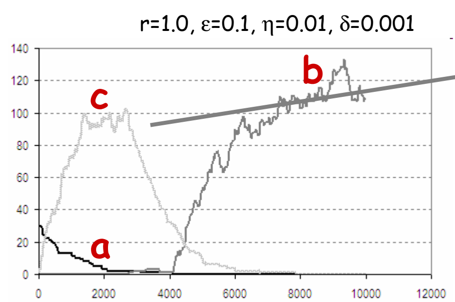
Signal Normalization

neg(a,b) |
neg(b,c)



$$neg(a,b) \triangleq ?a_r;\ \tau_h;\ neg(a,b) + \tau_\varepsilon;\ (tr(b)\ |\ neg(a,b))$$

$$tr(p) \triangleq (!p_r;\ tr(p)) + \tau_\delta$$

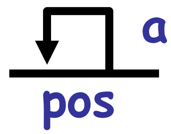


30*tr(a) | neg(a,b) | neg(b,c)

a non-zero input level, a, whether weak or strong, is renormalized to a standard level, c.

Self Feedback Circuits

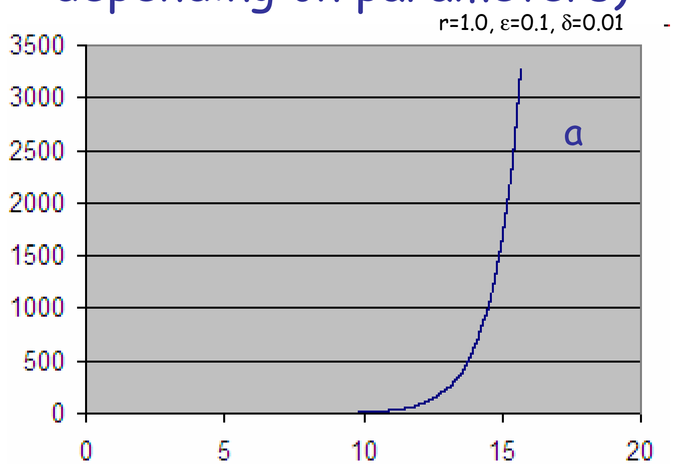
pos(a,a)



$$pos(a,b) \triangleq$$
$$?a_r;\ (tr(b) \mid pos(a,b)) +$$
$$\tau_\varepsilon;\ (tr(b) \mid pos(a,b))$$

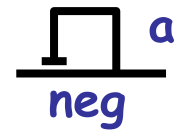
$$tr(p) \triangleq (!p_r;\ tr(p)) + \tau_\delta$$

(Can overwhelm degradation, depending on parameters)



pos(a,a)

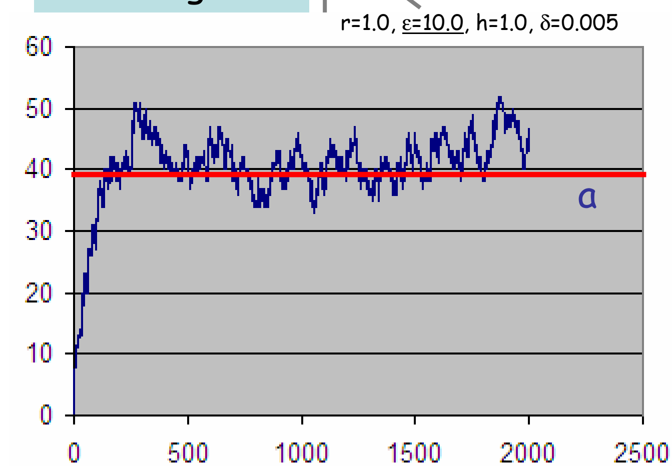
neg(a,a)



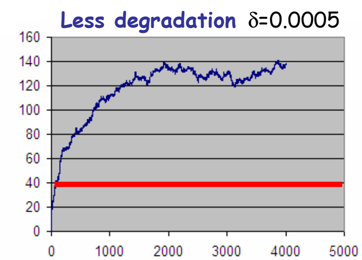
$$neg(a,b) \triangleq$$
$$?a_r;\ \tau_h;\ neg(a,b) +$$
$$\tau_\varepsilon;\ (tr(b) \mid neg(a,b))$$

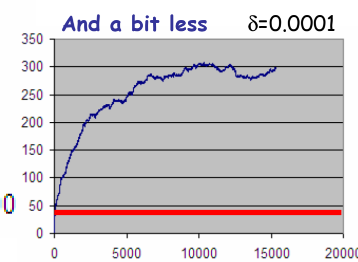
$$tr(p) \triangleq (!p_r;\ tr(p)) + \tau_\delta$$

high, to raise the signal



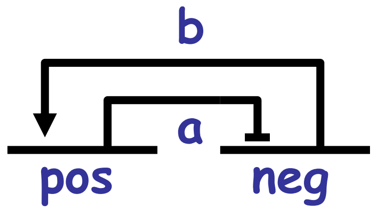
neg(a,a)



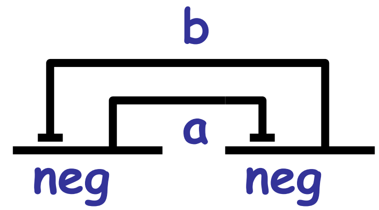


Two-gate Feedback Circuits

pos(b,a) | neg(a,b)

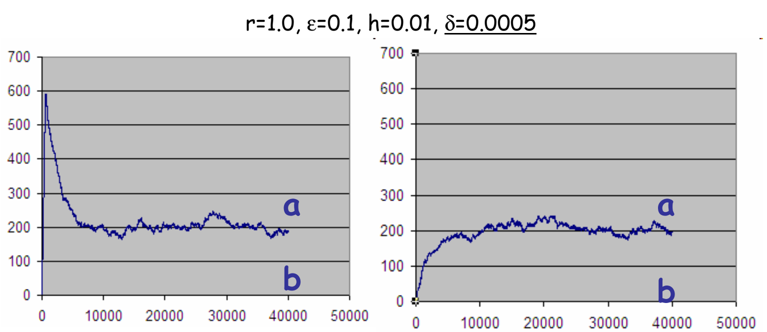


neg(b,a) | neg(a,b)



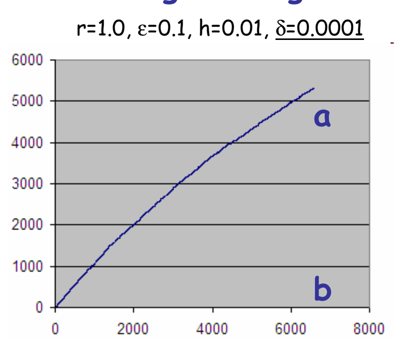
Monostable:

For some degradation rates is quite stable:



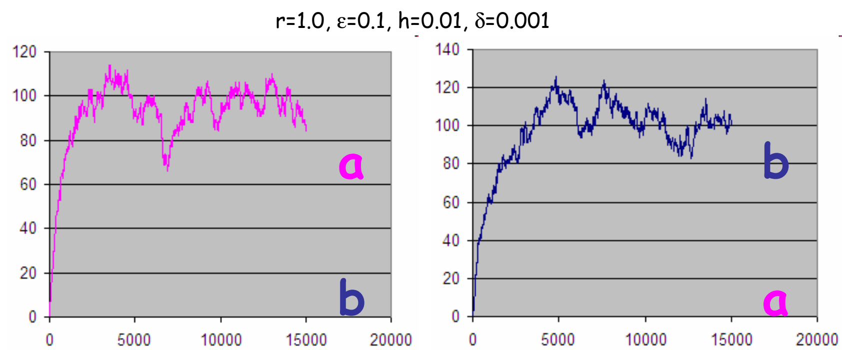
pos(b,a) | neg(a,b)

But with a small change in degradation, it goes wild:

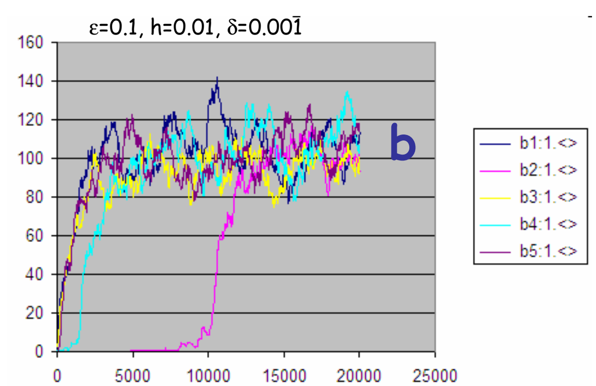


pos(b,a) | neg(a,b)

Bistable:



neg(b,a) | neg(a,b)

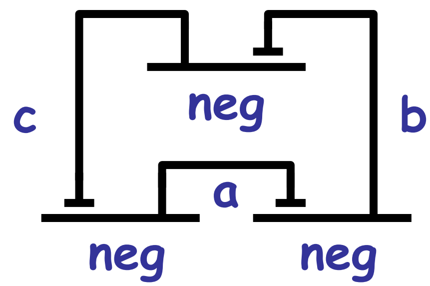


5 runs with r(a)=0.1, r(b)=1.0 shows that circuit is now biased towards expressing b

Repressilator

Repressilator

$neg(a,b)\ |$
$neg(b,c)\ |$
$neg(c,a)$



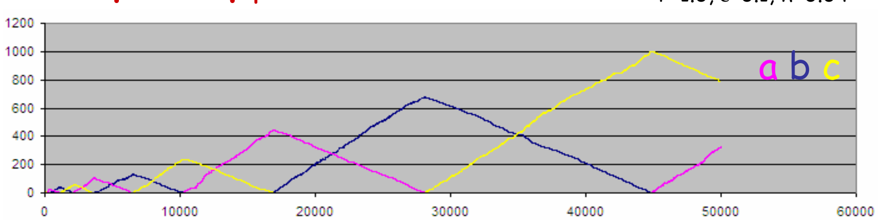
$neg(a,b) \triangleq$
$?a_r;\ \tau_h;\ neg(a,b)\ +$
$\tau_\varepsilon;\ (tr(b)\ |\ neg(a,b))$

Same circuit, three different degradation models by changing the tr component:

$tr(p) \triangleq\ !p_r$

interact once and die
otherwise stick around

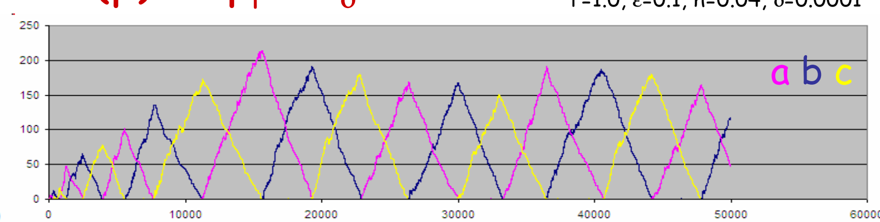
r=1.0, ε=0.1, h=0.04



$tr(p) \triangleq\ !p_r + \tau_\delta$

interact once and die
otherwise decay

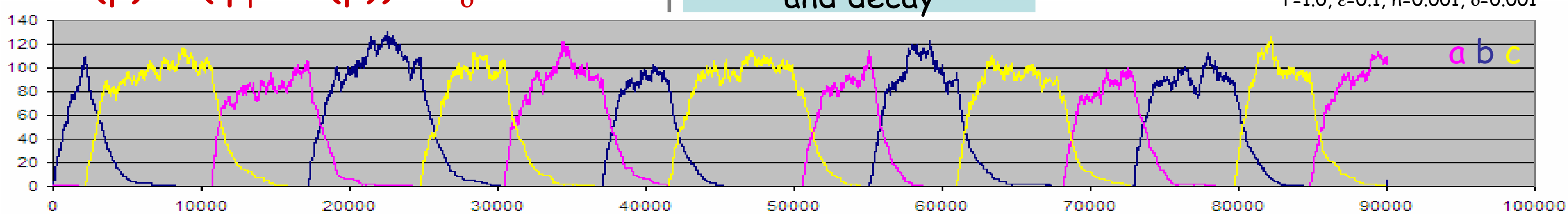
r=1.0, ε=0.1, h=0.04, δ=0.0001



$tr(p) \triangleq (!p_r;\ tr(p)) + \tau_\delta$

interact many times
and decay

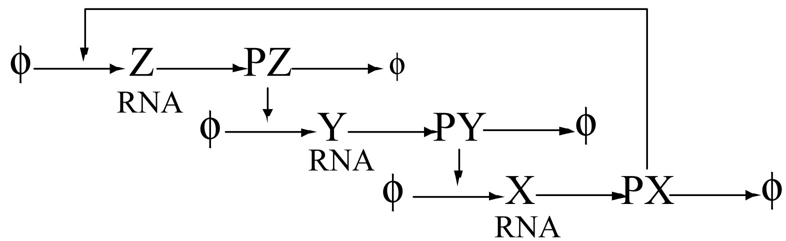
r=1.0, ε=0.1, h=0.001, δ=0.001



Subtle… at any point one gate is inhibited and the other two can fire constitutively. If one of them fires first, nothing really changes, but if the other one fires first, then the cycle progresses.

Repressilator ODE Model and Simulation

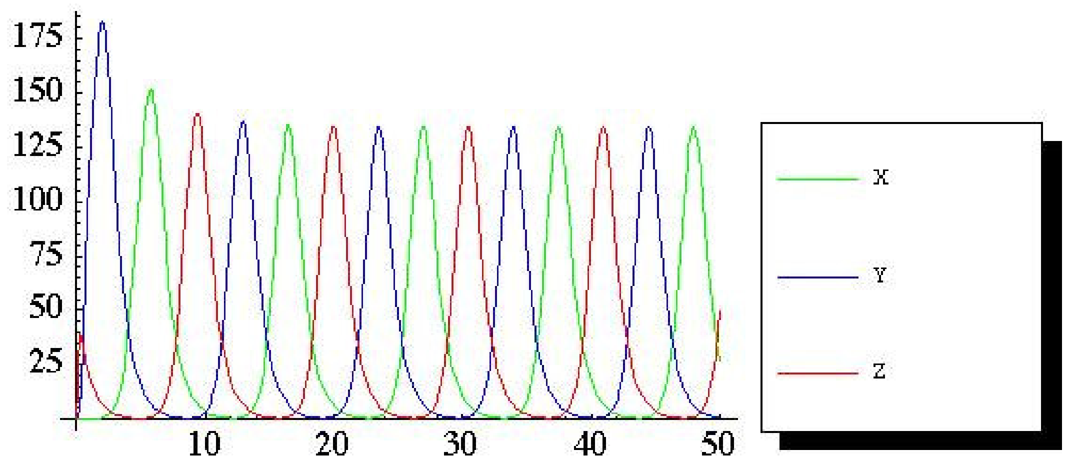
Bruce E Shapiro
Cellerator



$$\frac{d[\mathrm{X}]}{dt}=\alpha_0+\frac{\alpha+\alpha_1[\mathrm{PY}]^n}{K^n+[\mathrm{PY}]^n}-k[\mathrm{X}],\quad \frac{d[\mathrm{PX}]}{dt}=\beta\{[\mathrm{X}]-[\mathrm{PX}]\}$$

$$\frac{d[\mathrm{Y}]}{dt}=\alpha_0+\frac{\alpha+\alpha_1[\mathrm{PZ}]^n}{K^n+[\mathrm{PZ}]^n}-k[\mathrm{Y}],\quad \frac{d[\mathrm{PY}]}{dt}=\beta\{[\mathrm{Y}]-[\mathrm{PY}]\}$$

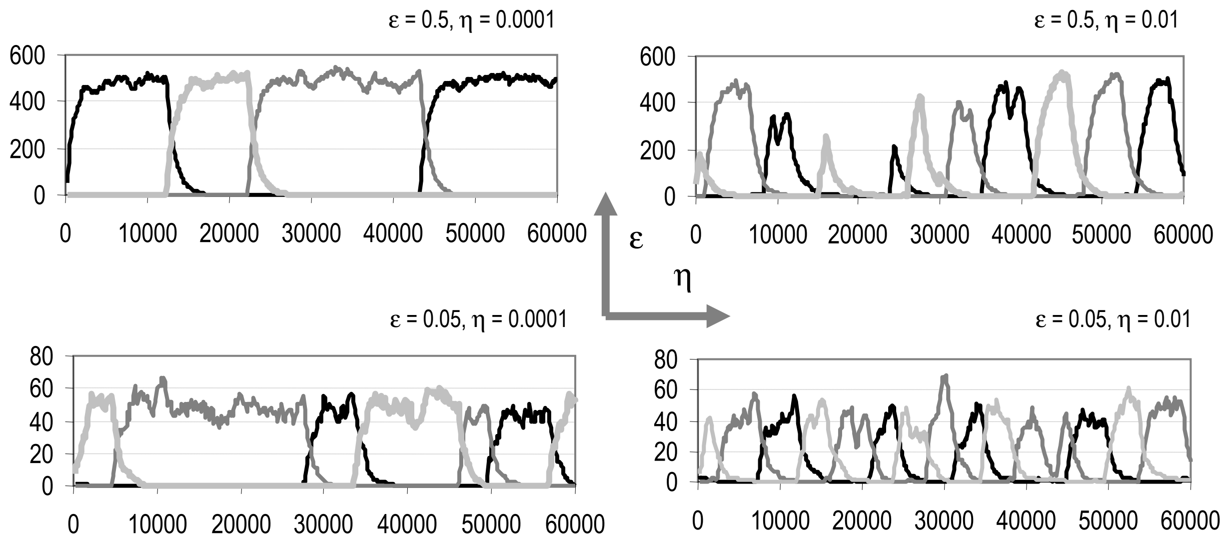
$$\frac{d[\mathrm{Z}]}{dt}=\alpha_0+\frac{\alpha+\alpha_1[\mathrm{PX}]^n}{K^n+[\mathrm{PX}]^n}-k[\mathrm{Z}],\quad \frac{d[\mathrm{PZ}]}{dt}=\beta\{[\mathrm{Z}]-[\mathrm{PZ}]\}$$



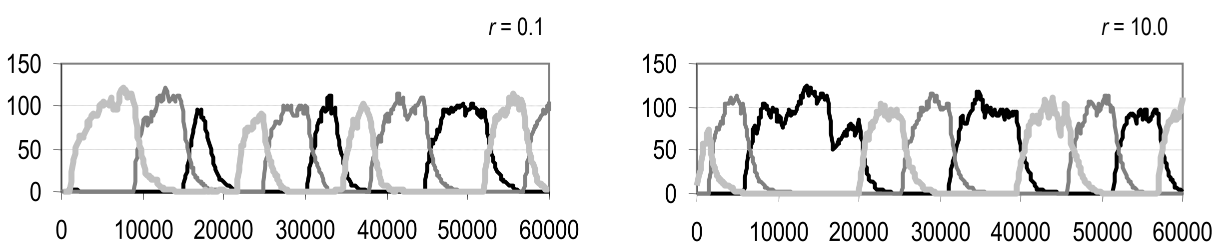
Repressilator in SPiM

```
val dk = 0.001      (* Decay rate *)
val eta = 0.001     (* Inhibition rate *)
val cst = 0.1       (* Constitutive rate *)

let tr(p:chan()) =
  do !p; tr(p)
  or delay@dk

let neg(a:chan(), b:chan()) =
  do ?a; delay@eta; neg(a,b)
  or delay@cst; (tr(b) | neg(a,b))

(* The circuit *)
val bnd = 1.0       (* Protein binding rate *)
new a@bnd: chan()
new b@bnd: chan()
new c@bnd: chan()

run (neg(c,a) | neg(a,b) | neg(b,c))
```

System Properties: Oscillation Parameters



The constitutive rate ε (together with the degradation rate) determines oscillation amplitude, while the inhibition rate η determines oscillation frequency.



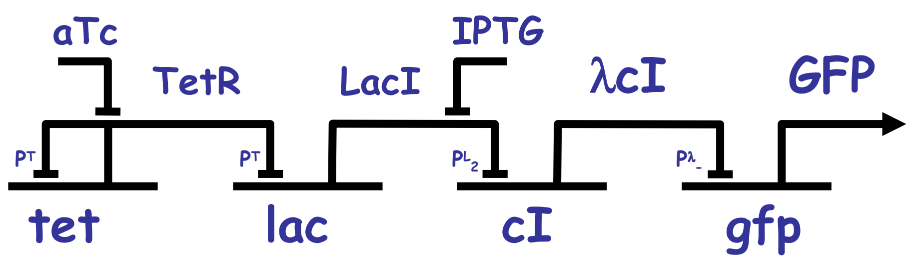
We can view the interaction rate *r* as a measure of the volume (or temperature) of the solution; that is, of how often transcription factors bump into gates. Oscillation frequency and amplitude remain unaffected in a large range of variation of *r*.

Guet et al.

Guet et al.

Combinatorial Synthesis of Genetic Networks, **Guet, Elowitz, Hsing, Leibler, 1996,** *Science*, May 2002, 1466-1470.

They engineered in E.Coli all genetic circuits with four single-input gates; such as this one:



Then they measured the GFP output (a fluorescent protein) in presence or absence of each of two inhibitors (aTc and IPTG).

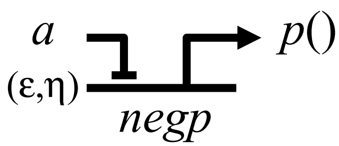
Experiment:	
aTc	0101
IPTG	0011
GFP	0100

The output of some circuits did not seem to make any sense…

Here "1" means "high brightness" and "0" means "low brightness" on a population of bacteria after some time. (I.e. integrated in space and time.)

Further Building Blocks

Negp Gate



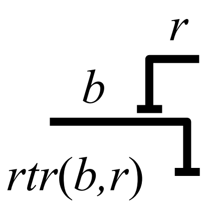
regulatory input

rates

product

$$negp(a,(\varepsilon,\eta),p) = ?a.\ \tau_{\eta}.\ negp(a,(\varepsilon,\eta),p) + \tau_{\varepsilon}.\ (p() \mid negp(a,(\varepsilon,\eta),p))$$

product generation



repressible factor

$$rtr(b,r) = !b.\ rtr(b,r) + !r.\ 0 + \tau_{\delta}.\ 0$$

interaction — binding

interaction — repression

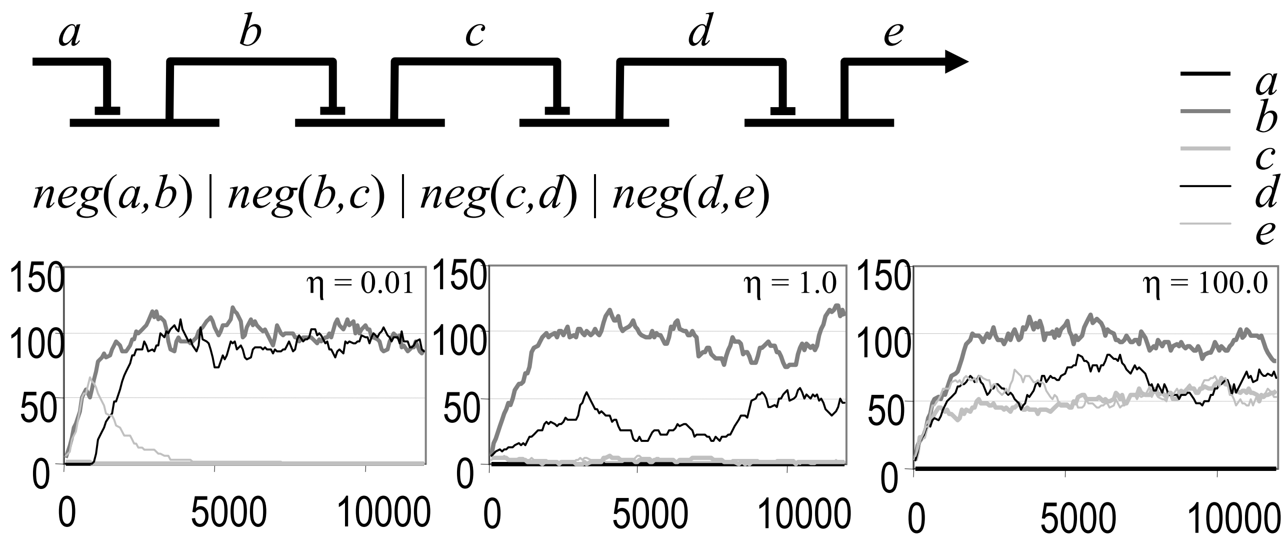
delay — degradation

repressor

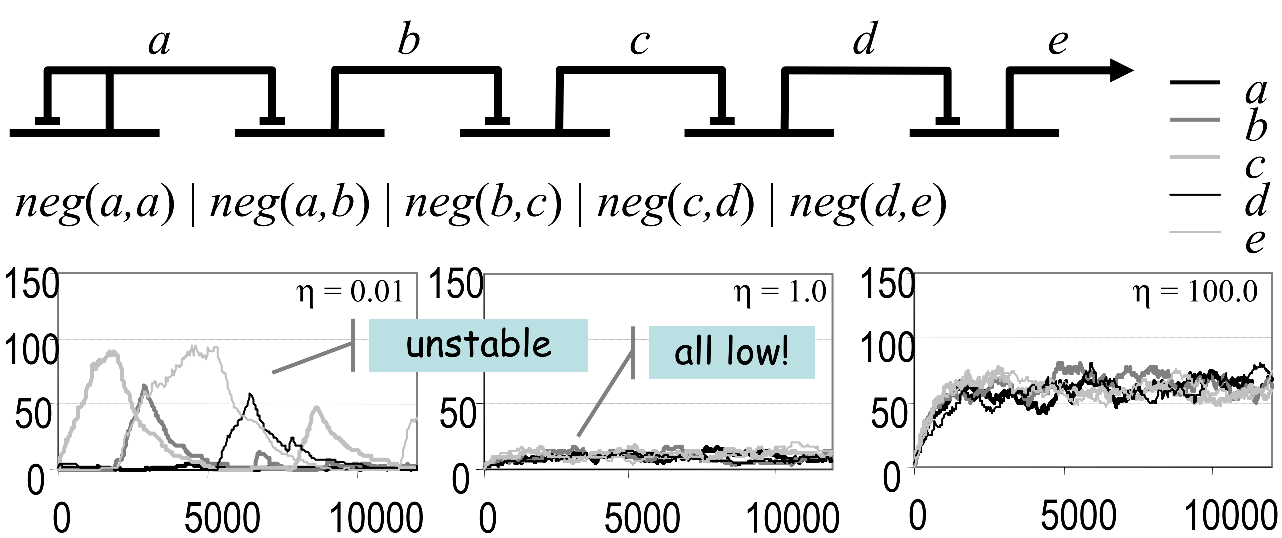
$$rep(r) = ?r.\ rep(r)$$

arbitrary amounts of..

System Properties: Fixpoints



A sequence of neg gates behaves as expected, with alternating signals, (less "Booleanly" depending on attenuation).

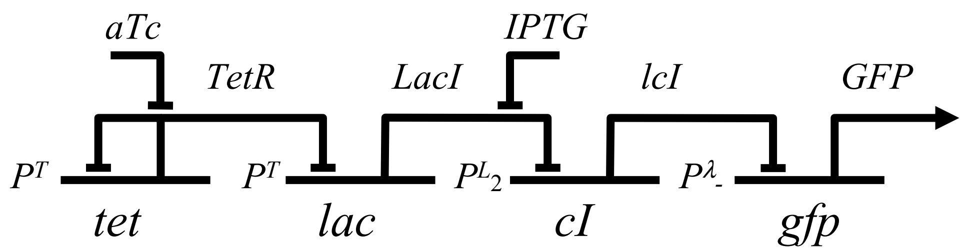


Now add a self-loop at the head. Not a Boolean circuit!

No more alternations, because... each gate is at its fixpoint.

D038/lac⁻

D038/lac⁻



Experment:	
aTc	0101
IPTG	0011
GFP	0100

Naïve "Boolean" analysis would suggest GFP=0.5 (oscillation) because of self-loop.

GFP=0 is consistent only with (somehow) the head loop setting TetR=LacI=0 (they have the same promoter P^T). But in that case, aTc should have no effect (it can only subtract from those signals) but instead adding aTc sets GFP=1.

Hence we need to understand better the "dynamics" of this network.

<u>*channels*</u> $\boxed{TetR{:}\mathbf{r_1},\ LacI{:}\mathbf{r_2},\ lcI{:}\mathbf{r_3},\ GFP{:}\mathbf{r_4},\ aTc{:}\mathbf{r_5},\ IPTG{:}\mathbf{r_6}}$

molecules

$$\boxed{P^T = (\varepsilon_1, \eta_1) \quad P^L{}_2 = (\varepsilon_2, \eta_2) \quad P^\lambda{}_- = (\varepsilon_3, \eta_3)}$$

promoters

$$tet = negp(TetR,\ P^T,\ rtr(TetR, aTc))$$
$$lac = negp(TetR,\ P^T,\ rtr(LacI, IPTG))$$
$$cI = negp(LacI,\ P^L{}_2,\ tr(lcI))$$
$$gfp = negp(lcI,\ P^\lambda{}_-,\ tr(GFP))$$

genes

repressors (when present)

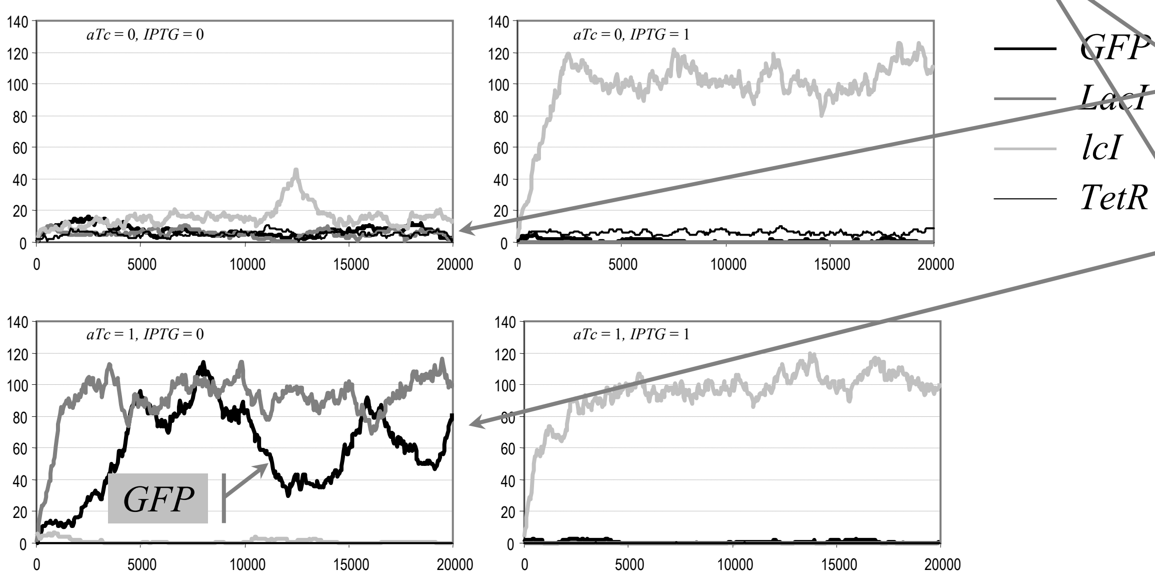
$$D038lac^- = \boxed{tet \mid lac \mid cI \mid gfp} \ \mid\mid\ \boxed{rep(aTc) \mid rep(IPTG)}$$

Simulation results for D038/lac$^-$

Experiment:

aTc	0101
IPTG	0011
GFP	0100

r=1.0, ε=0.1, h=1.0, δ=0.001

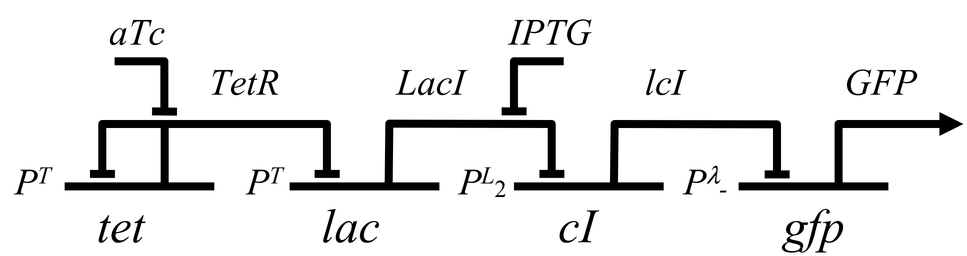


The fixpoint effect can explain this (all signals set very low).

Then, aTc can destabilize the fixpoint, explaining GFP high (oscillating)

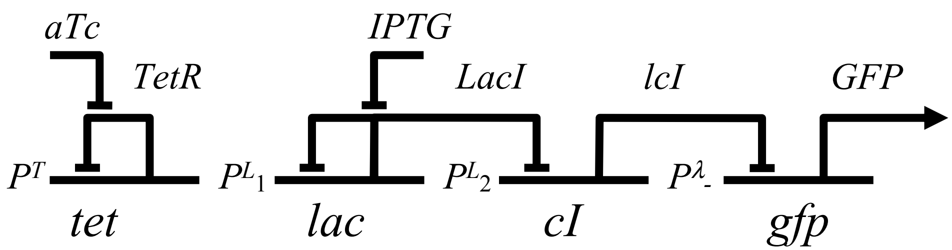
$r = 1.0$, $\varepsilon = 0.1$, $\eta = 0.25$ (P^T), $\eta = 1.0$ ($P^L{}_2$, $P^\lambda{}_-$), $\delta = 0.001$

D038/lac$^-$



D016/lac⁻

D016/lac⁻



Experiment:

aTc	0101
IPTG	0011
GFP	1000

How can aTc affect the result??

One theory: aTc prevents the self-inhibition of tet, so that a very large quantity of TetR is produced. That then overloads the overall degradation machinery of the cell, affecting the rest of the circuit.

Even so, how can GFP be high here?

Even the fixpoint explanation fails here, unless we assume that the lac gate is operating in its instability region.

$\underline{channels}$ $\boxed{TetR{:}\mathbf{r_1},\ LacI{:}\mathbf{r_2},\ lcI{:}\mathbf{r_3},\ GFP{:}\mathbf{r_4},\ aTc{:}\mathbf{r_5},\ IPTG{:}\mathbf{r_6}}$

$$\boxed{P^T = [\varepsilon_1, \eta_1] \quad P^L{}_2 = [\varepsilon_2, \eta_2] \quad P^\lambda{}_- = [\varepsilon_3, \eta_3] \quad P^L{}_1 = [\varepsilon_4, \eta_4]}$$

promoters

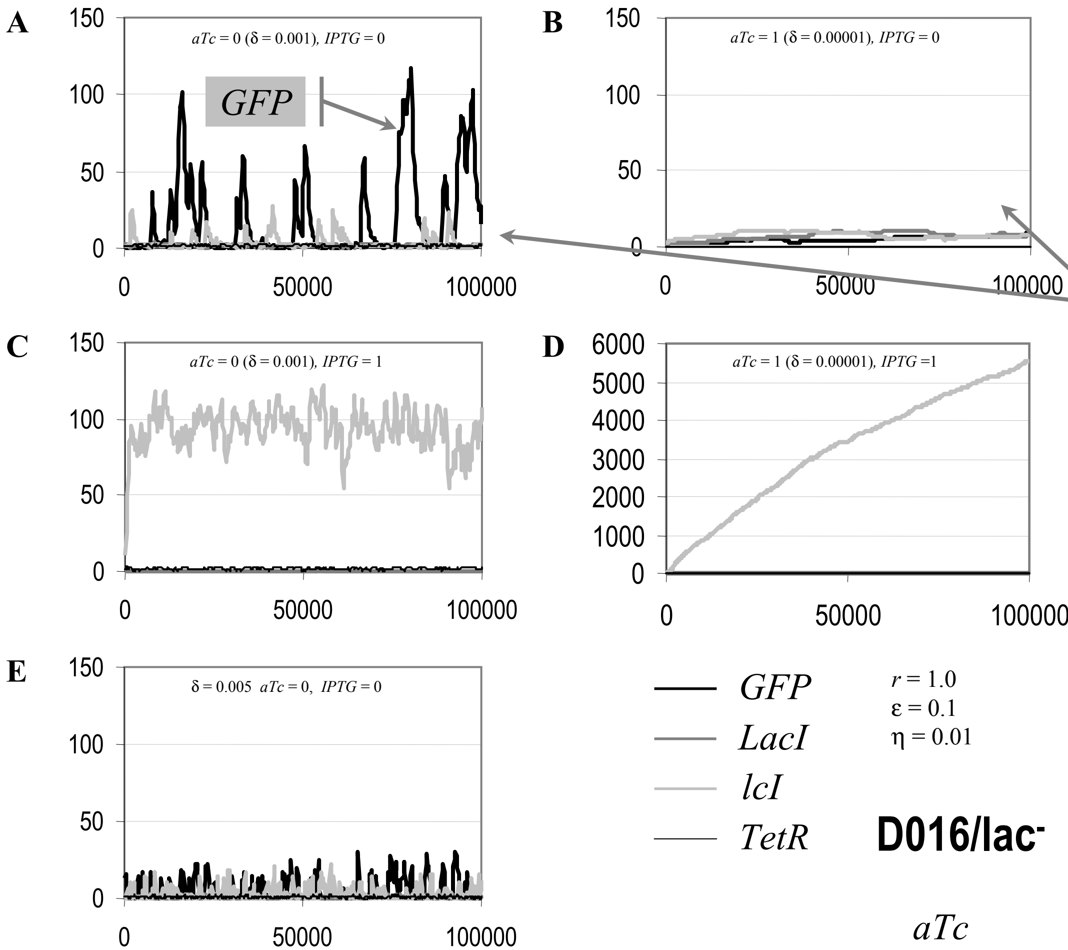
$$tet = negp[TetR, P^T, rtr[TetR,aTc]]$$
$$lac = negp[LacI, P^L{}_1, rtr[LacI,IPTG]]$$
$$cI = negp[LacI, P^L{}_2, tr[lcI]]$$
$$gfp = negp[lcI, P^\lambda{}_-, tr[GFP]]$$

genes

repressors

$$D016lac^- = \boxed{tet \mid lac \mid cI \mid gfp} \mid \boxed{rep[aTc] \mid rep[IPTG]}$$

Simulation results for D016/lac$^-$



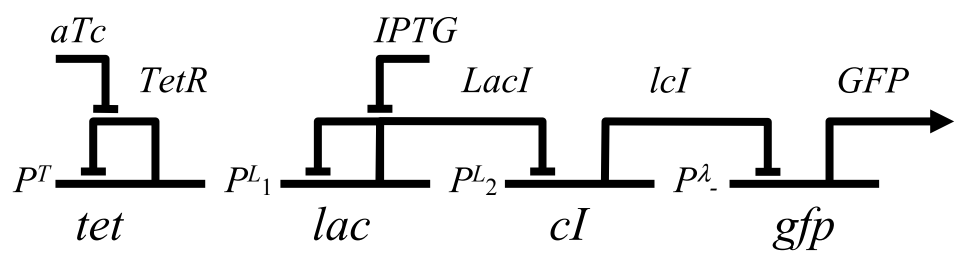
Experiment:

aTc	0101
IPTG	0011
GFP	1000

The fixpoint effect, in instability region, explains this: GFP high because wildly oscillating.

Overloading of degradation machinery, induced by aTc, can reinstate the fixpoint regime.

D016/lac$^-$



Summary

- Combinatorial components
  - A "library" of gates that can be used to build circuits.
- Repressilator
  - A first example of engineered genetic circuits.
- Combinatorial circuits
  - Trying to analyze the surprising cases.
- What was the point?
  - Deliberately pick a controversial/unsettled example to test the methodology.
  - Show that we can easily "play with the model" and run simulations.
  - Get a feeling for the kind of subtle effects that may play a role.
    - In particular, stochastic effects (wild oscillations) seem essential to some explanations.
  - Get a feeling for kind of analysis that is required to understand the behavior of these systems.
  - Building theories/models that support of contradict experiments (and that suggest further experiments).

Q?