Computing in biomolecular systems

Mikhail Ivanchenko
Lobachevsky State University of Nizhny Novgorod

Contents

- Introduction: from bio to maths
- Mathematical models of gene regulation
- Dynamics of gene regulation
- Logic circuits
- Synthetic gene classifier

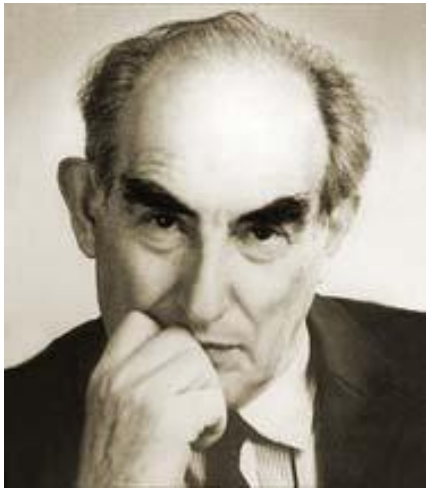
Апрель 1999 г. *Том 169, № 4*

УСПЕХИ ФИЗИЧЕСКИХ НАУК

ФИЗИКА НАШИХ ДНЕЙ

Какие проблемы физики и астрофизики представляются сейчас особенно важными и интересными (тридцать лет спустя, причем уже на пороге XXI века)?

В.Л. Гинзбург

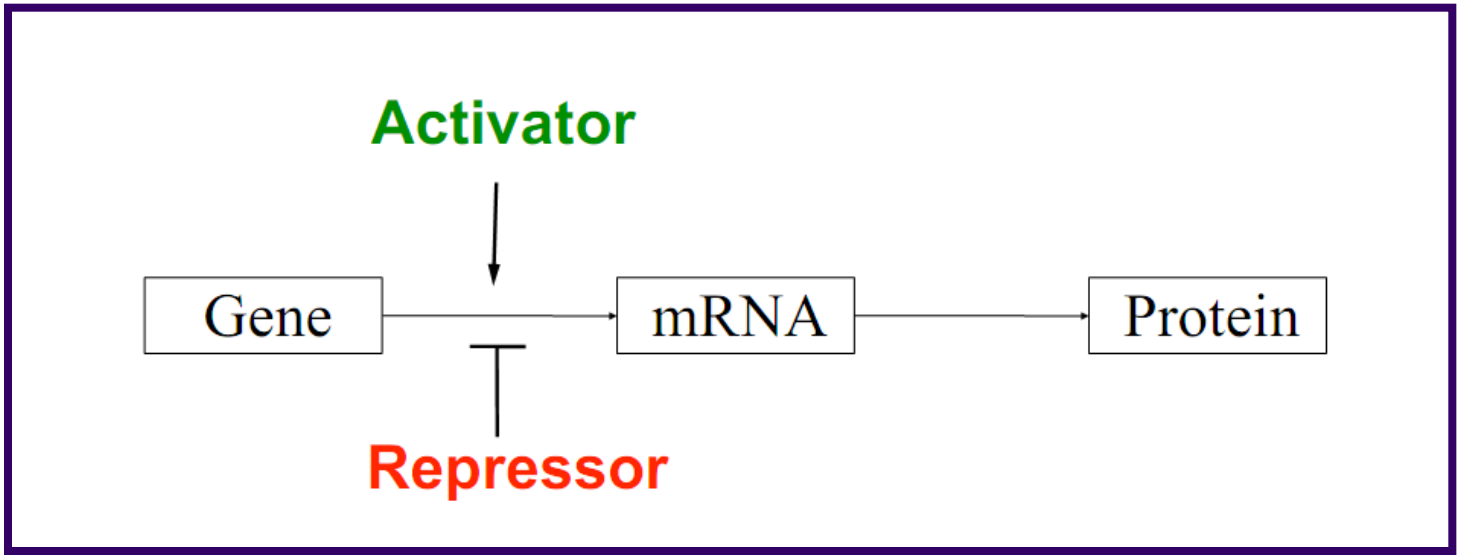


30 + 3 «великих» проблемы науки XXI века

Последняя "великая" проблема, которая будет здесь затронута, касается связи физики с биологией. С конца

основе физики, уже известной физики. Конкретно, основными являются вопросы о происхождении жизни и появлении сознания (мышления). Образование в усло-

«Newton's Law» for living systems

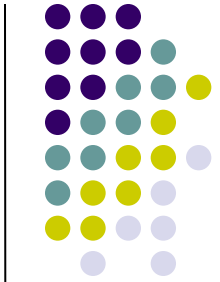


Transcriptional regulation:

Activator - increases reaction rate

Repressor - decreases reaction rate

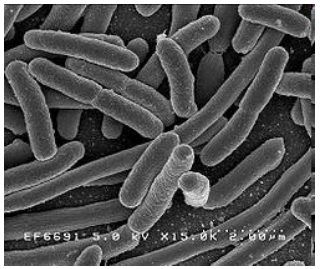
Dynamics



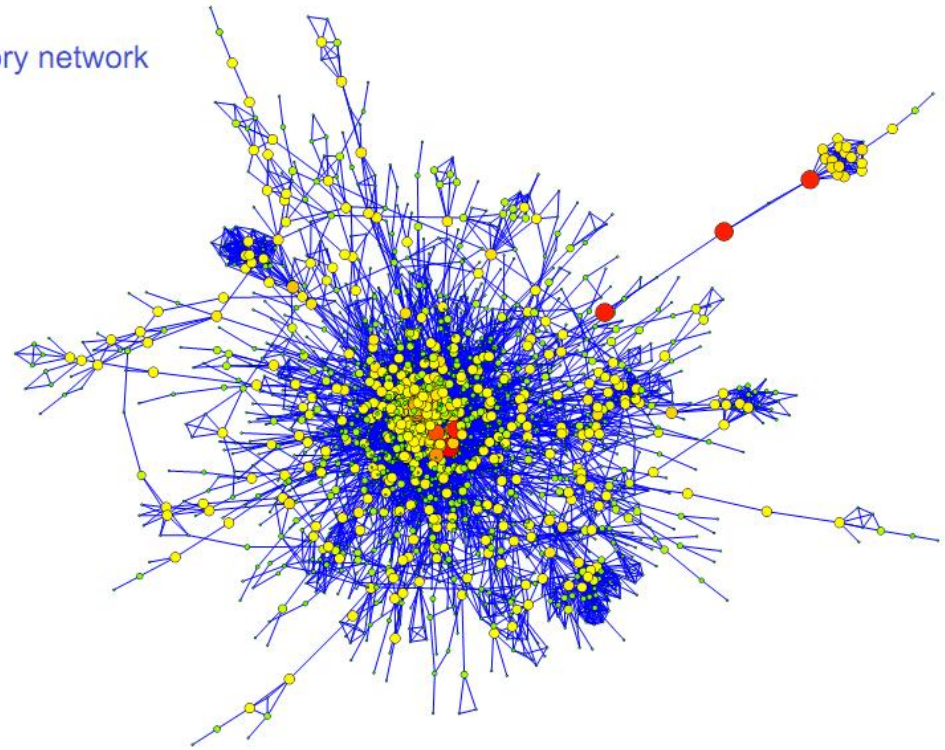
Gene expression is never stationary:

- External signals
- Extrinsic noise
- Intrinsic noise
- Oscillations (circadian and ultradian clocks, cell cycle, etc.)

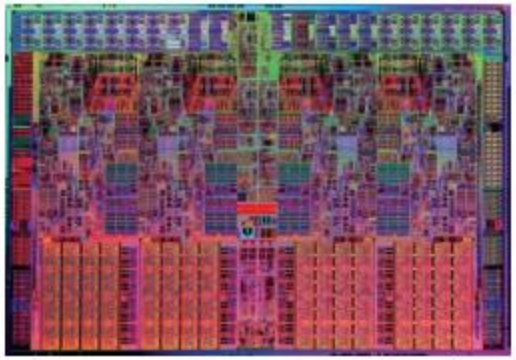
Gene regulation networks



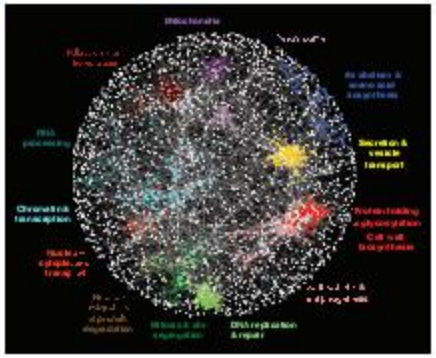
E.Coli regulatory network



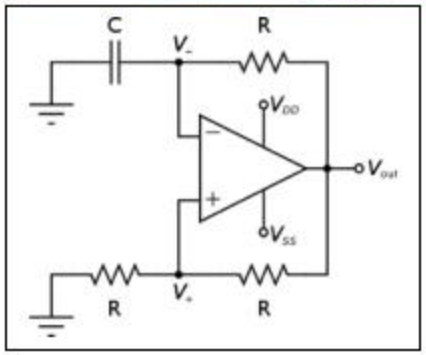
- «read» all genes (genomics)
- determine gene and protein interactions (network reconstruction)
- network dynamics (complex systems)

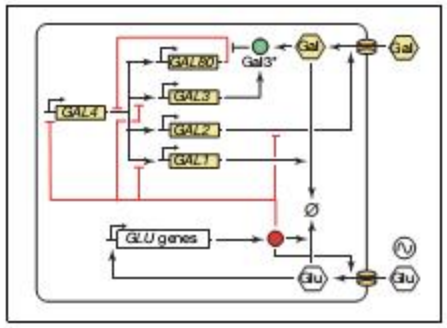


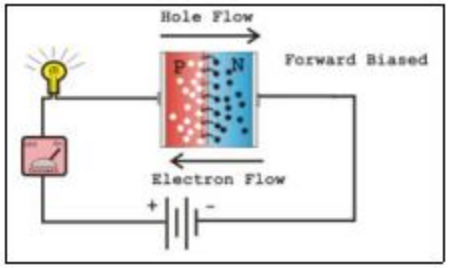
Systems Biology



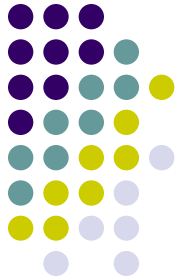
scale





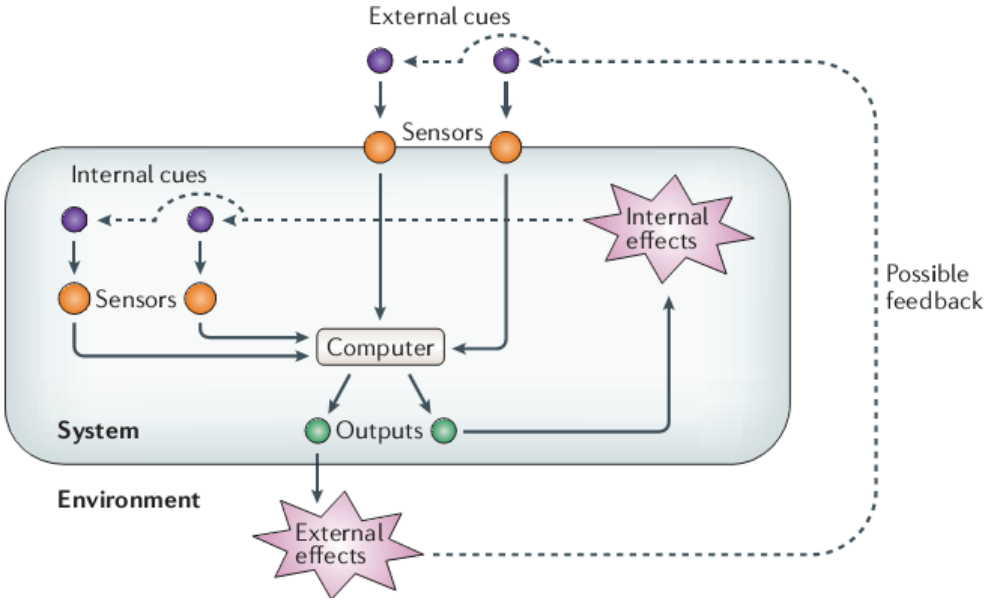


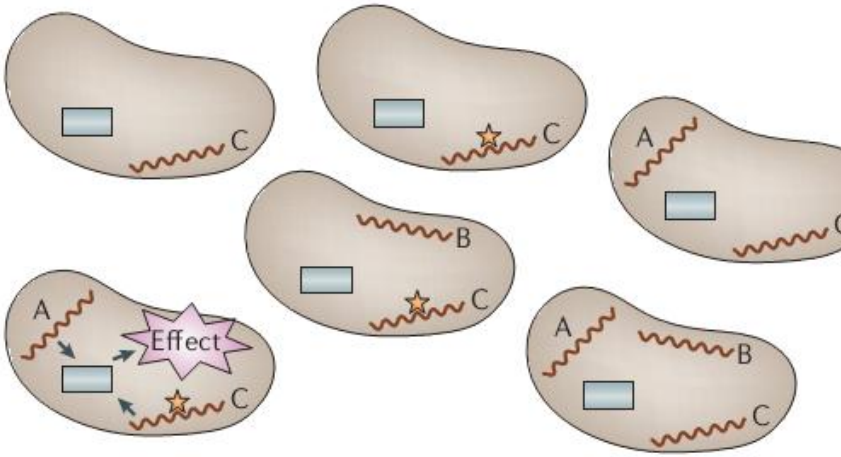
Synthetic Biology



- Do gene networks solve computational tasks?
- Can one construct gene networks for computing?

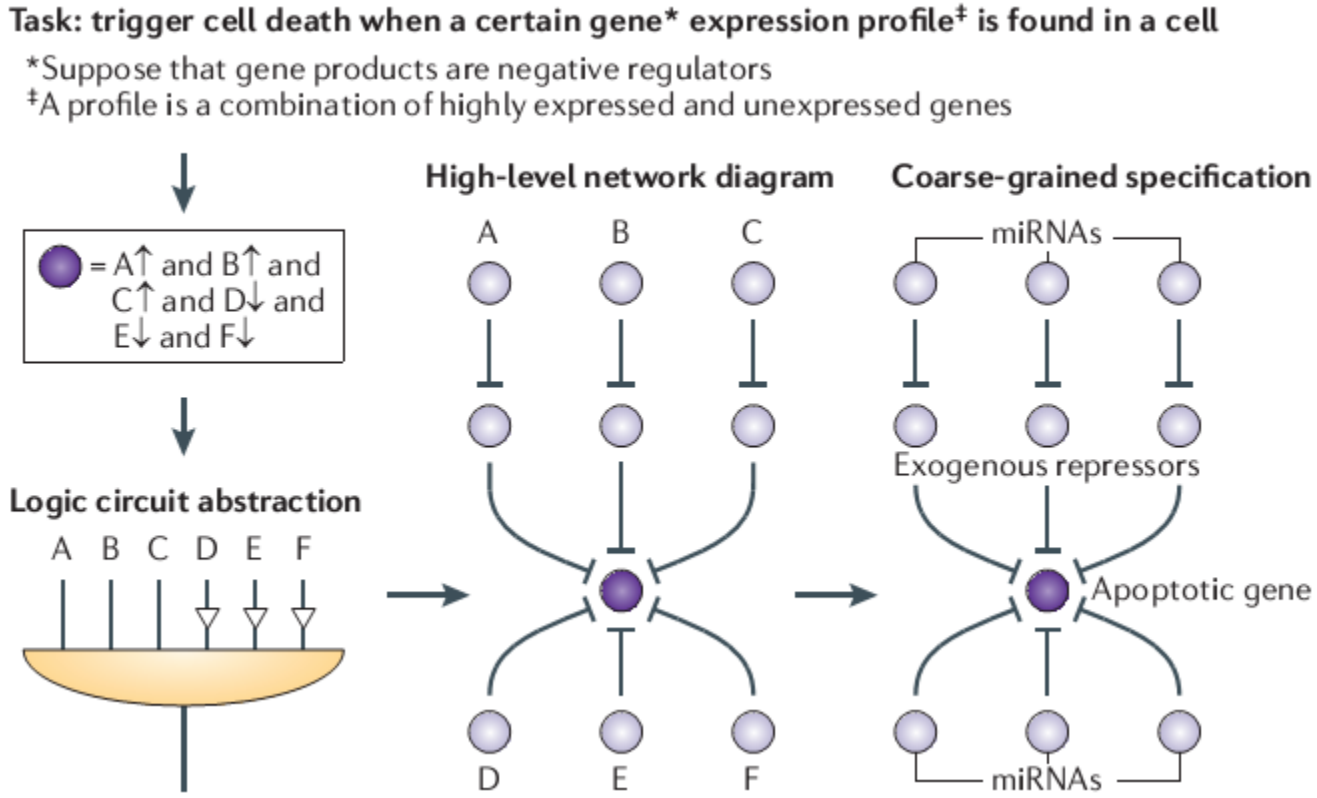
Do regulatory networks solve computational tasks?



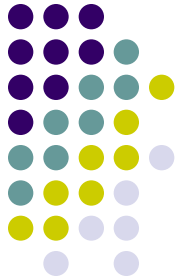


Yes, in the sense of information processing

Can one construct gene networks for computing?

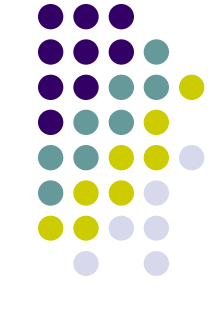


Devil in details: find components!



Mathematical models of gene regulation

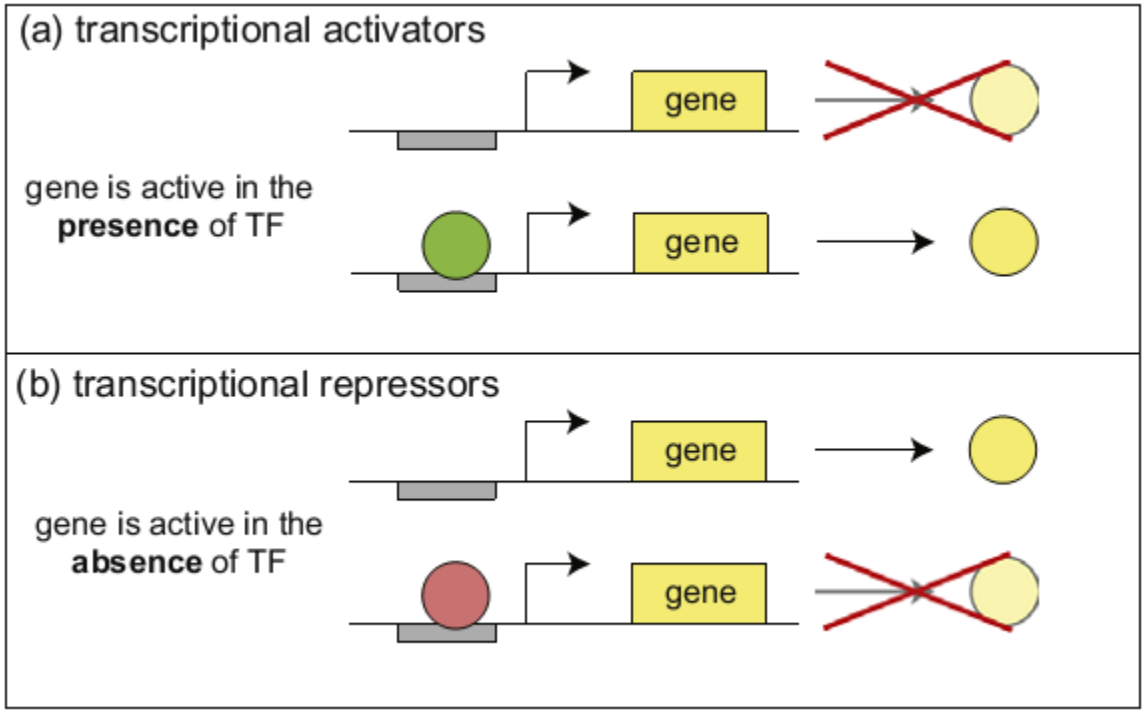
Transcriptional regulation and dynamics



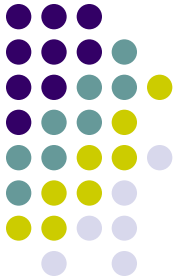
• Let TF be active only as a dimer.
• Let P be a concentration of a monomer TF, P2 — concentration of a dimer,
O_u — concentration of unbound operator sites ,
O_b — of bound sites.
• Then for equilibrium DNA and TF concentrations one gets:

$$P + P \underset{k_{-d}}{\overset{k_d}{\rightleftharpoons}} P_2,$$

$$O_u + P_2 \underset{k_{-b}}{\overset{k_b}{\rightleftharpoons}} O_b,$$



Transcriptional regulation and dynamics



In the equilibrium state

$$k_d P^2 = k_{-d} P_2,$$
$$k_b O_u P_2 = k_{-b} O_b.$$

Since DNA concentration is constant, Therefore

$$O_u + O_b = N$$

$$O_u = \frac{N}{1 + (P/K)^2},$$

Hill equations

Hill constant

where $K = \sqrt{k_{-b} k_{-d} / k_b k_d}$

Hill coefficient (not necessarily =2, not even a round number, e.g. in case of simultaneous binding to several operons)

And similarly $$O_b = \frac{N(P/K)^2}{1 + (P/K)^2}$$

Transcriptional regulation and dynamics

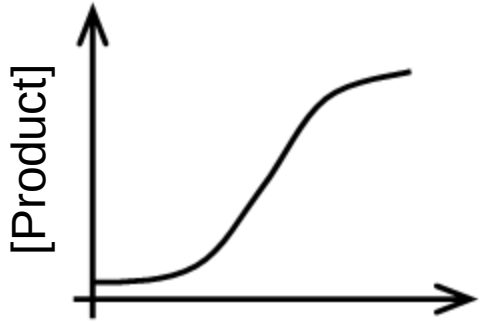
If gene promoter 'works' with rate α_u when unbounded by TF, and with rate α_b, when bound, then the net kinetic coefficient of gene expression reads

$$v_{init}(P) = \alpha_u O_u + \alpha_b O_b = \alpha_u \frac{N}{1+(P/K)^2} + \alpha_b \frac{N(P/K)^2}{1+(P/K)^2}.$$

$\alpha_b = 0$

[Transcriptional repressor]

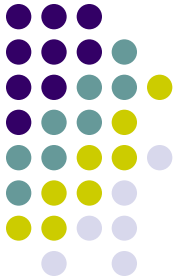
$\alpha_u = 0$



[Transcriptional activator]

Already a sense of a basic threshold (yes/no) detector!

Transcriptional regulation and dynamics



Taking into account protein degradation, we arrive at the differential equation for dynamics of protein concentration

$$\dot{x} = \alpha_u \frac{N}{1 + (P/K)^2} + \alpha_b \frac{N(P/K)^2}{1 + (P/K)^2} - r_{deg}(x)$$

Note:
In fact, the number of molecules is discrete and molecular events are discrete too. Hence ODE gives an approximate description only. A more precise description would employ random processes with probabilities proportional to kinetic coefficients. When the number of molecules is small or production events are rare, that model is much more adequate.

Example 1

- Let TF be active only as a **tetramer.**
- Let P be a concentration of a monomer TF, P2 — concentration of a dimer, O_u — concentration of unbound operator sites , O_b — of bound sites.
- Then for equilibrium DNA and TF concentrations one gets:

$$P + P + P + P \underset{k_{-d}}{\overset{k_d}{\rightleftharpoons}} P_4 ,$$

$$O_u + P_4 \underset{k_{-b}}{\overset{k_b}{\rightleftharpoons}} O_b ,$$

Find equation for the protein dynamics

In equilibrium

$$k_d P^4 = k_{-d} P_4 ,$$
$$k_b O_u P_4 = k_{-b} O_b .$$

$$O_u + O_b = N$$

Th

$\dot{x}$ =

$r_{deg}(x)$

В равновесном состоянии будет выполняться

$$k_d P^2 = k_{-d} P_2 ,$$
$$k_b O_u P_2 = k_{-b} O_b .$$

Раз суммарная концентрация ДНК постоянна, $O_u + O_b = N$

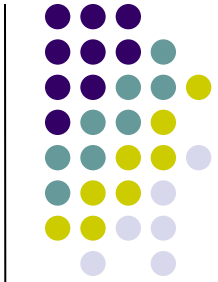
Отсюда

$$O_u = \frac{N}{1 + (P/K)^2} ,$$

где $K = \sqrt{k_{-b} k_{-d} / k_b k_d}$

и аналогично $O_b = \frac{N (P/K)^2}{1 + (P/K)^2} .$

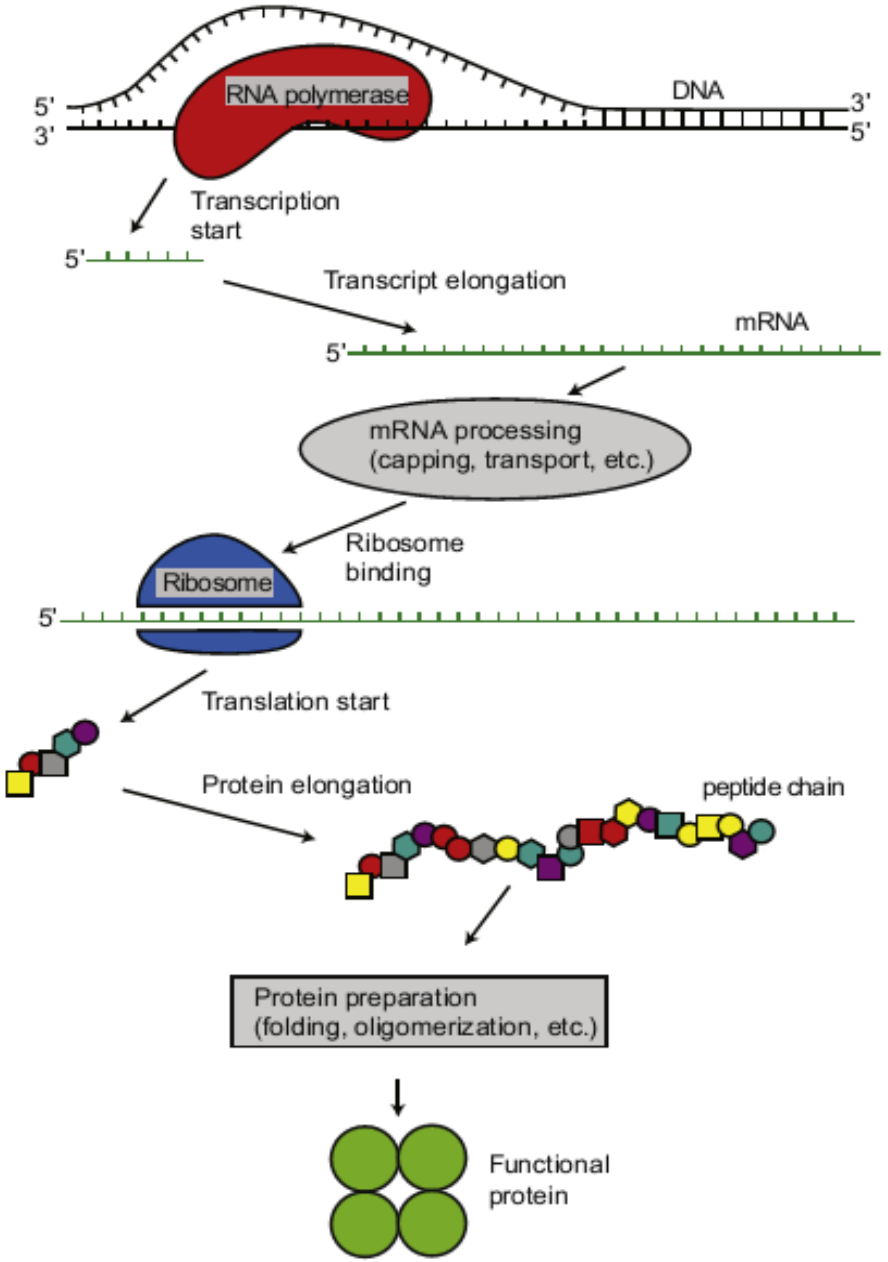
Transcriptional regulation and dynamics



Equations get more complicated if:

- A gene promotor consists of several operators (e.g., a tandem promotor of two operators: for a repressor TF and an activator TF)
- In eucaryotic cells TFs must enter cell nucleus first; coming in and out may add to dynamics
- Often TF are modulated by light active molecules – ligands. E.g. IPTG molecules can bind to TF-repressor LacI – sensitive promotor, and prevent binding of the latter. TF-activator AraC binds to the sensitive promotor only in presence of arabinose.

Detailed protein synthesis scheme



More detailed models

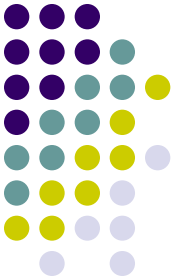
Taking into account mRNA synthesis as a separate process

$$\dot{m} = \frac{\alpha_u N}{1 + (P/K)^2} + \frac{\alpha_b N (P/K)^2}{1 + (P/K)^2} - r_m(m),$$
$$\dot{x} = r_{tl} m - r_x(x),$$

Phenomenological time delay equation

$$\dot{x} = \frac{\alpha_u N}{1 + (P(t-\tau)/K)^2} + \frac{\alpha_b N (P(t-\tau)/K)^2}{1 + (P(t-\tau)/K)^2} - r_{deg}(x)$$

Protein degradation



- «Exponential» (generic mechanisms of cellular degradation)

$$r_{deg}(x) = \gamma x$$

- Enzyme (by a specific enzyme that acts on tagged proteins)

$$r_{deg,enz} = \frac{V_{max} x}{K_m + x}$$

- If an enzyme degrades several proteins, the rate becomes

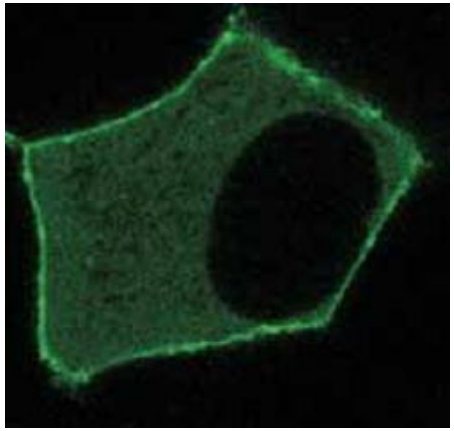
$$r_{deg,i}(x_i, T) = \frac{V_{max,i} x_i}{K_m + \sum_j x_j} - \gamma x_i$$

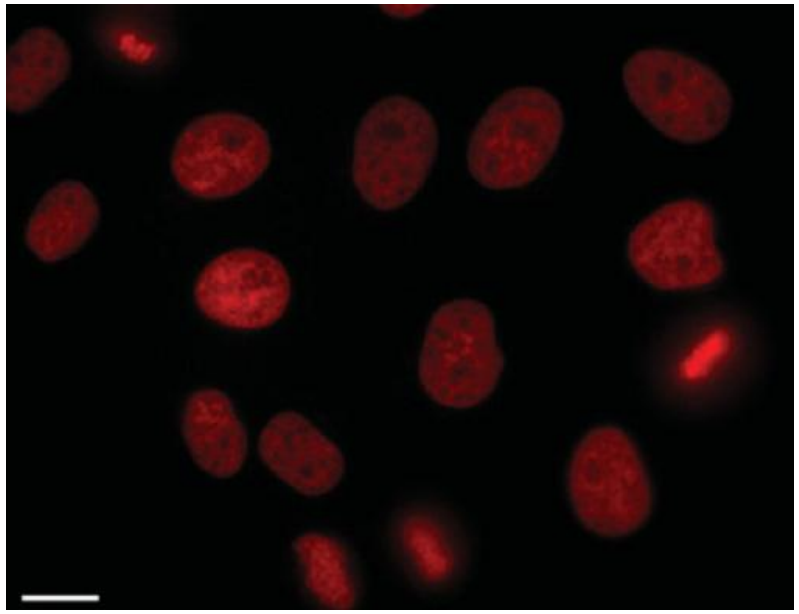
Synthetic biology: plugging in new genes and constructing regulatory networks

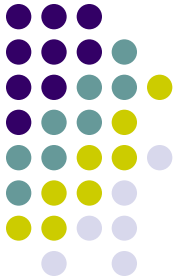
- Novel functionality

- «well-orthogonal» signals = an opportunity to study dynamics of networks isolated to a good degree

- Visualising gene expression by fluorecent proteins

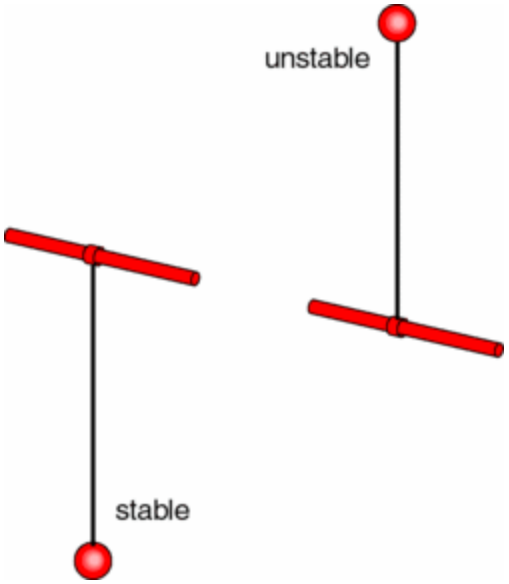






Dynamics of gene regulation

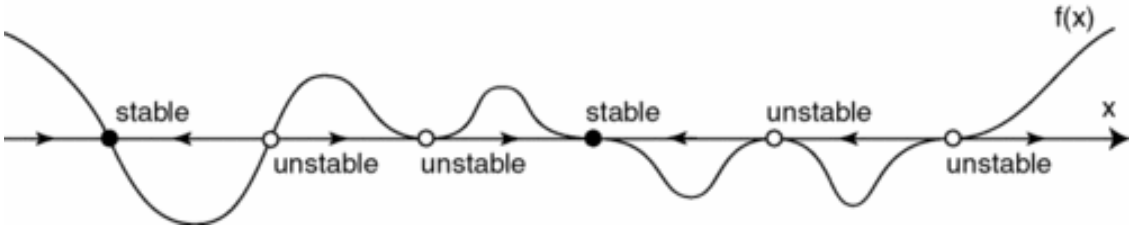
Nonlinear dynamics: equilibrium states



$$\dot{x} = f(x)$$ 1D dynamical system

$$f(x^*) = 0$$ Equilibrium

$$f'(x^*) < 0$$ Stability criterion



Nonlinear dynamics: equilibrium states

2D dynamical system

$$\dot{x}_1 = f(x_1, x_2)$$
$$\dot{x}_2 = f(x_1, x_2)$$

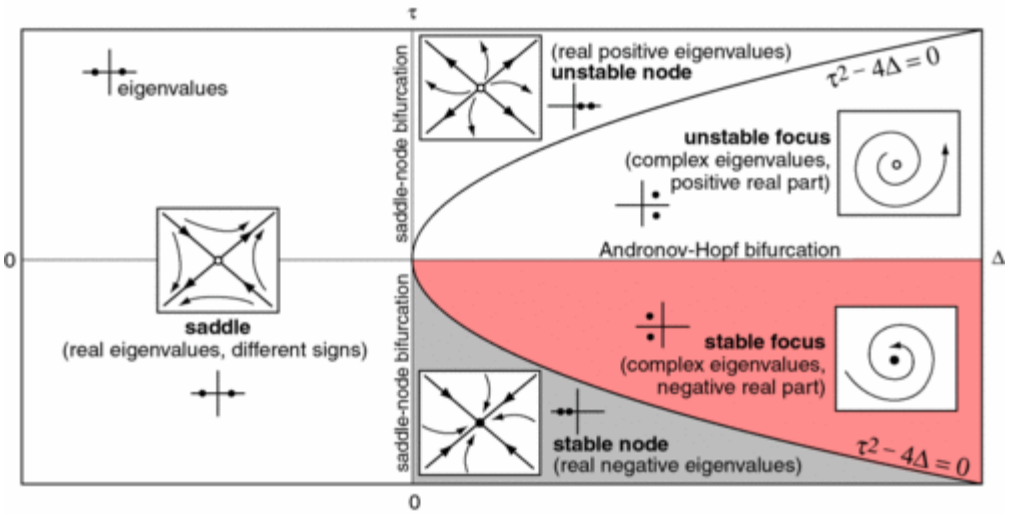
Equilibrium

$$f_1(x_1, x_2) = 0$$
$$f_2(x_1, x_2) = 0$$

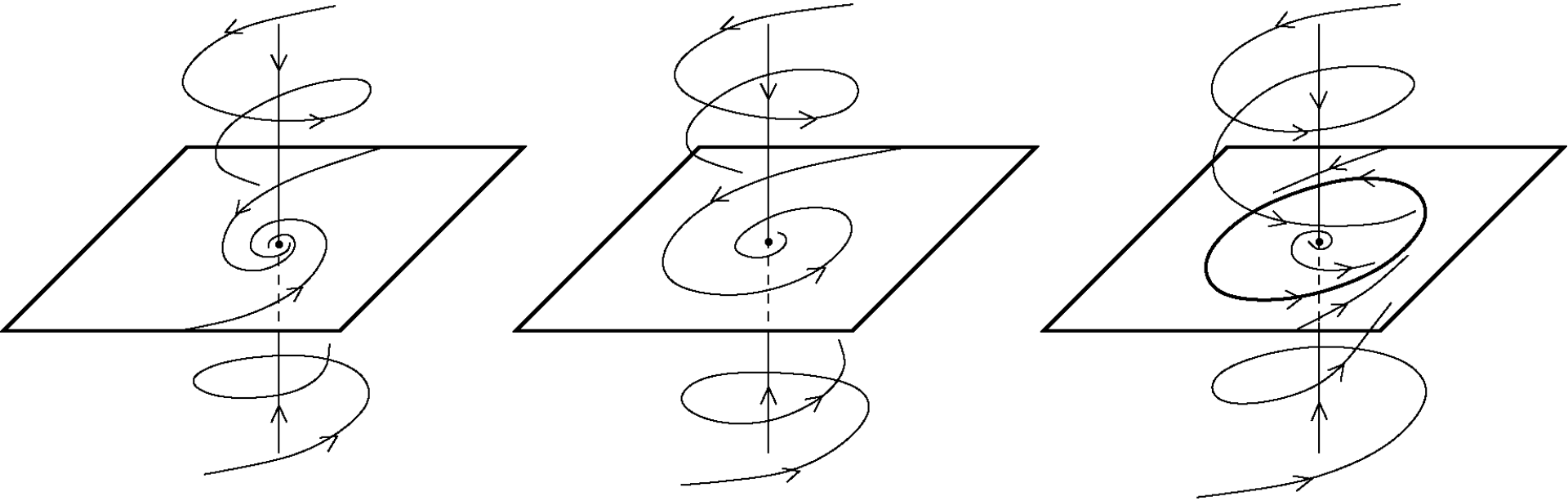
Stability criterion

$$\mathrm{Det}\begin{pmatrix} \partial_{x_1} f_1 - \lambda & \partial_{x_2} f_1 \\ \partial_{x_1} f_2 & \partial_{x_2} f_2 - \lambda \end{pmatrix} = 0$$

$$\mathrm{Re}\,\lambda_{1,2} < 0$$

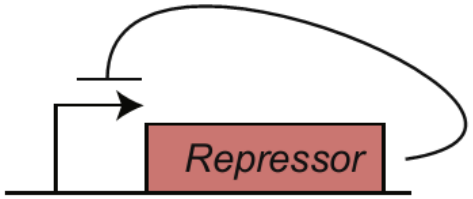


Nonlinear dynamics: oscillations



$\beta < 0$ $\beta = 0$ $\beta > 0$

Example 2: self-repressor



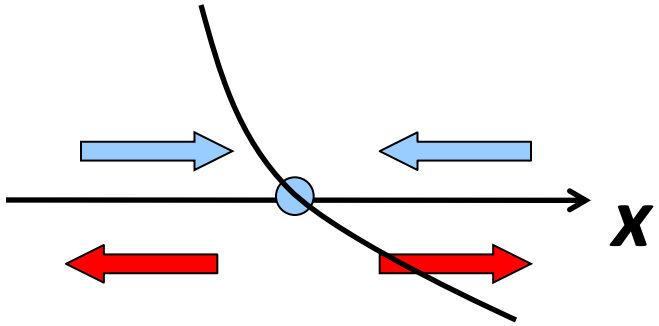
Differential equation

$$\dot{x} = \frac{\alpha}{1 + (x/K)^n} - \gamma x$$

Equilibrium

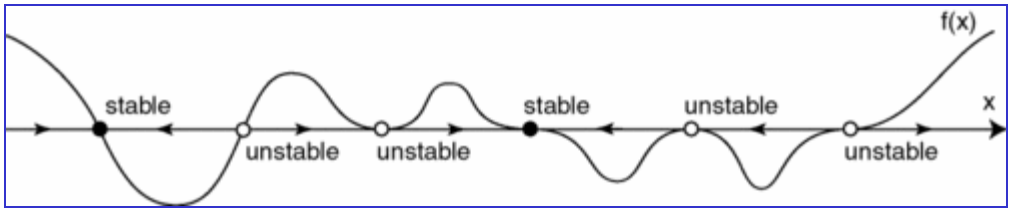
$$\gamma x = \frac{\alpha}{1 + (x/K)^n}$$

Stability



?

$$\frac{\alpha}{1 + (x/K)^n} - \gamma x$$



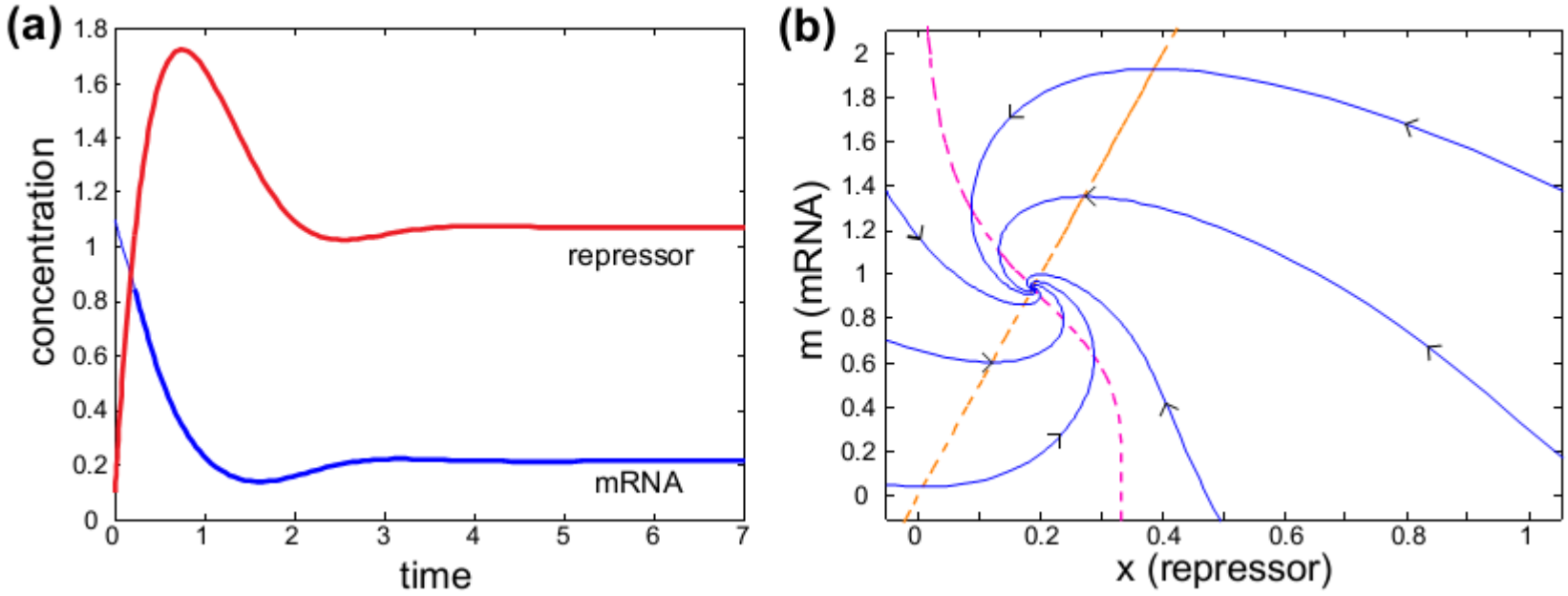
Self-repressor: more complex dynamics?

- mRNA dynamics:

$$\dot{m} = \frac{\alpha}{1 + (x/K)^n} - \delta m,$$

$$\dot{x} = \beta m - \gamma x, ,$$

- Still an equilibrium, though with oscillatory approach (stable focus)



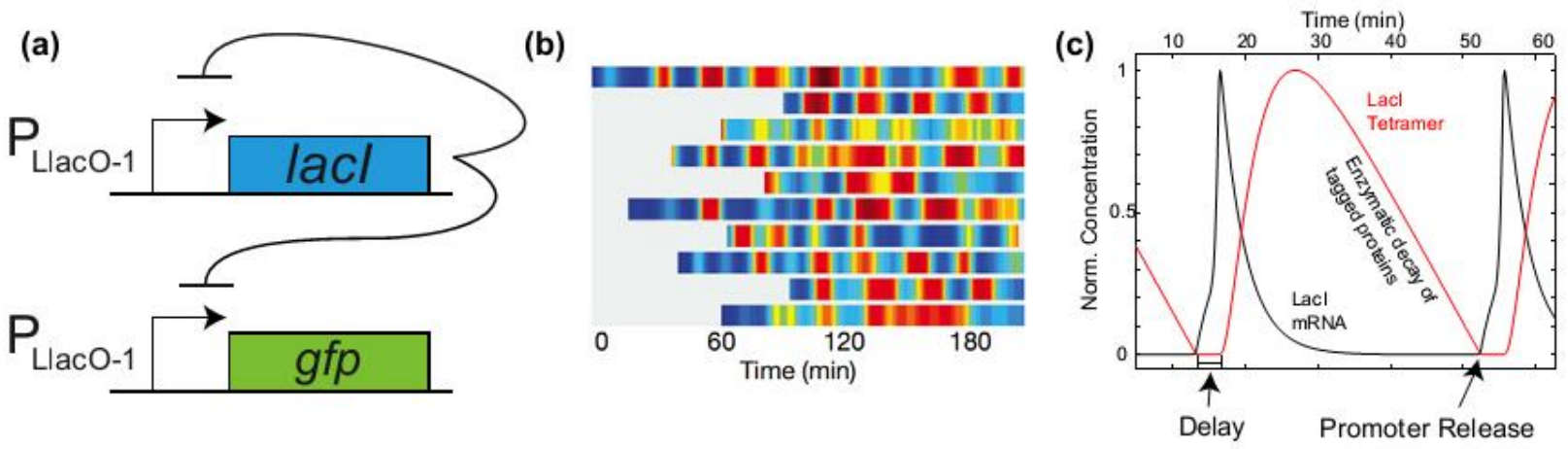
Self-repressor with time-delay: self-sustained oscillations!

- Mathematical model and biological experiment

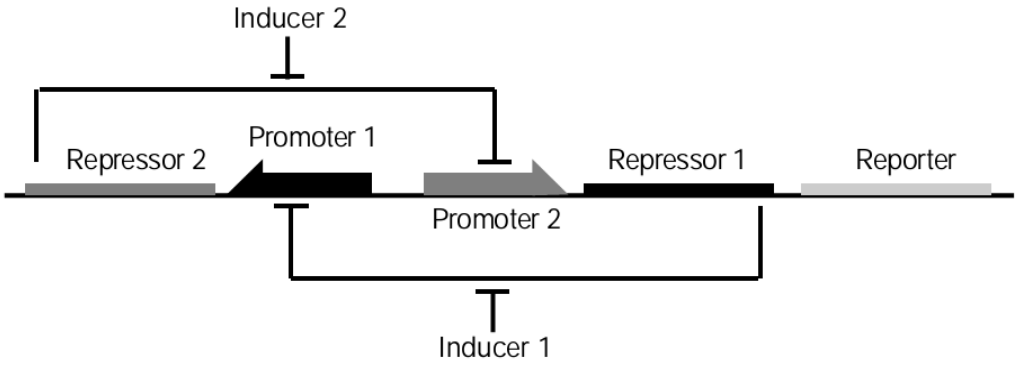
  (*Striker et al., Nature, 2008*)
- Biological clocks

$$\dot{x} = \frac{\alpha'}{1 + (x_\tau / K)^\beta} - \frac{Vx}{K_m + x} - \gamma x$$

$$x_\tau = x(t - \tau)$$



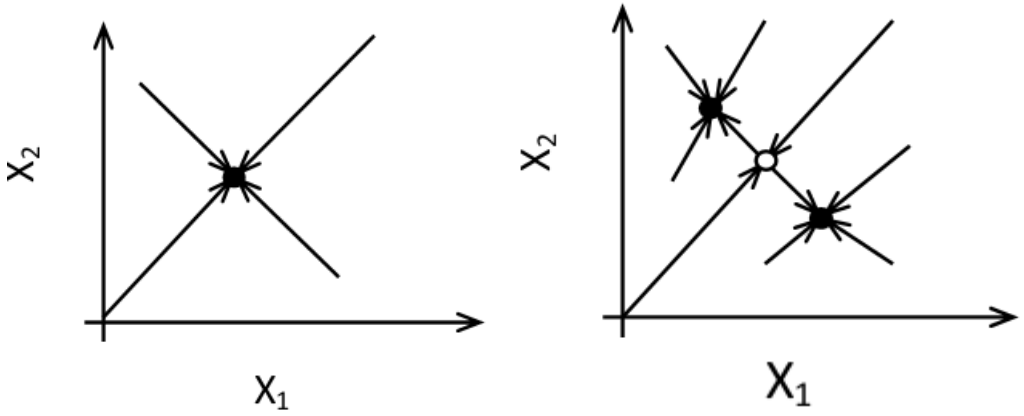
Genetic toggle-switch

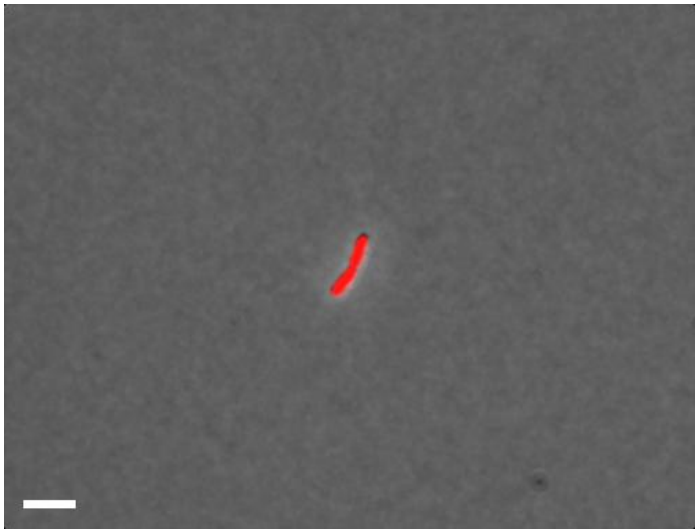


$$\dot{x}_1 = \frac{K_1}{1 + x_2^{n_2}} - x_1$$

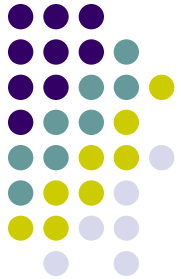
$$\dot{x}_2 = \frac{K_2}{1 + x_1^{n_1}} - x_2$$

(*Gardner, Kantor & Collins, Nature, 2000*)



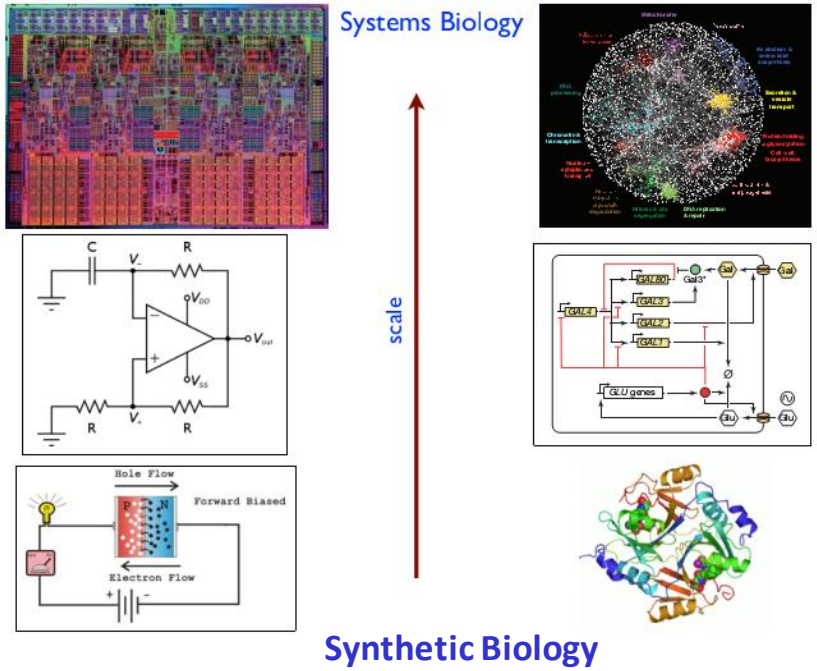


Noise-driven separation



Collective dynamics

Towards distributed circuits



Single cell limitations:

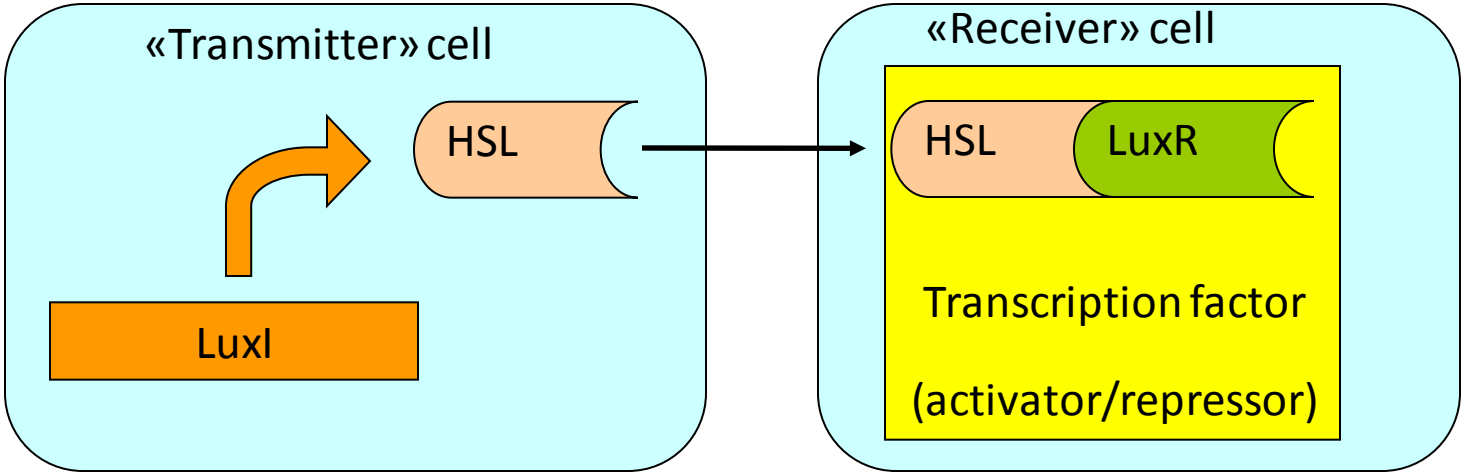
- Number of exogeneous constructs (robustness of circuit, cell viability)
- Complexity of circuit
- Useful functionality, regulated by the circuit (e.g. oscillatory)

Solution:

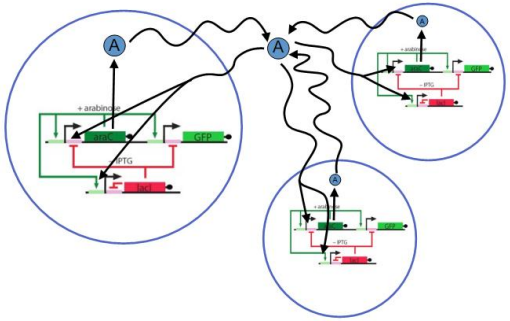
- Distributed circuits and intercellular communication

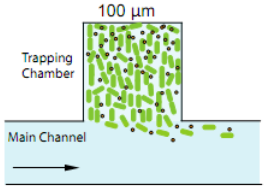
Intercellular communication

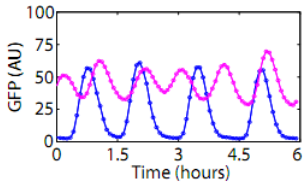
Quorum-sensing mechanisms:
HSL (homoserine-lacton) family ligands

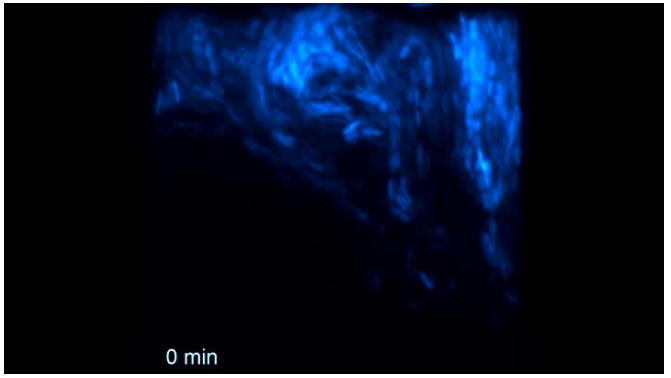


Intercellular communcation: synchronization of genetic clocks

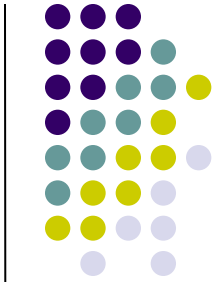






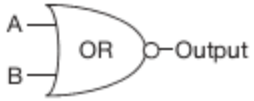


Danino et al., Nature, 2010

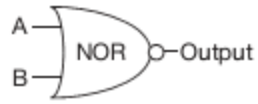


Logic circuits

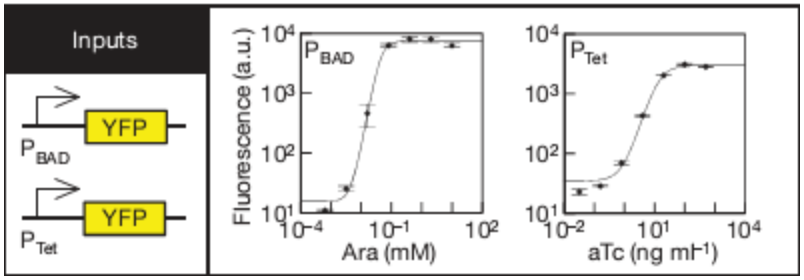
Example 3. Logic circuits

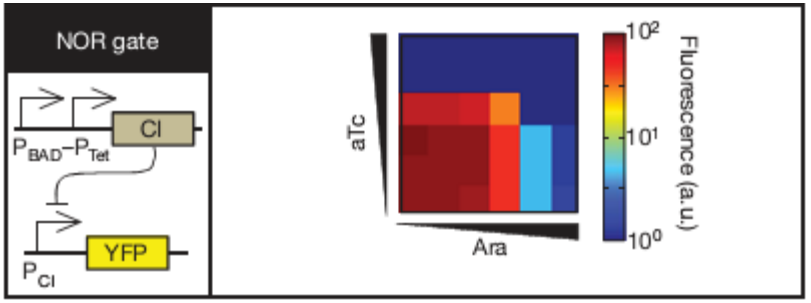


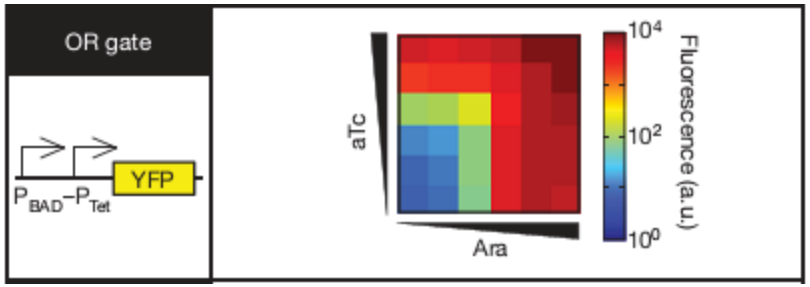
Inputs		
in1	in2	Output
0	0	0
0	1	1
1	0	1
1	1	1



Inputs		
in1	in2	Output
0	0	1
0	1	0
1	0	0
1	1	0

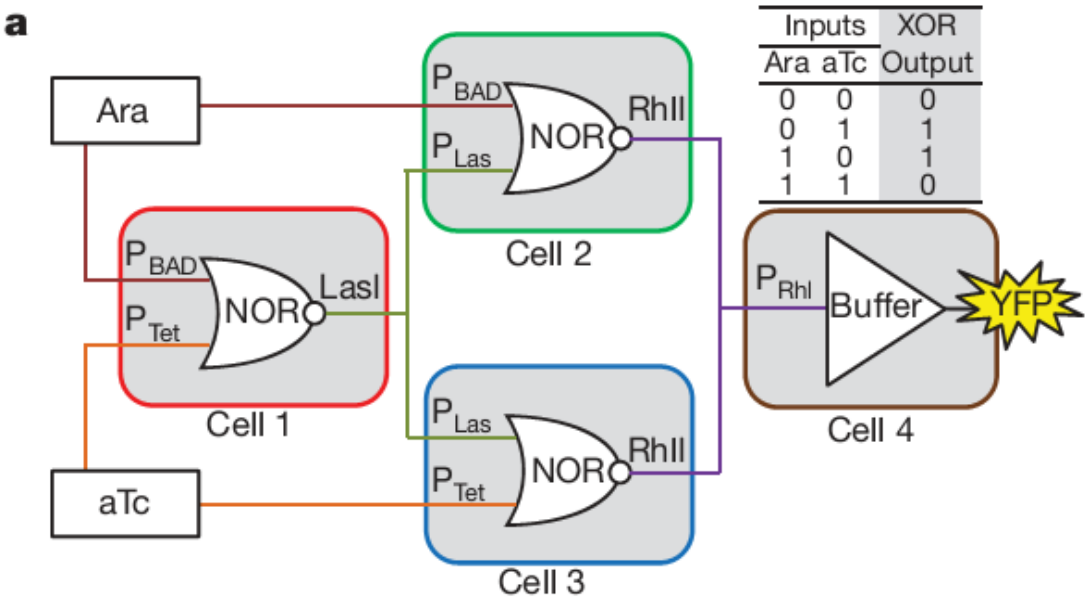


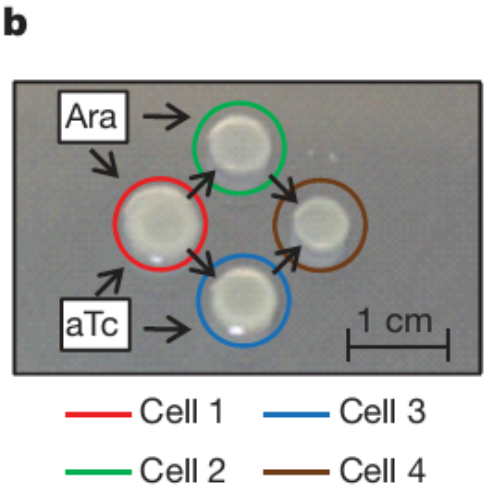




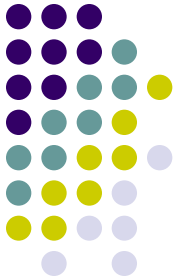
Tamsir, Tabor & Voigt, Nature, 2011

Distributed logic circuits





Tamsir, Tabor & Voigt, Nature, 2011



Next generation gene computing: learning

Distributed genetic classifier

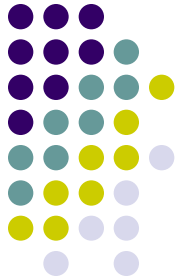
http://arxiv.org/abs/1405.5328

Andriy Didovyk,[†] Oleg I. Kanakov,[‡] Mikhail V. Ivanchenko,[¶] Jeff Hasty,[†,§,‖]

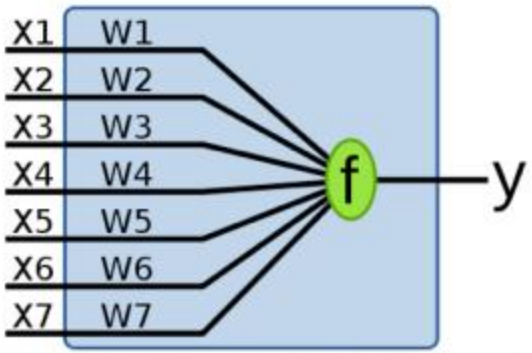
Ramón Huerta,[*,†] and Lev Tsimring[*,†]

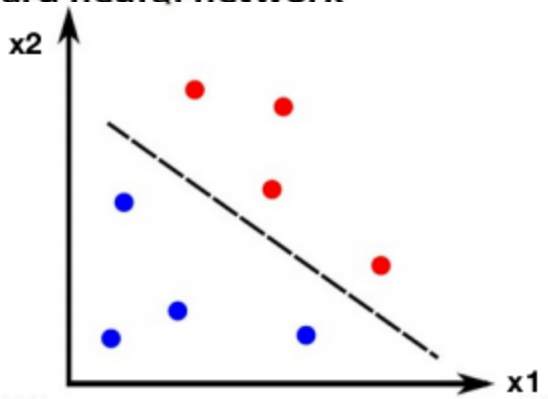
BioCircuits Institute, University of California San Diego, La Jolla CA, USA, Department of Radiophysics, Lobachevsky State University of Nizhniy Novgorod, Nizhniy Novgorod, Russia, Department for Bioinformatics, Lobachevsky State University of Nizhniy Novgorod, Nizhniy Novgorod, Russia, Department of Bioengineering, University of California San Diego, La Jolla CA, USA, and Molecular Biology Section, Division of Biological Science, University of California San Diego, La Jolla CA, USA

Machine learning: classical ideas



Perceptron- one layer feedforward neural network





1. Initialise weights and threshold. Note that weights may be initialised by setting each weight node $w_i(0)$ to 0 or to a small random value. In the example below, we choose the former.

2. For each sample j in our training set D, perform the following steps over the input $\mathbf{x}_j$ and desired output d_j:

2a. Calculate the actual output:

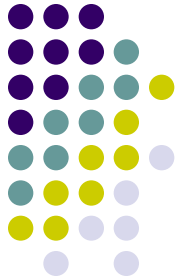
$$y_j(t) = f[\mathbf{w}(t) \cdot \mathbf{x}_j] = f[w_0(t) + w_1(t)x_{j,1} + w_2(t)x_{j,2} + \cdots + w_n(t)x_{j,n}]$$

2b. Adapt weights:

$$w_i(t+1) = w_i(t) + \alpha(d_j - y_j(t))x_{j,i}, \text{ for all nodes } 0 \leq i \leq n.$$

Step 2 is repeated until the iteration error $d_j - y_j(t)$ is less than a user-specified error

Molecular perceptron?



Perceptron inside the cell??

Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 10983–10987, December 1991
Chemistry

Chemical implementation of neural networks and Turing machines

ALLEN HJELMFELT†, EDWARD D. WEINBERGER†, AND JOHN ROSS‡

†Max-Planck-Institut für Biophysikalische Chemie, D-3400 Göttingen, Federal Republic of Germany; and ‡Department of Chemistry, Stanford University, Stanford, CA 94305

972 Biophysical Journal Volume 66 April 1994 972–977

Computer Simulated Evolution of a Network of Cell-Signaling Molecules

Dennis Bray* and Steven Lay†

*Department of Zoology and †Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, United Kingdom

Journal of Theoretical Biology 249 (2007) 58–66

www.els

Associative learning in biochemical networks

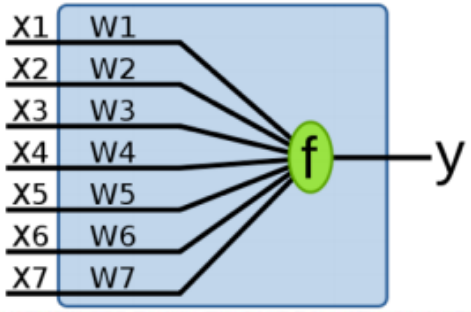
Nikhil Gandhi[a], Gonen Ashkenasy[b,*], Emmanuel Tannenbaum[b,*]

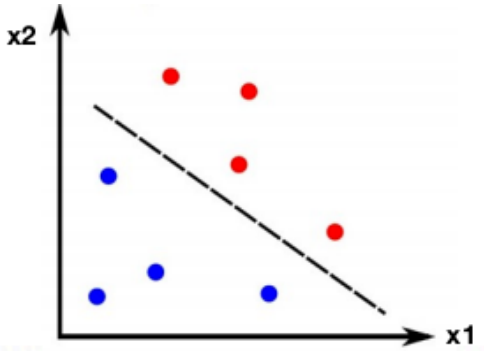
[a]*College of Computing, Georgia Institute of Technology, Atlanta, GA 30332, USA*
[b]*Department of Chemistry, Ben-Gurion University of the Negev, Be'er-Sheva 84105, Israel*

Molecular classifier?

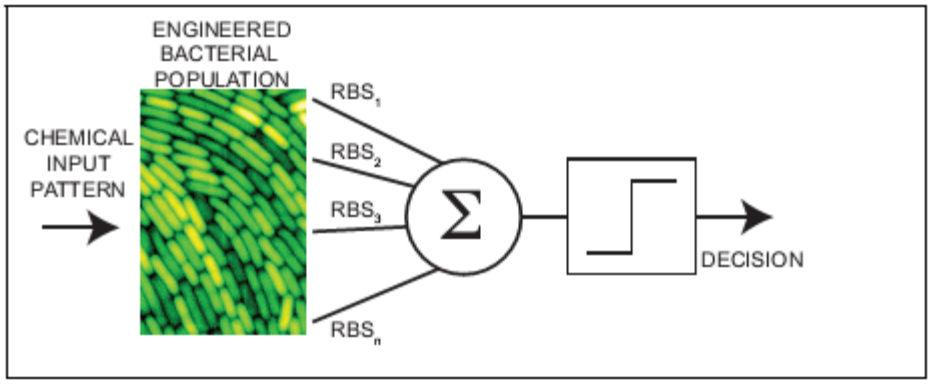
Weak points:

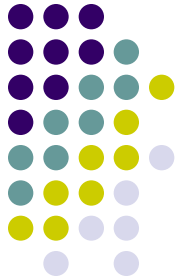
- Actual constructs
- Memory





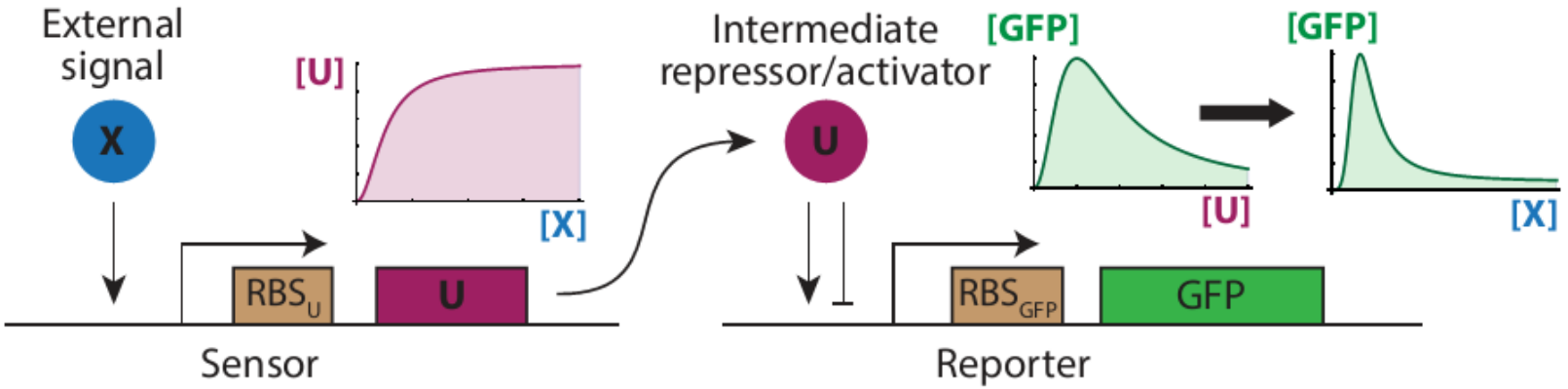
- Our solution: population classifier





Single-input classifier

Single cell synthetic circuit



$$\frac{du}{dt} = r_u(x; m_u) - \mu_u u \qquad r_u(x; m_u) = m_u \cdot \frac{\alpha A_u^{p_u} + x^{p_u}}{A_u^{p_u} + x^{p_u}}$$

$$\frac{dz}{dt} = r_g(u; m_g) - \mu_g z \qquad r_g(u; m_g) = m_g \cdot \frac{A_g^{p_g} u^{p_g}}{(A_g^{p_g} + u^{p_g})^2}$$

Ensemble of cells – ‘pixel’ classifier

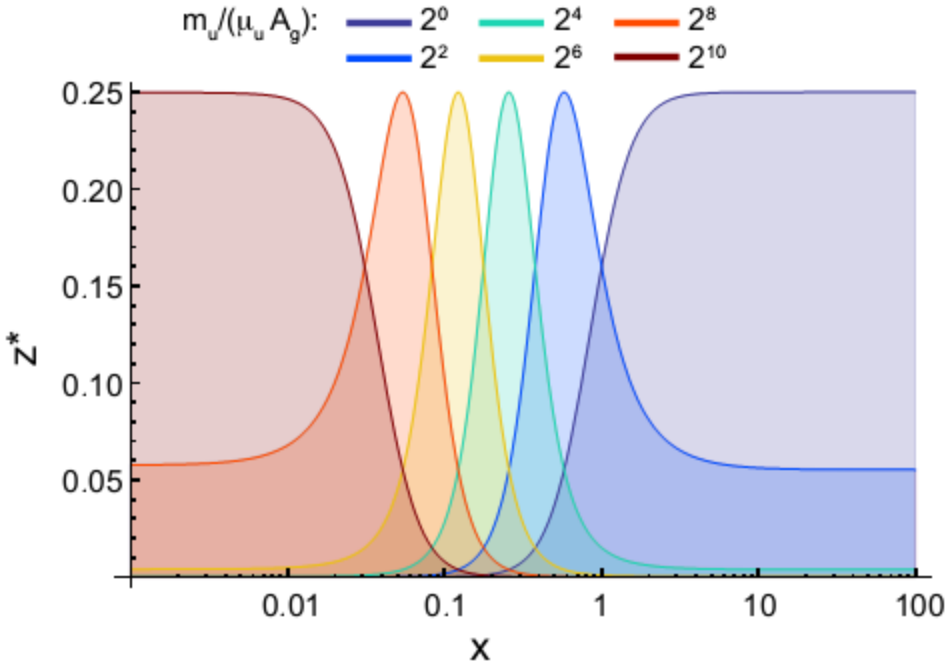
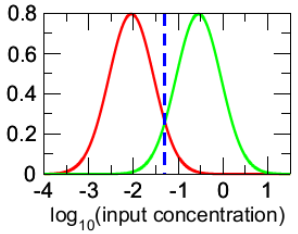
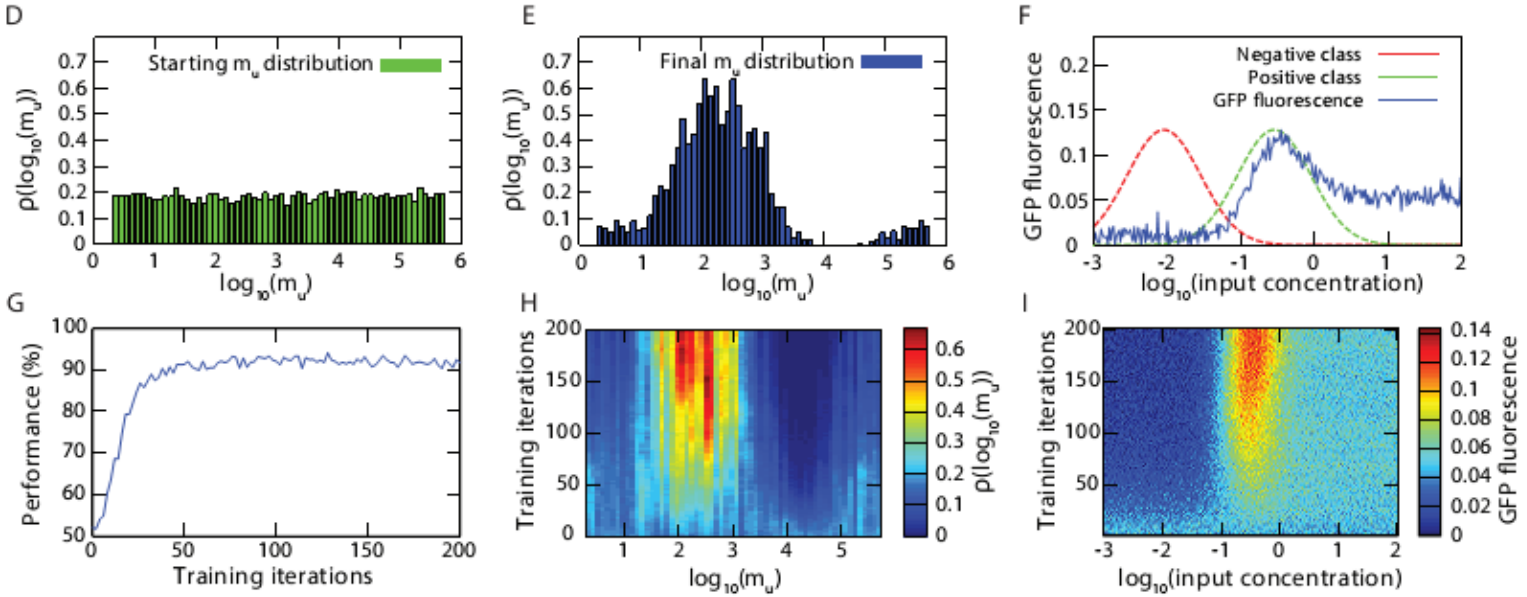


Figure 2: Steady state GFP concentration (z^*) as a function of the concentration of the external chemical signal X for the modular classifier circuit shown for a range of m_u values representing a range of the relative strengths of the sensor promoter (Fig. 1). Non-dimensional circuit parameters are $\mu_u = \mu_g = m_g = A_u = 1$, $A_g = 20$, $p_g = p_u = 2$, $\alpha = 10^{-3}$.

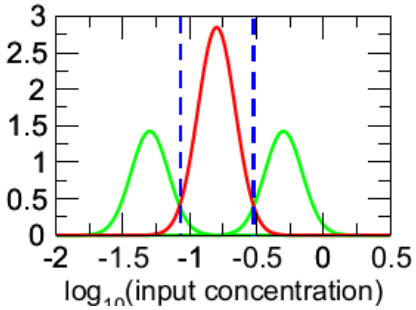
Training by examples through cell elemination

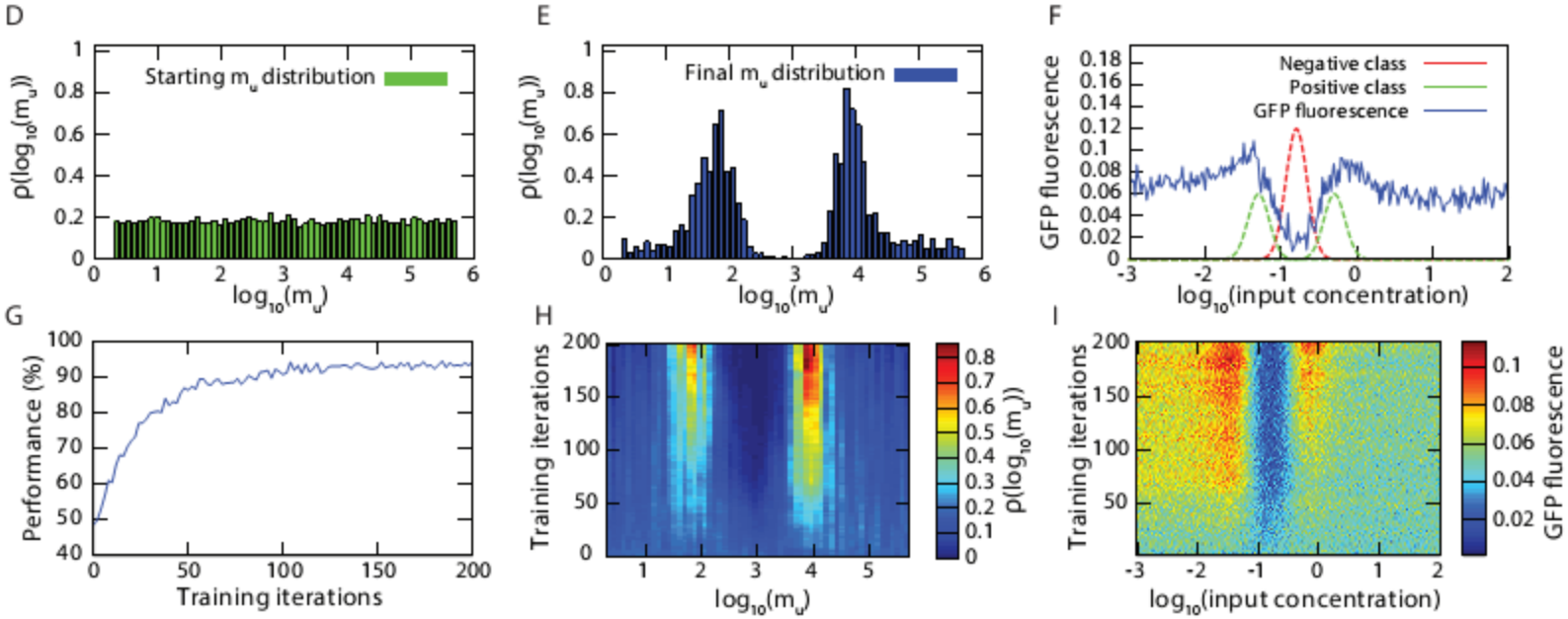
Discrimination of two unimodal (overlapping!) classes

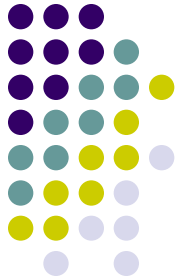




Unimodal vs. Bimodal classes







Two- (multiple) input classifier

Two- (multiple) input classifier

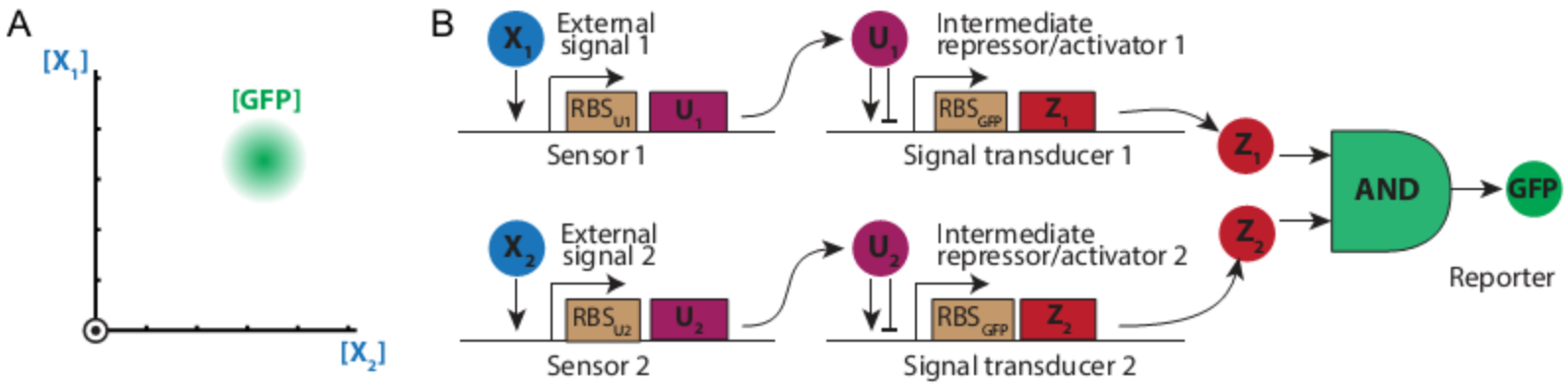
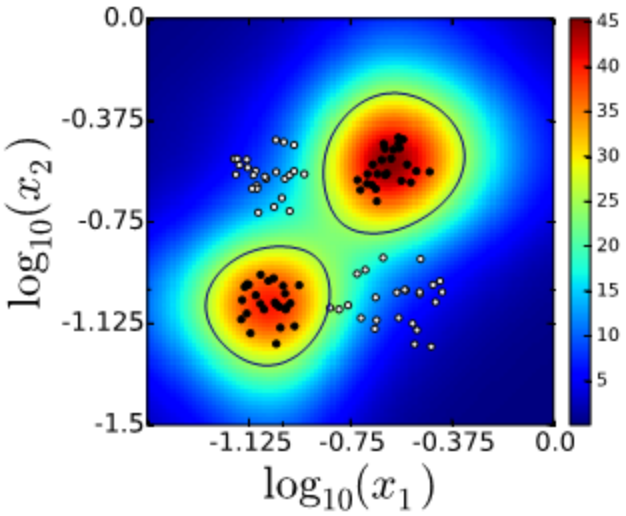


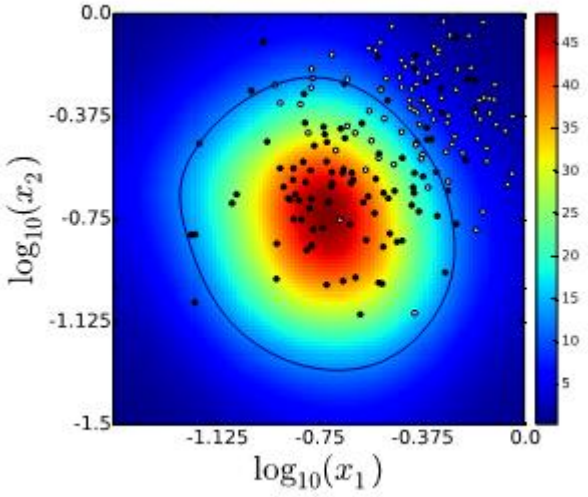
Figure 8: (A) The desired bell-shaped response of the multidimensional classification circuit. (B) The genetic circuit proposed for use in a distributed genetic classifier with multiple inputs per cell. Independent sensing and response functionalities are combined using an appropriate biological AND gate. The resulting response function of the entire two-promoter circuit is bell-shaped with respect to the inputs X_1 and X_2 for the relevant choice of parameters.

Results

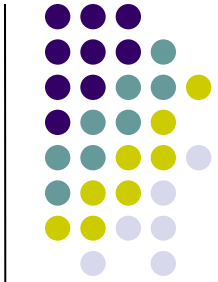
- Linearly non-separable classes



- Overlapping classes



Conclusions



- Synthetic gene regulatory network – a powerful approach in biocomputing
- Equations of molecular kinetics
- Already minimal gene circuits display stationary states, oscillations, bistability
- Distributed regulation networks through quorum sensing communications
- Logical circuits
- Learning. Distributed synthetic gene classifier.

Collaborations

- Lev Tsimring, UCSD

- Alexey Zaikin, UCL

- Oleg Kanakov, UNN

Acknowledgements

- Russian Foundation for Basic Research RFBR 13-02-00918
- Russian Foundation for Basic Research RFBR 14-02-01202
- The Ministry of education and science of the Russian Federation grant 02.B.49.21.0003