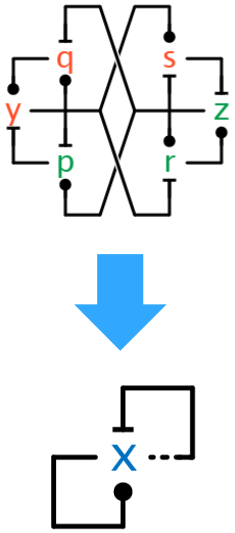
Microsoft Research

UNIVERSITY OF OXFORD

Morphisms of Reaction Networks

Luca Cardelli, Microsoft Research & Oxford University

Workshop on Molecular Walkers, Oxford, 2014-07-15

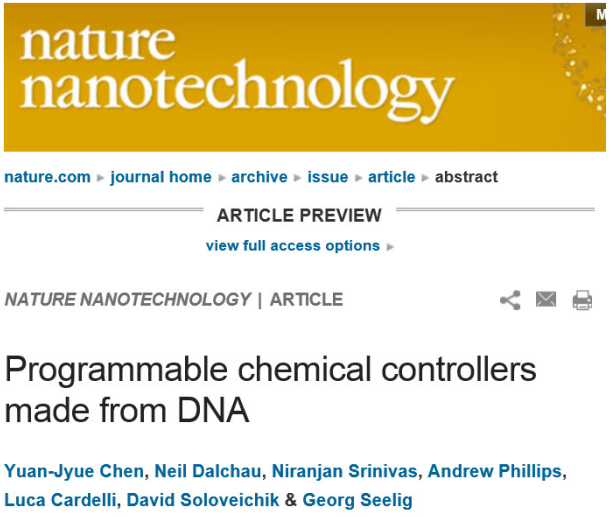


Microsoft Research

Introduction

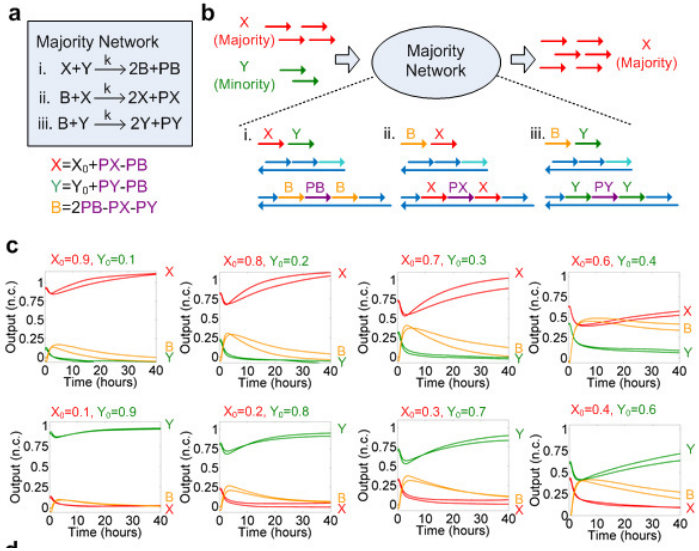
DNA Computing

- Programmable controllers for embedded DNA systems



Chemical Reaction Networks

- In DNA Strand Displacement we can implement arbitrary chemical reaction networks (CRN)
- CRN has become our "general purpose programming language" for nanotechnology



Engineered CRNs

- What is the meaning/purpose/effect of an engineered CRN program?
- How can we represent desired behavior (algorithms) in the CRN language?
- How can we correctly transform programs written in the CRN language?

Natural CRNs

- What is the meaning/purpose/effect of a natural CRN program?
- How can nature represent desired behavior (algorithms) in the CRN language?
- How can nature correctly transform programs written in the CRN language?

CRN Morphisms

When are two reaction networks related?

For example:

- When do they produce the same behavior?
- When is one more robust than another?
- When has one evolved from another?
- When is one a simplified but representative version of another?
- When are there hidden symmetries within one network?

- A morphism (map) relates two networks
  - Study conditions on morphisms that answer the above questions

Microsoft Research

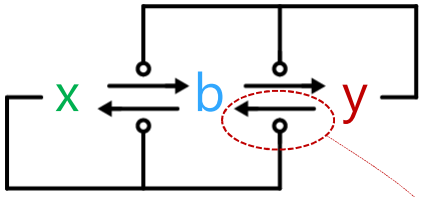
UNIVERSITY OF OXFORD

Algorithms

A Consensus Algorithm

- Population Protocols
  - Finite-state identity-free agents (molecules) interact in randomly chosen pairs
  - Each interaction (collision) can result in state changes
  - Complete connectivity, no centralized control (well-mixed solution)
- A Population Consensus Problem
  - Find which state x or y is in majority in the population
  - By converting the *whole* population to x or y
- Approximate Majority (AM) Algorithm
  - Uses a third "undecided" state b
  - Disagreements cause agents to become undecided
  - Undecided agents believe any non-undecided agent
- With high probability, for *n* agents
  - The total number of interactions is $O(n \log n) \Rightarrow$ fast (optimal)
  - Correct outcome if the initial disparity is $\omega(\sqrt{n} \log n) \Rightarrow$ robust
  - In parallel time, converges in $O(\log n)$

catalysis



chemical reaction network

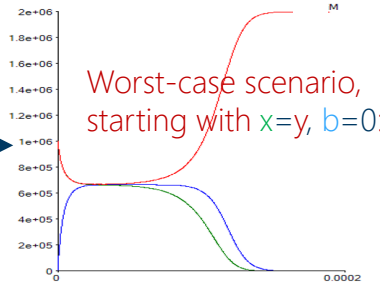
$x + y \rightarrow^{r} y + b$

$y + x \rightarrow^{r} x + b$

$b + x \rightarrow^{r} x + x$

$b + y \rightarrow^{r} y + y$

x=y



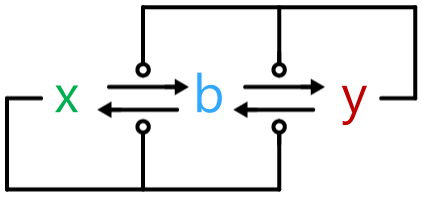
Worst-case scenario, starting with x=y, b=0:

Dana Angluin · James Aspnes · David Eisenstat

A Simple Population Protocol for Fast Robust Approximate Majority

A Plain Biological Implementation

Approximate Majority (AM)



Dana Angluin · James Aspnes · David Eisenstat

A Simple Population Protocol for Fast Robust Approximate Majority

2007

Epigenetic Switch

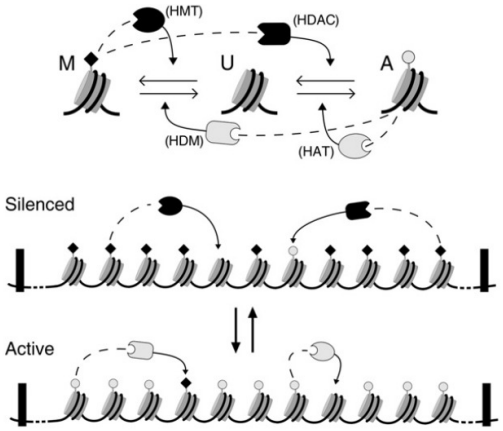


Figure 1. Basic Ingredients of the Model

Theory

Cell

Theoretical Analysis of Epigenetic Cell Memory by Nucleosome Modification

Ian B. Dodd,[1,2] Mille A. Micheelsen,[1] Kim Sneppen,[1,*] and Geneviève Thon[3]

[1]Center for Models of Life, Niels Bohr Institute, Blegdamsvej 17, DK-2100, Copenhagen Ø, Denmark

[2]Department of Molecular and Biomedical Sciences (Biochemistry), University of Adelaide SA 5005, Australia

[3]Department of Molecular Biology, University of Copenhagen Biocenter, Ole Maaløes Vej 5, DK-2200 Copenhagen N, Denmark

*Correspondence: sneppen@nbi.dk

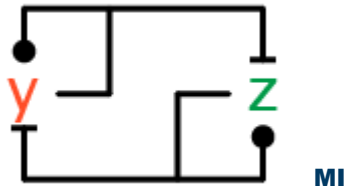
DOI 10.1016/j.cell.2007.02.053

2007

Obfuscated Implementations?

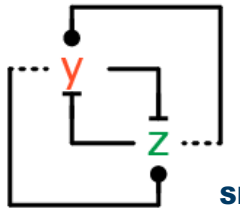
activation
inhibition

Mutual Inhibition & Self Activation



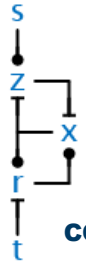
MI

Mutual Inhibition & Mutual Anti-activation



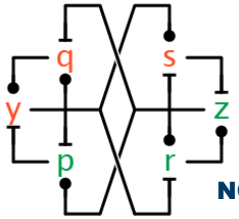
SI

Cell Cycle Switching



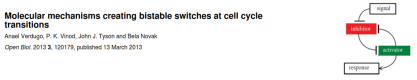
CC

Better Switching



NCC

Cell cycle transitions

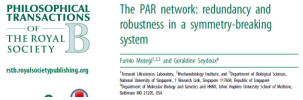


Molecular mechanisms creating bistable switches at cell cycle transitions

Angel Verdugo, P. K. Vinod, John J. Tyson and Bela Novak

Open Biol. 2013 **3**, 120179, published 13 March 2013

Polarity establishment

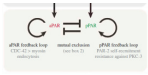


PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

rstb.royalsocietypublishing.org

The PAR network: redundancy and robustness in a symmetry-breaking system

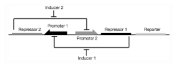
Fumio Motegi[1,2,3] and Geraldine Seydoux[4]



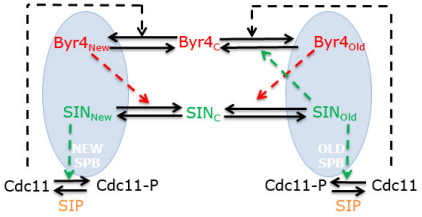
Gene networks

Construction of a genetic toggle switch in *Escherichia coli*

Timothy S. Gardner[1,2], Charles R. Cantor[1] & James J. Collins[1,2]



Septation Initiation



SIN inhibiting Byr4,
absence of SIN activating Byr4

Dynamics of SIN Asymmetry Establishment

Archana Bajpai[1], Anna Feoktistova[2], Jun-Song Chen[2], Dannel McCollum[2], Masamitsu Sato[3,4], Rafael E. Carazo-Salas[5], Kathleen L. Gould[2], Attila Csikasz-Nagy[1,6,*]

PLOS Computational Biology | www.ploscompbiol.org 1 July 2013 | Volume 9 | Issue 7 | e1003147

The G_2/M cell cycle switch

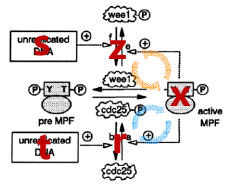
Journal of Cell Science 106, 1153-1168 (1993)
Printed in Great Britain © The Company of Biologists Limited 1993

Numerical analysis of a comprehensive model of M-phase control in *Xenopus* oocyte extracts and intact embryos

Bela Novak* and John J. Tyson†

Department of Biology, Virginia Polytechnic

*Permanent address: Department of Agricultural Chemic
†Author for correspondence



Nature **344**, 503 - 508 (05 April 1990); doi:10.1038/3

Universal control mechanism regulating onset of M-phase

PAUL NURSE

ICRF Cell Cycle Group, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

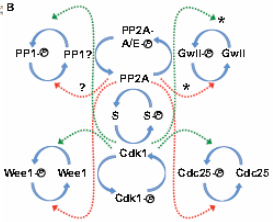
The "new" cell cycle switch

Phosphorylation network dynamics in the control of cell cycle transitions

Daniel Fisher[1,*], Liliana Krasinska[1,2], Damien Coudreuse[3,4] and Béla Novák[3,*]

[1]Institut de Génétique Moléculaire de Montpellier, IGMM, CNRS UMR 5535, Université Montpellier I and II, 34293 Montpellier, France
[2]Institute of Genetics and Development of Rennes, CNRS UMR 6290, 35043 Rennes, France
[3]Oxford Centre for Integrative Systems Biology, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK
*Author for correspondence (bela@bioc.ox.ac.uk)
†These authors contributed equally to this work

Journal of Cell Science 122, 69 - 73 © 2012. Published by The Com
doi: 10.1242/jcs.100321

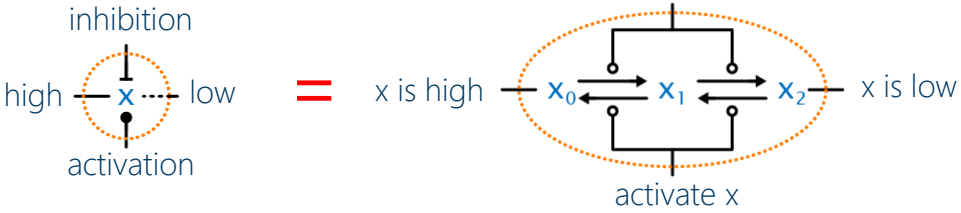


Microsoft Research

Networks and Morphisms

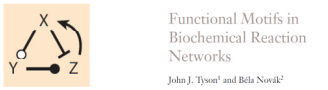
Influence Networks

activation
inhibition
catalysis



Usually modeled by sigmoid (e.g. Hill or Reinitz) functions

We model them by 4 mass action reactions over 3 species x_0, x_1, x_2



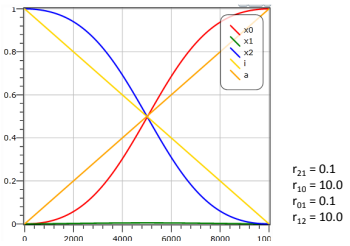
Functional Motifs in Biochemical Reaction Networks

John J. Tyson[1] and Béla Novák[2]

$$\frac{dX_i}{dt} = \gamma_i \frac{[A_i(1-X_i) - B_i X_i]}{A_i + B_i}, \quad i = 1, \ldots, N,$$

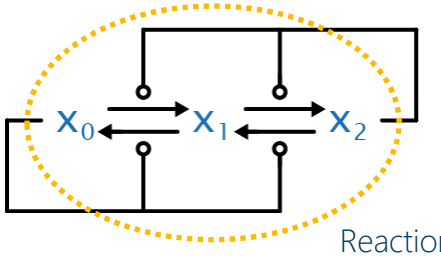
$$A_i = \exp\left\{\sigma_i\left(\alpha_{i0} + \sum_{j=1}^{N} \alpha_{ij} X_j\right)\right\}, \quad B_i = \exp\left\{\sigma_i\left(\beta_{i0} + \sum_{j=1}^{N} \beta_{ij} X_j\right)\right\}, \qquad (4)$$

They actually implement a Hill function of coefficient 2:



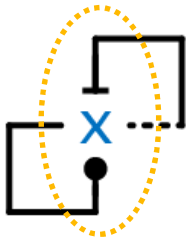
$r_{21} = 0.1$
$r_{10} = 10.0$
$r_{01} = 0.1$
$r_{12} = 10.0$

Approximate Majority



Reaction Network

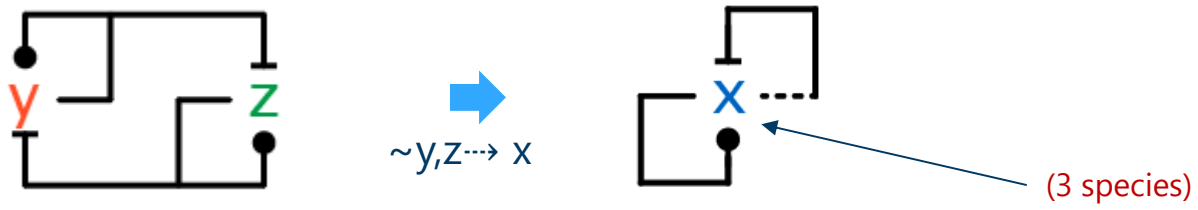
=

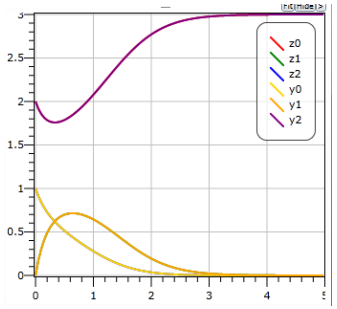


Influence Network

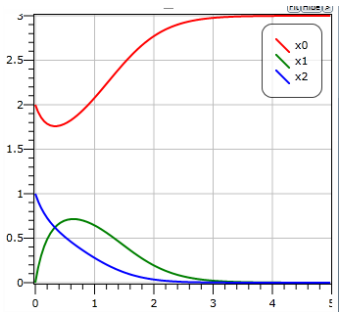
Network Emulation: MI emulates AM

- For *any* rates and initial conditions of **AM**, we can find *some* rates and initial conditions of **MI** such that the (6) trajectories of **MI** retrace those (3) of **AM**:





(6 species on 3 trajectories)



(3 species on 3 trajectories)

initialize:
$z = x$
$\sim y = x$

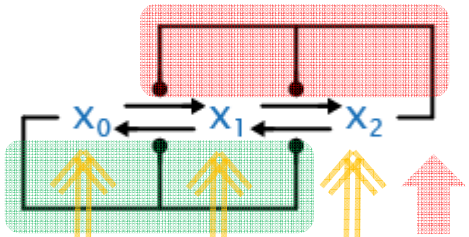
($y_2 = x_0$
$y_1 = x_1$
$y_0 = x_0$)

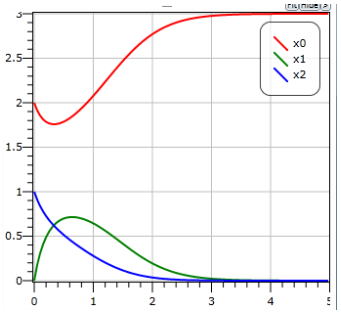
- How do we find these matching parameters? By a network morphism!

Emulation is a Network Morphism

A mapping of species and reactions

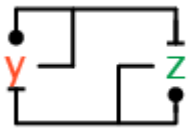
AM



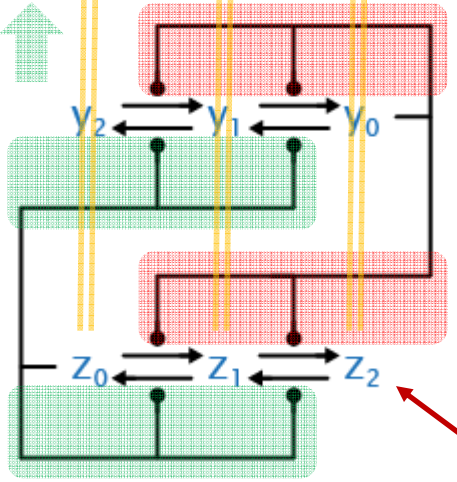


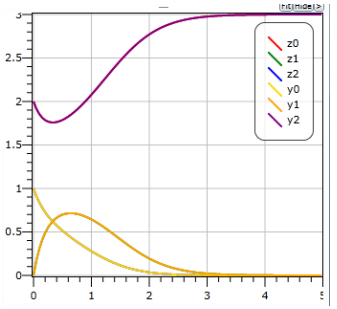
any initial conditions

homomorphic mapping



MI





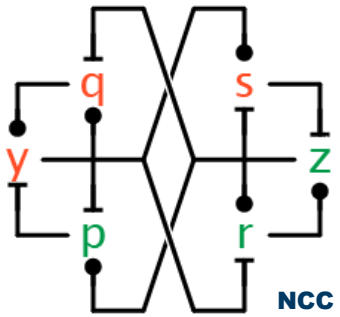
initial conditions:

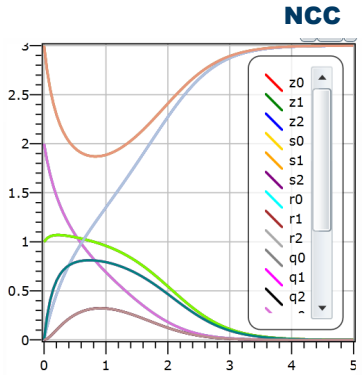
$z_0 = y_2\ (= x_0)$
$z_1 = y_1\ (= x_1)$
$z_2 = y_0\ (= x_2)$

less trivial than you might think:
it need not preserve the out-degree of a node!

Network Emulation: NCC emulates MI

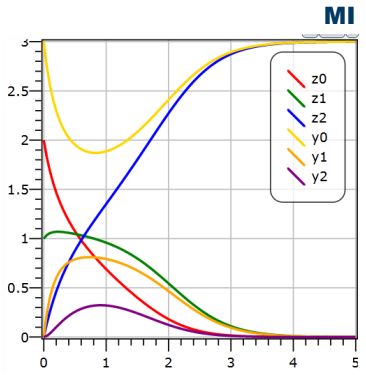
- For *any* rates and initial conditions of MI we can find *some* rates and initial conditions of NCC such that the (18) trajectories of NCC retrace those (6) of MI





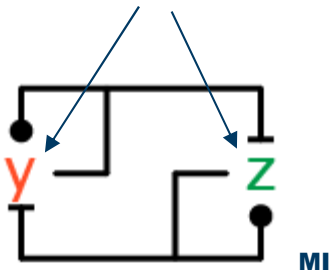
(18 species on 6 trajectories)

z,r,p ⇢ z
y,q,s ⇢ y



(6 species on 6 trajectories)

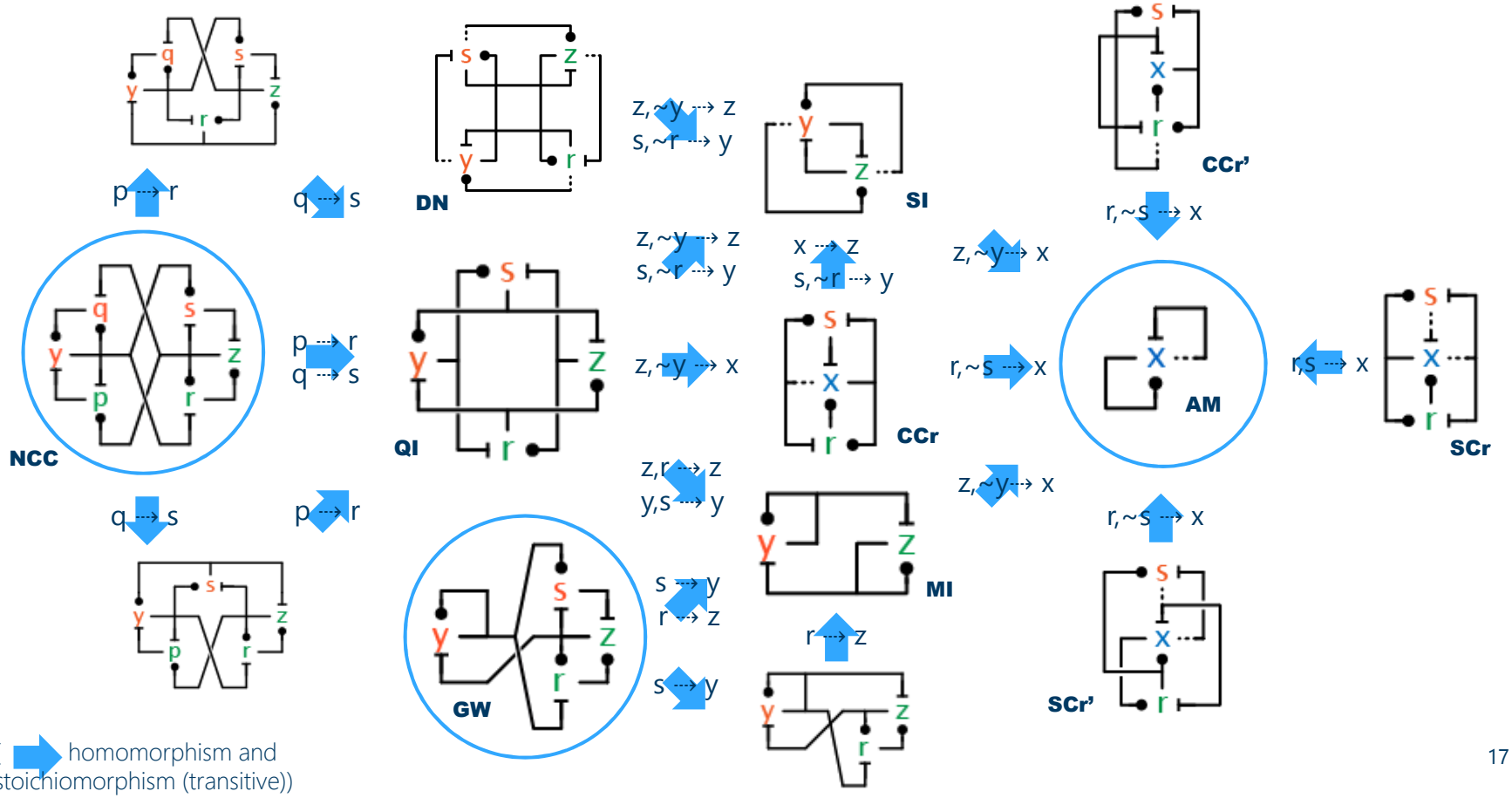
(3 species each)



initialize
z,r,p = z
y,q,s = y

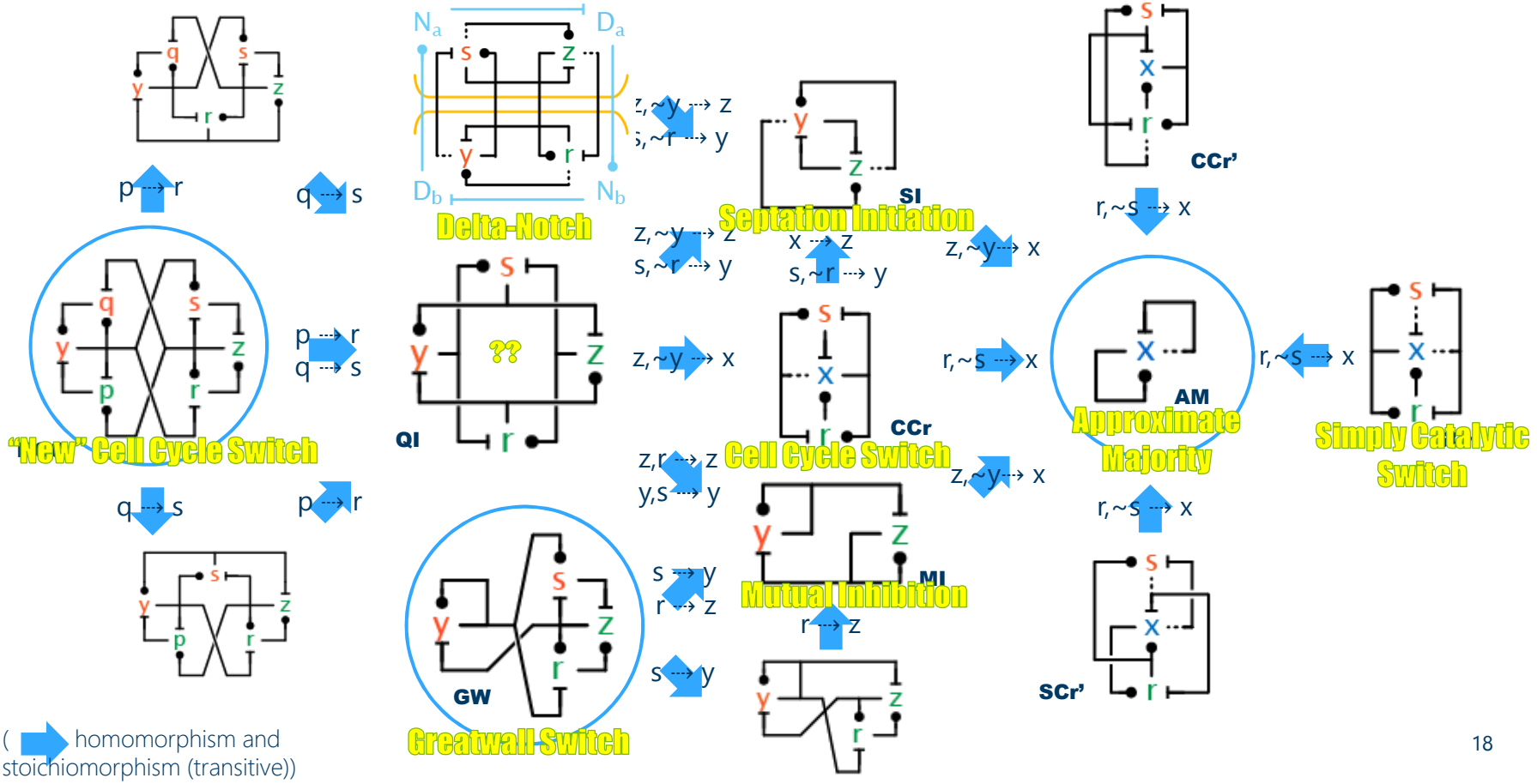
- Why does this work so well?

Approximate Majority Emulation Zoo



(➡ homomorphism and stoichiomorphism (transitive))

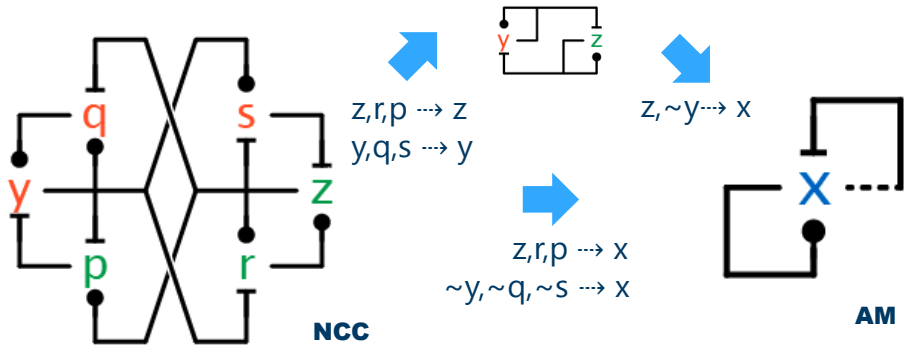
Approximate Majority Emulation Zoo

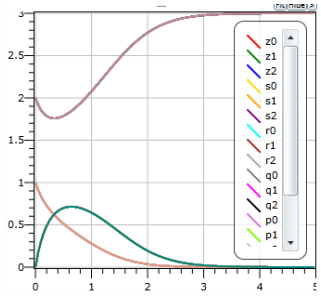


(homomorphism and stoichiomorphism (transitive))

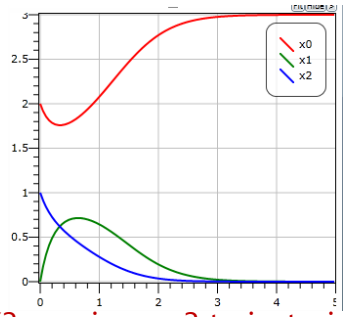
Emulations Compose: NCC emulates AM

- The (18) trajectories **NCC** can *always* retrace those (3) of **AM**



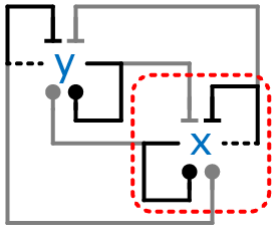


(18 species on 3 trajectories)

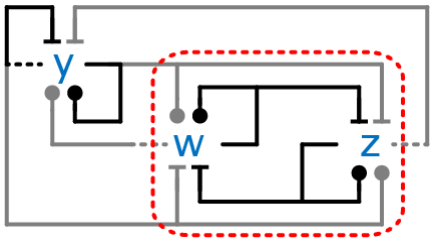


(3 species on 3 trajectories)

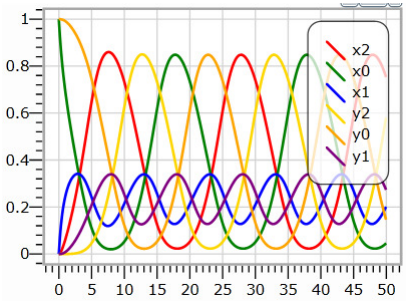
Emulation in Context

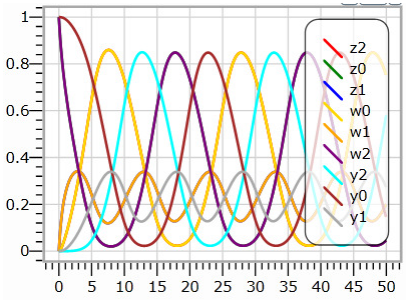


AM-AM Oscillator



AM-MI Oscillator





$m \in \mathbf{MI} \to \mathbf{AM}$ is an emulation:
it maps $z \to x$ and $\sim w \to x$

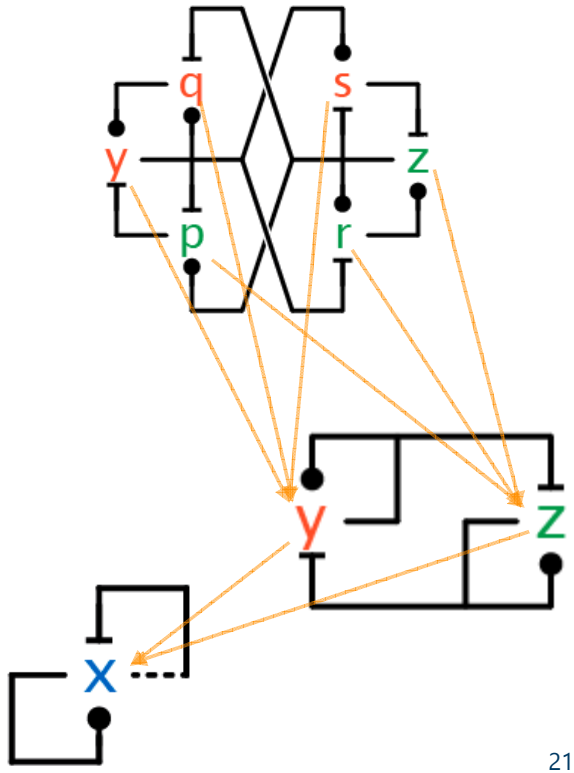
We can replace AM with MI in a context. The mapping m tells us how to wire MI to obtain an overall emulation:

Each influence crossing the dashed lines into x is replaced by a similar influence into *both* z *and* $\sim w$. The latter is the same as an opposite influence into w (shown).

Each influence crossing the dashed lines out of x is replaced by a similar influence from the same side of *either* z *or* $\sim w$. The latter is the same as a similar influence from the opposite side of w (shown), and the same as an opposite influence from the same side of w.

When can a Network Emulate Another?

- What kind of morphisms guarantee emulation?
  - do they preserve network structure?
  - do they preserve stoichiometry?



Emulation Theorem

Theorem: If $m \in (S, R) \to (\hat{S}, \hat{R})$ is a CRN reactant morphism and stoichiomorphism then it is a CRN emulation

reactant morphism $\boldsymbol{m_S}^{\mathbf{T}} \cdot \boldsymbol{\rho} = \hat{\boldsymbol{\rho}} \cdot \boldsymbol{m_R}^{\mathbf{T}}$ preserve enough network structure

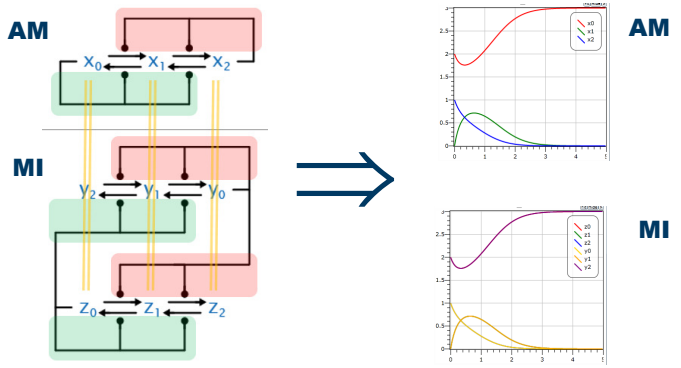
stoichiomorphism $\boldsymbol{\varphi} \cdot \boldsymbol{m_R} = \boldsymbol{m_S} \cdot \hat{\boldsymbol{\varphi}}$ preserve enough chemical stoichiometry

$\Downarrow$

emulation $\forall \hat{\boldsymbol{v}}.\ F(\hat{\boldsymbol{v}} \circ m_S) = \hat{F}(\hat{\boldsymbol{v}}) \circ m_S$ preserve derivatives

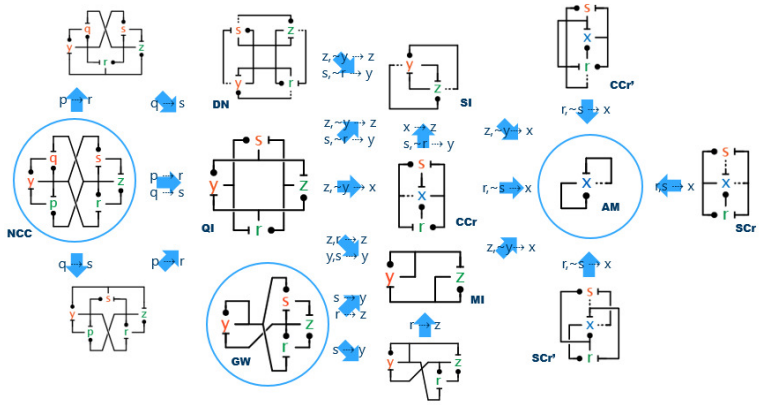
F is the *differential system* of (S, R), given by the law of mass action, $\hat{\boldsymbol{v}}$ is a state of $(\hat{S}, \hat{R})$. $\boldsymbol{\varphi}$ is the stoichiometric matrix and $\boldsymbol{\rho}$ is the related reactant matrix. $\boldsymbol{m_S}$ and $\boldsymbol{m_R}$ are the characteristic 0-1 matrices of the morphism maps m_S (on species) and m_R (on reactions). $-^{\mathbf{T}}$ is transpose. Homomorphism implies reactant morphism.

Thus, for *any initial conditions* of $(\hat{S}, \hat{R})$ we can initialize (S, R) to match its trajectories. And also (another theorem), for *any rates* of $(\hat{S}, \hat{R})$ we can choose rates of (S, R) that lead to emulation.

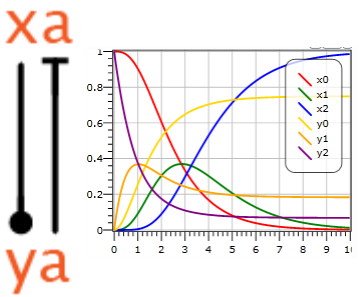


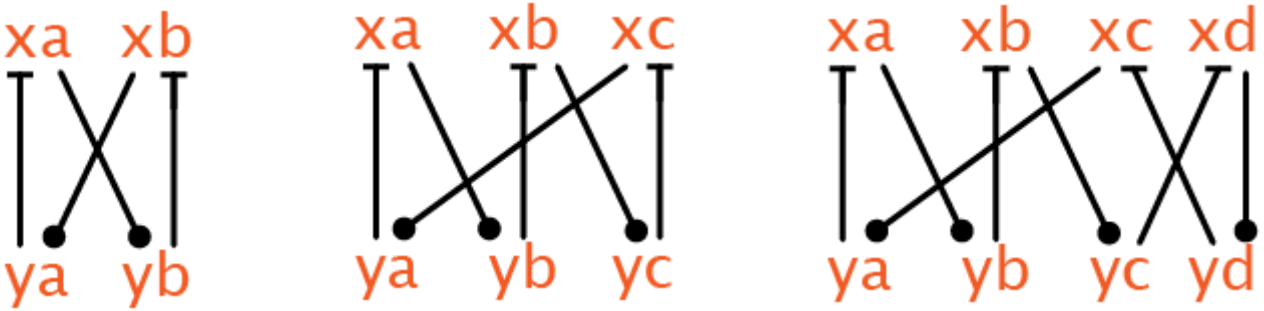
Corollaries

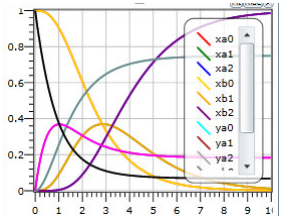
- By checking only static network and morphism properties we can learn that:
  - All these networks are (at least) bistable
  - (We do not have to reanalyze the steady states of all these dynamical systems)
  - All these networks can perform *exactly* as fast as AM
  - (We do not have to reprove the complexity bounds for all these networks)

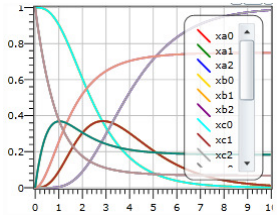


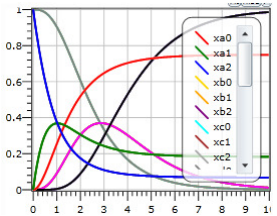
Another Zoo


xa
ya
x0
x1
x2
y0
y1
y2


xa xb
ya yb
xa xb xc
ya yb yc
xa xb xc xd
ya yb yc yd

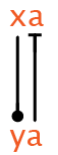

xa0
xa1
xa2
xb0
xb1
xb2
ya0
ya1
ya2

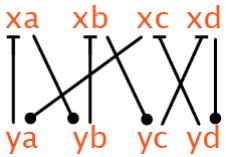

xa0
xa1
xa2
xb0
xb1
xb2
xc0
xc1
xc2


xa0
xa1
xa2
xb0
xb1
xb2
xc0
xc1
xc2

Network Perturbations

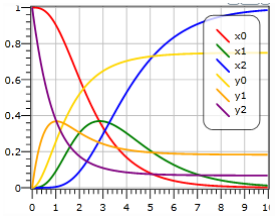
Network

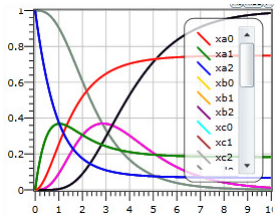




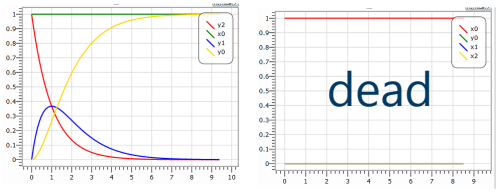
A complex but robust implementation of the simple network

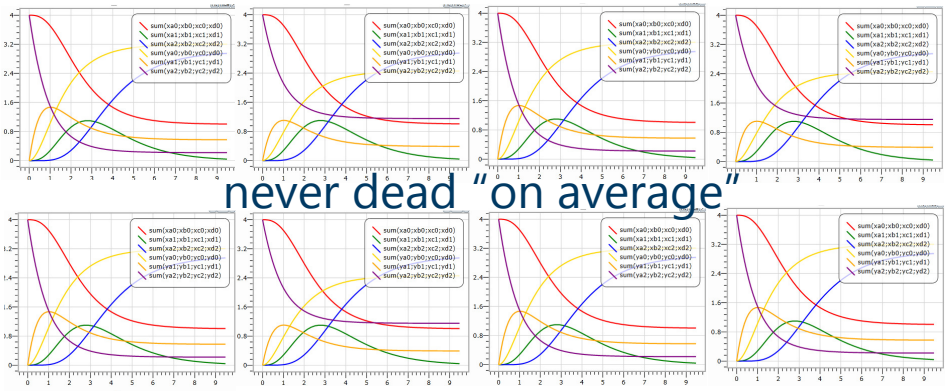
Normal Behavior





Removing each link in turn





Microsoft Research

UNIVERSITY OF OXFORD

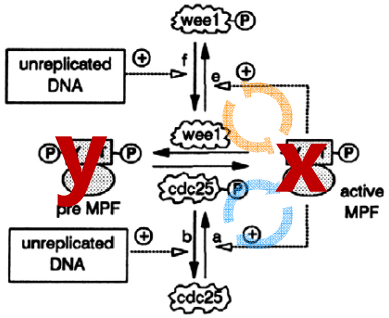
Cell Cycle Switch

The Cell Cycle Switch

Universal control mechanism regulating onset of M-phase

Paul Nurse

- This basic network is universal in Eukaryotes [P. Nurse]
  - The *switching function* and the *basic network* is *the same* from yeast to us.
    The human cdc2 gene can be replaced for the yeast one, and it works!
  - In particular detail, in frog eggs:



Double positive feedback on x
Double negative feedback on x
No feedback on y
Why ???

Journal of Cell Science 106, 1153-1168 (1993)
Printed in Great Britain © The Company of Biologists Limited 1993

Numerical analysis of a comprehensive model of M-phase control in *Xenopus* oocyte extracts and intact embryos

Bela Novak* and John J. Tyson†

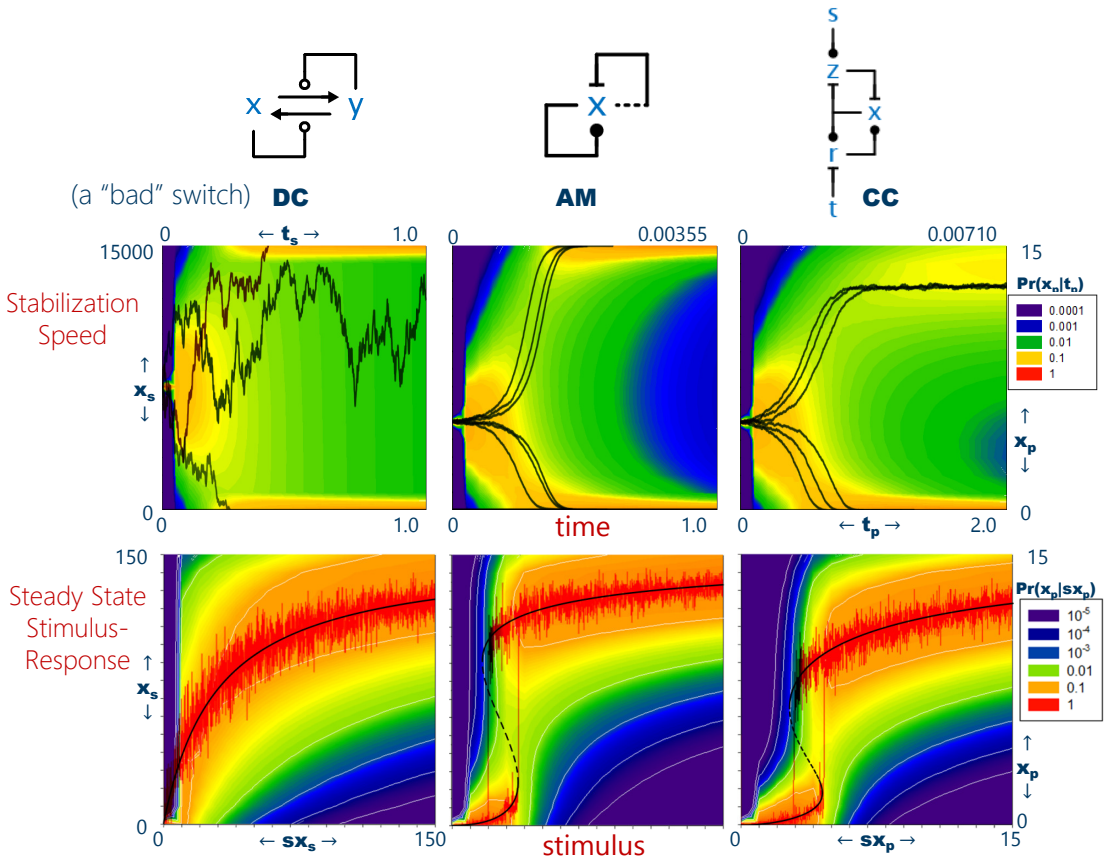
Department of Biology, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24060-0406, USA

*Permanent address: Department of Agricultural Chemical Technology, Technical University of Budapest, 1521 Budapest Gellert Ter 4, Hungary
†Author for correspondence

- The function is very well-studied. But why this network structure?
- That is, why this peculiar algorithm?

Cell Cycle vs AM

activation
inhibition
catalysis



The "classical" Cell Cycle Switch CC approximates AM performance

SCIENTIFIC REPORTS

OPEN

The Cell Cycle Switch Computes Approximate Majority

SUBJECT AREAS: COMPUTATIONAL BIOLOGY

Luca Cardelli[1] & Attila Csikász-Nagy[2,3]

CC converges in O(log n) time (like AM)
(but 2x slower than AM, and does not fully switch)

Symmetrical initial conditions ($x_0=x_1=x_2$)

Black lines: high-count stochastic simulation traces
Color: full probability distribution of low-count system

Hor axis is *time*.

AM shows hysteresis (like CC)

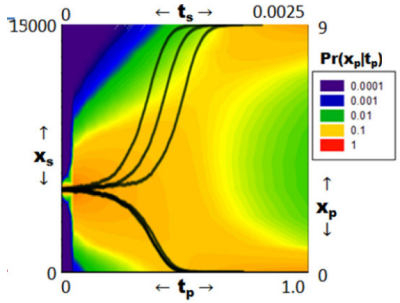
Black lines: deterministic ODE bifurcation diagrams
Red lines: medium-count stochastic simulations
Color: full probability distribution of low-count system

Hor axis is *stimulus* pushing towards x_0 against fixed bias.

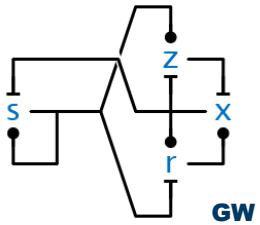
There is an obvious bug in CC performance: let's fix it!

Cell Cycle vs AM

- But GW is better!
  - Fully switchable, just as fast as AM
  - GW *emulates* AM



- That same week:
  - The Greatwall loop is a necessary component of the switch
  - So, nature fixed CC!



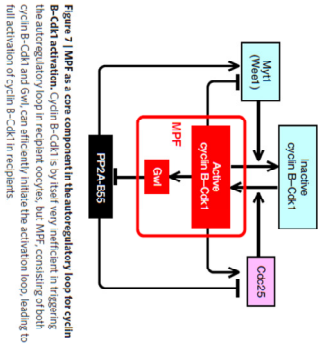


Figure 7 | MPF as a core component in the autoregulatory loop for cyclin B–Cdk1 activation. Cyclin B–Cdk1 s by itself very inefficient in triggering the autoregulatory loop in recipient oocytes, but MPF, consisting of both cyclin B–Cdk1 and Gwl, can efficiently initiate the activation loop, leading to full activation of cyclin B–Cdk1 in recipients

SCIENTIFIC REPORTS

OPEN

The Cell Cycle Switch Computes Approximate Majority

SUBJECT AREAS: COMPUTATIONAL BIOLOGY

Luca Cardelli[1] & Attila Csikász-Nagy[2,3]

nature COMMUNICATIONS

OPEN

ARTICLE

Received 6 Jul 2012 | Accepted 14 Aug 2012 | Published 11 Sep 2012

DOI:10.1038/ncomms2062

Greatwall kinase and cyclin B-Cdk1 are both critical constituents of M-phase-promoting factor

Masatoshi Hara[1,†], Yusuke Abe[1,†], Toshiaki Tanaka[2], Takayoshi Yamamoto[1,†], Eiichi Okumura[1] & Takeo Kishimoto[1]

A Theory of Network Emulation

(with thanks to David Soloveichik)

- So far, evidence is empirical
  - Simulations based on a choice of parameters
- But indeed…
  - We can show that, GW, NCC, etc. are *exactly* and *always* as good as AM
  - Where *exactly* means *numerically* as good, not just in the same complexity class
  - And *always* means for *any* choice of rates and initial conditions

Microsoft Research

UNIVERSITY OF OXFORD

Conclusions

Nature likes a good algorithm

Simulation

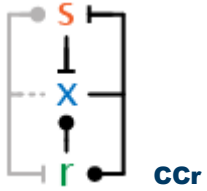
Approximate

"default" rates and initial conditions

Morphisms

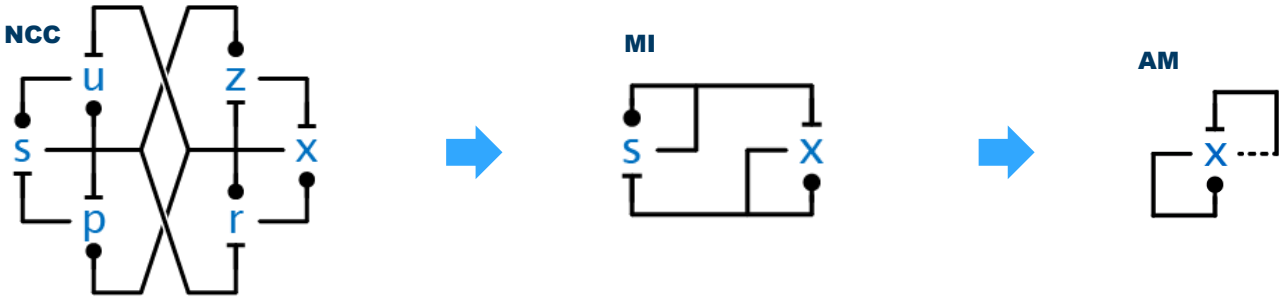
Exact

any rates and initial conditions



These additional feedbacks *do exist* in real cell cycles (via indirections)

Even the most recent, most complex, cell cycle switch *can* exactly emulate AM

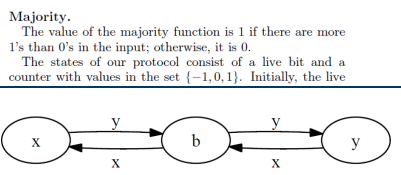


Interpretations of Network Morphisms

- Explanation of network structure
  - E.g. we know that the main function of Delta-Notch is to stabilize the system in one of two states. AM is the quintessential network that embodies fast robust bistability. The stoichiomorphism from Delta-Notch to AM "explains" what Delta-Notch (normally) does, and exactly how well it can do it.
- Robust implementation of simpler function
  - Redundant symmetries are implicit in the stoichiomorphism relationships
- Neutral paths in network space (evolution)
  - If an evolutionary event happens to be a stoichiomorphism, or close to it, it will not be immediately selected against, because it is "kinetically neutral".
  - This allows the network to increase its complexity without kinetic penalty.
  - Later, the extra degrees of freedom can lead to kinetic differentiation.
  - But meanwhile, the organism can explore variations of network structure.
- Network implementation (not abstraction!)
  - Stoichiomorphisms are not about abstraction / coarse-graining that preserve behavior, on the contrary, they are about *refinement* / *fine-graining* that preserve behavior.
  - They describe *implementations* of abstract networks, where the abstract networks themselves may not be (biologically) implementable because of excessive demands on species interactions.

Population Majority

2004: **Computation in networks of passively mobile finite-state sensors.** Dana Angluin, James Aspnes, Zoë Diamadi, Michael J. Fischer, René Peralta. PODC'04.

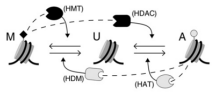


Exact Majority - 6-state
Nondeterministic.
(population protocol)

2007: **A Simple Population Protocol for Fast Robust Approximate Majority.** Dana Angluin, James Aspnes, David Eisenstat. DISC'07.

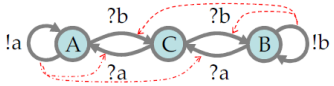
Approximate Majority - 3-state
Stochastic, discrete time
(DTMC) Fundamental results.

2007: **Theoretical Analysis of Epigenetic Cell Memory by Nucleosome Modification.** Ian B. Dodd, Mille A. Micheelsen, Kim Sneppen, Genevieve Thon. Cell.



Approximate Majority - 3-state
Stochastic, discrete time
(ad-hoc)

2009. **Artificial Biochemistry.** Luca Cardelli: Algorithmic Bioprocesses, Springer.



Approximate Majority - 3-state
Stochastic, **continuous time**
(CTMC)

2009: **Robust Stochastic Chemical Reaction Networks and Bounded Tau Leaping (Appendix 4).** David Soloveichick. J.Comput.Biol.

Transfer complexity results from discrete time population protocols to continuous time **stochastic chemical reaction networks**.

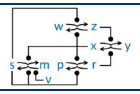
2009. **Using Three States for Binary Consensus on Complete Graphs.** Etienne Perron, Dinkar Vasudevan, and Milan Vojnovic. IEEE Infocom.

Approximate Majority - 3-state
Stochastic, **continuous time**
(CTMC) Fundamental results.

2010: **Convergence Speed of Binary Interval Consensus.** Moez Draief, Milan Vojnovic. Infocom'10.

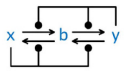
Exact Majority - 4-state
Stochastic, **continuous time**.

2012: **The Cell Cycle Switch Computes Approximate Majority.** Luca Cardelli, Attila Csikász-Nagy. Scientific Reports.



The biological cell cycle switch
is a (non-obvious) implementation of
approximate majority.

2014: **Morphisms of Reaction Networks that Couple Structure to Function.** Luca Cardelli.



Approximate Majority - 3-state
Continuous space, continuous time
(Deterministic ODE)