Microsoft Research

UNIVERSITY OF OXFORD

Morphisms of Reaction Networks

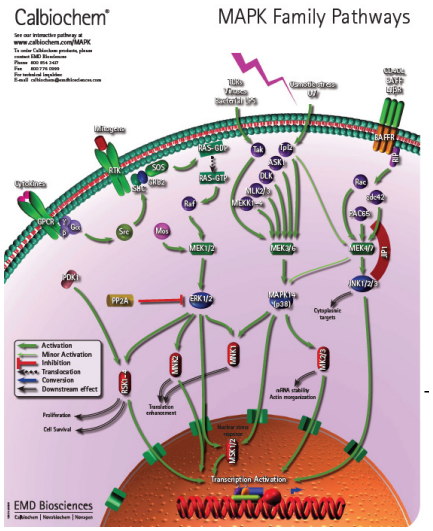
Luca Cardelli, Microsoft Research & Oxford University

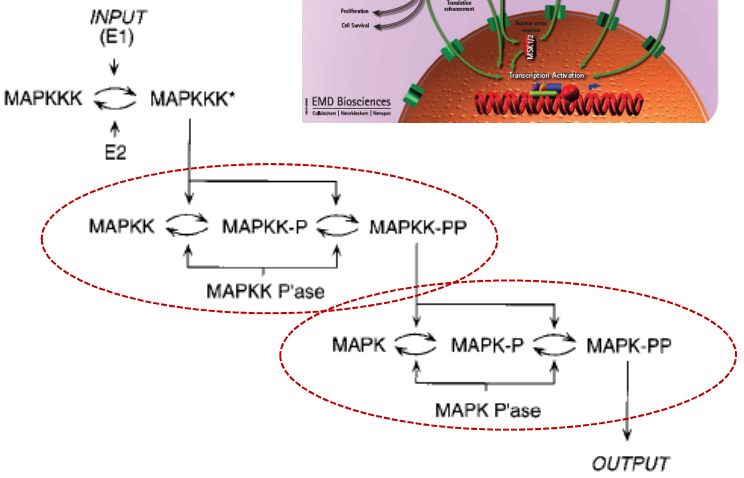
related work: Attila Csikász-Nagy, Neil Dalchau, Mirco Tribastone, Max Tschaikowski, Andrea Vandin

HSB, Madrid 2015-09-04

Cellular Computation

- No survival without computation!
  - Finding food
  - Avoiding predators
- How do cells compute?
  - *Clearly* doing "information processing"
  - What are their computational principles?

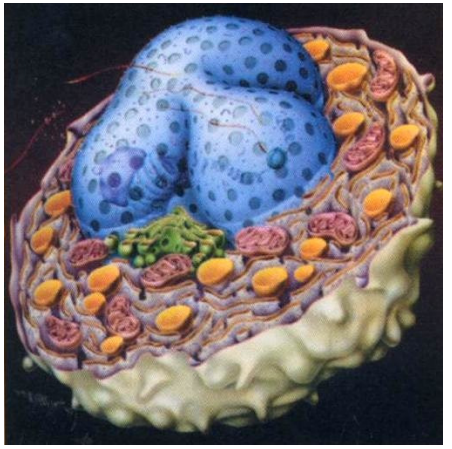




Ultrasensitivity in the mitogen-activated protein cascade, Chi-Ying F. Huang and James E. Ferrell, Jr., 1996, *Proc. Natl. Acad. Sci. USA*, 93, 10078-10083.

More concretely

- Give substance to the claim that "cells compute"
  - Yes, but *what* do they compute?
- Catch nature red-handed in the act of running a computational task
  - Something that a computer scientist would recognize as an *algorithm*



Chemical Algorithms

Can *Chemistry* Compute?

- If we believe that biology can do computation…
  - It must be somehow based on chemistry
- So, can chemistry compute, and how?
  - That is in itself a very interesting question with non-trivial answers

Chemical Programming Examples

specification

Y := min(X1, X2)

Y := max(X1, X2)

program

X1 + X2 -> Y

X1 -> L1 + Y
X2 -> L2 + Y
L1 + L2 -> K
Y + K -> 0

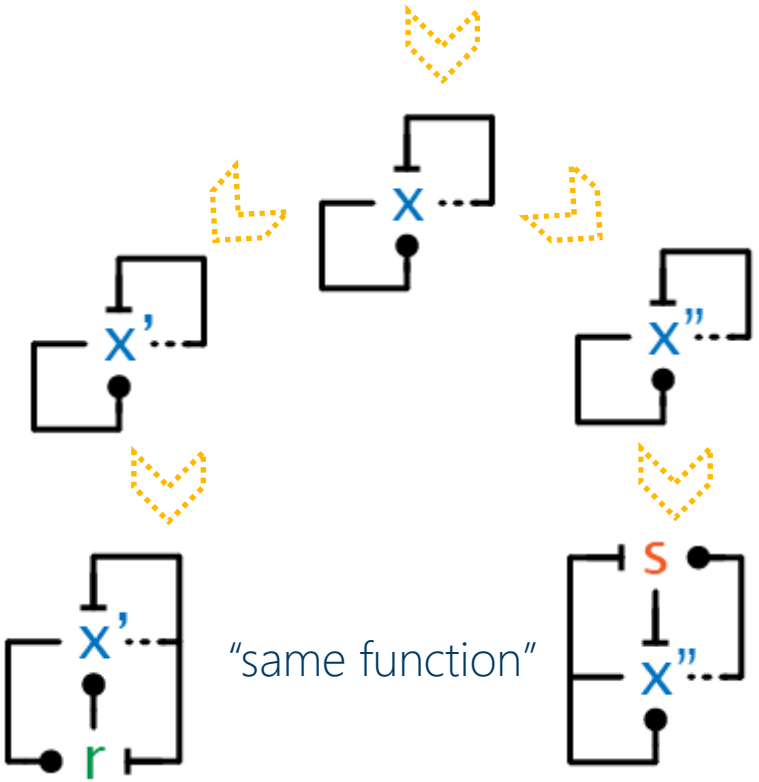
max(X1,X2)=
(X1+X2)-min(X1,X2)

(but is not computed "sequentially": it is a form of concurrent computation)

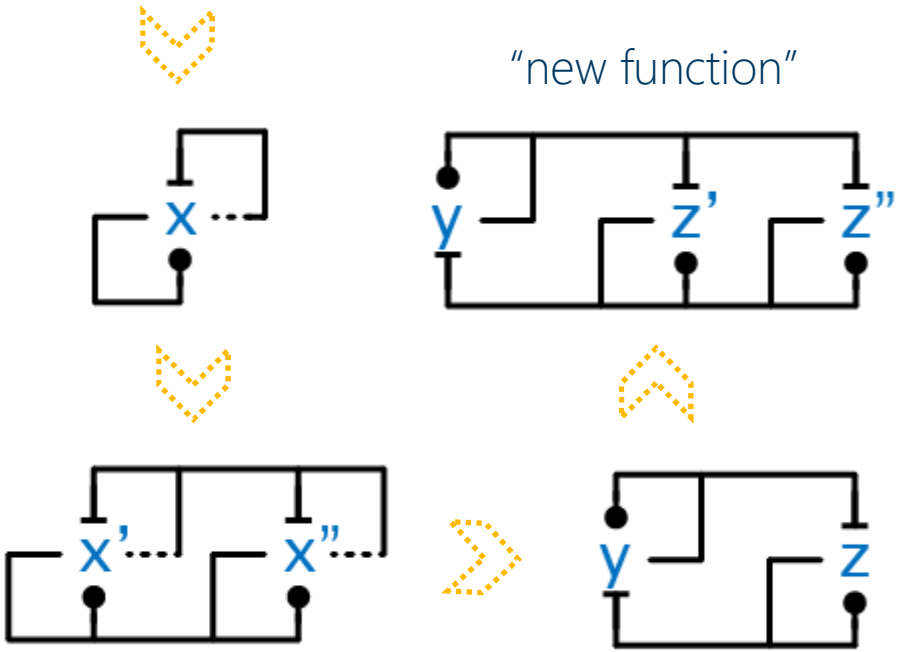
(bio-)chemical reaction network

Biochemical Networks

Across species: *Ortholog genes*



Within species: *Paralog genes*

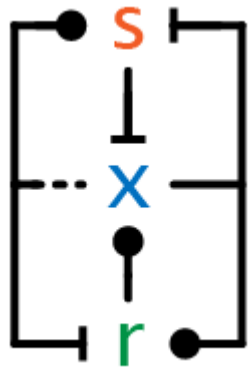


How do we know networks exist?

- If you can break it, it must exist
  - Genome sequencing identifies genes (their "coding" regions)
  - Sequence comparison identifies orthologs and paralogs
  - Gene-produced proteins are isolated or synthetically produced in vitro or in vivo (all difficult)
  - Their qual/quant interactions are studied (often only in vitro)
  - Their 3D structure is determined (may take decades)
  - Networks are hypothesized, often qualitatively
  - Models are build, quantitative function is inferred
  - Further experiments (such as gene knockouts) are performed to break the network.

- Genes and networks are compared across and within species
  - High-value activity: 2001 Nobel prize in Physiology for the discovery of *"Key regulators of the cell cycle … they have identified key molecules that regulate the cell cycle in all eukaryotic organisms, including yeast, plants, animals, and human."* These are actually not (currently) the same molecules, but it is (still) "the same network" in all of them.

Simplified example



- Genes for x, s, r identified
- Say protein x exists in high quantity
  - Knock gene-x out: one protein goes missing, that must be x's protein
- Say proteins s exists in "undetectable" quantities
  - Maybe 10~100 copies per cell on average: it *cannot be found*
  - Knock gene-s out: nothing seems to go missing, but the network's function stops
  - Then we know protein-s must be in the network, although we don't know "where"
- Heterogeneous system
  - It is indeed the case (in this cell-cycle-switch example) that x is "deterministic" (high copy count), while s,r, are "stochastic" (very low copy count) and yet s,r control x.

Consensus Networks

A Consensus Problem

- Population Consensus
  - Given two populations of x and y "agents"
  - We want them to "reach consensus"
  - By converting *all* agents to x or to y depending on which population was in majority initially
- Population Protocols Model
  - Finite-state identity-free agents (molecules) interact in randomly chosen pairs ($\Rightarrow$ stochastic symmetry breaking)
  - Each interaction (collision) can result in state changes
  - Complete connectivity, no centralized control (well-mixed solution)

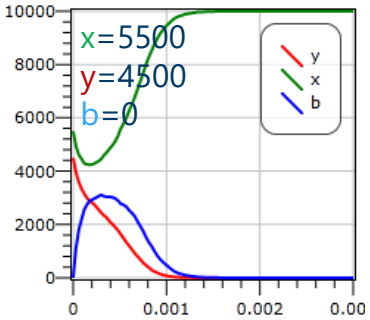
specification

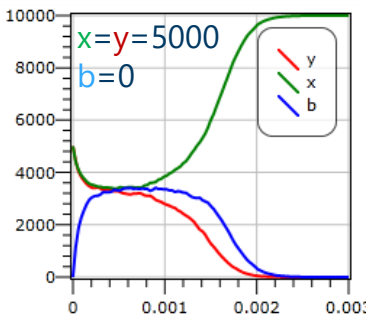
$X,Y := X+Y, 0 \quad \text{if} \quad X_0 \geq Y_0$

$X,Y := 0, X+Y \quad \text{if} \quad Y_0 \geq X_0$

A Consensus Algorithm

- Approximate Majority (AM) Algorithm
  - Uses a third "undecided" population b
  - Disagreements cause agents to become undecided
  - Undecided agents agree with any non-undecided agent

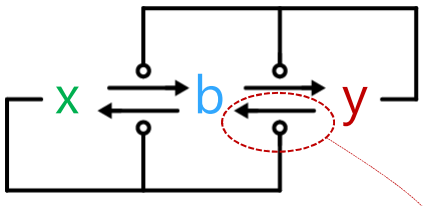




Dana Angluin · James Aspnes · David Eisenstat

A Simple Population Protocol for Fast Robust Approximate Majority

catalysis



chemical reaction network

$x + y \to^r y + b$

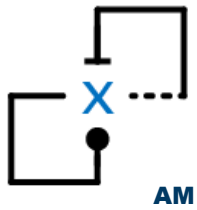
$y + x \to^r x + b$

$b + x \to^r x + x$

$b + y \to^r y + y$

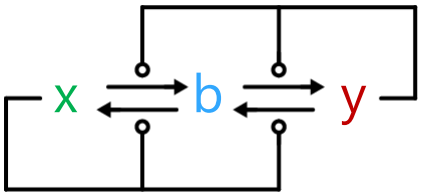
activation

inhibition



A Biological Implementation

Approximate Majority (AM)



1) Bistable
   Even when initially x=y (stochastically)

2) Fast (asymptotically optimal)
   O(*log n*) convergence time

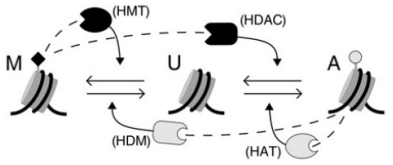
3) Robust to perturbation
   above a threshold, initial majority wins *whp*

Dana Angluin · James Aspnes · David Eisenstat

A Simple Population Protocol for Fast Robust Approximate Majority

2007

Epigenetic Switch



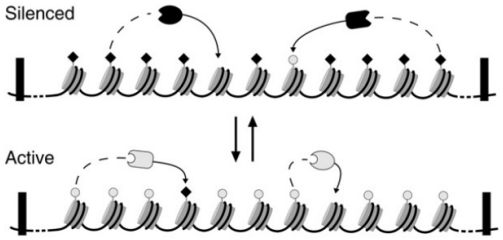
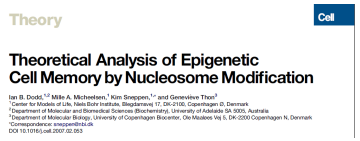


Figure 1. Basic Ingredients of the Model


Theory

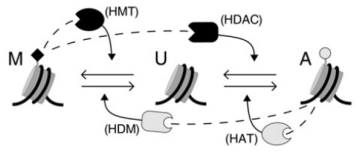
Cell

Theoretical Analysis of Epigenetic Cell Memory by Nucleosome Modification

2007

Here We Got Lucky

- We can claim that the epigenetic switch is a *direct* biological implementation of an algorithm
  - Although we may have to qualify that with some notion of approximation of the (enzymatic) kinetics



- In most cases the biological implementation seems more *indirect* or *obfuscated*
  - *"Nature is subtle but not malicious - Einstein"* Ha! think again!
  - Other implementations of Approximate Majority seem more convoluted and approximate

How to model "Influence"

"True" molecular interactions.

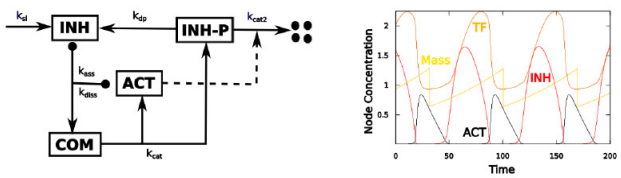


Figure 3: a) Schematic diagram of a simplified SIMM model [17]. The activa-

"Equivalent" influence interactions.

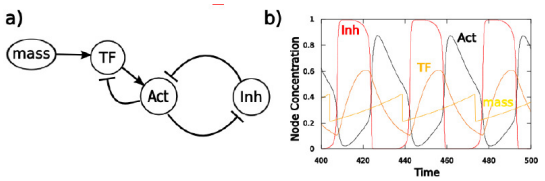


Figure 4: a) Schematic diagram of a primitive cell cycle in the reinitz framework.

Chemical Reaction Network ⟷ Influence Network

Evolving a Primitive Eukaryotic Cell Cycle Model

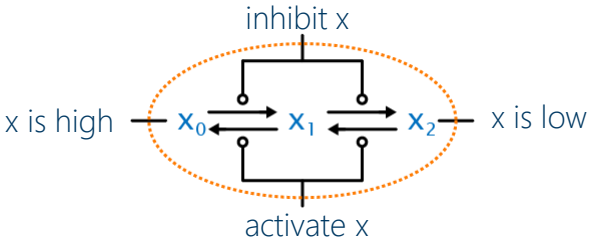
Malte Lücken, Jotun Hein, Bela Novak

Instead of modeling basic interactions, such as binding, synthesis, and degradation of molecular components, this framework models interactions simply as activation or inhibition. This approach also reduces the number of nodes necessary in the network, as e.g. the inhibitor binding tightly to the activator to form a complex, which produces phosphorylated inhibitor to be degraded under catalysis by the activator, is now simply a double negative feedback loop shown in Figure 1. This type of interaction is the basis of both aforementioned molecular model, therefore they can both be summarized in a single Reinitz model.

The Triplet Model of Influence

activation
inhibition
catalysis

=



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

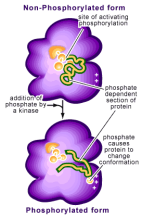
Functional Motifs in Biochemical Reaction Networks

John J. Tyson[1] and Béla Novák[2]

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

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

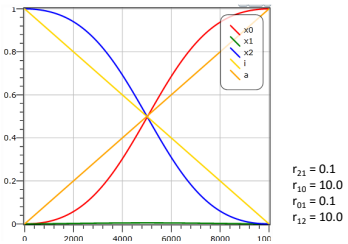
biological mechanism:
(e.g.:) multisite phosphorylation



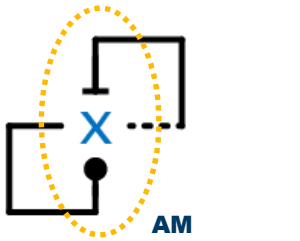
triplet motif

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

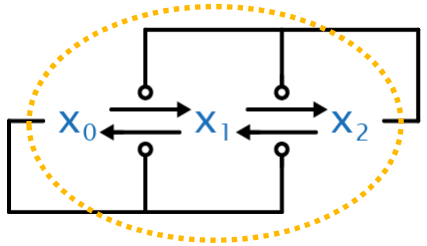
They actually implement a
Hill function of coefficient 2:



For example:



=



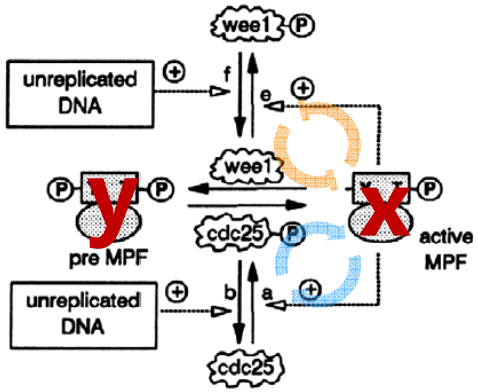
Approximate Majority

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.
  - In particular detail, in frog eggs, G_2/M transition:

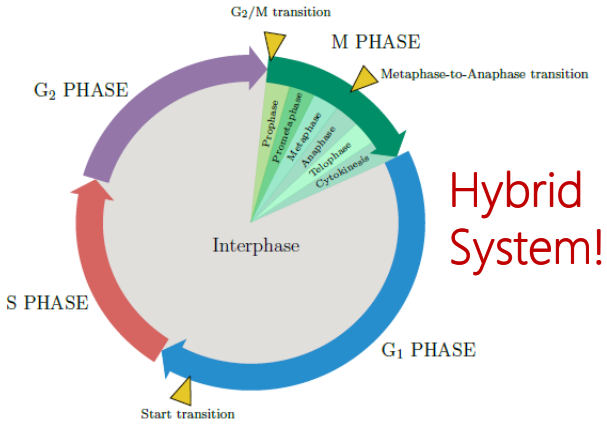


Double positive feedback on x
Double negative feedback on x
No feedback on y. Why ???

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



Hybrid System!

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

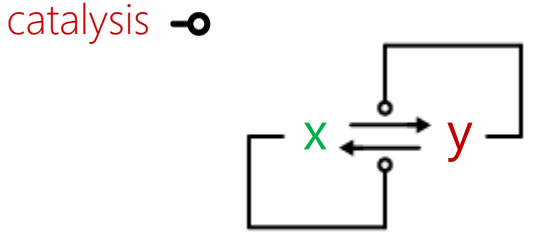
How to Build a Good Switch

- We need first a bistable system: one that has two *distinct* and *stable* states. I.e., given any initial state the system must settle into one of two states
- The settling must be fast (not get stuck in the middle for too long) and robust (must not spontaneously switch back)
- Finally, we need to be able to flip the switch by external inputs

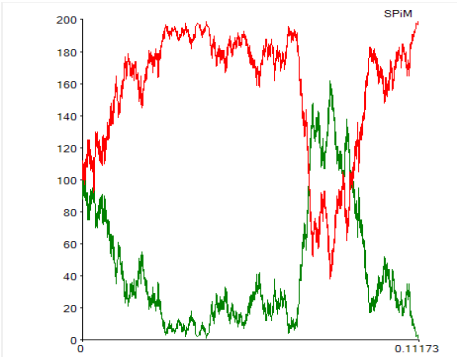
A Bad Algorithm

- Direct Competition
  - x catalyzes the transformation of y into x
  - y catalyzes the transformation of x into y
  - when all-x or all-y, it stops

- This system has two end states, but
  - Convergence to an end state is slow (a random walk)
  - Any perturbation of an end state can start a random walk to the other end state (hence not really *bistable*)



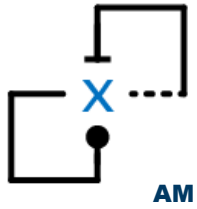
$$y + x \rightarrow x + x$$
$$x + y \rightarrow y + y$$



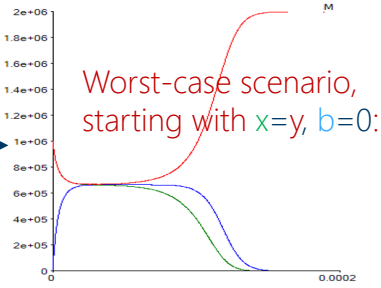
A Good Algorithm

- Approximate Majority (AM)
  - 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 before converging is $O(n \log n)$
    ⇒ fast (optimal)
  - The final outcome is correct if the initial disparity is $\omega(\sqrt{n} \log n)$
    ⇒ solution states are robust to perturbations
- Logarithmic time bound in parallel time
  - *Parallel time* is the number of steps divided by the number of agents
  - In parallel time the algorithm converges with high probability in $O(\log n)$

activation
inhibition



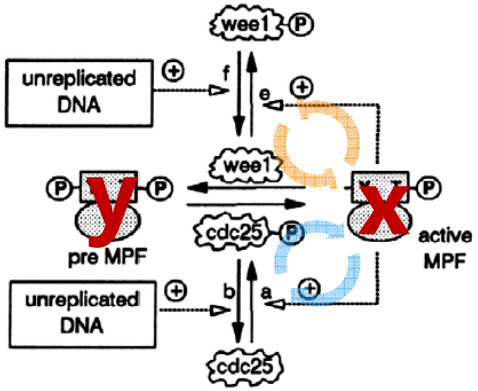
x=y



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

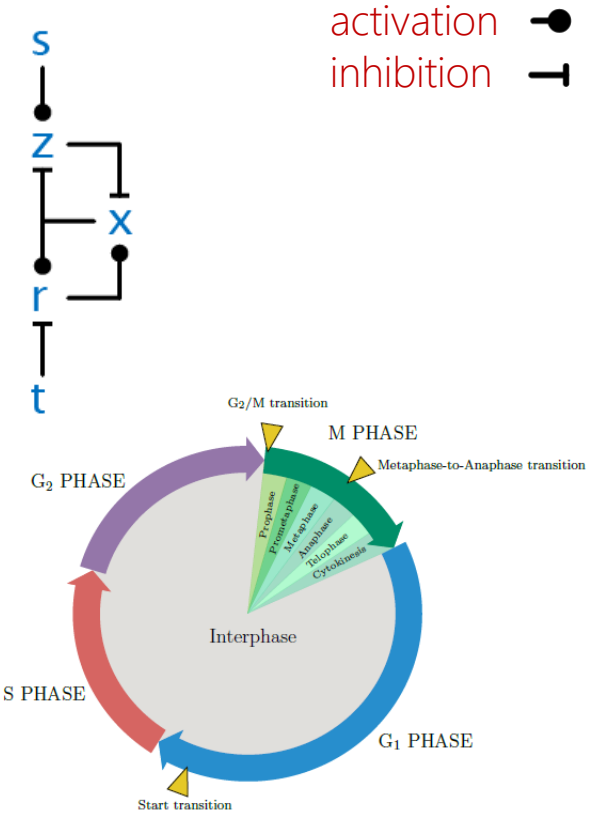
Dana Angluin · James Aspnes · David Eisenstat
A Simple Population Protocol for Fast Robust Approximate Majority

An "Ugly" Algorithm: Cell Cycle Switch



Nobel-prize winning network

Obfuscation of a distributed algorithm?

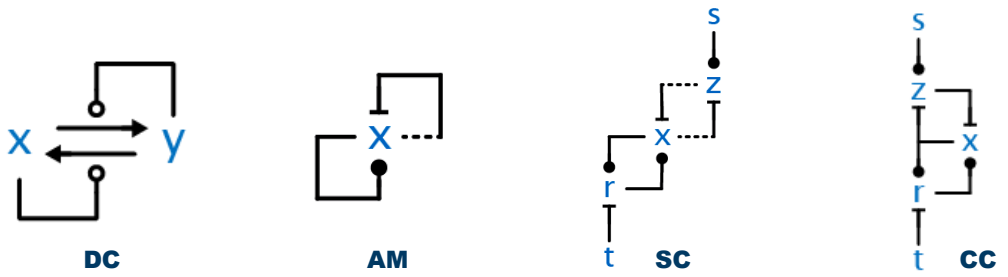


- Is it a good algorithm? Is it bad?
- Is it optimal or suboptimal?

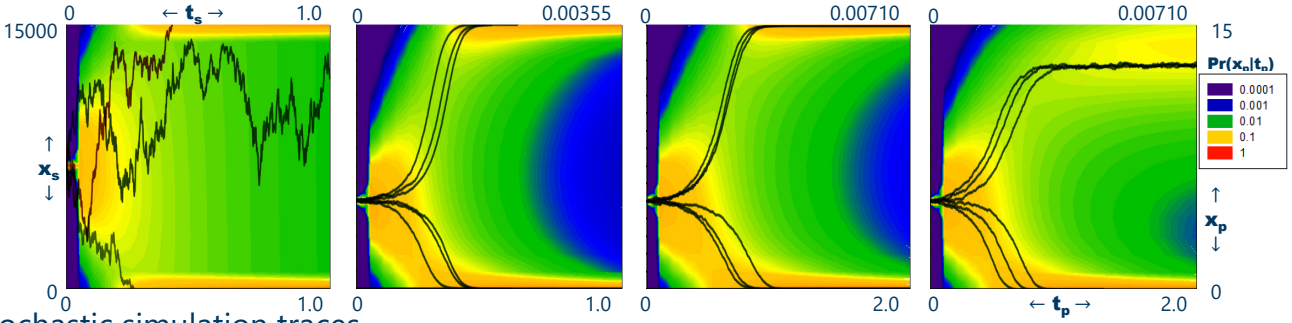
Convergence Analysis - CONSENSUS

- Switches as computational systems

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



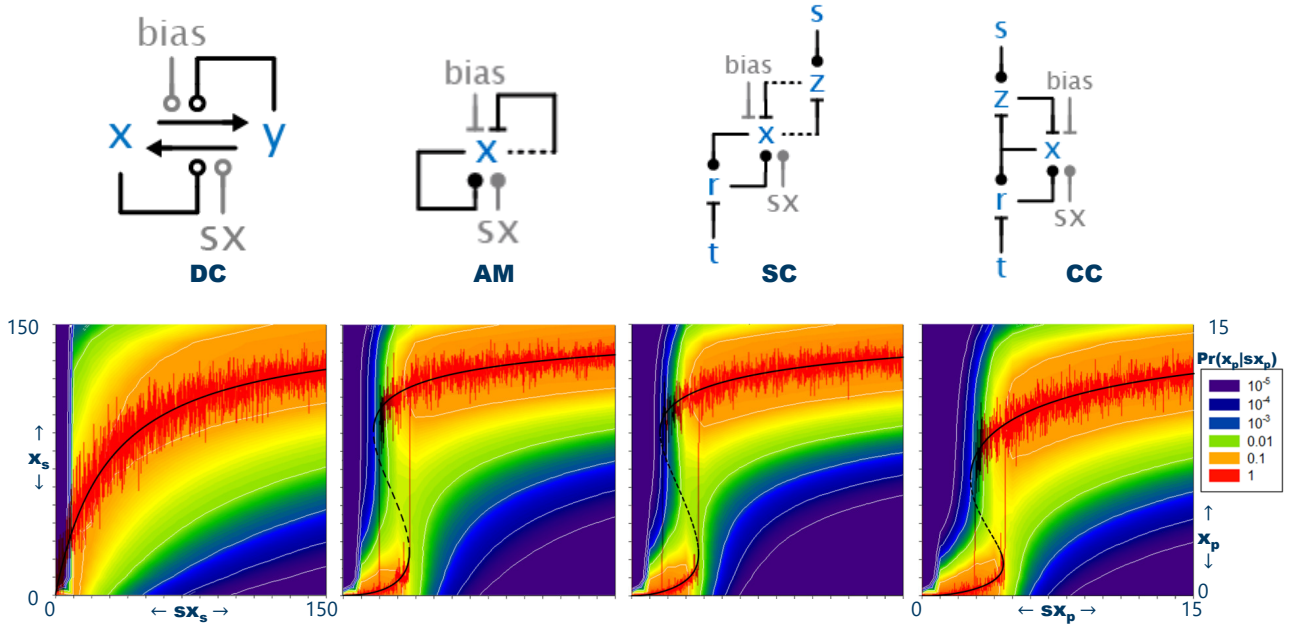
Start symmetrical
(x_0=x_1=x_2 etc.)



Black lines: several stochastic simulation traces
Color: full probability distribution of small-size system

Steady State Analysis - SWITCH

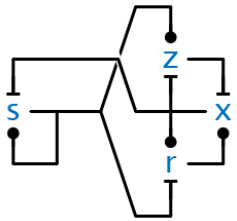
- Switches as dynamical systems



Black lines: deterministic ODE bifurcation diagrams
Red lines: noisy stochastic simulations
Color: full probability distribution of small-size system

The Cell Cycle Switch Computes AM

- Our paper appeared:
  - Suggesting GW is a better switch than CC. *September 2012*
- Another paper that same week:
  - Showing experimentally that the Greatwall loop is a necessary component of the switch, i.e. the not-as-good-as-AM network has been 'refuted'



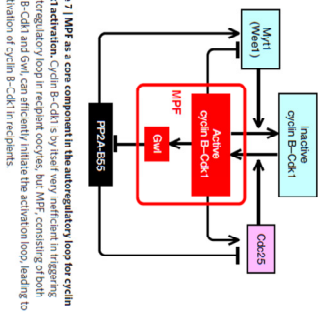


Figure 7 | MPF as a core component in the autoregulatory loop for cyclin B–Cdk1 activation. Cyclin B–Cdk1 is synthesized very inefficient in triggering the autoregulatory loop in recipient oocytes, but MPF, consisting of both cyclin B–Cdk1 and Gwl, can efficiently initiate the activation loop, leading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

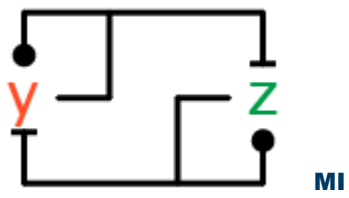
Greatwall kinase and cyclin B-Cdk1 are both critical constituents of M-phase-promoting factor

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

Antagonistic Networks

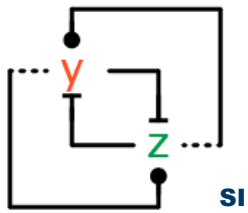
activation
inhibition

1 vs. 1
Mutual Inhibition & Self Activation



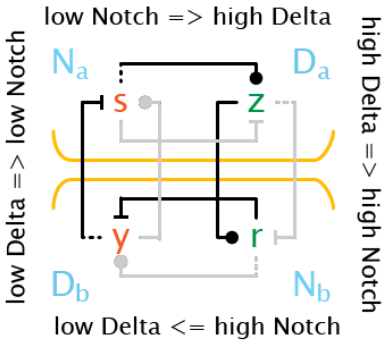
MI

1 vs. 1
Mutual Inhibition & Mutual Anti-activation

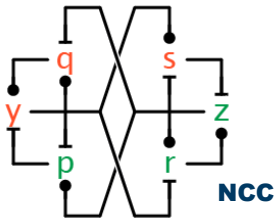


SI

2 vs. 2

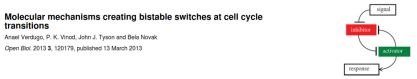


3 vs. 3



NCC

Cell cycle transitions

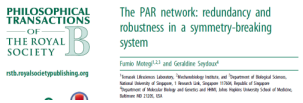


Molecular mechanisms creating bistable switches at cell cycle transitions

Anael Verdugo, P. K. Vinod, John J. Tyson and Béla Novak

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

Polarity establishment

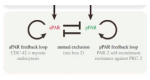


PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

rstb.royalsocietypublishing.org

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

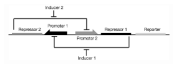
Fumio Motegi[1,2] and Geraldine Seydoux[3]



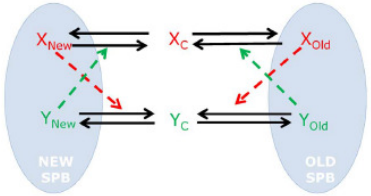
Gene networks

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

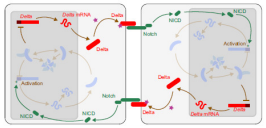


Dynamics of SIN Asymmetry Establishment

Archana Bajpai[1], Anna Fookstsova[2], Jun-Song Chen[2], Dannel McCollum[3], Masamitsu Sato[4,5], Rafael E. Carazo-Salas[6], Kathleen L. Gould[2], Attila Csikasz-Nagy[1,7]*

PLOS Computational Biology | www.ploscompbiol.org

Delta-Notch



Patterning embryos with oscillations: structure, function and dynamics of the vertebrate segmentation clock

Andrew C. Oates[1,2], Luis G. Morelli[1,2] and Saúl Ares[3,4,*]

Lateral Inhibition through Delta-Notch Signaling: A Piecewise Affine Hybrid Model*

Ronojoy Ghosh and Claire J. Tomlin

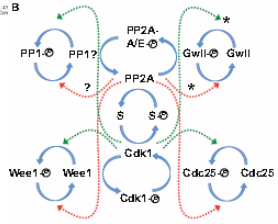
M.D. Di Benedetto, A. Sangiovanni-Vincentelli (Eds.): HSCC 2001, LNCS 2034, pp. 232–246, 2001.
© Springer-Verlag Berlin Heidelberg 2001

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[1,2]

[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 (daniel.fisher@igmm.cnrs.fr)
[†]These authors contributed equally to this work

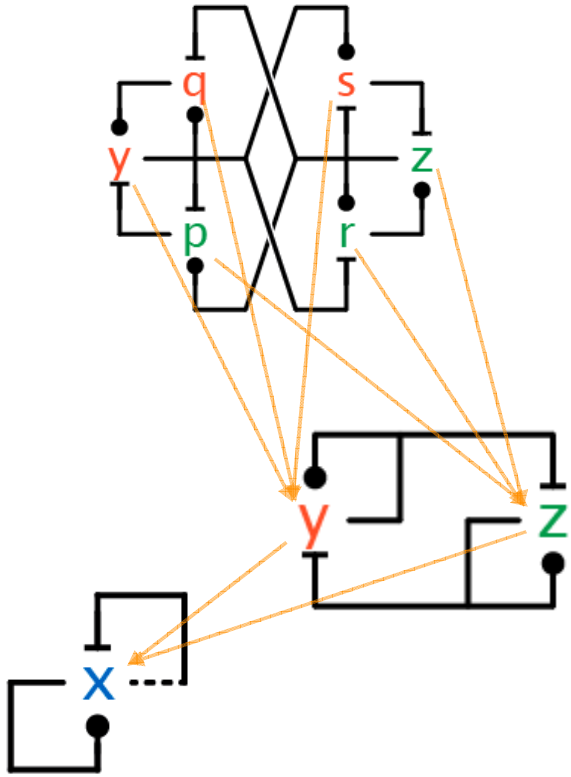


Network Morphisms

When does a (complex) network implement a (simpler) algorithm?

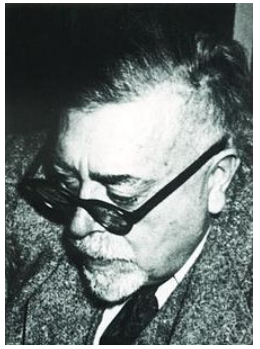
Comparing networks

- How can we compare different networks?
  - Different number of species
  - Different number of reactions
  - Apparently unrelated connectivity
- So that we can compare their function?
  - Does antagonism (in network structure) guarantee bistability (in function)?
- We do it by *mapping* networks onto one another so that they *emulate* each other
  - Deterministic semantics version of "simulation" of systems
  - (Stochastic semantics was the starting point, but too difficult/demanding for typical biological networks.)



Mapping one network into another

- Notion is strangely missing from the literature
  - Seen in Biology: single-network analysis (e.g. structure of feedback loops) and network reduction (e.g. while preserving steady states). Study of common or frequent subnetworks.
  - Seen in C.S.: comparing network *behaviors* (e.g. morphisms of event structures).
  - Nothing much resembling (bi)simulation "on the syntax" (structure) of whole biochemical networks.
- Model reduction is unavoidable and pervasive, but
  - Often criticized/ignored by biologists when it leads to quantities that are "not biologically meaningful". E.g. a fusion or change a variables in the ODEs where the new variables do not correspond to biological parts. The reduced model should "inform" the original one.
- Science's ethos
  - The "truth" is the big network, not the small one!
    If you depart from the truth in any way, you have to explain how you can get back to it.
  - The point is not to reduce the size of the network (although that's neat),
    but to understand aspects of *the big network* by reference to a smaller one.
  - The mapping is more important than either networks.



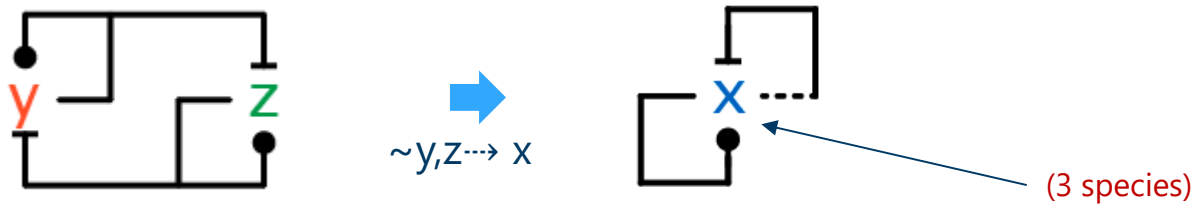
Norbert Wiener

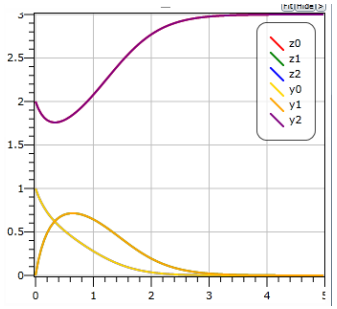
Pioneer of stochastic processes and inventor of Cybernetics.

"The best material model of a cat is another, or preferably the same, cat"

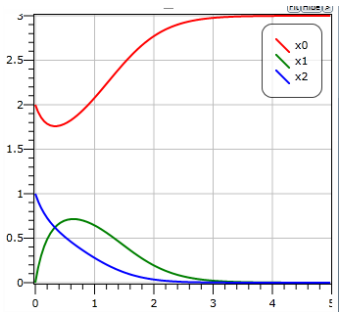
Network Emulation E.g.: MI emulates AM

- 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!

CRN Morphisms

A *CRN morphism* from (S, R) to $(\hat{S}, \hat{R})$
written $m \in (S, R) \to (\hat{S}, \hat{R})$

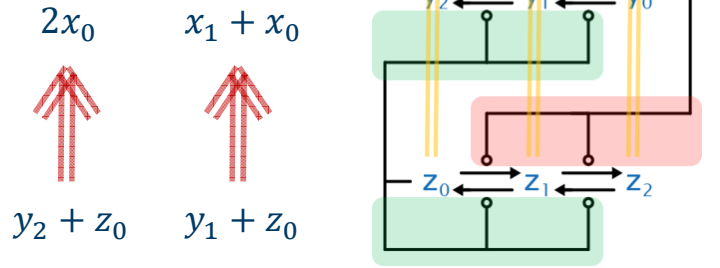
is a pair of maps $m = (m_S, m_R)$

a *species map* $m_S \in S \to \hat{S}$

a *reaction map* $m_R \in R \to \hat{R}$

extended to a *complex map* $m_S \in \mathbb{N}^S \to \mathbb{N}^{\hat{S}}$
linearly: $m_S(\rho)_{\hat{s}} = \Sigma_{s \in m_S^{-1}(\hat{s})} \rho_s$

Mappings (symmetries) between two networks



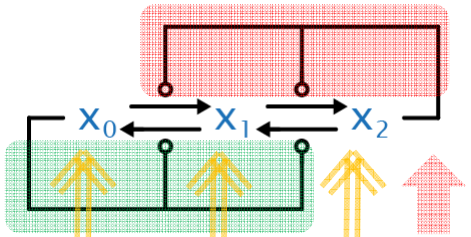
How to check emulations

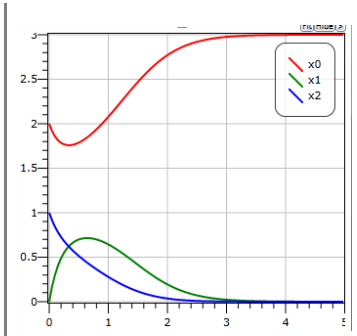
- How do we check a potential emulation morphism for all possible initial conditions of the target?
  - Statically! Check conditions on the joint stoichiometric matrices of the two networks under the mapping.
- How do we check a potential emulation morphism for all possible rates of the target?
  - Can't; but if one emulation is found, then the rates of the target network can be changed *arbitrarily* and a related emulation will again exist.

Network Emulation: MI emulates AM

A mapping of species and reactions

AM



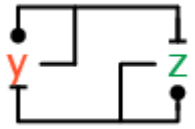


any initial conditions

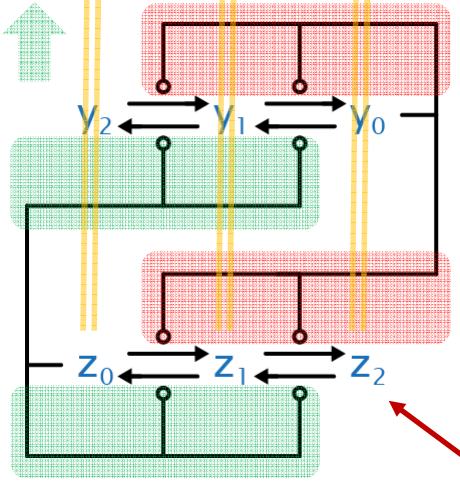
homomorphic mapping

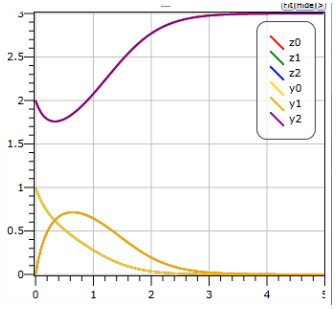
z -> x

~y -> x



MI





initial conditions:

$z_0 = y_2 = x_0$

$z_1 = y_1 = x_1$

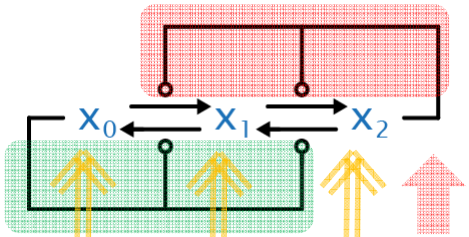
$z_2 = y_0 = x_2$

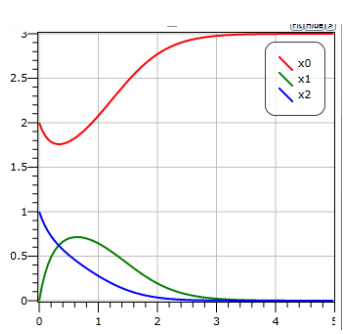
less trivial than you might think:
it need not preserve the out-degree of a node!

Network Emulation: SI emulates AM

A mapping of species and reactions

AM



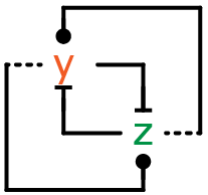


any initial conditions

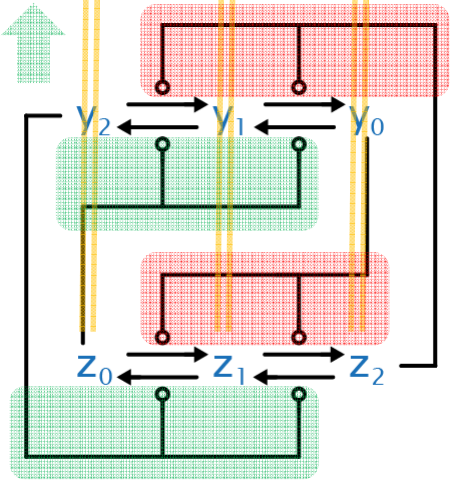
homomorphic mapping

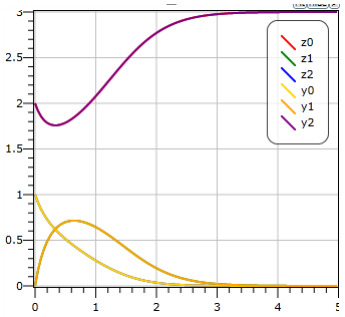
z -> x

~y -> x



SI





initial conditions:

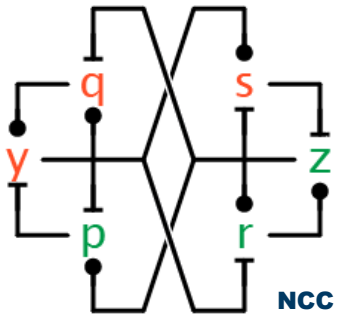
$z_0 = y_2 = x_0$

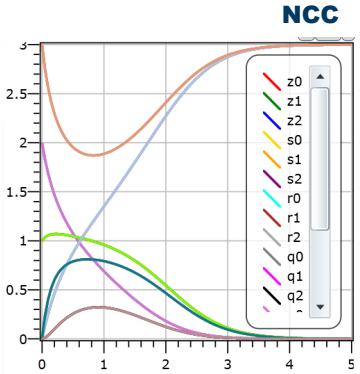
$z_1 = y_1 = x_1$

$z_2 = y_0 = x_2$

Network Emulation: NCC emulates MI

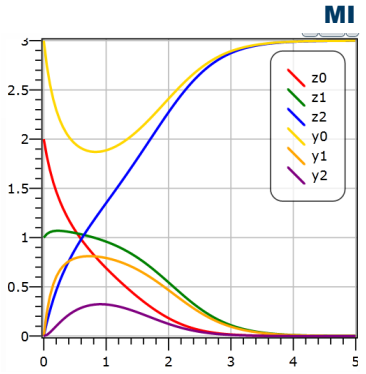
- For *any* rates and initial conditions of MI we can find *some* rates and initial conditions of NCC such that the (18) trajectories of NCC retrace those (6) of MI





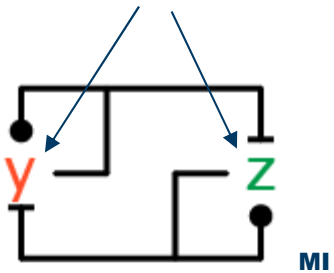
(18 species on 6 trajectories)

z,r,p ⇢ z
y,q,s ⇢ y



(6 species on 6 trajectories)

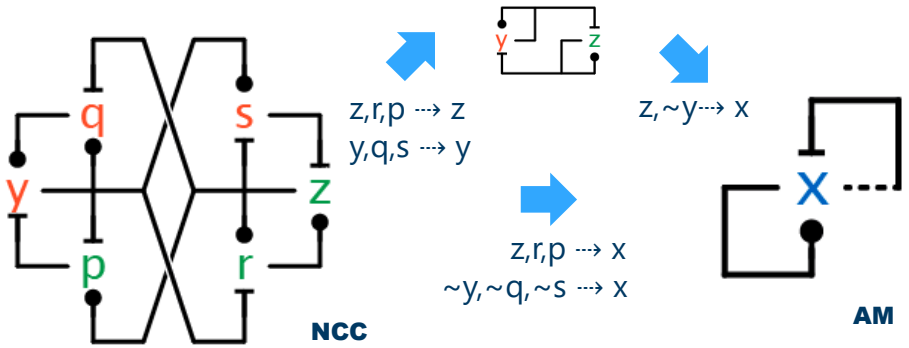
(3 species each)

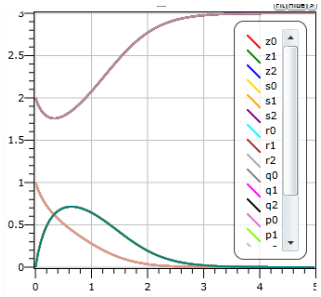


initialize
z,r,p = z
y,q,s = y

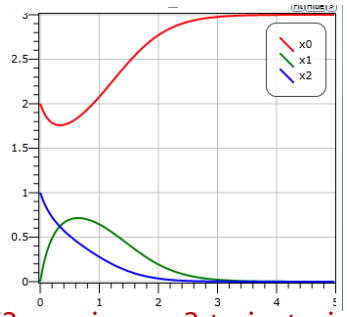
Emulations Compose

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





(18 species on 3 trajectories)



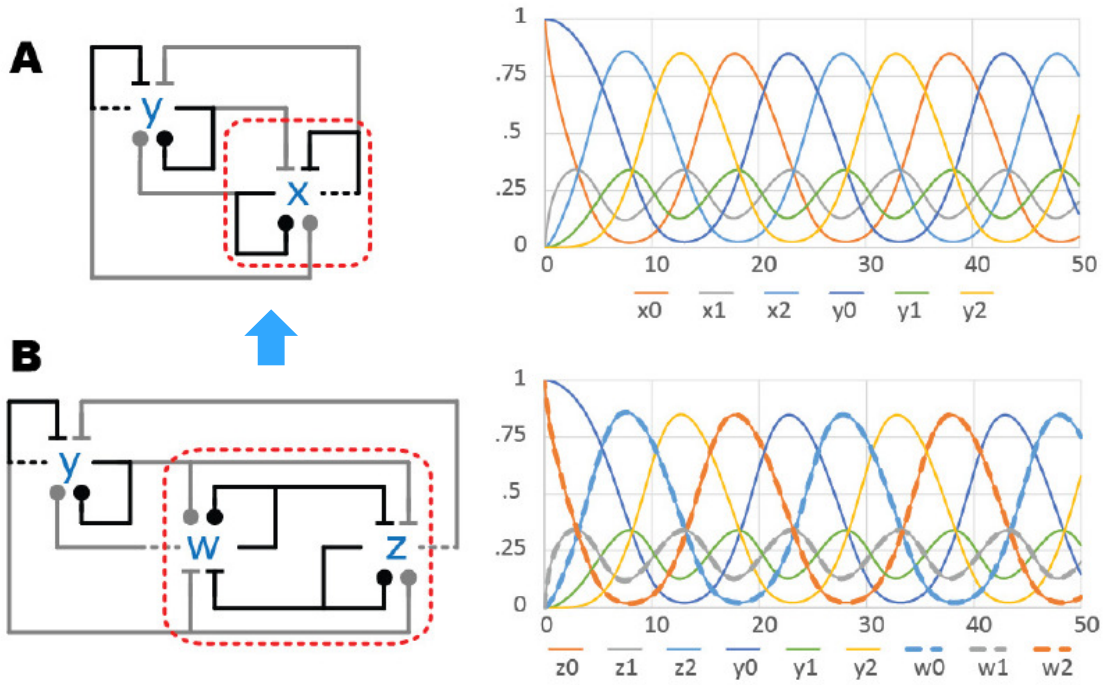
(3 species on 3 trajectories)

The new cell cycle switch can emulate AM *exactly*.

For *any* initial conditions of AM.

And for *any* rates of AM.

Emulations are Modular



Static Test for Emulation

Emulation 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.

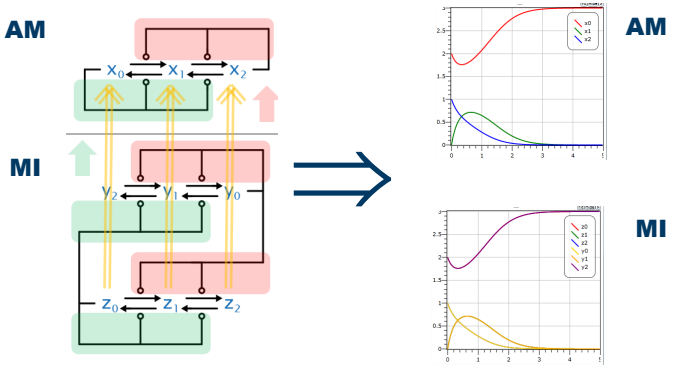
Cardelli *BMC Systems Biology* 2014, **8**:84
http://www.biomedcentral.com/1752-0509/8/84

BMC Systems Biology

RESEARCH ARTICLE Open Access

Morphisms of reaction networks that couple structure to function

Luca Cardelli[1,2]



Stoichiomorphims condition is *sufficient* for "networks of interest" and actually "close" to a *necessary* condition.

Emulation is (Backward) Bisimulation

▶ **Definition 13** (Cumulative flux rate). Let (S, R) be a CRN, $X \in S$, $\rho \in \mathcal{MS}(S)$, and $\mathcal{M} \subseteq \mathcal{MS}(S)$. Then, we define

$$\mathbf{fr}(X, \rho) := \sum_{\rho \xrightarrow{\alpha} \pi \in R} (\pi(X) - \rho(X)) \cdot \alpha, \qquad \mathbf{fr}[X, \mathcal{M}] := \sum_{\rho \in \mathcal{M}} \mathbf{fr}(X, \rho).$$

We call $\mathbf{fr}(X, \rho)$ and $\mathbf{fr}[X, \mathcal{M}]$ *ρ-flux rate* and *cumulative $\mathcal{M}$-flux rate* of X, respectively.

▶ **Definition 14** (Backward CRN bisimulation). Let (S, R) be a CRN, $\mathcal{R}$ an equivalence relation over S, $\mathcal{H} = S/\mathcal{R}$ and μ the choice function of $\mathcal{H}$. Then, $\mathcal{R}$ is a backward CRN bisimulation (BB) if for any $(X, Y) \in \mathcal{R}$ it holds that

$$\mathbf{fr}[X, \mathcal{M}] = \mathbf{fr}[Y, \mathcal{M}] \quad \text{for all} \quad \mathcal{M} \in \{\rho \mid \rho \xrightarrow{\alpha} \pi \in R\}/\approx_{\mathcal{H}}, \tag{2}$$

where any two $\rho, \sigma \in \mathcal{MS}(S)$ satisfy $\rho \approx_{\mathcal{H}} \sigma$ if $\mu(\rho) = \mu(\sigma)$.

▶ **Theorem 17** (Backward bisimulation characterizes exact fluid lumpability). *Let (S, R) be a CRN. Then, $\mathcal{H}$ is an exactly fluid lumpable partition of S if and only if $\mathcal{H}$ is a BB of S.*

Forward and Backward Bisimulations for Chemical Reaction Networks

Luca Cardelli[1], Mirco Tribastone[2], Max Tschaikowski[3], and Andrea Vandin[4]

1 Microsoft Research & University of Oxford, UK
luca@microsoft.com
2-4 University of Southampton, UK
{m.tribastone,m.tschaikowski,a.vandin}@soton.ac.uk

Concur 2015

An emulation between two CRNs can be understood in terms of a backward CRN bisimulation over the species of a "union CRN" that contains all the species and reactions of the two CRNs of interest.

Applications of BB

- Model Reduction
  - Find reduced networks
  - Compute quotient CRNs
  - Find network symmetries that may be of biological interest
- Morphism Generation
  - Find morphisms between networks (e.g. all the ones for a fixed rate assignment)

Forward and Backward Bisimulations for Chemical Reaction Networks

Luca Cardelli[1], Mirco Tribastone[2], Max Tschaikowski[3], and Andrea Vandin[4]

1 Microsoft Research & University of Oxford, UK
luca@microsoft.com
2-4 University of Southampton, UK
{m.tribastone,m.tschaikowski,a.vandin}@soton.ac.uk

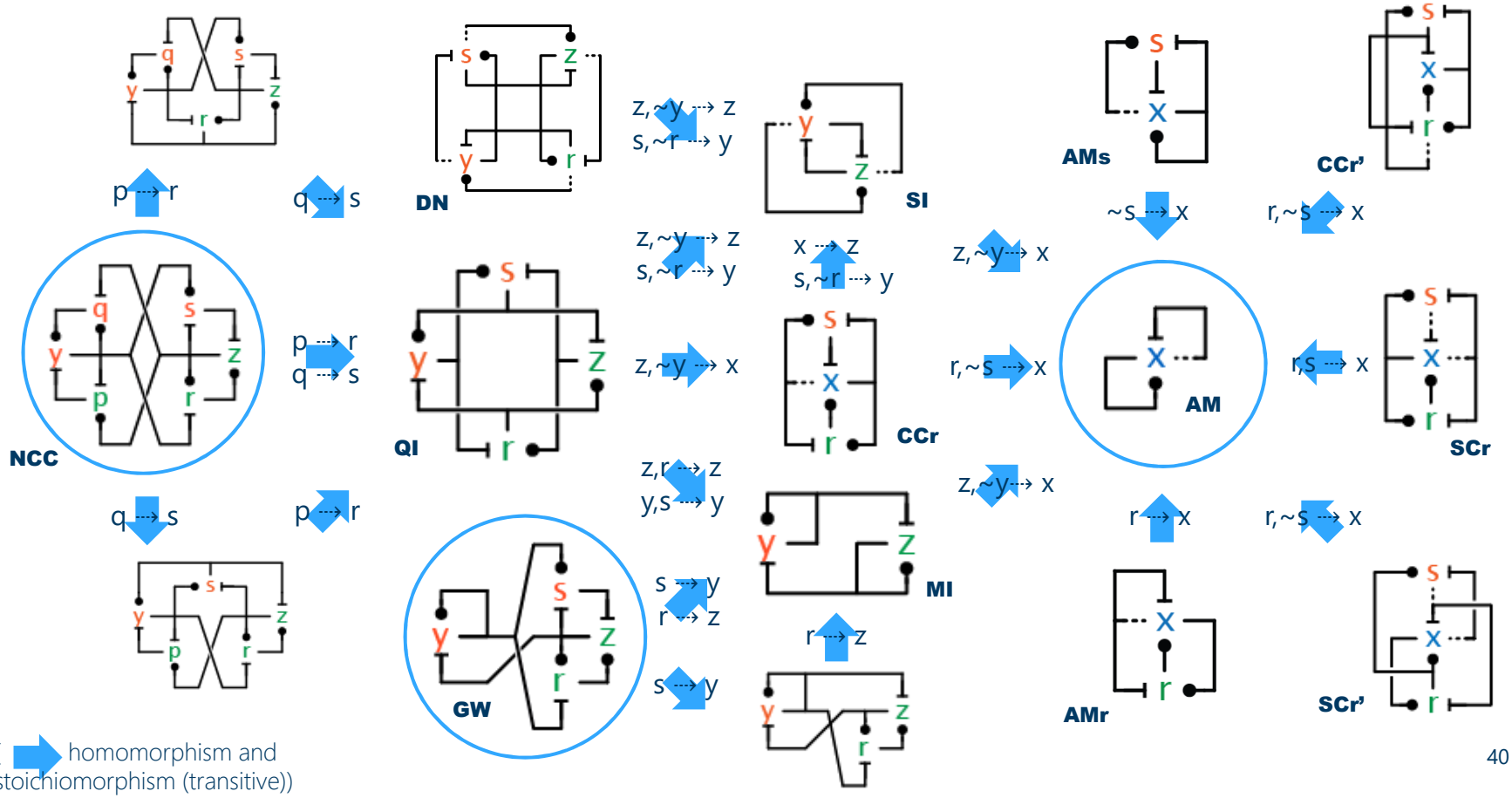
Concur 2015

Benchmarks from
Sneddon et al., Nature Methods, 2011

Model	Reactions	Species	FB	Time (s)	BB	Time (s)
e9	3538944	262146	222	4.61E+4	222	7.65E+4
e8	786432	65538	167	1.92E+3	167	3.68E+3
e7	172032	16386	122	8.15E+1	122	1.77E+2
e6	36864	4098	86	3.00E+0	86	7.29E+0
e5	7680	1026	58	1.54E-1	58	4.06E-1
e4	1536	258	37	9.00E-3	37	1.09E-1
e3	288	66	22	1.00E-3	22	3.00E-3
e2	48	18	12	1.00E-3	12	2.00E-3

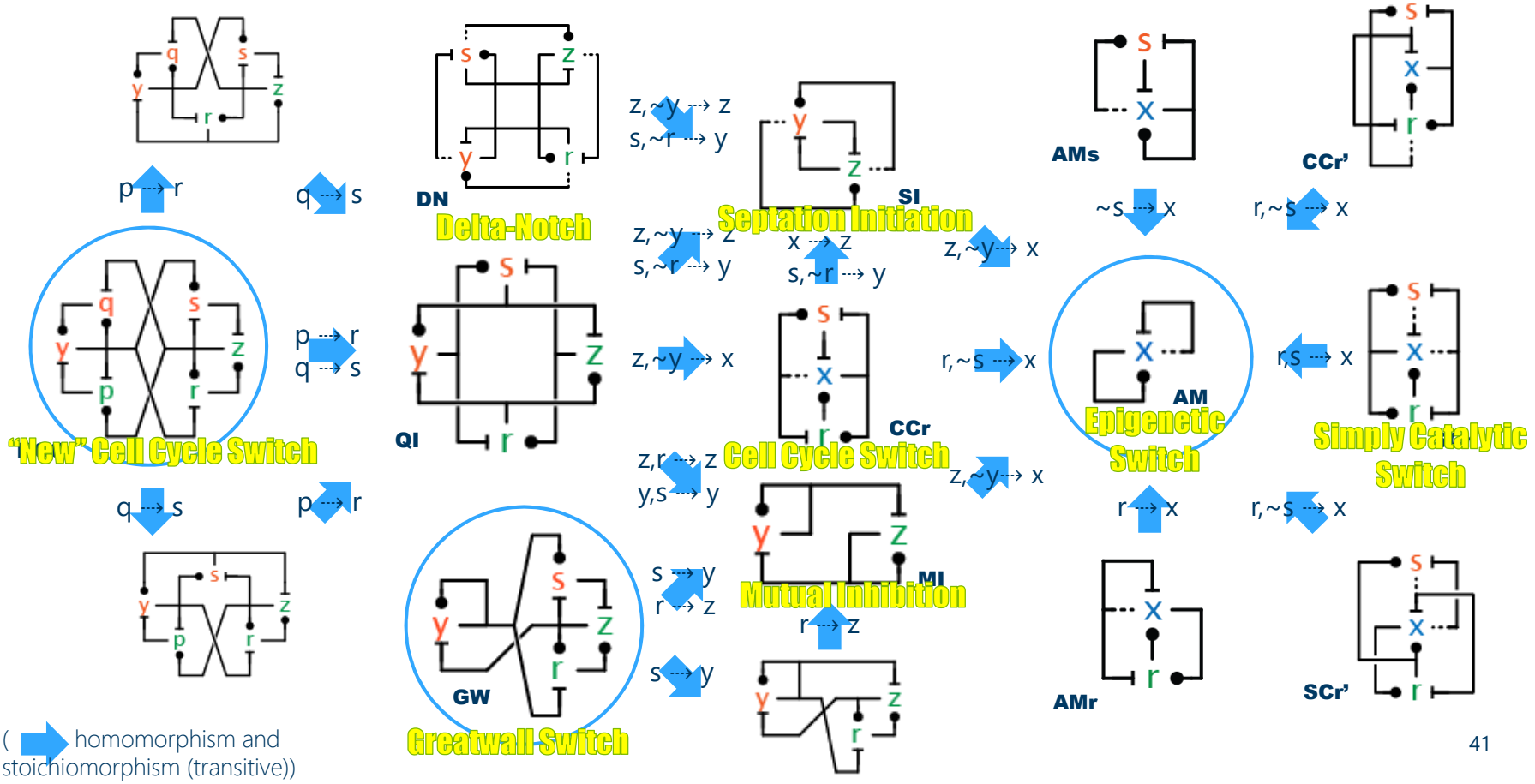
Aggregation reduction | Emulation reduction

Approximate Majority Emulation Zoo



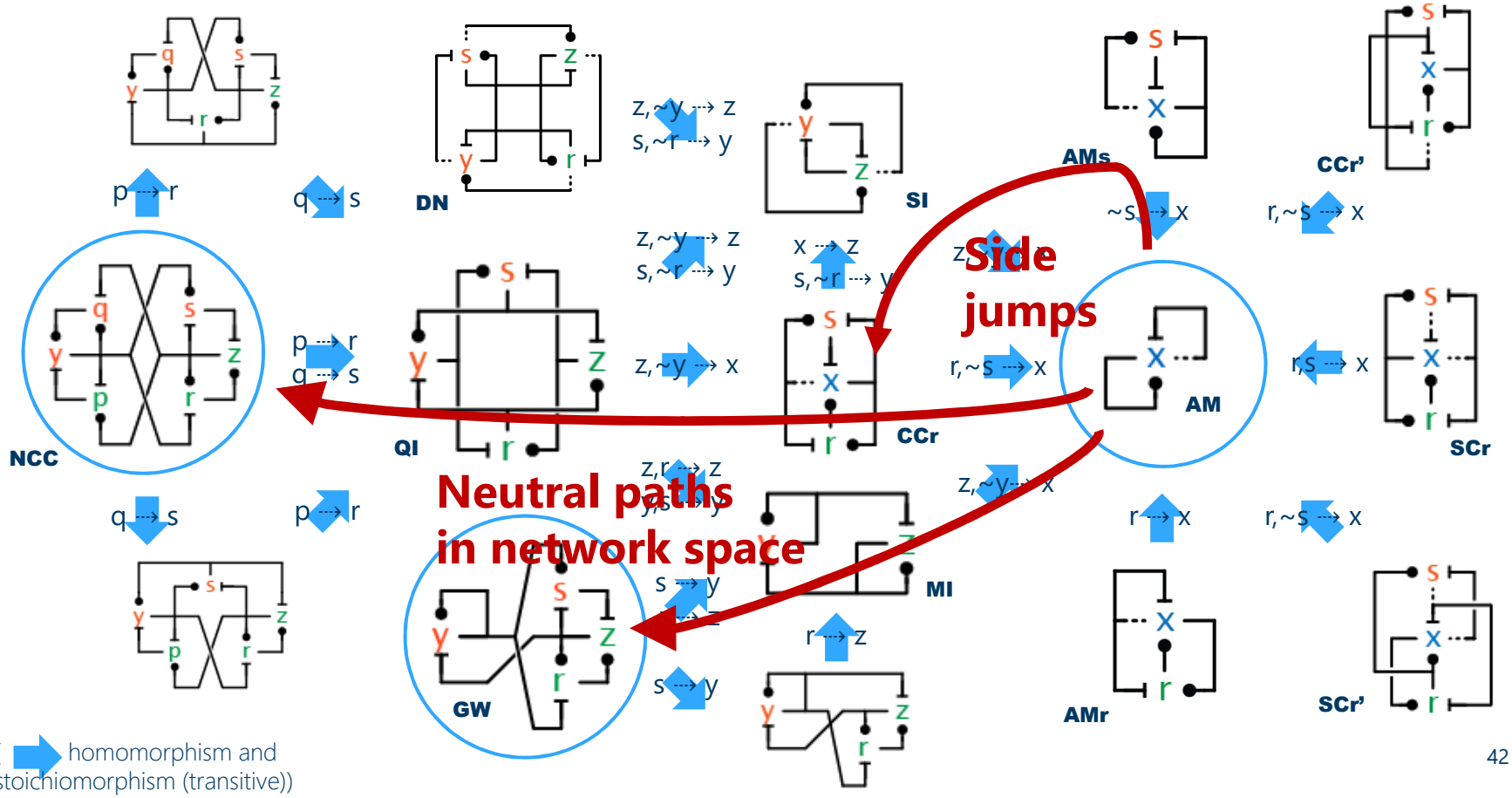
(homomorphism and stoichiomorphism (transitive))

Approximate Majority Emulation Zoo



(homomorphism and stoichiomorphism (transitive))

Approximate Majority Emulation Zoo

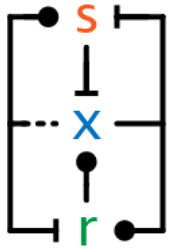


(homomorphism and stoichiomorphism (transitive))

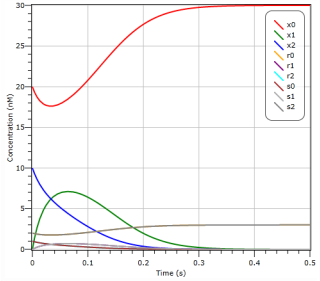
Stochasticity

The switch is noisy

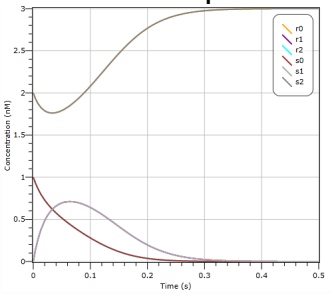
CCr with r,s at 1/10 of x,
r_0,s_0 rates 10* the rates of x_0,x_2



- Biological conditions:
  - x is abundant, r,s are "undetectable"
- This situation does not emulate AM
  - Because of the extra low-count r/s traces
  - BUT it emulates two separate copies of AM: one for x and one (low-count) copy for r/s
  - Hence it is still (deterministically) a good switch in the AM family
  - In particular, the low count species can be effective regulators even though they are present in "undetectable" quantities.
- But, we can expect significant noise
  - On r/s because they are in low-count
  - Likely on x because it is regulated by r/s

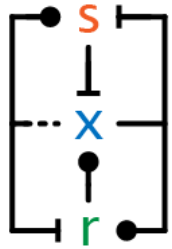


Showing just the low-count species:

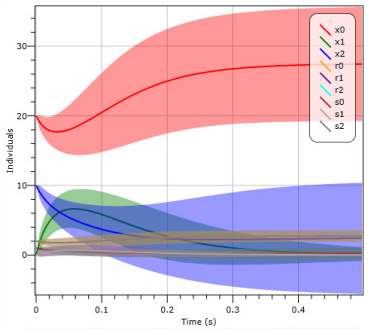


Stochastic behavior

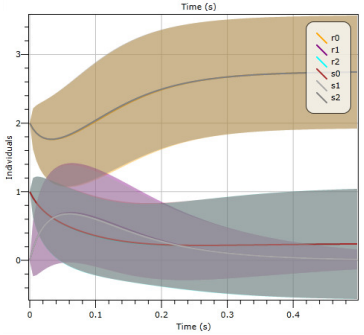
CCr with r,s at 1/10 of x,
r_0,s_0 rates 10* the rates of x_0,x_2



- We can in fact study the Chemical Master Equation

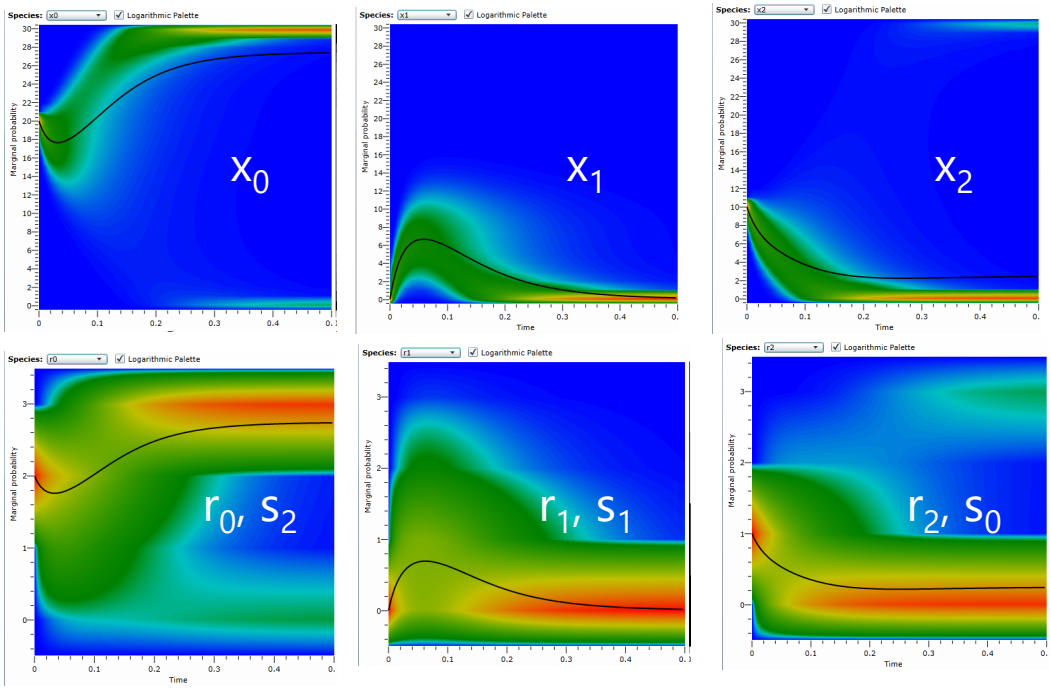


all species



zoom on r,s species

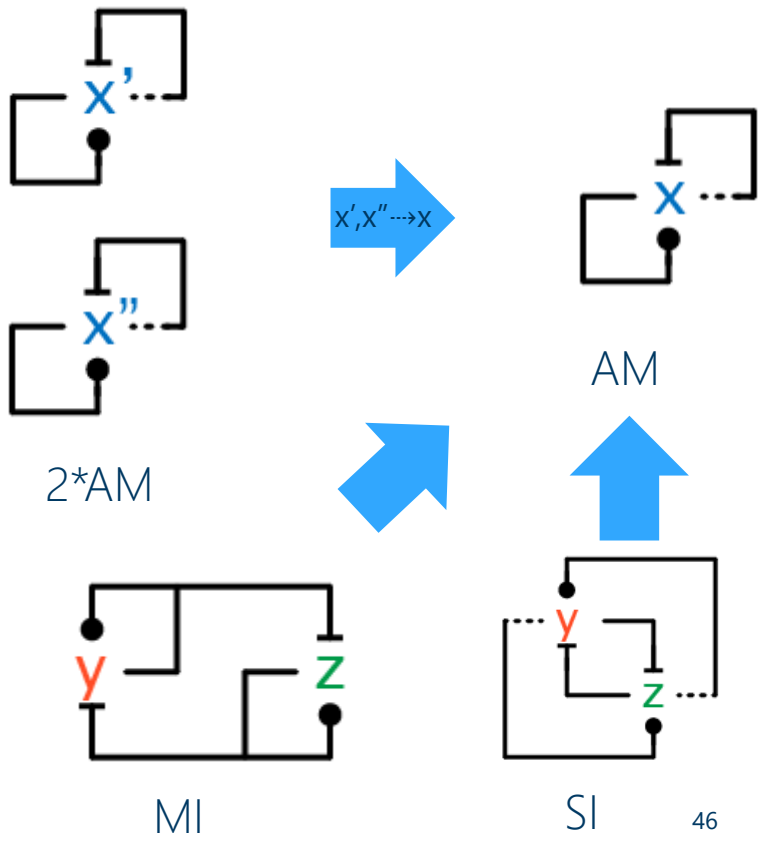
Mean and ±s.d. of species over time.



Mean and probability distribution of one species over time.

Trivial Example: AM vs. 2*AM

- Usually "more molecules" means "less noise"
- But not always
  - 2*AM emulates AM, hence the mean trajectories of 2*AM are the same as AM
  - The noise (s.d.) of 2*AM is also the same as AM
  - So, 2*AM has twice as many molecules, but noise is not reduced
- And not uniformly
  - MI,SI are two "intertwined" copies of AM
  - Are MI,SI less noisy than 2*AM?
  - Are MI,SI equally noisy? (They have the same number of molecules and reactions.)



Stochastic Switches

- Disentangle the contribution of complexity to stochasticity
  - Compare network noise on the baseline of deterministic emulation, across networks of different size and structure

Noise Reduction in Complex Biological Switches

Luca Cardelli[1,2,¶,*], Attila Csikász-Nagy[3,4,¶], Neil Dalchau[1,¶], Mirco Tribastone[5,¶], Max Tschaikowski[5,¶]

(To appear.)

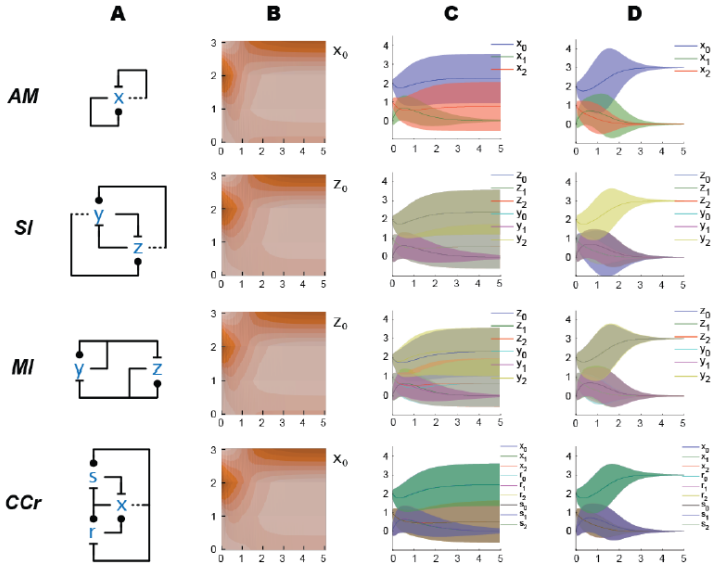


Fig 3 – Basic switching networks: stochastic solution. *Horizontal axes represent time, vertical axes represent number of molecules.* ***(A)*** *Influence networks.* ***(B)*** *Chemical Master Equation solution: probability distribution, with color (in 10 bands from light = 0 to dark = 1) indicating the probability that at time t there are y molecules of the single indicated species.* ***(C)*** *Chemical Master Equation solution: mean (solid lines) and standard deviation (color bands) for the species in the network.* ***(D)*** *Central Limit Approximation solution: mean (solid lines) and standard deviation (color bands) for the species in the network. Simulation scripts are in S5 Appendix.*

Stochastic Switches

- Network complexity *intrinsically* reduces noise

Noise Reduction in Complex Biological Switches

Luca Cardelli[1,2,¶,*], Attila Csikász-Nagy[3,4,¶], Neil Dalchau[1,¶], Mirco Tribastone[5,¶], Max Tschaikowski[5,¶]

(To appear.)

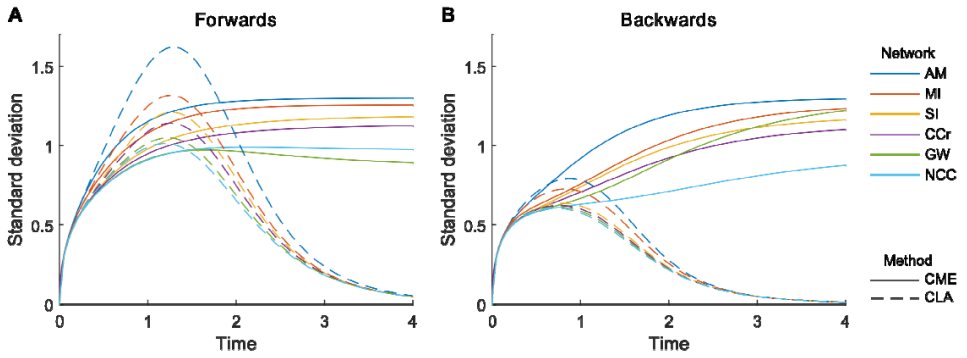


Fig 6 – Complexity improves overall performance of the cell cycle switch. *The performance of different networks was evaluated by calculating the standard deviation of the main molecular states (x_0 or z_0, depending on the network) over time. Standard deviations are calculated via numerical integration of the chemical master equation (CME) using the Visual GEC software, and via numerical integration of the central limit approximation (CLA) in Matlab. We investigate switching in one direction or the other by providing different initial conditions that settle (more likely) in different steady states.* ***(A)*** *In the forward direction, principal molecular states were initialised at 2 copies, and complementary molecular states were initialised at 1 copy, as shown in Fig 2C and Fig 5B.* ***(B)*** *In the reverse direction, principal molecular states were initialised at 1 copy, and complementary molecular states were initialised at 2 copies.*

"Stochastic Network Morphisms"

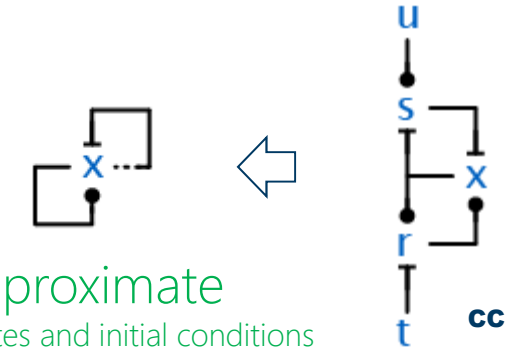
- ?

Conclusions

Networks are Algorithms

- They are *methods* for achieving a function
  - We need to understand how these methods relate to each other
  - In addition to how and how well they implement function
  - Algorithms can be obfuscated, and nature can obfuscate networks (to what end?)
- Network emulation can be checked *statically*
  - By stoichiometric/reaction-rate (*structural*) properties
  - That is, no need to compare ODE (*functional*) properties
  - For *any* initial conditions and rates of (one of) the networks
- We can efficiently discover emulations
  - Automatic model reduction of large networks

Nature likes good algorithms



Approximate

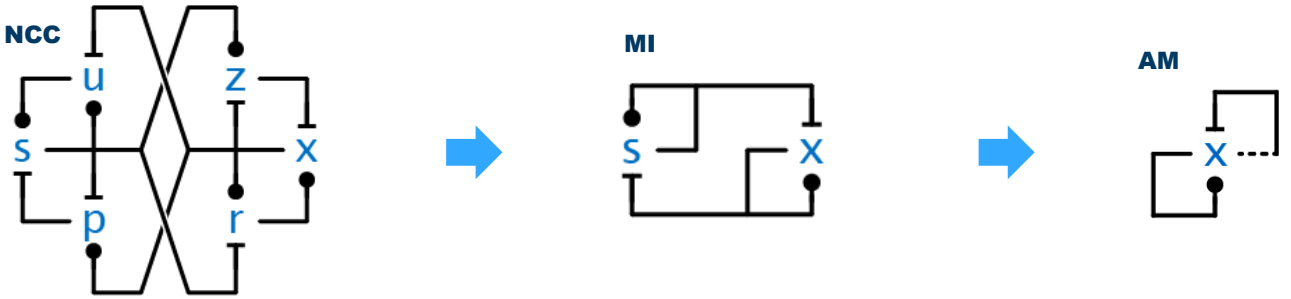
"default" rates and initial conditions

Exact

any rates and initial conditions

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

The cell cycle switch *can* exactly emulate AM



Feynman's Blackboard