Microsoft Research

UNIVERSITY OF OXFORD

The Cell Cycle Switch Computes Approximate Majority

Luca Cardelli, Microsoft Research & Oxford University

Mestre, 2015-04-08

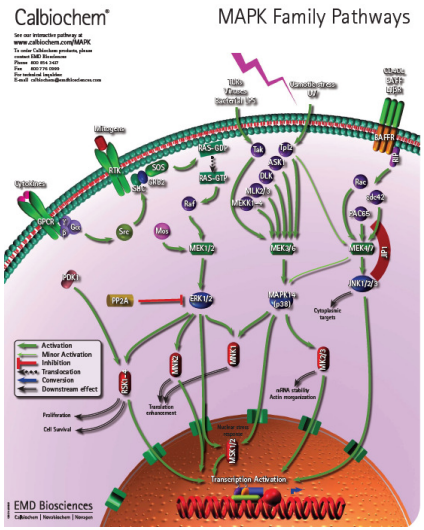
Outline

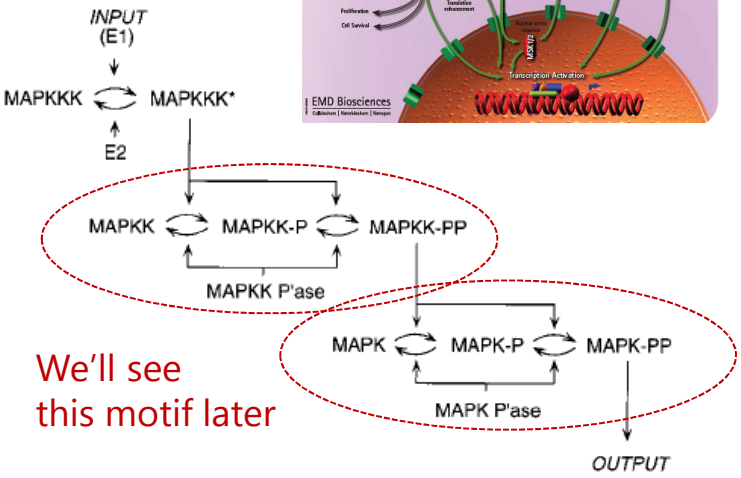
- Cellular Computation
  - Computational capabilities of biochemical mechanisms that may (or may not) be used by biological entities
- Chemical Algorithms
  - Specific instances of (bio-)chemical computation
  - Particularly, *consensus* and the cell cycle switch
- Obfuscation
  - How to hide a simple algorithm in a complex network
  - How to understand a complex network by a simple algorithm (de-obfuscation)

Cellular Computation

Cellular Computation

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

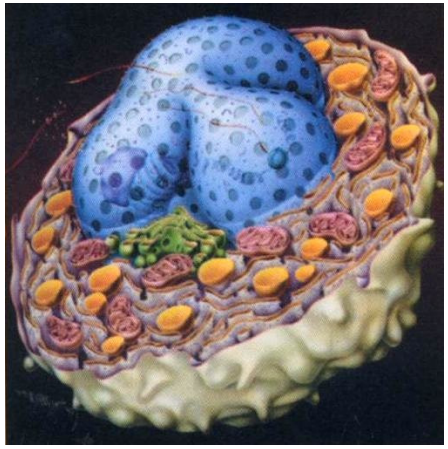


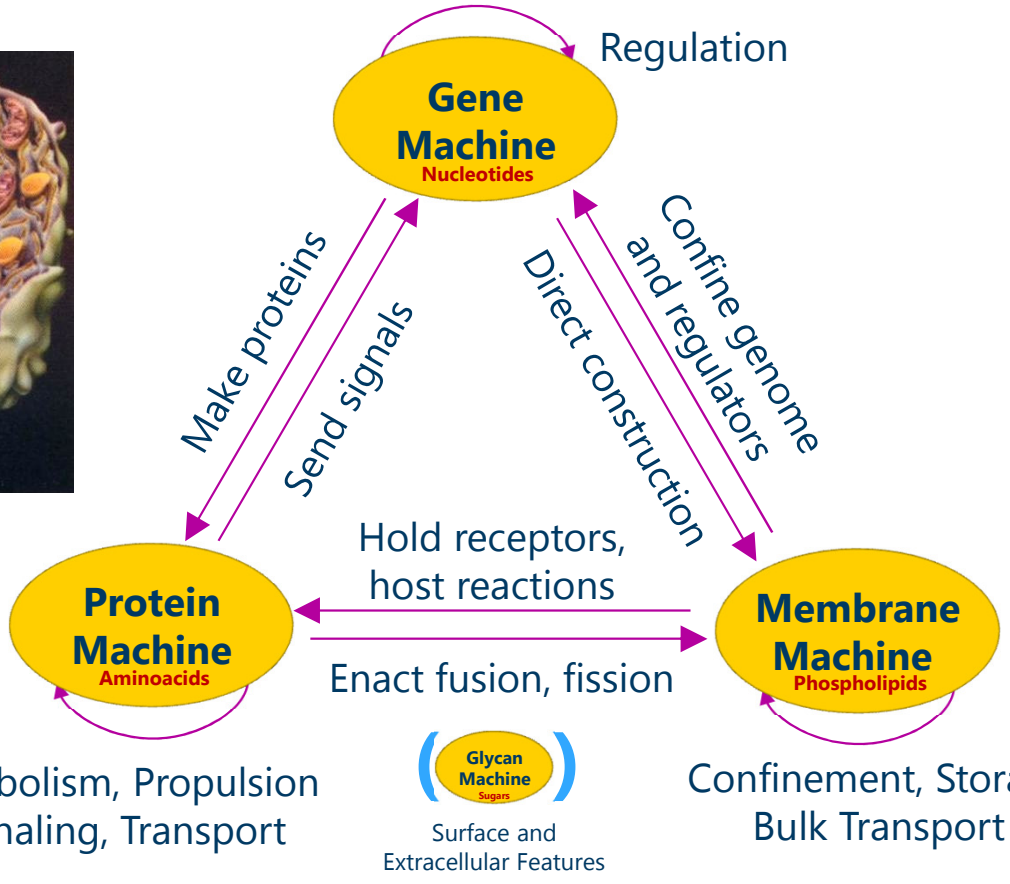


We'll see this motif later

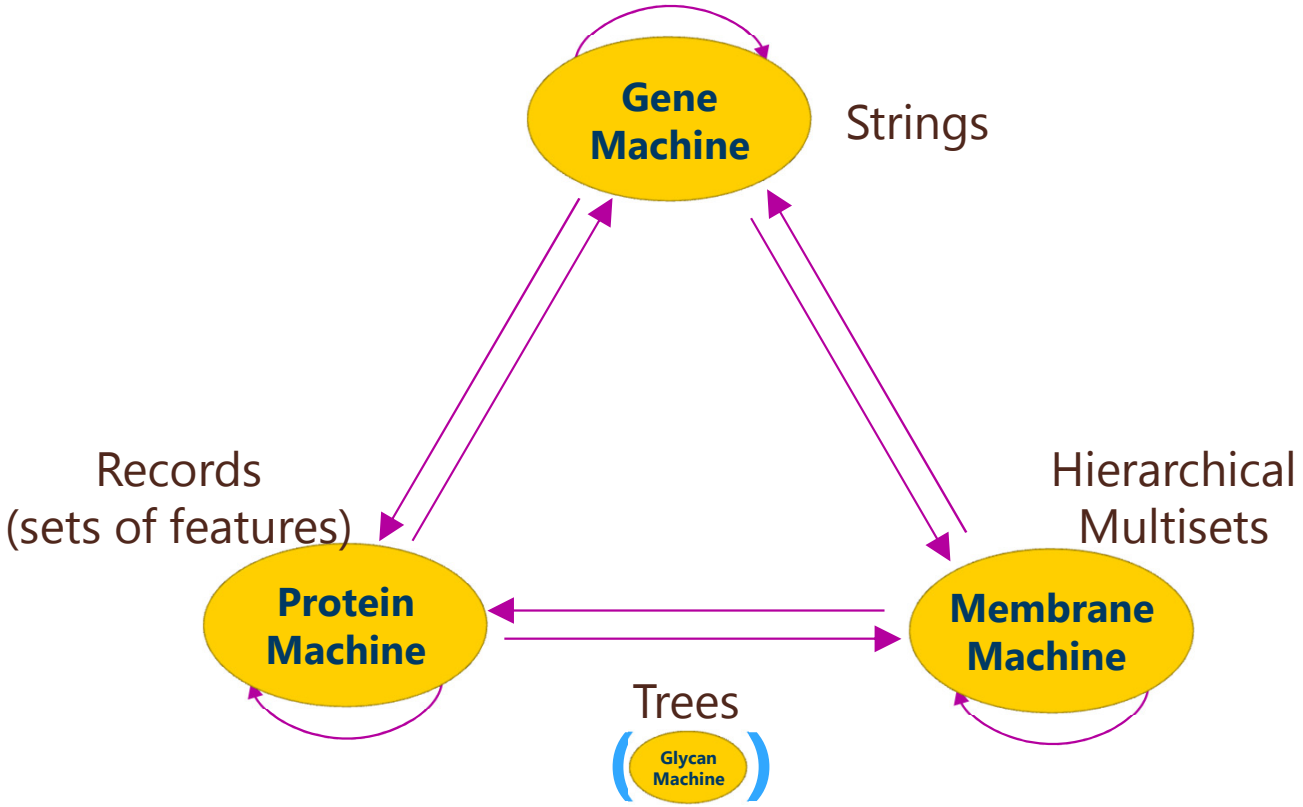
Ultrasensitivity in the mitogen-activated protein cascade, Chi-Ying F. Huang and James E. Ferrell, Jr., 1996, *Proc. Natl. Acad. Sci. USA*, 93, 10078-10083.

Abstract Machines of Biochemistry

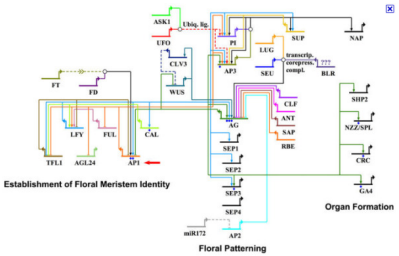




Bioinformatics View (Data Structures)



Systems Biology View (Networks)

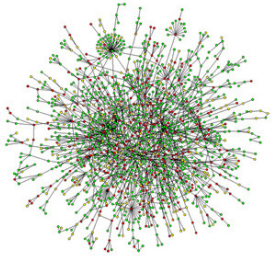


Gene Machine

Gene Regulatory Networks

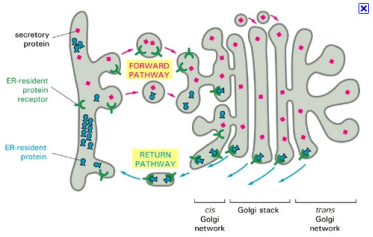
Biochemical Networks

Transport Networks

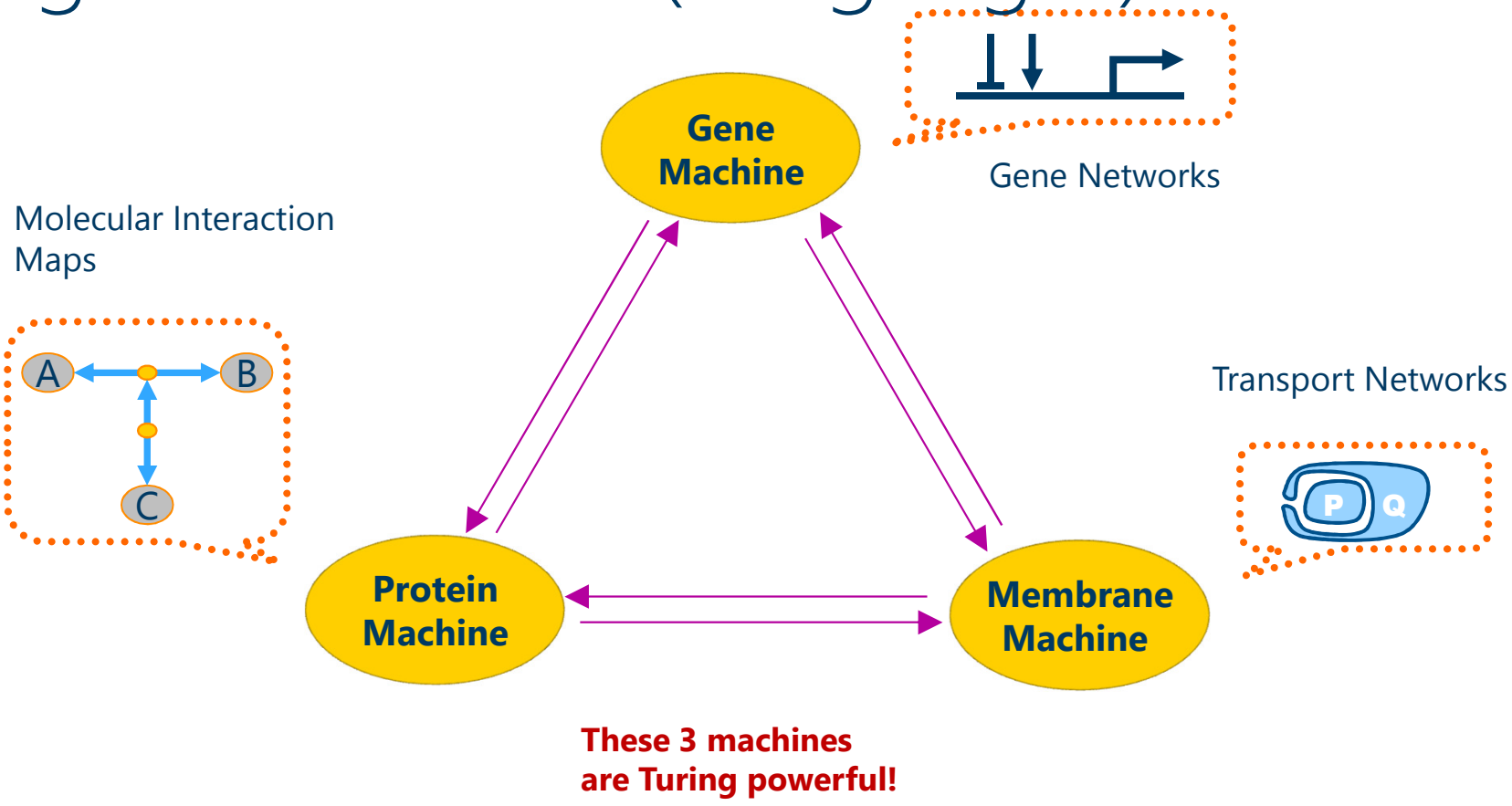


Protein Machine

Membrane Machine

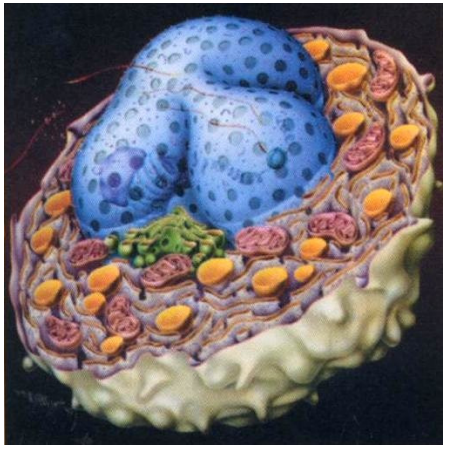


Algorithmic View (Languages)



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

spec	*program*	
Y := 2X	X -> Y + Y	(extra mass comes from "somewhere")
Y := ⌊X/2⌋	X + X -> Y	
Y := X1 + X2	X1 -> Y X2 -> Y	

Advanced Programming Examples

spec

Y := min(X1, X2)

Y := max(X1, X2)

program

X1 + X2 -> Y

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

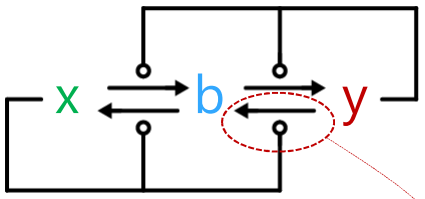
max(X1,X2)=
(X1+X2)-min(X1,X2)

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

A Consensus Algorithm

- A Population Consensus Problem
  - Given two populations of x and y "agents" (entities/molecules)
  - we want them to "reach consensus"
  - by converting *all* agents to x or to y
    depending on which population was in majority initially
- Approximate Majority (AM) Algorithm
  - Uses a third "undecided" population b
  - Disagreements cause agents to become undecided
  - Undecided agents agree with any non-undecided agent
- Population Protocols Model
  - Finite-state identity-free agents (molecules) interact in randomly chosen pa
  - Each interaction (collision) can result in state changes
  - Complete connectivity, no centralized control (well-mixed solution)

catalysis



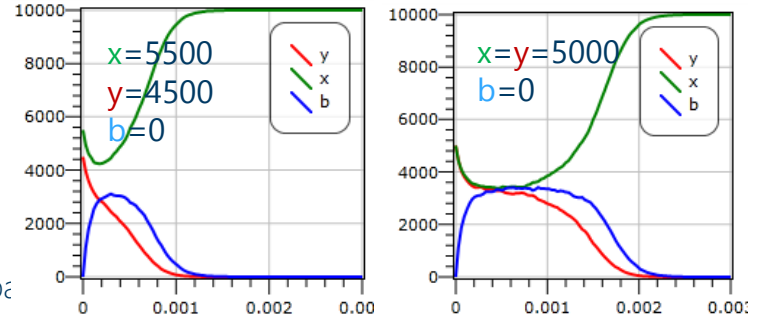
chemical reaction network

$x + y \xrightarrow{r} y + b$

$y + x \xrightarrow{r} x + b$

$b + x \xrightarrow{r} x + x$

$b + y \xrightarrow{r} y + y$

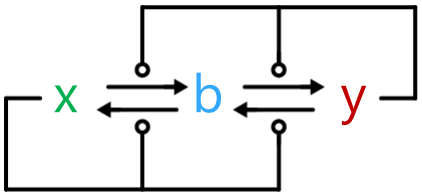


Dana Angluin · James Aspnes · David Eisenstat

A Simple Population Protocol for Fast Robust Approximate Majority

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

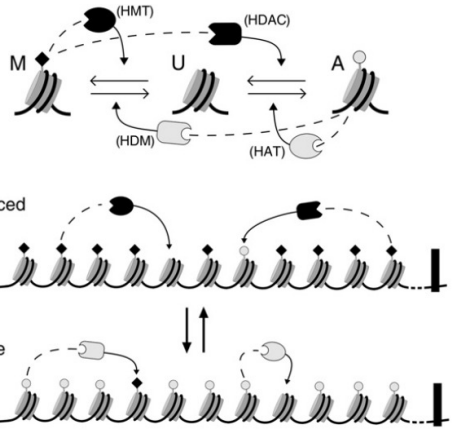
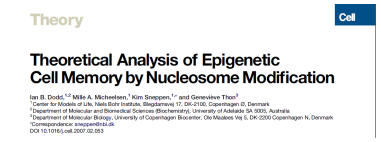


Figure 1. Basic Ingredients of the Model


Theory

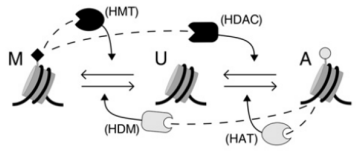
Cell

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



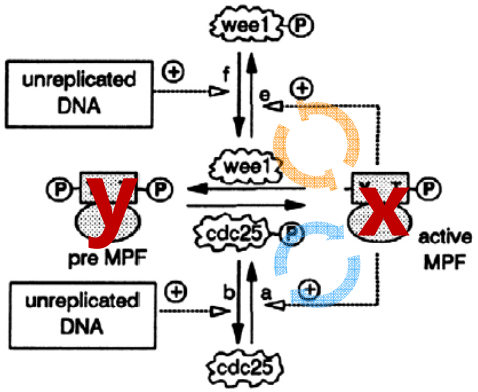
- In most cases the biological implementation seems more *indirect* or *obfuscated*
  - *"Nature is subtle but not malicious - Einstein"* Ha! think again!
  - Other implementations of Approximate Majority seem more convoluted and… approximate

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:



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*?

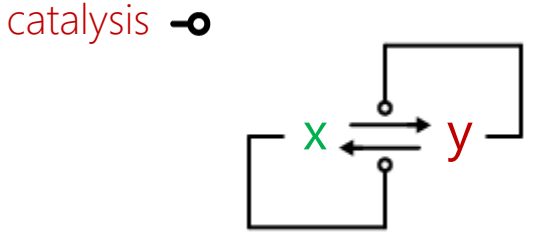
How to Build a Good Switch

- 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

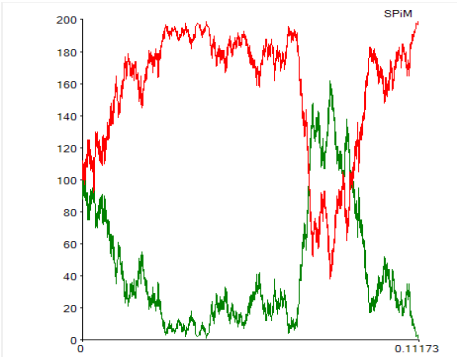
A Bad Algorithm

- 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*)



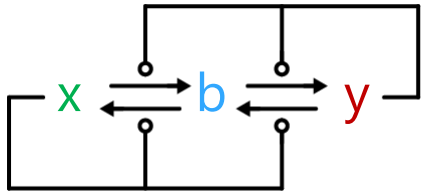
$$y + x \rightarrow x + x$$
$$x + y \rightarrow y + y$$



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 $O(\log n)$

catalysis

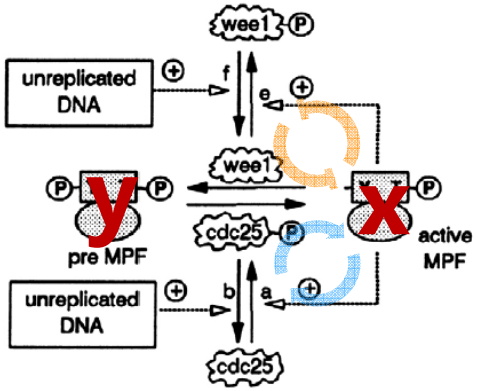


$$x + y \rightarrow y + b$$
$$y + x \rightarrow x + b$$
$$b + x \rightarrow x + x$$
$$b + y \rightarrow y + y$$

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

Variation on a distributed algorithm?

Need to explain this network notation!

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

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

How to model "Influence"

"True" molecular interactions.

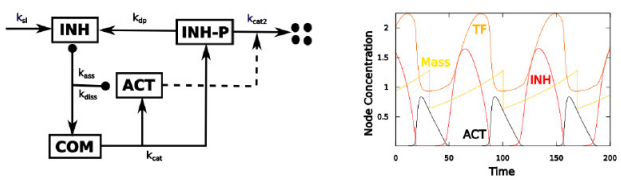


Figure 3: a) Schematic diagram of a simplified SIMM model [17]. The activa-

"Equivalent" influence interactions.

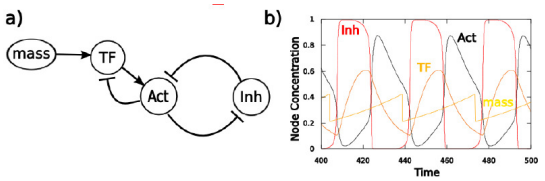


Figure 4: a) Schematic diagram of a primitive cell cycle in the reinitz framework.

Chemical Reaction Network ⟷ Influence Network

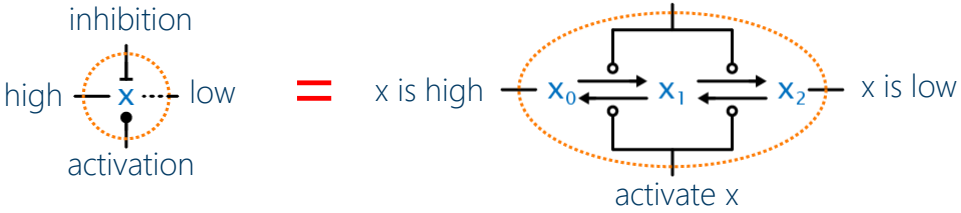
Evolving a Primitive Eukaryotic Cell Cycle Model

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

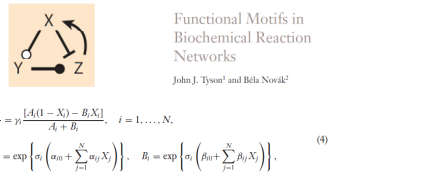


Usually modeled by sigmoid (e.g. Hill or Reinitz) functions

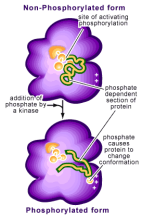
triplet motif

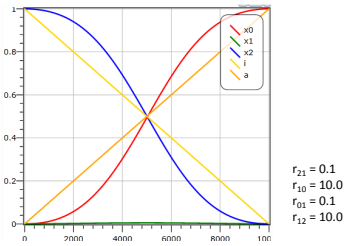
We model them by 4 mass action reactions over 3 species x_0, x_1, x_2

They actually implement a Hill function of coefficient 2:



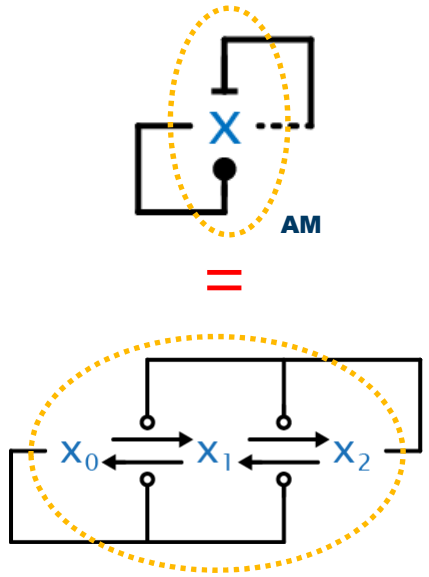
biological mechanism: (e.g.:) multisite phosphorylation





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

For example:

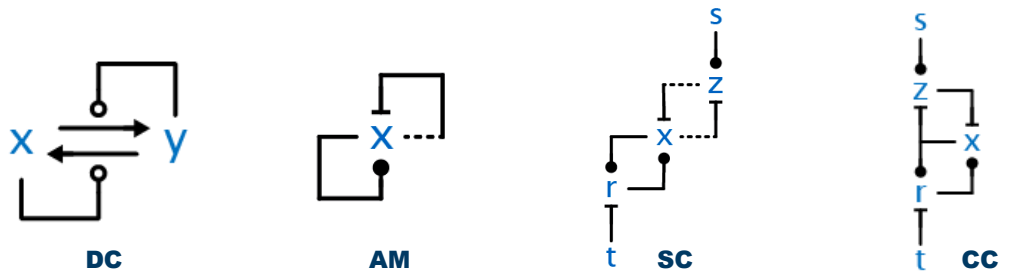


Approximate Majority

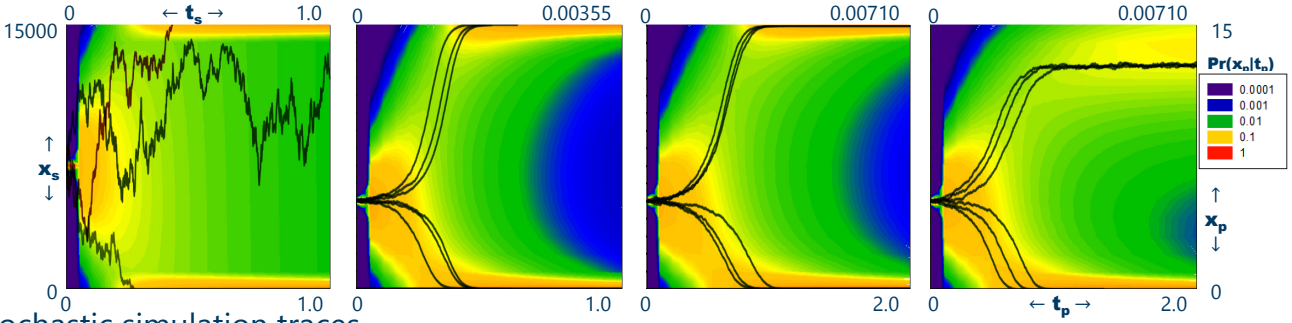
Convergence Analysis

- Switches as computational systems

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



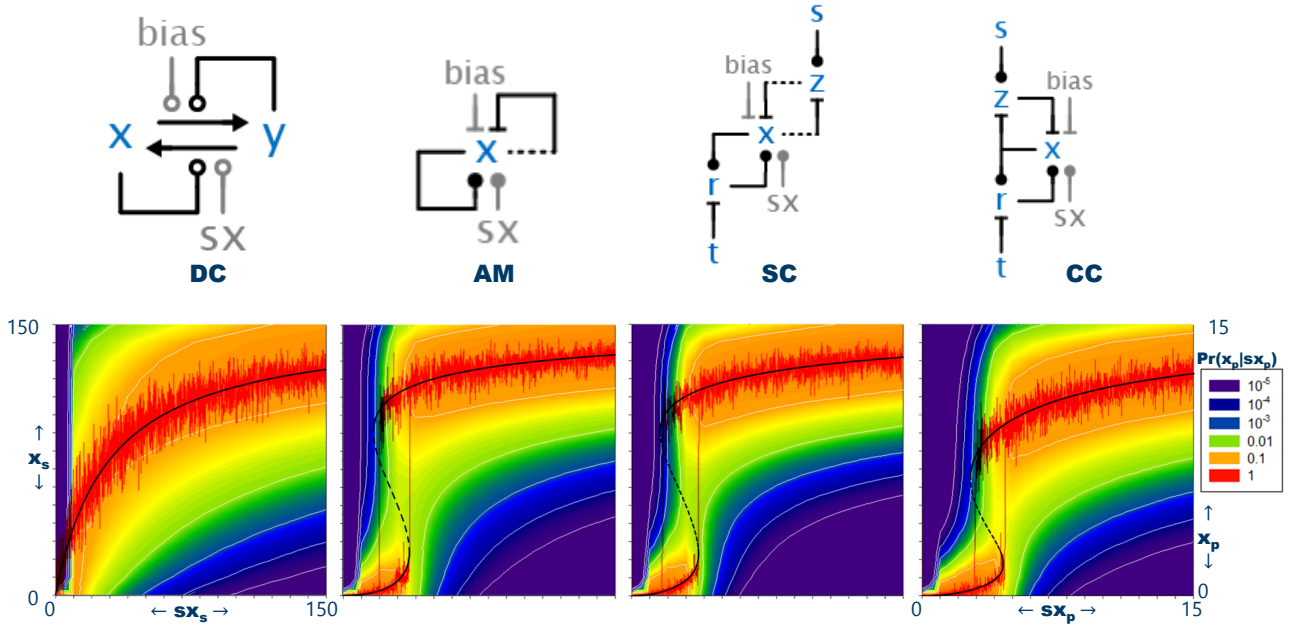
Start symmetrical
(x_0=x_1=x_2 etc.)



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

Steady State Analysis

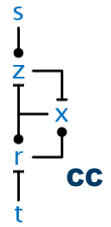
- Switches as dynamical systems

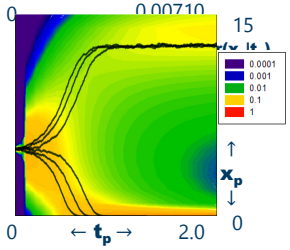


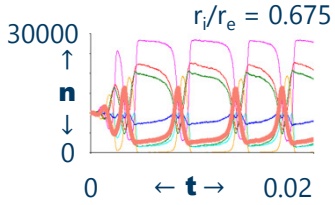
Black lines: deterministic ODE bifurcation diagrams
Red lines: noisy stochastic simulations
Color: full probability distribution of small-size system

Why is CC worse than AM?

- The classical CC has an algorithmic "bug"
  - It works ok but never as well as AM
  - Because s continuously inhibits x through z, so that x cannot fully express





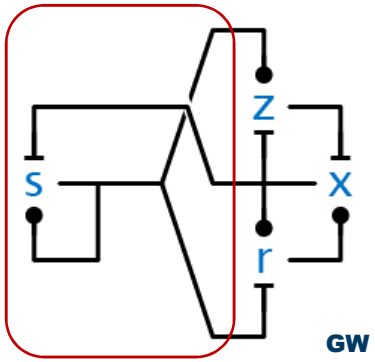


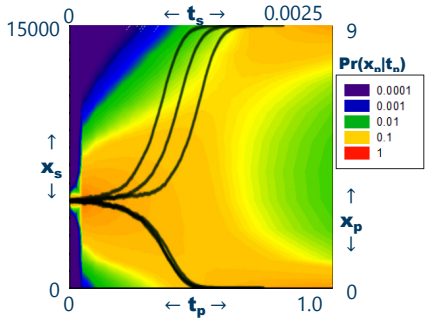
The corresponding cell cycle oscillator is also depressed

- So let's fix the bug!
  - Easy: let x inhibit s and t "in retaliation"
  - Q: Why didn't nature fix it?

Nature fixed it!

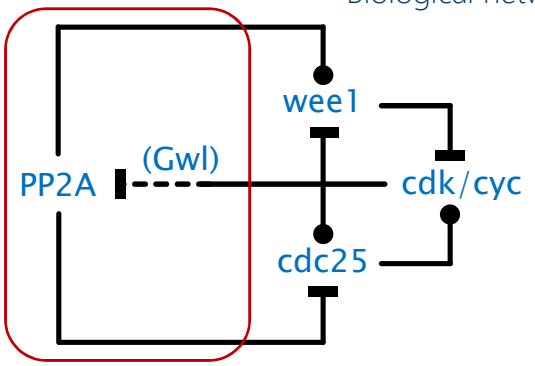
- There is another known feedback loop
  - By which x suppresses s "in retaliation" via the so-called Greatwall loop
  - Also, s and t happen to be the same molecule (=s)





Full activation!

Biological network



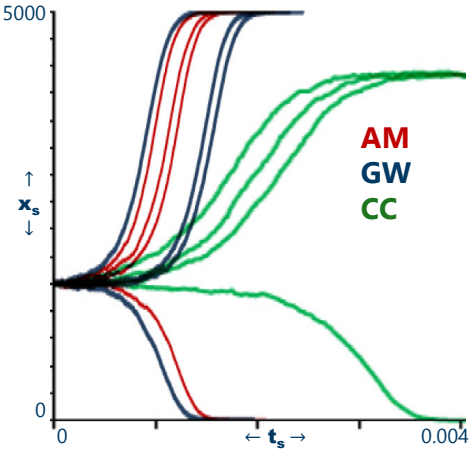
- s and x now are antagonists: they are the two halves of the switch, mutually inhibiting each other (through intermediaries).

More surprisingly

- The fix makes it faster too!
  - The extra feedback also speeds up the decision time of the switch, making it about as good as the 'optimal' AM switch:

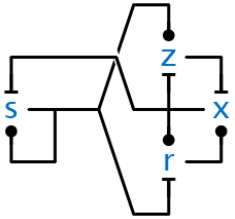
Conclusion:
Nature is trying as hard as it can to implement an AM-class algorithm!

The "classical" cell cycle switch is only half of the picture: the extra feedback completes it *algorithmically*.



Publications

- Our paper appeared:
  - Suggesting GW is a better switch than CC. *September 2012*



OPEN

The Cell Cycle Switch Computes Approximate Majority

SUBJECT AREAS:
COMPUTATIONAL BIOLOGY

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

- 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'

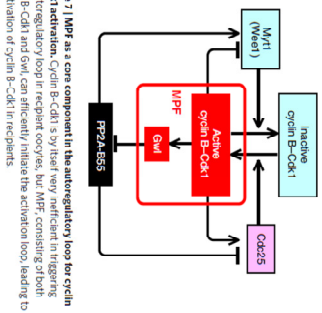


Figure 7 | MPF as a core component in the autoregulatory loop for cyclin B–Cdk1 activation. Cyclin B–Cdk1 's synthesis may not be efficient in triggering the autoregulatory loop in recipient oocytes, but MPF, consisting of both cyclin B–Cdk1 and Gwl, can efficiently initiate the activation loop, leading to full activation of cyclin B–Cdk1 in recipients.

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]

What we learned

- The network structure of AM implements an input-driven switching function (in addition to the known majority function).
- The network structure of CC/GW implements a input-less majority function (in addition to the known switching function).

- The behavior of AM and CC/GW in isolation are related.
- The behavior of AM and CC/GW in oscillator contexts are related (not shown).

- A refinement (GW) of the core CC network, known to occur in nature, improves its switching performance and brings it in line with AM performance.

But again, is CC (or GW) the "same" as AM?

- Our evidence for computational content of biochemical networks is so far
  - Quantitative, covering both kinetic and steady state behavior of *what* networks do
  - But empirical (based on simulations/numerical solutions)
  - And it does not yet explain *how* the CC/GW network relates to the AM network, that is, how each *piece* of CC/GW corresponds to each *piece* of AM
- Analytical evidence is harder to obtain
  - The proofs of the computational properties (optimality etc.) for the AM algorithm are hard and do not generalize easily to more complex networks
  - Quantitative theories of behavioral equivalence and behavioral approximation, e.g. in process algebra, are still lacking (although rich qualitative theories exist)

Obfuscation

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

Antagonistic Networks

- Let's generalize:
  - AM is based on antagonism between two species (inside the triplet)
  - So (essentially) is GW
  - So (essentially) are many standard biological networks

- Are they somehow related?
  - We could try the same empirical analysis as for CC/AM
  - But we can do better

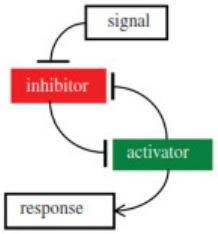
Mutual Inhibition (1 vs. 1)

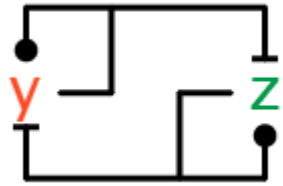
- "All" cellular switches in all phases of the cell cycle follow (abstractly) a mutual inhibition pattern:

Molecular mechanisms creating bistable switches at cell cycle transitions

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

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





MI

- Also found in other areas (cell polarity establishment):

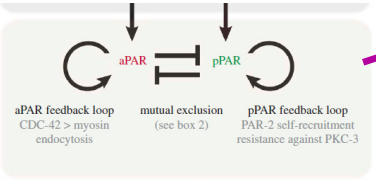
PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

rstb.royalsocietypublishing.org

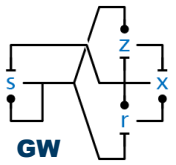
The PAR network: redundancy and robustness in a symmetry-breaking system

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

[1]Temasek Lifesciences Laboratory, [2]Mechanobiology Institute, and [3]Department of Biological Sciences, National University of Singapore, 1 Research Link, Singapore 117604, Republic of Singapore
[4]Department of Molecular Biology and Genetics and HHMI, Johns Hopkins University School of Medicine, Baltimore MD 21205, USA



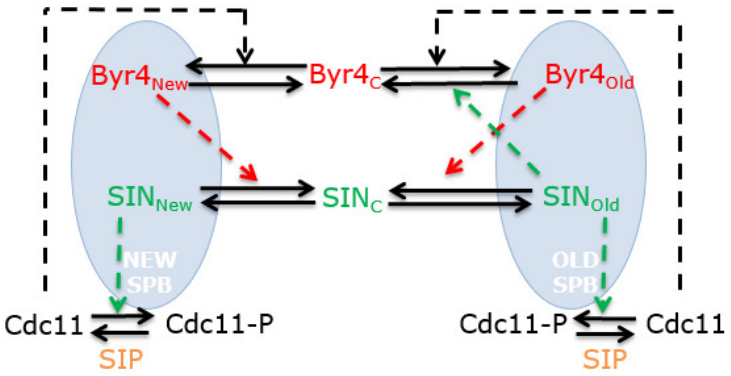
cf.:



GW

Septation Initiation (1 vs. 1)

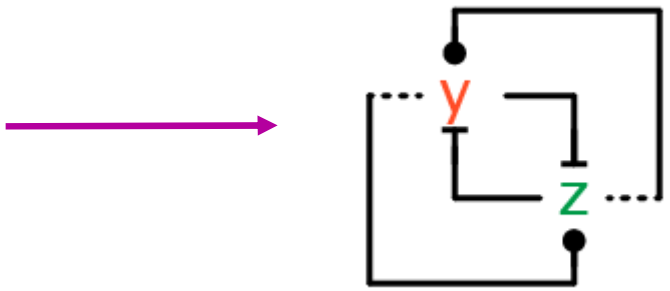
- Other (inherently different) biological networks are based on mutual inhibition, and share characteristics with AM



OPEN ACCESS Freely available online
PLOS | COMPUTATIONAL BIOLOGY

Dynamics of SIN Asymmetry Establishment

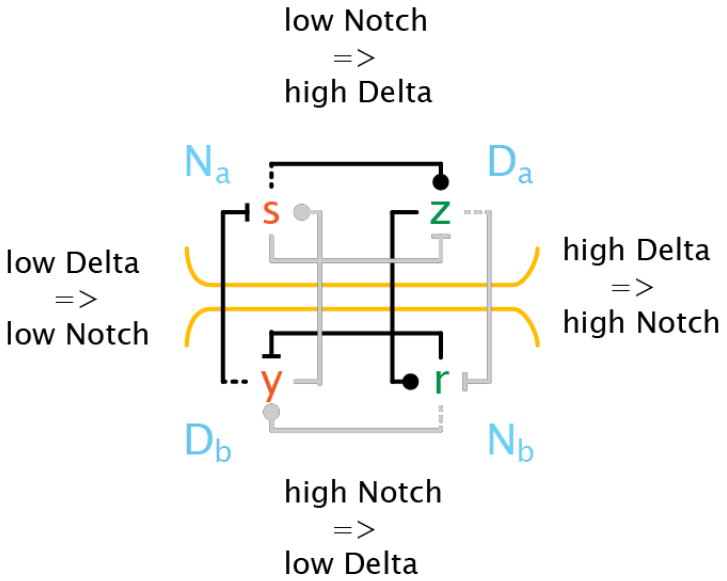
Archana Bajpai[1], Anna Feoktistova[2], Jun-Song Chen[2], Dannel McCollum[3], Masamitsu Sato[4,5], Rafael E. Carazo-Salas[6], Kathleen L. Gould[2], Attila Csikász-Nagy[1,7,8*]



SIN inhibiting Byr4,
absence of SIN promoting Byr4
Byr4 inhibiting SIN,
absence of Byr4 promoting SIN

Delta-Notch (2 vs. 2)

- A mutual inhibition pattern
  - Involving *two* species in each cell
- In two cells a,b
  - D_a,N_b antagonize D_b,N_a



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

New Cell Cycle Network (3 vs. 3)

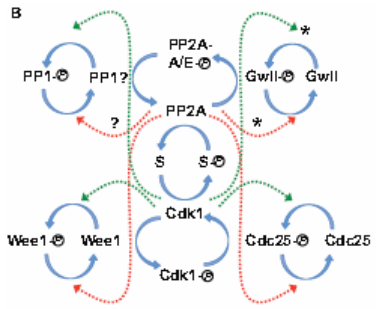
- A recent paper presents a more complete view of the cell cycle switch
- N.B. "phosphorylation network dynamics" here is the same as our x_0-x_1-x_2 motif

Phosphorylation network dynamics in the control of cell cycle transitions

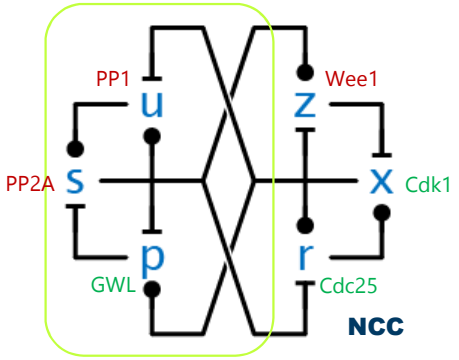
Daniel Fisher[1,*], Liliana Krasinska[1,‡], Damien Coudreuse[2,‡] and Béla Novák[3,‡]
[1]Institut de Génétique Moléculaire de Montpellier, IGMM, CNRS UMR 5535, Université Montpellier I and II, 34293 Montpellier, France
[2]Institute of Genetics and Development of Rennes, CNRS UMR 6290, 35043 Rennes, France
[3]Oxford Centre for Integrative Systems Biology, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK
*Author for correspondence (daniel.fisher@igmm.cnrs.fr)
‡These authors contributed equally to this work

Journal of Cell Science 125, 4703–4711

doi: 10.1242/jcs.106351

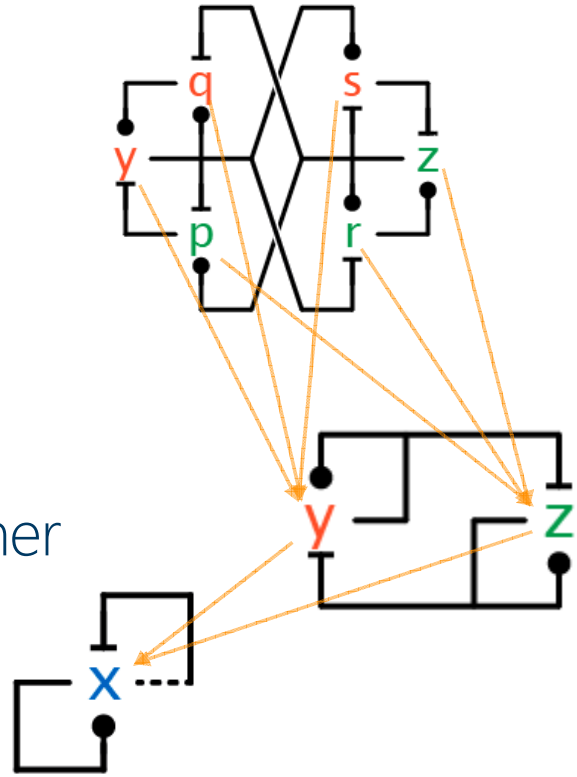


Mutual inhibition between *three* species each



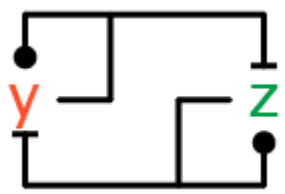
Comparing networks

- 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



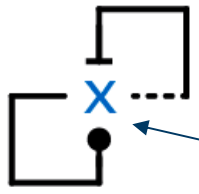
Network Emulation: MI emulates AM

- 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**:



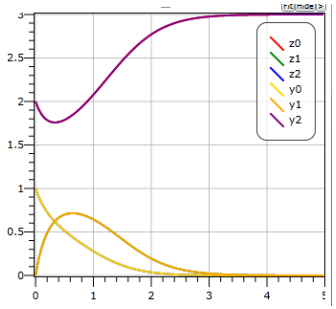
MI

~y,z⇢ x

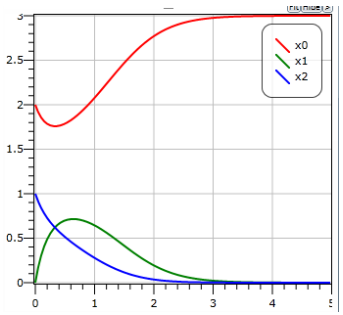


AM

(3 species)



(6 species on 3 trajectories)



(3 species on 3 trajectories)

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

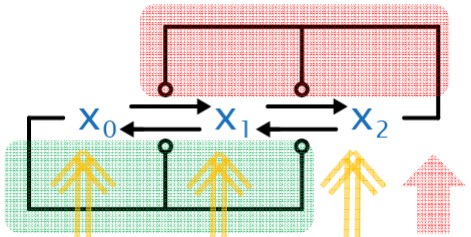
($y_2 = x_0$
$y_1 = x_1$
$y_0 = x_0$)

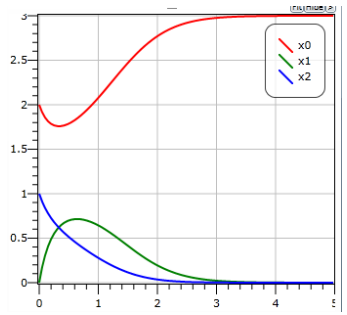
- How do we find these matching parameters? By a network morphism!

Network Emulation: MI emulates AM

A mapping of species and reactions

AM



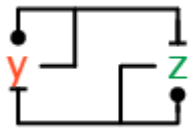


any initial conditions

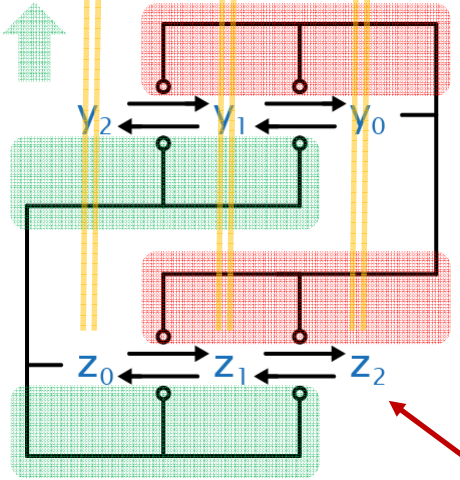
homomorphic mapping

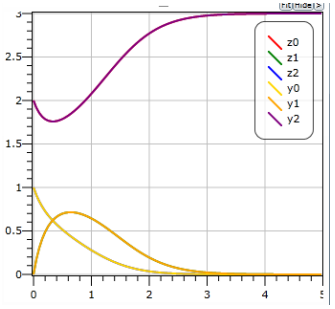
z -> x

~y -> x



MI





initial conditions:

$z_0 = y_2 = x_0$

$z_1 = y_1 = x_1$

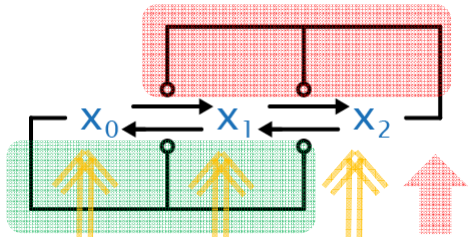
$z_2 = y_0 = x_2$

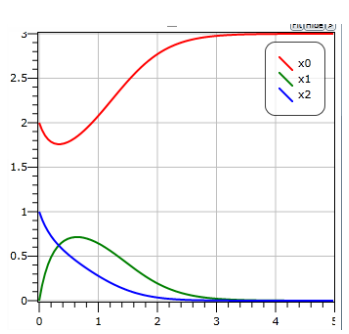
less trivial than you might think:
it need not preserve the out-degree of a node!

Network Emulation: SI emulates AM

A mapping of species and reactions

AM

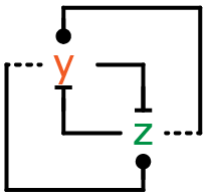




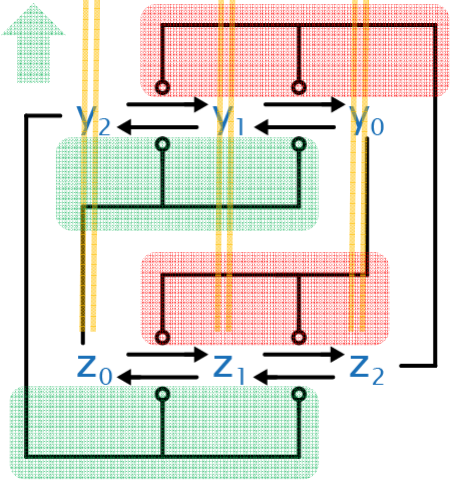
any initial conditions

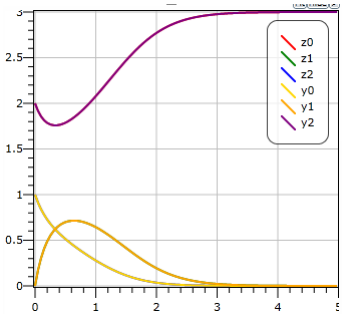
homomorphic mapping

z -> x
~y -> x



SI



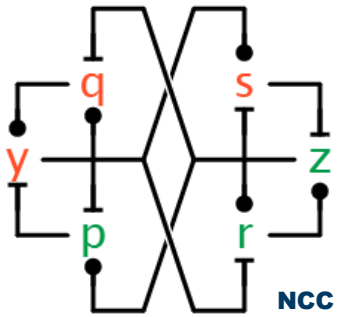


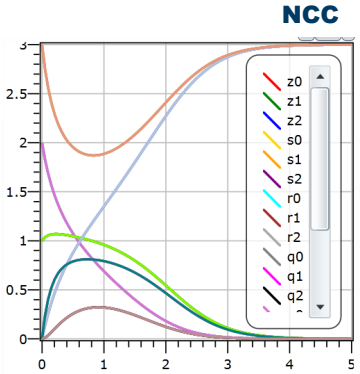
initial conditions:

$z_0 = y_2 = x_0$
$z_1 = y_1 = x_1$
$z_2 = y_0 = x_2$

Network Emulation: NCC emulates MI

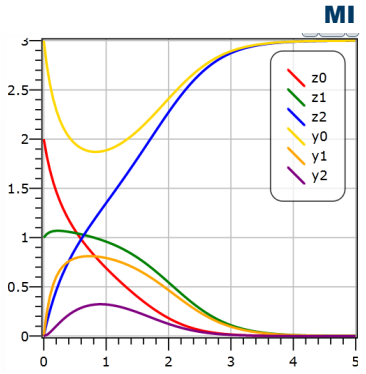
- For *any* rates and initial conditions of MI we can find *some* rates and initial conditions of NCC such that the (18) trajectories of NCC retrace those (6) of MI





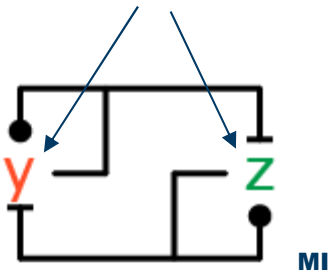
(18 species on 6 trajectories)

z,r,p ⇢ z
y,q,s ⇢ y



(6 species on 6 trajectories)

(3 species each)

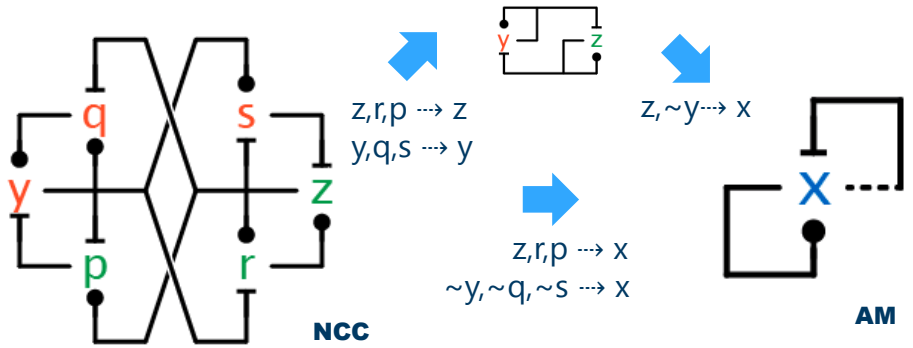


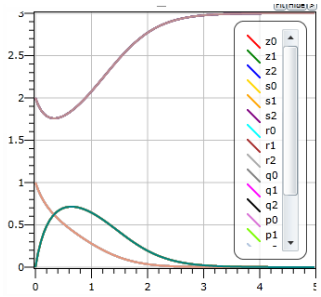
initialize
z,r,p = z
y,q,s = y

- Why does this work so well?

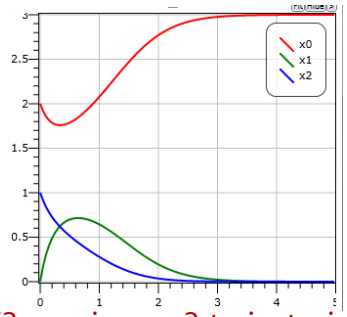
Emulations Compose: NCC emulates AM

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





(18 species on 3 trajectories)



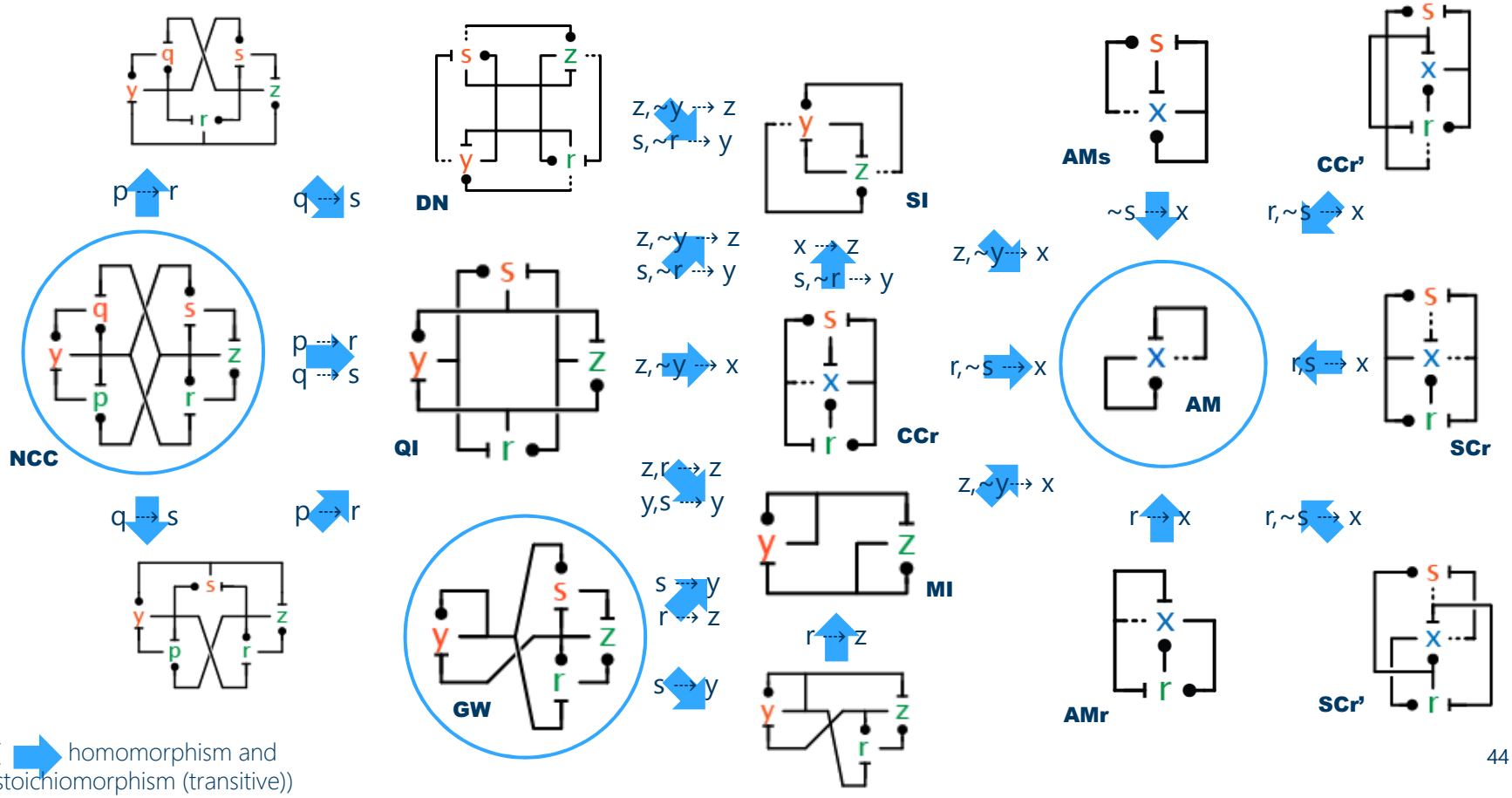
(3 species on 3 trajectories)

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

For *any* initial conditions of AM.

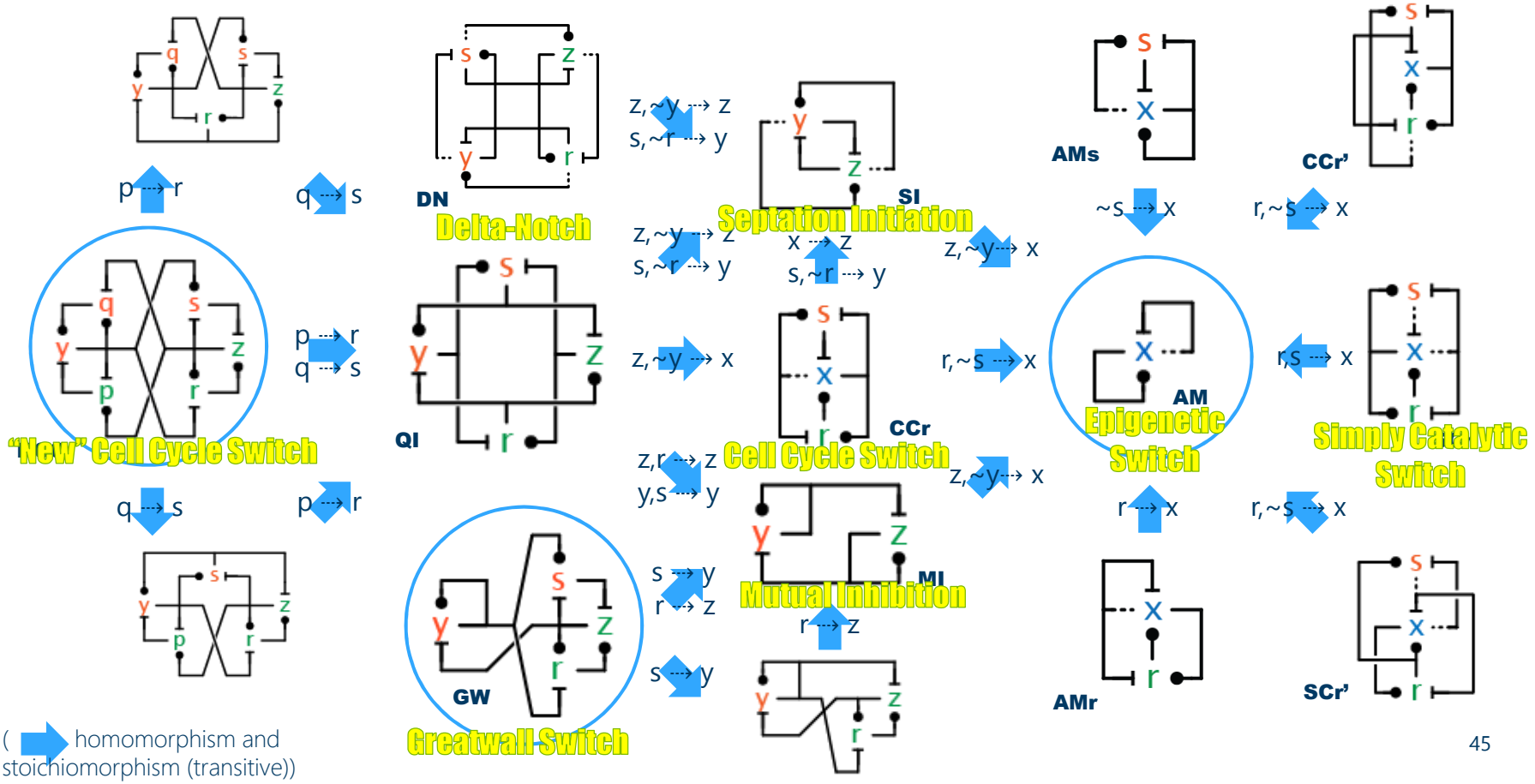
And for *any* rates of AM.

Approximate Majority Emulation Zoo



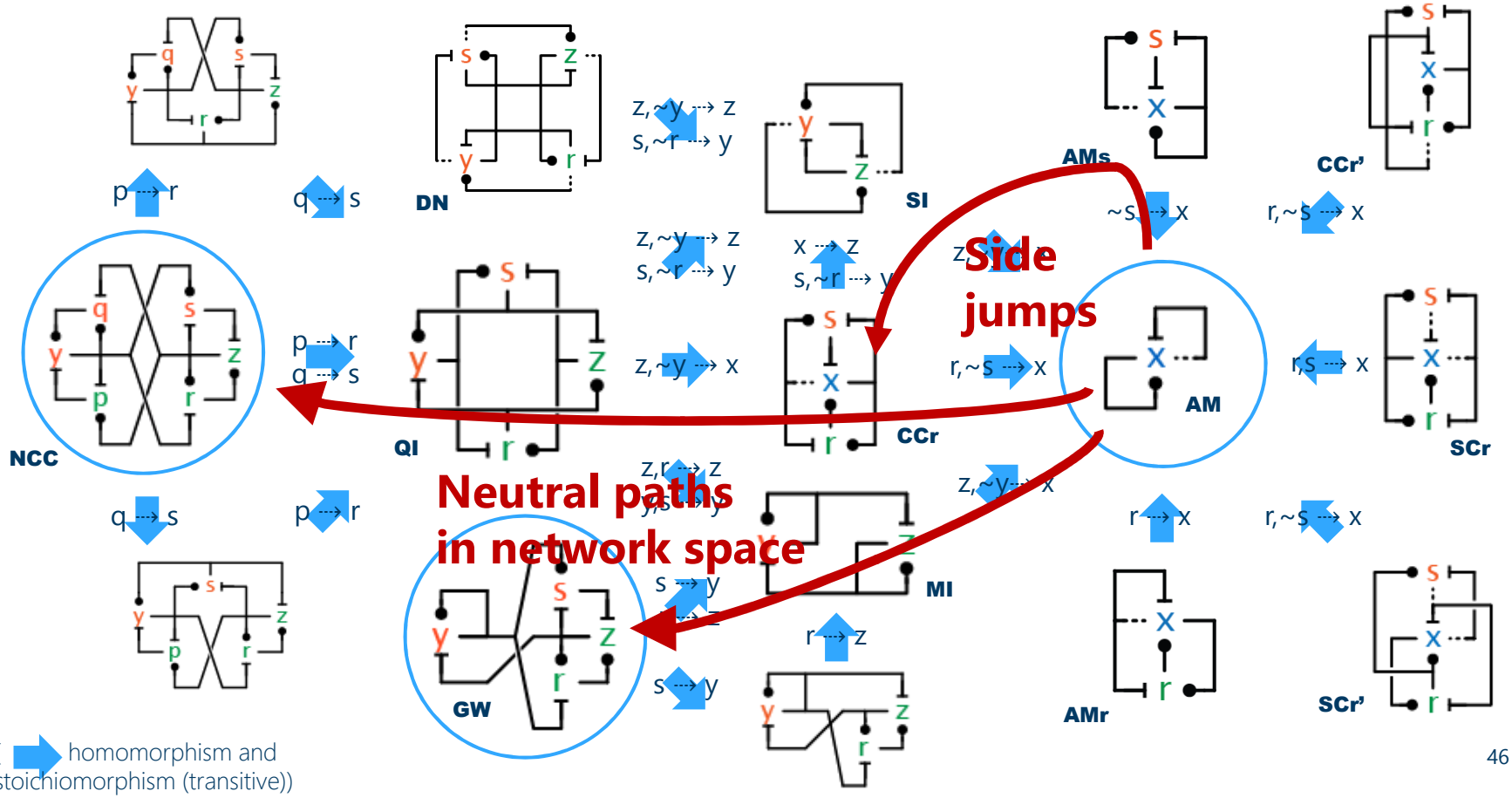
(homomorphism and stoichiomorphism (transitive))

Approximate Majority Emulation Zoo



(homomorphism and stoichiomorphism (transitive))

Approximate Majority Emulation Zoo



(⇢ homomorphism and stoichiomorphism (transitive))

Conclusions

Relating Networks

- Real biological networks
  - Are of course much more complex than these simple patterns
  - How much of that is obfuscation and how much is functional?
- 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
- Efficient algorithms can find emulations
  - Automatic model reduction of large networks

Computational Approach

- Q (traditional): What kind of dynamical system is the cell-cycle switch?
- A (traditional): Bistability – ultrasensitivity – hysteresis …
  - Focused on how sub-populations change over time.

- Q (computational): What kind of algorithmic system is the cell-cycle switch?
- A (computational): Interaction – complexity - convergence …
  - Focused on how individual molecules interact as algorithmic components.

- Leading to a better understanding of not just the *function* but also the *network* (algorithm).
  - If there is some clever population algorithm in nature that we have not invented yet (unlike AM) how shall we recognize it?