## The Cell Cycle Switch Computes Approximate Majority <br> Luca Cardelli, Microsoft Research \& Oxford University <br> Joint work with Attila Csikász-Nagy, Fondazione Edmund Mach \& King's College London <br> Northwestern University, CS+X Colloquium, 2014-04-29 <br> 

## Cells Compute

- No survival without computation!
- Finding food
- Avoiding predators
- How do they compute?
- Clearly doing "information processing"
- But can we actually catch nature running an (optimal) algorithm?


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

## Abstract Machines of Biology

Regulation


## Biological Languages



## Biological Networks



## How to build a good switch

## The Cell Cycle Switch

## Universal control mechanism regulating

 onset of M-phasePaul 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:


Double positive feedback on $x$ Double negative feedback on $x$

No feedback on y
Why ???


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

## Bela Novak ${ }^{+}$and John J. Tyson ${ }^{+}$

Separment of Biology, Virginia Polytechnici Institute and Stata Universit, Blacksburg. Virinina 20060.0406, USA


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


## How to Build a Good Switch

## -What is 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
- "Population protocol" switches
- Identical agents ('molecules') in a population start in some state, say x or y
- A pair of agents is chosen randomly at each step, they interact ('collide') and change state
- The whole population must eventually agree on a majority value (all-x or all-y) with probability 1


## A Bad Algorithm

- Direct Competition


## catalysis -o



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

$$
\begin{aligned}
& y+x \rightarrow x+x \\
& x+y \rightarrow y+y
\end{aligned}
$$

- 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)



## A Very 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 $O(n \log n)$ $\Rightarrow$ fast (optimal)

- The final outcome is correct if the initial disparity is $\omega(\operatorname{sqrt}(n) \log n)$ $\Rightarrow$ 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 $O(\log n)$

Dana Angluin • James Aspnes • David Eisenstat

## A Biological Implementation

## Approximate Majority (AM)



Bistable
Even when $x=y$ (stochastically)
Fast
$O(\log n)$ convergence time
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


Figure 1. Basic Ingredients of the Model
Theory
Theoretical Analysis of Epigenetic
Cell Memory by Nucleosome Modification

.

## Back to Biology

- The AM algorithm has ideal properties for settling a population into one of two states
- Seems like this would be useful in Biology
- Can we find biological implementations of this algorithm?
- Could it be related to the cell cycle switch?

Algorithms and Dynamical Systems

## Influence Nodes



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

Functional Motifs in Biochemical Reaction Networks

$\frac{d X_{i}}{d t}=r \frac{\left[A\left(1-X_{i}\right)-B_{1} X_{]}\right]}{A_{i}+B_{t}}, i=1, \ldots, N$,

activation
inhibition
catalysis
Approximate Majority



## Nature fixed the bug!

- There is another known feedback loop by which x suppresses s "in retaliation" via the so-called Greatwall loop; s and $x$ are antagonists: they are the two halves of the switch, mutually inhibiting each other (through intermediaries).
- Also, $s$ and $t$ happen to be the same molecule (=s)

- The "classical" cell cycle switch seems to be only half of the picture: the extra feedback completes it algorithmically and makes it as good as AM.


## In Previous Work

- 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!


Networks and Morphisms

## A Theory of Network Emulation <br> (with thanks to David Soloveichik)

- So far, evidence is empirical
- Specific 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
- A network emulates another network:
- When it can exactly reproduce the kinetics of another network for any choice of rates and initial conditions
- We aim to show that the cell cycle switch can emulate AM in that sense
- And moreover that the emulation is algorithmic: it is determined by network structure


## 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)

$$
\begin{gathered}
\text { initialize: } \\
z=x \\
\sim y=x \\
\left(y_{2}=x_{0}\right. \\
y_{1}=x_{1} \\
\left.y_{0}=x_{0}\right)
\end{gathered}
$$

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


## Emulation is a Network Morphism

A mapping of species and reactions

any initial conditions

$z_{0}=y_{2}=x_{0}$
$z_{1}=y_{1}=x_{1}$
$z_{2}=y_{0}=x_{2}$
initial conditions:
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

- Why does this work so well?

Kinetic Emulation

## When can a Network Emulate Another?

-What kind of morphisms guarantee emulation?

- do they preserve network structure?
- do they preserve stoichiometry?



## Chemical Reaction Networks

- A CRN is a pair $(S, R)$ where
- $S=\left\{s_{1}, \ldots, s_{n}\right\} \quad$ a finite set of species

$$
\begin{aligned}
& S=\{A, B, C\} \\
& R=\{r\}
\end{aligned}
$$

- $R=\left\{r_{1}, \ldots, r_{m}\right\} \quad$ a finite set of reactions
$r=2 A+B \rightarrow^{k} A+3 C$

$$
\rho \rightarrow^{k} \pi
$$

with stoichiometric numbers $\rho, \pi \in \mathbb{N}^{S}$

- The stoichiometry of $s$ in $\rho \rightarrow^{k} \pi$ is:

$$
\begin{aligned}
\eta\left(s, \rho \rightarrow^{k} \pi\right) & =\pi_{s}-\rho_{s} \\
\varphi\left(s, \rho \rightarrow^{k} \pi\right) & =k \cdot\left(\pi_{s}-\rho_{s}\right)
\end{aligned}
$$

$\eta(A, r)=-1 \quad$ net stoichiometry
$\varphi(A, r)=-k \quad$ (instantaneous) stoichiometry

## CRN Morphisms

A CRN morphism from $(S, R)$ to $(\hat{S}, \hat{R})$
written $m \in(S, R) \rightarrow(\hat{S}, \widehat{R})$
is a pair of maps $m=\left(m_{\mathcal{S}}, m_{\mathcal{R}}\right)$
a species map $m_{\mathcal{S}} \in S \rightarrow \hat{S}$
a reaction map $m_{\mathcal{R}} \in R \rightarrow \hat{R}$
(sometimes omitting the subscripts on $m$ )
We are interested in morphisms that are not injective, that represent refinements of simpler networks

Mappings (symmetries) between two networks


## 3 Key Morphisms

- A morphism $m \in(S, R) \rightarrow(\hat{S}, \hat{R})$ is
- a CRN homomorphism
if $m_{\mathcal{R}}$ is determined by $m_{\mathcal{S}}$ :

$$
m_{\mathcal{R}}\left(\rho \rightarrow^{k} \pi\right)=m_{\mathcal{S}}(\rho) \rightarrow^{k} m_{\mathcal{S}}(\pi) \quad \Rightarrow \quad \boldsymbol{m}_{\mathcal{S}}{ }^{\mathrm{T}} \cdot \boldsymbol{\varphi}=\widehat{\boldsymbol{\varphi}} \cdot \boldsymbol{m}_{\mathcal{R}}{ }^{\mathrm{T}}
$$

- a CRN reactant morphism if $m_{\mathcal{R}}$ is determined by $m_{\mathcal{S}}$ on reactants. $\exists \hat{k}, \hat{\pi}$ :

$$
m_{\mathcal{R}}\left(\rho \rightarrow^{k} \pi\right)=m_{\mathcal{S}}(\rho) \rightarrow^{\hat{k}} \hat{\pi} \quad \Leftrightarrow \quad \boldsymbol{m}_{\mathcal{S}}{ }^{\mathbf{T}} \cdot \boldsymbol{\rho}=\hat{\boldsymbol{\rho}} \cdot \boldsymbol{m}_{\mathcal{R}}{ }^{\mathrm{T}}
$$

- a CRN stoichiomorphism if:
def.
$\boldsymbol{\varphi} \cdot \boldsymbol{m}_{\mathcal{R}}=\boldsymbol{m}_{\boldsymbol{S}} \cdot \hat{\boldsymbol{\varphi}}$
$m_{S}(\rho)_{\hat{s}}=\Sigma_{s \in m_{S}-1(\hat{s})} \rho_{s}$


## Checking the Stoichiomorphism Condition



All unit rates (sufficient because of another theorem)
This is both a homomorphism and a stoichiomorphism


## CRN Kinetics

A state of a CRN $(S, R)$ is a $\boldsymbol{v} \in \mathbb{R}_{+}^{S}$
The differential system of a $\operatorname{CRN}(S, R), F \in \mathbb{R}_{+}^{S} \rightarrow \mathbb{R}^{S}$


Given by the law of mass action:

$$
F(\boldsymbol{v})(s)=\Sigma_{r=\left(\rho \rightarrow \rightarrow^{k} \pi\right) \in R} \varphi(s, r) \cdot \Pi_{\dot{S} \in S} \boldsymbol{v}_{\dot{S}}^{\rho_{s}}
$$

Usually written as a system of coupled concentration
ODEs, integrated over time: $\quad \frac{d v_{s}}{d t}=F(\boldsymbol{v})(s)$
a vector of concentrations for each species
$F(\boldsymbol{v})(s)$ gives the instantaneous change of concentration of a species in a given state
sum over all reactions of the stoichiometry of species in reaction times the product of reagent concentrations according to their stoichiometric numbers

## Kinetic Emulation

A morphism $m \in(S, R) \rightarrow(\hat{S}, \hat{R})$ is a $C R N$ emulation if for the respective differential systems $F, \hat{F}, \forall \hat{v} \in \mathbb{R}_{+}^{\hat{S}}$ :

$$
\begin{aligned}
& F(\widehat{\boldsymbol{v}} \circ m)=\hat{F}(\widehat{\boldsymbol{v}}) \circ m
\end{aligned}
$$

That is: $\forall s \in S . F(\widehat{\boldsymbol{v}} \circ m)(s)=\hat{F}(\widehat{\boldsymbol{v}})(m(s))$

if the derivative of $s$ (in state $\widehat{\boldsymbol{v}} \circ m$ )
equals the derivative of $m(s)$ (in state $\widehat{\boldsymbol{v}}$ )
if we start the two systems in states $\boldsymbol{v}=\widehat{\boldsymbol{v}} \circ m$ (which is a copy of $\widehat{\boldsymbol{v}}$ according to $m$ ) and $\widehat{\boldsymbol{v}}$ resp., for each $s$ the solutions are equal and the derivatives are equal, hence they will have identical trajectories by determinism

## Emulation Theorem

Theorem: If $m \in(S, R) \rightarrow(\hat{S}, \hat{R})$ is a CRN reactant morphism and stoichiomorphism then it is a CRN emulation
reactant morphism $\quad \boldsymbol{m}_{s}{ }^{\mathrm{T}} \cdot \boldsymbol{\rho}=\widehat{\boldsymbol{\rho}} \cdot \boldsymbol{m}_{\boldsymbol{R}}{ }^{\mathrm{T}}$
stoichiomorphism

$$
\varphi \cdot m_{\mathcal{R}}=m_{s} \cdot \widehat{\varphi}
$$

$\Downarrow$
emulation

$$
F(\widehat{v} \circ m)=\widehat{F}(\widehat{v}) \circ m
$$

N.B. homomorphism implies reactant morphism, implies $\boldsymbol{m}_{\boldsymbol{S}}{ }^{\mathrm{T}} \cdot \boldsymbol{\rho}=\widehat{\boldsymbol{\rho}} \cdot \boldsymbol{m}_{\mathcal{R}}{ }^{\mathrm{T}}$.
thus, for any initial conditions of $(\hat{S}, \hat{R})$ we can match trajectories


## Change of Rates Theorem

A change of rates for $(S, R)$ is morphism $\iota \in(S, R) \rightarrow\left(S, R^{\prime}\right)$ such that $\iota(S)$ is the identity and $\iota(\rho, \pi, k)=\left(\rho, \pi, k^{\prime}\right)$.

Theorem: If $m \in(S, R) \rightarrow(\hat{S}, \hat{R})$ is a stoichiomorphism, then for any change of rates $\hat{\imath}$ of $(\hat{S}, \hat{R})$ there is a change of rates $\iota$ of $(S, R)$ such that $\hat{\imath} \circ m \circ \iota^{-1}$ is a stoichiomorphism.
thus, for any rates of $(\hat{S}, \hat{R})$ we can match trajectories

In fact, $\iota$ changes rates by the ratio with which $\hat{\imath}$ changes rates:

$$
\iota(\rho, \pi, k)=\left(\rho, \pi, k \cdot \frac{\hat{k}^{\prime}}{\hat{k}}\right) \text { where } m(\rho, \pi, k)=(\hat{\rho}, \hat{\pi}, \hat{k}) \text { and } \hat{\imath}(\hat{\rho}, \hat{\pi}, \hat{k})=\left(\hat{\rho}, \hat{\pi}, \hat{k}^{\prime}\right)
$$



Network Zoos

## Approximate Majority Emulation Zoo



## Approximate Majority Emulation Zoo


( $\quad$ homomorphism and
stoichiomorphism (transitive))


## Emulations Compose: NCC emulates AM

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



## Emulation in Context



AM-AM Oscillator


AM-MI Oscillator
$m \in \mathrm{MI} \rightarrow \mathrm{AM}$ is an emulation:
it maps $z \rightarrow x$ and $\sim w \rightarrow x$

We can replace AM with MI in a context. The mapping $m$ tells us how to wire Ml 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 $\boldsymbol{z}$ 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 $w$.

## Another <br> Zoo




## Network Perturbations

Networ


A complex but robust implementation of the simple network

Normal Behavior

never dead "on average



## Conclusions

## Interpretations of Stoichiomorphism

- 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.


## Nature likes a good algorithm



Second part

any rates and initial conditions
These additional feedbacks do exist in real cell cycles (via indirections)

The cell cycle switch can exactly emulate AM


## In separate work...

- We produced a chemical implementation of AM using DNA gates
- I.e., a 'synthetic reimplementation' of the central cell-cycle switch.


## nature nanotechnology

nature.com $\stackrel{\text { - journal }}{ }$ home > archive $>$ issue > article - abstrac
ARTICLE PREVIEW
view full access options .
NATURE NANOTECHNOLOGY | ARTICLE
Programmable chemical controllers made from DNA

Yuan-Jyue Chen, Neil Dalchau, Niranjan Srinivas, Andrew Phillips, Luca Cardelli, David Soloveichik \& Georg Seelig


