Morphisms of Reaction Networks

Luca Cardelli, Microsoft Research & Oxford University

IMT Lucca, 2014-07-08
Motivation

- 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*
A Consensus Algorithm

- Population Protocols
  - Finite-state identity-free agents (molecules) interact in randomly chosen pairs
  - Each interaction (collision) can result in state changes
  - Complete connectivity, no centralized control (well-mixed solution)

- A Population Consensus Problem
  - Find which state $x$ or $y$ is in majority in the population
  - By converting the whole population to $x$ or $y$

- Approximate Majority (AM) Algorithm
  - Uses a third “undecided” state $b$
  - Disagreements cause agents to become undecided
  - Undecided agents believe any non-undecided agent

- With high probability, for $n$ agents
  - The total number of interactions is $O(n \log n)$ ⇒ fast (optimal)
  - Correct outcome if the initial disparity is $\omega(\sqrt{n} \log n)$ ⇒ robust
  - In parallel time, converges in $O(\log n)$

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

\[
\begin{align*}
x + y & \rightarrow^r y + b \\
y + x & \rightarrow^r x + b \\
b + x & \rightarrow^r x + x \\
b + y & \rightarrow^r y + y
\end{align*}
\]
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 \textit{whp}

Epigenetic Switch

Figure 1. Basic Ingredients of the Model
Motivation (cont’d)

- 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 convoluted and... approximate.
  - Like finding an algorithm in a haystack...
Obfuscated Implementations

- GW is a better cell cycle switch than [the traditional switch]

GW = AM "obfuscated"

switch reset

new feedback

GW

GW is how the cell cycle switch “really works”
Motivation (cont’d)

• When does a biologically messy network X “implement” some ideal algorithm Y?
  • Pushed coauthors into thinking about approximate stochastic bisimulation metrics for CTMCs
  • But they didn’t come back...

• Some networks behave similarly because “their ODEs are just equivalent” [David S.]
  • Epiphany! Forget stochastic! Forget approximate! When are CRNs “deterministically equivalent”? 
  • Or better, when do trajectories of one CRN “collapse” into trajectories of another?
  • Much simpler! And can be solved on the static structure of CRNs as opposed to their kinetics.
  • Independently on rates and initial conditions (of one of the two networks).
Influence Networks

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

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

They actually implement a Hill function of coefficient 2:

\[
\text{Inhibit } x 
\xrightarrow{\text{inhibit } x} \text{Activate } x
\]

Approximate Majority

Reaction Network

Influence Network
**Biological Influence Networks**

- **Mutual Inhibition & Self Activation**
- **Mutual Inhibition & Mutual Anti-activation**
- **Cell Cycle Switching**
- **Better Switching**

**Cell cycle transitions**

**Polarity establishment**

**Gene networks**

Construction of a genetic toggle switch in *Escherichia coli*

**Septation Initiation**

**The G2/M cell cycle switch**

Numerical analysis of a comprehensive model of G2-phase control in *Xenopus* oocytes extracts good/bad switches

**The “new” cell cycle switch**

Phosphorylation network dynamics in the control of cell cycle transitions
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:

  \[ \begin{align*}
  y_0 &= x_0 \\
  y_1 &= x_1 \\
  y_2 &= x_0
  \end{align*} \]

- How do we find these matching parameters? By a network morphism!
MI to AM Emulation: Network Morphism

A mapping of species and reactions

\[ \text{homomorphic mapping} \]

\[ \begin{align*}
    z &\rightarrow x \\
    \sim y &\rightarrow x
\end{align*} \]

\[ \begin{align*}
    z_0 &= y_2 = x_0 \\
    z_1 &= y_1 = x_1 \\
    z_2 &= y_0 = x_2
\end{align*} \]

any initial conditions

less trivial than you might think: it need not preserve the out-degree of a node!
SI to AM Emulation: Network Morphism

A mapping of species and reactions

*homomorphic* mapping

\[ z \rightarrow x \]
\[ \sim y \rightarrow x \]

Any initial conditions:

\[ z_0 = y_2 = x_0 \]
\[ z_1 = y_1 = x_1 \]
\[ z_2 = y_0 = x_2 \]
AMr to AM Emulation: Network Morphism

A mapping of species and reactions

**Homomorphic mapping**

\[ x \rightarrow x \]

\[ r \rightarrow x \]

(AMr adds an indirection to the x positive loop; if we also add an indirection to the x negative loop, we obtain a prototypical cell cycle switch that also emulates AM: CCR)

Any initial conditions:

\[ r_0 = x_0 = x_0^{AM} \]

\[ r_1 = x_1 = x_1^{AM} \]

\[ r_2 = x_2 = x_2^{AM} \]
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?
Emulations Compose: NCC emulates AM

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

![Diagram of NCC and AM trajectories]
Approximate Majority Emulation Zoo

( homomorphism and stoichiomorphism (transitive))
Approximate Majority Emulation Zoo

( homomorphism and stoichiomorphism (transitive))
Approximate Majority Emulation Zoo

Neutral paths in network space

Side jumps
\textbf{Emulation in Context}

\textit{m} \in \text{MI} \rightarrow \text{AM} \text{ is an emulation: it maps } z \rightarrow x \text{ and } \sim w \rightarrow x

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

Each influence crossing the dashed lines into \( x \) is replaced by a similar influence into \textit{both} \( z \) \textit{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 \textit{either} \( z \) \textit{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 Zoo
Network Perturbations

Network               Normal Behavior         Removing each link in turn

A complex but robust implementation of the simple network
Morphisms of CRNs
A Theory of Network Emulation
(with thanks to David Soloveichik)

- So far, evidence is empirical
  - Specific simulations based on a choice of parameters

- But indeed...
  - We can show analytically 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 (of the target network)

- A network emulates another network:
  - When it can exactly reproduce the kinetics of another network for any choice of rates and initial conditions (of the other network)
  - We aim to show that e.g. the cell cycle switch can emulate AM in that sense
  - And moreover that the emulation is algorithmic: it is determined by static network structure (including rate constants and stoichiometric constants), not by random kinetic
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
  - $S = \{s_1, ..., s_n\}$ a finite set of species
  - $R = \{r_1, ..., r_m\}$ a finite set of reactions

- Reactions $r = \rho \rightarrow^k \pi \in R$
  with complexes $\rho, \pi \in \mathbb{N}^S$
  stoichiometric numbers $\rho_s, \pi_s$ for $s \in S$
  and rate constants $k > 0$

- The stoichiometry of $s$ in $\rho \rightarrow^k \pi$ is:
  $\eta(s, \rho \rightarrow^k \pi) = \pi_s - \rho_s$
  $\varphi(s, \rho \rightarrow^k \pi) = k \cdot (\pi_s - \rho_s)$

- $S = \{A, B, C\}$
  $R = \{r\}$

- $r = 2A + B \rightarrow^k A + 3C$
  $\rho_A = 2, \rho_B = 1, \rho_C = 0$
  $\pi_A = 1, \pi_B = 0, \pi_C = 3$

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

- $\rho \rightarrow^k \pi, \rho \rightarrow^k' \pi \in R \Rightarrow k = k'$
CRN Morphisms

A *CRN morphism* from \((S, R)\) to \((\hat{S}, \hat{R})\)
written \(m \in (S, R) \rightarrow (\hat{S}, \hat{R})\)
is a pair of maps \(m = (m_S, m_R)\)

- a *species map* \(m_S \in S \rightarrow \hat{S}\)
- a *reaction map* \(m_R \in R \rightarrow \hat{R}\)

extended to a *complex map* \(m_S \in \mathbb{N}^S \rightarrow \hat{\mathbb{N}}^\hat{S}\)
linearly: \(m_S(\rho)_{\hat{S}} = \sum_{S \in m_S^{-1}(\hat{S})} \rho_S\)

(sometimes omitting the subscripts on \(m\))
3 Key Morphisms

- A morphism \( m \in (S, R) \rightarrow (\hat{S}, \hat{R}) \) is
  - a **CRN homomorphism** if \( m_R \) is determined by \( m_S \):
    \[
    m_R(\rho \rightarrow^k \pi) = m_S(\rho) \rightarrow^k m_S(\pi)
    \]
  \[\Rightarrow\]
  \[
  m_S^T \cdot \varphi = \hat{\varphi} \cdot m_R^T
  \]
  - a **CRN reactant morphism** if \( m_R \) is determined by \( m_S \) on reactants. \( \exists \hat{k}, \hat{\pi} : \)
    \[
    m_R(\rho \rightarrow^k \pi) = m_S(\rho) \rightarrow^\hat{k} \hat{\pi}
    \]
  \[\Leftrightarrow\]
  \[
  m_S^T \cdot \rho = \hat{\rho} \cdot m_R^T
  \]
  - a **CRN stoichiomorphism** if:
    \[
    \varphi \cdot m_R = m_S \cdot \hat{\varphi}
    \]

\( \varphi, \hat{\varphi} \) are the respective stoichiometric matrices
\( \rho, \hat{\rho} \) are the respective reactant matrices
\( m_S, m_R \) are the characteristic 0-1 matrices of \( m_S, m_R \)
\[
\begin{align*}
    m_S(s, \hat{s}) = 1 & \text{ if } m_S(s) = \hat{s} \text{ else } 0 \\
    m_R(s, \hat{s}) = 1 & \text{ if } m_R(s) = \hat{s} \text{ else } 0
\end{align*}
\]
Checking the Stoichiomorphism Condition

\[ m \in \text{MI} \rightarrow \text{AM} \]

\[ \forall s \in S. \quad \forall \hat{r} \in \hat{R}. \quad \Sigma_{r \in m^{-1}(\hat{r})} \varphi(s, r) = \varphi(m(s), \hat{r}) \]

\( m^{-1}(am_0) \)

All unit rates (sufficient because of another theorem)

This is both a homomorphism and a stoichiomorphism

\[ \varphi(y_0, m_0) + \varphi(y_0, m_4) = -1 = \varphi(x_2, am_0) \]
CRN Kinetics

A state of a CRN \((S, R)\) is a \(\mathbf{v} \in \mathbb{R}_+^S\).

The differential system of a CRN \((S, R)\), \(F \in \mathbb{R}_+^S \rightarrow \mathbb{R}^S\)

Given by the law of mass action:

\[
F(\mathbf{v})(s) = \sum_{r \in R} \varphi(s, r) \cdot [r]_\mathbf{v}
\]

Usually written as a system of coupled concentration ODEs, integrated over time:

\[
\frac{d\mathbf{v}_s}{dt} = F(\mathbf{v})(s)
\]

\(F(\mathbf{v})(s)\) gives the instantaneous change of concentration of a species in a given state.

The mass action of a reaction in a state is the product of reagent concentrations according to their stoichiometric numbers:

\[
[\rho \rightarrow^k \pi]_\mathbf{v} = \mathbf{v}^\rho = \prod_{s \in S} \mathbf{v}_s^{\rho_s}
\]
Kinetic Emulation

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

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

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

if we start the two systems in states $v = \hat{v} \circ m$ (which is a copy of $\hat{v}$ according to $m$) and $\hat{v}$ 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 \( m \in (S, R) \to (\hat{S}, \hat{R}) \) is a CRN reactant morphism and stoichiomorphism then it is a CRN emulation. 

- **Reactant Morphism**
  \[ m_S^T \cdot \rho = \hat{\rho} \cdot m_R^T \]
  - Preserve enough network structure
- **Stoichiomorphism**
  \[ \varphi \cdot m_R = m_S \cdot \hat{\varphi} \]
  - Preserve enough chemical stoichiometry
- **Emulation**
  \[ \forall \hat{\mu}. \; F(\hat{\mu} \circ m_S) = \hat{F}(\hat{\mu}) \circ m_S \]
  - Preserve derivatives

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

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

A \textit{change of rates} for \((S, R)\) is morphism \(\iota \in (S, R) \to (S, R')\) such that \(\iota(S)\) is the identity and \(\iota(\rho, \pi, k) = (\rho, \pi, k')\).

\textbf{Theorem:} If \(m \in (S, R) \to (\hat{S}, \hat{R})\) is a stoichiomorphism, then for any change of rates \(\hat{\iota}\) of \((\hat{S}, \hat{R})\) there is a change of rates \(\iota\) of \((S, R)\) such that \(\hat{\iota} \circ m \circ \iota^{-1}\) is a stoichiomorphism.

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

\[\iota(\rho, \pi, k) = \left(\rho, \pi, k \cdot \frac{\hat{k}'}{\hat{k}}\right)\text{ where } m(\rho, \pi, k) = (\hat{\rho}, \hat{\pi}, \hat{k})\text{ and } \hat{\iota}(\hat{\rho}, \hat{\pi}, \hat{k}) = (\hat{\rho}, \hat{\pi}, \hat{k}').\]
Corollaries

- By checking only static network and morphism properties we can learn that:
  - All these networks are (at least) bistable
  - (We do not have to reanalyze the steady states of all these dynamical systems)

- All these networks can perform exactly as fast as AM
- (We do not have to reprove the complexity bounds for all these networks)
Examples of CRN morphisms

Circles are species and squares are reactions. Red arrows are species mappings \( m_S \) and blue arrows are reaction mappings \( m_R \). Solid arrows indicate morphisms that are emulations.

(A) A simple stoichiomorphism: the species in the source reactions are distinct. In general, multiple separate copies of a system will map to it via a trivial map that is a homomorphism and stoichiomorphism.

(B) This is a homomorphism, but is not a stoichiomorphism. For \( s_0, \hat{r}_0 : \Sigma_{r \in m_R(\hat{r}_0)} \varphi(s_0, r) = -2 \neq -1 = \varphi(m_S(s_0), \hat{r}_0) \).

(C) This is a stoichiomorphism, but is not a homomorphism or a reactant morphism. \( r_0 = \rho \rightarrow \pi \) with \( \rho s_0 = 1 \) but \( m_R(r_0) = \hat{r}_0 = \hat{\rho} \rightarrow \hat{\pi} \) with \( \hat{\rho} m_S(s_0) = \hat{\rho} s_0 = 2 \), so \( \hat{\rho} \neq m_S(\rho) \) and \( m_R(r_0) \neq m_S(\rho) \rightarrow \hat{\pi} \).

(D) This is a homomorphism but not a stoichiomorphism. For \( s_1, \hat{r}_0 : \Sigma_{r \in m_R(\hat{r}_0)} \varphi(s_1, r) = 1 \neq 2 = \varphi(m_S(s_0), \hat{r}_0) \).

(E) This stoichiomorphism is not a homomorphisms, but is a reactant morphism. \( r_0 = \rho \rightarrow \pi \) and \( m_R(r_0) = \hat{r}_0 = \hat{\rho} \rightarrow \hat{\pi} \) with \( \hat{\rho} = m_S(\rho) \) and \( m_R(r_0) = m_S(\rho) \rightarrow \hat{\pi} \).

(F) This reactant morphism is not a homomorphism but is a stoichiomorphism. E.g., for \( s_1, \hat{r}_0 : \Sigma_{r \in m_R(\hat{r}_0)} \varphi(s_1, r) = \varphi(s_1, r_0) + \varphi(s_1, r_1) = 2 \cdot k + 0 \cdot k = 1 \cdot 2k = \varphi(m_S(s_1), \hat{r}_0) \).
Conclusions
Network Emulation Morphisms

- **What guarantees emulation?**
  - Reactant morphism + stoichiomorphism: static, state-independent (*structural*) conditions

- **How do you find them?**
  - Emulation Theorem => they do not depend on initial conditions
  - Change of Rates Theorem => can look for rate-1 morphisms
  - E.g. test all possible rate-1 homomorphism between two networks to see if they are stoichiomorphisms

- **How common are they?**
  - Likely relatively rare, but still many useful ones => richness of networks space

- **How useful are they?**
  - Establish structural, algorithmic, (non-accidental) reasons for kinetic similarity
  - Explain simple behavior “facets” of complicated networks
  - Investigate evolutionary paths (maybe)

- **How brittle are they?**
  - Will a perturbed trajectory of the source network converge to a trajectory of the target network?
  - What about other reaction kinetics?

- **What about stochastic?**
  - Is there a CME Emulation Theorem?
# Population Majority

<table>
<thead>
<tr>
<th>Year</th>
<th>Title</th>
<th>Authors</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>Computation in networks of passively mobile finite-state sensors</td>
<td>Dana Angluin, James Aspnes, Zoe Diamadi, Michael J. Fischer, Renee Peralta</td>
<td>PODC'04</td>
</tr>
<tr>
<td>2007</td>
<td>Theoretical Analysis of Epigenetic Cell Memory by Nucleosome Modification</td>
<td>Ian B. Dodd, Mille A. Micheelsen, Kim Sneppen, Genevieve Thon</td>
<td>Cell.</td>
</tr>
<tr>
<td>2009</td>
<td>Artificial Biochemistry</td>
<td>Luca Cardelli</td>
<td>Algorithmic Bioprocesses, Springer</td>
</tr>
<tr>
<td>2009</td>
<td>Using Three States for Binary Consensus on Complete Graphs</td>
<td>Etienne Peron, Chirag Vasudevan, Milan Vojnovic</td>
<td>IEEE Infocom.</td>
</tr>
<tr>
<td>2010</td>
<td>Convergence Speed of Binary Interval Consensus</td>
<td>Moez Draief, Milan Vojnovic</td>
<td>Infocom'10</td>
</tr>
<tr>
<td>2012</td>
<td>The Cell Cycle Switch Computes Approximate Majority</td>
<td>Luca Cardelli, Attila Csikasz-Nagy</td>
<td>Scientific Reports.</td>
</tr>
<tr>
<td>2014</td>
<td>Morphisms of Reaction Networks that Couple Structure to Function</td>
<td>Luca Cardelli</td>
<td></td>
</tr>
</tbody>
</table>

**Majority**

- The value of the majority function is 1 if there are more 1's than 0's in the input; otherwise, it is 0.
- The states of our protocol consist of a live bit and a counter with values in the set \{-1, 0, 1\}. Initially, the live bit is set to 1.

**Exact Majority**

- 6-state Nondeterministic (population protocol)

**Approximate Majority**

- 3-state Stochastic, discrete time (DTMC) Fundamental results.
- 3-state Stochastic, discrete time (ad-hoc)
- 3-state Stochastic, continuous time (CTMC). Simulations.
- 3-state Stochastic, continuous time (DDE). Emulation theorem.

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

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

**3-state Continuous space, continuous time (Deterministic ODE): Emulation theorem.**