Morphisms of Reaction Networks

Luca Cardelli, Microsoft Research & Oxford University

BIRS Programming with CRNs, Banff, 2014-06-11
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 $O(\log n)$
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

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”

GW = AM “obfuscated”

GW is how the cell cycle switch “really works”

GW = AM “obfuscated”

GW = AM “obfuscated”
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

\[
\begin{align*}
\text{inhibition} &\quad \text{high} \quad x \quad \text{low} \\
\text{activation} &\quad \text{inhibit } x \\
&\quad x \text{ is high} \\
&\quad x \text{ is low}
\end{align*}
\]

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:

\[
\begin{align*}
\text{activation} \\
\text{inhibition} \\
\text{catalysis}
\end{align*}
\]

Approximate Majority

Reaction Network

Influence Network

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

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

A mapping of species and reactions

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

homomorphic mapping

$z \rightarrow x$
$\sim y \rightarrow x$

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

**Homomorphically mapping**

- \( z \rightarrow x \)
- \( \sim y \rightarrow x \)

**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

*homomorphics* 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)
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?

(3 species each)

initialize
z,r,p = z
y,q,s = y
Emulations Compose: NCC emulates AM

- The (18) trajectories NCC can always retrace those (3) of AM.
Approximate Majority Emulation Zoo

\[ \text{homomorphism and stoichiometry (transitive)} \]
Approximate Majority Emulation Zoo

Homomorphism and stoichiomorphism (transitive)
Approximate Majority Emulation Zoo

Neutral paths in network space

Side jumps

(homomorphism and stoichiomorphism (transitive))
$m \in \text{MI} \rightarrow \text{AM}$ 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 MI to obtain an overall emulation:

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

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

"Emulation in Context"
Another Zoo
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
  - \(S = \{s_1, \ldots, s_n\}\) a finite set of species
  - \(R = \{r_1, \ldots, r_m\}\) 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:
  \[
  \eta(s, \rho \rightarrow^k \pi) = \pi_s - \rho_s \\
  \varphi(s, \rho \rightarrow^k \pi) = k \cdot (\pi_s - \rho_s)
  \]

\(r = 2A + B \rightarrow^k A + 3C\)
\[
\rho_A = 2, \ \rho_B = 1, \ \rho_C = 0 \\
\pi_A = 1, \ \pi_B = 0, \ \pi_C = 3
\]

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

\(^(*)\) \(\rho \rightarrow^k \pi, \rho \rightarrow^{k'} \pi \in R \implies k = k'\)
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: \(m_S(\rho)_{\hat{S}} = \Sigma_{S \in m_S^{-1}(\hat{S})} \rho_S\)

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

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

\( \varphi, \hat{\varphi} \) are the respective stoichiometric matrices
\( \rho, \hat{\rho} \) are the respective reactant matrices
\( m_S, m_R \) are the characteristic 0-1 matrices of \( m_S, m_R \)

\[
m_S(s, \hat{s}) = \begin{cases} 1 & \text{if } m_S(s) = \hat{s} \\ 0 & \text{else} \end{cases}
\]

\[
\begin{align*}
\Rightarrow & \quad m_S^T \cdot \varphi = \hat{\varphi} \cdot m_R^T \\
\iff & \quad m_S^T \cdot \rho = \hat{\rho} \cdot m_R^T \\
\text{def.} & \quad \varphi \cdot m_R = m_S \cdot \hat{\varphi}
\end{align*}
\]
Checking the Stoichiomorphism Condition

\[ m \in MI \rightarrow AM \]

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

\[ \varphi(y_0, m_{i_0}) + \varphi(y_0, m_{i_4}) = -1 = \varphi(x_2, a_{m_0}) \]

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 \(v \in \mathbb{R}^S_+\).

The *differential system* of a CRN \((S, R)\), \(F \in \mathbb{R}^S_+ \rightarrow \mathbb{R}^S\) given by the law of mass action:

\[
F(v)(s) = \Sigma_{r \in R} \varphi(s, r) \cdot [r]_v
\]

Given by the law of mass action:

\[
F(v)(s) = \Sigma_{r \in R} \varphi(s, r) \cdot [r]_v
\]

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

\[
\frac{dv_s}{dt} = F(v)(s)
\]

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

Sum over all reactions of the stoichiometry of the species in the reaction times the mass action of the reaction in the state.

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

\[
[r \rightarrow^k \pi]_v = v^\rho = \Pi_{s \in S} v^\rho_s
\]
Kinetic Emulation

A morphism $m \in (S, R) \to (\hat{S}, \hat{R})$ 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 the derivative of $s$ (in state $\hat{v} \circ m$) equals the derivative of $m(s)$ (in state $\hat{v}$) 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 $m \in (S, R) \rightarrow (\hat{S}, \hat{R})$ is a CRN reactant morphism and stoichiomorphism then it is a CRN emulation.

- Reactant morphism: $m_S^T \cdot \rho = \hat{\rho} \cdot m_R^T$
- Stoichiomorphism: $\varphi \cdot m_R = m_S \cdot \hat{\varphi}$
- Emulation: $\forall \hat{v} . \ F(\hat{v} \circ m_S) = \hat{F}(\hat{v}) \circ m_S$

$F$ is the differential system of $(S,R)$, given by the law of mass action, $\hat{v}$ 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 change of rates for \((S, R)\) is morphism \(\iota \in (S, R) \rightarrow (S, R')\) such that \(\iota(S)\) is the identity and \(\iota(\rho, \pi, k) = (\rho, \pi, k')\).

**Theorem:** If \(m \in (S, R) \rightarrow (\hat{S}, \hat{R})\) 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:

\[\iota(\rho, \pi, k) = \left(\rho, \pi, \frac{k'}{k}\right)\]

where \(m(\rho, \pi, k) = (\hat{\rho}, \hat{\pi}, \hat{k})\) and \(\hat{\iota}(\hat{\rho}, \hat{\pi}, \hat{k}) = (\hat{\rho}, \hat{\pi}, \hat{k}')\).

Thus, for any rates of \((\hat{S}, \hat{R})\) we can match trajectories.
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 $m_R(r_0) \neq m_S(\rho) \rightarrow \hat{\rho}$.

(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 $m_R(r_0) = m_S(\rho) \rightarrow \hat{\rho}$.

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

34
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

Dana Angluin, James Aspnes, Zoë Diamadi, Michael J. Fischer, René Peralta. POCDC'04.

Dana Angluin, James Aspnes, David Eisenstat. DISC'07.

## 2007: Theoretical Analysis of Epigenetic Cell Memory by Nucleosome Modification.

Luca Cardelli. Algorithmic Bioprocesses, Springer.

## 2009: Robust Stochastic Chemical Reaction Networks and Bounded Tau Leaping (Appendix 4).

Etienne Perron, Dinkar Vasudevan, and Milan Vojnovic. IEEE Infocom.

## 2010: Convergence Speed of Binary Interval Consensus.
Milan Vojnovic. Infocom'10.

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

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

<table>
<thead>
<tr>
<th>Year</th>
<th>Title</th>
<th>Authors</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>Computation in networks of passively mobile finite-state sensors.</td>
<td>Dana Angluin, James Aspnes, Zoë Diamadi, Michael J. Fischer, René Peralta</td>
<td>POCDC'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</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 Perron, Dinkar Vasudevan, and Milan Vojnovic</td>
<td>IEEE Infocom</td>
</tr>
<tr>
<td>2010</td>
<td>Convergence Speed of Binary Interval Consensus.</td>
<td>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 Csikász-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>

**Notes:**
- $\text{Approximate Majority - 3-state}$
- $\text{Stochastic, discrete time}$
- $\text{(DTMC) Fundamental results.}$
- $\text{Transfer complexity results from discrete time population protocols to continuous time stochastic chemical reaction networks.}$
- $\text{Exact Majority - 6-state}$
- $\text{Nondeterministic}$
- $\text{(population protocol)}$
- $\text{Approximate Majority - 3-state}$
- $\text{Stochastic, discrete time}$
- $\text{(ad-hoc)}$
- $\text{Approximate Majority - 3-state}$
- $\text{Stochastic, continuous time}$
- $\text{(CTMC). Simulations.}$
- $\text{Approximate Majority - 3-state}$
- $\text{Stochastic, continuous time}$
- $\text{(CTMC) Fundamental results.}$
- $\text{Exact Majority - 4-state}$
- $\text{Stochastic, continuous time}$
- $\text{(similar to 2004 paper)}$
- $\text{Approximate Majority - 3-state}$
- $\text{Continuous space, continuous time}$
- $\text{(Deterministic ODE: Emulation theorem.}$