Morphisms of Reaction Networks

Luca Cardelli, Microsoft Research & Oxford University

with: Mirco Tribastone, Max Tschaikowski, Andrea Vandin
IMT Institute for Advanced Studies, Lucca
Attila Csikász-Nagy
King’s College London
Neil Dalchau
Microsoft Research Cambridge

Beilstein Symposium 2018-06-07
Mapping one network into another

- We would like *simple* understanding of *complex* systems
  - Subnetworks, motifs, model reduction, ...

- But we also want to preserve meaning
  - What is a good model of a cat?

- Understanding how complex systems may arise from simpler systems

- How to reconcile?
  - Look for relationships between large (complex) and small (simple) networks that preserve *structure* and *function*.

---

Norbert Wiener
Pioneer of stochastic processes and inventor of Cybernetics.

“The best material model of a cat is another, or preferably the same, cat”
Comparing networks by morphing them

- How can we compare different networks?
  - Different number of species
  - Different number of reactions
  - Apparently unrelated connectivity

- How is structure related to function and performance?
  - Does antagonism (in network structure) guarantee bistability (in function)?

- We *morph* networks onto one another (*structurally*) so that they *emulate* each other (*’s function*)
  - Deterministic version of simulation of reactive systems
Morphisms of Antagonistic Networks
Antagonistic Networks

1 vs. 1
Mutual Inhibition & Self Activation

1 vs. 1
Mutual Inhibition & Mutual Anti-activation

2 vs. 2
low Notch -> high Delta
low Delta =< high Notch

3 vs. 3
high Delta = high Notch

Cell cycle transitions
Polarity establishment
Gene networks

Septation Initiation
Delta-Notch

The “new” cell cycle switch
Phosphorylation network dynamics in the control of cell cycle transitions
A Consensus Algorithm

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

\[
\begin{align*}
\text{x} + \text{y} &\rightarrow \text{r} \quad \text{y} + \text{b} \\
\text{y} + \text{x} &\rightarrow \text{r} \quad \text{x} + \text{b} \\
\text{b} + \text{x} &\rightarrow \text{r} \quad \text{x} + \text{x} \\
\text{b} + \text{y} &\rightarrow \text{r} \quad \text{y} + \text{y}
\end{align*}
\]

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 \textit{whp}

Epigenetic Switch

![Epigenetic Switch Diagram]

Figure 1. Basic Ingredients of the Model

Dana Angluin - James Aspnes - David Eisenstat
A Simple Population Protocol for Fast Robust Approximate Majority 2007
Not always that simple

- The epigenetic switch seems 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 Triplet Model of Influence

**Inhibition**
- High (modified)
- Low (unmodified)

**Activation**
- Usually modeled by sigmoid (e.g. Hill or Reinitz) functions
- Biological mechanism: (e.g.) multisite phosphorylation

**Triplet Motif**
- We model them by 4 mass action reactions over 3 species $x_0$, $x_1$, $x_2$
- They actually implement a Hill function of coefficient 2:

$$
\begin{align*}
&\text{activation} & \rightarrow & \text{inhibition} \\
&\text{(mass action) catalysis} & \text{activation} & \text{inhibition}
\end{align*}
$$

**For example:**

$$
\begin{align*}
x_0 + x_2 & \rightarrow^{0.1} r_01 x_2 + x_1 \\
\end{align*}
$$

Approximate Majority
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:

  ![Diagram](image)

  - (6 species on 3 trajectories)
  - (3 species on 3 trajectories)

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

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

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

Mappings (symmetries) between two networks

- 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

\(\varphi\) 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.
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

\[ z, r, p \rightarrow z \]
\[ y, q, s \rightarrow y \]
The (18) trajectories NCC can *always* retrace those (3) of AM.

The new cell cycle switch can emulate AM *exactly*. For *any* initial conditions of AM. And for *any* rates of AM.
Emulations are Modular
Nature seems to like good algorithms

(G) Close but
No Emulation

Exact Emulation

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

The cell cycle switch can exactly emulate AM
Nature seems to like good algorithms

- The cell cycle switch emulates approximate majority
  - Hence it can switch as fast as Approximate Majority (it can follow the same trajectories)
  - And Approximate Majority is optimal!
- And it is as robust to perturbation as Approximate Majority
  - Which can resist large fluctuations
How to check for emulation

• How do we check a potential emulation morphism for all possible initial conditions of the target?
  • Statically! Check conditions on the joint stoichiometric matrices of the two networks under the mapping.

• How do we check a potential emulation morphism for all possible rates of the target?
  • Can’t; but if one emulation is found, then the rates of the target network can be changed \textit{arbitrarily} and a related emulation will again exist.
Applications of Emulation

- Model Reduction
  - Find reduced networks
  - Compute quotient CRNs
  - Find network symmetries that may be of biological interest

- Morphism Generation
  - Find morphisms between networks (e.g. all the ones for a fixed rate assignment)

Benchmarks from Sneddon et al., Nature Methods, 2011

<table>
<thead>
<tr>
<th>Model</th>
<th>Reactions</th>
<th>Species</th>
<th>FB</th>
<th>Time (s)</th>
<th>BB</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>e9</td>
<td>3538944</td>
<td>262146</td>
<td>222</td>
<td>4.61E+4</td>
<td>222</td>
<td>7.65E+4</td>
</tr>
<tr>
<td>e8</td>
<td>786432</td>
<td>65538</td>
<td>167</td>
<td>1.92E+3</td>
<td>167</td>
<td>3.68E+3</td>
</tr>
<tr>
<td>e7</td>
<td>172032</td>
<td>16386</td>
<td>122</td>
<td>8.15E+1</td>
<td>122</td>
<td>1.77E+2</td>
</tr>
<tr>
<td>e6</td>
<td>36864</td>
<td>4098</td>
<td>86</td>
<td>3.00E+0</td>
<td>86</td>
<td>7.29E+0</td>
</tr>
<tr>
<td>e5</td>
<td>7680</td>
<td>1026</td>
<td>58</td>
<td>1.54E-1</td>
<td>58</td>
<td>4.06E-1</td>
</tr>
<tr>
<td>e4</td>
<td>1336</td>
<td>253</td>
<td>37</td>
<td>9.00E-3</td>
<td>37</td>
<td>1.09E-1</td>
</tr>
<tr>
<td>e3</td>
<td>288</td>
<td>66</td>
<td>22</td>
<td>1.00E-3</td>
<td>22</td>
<td>3.00E-3</td>
</tr>
<tr>
<td>e2</td>
<td>48</td>
<td>18</td>
<td>12</td>
<td>1.00E-3</td>
<td>12</td>
<td>2.00E-3</td>
</tr>
</tbody>
</table>
Network Morphisms as Evolutionary Paths
Network Evolution

“same function”

“new function”
Emulation Zoo

emulation (transitive)
Walks in Network Space

Neutral paths in network space

Side jumps

emulation (transitive)
Another Zoo
### Network Perturbations

<table>
<thead>
<tr>
<th>Network</th>
<th>Normal Behavior</th>
<th>Removing each link in turn</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Network Diagram" /></td>
<td><img src="image" alt="Normal Behavior Graph" /></td>
<td><img src="image" alt="Removing each link in turn" /></td>
</tr>
</tbody>
</table>

A complex but robust implementation of the simple network

- **Network**: A complex but robust implementation of the simple network.
- **Normal Behavior**: Graph showing typical behavior of the network.
- **Removing each link in turn**: Graphs illustrating the effect of removing each link in turn, with "dead" indicating failure and "never dead "on average"" indicating resilience.

24
Noise Reduction in Biochemical Switches
Basic Switches (deterministic)

(A) Influence network diagrams
(B) Chemical reaction network diagrams and feedback loops
(C) Numerical solutions of the deterministic kinetics of the networks:
    Horizontal axis is time
    Vertical axis is species concentration

First some arbitrary initial conditions are chosen for AM.
Then the initial conditions of the other networks are chosen in such a way that each trace of each of the other networks retraces exactly one trace of AM.
This can be done for any initial conditions chosen for AM, and indicates the potential of each of the other networks to operate as a simpler switch.

Noise Reduction in Complex Biological Switches

Luca Cardelli\textsuperscript{1,\textbullet,\textbullet,}, Attila Csikász-Nagy\textsuperscript{1,\textbullet,\textbullet,}, Neil Dalchau\textsuperscript{1,\textbullet,\textbullet,}, Miro Tribastone\textsuperscript{1,\textbullet,\textbullet,}, Max Tschalikowski\textsuperscript{1,\textbullet,\textbullet,}

(To appear.)
Basic Switches (stochastic)

Horizontal axes is time
Vertical axes is number of molecules.

(A) Influence networks.
(B) Chemical Master Equation solution: probability distribution, with color (in 10 bands from light = 0 to dark = 1) indicating the probability that at time t there are y molecules of the single indicated species.
(C) Chemical Master Equation solution: mean (solid lines) and standard deviation (color bands) for the species in the network.
(D) Central Limit Approximation solution: mean (solid lines) and standard deviation (color bands) for the species in the network.

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

Horizontal axes are time, vertical axes are number of molecules.

(A) Influence networks.
(B) ODE solutions for comparison
(C) Chemical Master Equation solution: mean (solid lines) and standard deviation (color bands) for the species in the network.
(D) Central Limit Approximation solution: mean (black lines) and standard deviation (color bands) for the species in the network.
Complexity improves overall performance of the cell cycle switch. The performance of different networks was evaluated by calculating the standard deviation of the main molecular states over time. Standard deviations are calculated via numerical integration of the chemical master equation (CME) using the Visual GEC software, and via numerical integration of the central limit approximation (CLA) in Matlab.
Extrinsic Noise

Extrinsic noise can confer robustness to extrinsic noise. Extrinsic noise is introduced by randomly perturbing all the reaction rates (separately but from the same distribution) of each model. (So the total variation in more complex models is actually higher.)

Variations in network behaviour is assessed in comparison to the default parameters, in which all reaction rates are set equal to 1. Network variation is quantified using the summed Wasserstein metric over the whole probability distribution over time.

MI and SI have the same number of species and reactions.
Noise vs. Complexity

- With corresponding initial conditions, all studied networks show the same mean behavior.
- CCr emulating AM is the simplest explanation of the core cell cycle switching function.
- Many other biological switches can be so reduced to an algorithm with well-understood properties.
- On the basis of kinetic similarity of mean behavior, we show variations in noise behavior (both intrinsic and extrinsic).
- Noise tends to decrease with complexity, but this also depends on network structure and not directly on total molecular counts.
Conclusions
Networks are Algorithms

- They are *methods* for achieving a function
  - We need to understand how these methods relate to each other
  - In addition to how and how well they implement function
  - Algorithms can be obfuscated, and nature can obfuscate networks

- Network emulation can be checked *statically*
  - By stoichiometric/reaction-rate (*structural*) properties
  - That is, no need to compare ODE (*functional*) properties
  - For *any* initial conditions and rates of (one of) the networks

- We can efficiently discover emulations
  - Automatic model reduction of large networks
Interpretations of Network Morphisms

- **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 refinement**
  - Emulations are not about abstraction / coarse-graining that preserve behavior, on the contrary, they are about refinement / fine-graining that preserve behavior.
  - They map out successive refinements of simple networks.
Network Emulation Morhpism FAQ

- What guarantees emulation?
  - Reactant morphism + stoichiomorphism: static, state-independent \textit{(structural)} conditions

- How do you find them?
  - Emulation Theorem $\Rightarrow$ they do not depend on initial conditions
  - Change of Rates Theorem $\Rightarrow$ 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 $\Rightarrow$ richness of networks space
  - Approximate emulations exist too

- How useful are they?
  - Establish structural, algorithmic, (non-accidental) \textit{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?