Artificial Biochemistry

Luca Cardelli
Microsoft Research

BIOWIRE 2007
Cambridge 2007-04-03

http://LucaCardelli.name
Intro

- Understanding how cells compute
  - How do signaling networks work?
  - Much is understood, and much is not.

- An unusual computational paradigm
  - By protein interactions (mostly)
  - Is it related to:
    - Electronic circuits?
    - Automata?
    - Process Algebra?

- Why study signaling networks?
  - It’s “just chemistry”, we should be able to cope with it.
    - Simpler than gene networks, neural networks, ants, and bees!
  - Yet non-trivial; general principles and algorithms may apply.

Computing by Stochastic Collectives

- "Collective":
  - A large set of interacting finite state automata:
    - Not quite language automata ("large set")
    - Not quite cellular automata ("interacting" but not on a grid)
    - Not quite process algebra ("collective behavior")
    - Cf. multi-agent systems and swarm intelligence

- "Stochastic":
  - Interactions have rates
    - Not quite discrete (hundreds or thousands of components)
    - Not quite continuous (non-trivial stochastic effects)
    - Not quite hybrid (no "switching" between regimes)

- Very much like biochemistry
  - Which is a large set of stochastically interacting molecules/proteins
  - Are proteins finite state and subject to automata-like transitions?
    - Let’s say they are, at least because:
      - Much of the knowledge being accumulated in Systems Biology is described as state transition diagrams [Kitano].
Compositionality (NOT!)

http://www.expasy.ch/cgi-bin/show_thumbnails.pl
Interacting Automata

Communicating automata: a graphical FSA-like notation for “finite state restriction-free \( \pi \)-calculus processes”. Interacting automata do not even exchange values on communication. The stochastic version has rates on communications, and delays.

“Finite state” means: no composition or restriction inside recursion. Analyzable by standard Markovian techniques, by first computing the “product automaton” to obtain the underlying finite Markov transition system. [Buchholz]
Suppose this is the next interaction (stochastically chosen)

One lonely automaton cannot interact
Interactions in a Population
Interactions in a Population

All-A stable population
Suppose this is the next interaction
Interactions in a Population (2)

All-B stable population

Nondeterministic population behavior ("multistability")
Groupies and Celebrities

Groupie (wants to be like somebody different)

<table>
<thead>
<tr>
<th>a@1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>b@1.0</td>
</tr>
</tbody>
</table>

B

!a

?b

?a

!b

Celebrity (does not want to be like somebody else)

<table>
<thead>
<tr>
<th>a@1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>b@1.0</td>
</tr>
</tbody>
</table>

directive sample 0.1 200

directive plot A(); B()

new a@1.0:chan()

new b@1.0:chan()

let A() = do !a; A() or ?a; B()

and B() = do !b; B() or ?b; A()

run 100 of (A() | B())

A stochastic collective of celebrities:

Unstable because within an A majority, an A has difficulty finding a B to emulate, but the few B's have plenty of A's to emulate, so the majority may switch to B. Leads to deadlock when everybody is in the same state and there is nobody different to emulate.

A stochastic collective of groupies:

Always eventually deadlock

Stable because as soon as a A finds itself in the majority, it is more likely to find somebody in the same state, and hence change, so the majority is weakened.
Both Together

A way to break the deadlocks: Groupies with just a few Celebrities

Many Groupies

A few Celebrities

A tiny bit of “noise” can make a huge difference

Luca Cardelli
2007-04-03

14
Hysteric Groupies

We can get more regular behavior from groupies if they “need more convincing”, or “hysteresis” (history-dependence), to switch states.

Regularity can arise not far from chaos

N.B.: It will not oscillate without doping (noise)

(With doping to break deadlocks)

“regular” oscillation
Semantics of Collective Behavior
These diagrams commute via appropriate maps.

L. Cardelli: “On Process Rate Semantics”
From Automata to Chemistry

<table>
<thead>
<tr>
<th>Automata</th>
<th>Discrete Chemistry</th>
<th>Continuous Chemistry</th>
<th>( \gamma = N_A V )</th>
</tr>
</thead>
<tbody>
<tr>
<td>initial states</td>
<td>initial quantities</td>
<td>initial concentrations</td>
<td>( \gamma = 1 )</td>
</tr>
<tr>
<td>( A \mid A \mid \ldots \mid A )</td>
<td>#( A_0 )</td>
<td>( [A]_0 ) with ( [A]_0 = #A_0 / \gamma )</td>
<td></td>
</tr>
</tbody>
</table>

\[ A \rightarrow^r A' \quad A \rightarrow^k A' \quad \text{with } k = r \]

\[ A+B \rightarrow^r A'+B' \quad A+B \rightarrow^k A'+B' \quad \text{with } k = r\gamma \]

\[ A+A \rightarrow^{2r} A'+A'' \quad A+A \rightarrow^{2k} A'+A'' \quad \text{with } k = r\gamma/2 \]

Think \( \gamma = 1 \)
i.e. \( V = 1/N_A \)
Quantitative Process Semantics

Continuous-state Semantics (Generalized Mass Action)

\[ [X]^* = (\Sigma (Y \in E) \text{Accr}_E(Y,X) [Y]) - \text{Depl}_E(X) [X] \quad \text{for all } X \in E \]

Process Rate Equation

\[ \frac{\partial p_r(p,t)}{\partial t} = \Sigma (q(p,v) p_{r(v,t)} - q(p) p_r(p,t)) \quad \text{for all } p \in \text{States}(E) \]

Process Master Equation

Discrete-state Semantics (Chemical Master Equation)

Continuous Chemistry

Discrete Chemistry

CTMC

Process Algebra

Nondeterministic Semantics

Stochastic Semantics

 Defined over the syntax of processes

Accretion
Depletion
Interactions
Propensity

CTMC

ODE

ODE

Luca Cardelli
2007-04-03

19
Further Examples
Second-order and Zero-order Regime

Second-Order Regime

\[ [S]^* = -r[E][S] \]

directive sample 1000.0

directive plot S(); P(); E()

new a@1.0:chan()

let E() = !a; E()
and S() = ?a; P()
and P() = ()

run (1 of E() | 1000 of S())

Zero-Order Regime

\[ [S]^* \cong -1 \quad \text{(by assuming [ES]^* = 0)} \]

directive sample 1000.0

directive plot S(); P(); E()

new a@1.0:chan()

let E() = !a; delay@1.0; E()
and S() = ?a; P()
and P() = ()

run (1 of E() | 1000 of S())
Ultrasensitivity

Zero-Order Regime
A small E-F inbalance causes a much larger S-P switch.

Second-Order Regime
Second-Order Regime cascade: a signal amplifier (MAPK)
\[ a_{Hi} > 0 \implies c_{Hi} = \text{max} \]

Zero-Order Regime cascade: a signal divider!
\[ a_{Hi} = \text{max} \implies c_{Hi} = \frac{1}{3} \text{max} \]
Nonlinear Transition (NLT)

A = ?c(s); B
B = !c(s); B
A + B → s B + B

[A]⁺ = -s[A][B]
[B]⁺ = s[A][B]

N.B.: needs at least 1 B to "get started".
Two NLTs: Bell Shape

\[ [B]^* = [B][A] - [C] \]

A = ?b; B
B = !b; B ⊕ ?c; C
C = !c; C

A + B →¹ B + B
B + C →¹ C + C

\[ [A]^* = -[A][B] \]
\[ [B]^* = [A][B] - [B][C] \]
\[ [C]^* = [B][C] \]
NLT in a Cycle: Oscillator

A = !a(s):A \oplus ?b(s):B
B = !b(s):B \oplus ?c(s):C
C = !c(s):C \oplus ?a(s):A

A+B \rightarrow s B+B
B+C \rightarrow s C+C
C+A \rightarrow s A+A

\begin{align*}
[A]^* &= -s[A][B]+s[C][A] \\
[B]^* &= -s[B][C]+s[A][B] \\
[C]^* &= -s[C][A]+s[B][C]
\end{align*}
Developing the Use of Process Algebra in the Derivation and Analysis of Mathematical Models of Infectious Disease

R. Norman and C. Shankland
Department of Computing Science and Mathematics, University of Stirling, UK.
{ces,ran}@cs.stir.ac.uk

Abstract. We introduce a series of descriptions of disease spread using the process algebra WS-CCS and compare the derived mean field equations with the traditional ordinary differential equation model. Even the preliminary work presented here brings to light interesting theoretical questions about the “best” way to define the model.
ODE

\[ S = i(t)I \]
\[ I = i(t)I \oplus i(t)I \oplus \tau_r; R \]
\[ R = i(t)R \]

\[ S + I \rightarrow ^{\gamma} I + I \]
\[ I + I \rightarrow ^{\gamma} I + I \]
\[ I \rightarrow ^{r} R \]
\[ R + I \rightarrow ^{\gamma} R + I \]

\[ [S]^* = -\gamma[S][I] \]
\[ [I]^* = \gamma[S][I] - r[I] \]
\[ [R]^* = r[I] \]

"useless" reactions

Differentiating processes!

Automata produce the standard ODEs!

\[
\frac{dS}{dt} = -aIS \\
\frac{dI}{dt} = aIS - bI \\
\frac{dR}{dt} = bI
\]

(the Kermack-McKendrick, or SIR model)!
Conclusions
Conclusions

- **Compositional Models**
  - Accurate (at the “appropriate” abstraction level).
  - Manageable (so we can scale them up by composition).

- **Interacting Automata**
  - Complex global behavior from simple components.
  - Bridging individual and collective behavior.
  - Connections to classical Markov theory, chemical Master Equation, and Rate Equation.

- **An “artificial biochemistry”**
  - A scalable mathematical and computational modeling framework.
  - To investigate “real biochemistry” on a large scale.
Q?