A structured approach for the engineering of biochemical network models, illustrated for signalling pathways

Rainer Breitling\textsuperscript{3}, David Gilbert\textsuperscript{1}, Monika Heiner\textsuperscript{2}, Robin Donaldson\textsuperscript{1}

\textsuperscript{(1)} Bioinformatics Research Centre University of Glasgow, Glasgow, UK  
\textsuperscript{(2)} Computer Science Department, Brandenburg University of Technology, Cottbus, Germany  
\textsuperscript{(3)} Groningen Bioinformatics Centre, University of Groningen, Groningen, Netherlands.

www.brc.dcs.gla.ac.uk/~drg/workshops/ismb08
Tutorial outline

I. Biological introduction  
   *Rainer Breitling*

II. Petri net introduction  
   *Monika Heiner*

III. Biological applications  
   *David Gilbert*

IV. Model checking  
   *Robin Donaldson*

(each 50 min + 10 min break/discussion)
A structured approach …

Part I

Biology

Rainer Breitling

Groningen Bioinformatics Centre, University of Groningen, Groningen, Netherlands.
Outline

• Part 1: Why modelling?
• Part 2: The statistical physics of modelling: $A \rightarrow B$
  (where do differential equations come from?)
• Part 3: Translating biology to mathematics
  (finding the right differential equations)
Biology = Concentrations
Humans think small-scale...  
(the “7 items” rule)

- phone number length  
  (memory constraint)
- optimal team size  
  (manipulation constraint)
- maximum complexity for rational decision making

...but biological systems contain (at least) dozens of relevant interacting components!
Humans think linear...

...but biological systems contain:

• non-linear interaction between components
• positive and negative feedback loops
• complex cross-talk phenomena
Biochemical Pathway Simulation

- What is the best formalism?
- How to deal with lack of information?
- Predictions on what?
- How to collect quantitative measurements in vivo?
- How to manipulate regulatory mechanisms?
The simplest chemical reaction

\[ A \rightarrow B \]

- irreversible, one-molecule reaction
- examples: all sorts of decay processes, e.g. radioactive, fluorescence, activated receptor returning to inactive state
- any metabolic pathway can be described by a combination of processes of this type (including reversible reactions and, in some respects, multi-molecule reactions)
The simplest chemical reaction

A $\rightarrow$ B

various levels of description:

- homogeneous system, large numbers of molecules = ordinary differential equations, **kinetics**
- small numbers of molecules = probabilistic equations, **stochastics**
- spatial heterogeneity = partial differential equations, **diffusion**
- small number of heterogeneously distributed molecules = single-molecule tracking (e.g. cytoskeleton modelling)
Kinetics Description

Main idea: Molecules don’t talk

- Imagine a box containing N molecules.
  How many will decay during time t? $kN$
- Imagine two boxes containing $N/2$ molecules each.
  How many decay? $kN$
- Imagine two boxes containing N molecules each.
  How many decay? $2kN$
- In general:

\[
- \frac{dn(t)}{dt} = \lambda n(t) \quad \Leftrightarrow \quad n(t) = N_0 e^{-\lambda t}
\]

differential equation (ordinary, linear, first-order)

exact solution (in more complex cases replaced by a numerical approximation)
Kinetics Description

If you know the concentration at one time, you can calculate it for any other time! (and this really works)
Probabilistic Description

Main idea: Molecules are isolated entities without memory

Probability of decay of a single molecule in some small time interval:

\[ p_1 = \lambda \Delta t \]

Probability of survival in \( \Delta t \):

\[ p_2 = 1 - p_1 = 1 - \lambda \Delta t \]

Probability of survival for some time \( t \):

\[ p = \lim_{x \to \infty} (1 - \lambda \frac{t}{x})^x = e^{-\lambda t} \]

Transition to large number of molecules:

\[ n(t) = N_0 e^{-\lambda t} \quad \text{or} \]

\[ \frac{dn(t)}{dt} = -\lambda N_0 e^{-\lambda t} = -\lambda n(t) \]
Probabilistic Description – 2

Probability of survival of a single molecule for some time $t$:

$$p = \lim_{x \to \infty} (1 - \lambda \frac{t}{x})^x = e^{-\lambda t}$$

Probability that exactly $x$ molecules survive for some time $t$:

$$p_x = (e^{-\lambda t})^x (1 - e^{-\lambda t})^{N_0 - x} \binom{N_0}{x}$$

Most likely number to survive to time $t$:

$$\max(x \mid p_x) = N_0 e^{-\lambda t}$$
Probabilistic Description – 3

Markov Model (pure death!)

Decay rate:
Probability of decay:
Probability distribution of n surviving molecules at time t:

\[ \Lambda(n, t) = n\lambda \]
\[ p = \Lambda(n, t)dt \]
\[ P(n, t) \]

Description:

Time: \( t \rightarrow \text{wait } dt \rightarrow t+dt \)
Molecules:
\( n \rightarrow \text{no decay } \rightarrow n \)
\( n+1 \rightarrow \text{one decay } \rightarrow n \)

\[ P(n, t + dt) = \]
\[ P(n + 1, t)\Lambda(n + 1, t)dt \]
\[ + P(n, t)[1 - \Lambda(n, t)dt] \]

Final Result (after some calculating): The same as in the previous probabilistic description
Spatial heterogeneity

- concentrations are different in different places, \( n = f(t,x,y,z) \)
- diffusion superimposed on chemical reactions:

\[
\frac{\partial n(t)}{\partial t}_{xyz} = -\lambda n(t)_{xyz} \pm \text{diffusion}
\]

- partial differential equation
Spatial heterogeneity

• one-dimensional case (diffusion only, and conservation of mass)

\[ \frac{\partial n(t, x)}{\partial t} \Delta x = \text{inflow} - \text{outflow} \]

\[ \text{outflow} = -K \frac{\partial n(t, x + \Delta x)}{\partial x} \]

\[ \text{inflow} = -K \frac{\partial n(t, x)}{\partial x} \]
Spatial heterogeneity – 2

\[ \frac{\partial n(t, x)}{\partial t} \frac{\Delta x}{\partial x} = K \frac{\partial n(t, x + \Delta x)}{\partial x} - K \frac{\partial n(t, x)}{\partial x} \]

Transition to differential equation to get diffusion equation:

\[ \frac{\partial n(t, x)}{\partial t} = K \frac{\partial^2 n(t, x)}{\partial x^2} \]

Shorthand for three dimensions:

\[ \frac{\partial n(t, x, y, z)}{\partial t} = K \nabla^2 n(t, x, y, z) \]

Combination with chemical reaction:

\[ \frac{\partial n(t)}{\partial t} = -\lambda n(t) + K \nabla^2 n(t) \]
Summary of Physical Chemistry

- Simple reactions are easy to model accurately
- Kinetic, probabilistic, Markovian approaches lead to the same basic description

\[
\frac{dn(t)}{dt} = -\lambda n(t) \iff n(t) = N_0 e^{-\lambda t}
\]

- Diffusion leads only to slightly more complexity
- Next step: Everything is decay...
Some (Bio)Chemical Conventions

Concentration of Molecule A = \([A]\), usually in units mol/litre (molar)
Rate constant = \(k\), with indices indicating constants for various reactions (\(k_1, k_2\ldots\))

Therefore:

\[ A \rightarrow B \]

\[
\frac{d[A]}{dt} = -\frac{d[B]}{dt} = -k_1[A]
\]
Reversible, Single-Molecule Reaction

A $\leftrightarrow$ B, or A $\rightarrow$ B $\|$ B $\rightarrow$ A, or

Differential equations:

$$\frac{d[A]}{dt} = -k_1[A] + k_2[B]$$

$$\frac{d[B]}{dt} = k_1[A] - k_2[B]$$

Main principle: Partial reactions are independent!
Reversible, single-molecule reaction

Differential Equation:
\[
\frac{d[A]}{dt} = -k_1[A] + k_2[B]
\]
\[
\frac{d[B]}{dt} = k_1[A] - k_2[B]
\]

Equilibrium (=steady-state):
\[
\frac{d[A]_{equi}}{dt} = \frac{d[B]_{equi}}{dt} = 0
\]
\[-k_1[A]_{equi} + k_2[B]_{equi} = 0\]
\[
\frac{[A]_{equi}}{[B]_{equi}} = \frac{k_2}{k_1} = K_{equi}
\]
Irreversible, two-molecule reaction

A + B → C

Differential equations:

\[
\frac{d[A]}{dt} = \frac{d[B]}{dt} = -\frac{d[C]}{dt}
\]

\[
\frac{d[A]}{dt} = -k[A][B] \quad \text{Non-linear!}
\]

Underlying idea: Reaction probability = Combined probability that both [A] and [B] are in a “reactive mood”:

\[
p(AB) = p(A)p(B) = k_1^*[A]k_2^*[B] = k[A][B]
\]
### A simple metabolic pathway

$A \rightarrow B \leftrightarrow C + D$

**Differential equations:**

<table>
<thead>
<tr>
<th>d/dt</th>
<th>decay</th>
<th>forward</th>
<th>reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[A]=$</td>
<td>$-k_1 [A]$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$[B]=$</td>
<td>$+k_1 [A]$</td>
<td>$-k_2 [B]$</td>
<td>$+k_3 [C] [D]$</td>
</tr>
<tr>
<td>$[C]=$</td>
<td></td>
<td>$+k_2 [B]$</td>
<td>$-k_3 [C] [D]$</td>
</tr>
<tr>
<td>$[D]=$</td>
<td></td>
<td>$+k_2 [B]$</td>
<td>$-k_3 [C] [D]$</td>
</tr>
</tbody>
</table>
Metabolic Networks as Bigraphs

A → B ↔ C + D

\[
\begin{align*}
\text{d/dt} & \quad \text{decay} & \quad \text{forward} & \quad \text{reverse} \\
[A] & \quad -k_1 [A] & \quad & \\
[B] & \quad +k_1 [A] & \quad -k_2 [B] & \quad +k_3 [C] [D] \\
[C] & \quad & \quad +k_2 [B] & \quad -k_3 [C] [D] \\
[D] & \quad & \quad +k_2 [B] & \quad -k_3 [C] [D]
\end{align*}
\]
Biological description → bigraph → differential equations
Biological description → bigraph → ODEs

substance A

EC 1.1.1.2

substance B

A

k1

B
Biological description → bigraph → ODEs

substance A

EC 1.1.1.2

substance B

A → k → B

k1

k2

k*

A special case: enzyme reactions

\[ \frac{k_1}{k_2} \quad E + S \xleftrightarrow{\kappa_1} ES \xrightarrow{k_2} E + P \]

In a quasi steady state, we can assume that [ES] is constant. Then:

\[ [ES] = \frac{k_1[E][S]}{k_1 + k_2} \]

If we now define a new constant \( K_m \) (Michaelis constant), we get:

\[ [ES] = \frac{[E][S]}{K_m} \]

\[ K_m = \frac{k_1 + k_2}{k_1} \]
A special case: enzyme reactions

Substituting \([E]\) (free enzyme) by the total enzyme concentration we get:

\[
[ES] = \frac{([E_0] - [ES])[S]}{K_m}
\]

\[
[ES] = [E_0] \frac{1}{1 + \frac{K_m}{[S]}}
\]

Hence, the **reaction rate** is:

\[
V = \frac{d[P]}{dt} = k_2[ES]
\]

\[
\frac{d[P]}{dt} = k_2[E_0] \frac{[S]}{K_m + [S]} = V_{max} \frac{[S]}{K_m + [S]}
\]
A special case: enzyme reactions

Underlying assumptions of the Michaelis-Menten approximation:

• Free diffusion, random collisions
• Irreversible reactions
• Quasi steady state

In cell signaling pathways, all three assumptions will be frequently violated:

• Reactions happen at membranes and on scaffold structures
• Reactions happen close to equilibrium and both reactions have non-zero fluxes
• Enzymes are themselves substrates for other enzymes, concentrations change rapidly, \( \frac{d[ES]}{dt} \approx \frac{d[P]}{dt} \)
Metabolic pathways vs Signalling Pathways
(can you give the mass-action equations?)

Metabolic
(initial substrate)
S
E1 → S'
E2 → S''
E3 → S'''
(final product)

Signalling cascade
Input Signal
X
S1 → P1
S2 → P2
S3 → P3
Output

Classical enzyme-product pathway
Product become enzyme at next stage
Figure 1. Signal transduction pathways in *Paracoccidioides brasiliensis*. Cell adhesion (orange), pheromone response (green), calcium/calcmodulin (pink), cell integrity (blue), high osmotic growth stress response (brown), and TOR (purple) pathways are depicted.
Cell signaling pathways

Neurogenesis

Signaling Pathways: Notch | Shh | β-catenin
Cell signaling pathways
Cell signaling pathways

• Common components:
  – Receptors binding to ligands
    • R(inactive) + L \rightarrow RL(active)

  – Proteins forming complexes
    • P1 + P2 \rightarrow P1P2-complex

  – Proteins acting as enzymes on other proteins (e.g., phosphorylation by kinases)
    • P1 + K \rightarrow P1^* + K
Cell signaling pathways

Nature Reviews | Cancer
Cell signaling pathways

Fig. courtesy of W. Kolch
Cell signaling pathways

Fig. courtesy of W. Kolch
Cell signaling pathways

Fig. courtesy of W. Kolch