

Stochastic Biomodelling

Andrzej Mizera (andrzej.mizera@uni.lu)
Faculty of Science, Technology and Communication
University of Luxembourg



Course content

1. Introduction

- Stochasticity in biological processes
- Deterministic vs stochastic biomodelling

2. Prerequisite:

- Crash course on probability theory

3. Stochastic modelling of chemical kinetics: the chemical master equation (CME)

4. Stochastic simulation of the CME – Gillespie's direct method algorithm

➔ 5. Practicals:

- Implementing the Gillespie's algorithm in MATLAB and investigating its characteristics on various biochemical systems
- Comparing the obtained simulation results with the solutions in the deterministic formulation

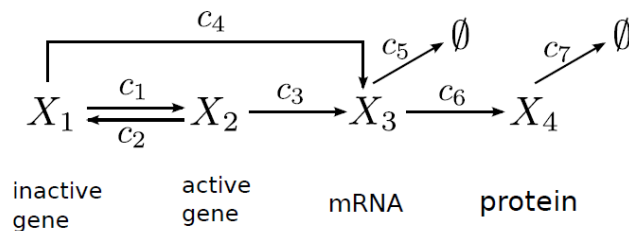
Some words on what is the role of mathematical modelling in Systems Biology

Stochastic Biomodelling

Practicals:

- Implementing the Gillespie's algorithm in MATLAB and investigating its characteristics on various biochemical systems
- Comparing the obtained simulation results with the solutions in the deterministic formulation

Task 1: Gillespie's algorithm



j	R_j	$a_j(\mathbf{X})$	\mathbf{v}_j
1	$X_1 \rightarrow X_2$	$c_1 X_1$	$(-1, 1, 0, 0)$
2	$X_2 \rightarrow X_1$	$c_2 X_2$	$(1, -1, 0, 0)$
3	$X_2 \rightarrow X_2 + X_3$	$c_3 X_2$	$(0, 0, 1, 0)$
4	$X_1 \rightarrow X_1 + X_3$	$c_4 X_1$	$(0, 0, 1, 0)$
5	$X_3 \rightarrow \emptyset$	$c_5 X_3$	$(0, 0, -1, 0)$
6	$X_3 \rightarrow X_3 + X_4$	$c_6 X_3$	$(0, 0, 0, 1)$
7	$X_4 \rightarrow \emptyset$	$c_7 X_4$	$(0, 0, 0, -1)$

$$c_1 = 0.4, c_2 = 0.27, c_3 = 0.7, c_4 = 0.05, c_5 = 0.65, c_6 = 0.75, c_7 = 0.5$$

Start state: $(1, 0, 0, 0)$

- See the MATLAB m-file: `gene_expression.m`

Task 2: Gillespie's algorithm

Consider the following system of reactions:



We assume that the volume of the system is $V = 1$. Let the initial state of the system be $\mathbf{X}(t=0) = (\#A = 15, \#B = 5, \#C = 3)$.

- Implement the Gillespie algorithm to model the time evolution of the system during 1000 time units. See the MATLAB m-file: ABC.m.

Task 2: Gillespie's algorithm

- How will the program change, if we modify the first reaction to $A + A \rightarrow C$?

We consider the volume of the system to be $V=1$. k_1 , k_2 , k_3 , and k_4 are **kinetic rate constants**. They need to be transformed into the **stochastic rate constants**. However, the original set of equations is such that in fact the values for the stochastic and the kinetic rate constants are the same. The situation changes, when the first reaction is replaced with $A + A \rightarrow C$. In this case, the value for the corresponding stochastic rate constant for the first reaction is $2 \cdot k_1$. The values for the stochastic rate constants for the remaining reactions remain the same as the values for the corresponding kinetic rate constants. In consequence, the MATLAB code needs to be changed as follows (**red text**):

1. $\text{nu} = [-2 \ 0 \ 1; 1 \ 1 \ -1; -1 \ 1 \ 0; 0 \ 0 \ -1];$
2. $\text{ais} = [2 * c(1) * x(1,i-1) * (x(1,i-1)-1) / 2 \ c(2) * x(3,i-1) \ c(3) * x(1,i-1) \ c(4) * x(3,i-1)];$

Stochastic Biomodelling

Some words on the role of mathematical modelling
in Systems Biology

A model

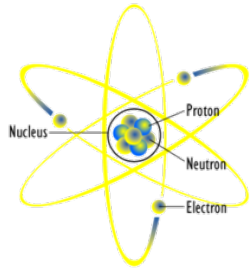
- What is a model?
 - A simplified view of complex reality
 - A (partial) view of the reality
 - An abstraction of the reality
 - A representation of the (supposedly) main features of the reality, including the connections among them
 - For a given object of study, many models may be given, possibly focusing on different features/aspects of the object
 - In a broad sense, an abstract representation of objects or processes that explains features of these objects or processes
- What a model is not
 - A model is not the reality!
 - A model is not certain!

A model

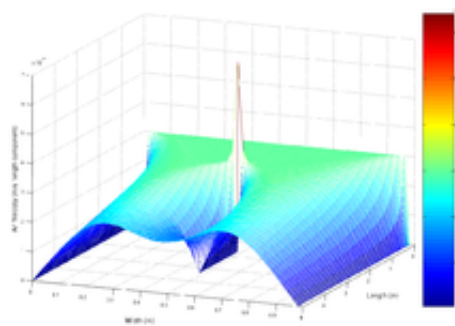
“All models are wrong, but some are useful.”

Box, G.E.P., Robustness in the strategy of scientific model building, in Robustness in Statistics, R.L. Launer and G.N. Wilkinson, Editors. 1979, Academic Press: New York.

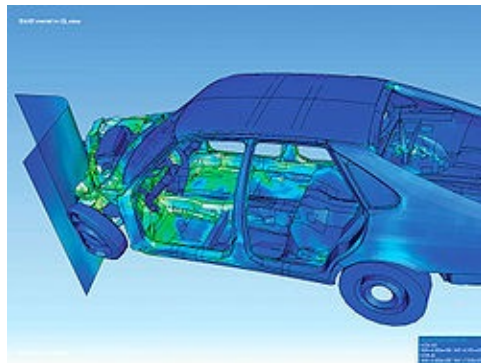
A model



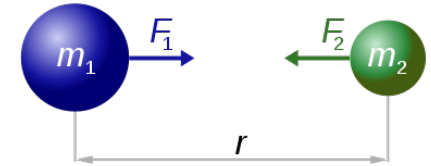
Generic atomic planetary model (E. Rutherford, N. Bohr)



Navier–Stokes differential equations used to simulate airflow around an obstruction.

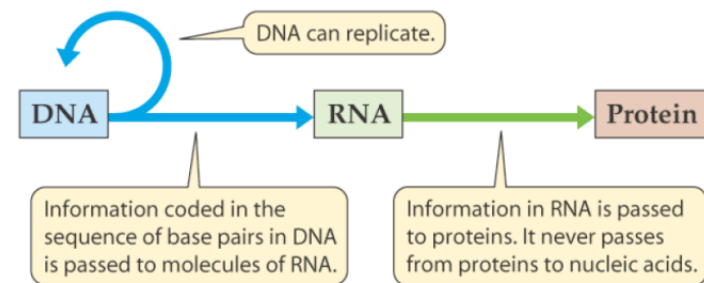


Car crash model



$$F_1 = F_2 = G \frac{m_1 \times m_2}{r^2}$$

Newton's law of universal gravitation



Central dogma of molecular biology

Mathematical models

(based on [J. Gunawardena. Models in biology: ‘accurate descriptions of our pathetic thinking’. *BMC Biology*, **12**:29, 2014].)

- Biology rests on **physics**.
- Physicists have worked out the **fundamental laws** governing the behaviour of matter at the length scales and timescales relevant to biology.

‘all things are made of atoms and ... everything that living things do can be understood in terms of the jiggling and wiggling of atoms’

R. P. Feynman, R. B. Leighton, and M. Sands. *The Feynman Lectures on Physics*. Vol. 1. Addison-Wesley, Reading, MA, USA, 1963. (Chapter 3-3)

- If our **assumptions** can be grounded in physics, then our model should be **predictive**, in the sense that they are **not subject to falsification** and we can be confident of the conclusions drawn.
- Models based on fundamental physical laws seem to be **objective descriptions of reality**, which can be interrogated to understand reality.

Mathematical models

Can we ground biological models on fundamental physics?

- The Schrödinger equation is too complicated to solve directly even for a single protein!
- On the move from atoms to molecules, i.e., upward the biological scale, we leave the realm of physics and enter the realm of biochemistry
 - Chemical reaction: $A+B \rightarrow C$
 - How to study the reaction quantitatively?
 - Use the **mass action law**: $d[C]/dt = k \cdot [A] \cdot [B]$
 - **But**: a chemist might point out that, e.g., the reaction $H_2 + Br_2 \rightarrow 2 HBr$ has a rate given by $d[HBr]/dt = (k_1 \cdot [H_2] \cdot [Br_2]^{3/2}) / ([Br_2] + k_2 \cdot [HBr])$
 - **Mass action is not physics or chemistry**. It is **phenomenology**: *a mathematical formulation, which may account for observed behaviour but which is not based on fundamental laws!*

Mathematical models

- **What are the consequences of phenomenology?**

- The models are no longer objective descriptions of reality.
- The models can no longer be considered predictive, in the sense of physics or even of molecular dynamics.
- The models are **falsifiable**.

- **What is the role of such models?**

- If the model is not going to be predictive and if we are not certain of its assumptions, then there is no justification for the model other than as a test of its assumptions.

‘They are meant to expose assumptions, define expectations and help us to devise new tests’

James Black, pharmacologist, Nobel Prize in Physiology or Medicine, 1988

Mathematical models

Summary

After [J. Gunawardena. Models in biology: ‘accurate descriptions of our pathetic thinking’. *BMC Biology* **12**:29, 2014].

- A mathematical model: a logical machine for converting assumptions into conclusions.
- If the model is correct and we believe its assumptions then we must, as a matter of logic, believe its conclusions.
- This logical guarantee allows a modeler, in principle, to navigate with confidence far from the assumptions, perhaps much further than intuition might allow and reach surprising conclusions.
- But the certainty is always relative to the assumptions.

Do we believe our assumptions?

Mathematical models

Summary (continued)

After [J. Gunawardena. Models in biology: ‘accurate descriptions of our pathetic thinking’. *BMC Biology* **12**:29, 2014].

- We believe fundamental physics on which biology rests. We can deduce many things from physics but not all. **This leaves us in the hands of phenomenology.**
- There is nothing wrong with that but we should not fool ourselves that our models are objective and predictive, in the sense of fundamental physics.

Building a model

- Start building the model with dividing the world into **3 parts**
 1. Things whose effects are neglected – ignore them in the model
 2. Things that affect the model but whose behaviour the model is not designed to study (e.g., external variables that are considered as parameters, input, or independent variables)
 3. Things the model is designed to study the behaviour of (internal <or dependent> variables of the model)
- Deciding what to model and what not to model **is difficult!**
 - Wrong things neglected: the model is no good
 - Too much included: hopelessly complex model
 - Choose the internal variables wrongly: the model will not capture its target
- How general should the model be?

Building a model (continued)

- Formulate the problem
 - What do you want to know/understand?
- Outline the model
 - Divide the world into the three categories given in the previous slide
 - Specify the interrelations among the variables
 - A challenging stage: not always clear what is important and what is not
- How useful is the resulting model?
 - Can you obtain the needed data for the model?
 - Can the available data be used in the model?
 - If not, reformulate the model and perhaps even the problem
 - Note: sometimes a model needs no data; all included in the assumptions

Building a model (continued)

- Test the model
 - Use the model to make predictions that can be checked against known data or common-sense
 - Note: sometimes difficult/expensive/impossible to test a model
- Use the model
 - **Careful** that the situations where it is used were captured in the modelling process

Every model makes certain assumptions that need to be satisfied in order for the model to be valid/useful. Make sure all the assumptions are satisfied or can be justified in the context of the study. Other wise the model is of no use! Critical in engineering!
 - Every application of the model is a test for the model

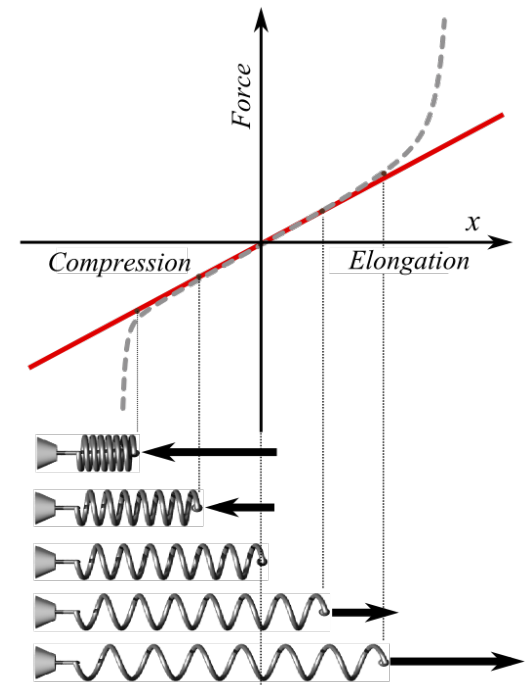


Image source: Wikipedia
February 2016

Building a model (continued)

