A Teaching Resource for Complex Systems, Machine Learning and Computational Biology

Soumya Banerjee, Joyeeta Ghose

University of Oxford

soumya.banerjee@maths.ox.ac.uk

Abstract

In this work, we present all teaching materials related to complex systems, machine learning and computational biology.

Introduction

This work presents a collection of teaching materials related to complex systems, machine learning, computational biology and computational immunology.

1) My class presentation for November 2010 on the generality of the SIR (Susceptible-Infected-Recovered) model.

2) Instructions and more material:
Download the Berkeley Madonna solver.
After downloading it, you can run everything in "Demo Mode". After installing Berkeley Madonna, simply clicking on the .mmd files (file1, file2)
will open up the graphical user interface. Simply click on "Run" and go to the window (from Window in the menu bar) to see the output. The
sliders with parameters are already set up.

If you want to include more parameters in them, simply click on the slider and it will bring up another window where you can add new parameters.
The knockout mice TCL.mmd file uses the virus data from the file
igm_diamond_mice_knockout. If you want to see the solver fit the model to data,
go to Parameters -> Curve Fit. Click the "Import Dataset" button and select the virus data file. You can try changing the initial guesses of all the
parameters and then click on OK and the solver will get into action and do the fitting. The user just needs to give the equations and the initial values of the populations.

Lastly, there is a paper on spreading of ideas which also uses a SIR model.

3) I gave another lecture on mathematical modelling to a non-mathematical audience.
I started with a very simple SIR model, and explained the model and the effect various parameters have on the behaviour of the model.
This is very easy to do in Berkeley Madonna, where you can create sliders and visually investigate the effect of increasing or decreasing some parameter.
Here is a Berkeley Madonna model file to play around with these parameters (simple model).

I then generated some data on infections (data file) and then demonstrated how to fit the model to the data using the "curve fitter" option in Berkeley Madonna.
The curve fitter option allows you to match your model predictions to data and thereby estimate (find) the values of your model parameters.
Here is a Berkeley Madonna file that fits a simple SIR model to infection data (fitting model to data).

I then pointed out that this fitting procedure is usually sensitive to the initial guesses you give to the curve fitter, i.e. choosing different values for model parameters
can give you very different results. This usually complicates the fitting process which means you may have to more sophisticated programs written in MATLAB, like the ones here (sophisticated fitting programs).
Here is a Berkeley Madonna file that gives different initial guesses and ends up producing a bad fit to the data (fitting model to data issues).

4) A Berkeley Madonna file for the Lotka-Volterra model (file) (Berkeley Madonna model file, more advanced version with sliders). Copy the contents of this file into Berkeley Madonna and hit Run.

5) Berkeley Madonna files implementing the basic target cell limited model (file1, file2)

6) Berkeley Madonna files for implementing an immune response model with T-cell responses (file1, file2) and associated paper

7) In another recent lecture I explained the law of mass-action and showed demo examples using NetLOGO example models. The law of mass action underlies many ODE models in biology.

8) I also taught the basics of biology and immunology using materials here and NetLOGO models of DNA protein synthesis

9) NetLOGO model of ants collecting food and the B-Z reaction (reaction diffusion reaction which can produce persistent patterns in media)

10) Teaching resource for students and the general public on a computational framework of the value of information in origin of life questions.

Information plays a critical role in complex biological systems. This paper is a teaching resource that explains the role of information processing in questions around the origin of life and suggests how computational simulations may yield insights into questions related to the origin of life. (teaching resource) Also available from link here

11) Code and teaching resource for using open data to model complex collaboration networks (link) (other projects on bitbucket)

12) An introduction to oscillators and circuits in biological systems and bifurcation diagrams (paper) and using xppauto to generate phase plots and bifurcation diagrams

13) Agent based modelling in python (mesa, pyabm, more resources and tutorials)

14) Agent based models of the immune system

Download the CyCells Agent-Based Model (ABM) simulator (by Christy Warrender and Drew Levin)
(CyCells) http://sourceforge.net/projects/cycells/
Here is a sample definition file and a sample init file. To execute simply download cycells and save the
two files in the same folder and type the following in the command line:
./CyCells -d mys29.def -i mys29.init -t 10000
The -t specifies the number of time steps that it has to run.
The simulation produces a file (test.history) which has the data from the run.

Play around with different kinds of immune system models and immune system cells (by changing the .def and .init files) and analyze the output in the test.history files

15) Stochastic simulations using StochPy in Python (link to demo examples)

16) Machine learning and data science basics coupled to analysis of complex systems (code, open source data science projects)

17) Machine learning and data science

   Machine learning resources (link, link to playlists)

   Data science tools (link)

18) A full-fledged course will include basic programming (using scripting languages), basic biostatistics, machine learning, bioinformatics, non-linear dynamical systems, complex systems and data science.

19) More teaching resources

   Complex Systems

   Complexity Explorer

   A course on practical applications of mathematics

   Course on PDEs and ODEs (link, link, teaching resources, annotated textbooks and resources on ODEs)

   Course on probability and statistics

   SAGANet on origins of life and astrobiology

   Play with Conway's Game of Life simulations (link)

University of Vermont Complex Systems Courses ([link](#))

Introduction to Natural Computation ([link](#))

Origins of life teaching resource
The associated material is available from https://doi.org/10.5281/zenodo.885371

References

[1] A spatial model of the efficiency of T cell search in the influenza-infected lung L Drew, F Stephanie, B Soumya, C Candice, C Judy, M Melanie. Journal of Theoretical Biology 398 (7), 52-63

[2] Banerjee, S., Guedj, J., Ribeiro, R. M., Moses, M., & Perelson, A. S. 2016. Estimating biologically relevant parameters under uncertainty for experimental within-host murine West Nile virus infection. Journal of the Royal Society Interface, 13(117), 20160130-.http://doi.org/10.1098/rsif.2016.0130

[3] Science and technology consortia in US biomedical research: A paradigm shift in response to unsustainable academic growth. Curt Balch, Hugo Arias-Pulido, Soumya Banerjee, Alex K. Lancaster. BioEssays 37 (2), 119-122

[4] Soumya Banerjee and Joshua Hecker. A Multi-Agent System Approach to Load-Balancing and Resource Allocation for Distributed Computing, arXiv preprint arXiv:1509.06420, 2015

[5] Soumya Banerjee and Melanie Moses. Immune System Inspired Strategies for Distributed Systems. arXiv preprint arXiv:1008.2799, 2010

[6] Soumya Banerjee and Melanie Moses. Scale Invariance of Immune System Response Rates and Times: Perspectives on Immune System Architecture and Implications for Artificial Immune Systems. Swarm Intelligence 4, 301–318 (2010). URL http://www.springerlink.com/content/w67714j724486533l/

[7] Soumya Banerjee, A Roadmap for a Computational Theory of the Value of Information in Origin of Life Questions, Interdisciplinary Description of Complex Systems, 14(3), 314-321, 2016

[8] Soumya Banerjee, Jeremie Guedj, Ruy Ribeiro, Melanie Moses, Alan Perelson (2016). Estimating biologically relevant parameters under uncertainty for experimental within-host murine West Nile virus infection. Journal of the Royal Society Interface, 13(117), 20160130-. http://doi.org/10.1098/rsif.2016.0130

[9] Soumya Banerjee. 2009. An Immune System Inspired Approach to Automated Program Verification, arXiv preprint arXiv:0905.2649, 2009

[10] Soumya Banerjee. 2013. Scaling in the immune system, PhD Thesis, University of New Mexico (2013)

[11] Soumya Banerjee. A Biologically Inspired Model of Distributed Online Communication Supporting Efficient Search and Diffusion of Innovation. Interdisciplinary Description of Complex Systems 14 (1), 10-22, 2016

[12] Soumya Banerjee. A computational technique to estimate within-host productively infected cell lifetimes in emerging viral infections. PeerJ Preprints 4 (e2621v2) 2017

[13] Soumya Banerjee. An artificial immune system approach to automated program verification: Towards a theory of undecidability in biological computing. PeerJ Preprints 5 (e2690v1) 2017

[14] Soumya Banerjee. An artificial immune system approach to automated program verification: Towards a theory of undecidability in biological computing. PeerJ Preprints 5 (e2690v1) 2017

[15] Soumya Banerjee. An Immune System Inspired Theory for Crime and Violence in Cities. Interdisciplinary Description of Complex Systems, 15(2):133-143, 2017

[16] Soumya Banerjee. Analysis of a Planetary Scale Scientific Collaboration Dataset Reveals Novel Patterns. arXiv preprint arXiv:1509.07313, 2015

[17] Soumya Banerjee. Optimal strategies for virus propagation. arXiv preprint arXiv:1512.00844, 2015

[18] Soumya Banerjee, Melanie Moses, Alan Perelson (2017). Modelling the effects of phylogeny and body size on within-host pathogen replication and immune response, 14:20170479. http://doi.org/10.1098/rsif.2017.0479