Documentation and User Manual for BioSIMI modeling framework

Miroslav Gasparek, Imperial College London
Professor Richard M. Murray, California Institute of Technology
Vipul Singhal, California Institute of Technology

30 August 2017

1. Introduction

The purpose of this document is to provide documentation and user guide to BioSIMI, which stands for **Biomolecular Subsystems Interconnection Modeling Instrument**. The aim of this MATLAB-based framework is to provide tool for rapid development and analysis of complex biomolecular circuits and systems arising from interconnection of modular subsystems, while accounting for phenomena related to interconnection of biomolecular subsystems, such as retroactivity and competition for resources. By providing the tools for fast assembly and analysis of interconnected synthetic biomolecular circuits with emphasis on analysis of Input/Output behavior of the whole system, BioSIMI aims to contribute to effort in building of genetically-engineered non-living artifical cells.

BioSIMI is based on MATLAB SimBiology toolbox which facilitates modeling and simulations of chemical reactions between species, which form basics of every biomolecular circuit. However, SimBiology module does not enable sufficiently fast modeling of more complex biomolecular systems, such as those that would be used in the design process of an artificial cell.

BioSIMI also aims to be compatible with a set of tools for cell-free synthetic biology developed by Professor Murray's group and other groups over last 5-10 years [2], especially with TX-TL modeling toolbox [citation needed] and with TX-TL method for implementing circuits and pathways *in vitro*

To conclude I hope that you will find this tool interesting and functional. If you have any suggestions, notes, and comments, please do not hesitate to contact authors at `miroslav.gasparek15@imperial.ac.uk`.

2. What is artificial cell?

The focus of the module is to enable fast analysis of behavior of artificial cells. So what is an artificial cell? This is a difficult question and there is not a single answer to it. One way to answer it is to draw analogy between living and artificial cell by implementing some of the defining features of living cells in artificially created biomolecular systems - which we will call, (very courageously!) artificial cells. These features of living cells include [citation needed]:

- *Homeostasis*
- *Structural organization*
- *Metabolism*
- *Growth*
- *Adaptation to the surrounding environment*
- *Ability to respond to environmental stimuli*
- *Reproduction*

Minimalistic model of an artificial cell can be represented by a set of biomolecular systems.

3. How to use BioSIMI

The BioSIMI modeling framework provides functions and subroutines allowing user to create SimBiology objects (species, reactions, rules...) using a set of pre-defined subsystems required for functioning of a cell. These subsystems can include transport systems (such as diffusion through the cell membrane), signaling pathways

(Double-Phosphorylation signaling pathway, MAPK pathway) or computational/regulatory pathways (Incoherent Feed-Forward Loop, Genetic Toggle Switch) or metabolic systems. Subsystems (belonging to the class of pre-defined types) can be created, added to the model, and connected in several lines of commands in MATLAB.

1. Creation of a mathematical model of simple artificial cell consisting of set of subsystems performing functions that are characteristic to living cell. These include sensing, signaling, regulation/computation, and the effect of metabolism.
2. Analytical and computational investigation of the artificial cell model, as the system accepting external input (i.e. extrinsic signaling molecule) and producing the corresponding output (i.e. specific protein or reporter protein).
3. Optimization of intrinsic processes in the sensing, signaling and regulatory subsystems so that desired system behavior (in terms of required transfer function from input to output, robustness, and stability) is achieved.

Modeling, analysis and simulations are specific to each of the objectives:
1. Creation of a mathematical model of simple artificial cell will encompass selection of appropriate cell subsystems enabling it to perform desired function of the artificial cell. This includes sensing of extrinsic signaling molecule, membrane receptor binding of the signaling molecule, signaling pathway from the artificial cell membrane to intracellular DNA circuit, binding to promoter region and transcription into mRNA, translation into desired protein at ribosome binding sites, and finally, export of the protein out of the artificial cell.
2. Analytical and computational investigation of the artificial cell model will include mathematical description of the subsystems and their mutual interactions, using established methods, such as those described in [3]. System will be first analyzed theoretically in general terms, and as many conclusions about its behavior will be drawn through approaches described in [3], [6], [7]. The goal of such analysis is to find the output of the system (concentration of the target protein or reporter) as a function of the input (extrinsic signaling molecule) and internal system parameters.
Subsequently, specific systems will be simulated in silico and information drawn from the theoretical analysis will be verified.
3. Optimization of intrinsic processes in the sensing, signaling and regulatory subsystems will require theoretical and computational analysis of designed artificial cell and manipulation of the parameters of sensing, signaling and regulatory subsystems, so that their interconnection produces required output corresponding to appropriate input. Using methods described in [3] and mathematical framework provided in [2], this optimization also aims to consider and address realistic imperfections in the artificial cell model, such as retroactivity, random noise, and competition for resources if multiple genetic circuits are present in the artificial cells. This should result into design of robust and stable biomolecular system.

The product of successful research project will be a realistic model of simple artificial cell composed of fine-tuned functional subsystems providing measurable response to the input signal, based on the features of desired output. It will also demonstrate optimal robustness to the retroactivity, effect of limited resources, and perturbation.

3. Subroutines

1. Creation of a mathematical model of simple artificial cell

Initially, the model of artificial cell will be treated as deterministic single input-single output model with dynamics defined by chemical reactions between individual parts and subsystems modelled by deterministic reaction rate equations.

Sensing and signal transduction: Binding of the extrinsic signaling molecule (input) to the membrane-bound receptor of artificial cell and subsequent signal transduction to the genetic circuit in the nucleus will be lumped together and modelled as MAPK cascade [3], [5]. Dynamics of the process will be described by set of ordinary differential equations.

Regulation/Computation: Activated output of intracellular signaling pathway binds to promoter/operator region of the DNA and regulates the transcription of DNA into mRNA and thus subsequent production of the protein. Implemented gene circuits can range from simple activator action, through various logic functions and oscillators.

Export: Export of the substances will be modelled through exocytosis of the produced protein, which would represent the final output. Alternatively, if this would not be feasible, output can be measured as concentration of the suitable reporter protein.

2. Analytical and computational investigation of the artificial cell model

The proposed model will be described by reaction rate equations with parameters chosen from sources such as *BioNumbers* [8] to represent biologically plausible scenario. Variety of methods from fields of nonlinear dynamics (such as stability analysis and timescale separation [3], [6]) and control theory (Laplace transform for linear systems and describing functions for nonlinear systems [7]). In silico simulations will be performed in Matlab (with which I have most experience) or another appropriate software as suggested.

3. Optimization of intrinsic processes in the sensing, signaling and regulatory subsystems

Analysis of perturbation due to parameter sensitivity will be performed and its effect on transfer function of the system will be investigated.

In terms of interconnection of components, the effects of retroactivity in sensing and signaling MAPK cascade and in gene circuits will be investigated and design approaches such as use of high-gain feedback will be employed [3].

Competition for the resources, such as ATP required for phosphorylation or mRNA will be modelled as in [3]. Finally, if time constraints will permit, analysis of random processes in frequency domain for deterministic linear processes in sensing, signaling and regulatory/computational subsystems will be performed and their overall effect on the output signal will be examined, along with options of minimizing detrimental effect of the noise (such as low-pass filtering) [3]. However, I am aware that I will need to acquire competency in modeling of the processes described in this section, which is something that I would like to do before the start of my research placement.

Accomplishment of objectives 1 and 2 should take approximately four weeks, while work towards accomplishment of objective 3 will be distributed throughout following six weeks of the internship.

From my perspective, the most difficult steps include those leading to accomplishment of objective 3, especially perturbation analysis, retroactivity attenuation, and stochastic modeling. I have been introduced into these topics through self-studying of [3], [6] and [7]. Although I have not had experiences with simulation of such processes, I am willing to acquire such experiences through practicing exercises in [3] and other

recommended sources, so that I am prepared for such tasks adequately prior to the start of my internship. The resources required to conduct the internship include mainly access to commercially available, high-performance computer required for computational simulations. Although I do not expect to take significant part in the wet lab research, I am absolutely open to such opportunity, especially considering that I will have taken two Wet Labs courses before the proposed internship start date.

I expect to be under supervision of Professor Richard M. Murray, Professor of Control & Dynamical Systems and Bioengineering and under supervision of graduate and/or postgraduate members of his group working on similar tasks.

4. Subsystems

This work plan outlines schedule that should lead to accomplishment of the objectives. Plan is only provisional and it might change based on the progress and supervisor's feedback.

Week 1: Introductory week. Design of the simple deterministic model of artificial cell with focus on sensing and signaling subsystem. Mathematical analysis through methods of control theory and nonlinear dynamics. Preliminary computational simulations.

Weeks 2-3: Computational simulations of design of sensing and signaling pathway. Verification of the mathematical results obtained in Week 1. Input/output analysis of the sensing and signaling pathway model. Implementation of required modifications based on results of simulations. Preliminary perturbation analysis of the sensing and signaling pathway.
Writing and submission of Progress report 1.

Week 4-5: Design of simple deterministic model of the intracellular gene circuit. Mathematical analysis and computational simulations of design of gene circuit. Verification of the mathematical results obtained in Week 1. Input/output analysis of the regulatory subsystem. Implementation of required modifications based on results of simulations. Perturbation analysis of the transcriptional and translational subsystem.

Weeks 6-7: Effect of retroactivity in sensing & signaling pathway and in transcriptional and translational circuits. Implementation of mechanisms leading to retroactivity attenuation.

Week 8: Influence of competition for limited and shared resources in the artificial cell. Analytical and computational optimization of amount of resources required for proper function of the artificial cell.
Writing and submission of Progress report 2.

Weeks 8-9: Analysis and modeling of stochastic effects in transcription, translation and retroactivity attenuation. Analysis of random processes in the frequency domain. If time permits, more thorough analysis of the exocytosis out of artificial cell and intercellular signaling. Final report writing.

Week 10: Finalization of the final report and presentation preparation. Presentation of the achieved results.

References

[1] Elani, Y. Construction of membrane-bound artificial cells using microfluidics: a new frontier in bottom-up synthetic biology. *Biochemical Society Transactions,* 44(3), pp.723-730. 2016.

[2] Murray, R. (2017). *Genetically-Programmed Artificial Cells and Multi-Cellular Machines.* 1st ed. [ebook] Available at: `http://www.cds.caltech.edu/~murray/wiki/index.php?title=SURF_2017:_Genetically`

−Programmed_Artificial_Cells_and_Multi−Cellular_Machines&redirect=no[Accessed   6 Jun. 2017].

[3] Del Vecchio, D. and Murray, R. (2015). *Biomolecular feedback systems.* 1st ed. Princeton: Princeton University Press.

[4] Baldwin, G., Dickinson, R. and Kitney, R. (2015). *Synthetic Biology - A Primer.* 1st ed. London: Imperial College Press.

[5] Alberts, B. (2014). *Essential cell biology.* 3rd ed. New York: Garland Science.

[6] Strogatz, S. (2015). *Nonlinear dynamics and chaos.* 2nd ed. Boulder: Westview.

[7] Slotine, J. and Li, W. (2005). *Applied nonlinear control.* 1st ed. Taipei: Pearson education Taiwan.

[8] Bionumbers.hms.harvard.edu. (2017). *BioNumbers - The Database of Useful Biological Numbers.* [online] Available at: `http://bionumbers.hms.harvard.edu/` [Accessed 6 Jun. 2017].