BIO/AMS 332: Computational Modeling of Physiological Systems — Project 2

Instructions

This project is to be completed using MATLAB. MATLAB is a powerful environment for mathematical modeling — allowing quite sophisticated models to be implemented without requiring a lot of knowledge about programming. In this project, you will learn the basics of how to use MATLAB, using as examples some of the models that we have discussed in class. MATLAB is available on most SINC site computers, and for installation on your own computer via SoftWeb. Please read through the "Introduction to Matlab" appendix in the class notes for a basic introduction to MATLAB commands.

The solutions to the problems should be presented as a computer-formatted document with answers given in full sentences. All figures should be integrated into the document, and should include a brief descriptive caption. All plots should include a title and axes labels (with units), and a legend where appropriate. An appendix including all Matlab scripts should be included. A final copy of the report should be submitted electronically as a single PDF document through BlackBoard.

All MATLAB work presented in your answers must be consistent with your submitted M-files. Separate M-files should be used for each part of the project; for the various questions within a part, you may either create a single M-file that solves all the sub-problems or split your work among multiple M-files. Regardless, a single "zip" file containing all Matlab scripts (M-file) should be submitted at the same time as the project PDF document. The zip file should contain the executable M-files (as unformatted text); this is NOT a substitute for including the code as an appendix to your report.

You are encouraged to work in groups, both in discussing how to implement the problems in MATLAB and in the interpretation of your results, and to use online resources. However, simply copying the work of a fellow student (or directly copying any material found online) is forbidden. All projects must include a statement that explains all resources that were used in the completion of all aspects of the project and lists all collaborators involved. A few examples are given below:

- *I worked on implementing the Matlab code with Tom Hanks and Brad Pitt; Tom really knew how to solve the problems, and Brad and I largely followed his guidance. However, I wrote the program myself, based on Tom's help. Tom, Brad and I also discussed the answers to the questions together, using Wikipedia for guidance on question 2, but we all made equal contributions. The written work is all my own. Signed: Sam Jackson*

- *I did the project completely on my own; I did not discuss any of the problems with anyone in the class, and did not share any of my work with others. However, I did make extensive use of both Wikipedia and MatlabCentral. Signed: Katy Perry*

Part A: Determinstic model of the *cro–cI* genetic network.

When a bacteria is infected by the bacteriophage lambda (λ), two fates are possible. First, the virus may replicate many progeny within the bacteria, ultimately resulting in the bursting of the bacteria to release new phage; this is known as the lytic pathway, or lysis. However, in some cases, the phage DNA remains in the bacteria without making new phage; the phage DNA is replicated along with the bacterial genome during cell division, and thus all descendents of the infected cell also carry the phage DNA. This state is known as lysogeny; the lysogenic state is generally stable, meaning that all progeny of a lysogenic bacteria remain in the lysogenic state indefinitely. However, exposure to stress (such as radiation) can convert a cell in the lysogenic state into the lytic phase.

The decision to enter lysis or lysogeny is based on a pair of mutually repressive transcription factors, cI and cro. When levels of cI are high and levels of cro are low, an infected bacteria will be in the lysogenic phase; when levels of cI are low and levels of cro are high, the lytic phase is prefered. A simple model of this process can be given by the equations:

$$\frac{d[cI_{prot}]}{dt} = \omega_{cI}[cI_{rna}] - \chi_{cI,prot}[cI_{prot}] \tag{1}$$

$$\frac{d[cI_{rna}]}{dt} = \mu_{cI}\left(1 - \frac{[cro_{prot}]^2}{K_{cro,1/2}^2 + [cro_{prot}]^2}\right) - \chi_{cI,rna}[cI_{rna}] \tag{2}$$

$$\frac{d[cro_{prot}]}{dt} = \omega_{cro}[cro_{rna}] - \chi_{cro,prot}[cro_{prot}] \tag{3}$$

$$\frac{d[cro_{rna}]}{dt} = \mu_{cro}\left(1 - \frac{[cI_{prot}]^2}{K_{cI,1/2}^2 + [cI_{prot}]^2}\right) - \chi_{cro,rna}[cro_{rna}] \tag{4}$$

1. Using the autoregulatory gene model of Project 1 as a base, implement this model in MATLAB, using the following constants:

$$\chi_{cI,rna} = \chi_{cI,prot} = 1.2\ s^{-1} \qquad \chi_{cro,rna} = \chi_{cro,prot} = 0.8\ s^{-1}$$

$$\omega_{cI} = \mu_{cI} = \omega_{cro} = \mu_{cro} = 50\ s^{-1} \qquad K_{cI,1/2} = K_{cro,1/2} = 10\ \text{molecules/cell} \cdot s^{-1}$$

   Use a time step of 0.01 seconds, and choose an appropriate simulation length based on what you observe. Run three simulations, one with all initial concentrations set to 0, one starting only with 20 molecules of cro RNA present, and one starting only with 50 molecules of cI RNA present. Present your results as concentrations of all components versus time.

2. Repeat the simulation with varying initial concentrations of cro and cI RNA, plotting all trajectories in the $[cro_{prot}]$ vs $[cI_{prot}]$ phase plane. Consider at least two sets of concentrations: (1) all combinations from 0 to 20 molecules in intervals of 1; (2) all combinations of 0 to 2000 molecules in intervals of 500. You may consider the initial concentration of both cro and cI protein to be zero in all cases.

3. Discuss the above results in the context of the biological system. Given this model, what are the minimal requirements for an infection to lead to lysis or lysogeny? What do you expect to happen for a typical infection?

4. The switch from lysogeny to lysis is mediated by stress-induced degradation of the cI protein. Explain how you would change your model to describe a bacteria under stress, and do so, starting from the stable lysogenic state. What must the rate of stress-induced degradation be in order to achieve the switch to lysis?

Part B: Stochastic model of the *cro–cI* genetic network.

The same fundamental model can be implemented as a stochastic model by use of the Gillespie algorithm. We use the same equations to define the fundamental rates of each reaction:

$$v_1 = \omega_{cI}[cI_{rna}] \qquad v_2 = \chi_{cI,prot}[cI_{prot}]$$

$$v_3 = \mu_{cI}\left(1 - \frac{[cro_{prot}]^2}{K_{cro,1/2}^2 + [cro_{prot}]^2}\right) \qquad v_4 = \chi_{cI,rna}[cI_{rna}]$$

$$v_5 = \omega_{cro}[cro_{rna}] \qquad v_6 = \chi_{cro,prot}[cro_{prot}]$$

$$v_7 = \mu_{cro}\left(1 - \frac{[cI_{prot}]^2}{K_{cI,1/2}^2 + [cI_{prot}]^2}\right) \qquad v_8 = \chi_{cro,rna}[cro_{rna}]$$

where the concentrations correspond the the numbers of molecules present. The algorithm can be generally outlined as:

- `for each step`
- `    find time increment`
- `    choose next reaction`
- `    increment or decrement molecule numbers by 1, as appropriate for the chosen reaction`

The time increment is chosen from an exponential distribution with coefficient R_{tot}^{-1}, where R_{tot} is the total of all reaction rates ($R_{tot} = \sum_{i=1}^{8} v_i$). The next reaction is chosen based on a probability proportional to v_i ($P(i) = \frac{v_i}{R_{tot}}$).

1. Implement this stochastic model in MATLAB, using the same constants as for the deterministic model. Beginning with all concentrations equal to zero, run a simulation for 50,000 steps, and plotting all concentrations versus time. Discuss what you observe.

2. Stochastic simulations will never give identical behavior each time, so repeat the simulation from (1) at least 20 times, plotting all simuluations on the same plot of concentration versus time; also plot the results on a single plot of the $[cro_{prot}]$ vs $[cI_{prot}]$ phase plane. Discuss in your observations in detail.

3. Repeat these simulations with starting concentrations of 20 molecules of either cro or cI RNA (do both, but independently), and discuss your results. Starting concentrations of both cro and cI protein should be set to zero.

4. Add the enhanced degradation rate found in examining the deterministic system (*i.e* the last question of Part 1), and begin a set of simulations (at least 20) from starting conditions corresponding to the lysogenic state. Describe the behavior you observe; do you think that the length of simulation is significant here? Repeat these simulations for a longer period of time (such as 100,000 steps) and discuss any differences.

5. Discuss how the results obtained from the stochastic model compare to those seen in the deterministic case. What are the implications of these differences for the biological system? Are there specific challenges in understanding the results of the stochastic simulations? Be as detailed as possible in your response

