MCB137L: Homework 12: Final Project

Student Name: Linfeng Hu

Student ID: 3039731775

Note:

The origin copy files can be found in GitHub:

<https://github.com/HULinfengHideki/MCB137_HW12>

For the estimations in question 2, they are written in Typora, a markdown compiler. And the screenshot is attached in the word file.

**1 Physical Biology of the Cell: Your Turn (syllabus design)**

MCB137L / MCB237L: Physical Biology of the Cell

Course Overview

To equip students with a comprehensive understanding of physical biology using quantitative models, and applying mathematics, physics, and computer science to solve biological problems.

Learning Orientations:

Students should leave the course with the ability to:

1. Make simple estimations of common parameters in cell biology, at least at the order of magnitude level.
2. Utilize literature and online resources (such as BioNumbers) to collect basic data, with a fundamental understanding of programming, data fitting, and data visualization tools, and basic skills in using large language models to assist programming.
3. Replicate existing biophysics models and validate empirical data, with advanced skills including questioning basic models and making simple optimizations.
4. Explain biological phenomena based on mathematical models, such as explaining cell fate decisions starting from mathematical models.
5. Have a general understanding of biological processes simulated based on statistical mechanics; know basic idea to analyze biological processes from energy pathways.

Top Insights:

Student should obtain some basic understanding and inspirations into:

1. The root properties of biological problems are essentially physics and chemical processes, which can be described by mathematical modelling (e.g. evolutionary principles).
2. The significance and validation of magnitude, diffusion, and dynamic states and processes in the context of cell biology.
3. The role of probability distribution in understanding biological stochastic processes: how the probability “maps” the real-world cases?
4. The sophistication of biological network: how to study their nature by dynamic approach?
5. The interpretability of the model: what is the models’ range of application? What is the portability of the model?

Course Format:

* Lectures: Twice weekly, focusing on general theory, with a mix of PowerPoint presentations and blackboard calculations.
* Discussion Session: Extension to the content in the lecture; discuss the assignment; hand-on data analysis and simulation.
* Homework: Weekly assignments on the topics covered in the corresponding lectures. There will be estimation calculations, theoretical model conductions, code interpretation, data analysis and visualization.
* Midterm Quiz: One midterm quiz in discussion, mainly focusing on quantification protocol in bounding energy.
* Final Project: Replication of experimental based on given data, experimental design, and guidance protocols. Also, course concepts integration can be incorporated.

Assessment:

|  |  |
| --- | --- |
| Item | Percentage |
| Attendance and Participation | 10% |
| Homework | 60% |
| Midterm Quiz | 10% |
| Final Project | 20% |

PowerPoints Arrangement:

|  |  |
| --- | --- |
| Lecture 1-3 | * Introduction to MCB 137/237 * Fundamentals: estimation on biology parameters, magnitude in cell biology * Tools tutorial: GPT-assisted coding, pseudo code, and basic prompt |
| Lecture 4-6 | * Simple modeling and reviews on statistics and probabilities * Bacterial growth model (sizer, adder, timer): expectations * Estimation: composition of single cell and synthetic factors * Review: Poisson distribution and Exponential curves |
| Lecture 7-8 | * Role of probability in biological processes * Real-world applications and theoretical understanding * Case studies: mutations, ion channel |
| Lecture 9-11 | * Diffusion as Biology’s Null Hypothesis for Dynamics * Detailed exploration of diffusion processes (e.g. axonal transport) * Case studies: calibrating fluorescent protein counts, FRAP measuring diffusion and establish limits for enzyme catalysis and other reactions |
| Lecture 12-15 | * Entropy Rules: basis of thermodynamics * Entropy maximization principles and their implications in biological systems * Ensemble: a set of imagination * Case study: entropic forces-DNA as an entropic spring |
| Lecture 16-17 | * Phase transition: theoretical and practical impacts * Entropy maximization and free energy maximization * Hands-on coding: phase transition diagram * Case study: F1B-1 and nucleolar size |
| Lecture 18-22 | * Introduction to the Boltzmann Distribution and its application to ion channels and two-state systems * “Energy favorable” approach: nature of binding problems in biology * Case study: simple repression in the lac operon * Kinetic proofreading and its role in ensuring biological specificity * Maxwell’s demon and local entropy reduction |
| Lecture 23-25 | * Simple graph theory and its application in biological network * Case study: lambda switch * Genetic circuit: Logic gate and positive/negative feedback |
| Lecture 26-29 | * Overview of Monte Carlo methods * Simple Monte Carlo simulation in Python: implementing a random walk as a Markovian process * Introduction to ligand-target docking and Markovian sampling * Introduction to AlphaFold 2 and protein structure prediction |
| Lecture 30 | * Comprehensive review of key concepts covered throughout the course, highlighting connections between topics * Discussion on potential future directions in physical biology * how current students can contribute to these areas |

Tentative syllabus:

Lectures 1 - 3: Introduction to Physical Biology

* Overview of MCB 137/237, introducing course goals and applications of mathematics, physics, and computer science in solving biological problems.
* Techniques for estimating biological parameters at the order of magnitude level.
* Introduction to GPT-assisted coding, pseudo code, and basic prompts for biological computations.

Lectures 4 - 6: Modeling and Statistical Fundamentals

* Introduction to basic models in biology, focusing on bacterial growth models such as sizer, adder, timer.
* Estimation of compositions in single cells and synthetic biological factors.
* Discussion on the application of Poisson distribution and exponential growth curves in biology.

Lectures 7 - 8: The Role of Probability in Biological Systems

* Exploration of how probability theories apply to biological systems, with real-world applications.
* Case studies on mutations and ion channel dynamics to demonstrate probability in action.

Lectures 9 - 11: Dynamics of Biological Systems

* Examination of diffusion processes, including axonal transport, as fundamental dynamics in biology.
* Case studies using fluorescent protein calibration and FRAP to measure diffusion limits and enzyme catalysis.

Lectures 12 - 15: Thermodynamics and Entropy in Biology

* Exploration of entropy rules and their implications in biological systems.
* Case study on DNA as an entropic spring, illustrating thermodynamic principles in biological contexts.

Lectures 16 - 17: Phase Transitions and Energy

* Examination of phase transitions in biological contexts and their theoretical and practical impacts.
* Hands-on activity with coding for phase transition diagrams and exploring entropy and free energy maximization.

Lectures 18 - 22: Boltzmann Distribution and Molecular Interactions

* Application of Boltzmann principles to ion channels and two-state systems.
* Exploration of kinetic proofreading, Maxwell’s demon, and local entropy reduction to ensure biological specificity.

Lectures 23 - 25: Networks and Systems in Biology

* Application of simple graph theory to understanding complex biological networks.
* Investigations into lambda switches and genetic circuits, emphasizing logic gates and feedback mechanisms.

Lectures 26 - 29: Monte Carlo Methods and Their Applications

* Introduction to Monte Carlo methods and their broad applications in biological systems.
* Implementation of simple Monte Carlo simulations in Python, ligand-target docking simulations, and introduction to AlphaFold 2 for protein structure prediction.

Lecture 30: Course Review and Future Directions

* Comprehensive review of key concepts covered throughout the course, emphasizing the integration of physical biology principles.
* Discussion on future directions in physical biology and how current students might contribute to the field.

**2 Order of Magnitude Estimation**

Question chosen: 4, 6, 8, 9, 10, 12, 18, 20, 22, 29, 30, 31, 34, 36, 38.

图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成

图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成

图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成

图形用户界面, 文本, 应用程序

描述已自动生成

文本, 应用程序

描述已自动生成

图形用户界面, 文本, 应用程序

描述已自动生成

图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成

图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成

文本, 日程表

中度可信度描述已自动生成

图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成

图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成

文本

描述已自动生成

图形用户界面, 文本, 应用程序

描述已自动生成

图形用户界面

描述已自动生成

文本, 应用程序

描述已自动生成

图形用户界面, 文本, 应用程序

描述已自动生成

**3 Your Philosophy of Biology**

**The Quantifiable Life:**

Harnessing Mathematical Models to Decode and Predict Biological Phenomena

Linfeng Hu

Essay on MCB 137 Final Project, UC Berkeley

Modern biology has fundamentally transformed from a non-quantitative science to a quantitative one, driven by advances in technology and methodology. During my elementary years, I learned how Robert Hooke used a microscope to discover cells, and in high school, I acknowledged that Aristotle had determined whales are viviparous. These early studies prioritized objective observations. In contrast, modern biology relies heavily on quantitative data. Noteworthy examples include the discovery of the genome of extinct species and human evolution. This shift reflects the influence of foundational sciences, which have long embraced quantification. Erwin Schrödinger's seminal question in 'What is Life?'—'How can events within a living organism be accounted for by physics and chemistry?'—highlights this inevitable transition in biological research from qualitative observation to quantitative precision.

Quantitative models serve as bridges between experimental data and theories, moving away from black box approaches to data interpretation. These models allow us to understand the underlying principles of a system, offering an approach to summarize, predict, and verify changes methodically. It enables systematic control over phenomena and even fosters the creation of new systems based on these principles. While trial and error method has its place, it will eventually be replaced by research and production guided by theoretical models.

In this course, we have explored several elegant yet simple quantitative models that demonstrate their practical value in biological contexts. For instance, concepts like bi-stability, hysteresis, and bifurcations explain why equilibrium states differ when approached from two different directions. These models show how biological systems can maintain stability or generate various response to different initial conditions. Furthermore, we've seen how binding probability and fold-change are used to describe repression mechanisms, providing a quantitative framework for understanding gene regulation. Additionally, graph theory has proved instrumental in managing models of simple activation, helping us visualize and analyze complex networks of interactions within biological systems.

Future advancements in biological technology will increasingly rely on quantitative data and models, with fields like synthetic biology poised for significant breakthroughs. Synthetic biology can now redesign genomics and metabolic pathways, transforming cells and even creating synthetic cells, offering boundless possibilities. Gradually, experimental biologists will require a theoretical basis to estimate how to engineer cells, a foundation that only quantitative biology can provide. While past modeling primarily focused on metabolic networks, modeling at the cellular and multicellular levels is still developing. The recent rise of AI for Science (AI4Sci) could also aid in extracting undiscovered scientific principles from data, optimizing and interpreting models, and gradually demystifying complex processes.

In conclusion, modern biology's transformation into a quantitative science underscores the necessity of using quantitative data and models. As discussed, these models are crucial for understanding complex biological systems and for pioneering advances in areas like synthetic biology. The mantra of this course, "quantitative data demands quantitative models," captures this shift effectively. This approach is essential, as it moves us beyond mere observation to a deeper, systematic understanding of life's complexities. As biology continues to evolve, quantitative models will remain indispensable in driving forward the next wave of scientific discoveries.

**4 A Feeling for the Organism: Your Turn**I wonder what the number of lipids is involved in hepatocytes apoptosis in an average person at any given time.

Knowing that under normal conditions, a very small fraction of hepatocytes are actively dividing. At the meantime, hepatocytes have the capability to enter the cell cycle in response to injury or loss.

**Data and source:**

Cells per gram of liver tissue: According to BNID 110544, there are approximately hepatocytes per gram of liver tissue.

Liver weight: According to BNID 110212, an average adult liver weighs about 1.5 kg.

Lipid content per gram of liver: According to BNID 105885, there is approximately 65 mg of lipid per gram of wet liver weight.

Apoptosis rate: According to BNID 112171, less than 0.1% of hepatocytes are undergoing apoptosis at any given time.

Length of edge of hepatocyte (in a roughly cubical simplification): According to BNID 113236, the length of edge of a typical hepatocytes is 15 .

Surface area per lipid molecule: According to BNID 106993, surface area per lipid molecule is .

“Rule of thumb” for the mass of lipid molecule in lipid bilayer: According to BNID 101838, "Rule of thumb" for the mass of lipid molecule in lipid bilayer is 800 Dalton.

**Assumptions:**

Most of the cells in liver is hepatocytes.

For a simplification model, each hepatocyte can be considered as a cube with a side length of 15 microns.

If a hepatocyte undergoes apoptosis, we consider all the lipids it has are involved in apoptosis.

Hepatocytes turnover is in a common steady state.

**Calculations:**

Total Hepatocytes in the Liver:

Knowing that hepatocytes occupies most of the mass in liver cell, the total lipid in the liver can be considered the lipids for hepatocytes.

Hepatocytes involved in apoptosis:

The surface area of a hepatocyte cell: given that the cell can be simplified in to a cubic with 15 length of side,

Consider that most of the lipid is used for plasma membrane composition, and the membrane is a bilayer lipids structure, we can estimate the number of lipids that each cell has:

Thus, the total number of lipids involved in apoptosis at any given time can be calculated:

If we consider all the lipids of a hepatocyte undergoing apoptosis means involvement in apoptosis, an estimation is:

At any given time, the number of lipids involved in apoptosis is . Since hepatocyte is the major cell, but not sole cell type in liver tissue, the magnitude of estimation could be approximately **few lipids**.

**Sanity check:**

According to BNID 110212, an average adult liver weighs about 1.5 kg.

According to BNID 105885, there is approximately 65 mg of lipid per gram of wet liver weight.

Total lipids weight:

According to BNID 101838, "Rule of thumb" for the mass of lipid molecule in lipid bilayer is 800 Dalton.

The mass of lipids involved in apoptosis can be calculated:

Considering only a very small fraction of hepatocytes are undergoing apoptosis at any given time, it lines up with our estimations and assumptions. Also, not only plasma membrane contains lipids, it also contributes to the small fraction.

End of the Final Project.

Thanks for your review and tutorials this semester.