

SECB4313 BioInformatics Modelling and Simulation
2024 April 15 - Quiz 1

Answer ALL Questions.

1) The discipline of modelling and simulation is increasingly being applied to **biological problems**. Based on the attached literature review paper, provide a summary (not more than 2 pages) on the application of modelling and simulation in biological systems. (5 marks)

Answer:

The discipline of modelling and simulation is increasingly being applied to biological problems. The issues involved are ranging from kinetic responses of atoms, complex reactions of biochemical pathways, and temporal behaviors of the biological systems. **Model** from the term “modelling” is the math formulation in mathematical terms of a real-world problem. **Simulation** defines as to run a model to predict the results of experimental changes in the system.

One of the application of modelling and simulation in biological systems is **Bio-molecular Simulation**. At the bio-molecular level, Simulations enable to examine biomolecular dynamics, interactions, and structures. Researchers can see the behaviour of proteins, nucleic acids, and other biomolecules at the atomic level thanks to methods like molecular dynamics (MD) simulations. The mechanisms of biological processes like protein folding, ligand binding, and enzymatic activities are explained by these simulations. Some of the Bio-molecular Simulation are DNA Modeling, Protein Modeling and Neuron Modeling.

Other than that, another application is **Metabolic Simulation** where understanding gene regulation networks, metabolic processes, and cellular signalling pathways requires modelling and simulation. In order to clarify intricate connections within cells and forecast emergent features, computational models might include experimental data. This method helps with synthetic biological system design, personalised medicine, and medication discovery. Besides Metabolic Simulation, there is **Ecosystem Simulation**. Modelling and simulation are used in ecology and evolutionary biology to examine species interactions, ecosystem dynamics, and population dynamics. With the movement and behaviour of individuals within populations, agent-based models, differential equations, and spatially explicit models shed light on ecological patterns, biodiversity, and conservation tactics.

Furthermore, with the **Artificial Life**, modeling and simulation play a crucial role in designing and optimizing medical devices, drug delivery systems, and tissue engineering constructs. Computational models help predict the performance of implants, assess drug efficacy and toxicity, and simulate physiological processes to guide therapeutic interventions. In fact, **Artificial Life** is deeply associated with **Drug Design**.

As of the **NeuroScience**, computational models of neural networks and brain circuits aid in understanding brain function, learning, and behavior. These models range from simplified neuronal models to large-scale simulations of brain regions. They contribute to the development of neuroprosthetics, brain-computer interfaces, and treatments for neurological disorders. Finally, the last application is **Epidemiology and Public Health**. Modeling and simulation are essential for studying the spread of infectious diseases, predicting epidemics, and evaluating public health interventions. Epidemiological models simulate disease transmission dynamics based on factors such as population demographics, social interactions, and healthcare infrastructure. These models inform disease control strategies, vaccination campaigns, and preparedness for future pandemics.

In summary, the use of modelling and simulation in biological systems has become essential for expanding our knowledge of living things at various scales. Computational methods allow scientists to investigate, assess, and forecast the behaviour of intricate biological systems, from molecular interactions to ecosystem dynamics, with the ultimate goal of finding applications in public health, biotechnology, medicine, and ecology.

2) Based on your answer in (1), focus/choose on ONE application only. Then find related case study (from the internet) which is similar to that chosen application and provide brief description of the case study. (4 marks)

-Next, briefly describe the mathematical modelling used in the modelling and simulation of the chosen case study. (5 marks)

Your answer must be less than 2 pages. State the resource/reference of your answer. Attached the article at the end of your answer sheet.

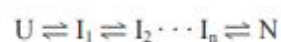
Answer:

Case study: Protein Folding: Matching Theory and Experiment

Link: [https://www.cell.com/biophysj/pdf/S0006-3495\(98\)77530-7.pdf](https://www.cell.com/biophysj/pdf/S0006-3495(98)77530-7.pdf)

From the answer above, I would like to focus on **Bio-molecular Simulation**. The case study that I found is **Protein Folding: Matching Theory and Experiment**. This case study mainly focus on matching theory and experiment is a major focus of studies on protein folding. The main concept involves how to model a protein molecule during folding so that simulations capture the key aspects of the process, and then using the simulations to learn how to interpret experimental data as realistically as possible. This case study also explores how computer modelling and experimental techniques can be used together to gain insights into protein folding. It discusses how simulations can help design experiments and how experimental data can be used to refine the simulation models. This back-and-forth approach is similar to how bio-molecular simulations are used to study other biological processes.

The mathematical modelling used in the modelling and simulation of the case study selected is shown as below:



According to this concept, folding is comparable to a typical chemical reaction or a metabolic pathway. It consists of a series of steps, each of which is distinguished by the generation of a particular intermediate. One step is typically rate-limiting, and folding's transition state I^\ddagger is ascribed to it. It is suggested that the process for figuring out the pathway is the same as for a metabolic pathway: the process is to isolate and characterise the intermediates. In practice, it is frequently challenging to populate and monitor the intermediates I_1, I_2, \dots . Specific parameters, also referred to as "strongly native" conditions, are necessary for folding in order to populate the intermediates. These requirements include low temperature, low denaturant concentration,

and the presence of a stabiliser like Na₂SO₄. It is possible to miss intermediates in less stabilising circumstances.

3) Give TWO points (with discussion) on how artificial intelligence may accelerates the modelling and simulation of your chosen case study (your answer in (2)). (4 marks)

Answer:

Some of the points on how artificial intelligence can or may accelerate the modelling and simulation of the chosen case study involved **“Protein Folding: Matching Theory and Experiment ”** involve **identifying the key folding steps and interactions** along with **predicting folding pathways and protein structures** .

For the first point, **identifying the key folding steps and interactions** shows the Complex interactions between different amino acids within the protein chain are necessary for protein folding. Simulations have always been time-consuming and computationally costly since they simulate every interaction. AI is capable of analysing enormous databases of protein structures and folding routes, particularly using deep learning algorithms. AI can aid in the creation of more precise simulations by spotting recurrent patterns and interactions that are essential for stable folded states. By focusing on the crucial interactions, these simulations would drastically lower the processing burden and speed up the procedure.

Apart from that, **predicting folding pathways and protein structures** shows current protein folding simulations often require a starting point, an initial guess of the protein's unfolded state. AI can be trained using known folding processes and protein structures. The AI can now understand the chemical and physical rules driving protein folding thanks to this training. With this information, even in the absence of an initial hypothesis, the AI can then forecast the most likely folding paths and ultimate structures for novel proteins. Researchers studying protein folding would be able to investigate a far wider variety of proteins and their activities with this capability, which would represent a major advancement in the field.

4) What is the role of “data” in simulation and modelling of biological systems? Suggest THREE Biological dataset (related to the case study that you have answered in Question (2)) that are publicly accessible. Provide description (including the URL) for each of the dataset. (8 marks)

Answer:

Data plays a fundamental role in simulating and modelling biological systems. It acts as the fuel that drives these simulations, providing the information needed to build realistic models and validate their accuracy. The categorization of the data is divided into **parameterization**, **model validation** and **calibration**. For the parameterization, biological models depend on a number of variables, including binding affinities, molecule concentrations, and reaction speeds. Accurate determination of these characteristics is aided by data from experiments such as microscopy, spectroscopy, and protein-protein interaction assays. For the model validation, after a model is constructed, its predictions are verified using information from separate experiments. This guarantees that the behaviour in the model is accurate. To validate a drug-target interaction model, for instance, experimental data on the drug's effect on the target protein would be compared to the model. Finally, the calibration shows the systems in biology are by nature complicated. Current models must be adjusted and improved upon when new data becomes available. Information from diverse sources, including as genetic investigations, data on the course of disease, and omics research (genomics, proteomics, metabolomics), aids in pinpointing areas in which models require enhancement.

The three biological dataset involve **PyMOL**. PyMOL is a potent molecular visualisation programme that is frequently used to see and examine biomolecular structures. It provides interactive molecular structure editing, sophisticated rendering capabilities, and a scripting language for task automation and specific visualisation customisation. PyMOL is frequently utilised for tasks including molecular docking, protein structure analysis, and molecular dynamics trajectory visualisation. **URL** : <https://pymol.org/>

Besides that, other biological dataset is **CHARMM** (Chemistry at HARvard Macromolecular Mechanics). Large-scale biomolecular system dynamics can be simulated using the CHARMM suite of molecular modelling and simulation software. With a variety of force fields and simulation techniques, it offers tools for energy minimization, molecular dynamics simulations, and structural analysis. CHARMM is especially useful for researching membrane proteins, protein-ligand interactions, and protein folding. URL: <https://www.charmm.org/>

The final biological dataset is **Rosetta**. A package of programmes called Rosetta is used for designing, predicting the structure, and modelling macromolecules. To forecast and optimise the structures of proteins and protein complexes, it combines a number of computational approaches, such as protein-protein docking, comparison modelling, and ab initio modelling. Protein structure prediction, investigation of protein-protein interactions, and protein design for biotechnological purposes are among the many uses of Rosetta. URL: <https://www.rosettacommons.org/>

5) Other than Python and R, suggest TWO other tools available in the market, that is suitable to simulate biological systems. Provide description for each tool (4 marks)

Answer:

Other than Python and R, some other tools available in the market suitable for simulating biological systems are **MATLAB/SimBiology** and **SimuLink**. **SimBiology** is a toolbox created especially for modelling and simulating biological processes, and it is available in the popular numerical computing environment **MATLAB**. Using a range of modelling techniques, such as ordinary differential equations (ODEs), stochastic simulation algorithms, and parameter estimation methods, SimBiology offers a graphical interface for creating complicated biological models. It is an effective tool for simulating many biological processes since it easily interacts with MATLAB's sophisticated computational and visualisation capabilities.

Finally, the **SimuLink** involves **MATLAB** that offers **Simulink**, a modelling and simulation platform for creating and simulating dynamic systems. Simulink is a popular engineering tool that may also be used to simulate biological systems, especially in the areas of biomechanics, control systems, and signal processing. It provides a block diagram interface for creating hierarchical models and is compatible with several types of simulations, such as hybrid, discrete-time, and continuous-time simulations.

Case study: Protein Folding: Matching Theory and Experiment

Link: [https://www.cell.com/biophysj/pdf/S0006-3495\(98\)77530-7.pdf](https://www.cell.com/biophysj/pdf/S0006-3495(98)77530-7.pdf)

[https://doi.org/10.1002/\(sici\)1521-3773\(19980420\)37:7%3C868::aid-anie868%3E3.0.co;2-h](https://doi.org/10.1002/(sici)1521-3773(19980420)37:7%3C868::aid-anie868%3E3.0.co;2-h)