COMPLEX NETWORKS IN SYSTEMS BIOLOGY
Course Overview

MATT MAHONEY

1. Course description

Systems biology is the study of biological organisms as networks of interacting components, from genes and molecules to tissues organs. These networks encode the complexity of biological systems; how different parts of the system influence each other to perform an astonishing array of functions. The mathematical theory of *complex networks* allows the systematic study of biological networks, and has become an indispensable tool of modern systems biology. The last 20 years have seen a surprising convergence of network concepts across many biological disciplines, where simple network ideas reappear in systems as disparate transcription networks of bacteria to neural connectivity patterns in the human brain. This course will introduce complex network theory through applications to real biological data. We will discuss the evolution and function of biological networks at multiple spatial scales, from molecules to the human brain. We will learn how the evolution of biological networks converges on universal patterns, and we will discuss what can go awry in networks to cause diseases.

2. Primary learning objectives

The purpose of this course is to survey the field of Network Biology, namely the study of biological systems using the mathematical theory of complex networks. Students will complete the course with fluency in the following areas:

- Programmatic manipulation of biological network data in multiple formats
- Basic topological features of networks (e.g. communities/modules, centrality, degree distribution, network motifs, and the scale-free and small-world phenomena) and their application to biological networks
- Random network ensembles and tests for statistical significance of network features
- Basic dynamical properties of networks, including network diffusion and reaction kinetics, and the role of network topology in modifying these dynamics

3. Secondary learning objectives

This course is designed as a 200-level elective for undergraduates in biomedical engineering, neuroscience, and computer science. It is designed as a 300-level elective for graduate students in the Neuroscience Graduate Program and the Complex Systems Certificate program. For this advanced population of students, a major focus on scientific communication will be maintained throughout. Students will be required to:

- Write well-commented, legible, concise computer programs to perform well-defined functions
- Provide cogent, written feedback on peers' work
- Engage productively in instructor-moderated in-class discussions
- Prepare rigorous papers and presentations conforming to modern scientific practices

4. Evaluation

4.1. **Weekly problem sets - 20% of final grade.** Students will solve a weekly problem set of ten questions requiring some combination of pencil-and-paper analysis, programming exercises, and essay responses.
Grading scale. Problems are scored as 0 (totally incorrect/not attempted), 1 (partially correct with significant errors/omissions), or 3 points (complete and correct). NB: A score of 2 is deliberately left out to reflect the austere ideology that complete and correct is substantially better than any alternative.
Grading. The instructor will automatically grade programming assignments according to whether student-generated programs output expected results on novel inputs. Programs that meet these benchmarks will be provisionally scored 3. All programs will be visually inspected and given a summary score (0, 1, 3) based on coding style, including commenting, clarity, and succinctness. Programs that score provisional 3's can be downgraded to 1's for poor style. The instructor will score non-programming exercises and provide written feedback.

4.2. **Midterm paper - 30% of final grade.** Students will prepare a midterm paper in the form of a standard scientific publication. The subject will be the same for all participants. The focus will be on the development of publication quality figures and concise long-form technical writing. The technical content of the midterm assignment will not be significantly harder than a weekly problem set.

4.3. **Final paper - 30% of final grade.** Students will prepare a final paper in the format of a standard scientific publication. The subject will be chosen by the student and will differ for each student. The focus will be on synthesizing multiple ideas from the course in a single research project.

4.4. **Final oral presentation - 10% of final grade.** Students will prepare a 15 minute oral presentation of their final research project. They will describe their motivate their problem, describe their findings (including showing all figures in their final paper), and describe future directions for that line of research.

4.5. **Peer-to-peer communication - 10% of final grade.** Students will participate in classroom discussions, in-class programming exercises, and peer grading.

Participation. Students are expected to attend class regularly, ask for help from and offer help to peers as needed during in-class assignments, and be engaged in subject matter discussions. The instructor will assess these activities and provide 5% or 0% of the final grade according to whether or not a student satisfactorily engaged with the material.

Peer grading. Students will grade the non-programming exercises of one other student each week and give a summary (0, 1, 3) score and provide a one sentence description explaining deficits in solutions for scores of 0 and 1. The instructor will assess peer grading and provide 5% or 0% of the final grade according to whether or not a student consistently provided cogent feedback to their fellow students, independent of the correlation between their scores and the instructors.

4.6. **Graduate student requirements.** Students taking this course for **graduate credit** will be graded according to a higher standards of scholarship. Problem sets, programming assignments, papers, and presentations will be assessed by how well they synthesize concepts into cogent scientific arguments and whether they conform to *professional standards* for scientific communication. In particular, factual assertions must be supported by sufficient relevant citations and/or *in silico* experimental evidence that is adequately displayed and described at a level appropriate for scientific publication. We will assume familiarity with the fundamentals of scientific writing and presentation. Papers and presentations are expected to demonstrate a clear understanding of how the project integrates into the larger literature on the topic. In particular, it is mandatory to develop a clear and rigorous rationale for performing the studies through a thorough discussion of prior literature. Similarly, it is required to systematically discuss both the limitations of the project methods and possible future directions, including avenues for addressing study limitations may be feasible in the near term.

4.7. **Undergraduate grade matrix.**

95+%	A
90-94%	A-
87-89%	B+
83-86%	B
80-92%	B-
77-79%	C+
73-76%	C
70-72%	C-
67-69%	D+
63-66%	D
60-62%	D-
≤ 59	F

4.8. **Graduate grade matrix.**

95+%	A
90-94%	A-
87-89%	B+
83-86%	B
80-92%	B-
77-79%	C+
73-76%	C
70-72%	C-
≤ 69	F

5. Syllabus

5.1. **Topological measures of importance.**

Centrality, peripherality, and hubs

- Embryonic lethality and protein networks
- Survivor bias and Mendelian disease
- Brain lesions and brain disease

Hierarchy, modularity, and communities

- Breaking down brains into functional subregions
- Breaking down molecular networks into functional pathways

5.2. **Global properties.**

Small worlds and wiring costs

- Synchronizability in neural circuits
- The small world of metabolism

Scale-free networks

- Evolution of molecular networks - the rich get richer

Degree correlations

- The rich club of the brain - a core processor

5.3. **Network motifs.**

Enrichment of structural elements

- Alon characterization of transcription networks
- Feed-forward inhibition in neural circuits

5.4. **Global properties of "stacked" networks.**

Pattern separation

- Biology of the primate visual system
- Multilayer perceptrons and deep neural networks
- Intracellular signaling cascades

5.5. **Reaction, Diffusion, and Reaction-Diffusion.**

The Laplacian

- Diffusion of brain activity

Normal mode analysis

- How the leopard got its spots