LS 50: Integrated Science

Course overview

LS50 is a two-semester, double course that introduces the natural sciences as an integrated whole. Its goal is to teach students how to solve scientific problems by drawing methods and concepts from biology, chemistry, physics, and mathematics. The course uses examples from biology as an integrating theme, principles from physics and mathematics to reduce complex problems to simpler forms, and computer simulation to allow students to develop their intuition about the behavior of the dynamical systems that control the physical and biological universe. The course includes bootcamps to introduce students to biological experiments and the computer language, MATLAB. Each semester will include a project lab, in which students will work in small teams to do original research on unsolved biological problems.

Natural science is a single intellectual enquiry into the universe of objects that surround us. Its components are linked by a common method: inducing hypotheses from a mixture of data and intuition, deducing predictions, and testing them by experiment and observation. The sciences depend on mathematics, from the simple act of counting to sophisticated methods required for computational chemistry and theoretical physics. The Integrated Science curriculum will introduce motivated freshmen to the concepts and methods needed to attack the life sciences in the 21^{st} century. For both semesters, students will take the equivalent of two courses, meeting for formal instruction every day, performing hands-on, original research, and using modern computer methods to simulate scenarios and analyze data.

Darwin and Wallace's theory of evolution revealed that the living things have a purpose: their structure, function, and behavior are integrated to leave as many progeny as possible. For much of the 20^{th} century, this difference, and the astonishing diversity of form and function, tended to separate biology from the other natural sciences: biology's complexity made it unappealing to many mathematicians, physicists, and chemists, and the "assume a spherical cow" flavor of theorists' simplifying assumptions made biologists skeptical about how useful theory was for understanding biology. Two advances have pushed biologists towards theorists and computer scientists: the need to test our understanding of biological processes by making explicit, mathematical models and the need to convert large datasets into information and, ultimately, knowledge.

We will teach students that the answer to "How will you solve this problem?" is "By any means necessary!" Our goal is to teach them how to find interesting problems, the means to solve them, and above all, the knowledge and courage to invent the new methods that make previously insoluble problems soluble. Coupling concepts and methods to problems that excite students and making them use these tools in their own research will embed the concepts in their working memory.

We will teach through iterated cycles of experiment and analysis, making use of experimental computation to simulate a system of interacting entities and explore the effect of parameter variation on the system's properties. Our goal is to complement the formal derivation of theorems, show the productive interplay between theory, simulation, and experiment, and show that computer systems and programs, like biological objects, have purposes. Mastering a restricted syntax to write algorithms will help students think about how biological systems use the restricted syntax of chemistry and genetics to accomplish tasks. Concepts like modularity, exploration with selection, error detection and correction, and recycling previous inventions are important in the function and evolution of both organisms and code. Six faculty will teach the course, working as three pairs of one life scientist and one physical scientist.

Students will use their knowledge to conduct original scientific research. To ensure that all the students start with abilities in laboratory experiments and computer simulations, we will have two, two week-long boot camps, one in computation and one in experimentation. After the boot camps, students will perform original research in project laboratories. The project labs will be based on the research of and run by the Bauer Fellows, independent scientists who spend five years at Harvard after their PhDs and run small research groups.

Class format

Lectures each weekday (MTuWThF) from 10-11:30 in Northwest Labs B108 (first basement level), 52 Oxford St.

Weekly discussion sections with teaching fellows, at locations and times to be determined based on student availability.

Weekly problem sets, to be submitted as indicated on each assignment.

Two in-class tests during semester, each covering the preceding month's worth of lecture material.

Cumulative final exams, with particular emphasis on material covered in the last month of the semester.

Two, two week bootcamps: one in programming in Matlab, one in experimental biology
Post-bootcamp, research-based laboratory: students work on original research projects in small groups, with open access to a sophisticated modern research laboratory. Students are expected to spend 6 hr/week working on their projects. Labs will meet Monday and Friday from 1-4 PM, beginning Monday, September 14th, in the teaching labs on the first basement level of Northwest Labs (52 Oxford St.). Bootcamps will include weekend sessions.

Faculty

An up-to-date list of contact information and office hours can be found on the course website at: https://canvas.harvard.edu/courses/3013/pages/course-staff-contact-information

Faculty

Ben de Bivort	OEB
Michael Desai	OEB
Cassandra Extavour	OEB
Erel Levine	Physics
Andrew Murray	MCB, Course Head
Mary Wahl	MCB

The year-long course will be divided into thirds, each team-taught by a pair of faculty, one from the physical and one from the life sciences. Murray will attend all the lectures to ensure continuity. We expect that the intense nature of the course will produce strong interactions both amongst the students and between them and the faculty.

Teaching Fellows and Assistants

Mariela Petkova	Biophysics	Early fall
Emma Nagy	MCB	Early fall
Cara Weisman	Biophysics	Late fall
Shu Wang	Biophysics	Late fall
Jim Valcourt	Systems Biology	Early spring
Parris Humphrey	OEB	Early spring
Zachary Werkoven	MCB	Late spring
Emily Hager	MCB	Late spring

Laboratory Heads

John Calarco	Bauer Fellow	Fall
Lauren O'Connell	Bauer Fellow	Spring

Laboratory Teaching Fellows and Assistants

Xico Gracida	Systems Biology	Fall
Adam Norris	Systems Biology	Fall

Lecture content

The list that follows was developed during a single semester Graduate Seminar in Undergraduate Education, in which the six faculty who will teach the course and seven graduate students, most of whom will serve as teaching fellows for the course, participated. Since the course will meet every day for two semesters, it will have roughly 125 lectures. As a result the description below is extremely telegraphic, including field-specific words, phrases, and acronyms. To provide more detail, we include detailed descriptions of the materials from three lectures: the opening lecture, the first lecture on probability and inference, and a lecture from roughly a quarter of the way into the course on energy. All three introduce areas, rather than going into areas in depth, since we thought it would be useful for the committee to see material that non-scientists could usefully opine on. For all three, we have provided a one page description of the key material the lecture covers (a roadmap) and the slides. This material is assembled in a single PDF file.

Lecture title	**Lecture content**
Introduction & course overview	What do cells & organisms do to survive & reproduce?
Biology as computation	The Central Dogma, coding & decoding, sensation & response
Inferences, Rules, Errors	The scientific method, measurement, measurement & statistical errors
Probability distributions and Bayesian analysis	Probability distributions, frequentist vs. Bayesian approaches, the Central Limit Theorem
Equilibrium constants and energy diagrams/equilibrium vs. steady-state	Equilibria, on & off rate constants, kinetics, activation energy
Lac operon I: Introduction	History, genetics & chemistry, non-genetic individuality via bistability
Lac operon II: Paradigm for gene regulation	Search times, introduction to 1-D diffusion, modern analysis
Quantization, elements, atomic orbitals	Atomic models: Bohr, Schrodinger
Periodic trends, molecular orbitals	Electronegativity, hybridization, bonds
Organic molecules and chirality	Structure of organic molecules, chirality, complexity
Protein structure, intermolecular forces	Hydrogen and van der Waals bonding, protein structure
Macromolecules and membranes	Sugars, lipids, nucleic acids, membranes, membrane properties
Introduction to ordinary differential equations	What is a differential equation, solutions to simple separable equation how to code numerical solutions
Dynamical systems I: fixed points and phase plane analysis	1-D dynamical systems: fixed points, phase plane analysis
Introduction to the cell cycle	Introduction, coordinating growth with division, yeast & embryos

Genetic and physiological analysis of the cell cycle	Embryonic oscillators vs. yeast dependent event chains, conserve regulators
Molecular mechanism of the cell cycle	Cell cycle engine mechanism & regulation, cell cycle checkpoints
DNA replication as an example of a cell cycle-controlled process	Biology of DNA replication and repair, kinetic proofreading
Cell cycle regulation by external events	Cell cycle regulation by environmental and developmental inputs
Cell cycle regulation of events in time and space	Controlling events in time and space, polarization, mitosis
Electrostatics and Debye screening	Managing backbone electrostatic repulsion during DNA compaction: introduction to electrostatics, Debye screening, nucle and chromosomal organization
Math primer: multivariate, Taylor series	Functions of more than one variable, partial derivatives, Taylor series
Many particle systems, approach to equilibrium	Definitions of system, environment, and entropy; mechanical, chemical, thermal
Equilibrium and entropy	Definition of entropy in terms of multiplicity, intensive vs. extensi variables, equations of state
Work and heat, energy, free energy	Work and heat; first law of thermodynamics; adiabatic and isothermal processes
Energy budget of the cell	Magnitude of overall ATP flux and distribution between processes efficiency
Boltzmann statistics	Finding extrema under constraints; Lagrange multipliers; the Boltzmann distribution; application to air pressure at altitude
Two-state systems	Ligand-receptor binding; Schottky; paramagnetism
Ideal gas, ideal solid, equipartition theorem	Average energy of molecule in a gas; equipartition theorem
Law of mass action	Reaction rate does not often reflect apparent stoichiometry: relationship between reaction rate and mechanism; rarity of termolecular processes; Arrhenius equation
Specificity: ligand, receptors, homologous recombination	Relative binding energy of correct and incorrect pairings; limitations on accuracy in translation; revisiting lac repressor movement
Biochemical approach to cooperativity	Derivation of Adair equation and Hill curve from expressions for chemical equilibria; MWC and KNF models and their biological motivations
Statistical mechanics approach to cooperativity	Gibbs distribution and grand partition function; application to derive equivalent expressions for cooperativity. Applied to nuclei acid denaturation.

Introduction to diffusion	Relationship between potentials and conservative forces; Fick's 1 and 2nd laws; Green's function for Brownian motion
Gaussian integrals, convolution, and applications to diffusion	Gaussian integrals and calculation of RMS distance traveled, convolution, applications to FRAP and other diffusion problems
Transport in the cell: diffusion, anomalous diffusion, and active transport	Detection of crowding/tethering and measurement of diffusion coefficient by FRAP, "speed" of diffusion, need for active transpor
Diffusion to detection (Berg and Purcell)	How many transporters are necessary? Steady-state solution for diffusion of particles towards absorbing regions on a sphere
Lac operon III: Application of rate equations and diffusion; revisiting classic system with modern techniques	Behavior of the lac operon explained via statistical mechanics, detailed analysis of diffusion and capture of the Lac repressor by the Lac operator
Molecular implementation of regulation: promoters, activators, repressors, chromatin	How transcription factors and cis-regulatory elements work (bacteria vs. eukaryotes), chromatin modification
Synthetic biological circuits	Design and implementation of synthetic circuits, e.g., genetic switches, repressilator, Collins toggle switch
Development I: discerning differences between cells	Methods for detecting differences in gene expression between cel epigenesis vs. predeterminism, overview of metazoan embryolog
Worms fate map (vulva as example)	Lineage, fate decisions, and noise in C. elegans vulval developmen
Chemical and phase equilibria	Clausius-Clapeyron, surface tension, mixing of hydrophobic molecules and water
Phase transitions	Phase diagrams, metastability, critical point, nucleation of solids
Solvation, solutions, and mixtures	Solvation, solutions, osmotic pressures, water and hydrophobicit
Electrostatics and the Nernst equation/intra-molecular interactions	The superposition principle; ionic bonds, hydrogen bonds, Londo dispersion forces can all be explained by Coulombic interactions; dipoles and dielectric effect; Bjerrum length calculation
Structure I	Modularity, primary through tertiary structure, turns, domains
Structure II	Self-assembly, Anfinsen, Levinthal, chaperones
Structure III	Driving forces of folding, exploration with selection (MCMC)
Introduction to optics	Reflection and refraction, Snell's law, lenses and microscopy; diffraction, Bragg's law
Fourier transforms	Definition of Fourier transform, examples, PSD for shot noise and relation to Poisson statistics, flicker and Brownian noise
Diffraction	Structure factor and relationship to electron density, diffraction grating demo, methods for phase determination
X-ray crystallography and EM	Interpretation of photo 51, introduction to EM

Membranes, permeability, and pumps	Lipid structure, surface tension and surfactants, permeability, partition coefficients, pumps and channels
Energy harvesting in photosynthesis	Introduction to photosynthesis with emphasis on light-dependen steps; production of proton gradient and ATP synthesis by relievi confinement.
Chemiosmotic hypothesis	Mitochondria, oxidative phosphorylation, generation of heat from brown fat, decoupling weight loss aids
Carbon fixation and consumption	Remaining steps of photosynthesis and glycolysis discussed
Enzymes I	Measurement, catalysis, local concentration
Enzymes II	Rate equations, separation of timescales, Michaelis-Menten kineti
Enzymes III	Orientation, arrow pushing, covalent intermediates
Two-component signaling	Introduction to signaling, crosstalk avoidance, arrow-pushing mechanism of HK-RR phosphotransfer
Linear algebra I	Vectors, scalars, vector space, dimensionality, component representation, matrices
Linear algebra II	Matrix multiplication, identity operator, and inverse
Linear algebra III	Change of basis, orthogonality
Linear algebra IV	Eigenvalues and eigenvectors
Principal Component Analysis (PCA)	Changing basis in a meaningful way, dimensionality reduction, application to genetic variation in Europeans
PCA application I: facial recognition	PCA applied to facial recognition
PCA application II: microbial communities	PCA applied to microbial communities, discussion of clustering methods
Newton's laws of motion	$F=ma$, second order differential equations, simple examples
Simple harmonic motion	Harmonic equation, solutions to this equation, ball on a spring.
Soft matter springs	Elasticity and entropic string
Conservation laws I	Energy and momentum
Conservation laws II	Example problems
Damped harmonic motion	Damped harmonic equation, overdamped and underdamped solutions
Driven harmonic motion	Driven harmonic equation, solution, resonance
Systems of harmonic equations	Normal modes of a system of harmonic equations
Examples of systems of harmonic equations	Coupled balls on a string, coupled pendulums, applications to modes of molecular oscillations
Motors and polarity	Directional motion of motors along polymers, dynein, sarcomere cycle

Thermodynamic cycles and molecular motor efficiency	Carnot cycle, calculation of efficiency of ATP synthase motor, Brownian ratchets and directed motion
Cytoskeleton I	Basics, polymerization, critical concentration, spindle assembly
Cytoskeleton II	Non-equilibrium polymers
Segregation	Pros and cons of putting different activities in different compartments
Endosymbiont theory	Cyanobacterial/alpha-proteobacterial origin of the chloroplast/mitochondrion, gradual gene transfer, non-Mendeliar inheritance
Import and export	Nuclear transport, protein secretion
Symmetry and symmetry breaking	Symmetry as invariance under transformation, CP violation and CPT symmetry, lateral inhibition and bud site determination
Development II	A-P axis patterning in Drosophila embryo, maternal effect/gap/segment polarity genes, morphogens, useful transgeni techniques
Dynamical systems II	Bifurcations; examples possibly including spruce budworm, firefl flashing, homochirality
Dynamical systems III	Two-dimensional dynamical systems, classifying fixed points, spi nodes
Dynamical systems IV	Limit cycles, Lotka-Volterra, Poincare-Bendixson
Introduction to neuroscience	Membrane potentials, chemical synapses, gap junctions, calcium compartments
Hodgkin and Huxley	Hodgkin-Huxley model in neuroscience, solutions by simulations.
Optogenetics, behavioral feedback loops	Using light to control neural activity and probe the function of nervous systems
Locomotion	Pattern recognition, inhibitory-excitatory alternation, proprioception, efference copy
Stochastic processes: from random walks to Markov chains	Formalizing random walks as a Markov chain, basic properties of Markov chains
Stochastic differential equations	Langevin equation for a stochastic process, interpretations, simulations
Biological examples of stochasticity	Mutation, excitability, plasmid partition
Master equation, birth-death processes	Using matrices to model the temporal evolution of complex syster
Noise in molecular biology	Stochasticity of gene expression, neuronal activity
Graph theory	Graphs as formal devices to represent connections within comple systems. Examples from control of neuronal activity and gene expression.
Network theory	Analysis of the structures and properties of large networks.

Worms II	Connectome, EM reconstruction, network properties
Metabolic networks	Metabolism as a dynamical system, flux balance analysis
Social interactions	Origins of complex behaviors (e.g. flocking) from simple rules
Classical genetics	Genes, alleles, sex, meiosis, mapping, screens
Molecular genetics I	Mutant selection, epistasis, chromosome elements, manipulation
Molecular genetics II	Diploids, transgenesis, CRISPR
Genomics	Sequencing, assembly, homology, BLAST
Evolutionary relationships	Phylogenetics, orthology, paralogy
Origin of life	Origin of life, RNA world
History of terrestrial biology and chemistry	History of life on earth, mass extinctions, major transitions in evolution
Diversity of life	Diversity of life, strange phyla
Population genetics I	Mutations and the origin and maintenance of diversity, selection, drift
Population genetics II	Sex: meiosis, sexual selection, Muller's ratchet
Population genetics III	Speciation and adaptive radiations
Population genetics IV	Detecting selection: MK, extended haplotype tests, bulk segregant/QTL
Evolutionary thinking	Drift, modification, recycling with modification, modularity
Molecular evolution of proteins	Evolution of protein function, extracting information from covariation
Evolution of development	Regulatory vs. coding region evolution, interpretation of heterologous expression experiments in Drosophilids, QTL applie to sticklebacks, duplication and divergence in coding regions
Diversifying mechanisms	Diversity inducing mech., contingency loci, switches, non genet individuality

Assessment

Assessment will be based on five criteria:

In-class tests
We will administer two in-class tests, each covering the material presented in the preceding month. The question format and content will be similar to the weekly problem sets.

Final exams
Since this is a double class, we will administer two, three-hour finals in each semester. For each the goal will be that well-prepared students should be able to complete the final in two hours and questions will focus on solving problems (mathematical analysis of problems, data interpretation, and experimental design), rather then regurgitation. The final exam will be cumulative but

special emphasis will be placed on topics from the last month of the semester (since no in-class test will cover that material).

Problem sets

Weekly problem sets will review the topics covered in lectures and discussion sections. We highly encourage students to collaborate to solve the problems, which may include a comparison of scratch work, methods used, and numerical results. All submitted work should be written independently and reflect the students' own understanding. For example, a student should not submit work which they would be unable to explain or reproduce on their own.

This policy reflects our intent in assigning the problem sets: to help students master the material under low-stress circumstances so they will not be blindsided by higher-stakes tests. While one should be wary of over-reliance, classmates are an excellent resource for help since they keep similar hours, live nearby, and remember their own learning process well enough to explain concepts appropriately. On the flip side, mentoring colleagues whenever possible is not merely kind and forward-thinking, but also helps crystallize your understanding and practice the mentorship and presentation skills needed for work in the scientific community.

Laboratory research

Students will be graded on the research projects they undertake, focusing on their ability to design and interpret experiments and their commitment to and engagement with their research more than on the quality of the results they produce (subject as these are to the whims of the Gods). Approximately half of the grade from laboratory research will reflect weekly participation in lab, with the remainder attributed to a final lab report.

Lecture participation

Students will be graded on their participation in class, which we expect to be highly interactive, weekly conferences, and the discussions associated with laboratory research.

These criteria will be accorded the following weights

In-class tests	20%
Final exams	30%
Problem sets	20%
Laboratory research	20%
Lecture participation	10%