NAME

FINAL EXAMINATION *of* SYSTEMS BIOLOGY

7.32 • 8.591 • 7.81

DEPARTMENT *of* PHYSICS
MASSACHUSETTS INSTITUTE *of* TECHNOLOGY

Instructions

a. **Write your name on the first page of the exam.**

b. Please do not turn the page until instructed to do so.

c. This exam is closed-book and closed-notes.

d. No calculators.

e. If the space is restricting your thoughts, you can continue your answer on the back side of each page.

f. You can use either pen or pencil to write your solutions. We trust you.

A preview of upcoming challenges

1 [15] Evolution of Virulence

Consider the set of equations to model the evolution of virulence, where x is the number of non-infected hosts, y_1 is the number of hosts infected with parasite 1, and y_2 is the number of hosts infected with parasite 2:

$$\begin{aligned}\frac{dx}{dt} &= k - ux - x(\beta_1 y_1 + \beta_2 y_2)\\ \frac{dy_1}{dt} &= y_1(\beta_1 x - u - \nu_1)\\ \frac{dy_2}{dt} &= y_2(\beta_2 x - u - \nu_2)\end{aligned}$$

a. [2] At equilibrium only one of the two parasites will typically survive. Why?

b. [3] Begin by considering the dynamics in the absence of parasite 2. Find an expression for equilibrium 1 (E1), in which parasite 1 survives and parasite 2 does not.

c. [3] Parasite 1 is endemic (non-zero) when R_1, the basic reproductive ratio of parasite 1, is larger than one. What is the expression for R_1 and how can it be interpreted?

d. [4] Show that if $R_2 > R_1 > 1$ then parasite 2 can invade equilibrium 1 and that parasite 1 cannot invade equilibrium 2 (when parasite 2 is present at non-zero number).

e. [3] How will virulence evolve over time if $\beta_1 = \alpha v_1$ and $\beta_2 = \alpha v_2$, where α is some coefficient?

2 [16] Cooperativity

Consider a DNA segment D containing two binding sites, A and B, to which protein X can bind in any order. The segment D is "activated" only when both sites are occupied. When neither site is occupied, the rate of binding to a site by a single molecule of X is r_1^{on}; when one of the sites is occupied, the rate of binding to the other site by a single molecule of X is r_2^{on}. The rate of unbinding of a single molecule of X is r_2^{off} when both sites are occupied and r_1^{off} when only one site is occupied. All the rates above are given for a specific concentration of the protein X: $[X] = X_0$ and all four rates have the dimension of $time^{-1}$.

a. [2] Some of these rates depend on the concentration X_0. Which of them and how?

b. [4] If $r_1^{on} = r_2^{on}$, $r_1^{off} = r_2^{off} = 0$, and at $t = 0$ both sites are unoccupied, what is the probability density function for the time it takes to occupy both sites? Plot it. What is the expected time it takes to occupy both sites?

Now, let's consider the most general case when r_1^{on}, r_1^{off}, r_2^{on} and r_2^{off} are not specified.

c. [3] Identify **all** conditions when the following statement must be true: the majority of the time the DNA segment D is bound by at least one X.

- □ $r_1^{on} = r_2^{off}, r_1^{off} \gg r_2^{on}$
- □ $r_1^{on} = r_2^{off}, r_1^{off} \ll r_2^{on}$
- □ $r_1^{off} = r_2^{on}, r_1^{on} \gg r_2^{off}$
- □ $r_1^{off} = r_2^{on}, r_1^{on} \ll r_2^{off}$
- □ $r_1^{on} = r_2^{on}, r_1^{off} \gg r_2^{off}$
- □ $r_1^{on} = r_2^{on}, r_1^{off} \ll r_2^{off}$
- □ $r_1^{on} \gg r_2^{on}, r_1^{off} = r_2^{off}$
- □ $r_1^{on} \ll r_2^{on}, r_1^{off} = r_2^{off}$
- □ $r_1^{on} = r_2^{on} = r_1^{off} = r_2^{off}$

d. [4] Identify **all** conditions that make the activation of D cooperative.

- □ $r_1^{on} = r_2^{off}, r_1^{off} \gg r_2^{on}$
- □ $r_1^{on} = r_2^{off}, r_1^{off} \ll r_2^{on}$
- □ $r_1^{off} = r_2^{on}, r_1^{on} \gg r_2^{off}$
- □ $r_1^{off} = r_2^{on}, r_1^{on} \ll r_2^{off}$
- □ $r_1^{on} = r_2^{on}, r_1^{off} \gg r_2^{off}$
- □ $r_1^{on} = r_2^{on}, r_1^{off} \ll r_2^{off}$
- □ $r_1^{on} \gg r_2^{on}, r_1^{off} = r_2^{off}$
- □ $r_1^{on} \ll r_2^{on}, r_1^{off} = r_2^{off}$
- □ $r_1^{on} = r_2^{on} = r_1^{off} = r_2^{off}$

e. [3] For any of the cases when the binding is cooperative, sketch the probability that D is activated as a function of $[X]$. Comment on the characteristic points on both axes. In particular, what is the concentration of X at which D is activated half of the time? If you remember the expression for this concentration, you don't have to derive it.

3 [17] Critical Transitions in the Toggle Switch

Consider the following system of differential equations describing a gene regulation network:

$$\frac{du}{dt} = \frac{\alpha v^n}{1+v^n} - u$$

$$\frac{dv}{dt} = \frac{\alpha u^n}{1+u^n} - v$$

Here, u and v are the concentrations of two genes expressed in dimensionless units and t is time expressed in dimensionless units. Assume that $\alpha > 0$, $n \geq 1$.

a. [2] Draw the gene regulation diagram of the system.

b. [1] If the transcription rate of the activated genes increases, α will

□ Increase □ Stay the same □ Decrease

c. [1] If the degradation of the proteins increases, α will

□ Increase □ Stay the same □ Decrease

d. [1] If the binding affinity of the transcription factors increases, α will

□ Increase □ Stay the same □ Decrease

e. [3] On the (u, v) plane draw the nullclines for $n = 2$ and $\alpha = 2$.

f. [2] The system is bistable when

□ $n \gg 1$ and $\alpha \ll 1$

□ $n \gg 1$ and $\alpha \gg 1$

□ $n = 1$ and $\alpha \gg 1$

□ $n = 1$ and $\alpha \ll 1$

g. [3] For $n = 2$, draw the bifurcation diagram for the system with the control parameter α on the x-axis and the state variable u on the y-axis. Where is the bifurcation? Draw all stable fixed points with a solid line and unstable fixed points with a dashed line.

h. [2] When the parameter α approaches a bifurcation point, how does the recovery rate after a small perturbation behave?

i. [2] How is the system in this question different from the toggle switch system discussed in class?

4 [11] Reynolds Number

a. [2] For an object of radius a, moving with velocity v in a liquid with density ρ and viscosity η, what is the Reynolds number?

b. [2] What do low/high Reynolds numbers physically mean?

c. [3] The critical force is a characteristic of a liquid and depends only on ρ and η. By performing dimensional analysis, find the critical force.

d. [4] Which of the following statements are true?

□ At high Reynolds numbers, the external force necessary to maintain a velocity v is much smaller than the critical force.

□ At low Reynolds numbers, the external force necessary to maintain a velocity v in various types of liquids is proportional to the critical force.

□ At low Reynolds numbers, the force necessary to maintain a velocity v is proportional to v^2.

□ At low Reynolds numbers, the force necessary to maintain a velocity v is proportional to the linear size of an object.

5 [18] Fitness Landscape

Consider a species which has a binary genome of length 100. The fitness of the genotype of all 0's is $1 + 80s_1$ and the fitness of the genotype of all 1's is $1 + 20s_2$. Let's assume the following fitness function:

$$f(n) = \begin{cases} 1 + (80 - n)s_1 & n < 80 \\ 1 + (n - 80)s_2 & n \geq 80 \end{cases},$$

where n is the number of 1's in the genotype. The population size is $N = 10000$ and $s_1 = 0.001$, $s_2 = 0.05$.

a. [2] Plot fitness as a function of n.

b. [2] For a state $n = 80$, plot the distribution of beneficial point mutations (the histogram of the selection coefficients). (A point mutation is a single substitution of either $1 \rightarrow 0$ or $0 \rightarrow 1$ in the entire genome).

c. [4] If for every individual in the population $n = 80$, how many times on average should a new beneficial mutation appear before fixation?

d. [4] An initially clonal population with $n = 80$ evolves for a long time. Eventually it will reach a steady state when the average n of a given population will not change anymore. For what mutation rates per base pair per replication is the population average n at the steady state bistable? Give the numeric values for the boundaries of the mutation rate.

e. [3] If for every individual in the population $n = 90$, for what mutation rates per base pair per replication can clonal interference be ignored?

f. [3] If for every individual in the population $n = 10$ and the mutation rate is low, what is the minimum population size so that population evolves in the direction of decreasing n?

6 [9] Host-Symbiont Interactions

Consider a host-symbiont interaction in which a large population of symbionts interacts with a single host. Both the host and the symbionts can produce an essential nutrient that everyone needs to grow. Each of the species has two strategies: (i) to produce the nutrient (cooperate) and (ii) not to produce the nutrient (defect). The production of the nutrient is costly, and if at least someone produces the nutrient, both species get the full benefit. The benefit of getting the nutrient is b; the cost of producing it is c; the generation times are t_h and t_s for a host and a symbiont respectively.

a. [2] What are the payoff matrices of a host and a symbiont in this game?

b. [2] Start by assuming that $t_h \approx t_s$. What is required of b and c so that there will be two stable states?

c. [3] If $t_s \ll t_h$, then one of these two equilibria becomes unstable. Which one and why?

d. [2] The stable state in this case is equivalent to a "sequential game" in which one of the players moves first and the other moves second. Which player moves first?

7 [14] Neutral Theory and Relative Species Abundance in Ecology

In the generalized birth-death model describing the mainland population, new species appear with rate b_0, individuals reproduce with rate b per individual and individuals die with rate d. Whenever the only individual of a given species dies, the species disappears. In this model, the mean number of species containing n ($n > 0$) individuals is given by

$$\langle \phi_n \rangle = \theta \frac{x^n}{n},$$

where $x = \frac{b}{d} < 1$.

a. [2] When $b_0 = 0$, what is θ? Explain.

b. [2] Write the master equations for the number of individuals in a given species.

c. [2] What is the distribution of the number of new species appearing during a time interval t? What is the mean and the variance?

d. [3] At a given time, how many individuals are there in the population on average?

e. [2] What is the probability that a randomly chosen individual is the only individual of its species?

f. [3] What is the probability that a randomly chosen individual belongs to a species that has a total of two individuals (i.e. $n = 2$ for the species)?

8 [11] Adaptive Prediction of Environmental Changes by Microorganisms

Harry Potter is to present the article "Adaptive prediction of environmental changes by microorganisms" by Amir Mitchell *et al.* during his lab's weekly journal club. For the past week, Mr. Potter has been busy handling one crisis after another. He did manage to throw together a presentation with some of the figures from the paper. However, he is quite confused about a few of the concepts and the presentation is just an hour away. Your task is to help Mr. Potter. Examine the figure below and answer the questions that follow.

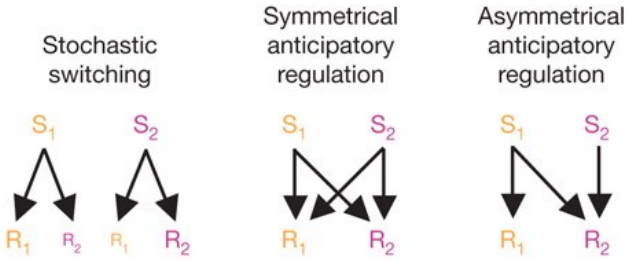

Figure 1: Three regulation strategies in response to environmental stimuli.

For each of the environments, mark the most favorable regulation strategy.

a. [1] S_1 and S_2 appear infrequently, but whenever S_1 appears, S_2 soon follows.
□ Stochastic □ Symmetric □ Asymmetric

b. [1] S_1 and S_2 appear infrequently, but when either appears the other signal soon follows.
□ Stochastic □ Symmetric □ Asymmetric

c. [1] The environment changes unpredictably.
□ Stochastic □ Symmetric □ Asymmetric

d. [4] Suppose a biological system is shown to exhibit an asymmetric regulatory response, such that the appearance of stimuli S_1 induces responses R_1 and R_2 while exposure to stimuli S_2 leads only to response R_2. Identify the criteria that suggest that this cross-regulation pattern evolved as an adaptive anticipatory response strategy.

□ Pre-exposure to stimuli S_1 of environment 1 increases fitness in environment 2, but pre-exposure to stimuli S_2 does not increase fitness in environment 1.

□ Pre-induction of genes needed to cope with S_2 is beneficial during S_1.

□ Pre-induction of genes needed to cope with S_2 is costly and not beneficial during S_1.

□ The conditioned response is specific to S_1 and not to other unrelated stimuli.

e. [2] Just before the start of the presentation, Draco Malfoy erased several data points from 2 figures in Harry Potter's power point presentation. Your task is to restore the missing data points by drawing them in.

In the first figure, the authors measured promoter activities in a wild-type strain in response to the addition of a sugar (lactose, maltose). They also measured promoter leakiness (no sugar). Here, LacZ is the lactose promoter while MalE, MalK, MalP, MalS, and MalZ are maltose promoters. The missing data points are:

- One green data point (promoter leakiness) for the LacZ promoter measurement.
- One red data point (added sugar lactose) per maltose promoter measurement for MalK, MalP, MalS and MalZ.

You know that when the bacteria *E. coli* travel through the digestive tract they are first exposed to lactose and then to maltose. Use your knowledge to restore the missing data points.

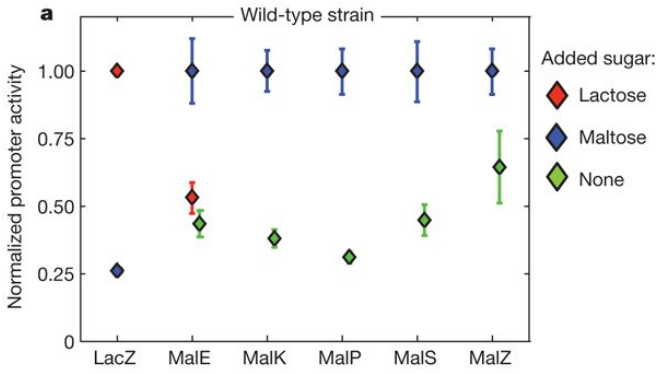

Figure 2

Courtesy of Nature Publishing Group. Used with permission.

f. [2] In the next figure, the authors examined laboratory evolved strains of *E. coli*, which grew for 500 generations on high levels of lactose but without maltose. The authors measured the promoter activity of the evolved strains in response to the addition of a sugar (lactose, maltose). Here, all six red data points (corresponding to promoter activity after addition of the sugar lactose) are missing. Draw them in.

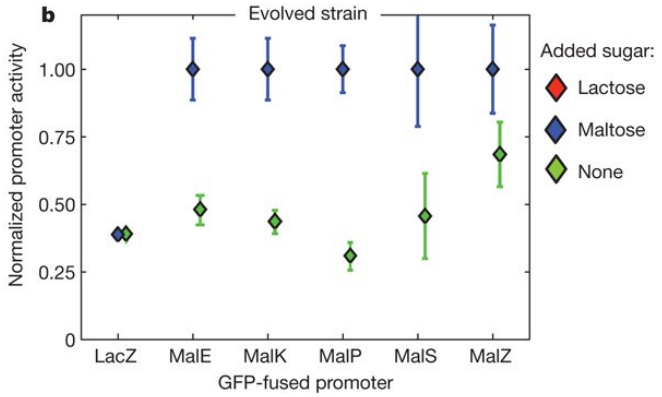

Figure 3

Courtesy of Nature Publishing Group. Used with permission.

9 [16] Rock-Paper-Scissors

Consider a 2 player, 3 strategy game with the following payoff matrix

$$\begin{pmatrix} 0 & -1 & b \\ 1 & 0 & -b \\ -1 & 1 & 0 \end{pmatrix}$$

$b > 0$.

a. [2] Are any of the pure strategies a Nash equilibrium?

b. [3] Consider two mixed strategies $p_1 = (x_1, x_2, 1 - x_1 - x_2)$ and $p_2 = (y_1, y_2, 1 - y_1 - y_2)$. What is your expected payoff if you play p_1 and your opponent plays p_2? How does the payoff behave as a function of x_1 and x_2 (linear, quadratic, exponential, etc.)?

c. [3] For strategy p_2 to be a Nash equilibrium, the inequality $E(p_2, p_2) \geq E(p_1, p_2)$ must hold for all p_1. Given the results of part b, it is sufficient to check the payoffs for three strategies p_1. What are those strategies?

d. [4] Show that if $y_1 = \frac{b}{1+2b}$ and $y_2 = \frac{b}{1+2b}$ then $p_2 = (y_1, y_2, 1 - y_1 - y_2)$ is a Nash equilibrium.

e. [4] Is the Nash equilibrium above an evolutionarily stable strategy? Explain.

8.591J / 7.81J / 7.32 Systems Biology
Fall 2014