

Problem 1: Lateral Inhibition: Epithelial Patterning:

In this problem you will study how patterns emerge in a system where cells mutually repress nearest neighbour cells. Such a process is called lateral inhibition and is responsible for the patterning of many tissues (e.g. hair cells on your skin, the 100's of photoreceptor cells in a fly's eye). In higher organisms it is carried out by the Notch-Delta signaling pathway. Delta is a surface bound ligand which binds to surface bound receptors on neighbouring cells called Notch. Binding of Delta, causes Notch to be cleaved and a portion of it then transmits signal to the nucleus where it regulates gene transcription. Activation of Notch leads to repression of *delta* within the activated cell. Thus two neighbouring cells have a negative feedback loop for the production of Delta – high delta in one cell forces the neighbouring cell to have low delta. A two cell system acts as a bistable switch. For a cell with Z nearest neighbours, a lateral inhibition model for the production of delta is:

$$\frac{d D_i}{dt} = \frac{\alpha}{1 + \left(\sum_{j=1}^Z D_j \right)^n} - D_i = G(\{D\}, t)$$

where D_k is the concentration of Delta in the k th cell, n is a cooperativity factor and the sum is just over nearest neighbours. To numerically solve do the following: use the Euler method to integrate the equation (i.e. $D_i(t + \Delta t) = D_i(t) + G(\{D(t)\}, t) \Delta t$) where Δt is sufficiently small ~ 0.0001 - 0.01 should be small enough. The production rate parameter α probably won't have much of an effect. Run it for a few hundred to thousand time steps – probably not necessary to check for convergence as it should reach a steady state fairly quickly. Start with all D 's equal to a constant ~ 0.5 with random noise introduced to break the spatial uniformity.

1. Solve the model numerically on a square lattice using $n=1, 2, 3$: (you'll find the $n=1$ case to be boring).
 - (a) Try a 4×4 square, 5×6 rectangle, and some bigger square using periodic boundary conditions (PBC)
 - (b) Fixed boundary conditions with $D = 0$ on the boundary for the same squares in (a).
2. Simulate this model on a triangular lattice using the same conditions with PBC and fixed boundary conditions.

What sort of patterns emerge? Differences between the square and triangular lattice?

3. Put in defects with some small probability, p . To do this, for each lattice site, make it a defect if a random number $< p$ and keep track of which sites are defects using an array of

flags = 0 or 1, one for each lattice site. Then when you do your nearest neighbour sum at each lattice site, those sites that are defects won't contribute to the sum. How do the patterns change? Does changing p have any effect?

Problem 2: Lotka-Volterra

This problem is a classic predator-prey model which is used in studying food webs where one organism is feeding on another. It is also relevant to some biochemical models where one chemical species is 'eating' another. In the model there are populations of foxes, $F(t)$, and rabbits, $R(t)$. The dynamics of these two populations are given by

$$\frac{dR}{dt} = aR - bRF$$

$$\frac{dF}{dt} = ebRF - cF$$

where a is the natural growth rate for rabbits, c is the natural death rate for foxes, b is the death rate of rabbits due to an encounter with a fox, and e is the efficiency of turning eaten rabbits into foxes.

- a) Explain in words what each of the two terms in each of the above equations represents.
- b) Solve for the equilibrium populations of foxes and rabbits. You should find two equilibrium values for each population.
- c) Try to figure out the stability of each equilibrium point by doing some graphical analysis. To do this draw a 2D graph where the x axis is the population of rabbits, R and the y axis is the population of foxes, F . On this graph, label the two equilibrium coordinate positions of (R, F) . Now in the vicinities of each of these positions, sketch the flow vectors. The x component of the flow vector is given by dR/dt and the y-component of the flow vector is given by dF/dt . Use the parameter values $a=0.1$, $b=0.05$, $c=0.5$ and $e=0.25$ so that you can have some numbers. By looking at the flow around the equilibrium points, are they stable or unstable?
- d) You should arrive at a conclusion about the stability of the two equilibrium points that may seem confusing. Think about it for a bit. Given your findings about the stabilities of the equilibrium points, what do you think the temporal behaviour of the populations of rabbits and foxes will look like?

Problem 3: Gillespie algorithm:

In this problem you will simulate a simple biological reaction including stochastic effects. The reaction is a simple model for the production of protein from mRNA. The two dynamic equations are:

$$\frac{dm}{dt} = v - \gamma m$$
$$\frac{dp}{dt} = \alpha m - \delta p$$

where v is the rate of mRNA molecule production, γ is the rate at which mRNA is degraded, α is the rate of protein production from mRNA and δ is the degradation rate for the protein. At steady state, the average number of mRNA molecules is

$\bar{m} = v/\gamma$, and the average number of proteins is $\bar{p} = (\alpha/\delta)\bar{m}$. We want to consider what happens when the molecular numbers are small, so set $\bar{m} = 20$ and set the average number of protein's produced from a single mRNA to be $\alpha/\delta = 10$. To set the kinetics, you can set $\delta = 1$ and we will assume that our mRNA is short-lived in comparison to our protein, so set $\gamma = 5\delta$.

1. Use the Gillespie method outlined in class to simulate the four reactions governing the production of mRNA and protein using the initial conditions that $m(0) = 0$ and $p(0) = 0$. Generate a time course for both $m(t)$ and $p(t)$ where the time runs from $t = 0$ to $t = t_f$, and t_f is chosen such that the initial transient is gone.
2. Now you will generate the probability distributions $P(m)$ and $P(p)$ which gives the probability that the system will have exactly m mRNAs and p proteins at time t_f . To get these probabilities, you'll need two arrays which will keep track of counting m and p . Run the procedure in (a) several hundred times, and for each run simply add to the counts for the m and p that you get at the last time point t_f . Your distribution for the mRNA better be Poisson with a variance of $\sqrt{20}$. Does the variance in the protein distribution go as $\sqrt{\bar{p}}$? or is it something different? The protein has a noisy source as well as other noise terms.

Problem 4: Morphogen Gradients:

During the process of tissue development in multicellular organisms, gradients of molecules called morphogens are set up which generate complex patterns of gene expression, that ultimately lead to the vast assortment of different cell types. A common way for an organism to set up a morphogen gradient, is to have a point source (i.e. a cell (or cell cluster)) that emit morphogens, which then diffuse out into the surrounding tissue. A given morphogen molecule diffuses with a diffusion coefficient, D , and is degraded at some rate, γ . For an organism setting up a morphogen gradient along its long axis, at steady state, the concentration of morphogen, $c(x)$, satisfies the reaction-diffusion

equation $D d^2 c(x)/dx^2 - \gamma c(x) = 0$.

- Given the units of D , c and x (i.e. length²/time, #/length³, and length), what must be the units of γ ?
- Consider the point source, that generates a flux of a morphogens at $x=0$. Show that $c(x) = A \exp(-x/\lambda)$ where $\lambda = \sqrt{\gamma/D}$ is a solution to the above reaction-diffusion equation. What are the units of λ ?
- Given that there is a flux of particles, α at $x=0$, use Fick's flux equation, $j = -D dc(x)/dx$ at $x=0$ to solve for A in terms of D , α , and γ .

Now we're going to look at the readout of this morphogen gradient by another factor and calculate the information in this process. **I just want you try to do as much of this problem as you can.** Let us say that there is an output gene, g that has c as an activator input. The average amount of g that gets produced for a given level of c is,

$$\bar{g}(c) = g_0 \frac{(c/K)^n}{1 + (c/K)^n}$$

How much information is transferred between the input c and the output g ? This can be quantified by the mutual information, given by,

$$I(g, c) = \int dg dc P(g|c) P(c) \log \left(\frac{P(g|c)}{P(g)} \right)$$

where $P(g|c)$ is the conditional probability of g given c , $P(g)$ is the distribution over output g , and $P(c)$ is the distribution over c . We'll now make some reasonable assumptions about the nature of the distributions for $P(g|c)$ and $P(c)$, in order to evaluate the mutual information. Let's assume that $P(g|c)$ is a Gaussian with mean

$$\bar{g}(c) \text{ and variance, } \sigma_g^2(c) = \left(\frac{d\bar{g}(c)}{dc} \right)^2 c + \bar{g}(c) \text{ . (Can you explain why the variance}$$

should have this form?). From the results of (a-c), the the input c is a function of

position along the length of the embryo, x , so we can write $P(c) = \frac{1}{L} \int_{x=0}^{x=L} dx P(c|x)$

where L is the length of the embryo. Again we'll assume that $P(c|x)$ is a Gaussian with mean $\bar{c}(x) = c_0 \exp(-x/\lambda)$ and variance, $\sigma_c^2(x) = \bar{c}(x)$. The last distribution that we need, $P(g)$ is found from $P(g) = \int_{c=0}^{c=\infty} dc P(g|c) P(c)$.

In the developing fruitfly, the output gene reads off the input and is turned on on $\frac{1}{2}$ the cell and off on the other. Fix the length of the embryo, L to be some number. Choose a value for λ . Then set $K = c_0 \exp(-L/2\lambda)$ so that the transition in the output, g , happens at $\frac{1}{2}$ the embryo length. Evaluate the dependence of the mutual information on the Hill coefficient, n at different values of λ . **Is there an optimal choice of n and λ that leads to the most mutual information given a g_0 and c_0 ?** You will have to do these integrals numerically – the functional form is specified for each of the integrands.