Principles of Synthetic Biology: Midterm Exam

October 28, 2010

1 Conceptual Simple Circuits

1.1 Consider the plots in figure 1. Identify all critical points with an x. Put a circle around the x for each critical point if it is stable. (3 pts)
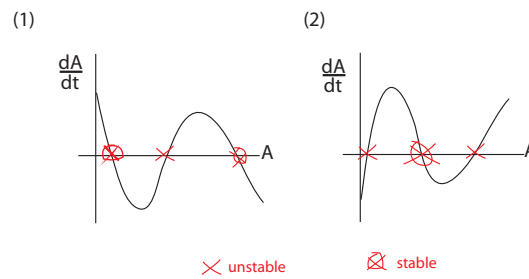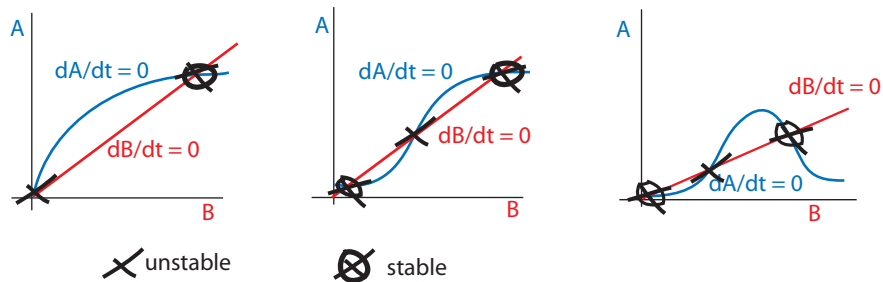

Figure 1: Schematic for questions 1.1.

Figure 2: Schematic for questions 1.2 through 1.3

1.2 Consider the nullclines in figure 2. Assume dA/dt is negative for large A and dB/dt is negative for large B. Identify all critical points with an x. Put a circle around the x for each critical point if it is stable. (5 pts)

1.3 What is the potential advantage of having the null-cline for A be high for intermediate values of B and low for high or low values of B? (3 pts)

The value of B in the high state does not change too dramatically even as the nullcline for B becomes less steep. Instead of just the max value of B at the high stable point being bounded, there a reasonable range of parameter space where the values of both A and B in the high state are bounded. This may place less metabolic load on the cell by avoiding overproduction of the pathway components.

1.4 Consider the cascades in figure 3. Which configuration allows for the largest effective Hill Coefficient to be achieved? (2 pts)

1.5 Describe two property you expect to differ between the circuits shown in figure 4 which may be useful to a bioengineer? (5 pts)

Switch-like response

- Version **B** has potential to be switch like. As long as the rate constants are appropriately matched, the hill coefficients grow. See HW3 problem 2 for a detailed analysis and mathematical justification.
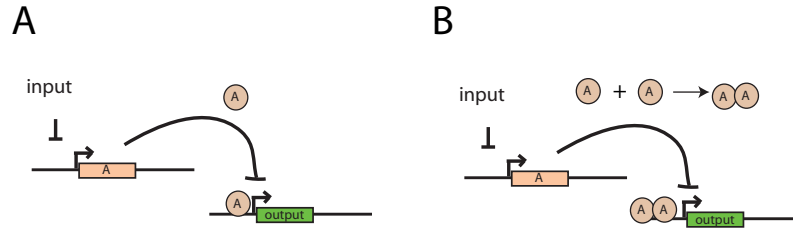
2

Figure 3: Schematic for question **??**

Speed

- Version **B** will in general take longer to propigate a signal.

Noise

- The improved cooperativity of version **B** may be useful in reducing noise, since it expands the region where small fluctuations have no effect.

- The delay created by the cascade can act to average out fluctuations in the input signal and thus reduce noise.

- If the independent components are reasonably noisy themselves relative to the noise in the input and output signals, the cascade will in fact amplify the noise and reduce reliability with increasing length.

- As an engineer one can use these counterbalancing effects to derive an optimal signal transduction cascade length, (See Thattai and van Oudenaarden *Biophys J* 2002).
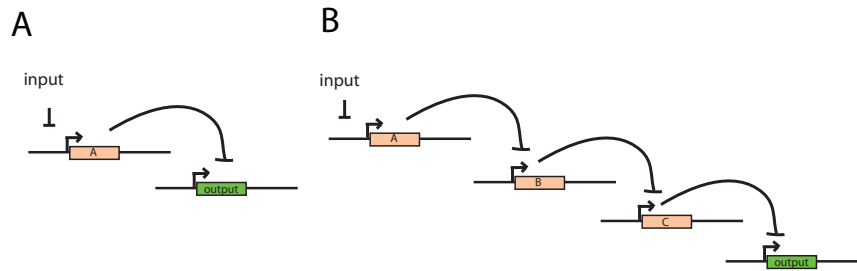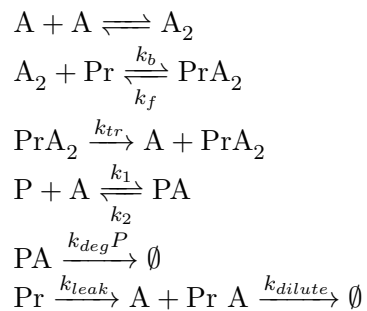
A   B

input   input
⊥   ⊥

Figure 4: Schematic for questions 1.4 and 1.5

2   Deterministic Systems Analysis

2.1   Consider a simple feedback system where protein A binds as a dimer, A_2, to activate its own promoter, as shown in the cartoon in figure 5. A is also degraded by a protease, P. Write down the chemical equations which represent this system. (3 pts)

$\text{A} + \text{A} \rightleftharpoons \text{A}_2$

$\text{A}_2 + \text{Pr} \underset{k_f}{\overset{k_b}{\rightleftharpoons}} \text{PrA}_2$

$\text{PrA}_2 \xrightarrow{k_{tr}} \text{A} + \text{PrA}_2$

$\text{P} + \text{A} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \text{PA}$

$\text{PA} \xrightarrow{k_{deg}P} \emptyset$

$\text{Pr} \xrightarrow{k_{leak}} \text{A} + \text{Pr} \quad \text{A} \xrightarrow{k_{dilute}} \emptyset$

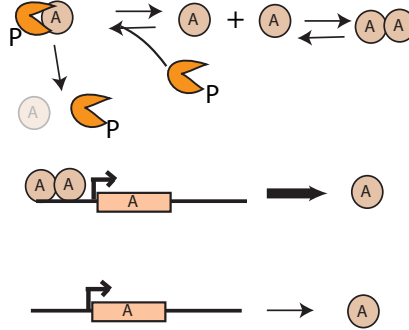where Pr is the Promoter and P_T the total concentration of protease.

Figure 5: Cartoon for question 2

2.2   Assume dimer formation and promoter binding are both fast relative to protein synthesis. Also assume the protease is in very low concentration and always saturated. What is the mass action equation for $\dot{A}$ as a function of A (remove any reference to transient states like A_2Pr) (15 pts)

$$\frac{dA}{dt} = k_{tr}PrA_2 + k_{leak}Pr - P_T k_{deg} \frac{A}{\frac{k_2+k_{deg}}{k_1} + A} - k_{dilute}A \tag{1}$$

Where degradation follows from simple michaelis menten kinetics, and is constant if the enzyme is saturated,

$$\frac{dA}{dt} = k_{tr}PrA_2 + k_{leak}Pr - P_T k_{deg} - k_{dilute}A \tag{2}$$

Add conservation:

$$Pr_T = [Pr] + [PrA_2] \tag{3}$$

And rapid equilibrium of dimer binding to promoter:

$$[PrA_2] = \frac{k_f}{k_b}[Pr][A_2] \tag{4}$$

$$[PrA_2] = \frac{k_f}{k_b}(Pr_T - [PrA_2])[A_2]$$

$$[PrA_2] + \frac{k_f}{k_b}[PrA_2][A_2] = \frac{k_f}{k_b}(Pr_T)[A_2]$$

5

$$[PrA_2](1 + \frac{k_f}{k_b}[A_2]) = \frac{k_f}{k_b}(Pr_T)[A_2]$$

$$[PrA_2] = \frac{\frac{k_f}{k_b}(Pr_T)[A_2]}{(1 + \frac{k_f}{k_b}[A_2])}$$

$$[PrA_2] = \frac{(Pr_T)[A_2]}{\frac{k_f}{k_b} + [A_2]} \tag{5}$$

$$[Pr] = [P_T] - [PrA_2] = Pr_T\left(1 - \frac{[A_2]}{\frac{k_f}{k_b} + [A_2]}\right) \tag{6}$$

$$[Pr] = Pr_T\left(\frac{\frac{k_f}{k_b} + [A_2]}{\frac{k_f}{k_b} + [A_2]} - \frac{[A_2]}{\frac{k_f}{k_b} + [A_2]}\right) = Pr_T\frac{\frac{k_f}{k_b}}{\frac{k_f}{k_b} + [A_2]} \tag{7}$$

Thus,

$$\frac{dA}{dt} = k_{tr}PrA_2 + k_{leak}Pr - P_Tk_{deg} - k_{dilute}A \tag{8}$$

becomes

$$\frac{dA}{dt} = Pr_T\frac{k_{leak} + k_{tr}A_2}{\frac{k_b}{k_f} + A_2} - P_Tk_{deg} - k_{dilute}A \tag{9}$$

Now we just need to get rid of A_2 using the rapid dimerization assumption: $A_2 = \frac{k_1}{k_2}A^2$

$$\frac{dA}{dt} = \frac{k_{tr}\frac{k_1}{k_2}A^2}{\frac{k_b}{k_f} + \frac{k_1}{k_2}A^2} - P_Tk_{deg} - k_{dilute}A \tag{10}$$

Multiplying through by $\frac{k_2}{k_1}$ over itself, and renaming $\frac{k_bk_2}{k_1k_f}$ as K_M^2

$$\frac{dA}{dt} = Pr_T\frac{k_{leak} + k_{tr}A^2}{K_M^2 + A^2} - P_Tk_{deg} - k_{dilute}A \tag{11}$$

2.3 How many unique steady states are possible? Describe qualitatively what are the steady states? (4 pts)

Linear decay and sigmoidal activation can intersect at most 3 times, giving three conditions: one stable, low A state, one stable high A state , and an intermediate, unstable state.

If decay is strictly in the saturated regime then there are only 2 intersections in our simple model. However at sufficiently high A cell division or or non protease mediated decay will check production so that there will be a stable 'on' state rather than unlimited production.

2.4 Consider next the case where the protein dimers form a tetramer before binding the promoter. Assume all multermization reactions are fast equilibrating. What is your new equation for $\dot{A}$? (4 pts)

This is much easier than analogous derivation for cooperative, sequential binding (which you did on the HW). Simply: $A_4 = k_{eq2} A_2^2$

$$\frac{dA}{dt} = Pr_T \frac{k_{leak} + k_{tr} A^4}{K_{M2}^4 + A^4} - P_T k_{deg} A \tag{12}$$

where $K_{M2}^4 = \frac{k_b k_2 k_4}{k_3 k_1 k_f}$

2.5 Explain a biological reason why a promoter with this tetramerization architecture might be useful to a cell or an engineer? (2 pts)

Increases the hill coefficient, giving a more switch like response

- System is more robust to fluctuations around the low or high point.

- Larger bistable region, easier to get bistability and more robust to keep it.

- Fewer intermediate values of promoter output are exhibited by the system

2.6 You are allowed to drive A or P expression by an inducible promoter in addition to the system, explain why this new system can be a prefect threshold detector for inducer of the new promoter for A or P. (7 pts)

Increasing the concentration of P can be used to flip the system into the low state. Decreasing the output of P might allow it to flip into the high state. An alternate promoter making A would function in the opposite way, increasing A flips the system into the high state by moving up the A nullcline, decreasing it may move the A production term in $\dot{A}$ below the degradation switching the system to the off state. See figure 6.
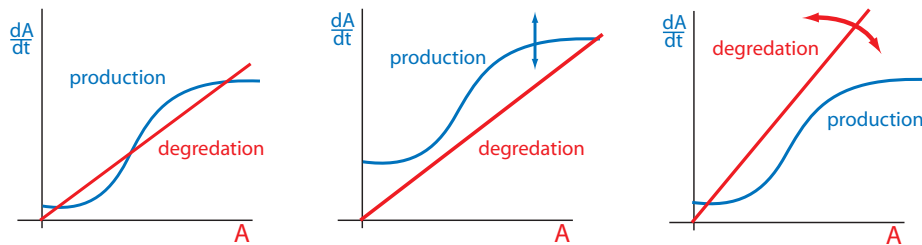

Figure 6: Switching state by changing A (middle panel) moves the sigmoid up and down, which can result in either monostable on, bistable, or monostable off conditions. Changing the concentration of the protease changes the degradation rate, which changes the slope of the red degradation line. This can sweep the line above, below, or intersecting the sigmoidal curve to again tune stability.

3 Efficient Engineering of Insulin Production

It is desired to control the release of insulin in an artificial beta cell. Let $I = 1$ denote release of insulin, and $I = 0$ to denote that the cell is not producing insulin. The four input conditions for determining the output condition: produce insulin should be 1 are:

A. Insulin should be controlled automatically during the hours of 8pm to 6am. Let the variable T be 0 during this period, and 1 otherwise. A timing device will be used to provide those inputs to the system, and a person can override this constraint manually.

B. It is not desired to produce insulin while there is low blood sugar from not eating because a person could die from too little glucose in the blood stream. Let the variable R be 0 if blood sugar is low and 1 blood sugar is high (presumably from eating).

C. The system is to be capable of manual ($M = 1$) or automatic operation ($M = 0$). If the blood sugar level becomes low while the system is in manual operation, the production of insulin is to be turned off.

D. We also want to be able to have the system have an override switch to turn it off completely. Let the variable P be 1 when the system is operating, either in manual or in automatic mode, and 0 when the system is turned off.

3.1 Obtain the truth table for the above system (5 pts)

.

P	R	M	T	I
0	0	0	0	0
0	0	0	1	0
0	0	1	0	0
0	0	1	1	0
0	1	0	0	0
0	1	0	1	0
0	1	1	0	0
0	1	1	1	0
1	0	0	0	0
1	0	0	1	0
1	0	1	0	0
1	0	1	1	0
1	1	0	0	0
1	1	0	1	1
1	1	1	0	1
1	1	1	1	1

3.2 Obtain the Boolean expression for the output, I, from the truth table. (4 pts)

$I = R \cdot P \cdot (T + M)$

3.3 Reduce the answer of to the simplest form, and again state the result as a Boolean expression and suitable English. (4 pts)

The whole system needs to be in the permissive state, sugar needs to be high, and either it needs to be the correct time of day or a manual mode activation in order for insulin to be produced.

3.4 Implement the design, using the fewest gates possible in terms of theoretical circuit elements. (10 pts)

3.5 Convert this simple Boolean expression to a genetic circuit design. (10 pts)

You can use things like a TF that becomes inactive when there isnt enough free glucose in the system. (We know glucose can not freely diffuse, but you can imagine a TF at the end of a signaling pathway that can detect blood glucose levels.)
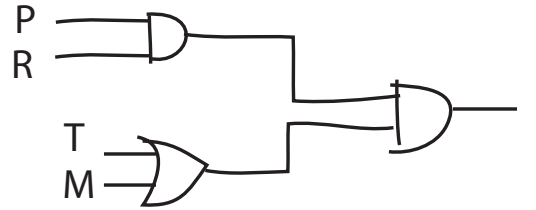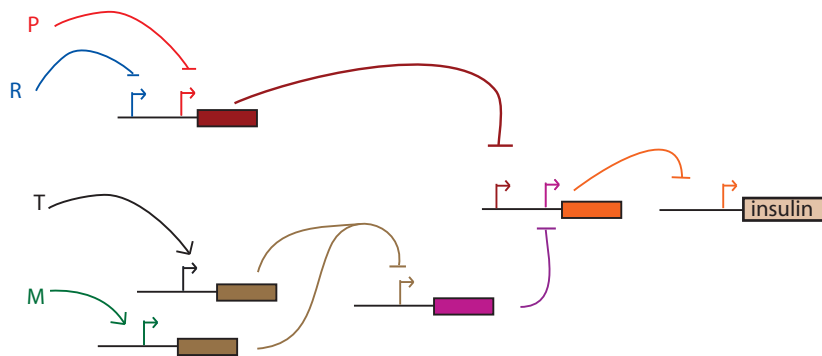
Figure 7: Circuit Diagram

Figure 8: Genetic implimentation. Promoters are colored by the genes which regulate them. Repressed promoters are assumed to be active unless repressor is present. Activated promoters are only active if activator is present.

3.5.1 List all the assumptions you make when designing the system. Be as realistic as possible. (2 pts)

Either we need about 8 different promoters or we need cell signaling pathways that can be separated. All should be reasonably cooperative for binary logic to work. Into different cells so we can reuse some components. We need our proteins expressed as reasonably high levels to avoid the effects of low number fluctuations. We also assume our input / output strengths are matched appropriately.

4 Stochastic System Analysis

Imagine we have a uniform field of 1000 genetically and chemically equivelant cells that we want to differentiate into two distinct populations, A sensors and B sensors. A sensors express some factor A, which is a very strong inhibitor of pathway B. B cells express factor B, which is in turn a strong inhibitor of A.

A naive cell has the ability to produce either A or B. The A promoter fires with a mean rate k_A and the B promoter a mean rate k_B. Because of the strong inhibition, the dynamics of this system are essentially governed by whichever promoter fires first.

4.1 What is the expected (i.e. mean) number of A sensor-cells? produced by this system? Justify your response. (6 pts)

This is just an exponentional race, as derived in HW3:

$$\frac{k_A}{k_A + k_B} \tag{13}$$

4.2 What is the standard deviation in the number of A cells? (Assume each cell is independent). Justify your response. (6 pts)

This is simply a binomial distribution since the states are A and NOT A (i.e. B), with probability of A being $p = \frac{k_A}{k_A+k_B}$ and the $n = 1000$ cells are all independent.

The binomial standard deviation is

$$\sqrt{np(1-p)} \tag{14}$$

plugging in the probabilities above,

$$\sqrt{1000\left(\frac{k_A}{k_A + k_B}\frac{k_B}{k_A + k_B}\right)} = \sqrt{1000\frac{k_A k_B}{(k_A + k_B)^2}} \tag{15}$$