BNFO601: Integrated Bioinformatics
Problem Set: Modeling

1. Modify ADPAs.py so that it uses metabolite and reaction classes in the same way as glycolysis.py. Feel free to copy shamelessly from that program.

2. You have made the discovery that adding the fatty acid linolenic acid to sea water containing sea urchin eggs produces hydrogen peroxide. This phenomenon has never been described before (who would have thought to do the experiment?) and it occurs to you that maybe the eggs use hydrogen peroxide as a defense against predators. Before you go deeper into fantasy, you want to make sure that it is really an enzyme from the eggs that is responsible for the reaction.

You therefore repeat the experiment with different concentrations of linolenic acid and get the results available from the module web site. Actually, you get two different sets of results (data set 1 and data set 2), depending on which parallel universe you live in. In each universe, decide whether the reaction is or is not enzyme-catalyzed, and in either case determine the constants associated with the reaction.

3. The liver enzyme cytochrome P450 is involved in the metabolism a large number of compounds, particularly those of foreign origin, like acetaminophen, caffeine, codeine, and AZT. Since and the effect of a drug is closely related to its lifetime in the body, understanding of the action of cytochrome P450 on a particular drug is often critical for predicting the drug's efficacy. You want to model the reaction below, catalyzed by cytochrome P450. Determine whether each of the proposed rate equations that follow is correct. If it isn't, make it correct.

\[ k_f \Rightarrow \quad k_c \Rightarrow \]

Reaction: \( \text{Pg} + \text{NADPH} + \text{O}_2 + \text{E} \rightleftharpoons \text{E-complex} \rightleftharpoons 6\text{-beta-OH-Pg} + \text{NADP}^+ + \text{E} \rightleftharpoons k_c \)

\( \text{Pg} = \text{progesterone} \)
\( 6\text{-beta-OH-Pg} = 6\text{-beta-hydroxy-progesterone} \)
\( \text{E} = \text{cytochrome P450} \)

3a. \( d[6\text{-beta-OH-Pg}]/dt = [\text{E-complex}] k_c - [6\text{-beta-OH-Pg}] [\text{NADP}] [\text{E}] k_c \)

3b. \( d[\text{E}]/dt = [\text{Pg}] [\text{NADPH}] [\text{O}_2] [\text{E}] k_f \)

3c. \( d[\text{E-complex}]/dt = - [\text{Pg}] [\text{NADPH}] [\text{O}_2] [\text{E}] + [\text{E-complex}] k_c \)

4. Alter the program Glycolysis.pl in the following ways:

4a. It may be disconcerting to run the program and have nothing happen for ~20 seconds. Do yourself a favor and change the program to report progress every 100th time step, printing the time and the current concentration of Glc6P (plus anything else that interests you)

**Hint:** Use the Python operator \( a \% b \) (which gives you the remainder of \( a/b \)) in conjunction with the function \( \text{round} \) (because \( \% \) doesn't work well with floating point numbers). If you aren't sure how \( \% \) and \( \text{round} \) work, play with them in short test programs.
4b. Modify the code you wrote in 3a to display the values of all reactions and metabolites, using the Print_all debugging tool (you'll find it in the subroutines of glycolysis.pl).

5. glycolysis.py has in the subroutine Model the following line:
   \[ \text{GbisP}.dx/dt = \text{Rxn7}.v - \text{Rxn8}.v \]

5a. What is the physical significance of this equation?

5b. Give the equation that produced the value of \( \text{Rxn7}.v \).

6. What happens when you decrease the time interval used in stepping through glycolysis.py? What happens when you increase the time interval? Be sure you try enough values to get a picture of the general trend.

7. Eisenthal & Cornish-Bowden (1998) did not use Eq. 2 of their article to model all the non-equilibrium reactions. If you look at the footnotes to Table 1, you'll find multiple exceptions. Consider the pyruvate kinase reaction.

7a. How does glycolysis.py handle that reaction? What justification is there for the specifics of that choice?

7b. Consider Eq. 5 in Eisenthal & Cornish-Bowden (1998) and the surrounding text. They claim that if the products of the pyruvate kinase reaction are zero, then the kinetics should be the same as described in Table 1. Do the algebra. What do you get by setting the products to zero? [Note that the exponent \( h \) is not explicitly defined in Eq. 5, but you'll find a description of it (not named as \( h \)) in the pertinent footnote for the reaction in Table 1]

7c. Modify the code so that glycolysis.py follows Eq. 5. Does the new model give results different from the old? Why (not)?

8. Alter the program glycolysis.py so as to test the effect of inhibiting hexokinase with a mythical compound that exhibits uncompetitive inhibition with respect to glucose. What level of inhibition is required to make a dent on ATP levels?

Suggested strategy:

- Note that Eisenthal and Cornish-Bowden (1998) have already done the enzymological derivation for you.
- Write a subroutine (perhaps called \( v_{of\_rxn2} \) that calculates the velocity of the reaction.
- Note that Eisenthal and Cornish-Bowden defines \( i \) as the concentration of the inhibitor divided by the constant of inhibition. This constant is equal to the concentration of inhibitor required to lead to half-maximal binding to the enzyme. Thus the ratio \( i \) is related (nonlinearly) to the occupancy of the enzyme by the inhibitor and hence the degree of inhibition. You don't have any value for the constant, but you can vary \( i \) over
reasonable values. I suggest you run the program multiple times, each time with different values of \( i \).

- To do this, I suggest making an instance variable (e.g. `Rxn2.inhibitory_constant` [you can think of a much better name!]), not a class variable. Don't initialize the instance variable inside the class definition (then it would be a class variable) but rather where you instantiate `Rxn2`.

9. Write an equation to describe the number of bacteria in a culture where the bacteria double every 20 minutes (the maximal growth rate of \( E. \ coli \)). Presume that there is no limitation by food or other resources. Now write a rate equation for the process.

10. Suppose that you are interested in studying the dynamics of how sexually transmitted diseases spread. From a study of the general population, it has been found:

- The disease is spread solely by heterosexual contact.
- The frequency of new infection is proportional to both the number of carriers in the population and the number of susceptible partners.
- Males and females are infected at different rates.
- Carrier males are treated and cured at a higher rate than carrier females (because the symptoms are more severe in males, hence they are more likely to seek treatment).
- They never learn. Cured males and females return to the susceptible population just as likely as before to contract the disease.

10a. Write rate equations describing the rates of change of the two carrier populations (\( C_m \) - carrier males, and \( C_f \) - carrier females).

You would like to test your equations on a test population. Accordingly you advertise for volunteers and select 600 healthy females and 899 healthy males for the experiment. You place the volunteers on a south sea island along with one infected male and wish them luck. You will return in 12 years to tabulate the results, but you'd like to know right now what to expect.

10b. Devise a model for the experiment, one that will predict the number of infected males and females over the 12-year duration of the experiment. Use the following constants:

\[
\begin{align*}
&k_1 = 0.000032 \text{ infected males per day per appropriate male per appropriate female} \\
&k_2 = 0.2 \text{ successfully treated males per day per appropriate male} \\
&k_3 = 0.00033 \text{ infected females per day per appropriate female per appropriate male} \\
&k_4 = 0.025 \text{ successfully treated females per day per appropriate female}
\end{align*}
\]

10c. Plot and interpret the results.
You are faced with the all too common problem of dividing a doughnut evenly amongst three people. The easy way is to slice it by three radial cuts (see picture to right).

But this would be WRONG. Doughnuts have holes, yes? This way, no holes. NOBODY gets a doughnut.

The correct approach is to slice the doughnut twice, parallel to its plane (see left). The problem is, exactly where do you position those two slices? A classical problem of calculus, but who knows calculus when you've got a knife in your hands and eyes staring at you? No, modeling is much simpler.

Here's a strategy. Consider just one of the two cross sectional circles of the doughnut. If you can slice that circle into three equal pieces by slices along the plane, that should do it. Where should the cuts go?

The solution is diagrammed to the right. I’ll model the cutting of the circle from the top, little by little, adding up the area of those slices, until the sum of the areas equals one-third of the area of the full circle. The point when that happens is the point where I should slice the doughnut to give one of the three equal pieces. The other cut should be at the equivalent position on the other side of the circle.

Some terms:
- \( r \) The radius of the circle
- \( h \) The distance from the center to the site of the slice
- \( w \) Half the width of the slice (the full width is the extended line to the other side of the circle). From the Pythagorean Theorum, \( w = \sqrt{r^2 - h^2} \).
- \( dh \) The thickness of the slices
- \( h_c \) The position of the last slice, the one that finally gives one-third of the circle.

So, I’ll start at the top (where \( w \) will be very small), adding up each slice (twice the width times the thickness of the slice), and I’ll let \( h \) go from \( r \) down towards zero. When the sum is equal to one-third the area of the circle, I’ll stop and note the value of \( h \). That value must be what I’m looking for, \( h_c \), the place I should cut the doughnut.

Write/steal a program that models the process described above and finds the value for \( h_c \).

(This is an example of numerical integration, common enough in computational biology, where the functions are often too complicated to integrate analytically)
12. Consider the following pathway,

\[
S + E \overset{k_f}{\rightleftharpoons} E_{\text{complex}} \overset{k_c}{\rightarrow} P + E
\]

12a. Derive the following equation for velocity:

\[
V = \frac{[S][E_{\text{total}}]k_c}{[S] + (k_c + k_r)/k_f}
\]

12b. Derive from this an equation for the maximal velocity.

"Enzyme complex" is somewhat of a black box. Let's look inside just a bit:

\[
S + E \overset{k_f}{\rightleftharpoons} E\cdot S_{\text{complex}} \overset{k_c}{\rightarrow} E\cdot P_{\text{complex}} \overset{k_c}{\rightarrow} P + E
\]

This model acknowledges that at some point during the enzymatically catalyzed reaction the substrate bound to the enzyme is transformed into the product, and this reaction (presumed to be reversible) has rate constants like any other reaction. You could derive an equation for velocity from this model in the same way you derived the equation in 11a and an equation for \(V_{\text{max}}\) in the same way you derived the equation in 11b,… but I'll give it to you:

\[
V_{\text{max}} = \frac{[E_{\text{total}}]k_{c2}}{((k_{r} + k_{c2})/k_{f}r) + 1)}
\]

12c. Suppose that the rates of interconversion of E\cdot S and E\cdot P are very fast compared to the rate of release of product from the enzyme. How then would you express \(V_{\text{max}}\)?

12d. Suppose that you, having no knowledge of black boxes, are trying to determine by actual experiment the value of \(k_c\) in a reaction. To do this, you measure the velocity of the reaction at very high substrate levels in the presence of known amounts of enzyme. The results are plotted below:

What is the value you calculate for \(k_c\)?

Suppose the enzyme operates as described in 11c (black box + fast interconversion). Then what quantity did you just calculate? What can you say about the rate at which the enzyme releases product?