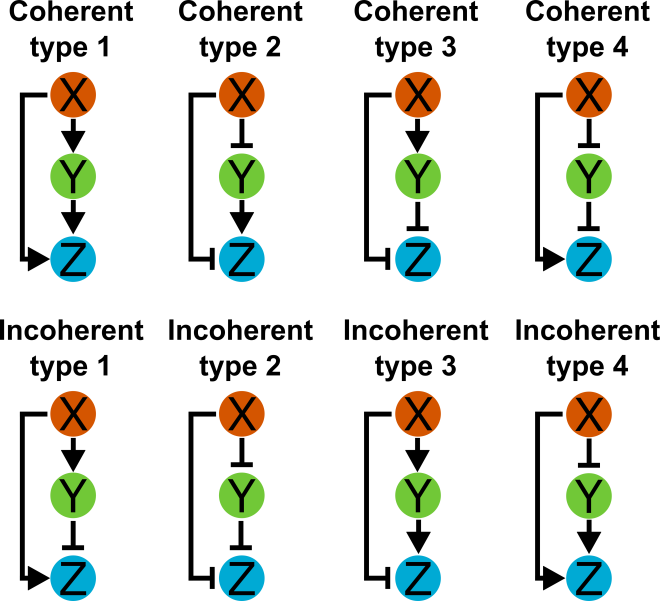
**PROBLEM 1 – Network motifs and qualitative/logic modeling**

Biological pathways, including signaling cascades and transcriptional regulation, consist of interactions between proteins, DNA, etc. whose behaviors are dictated by both the topology (i.e. structure) and the relative strengths of the interactions involved. An observed feature of large regulatory systems, for example yeast transcriptional regulatory networks, is that they are composed of small (3, 4, or 5 node) sub-modules called **network motifs**, which are recurring patterns that appear much more often than would be expected by chance. In this problem you will use **qualitative logical modeling** tools to analyze a variety of network motifs. Logical modeling methods apply theory from electrical engineering to biology by representing pathways as logical circuits. These incorporate logic gates (e.g. AND, OR) that describe how one or more inputs, either activating or inhibiting, induce a change in a target species (e.g. ON or OFF).

**Feed-forward loops** (FFLs) are common three-node network motifs that consist of a regulator X, which regulates species Y and Z, and where Z is also regulated by Y. There are eight possible feed-forward loops following this convention:

****

In this problem you will use a web-based qualitative biological modeling tool developed by Microsoft called **Bio Model Analyzer (BMA)** to create these eight FFLs and analyze their stabilities and steady-states. Visit <http://biomodelanalyzer.research.microsoft.com/> to access the tool.

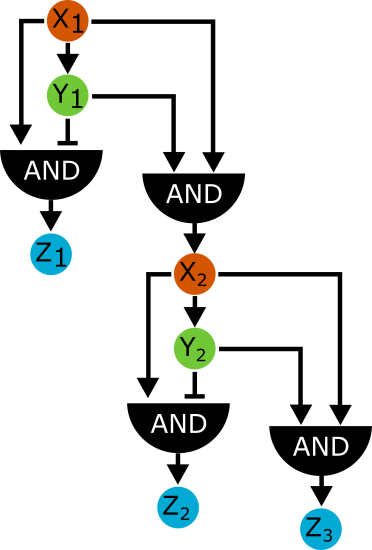
1. **Explore BMA**: At the top of the main BMA page, click on “Help” to access the F.A.Q. and example pages for the tool. Read through the F.A.Q. page to familiarize yourself with creating, editing, saving, and exporting models. Also pay attention to the basic target functions section.

Next, explore the two simple toy models in BMA (“Toy Example 1” and “Toy Example 2”) by clicking “Load model in tool” beneath the model headings. Click on the individual species to familiarize yourself with how they are defined and what parameters are associated (e.g. range of values, target function definition window, operators, etc.). Run the stabilization analysis tool (top window on right of page with a 2x2 grid of squares) and examine the output. Also explore the simulation tab (underneath the stabilization tool, image of a graph). Here you can examine time courses for the species in the model (bottom panel) and other variable information. Also in the simulation tab, expand the “variables” pane (click the box with the arrow pointing northwest). Here you will see that you can adjust the simulation parameters, including the number of time steps and the initial values for species. You can also export the time courses by clicking “export CSV.”

1. Use the BMA tool to create and analyze the steady-state species values for the eight FFLs listed above.
2. Right-click a species to access the parameters for each species. As you should have noticed from part (A), these parameters include the species name, the range, and the target function.
3. The ranges for all species should be set to [0, 1] (and they should be, by default).
4. Set the target function for all X species to 1 or equivalently, const(1). This initializes the cascade.
5. Set the target functions of the Y and Z species, using either AND logic or OR logic for the effect of X and Y on Z (see sub-questions). Model repression appropriately.

Hint: You can use one model pane to analyze all models in a single run by naming the species appropriately (for example, A, B, C for coherent type 1; D, E, F for coherent type 2 etc.) if you find it more convenient to do so.

1. Use AND logic for the interactions of X and Y on Z. A simple way to model AND logic is to use the minimum levels of the input functions.
   1. Turn in the steady-state results from the stability analysis. (You can export the image by clicking on ‘visual settings’). This should include the final values of each species. Do all the FFLs stabilize? Which of the eight FFLs reach a steady-state in which the output species Z is in the on state?
   2. Run the simulation tab to get the time courses for each species. In the simulation tab, if you expand variables, you can check that the initial values of all species have been set to 0 (and they should be, by default), and export the data to CSV if you wish. Turn in plots of time courses for all species in each FFL (one plot per FFL). Comment on how the time series of species Z compares to the final steady-state levels.
2. Change the logic on species Z to OR. To model OR logic, use the maximum of the two input levels. Repeat the stability and simulation analyses as in the previous question. Turn in the steady-state results and plot the time courses for all species in each FFL (one plot per FFL). Which FFLs change their steady-state behaviors following this change? Compare the results to the AND logic time courses from the previous question.
3. **Integrated FFL model**: Signaling pathways, transcriptional regulatory networks, etc. integrate these network motifs to produce more complex regulatory logic controlling cellular behaviors. The two most common of the eight FFLs you analyzed above are the coherent type 1 and incoherent type 1 motifs. The integrated FFL model below consists of two each of these FFL types and is a transcriptional regulatory network found in *Bacillus subtilis* that guides sporulation. The output species represent tens to hundreds of individual genes regulated by these processes.



In BMA, construct this integrated FFL model, report the steady-state values of Z1, Z2, and Z3, and export and plot the time series of the same three genes.

**PROBLEM 2** – **Feedback and dynamics**

Feedback loops are important regulatory elements that occur in both engineered and natural systems. Feedback occurs when an output of a system loops back and acts as an input, either as a positive or negative regulator. These types of systems can be particularly difficult to analyze without the help of models because simple logical reasoning on their behaviors is circular and because the relative strengths of feedbacks in the context of other circuit elements can determine the overall effect of such loops on the system as a whole. In this problem you will use qualitative logical modeling tools to analyze feedback circuits.

1. **Qualitative model of simple negative feedback circuit.** The simple negative feedback circuit below is implemented in BMA as “Toy Example 2.”



You should have noticed that this model does not stabilize in BMA due to the feedback loop. Export the output of a **20 step** BMA simulation of this model and describe the characteristics.

1. **Dynamic qualitative models**. The BMA model of the simple negative feedback loop represents a specific scenario where sustained oscillations in the participating species are observed. In more complex frameworks, time delays may be added to observe multiple states of the system to further analyze system behaviors. However, there many forms of negative feedback that these simple modeling frameworks often cannot capture, including time-delayed and threshold negative feedback which do occur in natural systems.

Another way of analyzing such network motifs is through dynamic analysis. Several methods exist to convert logical representations of networks to systems of qualitative differential equations. These methods also use simple activation/inhibition functions and logical operators (AND and OR) to describe the effects of multiple inputs on target species. These methods, however, use continuous transfer functions to describe the influence of a species on another. A common form of these transfer functions is the Hill function:

Where *X* is the influencing species, *act(X)* is the activating influence of *X* on an output species, *n* is the Hill coefficient, and *B* and *K* are constants which constrain the activation such that: *act(0)* = 0, *act(EC50)* = 0.5, and *act(1)* = 1, where EC50 is the fractional activation of the input required to induce half-maximal activation of the output. The Hill function can also be multiplied by a weight *w* to increases or decrease the influence of a particular input. Below are several plots of *act(X)* at various parameter settings for *w*, *n*, and *EC50*:

C:\Users\Anthony\Documents\MIT\20.440 Spring 2016 Psets\Mar7_PSet\activation_function_plots.emf

In the file run\_QDE.mwe provide the necessary code to create dynamic qualitative differential equation (QDE) models and provide an example implementation of the type 1 coherent FFL loop from earlier. There are four utility functions: act, which implements the activation Hill function above, inhib, which is 1-act, OR, which implements OR logic for two inputs (>2 inputs can be modeled with nested OR functions), and AND which implements AND logic on multiple inputs. Both the act and inhib functions take two inputs: the value of the activating species and a three-element parameter vector of the reaction weight, Hill coefficient *n*, and EC50.

From these four functions, you can generate QDEs for the type 1 coherent FFL with AND logic on Z as:

Each species has additional *ymax* and a *τ* parameters associated with them (set to 1 for simplicity, but can be manipulated to represent down/up-regulation or slowing/speeding of a reaction) and *stim* is a constant used to drive the activation of *X* in the cascade.

1. Simulate the **coherent type 1 FFL** using both AND and OR logic on output Z, applying stimulus (=1) for 10 steps, then removing (=0) for an additional 10 steps. Use as default reaction parameters for weight, *n*, and EC50 [1, 3, 0.5]. Describe the behavior of the two types of models in terms of when and how fast Z is activated/inactivated in relation of X and Y. How do these compare to the time courses you generated with BMA earlier?
2. Use the same code to implement the **incoherent type 1 FFL** as a dynamic model. You will need to use the inhib function to do this. Report time courses for all species using the same stimulus protocol above with both AND and OR logic. Describe the differences between the two and again compare to the time courses from earlier.
3. **Dynamic negative feedback loops.** You will now implement the simple negative feedback motif from part (A) as a QDE model.
4. Create the three-component negative feedback model using AND logic for the stimulus and feedback on species X and the same default reaction parameters for all reactions ([1, 3, 0.5]). Simulate for 50 time steps with the stimulus applied for the full time course. What is the behavior of the system at these settings?
5. Now you will model **fast negative feedback**, where the speed of the negative feedback is quick compared to the other activations. Set the weight parameter for the negative feedback reaction to 2 and plot the results. Describe the dynamics of the system at these settings.
6. Now model **slow negative feedback** by setting the weight of the negative feedback reaction to 0.3. Plot and describe the dynamics for this scenario.
7. Now model **threshold negative feedback** by adjusting the EC50 for the negative feedback reaction to 0.1 (low-threshold) and to 0.9 (high-threshold). Plot and describe the dynamics of these scenarios. Use all weights equal to one here.
8. **Dynamic positive feedback loops.** You will now implement simple positive feedback circuits as QDE models. For these simulations, increase all the *τ* values for the three species to 5, run the full simulation for 100 time steps, and apply the stimulus to X for 10 steps and remove for the remainder of the time course.
9. Run four simulations of the simple positive feedback loop shown below at four values of the positive feedback reaction weight: 0, 0.5, 0.75, and 1. Use OR logic for the stimulus and feedback reaction on X. Comment on the behavior of these four sets of time courses.



**PROBLEM 3**

1. For the first three spectra, using the manual validation procedure covered in class (and also described in the attached methods paper), manually validate the assigned peptide sequence and determine the site of phosphorylation. Each major peak should be labeled, and any unlabeled major peaks (e.g. those greater than 20% relative abundance) should be highlighted.
   1. For each peptide, figure out the identification of the protein that was phosphorylated and the residue number of the phosphorylation site.
   2. Briefly (1-3 sentences) describe the potential function of each phosphorylation site.
2. Using the 4th and 5th spectra (these are the same spectra, just displayed differently), identify the phosphorylated peptide by figuring out the sequence from the MS/MS spectrum.
   1. Briefly explain how you identified the peptide and protein. For instance, did you find a sequence tag? If so, what was it?