Read: Chapter 1: Mechanisms of gene regulation: Boolean network models of the lactose operon in *Escherichia coli*, by R. Robeva, B. Kirkwood, and R. Davis, pages 1–35.

Do: Create an account on either the Sage Notebook (`http://www.sagemath.org/`) or on the Clemson Sage Server (`https://sage.math.clemson.edu:34567`).

1. Recall the model of Stigler and Veliz-Cuba for the *lac* operon, where a global system state is a 9-varible Boolean vector $(M, P, B, C, R, A, A_\ell, L, L_\ell) = (x_1, x_2, \ldots, x_9) \in \mathbb{F}_2^9$:

$$\begin{array}{ll} f_M = \overline{R} \wedge C & \\ f_P = M & f_B = M \\ f_C = \overline{G_e} & f_R = \overline{A} \wedge \overline{A_\ell} \\ f_A = L \wedge B & f_{A_\ell} = A \vee L \vee L_\ell \\ f_L = \overline{G_e} \wedge P \wedge L_e & f_{L_\ell} = \overline{G_e} \wedge (L \vee L_e) \end{array}$$

   Give a well-written one sentence justification for each function. For example, $f_M = \overline{R} \wedge C$ could be: "mRNA is produced if the *lac* repressor protein is absent and the concentration of the catabolite activator protein (CAP) is high."

2. Consider the following system of polynomial equations:

$$\begin{aligned} x^2 + y^2 + xyz &= 1 \\ x^2 + y + z^2 &= 0 \\ x - z &= 0 \end{aligned}$$

   To compute a Gröbner basis for this system, type the following commands into Sage, one-by-one, and press Shift+Enter after each one:

```
P.<x,y,z> = PolynomialRing(RR, 3, order='lex'); P
I = ideal(x^2+y^2+xyz-1, x^2+y+z^2, x-z); I
B = I.groebner_basis(); B
```

   For the system above, use the Gröbner basis you computed to write a simpler systems of polynomial equations that has the same set of solutions. Sove that system *by hand* (it's not hard) to find all real solutions to the original system.

3. Repeat the steps of the previous problem for this system of polynomial equations:

$$\begin{aligned} x^2 y - z^3 &= 0 \\ 2xy - 4z &= 1 \\ z - y^2 &= 0 \\ x^3 - 4yz &= 0 \end{aligned}$$

4. Recall the first and most basic model of the *lac* operon that we saw:

$$f_M = \overline{G_e} \wedge (L \vee L_e), \qquad f_E = M, \qquad f_L = \overline{G_e} \wedge ((E \wedge L_e) \vee (L \wedge \overline{E})).$$

Propose a modified model by introducing a variable L_ℓ which denotes "*at least low levels of internal lactose.*" This model will have Boolean variables M, E, L, L_ℓ and parameters G_e, L_e.

(a) Write out each function as a Boolean polynomial.

(b) Use Gröbener bases and Sage to compute the fixed points of this model.

(c) Compute the entire phase space of your model with the help of the Analysis of Dynamic Algebraic Models (ADAM) toolbox, at `http://adam.plantsimlab.org/`. Use the Open Polynomial Dynamical System (oPDS) option to enter your model, and print your results for each of the four possibilities (G_e, L_e) of external glucose and lactose.

(d) Do the fixed points of this model make biological sense? Are you results biologically reasonable? Why or why not?

5. Go into Sage, and click "New Worksheet", give it a name, and then click "File" and "Load new worksheet from a file." There will be a box where you can enter the URL of an existing worksheet. Enter the following address, which is the file for the original 9-variable model of the *lac* operon:

`http://www.math.clemson.edu/~macaule/classes/s15_math4500/lac-operon.sws`

(a) In this 9-variable Boolean model, there is no variable to represent the cAMP receptor protein *cmp*. Could you justify this decision?

(b) Propose a new model with one additional variable C_{AMP} that represents *cmp*.

(c) Modify your Sage worksheet for your new model, and use Gröbener bases to compute the fixed points for each of the four possibilities (G_e, L_e) of external glucose and lactose. Are these fixed points biologically feasible?

(d) Use the ADAM software to determine if there are any periodic cycles that are not fixed points.

(e) Any final thoughts or comments? Do you notice any qualitative differences between this model and the original?