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

1. Consider the following simple model of the *lac* operon:

$$\begin{aligned} f_M &= \overline{R} & \qquad f_R &= \overline{A} \\ f_P &= M & \qquad f_A &= L \wedge B \\ f_B &= M & \qquad f_L &= P \end{aligned}$$

   For this problem, make the convention that $(x_1, x_2, x_3, x_4, x_5, x_6) = (M, P, B, R, A, L)$.

   (a) Justify each function in a single sentence. What other assumptions are made in this model? (E.g., presence or absense of external lactose and glucose?)

   (b) Sketch a (signed) wiring diagram for this model.

   (c) Write each function as a polynomial over $\mathbb{F}_2 = \{0, 1\}$. Then, write out the system of equations $\{f_i + x_i = 0,\ i = 1, \ldots, 6\}$, whose solutions are the fixed points of the Boolean network.

   (d) Compute the entire phase space of your model using the Cyclone program in AlgoRun, at `http://cyclone.algorun.org`. Include a print-out or screenshot. Describe each basin of attraction (connected component) by its size and limit cycle.

2. Consider a Boolean model of the *lac* operon, based on five variables: mRNA (M), β-galactosidase (B), allolactose (A), intracellular lactose (L), and *lac* permease (P), and the following transition functions:

$$\begin{aligned} f_M &= A \\ f_B &= M \\ f_A &= A \vee (L \wedge B) \\ f_L &= P \vee (L \wedge \overline{B}) \\ f_P &= M \end{aligned}$$

   This model does not have any parameters – it assumes that extracellular lactose is always available and extracellular glucose is always unavailable, and thus it is only able to describe the behavior of the system under the conditions.

   (a) As we saw in class, the dynamics do not accurately reflect the behavior of the biological system it is meant to model. Therefore, something is wrong. For each function, decide if it accurately reflects the underlying biology and/or the model assumptions.

   (b) Propose a modification of the transition functions aimed at eliminating the biologically infeasible fixed point. Give the rationale for your modification and specify the biological mechanism or model assumptions that justify the change.

   (c) Draw the wiring diagram and include print-out or screenshot of the phase space of your new model using AlgoRun.

   (d) Analyze your model. How many fixed points are there? Do they all correspond to biologically realistic situations? Note that there should be no limit cycles of size $k \geq 2$.

3. Consider the following model of the trpyptophanase (*tna*) operon in E. coli:

$$\begin{aligned} f_A &= M \wedge \overline{\gamma} \\ f_B &= M \\ f_C &= \overline{\gamma} \\ f_M &= C \wedge \overline{P} \\ f_P &= \overline{W} \wedge \overline{W}_m \\ f_W &= \omega_e \wedge B \\ f_{W_m} &= (\omega_{em} \wedge B) \vee \omega_e \vee W \end{aligned}$$

There are three parameters: γ is glucose, ω_e is high levels of extracellular tryptophan, and ω_{em} represents (at least) medium levels of extracellular tryptophan.

(a) Use Macaulay2 to convert these functions into polynomials over $\mathbb{F}_2 = \{0, 1\}$. Include the lines of code you used. Use $(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = (A, B, C, M, P, W, W_m)$.

(b) Use AlgoRun to compute the phase space when $(\gamma, \omega_e, \omega_{em}) = (0, 0, 1)$. Briefly summarize what you learn from this.

(c) For all three parameter vectors with $\gamma = 0$, use the BoolNet program in R to render the phase space. Downloading RStudio is recommended. For each one, make a table with the number of basins of attraction, and the size of the limit cycle of each. Include a screenshot of the phase space. (Everyone's in the class should look different because there is some randomness for how it arranges the nodes.)

(d) For each case where $\gamma = 0$, use BoolNet to find the attractors under an asynchronous update, and describe your findings.