Ascending biological counters

Matheus de Oliveira Saldanha[1], **Nícolas Goeldner**[1]

[1]Departamento de Informática e Estatística (INE) – Universidade Federal de Santa Catarina (UFSC)
Florianópolis – SC – Brazil

{matheus.saldanha,nicolas.goeldner}@grad.ufsc.br

Abstract. *The field of synthetic biology has expanded enormously for the past few years, mainly with the development of gene networks which can emulate digital circuits. One simple digital circuit that we tried to emulate is a counter. First, we present a gene network of an oscillator attached to AND gates that can count up to 6. For the second development, we took a completely different approach, creating the mathematical model of a biological switch that is essential to design a 3 bit counter. Both approaches have scalable models. Our simulations were in silico only and no biological implementation has been done.*

1. Introduction

An ascending counter, or simply a counter, is a circuit which stores and displays the number of times a particular event or process has occurred, often in relationship to a clock, i.e. an oscillator [A.K.Singh]. Creating a genetic counter is not something new in the field of synthetic biology. The work done by [Friedland et al. 2009] was fundamental in this area. Their counter can count up to three induction events, but it does not behave like a binary counter. This has the downside of the need to add a new biological component, e.g. a protein or a gene, to each new state to be added. In our paper, we model a more scalable genetic circuit in order to do not need a great amount of proteins or genes to build a counter with multiple states.

Counters are very related to oscillators due to the fact that an oscillator may be what provides stimulus to the counter. It always goes from a starting state, going through a set of state,s and then repeating this process again. With this in mind, we built in section 3 a biological counter using a oscillator as the main component, more specific the Repressilator [Elowitz and Leibler 2000]. It was chosen because it is a well replicated and studied work, leading to better chances of success when implementing the counter in a cell. We also used genetic AND gates in this approach to explore more states and benefit from it when the concentrations intersect, making it count up to six states. These genetic AND gates as well as the Repressilator are introduced in the related works section 2.

In the second counter, explained in details in section 4, we used the idea of a binary counter. A binary counter has several outputs, each of them with a different frequency. Taking a 3 bit binary counter for example, the MSB (Most Significant Bit) has a frequency f, as the middle bit has a frequency $2 \cdot f$ as the LSB (Least Significant Bit) has a frequency $4 \cdot f$, this behavior is shown in Figure 1. The Collins toggle switch [Gardner et al. 2000] was the prime component of the second counter as it gave us the foundations to implement the system. The preference of the Collins toggle switch was due to its simple implementation and capability to save data. Its simplicity scales well as we needed to use more than one switch in the 3 bit counter. Another fundamental factor

of this approach is the external input that enables the system to decide when to add one to the counting, going to the next state. The external factor is external to the circuit, i.e. it's not a protein or gene that is controlled inside the circuit. This factor could be a gene or a protein inside the cell that its concentration affects the circuit, e.g the increase of its concentration is direct related to proteins concentration inside the genetic circuit, leading to the next state. Due to the fact that we used the Collins toggle switch, we present it in the related works section 2.

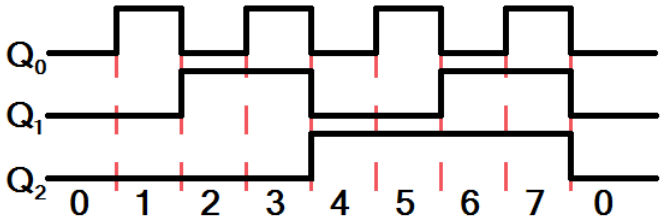

Figure 1. 3 bit counter timing diagram.

2. Related works

2.1. Repressilator

This genetic oscillator was first made by [Elowitz and Leibler 2000]. It uses three transcription repressors in the *E. coli*, where each protein represses the expression of the next gene (the proteins are: *TetR*, *LacI* and λ *cl*). This genetic circuit is illustrated by Figure 2.

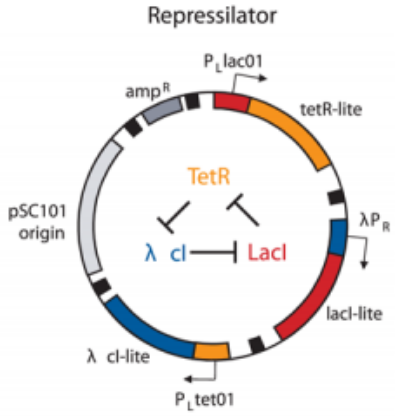

Figure 2. Repressilator genetic circuit
This image was taken from [Elowitz and Leibler 2000].

The main behavior of the system is described by the following differential equations, where p_i is the concentration of the protein i as m_i is the concentration of the mRNA i as a function of time:

$$\frac{dm_1(t)}{dt} = \alpha_0 + \frac{\alpha}{1 + p_3(t)^n} - m_1(t) \qquad \frac{dp_1(t)}{dt} = \beta \cdot m_1(t) - \beta \cdot p_1(t)$$

$$\frac{dm_2(t)}{dt} = \alpha_0 + \frac{\alpha}{1 + p_1(t)^n} - m_2(t) \qquad \frac{dp_2(t)}{dt} = \beta \cdot m_2(t) - \beta \cdot p_2(t)$$

$$\frac{dm_3(t)}{dt} = \alpha_0 + \frac{\alpha}{1 + p_2(t)^n} - m_3(t) \qquad \frac{dp_3(t)}{dt} = \beta \cdot m_3(t) - \beta \cdot p_3(t)$$

The parameter n is the Hill coefficient that represents the degree of cooperativity in binding to the DNA. α_0 is the transcription leak and α_0 + α is the maximal expression rate. Parameter β is the decay rate for the proteins [Ingalls 2013]. Time is rescaled in units of the mRNA lifetime. Protein concentrations are written in units of K_M, the number of repressors necessary to half-maximally repress a promoter; and mRNA concentrations are rescaled by their translation efficiency, the average number of proteins produced per mRNA molecule [Elowitz and Leibler 2000].

2.2. AND gates

There are several biological ways to build the AND logical operator [Sanassy et al. 2014]. Regarding gene regulatory networks in prokaryotes, the majority of gene regulation occurs through control of the initiation of transcription [Ingalls 2013]. Repressors and activators can thus be used as variables of a boolean function. Since boolean variables have only two values (true or false), we used the transcription factors concentrations and a threshold which is dependent of biochemical factors to represent the boolean value.

For example, if a gene g_1 is essentially expressed in the presence of the transcription factor p_1 and in the absence of p_2, logically it means that g_1 = p_1 AND (not p_2). This example is illustrated by Figures 3 (p_1 = false; p_2 = true); and 4 (p_1 = false; p_2 = true), both images were taken from [Ingalls 2013]. Yet, remains the cases where (p_1 = true; p_2 = true) and (p_1 = false; p_2 = false), in either of them no transcription occurs.

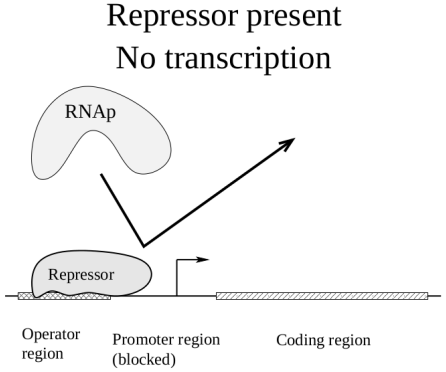

Figure 3. Biological AND gate

Here, the repressor p_2 is present and the activator p_1 is absent, thus, g_1 = p_1 AND (not p_2) =false AND (not true) = false AND false = false.

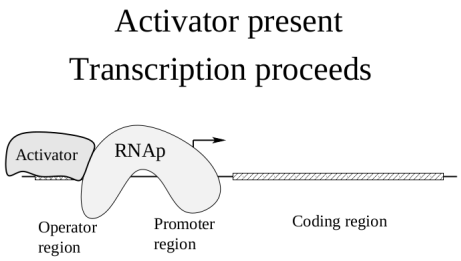

Figure 4. Biological AND gate

Here, the activator p_1 is present and the repressor p_2 is absent, thus, g_1 = p_1 AND (not p_2) = true AND (not false) = true AND true = true.

Genes are usually regulated by multiple trasncription factors and sometimes those factors have cooperativity when binding. Keeping our example of gene g_1, the fractions of the gene that may be transcripted without (1) and with (2) cooperativity are described below:

$$\frac{\frac{[p_1]}{K_{p_1}}}{1 + \frac{[p_1]}{K_{p_1}} + \frac{[p_2]}{K_{p_2}} + \frac{[p_1]\cdot[p_2]}{K_{p_1}\cdot K_{p_2}}} \quad (1) \qquad \frac{\frac{[p_1]}{K_{p_1}}}{1 + \frac{[p_1]}{K_{p_1}} + \frac{[p_2]}{K_{p_2}} + \frac{[p_1]\cdot[p_2]}{K_{p_1}\cdot K_{p_2}\cdot K_Q}} \quad (2)$$

where K_{p_i} is the dissociation constant for the binding event and K_Q is the multiplication factor for affinity. The promoter of the gene g_1 is supposed to have two non-overlapping operator sites and the binding events are considered to occur indifferently in any order. The association/disassociation between transcription factor and operator regions are on a much faster time-scale than gene expression, so it can be treated in equilibrium when modelling gene expression [Ingalls 2013].

2.3. The Collins toggle switch

The toggle switch is constructed from two repressible promoters arranged in a mutually inhibitory network, illustrated by Figure 5. It changes its states by transient chemical or thermal induction and exhibits a nearly ideal switching threshold [Gardner et al. 2000]. For our implementation the inducers are proteins. The main elements of this system are: repressors r_1 and r_2; promoter$_1$ and promoter$_2$; inducers i_1 and i_2.

In order to have the output of the r_2 high, the circuit needs a pulse of the input i_1 for sufficient amount of time to stabilize the system; If we want to turn off the output we may as well generate a pulse of i_2. A pulse of i_1 and a concomitant pulse of i_2 leads to undesired behaviors of the system.

The following differential equations describe these actions, where r_j is the concentration of the repressor j, β and γ indicate a degree of nonlinearity in the repression mechanisms and the i_j characterize the inducers [Gardner et al. 2000].

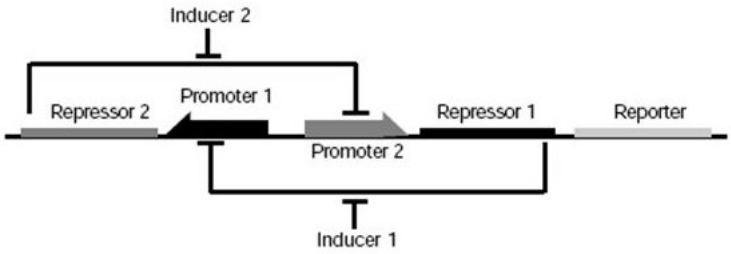

Figure 5. The Collins toggle switch
Image from [Gardner et al. 2000].

$$\frac{dr_1(t)}{dt} = \frac{\alpha_1}{1 + (\frac{r_2(t)}{1+i_2})^{\beta}} - r_1(t) \qquad \frac{dr_2(t)}{dt} = \frac{\alpha_2}{1 + (\frac{r_1(t)}{1+i_1})^{\gamma}} - r_1(t)$$

The functionality of this switch is very similar to the SR latch, a digital asynchronous bistable circuit [Ele 2012]. The SR latch has two inputs (S and R) that define the next state of the outputs (Q and Q'). The elements of the digital SR latch can be mapped to those of the toggle switch as follows: r_2 - Q; r_1 - Q'; i_1 - S; i_2 - R.

3. The repressilator counter

Our first biological counter uses the repressilator as its prime component. The main idea is to use the repressilator proteins concentration to point out states of our circuit. To create these states we used biological AND gates whose inputs are the proteins concentrations.

Another important aspect of this implementation is that it does not have an external signal, thus, when it starts to count it does not stop, acting as a clock with more than only two states.

3.1. Design of the genetic circuit

Analysing the repressilator plot, Figure 6, we realized that we could create states using, as explained previously, the proteins concentration.

The proteins concentration are used as the input of the AND gates, based on that we created a threshold, that can be seen in Figure 6 and was chosen considering where the concentrations intersect each other, e.g. state i: protein 1 is above this limit, protein 2 and 3 are below.

The AND gates receive one or two proteins that enforce their expression and one or two proteins that repress them, every AND gate depends on the three proteins of the repressilator. Notice that the three proteins concentration will never stay above or beneath the limit at the same time as they repress each other. Finally, this leave us to a total of six states, therefore, six AND gates, because the system needs one AND gate for each state, in our biological counter, that is modeled in Figure 7. The concentrations of the new proteins that are the outputs from the AND gates are shown in Figure 8 and Figure 9.

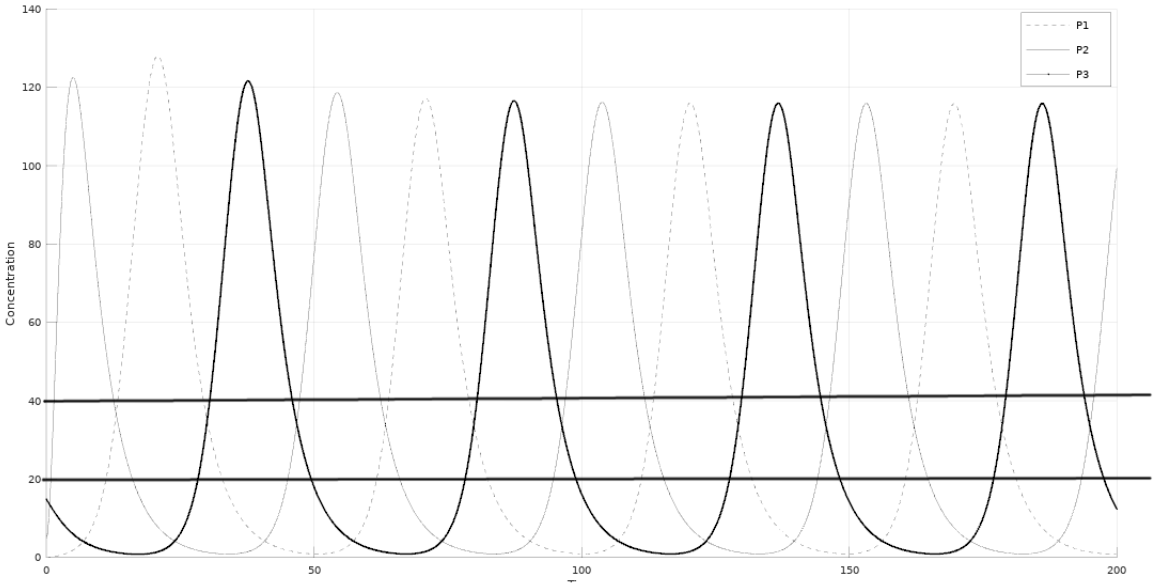

Figure 6. Repressilator plot

The parameters used to result in the above plot are: $\alpha_0 = 0.03$ (molecules per $cell \cdot min^{-1}$), $\alpha = 298.2$ $(molecules\ per\ cell \cdot min^{-1})$, $\beta = 0.2$ (min^{-1}) and $n = 2$. The initial concentrations of the mRNAs and the proteins are zero, regarding p_2 and p_3 that are 5 and 15 respectively.

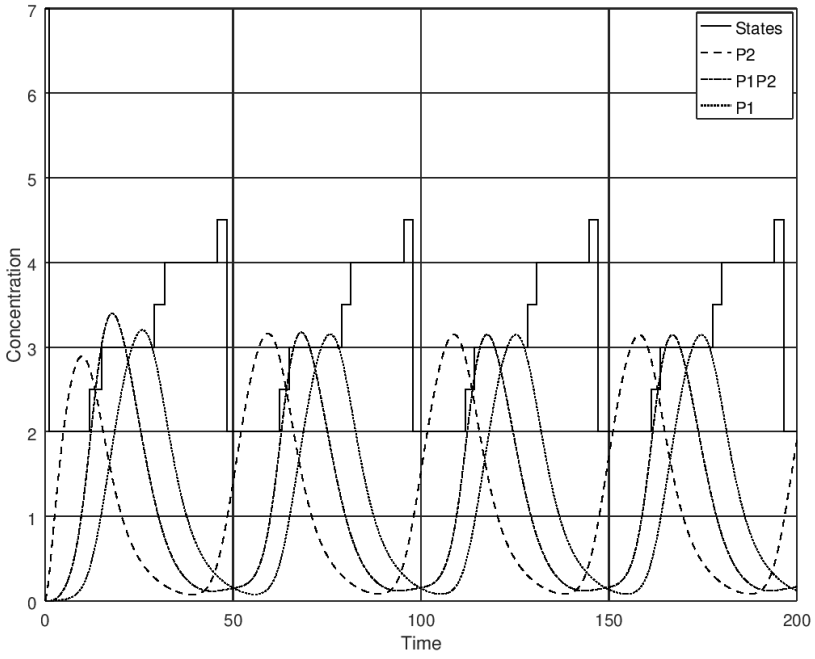

Figure 8. Repressilator with states.

The parameters used here are the same as Figure 6.

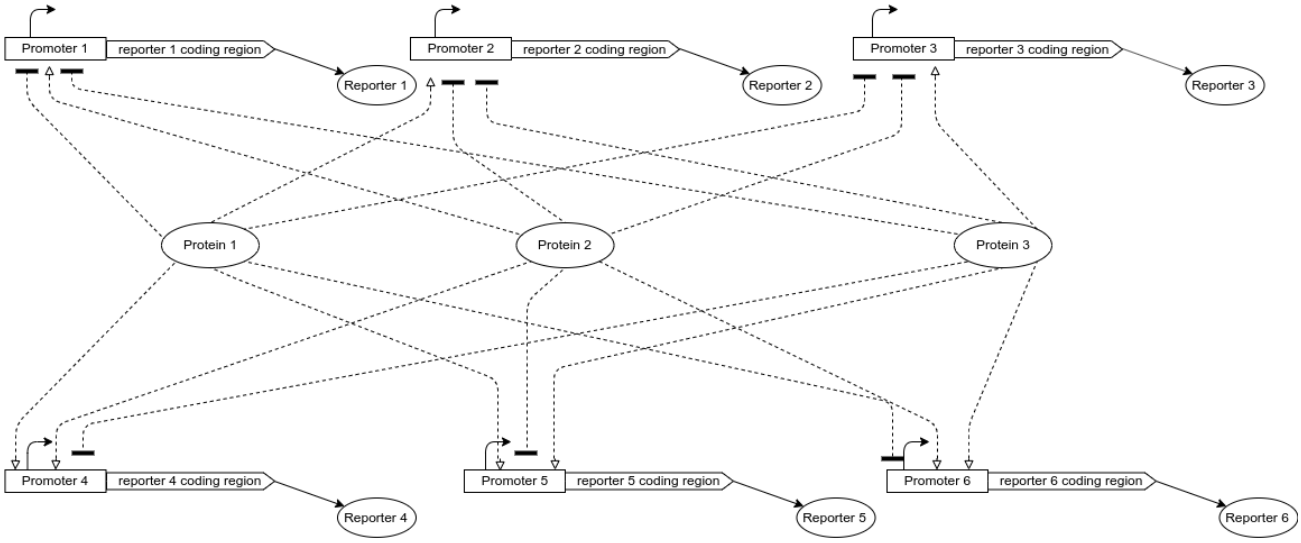

Figure 7. Repressilator AND gates.

The figure above shows how the three proteins are used as the inputs of the genetic AND gates.

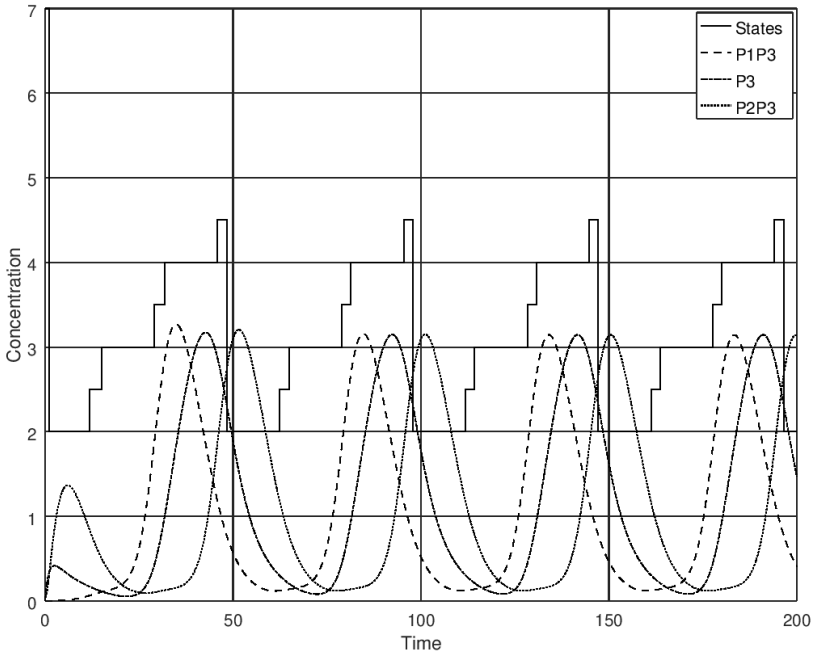

Figure 9. Repressilator with states.

The parameters used here are the same as Figure 6.

The Figure 8 and 9 show us the six possible states provided by the AND gates outputs. The curve P1P3 in Figure 9, for example, is in its peak when the protein$_1$ and protein$_3$ are in high concentrations and protein$_2$ is in low. This goes to all the other curves. When each of the curves is high, i.e. the output of the AND gate is in high concentration,

we have a new state. The states are represented by the ”ladder” States in the plot to give a distinguishable view.

To illustrate how this approach follows the idea of a counter we represent each output of an AND gate by a state number. Notice that the following example omits the proteins at low concentration. Starting from Figure 8, when protein$_2$ is in high concentration we have the state 0. When protein$_1$ and protein$_2$ are both in high concentration we have state 1. Following that is the state 2, when protein$_1$ is in high concentration. Analysing the Figure 9, we have the state 3 when both protein$_1$ and protein$_3$ are in high concentrations. The state 4 is described by the high concentration of protein$_3$ as the final state 5 is by the high concentration of protein$_2$ and protein$_3$. This behavior repeats itself multiple times since it has an oscillator as the main component.

4. The 3 bit counter

Binary counters are characterized by the fact that each bit$_i$ only changes its state when the previous bit$_{i-1}$ is switched off. This allows the counter to change its current state whenever the input event occurs. Thus, this implementation responds to external signals, however, it has limitations, naturally, the system needs time to stabilize, what requires the signals to have a minimum interval time between each two of them.

Here we first explain the one bit counter that is the toggle switch of one input. Then we explain how we connected three of them to form a three bit counter.

4.1. Design of the genetic circuit

4.1.1. The toggle switch of one input

We needed a system that has one input and when it receives a pulse, it changes its single output state, i.e. an one bit counter. Proceeding with our analogy to digital circuits, what we needed was a T flip-flop Figure 10 with the T input always high.

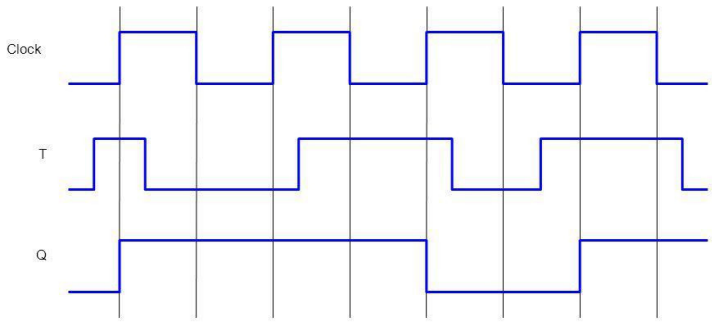

Figure 10. T Flip-Flip.

We merged another toggle switch with the previous one, resulting in Figure 11, this new attached component is equivalent to Figure 5, except by four differences:

(i) we exchanged r_1 with i_1 and r_2 with i_2;
(ii) promoter$_3$ is regulated by an AND between r_1 and (not r_2);

(iii) promoter$_4$ is regulated by an AND between r_2 and (not r_1);
(iv) we added a second gene after promoter$_3$ to be transcript along with i_1, called pulse$_2$.

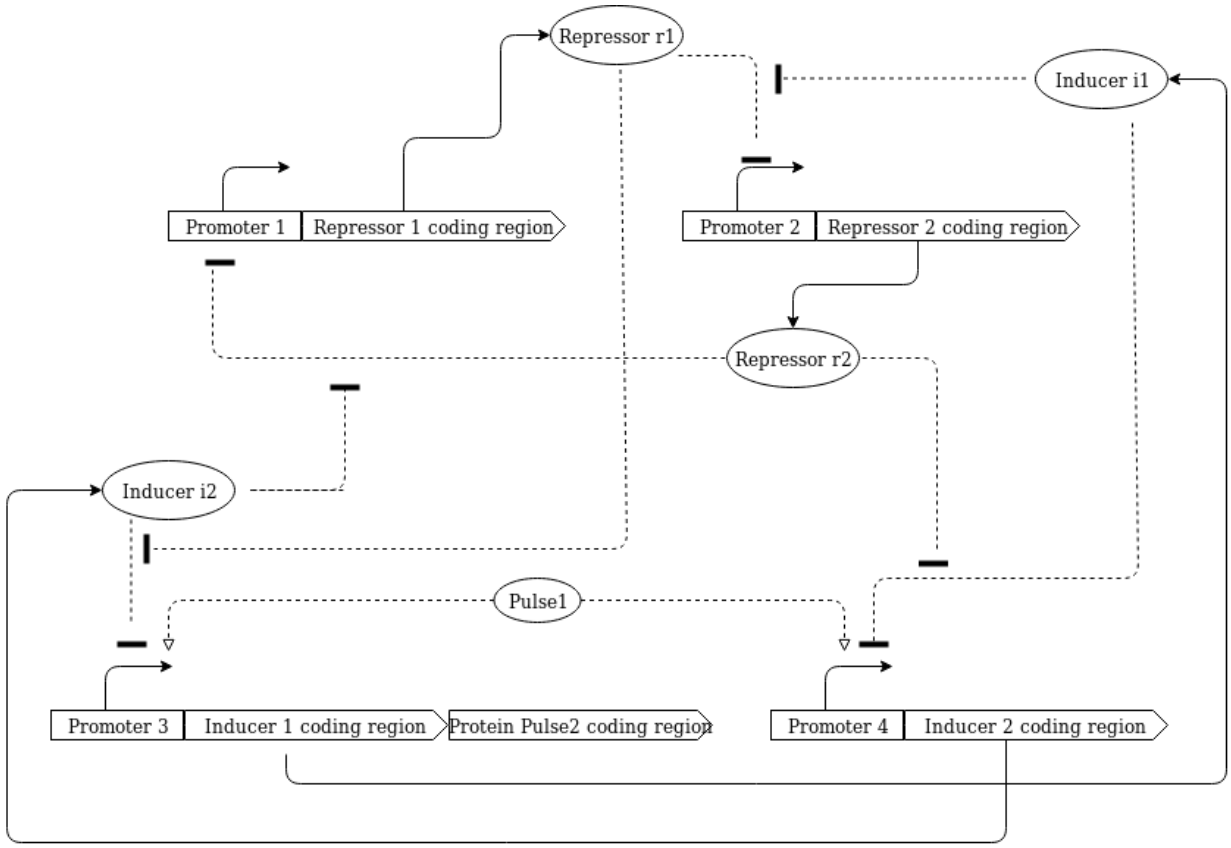

Figure 11. The toggle switch of one input model

Therefore, what regulates the system's overview is not the exogenous input of i_1 and i_2, but the single input that is represented by pulses of an activator of promoter$_3$ and promoter$_4$. This behavior is shown in Figure 12 and 13. When the system receives an exogenous pulse, the system itself generates an internal pulse of i_1 or i_2, it depends on which one is needed to change the concentration of r_1 and r_2. For example, the first pulse of Figure 12 and 13 generates an internal pulse of i_2 (Figure 13), which turns on the system by increasing the concentration of r_1 and decreasing the concentration of r_2. The mathematical model of this genetic network is:

Let

$$div = 1 + \frac{r_1(t)}{kp_1} + \frac{r_2(t)}{kp_2} + \frac{r_1(t) \cdot r_2(t)}{kp_1 \cdot kp_2}$$

$$\phi_1 = km_1 \cdot pulse \cdot \frac{\frac{1+r_1(t)}{kp_1}}{div} \qquad \phi_2 = km_2 \cdot pulse \cdot \frac{\frac{1+r_2(t)}{kp_2}}{div}$$

Then

$$\frac{di_1(t)}{dt} = \frac{\phi_1}{1 + (\frac{i_2(t)}{1+r_1(t)})^{\beta_2}} - i_1(t) \qquad \frac{di_2(t)}{dt} = \frac{\phi_2}{1 + (\frac{i_1(t)}{1+r_2(t)})^{\gamma_2}} - i_2(t)$$

$$\frac{dr_1(t)}{dt} = \frac{\alpha_1}{1 + (\frac{r_2(t)}{1+i_2(t)})^{\beta_1}} - r_1(t) \qquad \frac{dr_2(t)}{dt} = \frac{\alpha_2}{1 + (\frac{r_1(t)}{1+i_1(t)})^{\gamma_1}} - r_2(t)$$

ϕ_1 is the effective rate of synthesis of repressor$_1$. More precisely, the variation of ϕ_1 is the outcome of a genetic AND gate between r_1 AND (not r_2). i$_1$ is also proportional to the concentration of the pulse and to the constant km_1. The variation of concentration of i_1 is equivalent to r_1 as the above equations shows, except by the difference that α_1 is a fixed value representing the maximum expression rate [Ingalls 2013] and ϕ_1 is a dynamic value. ϕ_2 and i_2 are analogous to ϕ_1 and i_1 respectively. r_1 and r_2 behaviors have been already described in section 2.3. We still neglect the mRNA dynamics.

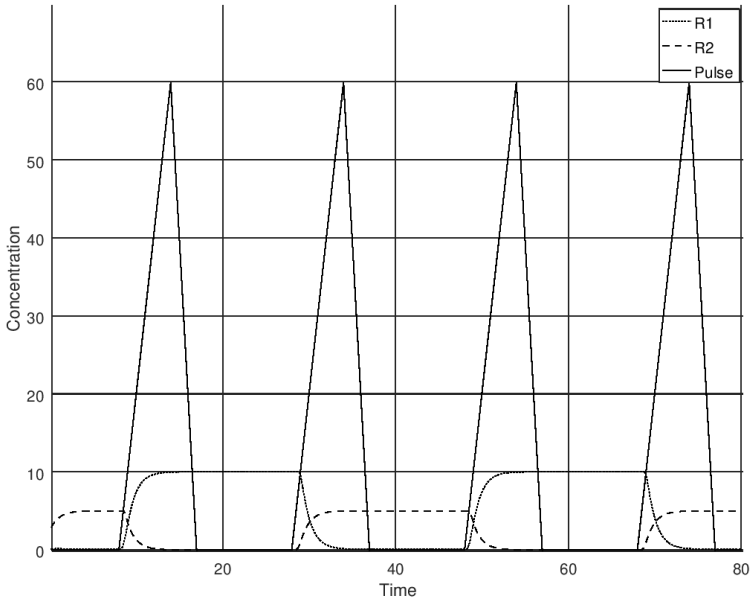

Figure 12. Toggle switch of one input plot.

The initial concentration of: r_1 is 0, r_2 is 2.5, i_1 and i_2 is 0. km_1 and km_2 are both equal to 4. kp_1 is 1 and kp_2 is 2.β_1 is 3, γ_1 is 4 β_2 and γ_2 are both equal to 4. And α_1 is 10, α_2 is 5.

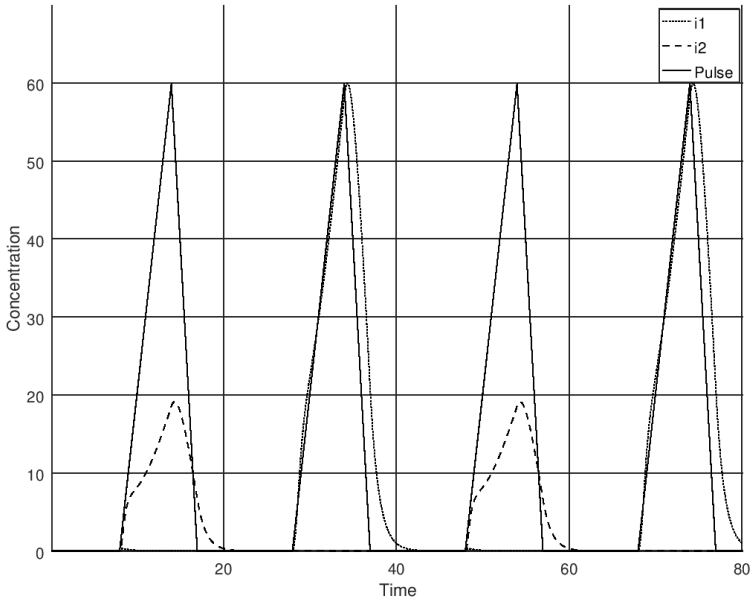

Figure 13. Toggle switch of one input

The parameters in this plot are the same as the above.

4.1.2. Increasing the number of bits

In order to develop a three bit counter, we first developed a two bit counter. Beginning with the two bit, we created two switches of one input each. The first switch, that controls the state of the less significant bit, is exactly what we described in the above section. The second switch, on the other hand, has the $pulse_2$ protein as the activator of promoter$_7$ and promoter$_8$. The $pulse_2$ concentration is proportional to i_1 concentration. The junction of these two switches enables the construction of the two bit counter.

The $pulse_2$ protein concentration increases every two pulses of the exogenous input, this action happens, consequently, when the concentration of the repressor r_1 is falling.

The same biochemical interaction is made between the bit$_1$ and bit$_2$ switches, Figure 14 illustrates the entire genetic system of the three bit counter. The result of the three bit counter is shown in Figure 15, r_1 represents bit$_0$, the less significant bit, also r_3 and r_5 represent bit$_1$ and bit$_2$ respectively. It is noticeable that each input pulse adds one to our previous state and when it reaches 111 its next state is 000.

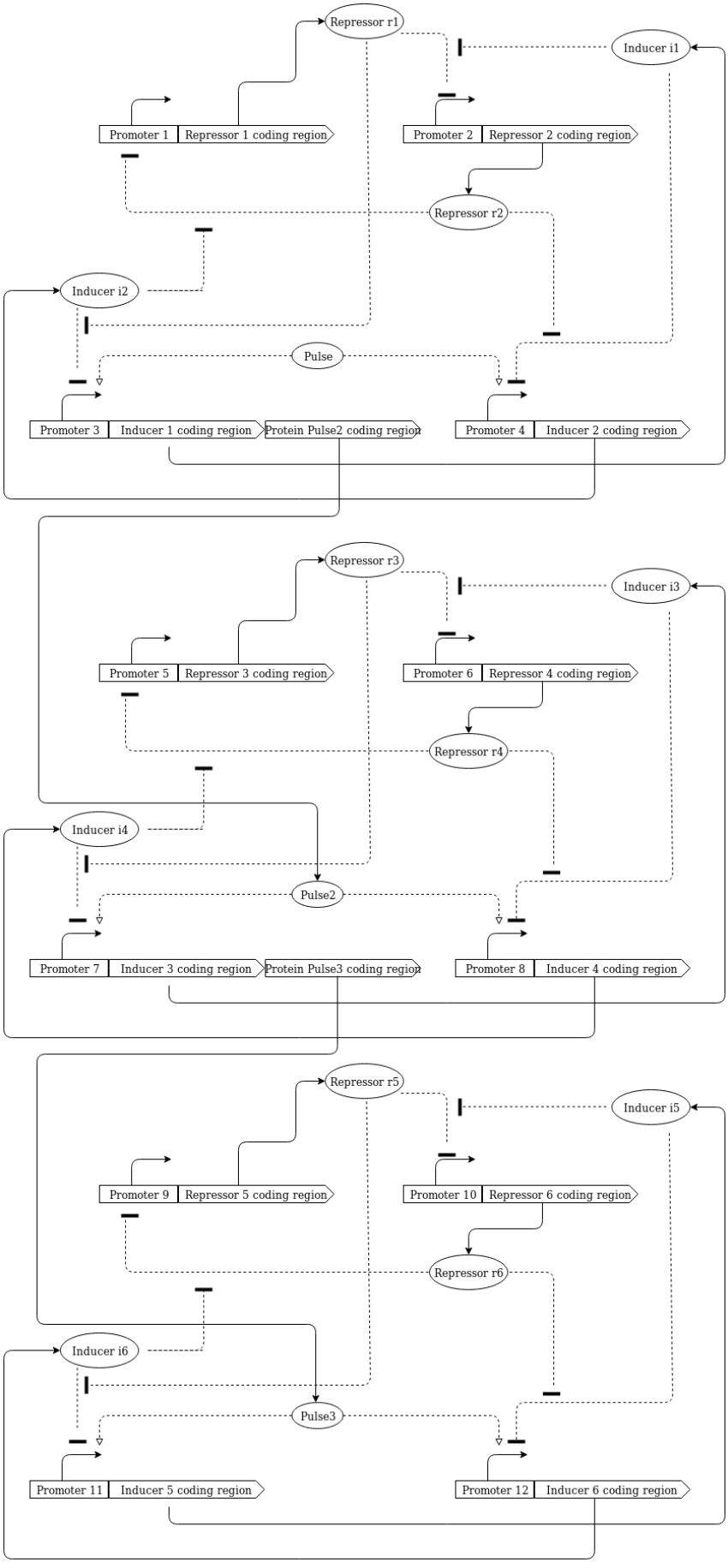

Figure 14. The complete 3 bit counter circuit

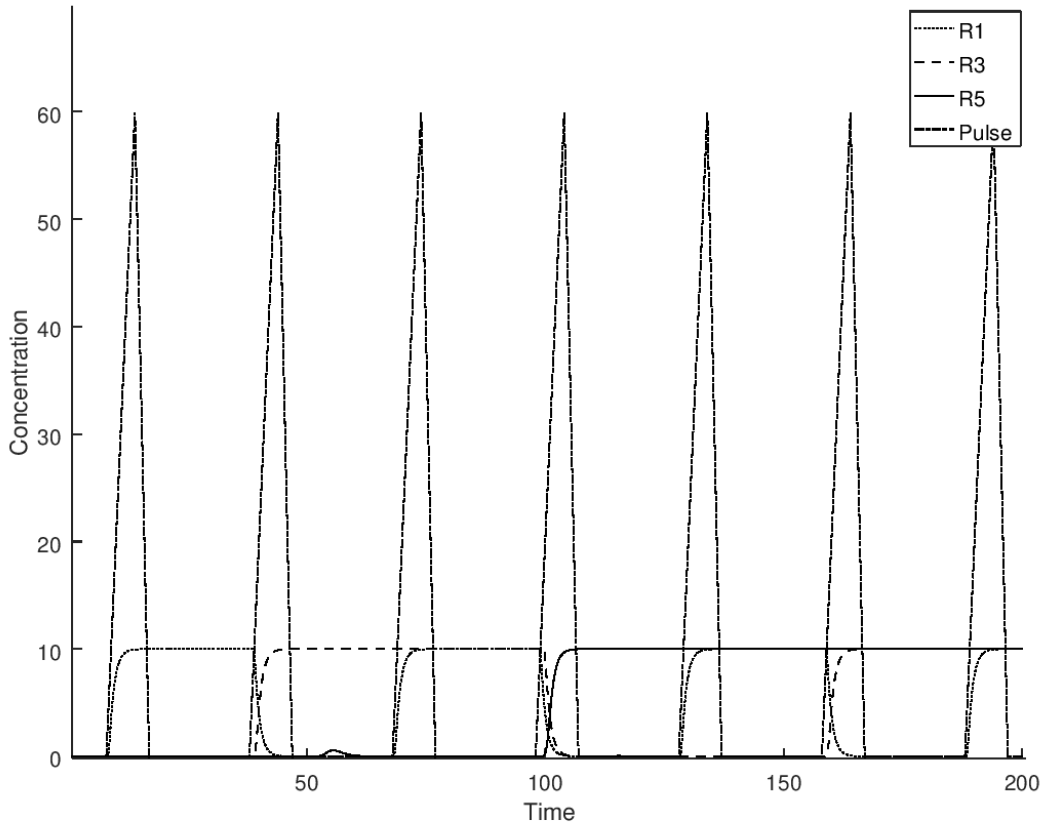

Figure 15. 3-bit counter.
The parameters are equivalent to the ones mentioned in this section.

5. Conclusion

This article establishes two main ideas in the process of developing genetic counters. The first one, implemented with the repressilator, uses AND gates to differentiate the six states, resulting in our counter. Note that this is simpler to build because it is made by a well studied oscillator as the prime component and it has fewer elements.

The second implementation main idea is to connect three switches of one input each in a way that they operate as a system of three bit counter. It also uses the AND gates, however, its most important feature is the capability to respond to an external input. This system is more complex and we focused on defining its mathematical model.

Regarding future works based on this paper, we hope to find a way to make the three bit counter biologically scalable. A good starting point we have in mind is to find some correspondent genetic circuits that implement our three bit model.

References

(2012). Elementos de Memória.

A.K.Singh. *Digital Principles Foundation of Circuit Design and Application*. New Age International Pvt Ltd Publishers.

Elowitz, M. B. and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338.

Friedland, A. E., Lu, T. K., Wang, X., Shi, D., Church, G., and Collins, J. J. (2009). Synthetic gene networks that count. *Science*, 324(5931):1199–1202.

Gardner, T. S., Cantor, C. R., and Collins, J. J. (2000). Construction of a genetic toggle switch in escherichia coli. *Nature*, 403(6767):339–342.

Ingalls, B. (2013). *Mathematical Modeling in Systems Biology*.

Sanassy, D., Fellermann, H., Krasnogor, N., Konur, S., Mierla, L. M., Gheorghe, M., Ladroue, C., and Kalvala, S. (2014). Modelling and stochastic simulation of synthetic biological boolean gates. In *2014 IEEE Intl Conf on High Performance Computing and Communications, 2014 IEEE 6th Intl Symp on Cyberspace Safety and Security, 2014 IEEE 11th Intl Conf on Embedded Software and Syst (HPCC, CSS, ICESS)*. IEEE.