Gillespie Algorithm Implementation

iGEM Costa Rica

October 20, 2021

1 Introduction

Our main motivation to do this document is to provide an easy-to-follow guide to implement the Gillespie Algorithm in your model. Sometimes the theory behind it could be perceived as hard to understand, so here we try to explain it in simple terms and provide as a example the equations we used for the ODE model. We hope some iGEM teams can find useful this document.

2 Step 1: Write down the differential equations of your system

The first step is to write your equations that depend on time, in this case we are going to use the ones that we reported in our wiki.

Following the central dogma of molecular biology, we can write down the following pair of ODE´s, describing the process of synthesis and degradation of a protein:

$$\frac{d[mRNA]}{dt} = K - \alpha_1[mRNA] \tag{1}$$

$$\frac{d[I]}{dt} = \beta_1[mRNA] - \alpha_2[I] \tag{2}$$

Where K is the promoter strength, α_1 the mRNA degradation constant, β_1 the protein synthesizing constant, α_2 the protein degradation constant and I is the integrase.

$$\frac{d[LRI_1]}{dt} = k_{+r}[PBI] - k_{-r1}[LRI_1] + k_{-syn}[LRI_2] - k_{+syn}[LRI_1] \tag{3}$$

$$\frac{d[PBI_1]}{dt} = k_{+r}[LRIR] - k_{-r2}[PBIR_1] + k_{-synr}[PBIR_2] - k_{+synr}[PBIR_1] \tag{4}$$

$$\frac{d[PB_{tot}]}{dt} = k_{-r1}[LRI_1] - k_{+r}[PBI] + k_{+r}[LRIR] - k_{-r2}[PBIR_1] \tag{5}$$

Independent reactions
$k_{+r}[PBI]$
$k_{-r1}[LRI_1]$
$k_{-syn}[LRI_2]$
$k_{+syn}[LRI_1]$
$k_{+r}[LRIR]$
$k_{-r2}[PBIR_1]$
$k_{-synr}[PBIR_2]$
$k_{+synr}[PBIR_1]$
K
$\alpha_1[mRNA]$
$\beta_1[mRNA]$
$\alpha_2[I]$

Table 1: Independent reactions in our system

3 Step 2: Identify independent events or terms across all the ODE´s

Some reactions may affect others, so some terms in the system of ODE´s may be repeated. In our system of 5 differential equations, there are 16 possible events that give rise to them, but only 12 of them are independent. For example, notice that the term $k_{-r1}[LRI_1]$ appears in equations 3 and 5, so if that reaction occurs it will affect both differential equations. Once we are done with this study write down all the independent events in a table. It is important to note that the sign of the reaction (plus or minus) it is not taken into account.

4 Step 3: Create event matrix

Now having identified the independent events in our system, we can construct the event matrix, that is basically how those events will affect the ODE´s if they are expressed.

Events	LRI1	PBIR1	PB_tot	mRNA	I
$k_{+r}[PBI]$	1	0	-1	0	0
$k_{-r1}[LRI_1]$	-1	0	1	0	0
$k_{-syn}[LRI_2]$	1	0	0	0	0
$k_{+syn}[LRI_1]$	-1	0	0	0	0
$k_{+r}[LRIR]$	0	1	1	0	0
$k_{-r2}[PBIR_1]$	0	-1	-1	0	0
$k_{-synr}[PBIR_2]$	0	1	0	0	0
$k_{+synr}[PBIR_1]$	0	-1	0	0	0
K	0	0	0	1	0
$\alpha_1[mRNA]$	0	0	0	-1	0
$\beta_1[mRNA]$	0	0	0	0	1
$\alpha_2[I]$	0	0	0	0	-1

Table 2: Event´s Matrix

The way to read table 2 is as follows, every row represent the independent reactions previously identified, and the columns represent the species for which the rate of change we described in the ODE´s. The way to fill the table is as follows: if the reaction or event occurs, you put a 1 in the columns that contain that term, and then aggregate the sign of the reaction depending on what it is described in the ODE´s.

5 Step 4: Identify initial conditions

Identify the initial quantity of molecules or cells for which you want to simulate the system. In our case that would be $[LRI1, PBIR1, PB\_tot, mRNA, I] = [0, 0, 50, 0, 0]$

6 Step 5: Choose two random numbers and iterate

The first number is to randomly pick the next reaction that will occur(a weighted choice based on the individual rates), each reaction has a probability associated with it, and the choice of next event is weighted by it. Think of this as if the value of every event is its associated probability, of course the events´values may be greater than 1, so we normalize them by dividing by the sum of all the events. Now imagine the probabilities as we put them one next to the other, so they form a chain from 0 to 1. Once we do this we just have to generate a random number between 0 and 1, see where it falls and pick the reaction it represents. MATLAB has a specific function to do this, it is randsample().

The second number is to pick when the next reaction is going to occur. We define it as: $\delta t = \frac{-lnu}{sum}$, where u is a uniformly distributed random number and sum is the sum of the possible events.

Next, adjust all the concentrations of the individual components to account for whatever reaction got chosen. For example if event 1 is chosen($k_{+r}[PBI]$), then based on the Event´s Matrix you have to update the initial concentration vector to sum 1 to LRI_1, subtract 1 to PB_{tot} and $PBIR_1$,$mRNA$ and I just leave it untouched, giving us $[1, 0, 49, 0, 0]$

As you can see some events depend directly on the time dependent species that we are calculating, so update the events´values because that would affect in choosing the next reaction to occur.

Iterate that from the starting conditions many times until a time criteria or any condition you want is met.

Find the implementation of the algorithm in MATLAB in our Github page, it is the file named GillespiePB.m