User Guide of AutoModel

October 16, 2014

Contents

1 What is this program? 2

2 How does it work? 2

3 How to use it? 2

3.1 Device 2

3.2 Network Document 2

3.3 Main Program 3

4 Demonstration 3

1 What is this program?

Synthetic biology has been developed for a decade and is being expected to evolve the academic society because of its standardization and convenience.[1,2] This novel technology has already yielded bunches of applications on the field of pharmacy[1,2] and therapeutics[3] while also showing its potential on energy industry.[5] However, a standardized model is eagerly required to be established because it will facilitate the exploration of synthetic biology. Therefore, we developed the software – AutoModel, which integrates dozens of models of simple devices and stipulates the I/O interface then provides user standardized modeling path with high degree of freedom.

This python script is being expected to help noviciate to construct their model and decide how much experimental data should be collected. In the coming section, we will demonstrate how this software works and how it help wet-lab experiments. Meanwhile, some instance are listed for illumination.

2 How does it work?

During the synthetic biological procedure, substantial biochemical reaction happens *in vivo* or *in vitro*. Generally speaking, those gene expression molecular level events could be represented by mass equation:[4]

$$k_1[A]_{aq} + k_2[B]_{aq} \rightleftharpoons k_3[A*B]_{aq}$$

where k_1, k_2 and k_3 are reaction coefficients. One of best mathematical representation for this reaction formula is differential equation. The differential equation is a mathematical equation about some continuously varying quantities and their rates of change in time that relates some function of one or more variables with its derivatives. Therefore, differential model is widely employed on the chemical modeling. However, constructing a precise model for any specific reaction is always complex because dozens of material and formulae has to be considered.

But now, a novel technique has been developed to help you figure out this problem. Inspired by the primary idea of synthetic biology, the AutoModel separates existed simple device as individual parts and integrates them into the device file which is a plug-in for main program. There are hundreds of mathematical expressions in the device file that represents those relationship between input(reactant) and output(resultant). Customers could add, adapt or amputate any functions, parameters and terms to optimally fitting their experimental data and theoretical model. Besides, preseted device functions could be called for predict your experimental result.

3 How to use it?

There are 3 kinds of documents included in the program package.

3.1 Device

All of differential equations are stored in the 'DEVICE.py' file which is a list of import subfunction of python. Those subfunction are defined in the same format that allowing user to recall them by the same method. The format is

```
def device's name (od, input_1, input_2, output, dt = 0.1):
    parameters
    input1 = input_1 + (differential equation of input_1)*od*dt
    input2 = input_2 + (differential equation of input_2)*od*dt
    output1 = output + (differential equation of output)*od*dt
    return (input1, input2, output)
```

where od is optical density and dt is integral time step.

3.2 Network Document

The 'network.txt' is the file of network which describes the connection among simulated model. This text file is a table. The first column is device's name which corresponds to the name in 'DEVICE.py'. The next 2 columns are input_1 and its initial value. It could be either inducer's name or device's name from first column. The 4th and 5th represent input_2. Last 2 columns are output's name and its value. If the word in any row is output, it means that the program will return the output value of this device.

3.3 Main Program

The main code for simulation is written in 'demo_python.py' which contains calculating function for growing curve, integrating function for ODE and a plotting part. Here we use an instance for illustration how to use this software.

First, there is a default growing curve provided by demo. The initial cell density(OD) is 0.2 while growing rate is 0.3 per 6 mins. You can change the maximum value of density by change the parameter k_1 in the function grow.

Second, function 'Calculate' is an integral function that calculating the output value while reading the network from network file from top to bottom and then loop this process until the end of simulation time. In this case, the demo network is invested by David L. Shis and Matthew R. Bennett.[6]

Final, 'Calculate' will return a matrix of time scalar, optical density and output to our plotting script.

4 Demonstration

It is reasonable to assume that our host is incubated in ideal environment that the curve of optical density (OD) perfectly matches the 'S' shape model. In other word, the growing curve is integrated by logistic model

$$\frac{dOD}{dt} = k_{grow} \times OD \times (1 - k_{max} \times OD)$$

where k_{grow} is producing rate of cells and k_{max} is maximun value of density of cells, i.e. $1/(OD_{max})$.

The AND gate with two devices achieved by Shis, David L., and Matthew R. Bennett[6] has been modeled for two separated parts that each one could be recalled as an individual function to represent the corresponding device. The ODEs for these 2 devices are demonstrated below.

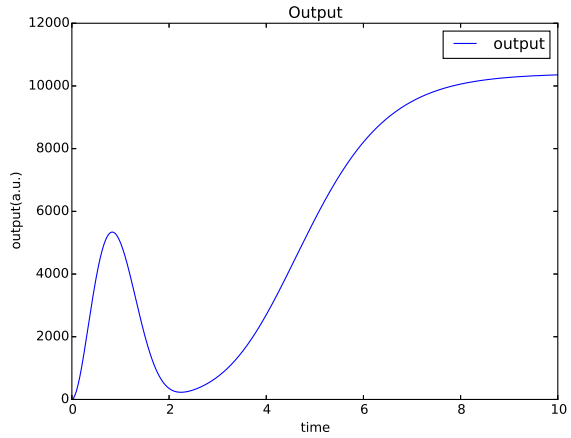
When pTara is induced by arabinose, this circuit produces LacI then stimulates its downstream device pET28. As result, the output of pTara is consumable inducer for circuit pET28. Therefore, we have 2 equations for pTara, which are

$$\frac{dx}{dt} = -20 \times x \times OD$$

and

$$\frac{dy}{dt} = (3 \times x + 5.2 \times OD^2 - 5 \times y) \times OD$$

where OD is optical density and x_1 is the input of device pTara, i.e., arabinose. Meanwhile, y is the output of device p-Tara and input of device pTara. The first equation represents the consumption of inducer that will give us an exponatial decay of the value of its input. The second equation is modeled for the output of pTara. The positive term corresponds to the producing rate of LacI whereas the negative term is of consuming rate.



As principle employed on modeling pTara, we also provide a differential equation for pET28 which is

$$\frac{dz}{dt} = (50000 \times y - 5 \times z) \times OD$$

where z is the output of device pET28, i.e., the value of GFP. Again the positive term is the producing rate of GFP but the negative term is of decomposion rate.

Finally, we link these 2 device to finalize the model of AND gate. The plot of the simulation result shows that our model approximately fittings the experimental data.

In this work, we successfully construct a new approaching of modeling for synthetic biology by applying the idea of this field which is about modulization. The modulized modeling functions which are using as bio-brick provide a novel method to academic society that allows circuit consturctors predict their experimental data. Furthermore, it gives us chance to develop the virtual laboratory so that dramatically reduces the laborious work in lab.

References

[1] Ellis, Tom, Xiao Wang, and James J. Collins. "Diversity-based, model-guided construction of synthetic gene networks with predicted functions." Nature biotechnology 27.5 (2009): 465-471.

[2] Andrianantoandro, Ernesto, et al. "Synthetic biology: new engineering rules for an emerging discipline." *Molecular systems biology* 2.1 (2006).

[3] Lu, Timothy K., and James J. Collins. "Engineered bacteriophage targeting gene networks as adjuvants for antibiotic therapy." *Proceedings of the National Academy of Sciences* 106.12 (2009): 4629-4634.

[4] Kaznessis, Yiannis N. "Models for synthetic biology." *BMC systems biology* 1.1 (2007): 47.

[5] Lee, Sung Kuk, et al. "Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels." *Current opinion in biotechnology* 19.6 (2008): 556-563.

[6] Shis, David L., and Matthew R. Bennett. "Library of synthetic transcriptional AND gates built with split T7 RNA polymerase mutants." *Proceedings of the National Academy of Sciences* 110.13 (2013): 5028-5033.