Continuous Modeling of Biomolecular Systems Based on Process Calculus

Yang Pu, Xiaoju Dong
BASICS Lab
Department of Computer Science and Engineering
Shanghai Jiaotong University, Shanghai 200240, China

Abstract

Process calculus is a kind of models for concurrent systems. Recent researches in Systems Biology have applied process calculus as the computational model to capture the dynamic behaviors of biomolecular systems. In this paper, we integrate a continuous framework with stochastic π-calculus to model the biomolecular systems. To verify the correctness of this approach, a modified stochastic Pi machine(SPiM) which was previously developed to simulate the systems described by stochastic π-calculus is proposed. From the consistency of the data obtained from the simulation of biomolecular system and the experimental data, it shows that the continuous framework introduced into Stochastic π-calculus is effective in simulating biomolecular systems.

1 introduction

Concurrency theory [1], especially the process calculus [2], has been used as a suitable tool to study Systems Biology [3, 4]. The main characters of biomolecular systems are interaction and the concurrence. The systems modeled by process calculus is the same as biomolecular systems to some extent. The main idea of using process calculus is to model molecular processes as interaction/communicating systems. That is to see biological components as concurrent processes and their interaction as process communication or process movement.

Using stochastic π-calculus [5] to formal model biological systems was first introduced by C.Priami [6]. The stochastic π-calculus enables its application to a wide variety of biomolecular systems in which quantitative aspects are key. In stochastic π-calculus, *kinetic constant* in biological reactions are abstracted as channels with base rates, and the actual reaction rate is calculated as an actual channel rate from the base rate and the number of processes offering communications. The selection of a time step and actual communication is based on these actual rates and follows Gillespie algorithm. Using the Gillespie algorithm [7], we can obtain the probability distribution for rules and times. The simulation of biomolecular processes which modeled by stochastic π-calculus is executed based on the Gillespie algorithm that calculates explicitly which reaction occurs next and how long it takes. But there are two problems in this simulation method. First, as we know, the biomolecular system is a concurrent system, all the reactions react independently. They are not sequential but parallel. But by using Gillespie algorithm, all the reactions occur one by one. It cannot describe the real biological system. Second, the most important benefit of process calculus is used to exhibit the concurrence of systems. By using Gillespie algorithm in the simulation, we cannot see the benefit of process calculus. Gillespie algorithm cannot simulate the systems described by process calculus accurately.

To address the above problems, we integrate continuous framework to simulate the biomolecular systems. This is inspired by the fact that, in vivo, biomolecular reactions evolve in a continuous way following a rate that depends on the concentration of the reactants. Therefore, we should deal with a non-integer number of processes in process calculus. Furthermore, we need to develop approximations in order to simulate the continuous reactions. Here we suppose time step $t_{i+1} - t_i$ is small enough to assume that the reaction rate and the concentrations of reactants remain constant. To address the concurrent problem, we use *Law of Mass Action* which states that the rate of a reaction is proportional to the product of the concentration of the reactants. And in every time step, the communications between processes are applied in parallel way according to the rate. In order to validate the efficiency of this continuous framework, we modified the stochastic Pi machine(SPiM) [8] to integrate the continuous framework, so called continuous SPiM. The original SPiM is a system used to simulate stochastic π-calculus based on Gillespie Algorithm. From the consistency of the data obtained from the simulation of Circadian Clock [9] by the continuous SPiM and the experimental data, it shows that the integration of this continuous framework with stochastic π-calculus is effective in modeling biomolecular systems.

The paper is organized as follows. The stochastic π-calculus used in this paper is introduced in the next section.

In section 3 the continuous simulation is described. We apply our approach to model Circadian Clock in Section 4. Finally, conclusions are given in section 5.

2 The Stochastic π-Calculus

In a biomolecular system, the molecules are abstracted as computational processes and the network of interacting molecules are abstracted as a mobile concurrent system in stochastic π-calculus. A complicated chemical process can always be decomposed into a set of many elementary bimolecular reactions, such as ($A + B \rightarrow \cdots$) or unimolecular reactions, such as ($A \rightarrow \cdots$). Two unimolecular reactions can be regarded as one bimolecular reactions, such as ($A + A \rightarrow \cdots$). Trimolecular reactions such as ($A + B + C \rightarrow \cdots$) are very rare. Therefore, the biochemical reactions between reactants can be abstracted as a communication between two channels with the same name. This kind of systems is composed by a community of co-existing computational process that communicate with each other and that change their interconnection structure at execution time. The stochastic π-calculus used in this paper is introduced in [8]. Here, we briefly introduce the syntax and the reduction of the stochastic π-calculus.

Definition 2.1 *The syntax of stochastic π-calculus is as follows*

$$\begin{aligned} P, Q ::= {} & \nu x P \quad \textit{Restriction} \\ & |P|Q \quad \textit{Parallel} \\ & \Sigma \quad \textit{Summation} \\ & \pi.P \quad \textit{Replication} \end{aligned}$$

where:

$$\begin{aligned} \Sigma ::= {} & 0 \quad \textit{Null} \\ & |\pi.P + \Sigma \quad \textit{Action} \end{aligned}$$

$$\begin{aligned} \pi ::= {} & x\langle n\rangle \quad \textit{Output} \\ & x(m) \quad \textit{Input}, x \neq m \end{aligned}$$

Definition 2.2 *The reduction of stochastic π-calculus is as follows, each channel x is associated with a corresponding reaction rate given by $rate(x)$:*

$$Q \equiv P \xrightarrow{r} P' \equiv Q' \quad \Rightarrow \quad Q \xrightarrow{r} Q'$$

$$P \xrightarrow{r} P' \quad \Rightarrow \quad \nu x P \xrightarrow{r} \nu x P'$$

$$P \xrightarrow{r} P' \quad \Rightarrow \quad P|Q \xrightarrow{r} P'|Q$$

$$x\langle n\rangle.P + \Sigma | x(m).Q + \Sigma' \xrightarrow{rate(x)} P|Q_{\{n/m\}}$$

According to Definition 2.1, the basic component is a summation Σ, which is a choice between zero or more output $x\langle n\rangle$or input $x(n)$ actions that the component can perform. $P|Q$ is parallel composition, and a given component P can contain a restricted reaction channel $\nu x P$. Replication $!\pi.P$ represents multiple copies of a given component $\pi.P$. π is either output or input actions. Two components in a biological system can react by performing complementary input and output actions on a common reaction channel. As show in Definition 2.2, summation containing an output $x\langle n\rangle.P$ can react with a parallel summation containing an input $x(m).Q$. The reaction occur with $rate(x)$, after which the name n is bound to m in process Q and processes P and $Q_{\{n/m\}}$ execute in parallel.

3 Continuous Simulation

A configuration of the biomolecular system described by stochastic π-calculus is a matrix of $\mathcal{M}_{n\times 2}(R^+)$ where $m_{i,1}$ represents the concentrations of interacting molecules which is described by input channel $x_i(m)$, while $m_{i,2}$ represents the concentrations of output channel $x_i\langle n\rangle$. The elements in the matrix are real numbers. An instantaneous configuration $E(t) = (m_{i,j}(t))_{1\leq i\leq n, j=1,2}$ with each instant $t \in R^+$.

To model the reactions we use the *Law of Mass Action* which states that the rate of a reaction is proportional to the product of the concentrations of the reactants. That is, if we have a reaction of the form $A + B \rightarrow \cdots$, then the rate of this reaction is $r = k \times |A| \times |B|$, for unimolecular reactions, such as ($A \rightarrow \cdots$), the rate is $r = \frac{1}{4} \times k \times |A|^2$, where k is called *kinetic constant*. In the stochastic π-calculus, we are using the rate of channel to represents *kinetic constant*, and the concentrations of the reactants can be obtained from the configuration $\mathcal{M}_{n\times 2}(R^+)$. In order to simulate evolution of biomolecular systems in computer, we need to develop approximations. Here for simplicity we use the rectangle rule; that is, we suppose $t_{l+1} - t_l = p$ is small enough to assume that the rate and the configuration remain constant. With this assumption we can approximate the effect of a communication during an interval of time of length p by $p \times r$. The implementation of continuous simulation based on stochastic π-calculus executes five steps as follows:

1. Initialize the configuration of the biomolecular system described by stochastic π-calculus
2. Initialize the absolute simulation time $t = 0$
3. Select out the communications which can be applied.
4. According to the rates of the communications, configuration and the small enough time interval p, calculate

the quantities of reactants and productions changed in each communication, and update the configuration.

5. Set the time $t = t + p$, and go to step 3.

4 Simulation of Circadian Clock

Circadian Clock phenomena are found in a large variety of organisms from cyanobacteria to mammals, and have probably evolved more than once. Recent works [9] suggest that the biomolecular mechanism of clocks shares common features over a wide range of organisms.The interaction consists of two interleaved feedback loops. In the positive loop, the activator element enhances its own expression. In the negative loop, the activator element enhances the expression of the negative element, which in turn sequesters the activator, as shown in Figure 1. The biological process involves two genes, an activator, A, and a repressor R, they are transcribed into mRNA and subsequently translated into protein. The activator A binds to the A and R promoters, and increases their basal transcription rates. Thus, A acts as the positive element in the system, whereas R acts as the negative element by sequestering the activator. This simple model is not intended to abstract any particular biomolecular system, but to capture the basic design principles shared by many systems, and believed to produce its basic functionality.

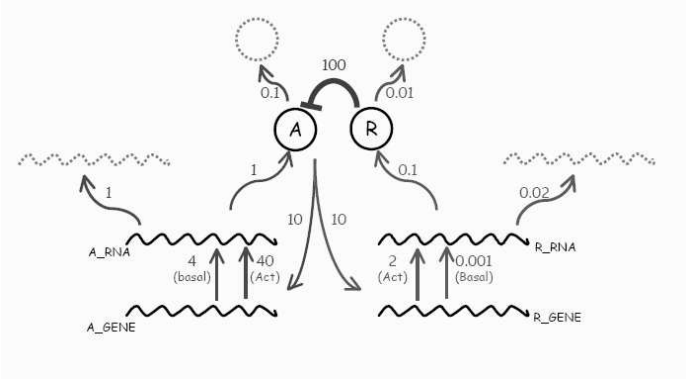

Figure 1: The network for a core Circadian Clock

We start by building a stochastic π-calculus abstraction of the Circadian Clock process. The reactions in the process are abstracted as communications on channels, and reaction rates as channel rates.

A-related process

$DNA\_A ::= tA().(DNA\_A|RNA\_A)|pA(u).DNA\_A2(u)$
$DNA\_A2(u) ::= tA'().(RNA\_A|DNA\_A2(u))|u\langle\rangle.DNA\_A$
$RNA\_A ::= trA().(RNA\_A|A)|drA()$
$A ::= pA\langle uA\rangle.uA().A|pR\langle uR\rangle.uR().A$
$\quad |\nu u(bind\langle u\rangle.A\_Bound(u))|dA()$
$A\_Bound(u) ::= dA().ReleaseR|u().A$
$ReleaseR ::= u\langle\rangle$

R-related process

$DNA\_R ::= tR().(DNA\_R|RNA\_R)|pR(u).DNA\_R2(u)$
$DNA\_R2(u) ::= tR'().(RNA\_R|DNA\_R2(u))|u\langle\rangle.DNA\_R$
$RNA\_R ::= trR().(RNA\_R|R)|drR()$
$A ::= bind\langle u\rangle.R\_Bound(u)|dR()$
$R\_Bound(u) ::= dR().ReleaseA|u().R$
$ReleaseA ::= u\langle\rangle$

In order to simulate our model abstracted by stochastic π-calculus we chosen the following basal channel rate: $rate(tA) = 4$, $rate(tR) = 0.001$, $rate(trA) = 1$, $rate(trR) = 0.1$, $rate(drA) = 1$, $rate(drR) = 0.02$, $rate(bind) = 100$, $rate(pA) = 10$, $rate(pR) = 10$, $rate(tA') = 40$, $rate(tR') = 2$, $rate(dA) = 0.1$, $rate(dR) = 0.01$. And we take the stochastic π-calculus abstraction of Circadian Clock as the input of continuous SPiM. The numbers of channels and processes obtained from continuous SPiM were plotted as a function of time and illustrate the oscillatory behavior.

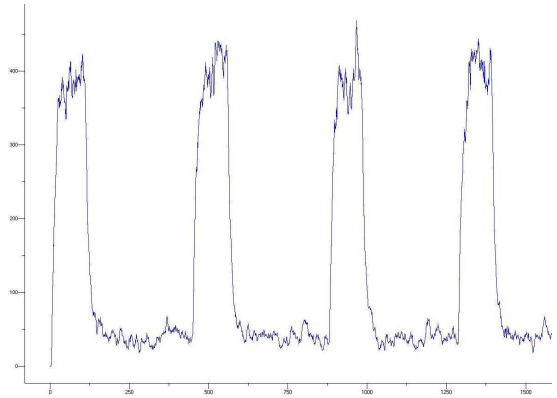

Figure2: The oscillatory behavior of $A\_Protein$

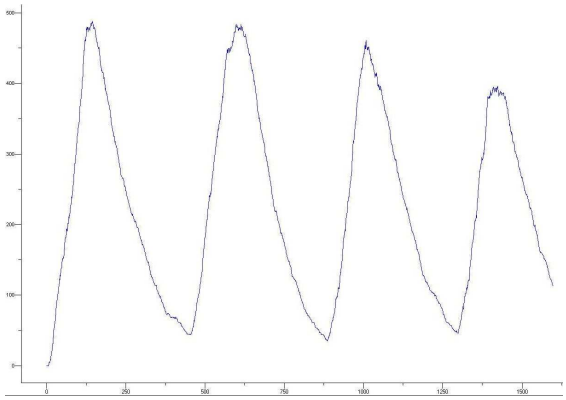

Figure3: The oscillatory behavior of $R\_Protein$

From Figure2, we can see the oscillatory behavior of A protein, and Figure3 exhibits us the oscillatory behavior of R protein. At first, the *kinetic constant* of A promoter's transcription is larger than R promoter's. With the accumulation of A proteins, it will activate A and R promoters, increase their transcription rates. It will result in more A and R proteins, meanwhile, because $R - A$ binding has the largest *kinetic constant*, more R proteins will bind to A proteins to repress the binding between A protein and A, R promoters, and decrease the number of A proteins. Small number of A proteins may lead to low transcription rates of A and R promoters because there will be few activated A and R promoters. The small *kinetic constant* of R promoters then leads to oscillations, which can be described as successive transitions between induced and repressed states.

As shown in Figure2 and Figure3, the results of the model yield the required oscillatory behaviors. In this we have reproduced the known result of [9], providing support for the correctness of the continuous abstraction framework based on stochastic π-calculus.

5 Conclusion

In this paper we integrate a continuous framework with the stochastic π-calculus. The numbers of channels and processes in the stochastic π-calculus are regarded as the real number to show continuous quantities of the substances. The communications in the stochastic π-calculus occurred in a parallel way in each time unit. This approach has been used to formal model the Circadian Clock. In order to validate our approach, we modified the Stochasitc Pi Machine(SPiM) to approximate the continuous simulation of biomolecular systems. From the continuous simulation of Circadian Clock, it can show us that the continuous framework of stochastic π-calculus is a reliable approach for simulating quantitive aspect of biomolecular systems.

Acknowledgements

This work is supported by The National Distinguished Young Scientist Fund of NNSFC (60225012), BDCC (03DZ14025), The National Nature Science Foundation of China (60473006), MSRA, and The BoShiDian Research Fund (20010248033).

References

[1] Milner R., Communication and Concurrency. *New York: Prentice Hall*, 1989.

[2] Milner R., Parrow J., Walker D., A Calculus of Mobile Processes. *Information and Computation*, 1992, 100(1): 1-40(Part I), 41-77 (Part II).

[3] Ciobanu G., Ciubotariu V., Tanasa B., A pi-Calculus Model of the Na Pump. *Genome Informatics*, 2002, 13: 469-471.

[4] A. Regev, W. Silverman, and E. Shapiro. Representation and simulation of biochemical processes using the pi- calculus process algebra. *In Pacific Symposium on Biocomputing*, volume 6, pages 459-470, Singapore, 2001. World Scientific Press.

[5] C.Priami, Stochasitc π-Calculus, *The Computer Journal*, 38:578-589, 1995.

[6] Priami C., Regev A., Shapiro E., Silverman W., Application of a Stochastic Name-passing Calculus to Representation and Simulation of Molecular Processes. *Information Processing Letters*, 2001, 80(1): 25-31.

[7] D.T.Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(25):2340−2361, 1977.

[8] Andrew Phillips and Luca Cardelli. A correct abstract machine for the stochastic pi-calculus. *In Bioconcur04. ENTCS*, August 2004.

[9] N.Barkai, Leibler S. Circadian clocks limited by noise. *Nature*, 403(6767):267−268, 2000.