Bio-calculus: Its Concept and Molecular Interaction

Masao Nagasaki [1235] masao@ims.u-tokyo.ac.jp
Shuichi Onami [15] sonami@symbio.jst.go.jp
Satoru Miyano [13] miyano@ims.u-tokyo.ac.jp
Hiroaki Kitano [14] kitano@symbio.jst.go.jp

[1] Systems Biology Group, Kitano Symbiotic Systems Project,
ERATO, JST, M31 6A, 6-31-15, Jingu-Mae, Shibuya-ku, Tokyo 150-0001, Japan
[2] Department of Information Science, University of Tokyo,
7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0034, Japan
[3] Human Genome Center, Institute of Medical Science, University of Tokyo,
4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
[4] Sony Computer Science Laboratory,
3-14-13 Higashi-Gotanda, Shinagawa, Tokyo 141 Japan
[5] These authors equally contributed to this paper.

Abstract

The way for expressing biological systems is a key element of usability. Expressions used in the biological society and those in the computer science society have their own merits. But they are too different for one society to utilize the expressions of the other society. In this paper, we design the *bio-calculus* that attempts to bridge this gap. We provide syntax which is similar to conventional expressions in biology and at the same time specifies information needed for simulation analysis. The information and mathematical background of bio-calculus is what is desired for the field of computer science. We show the practicality of *bio-calculus* by describing and simulating some molecular interactions with *bio-calculus*.

1 Introduction

Recent progress of molecular biology and genomics is rapidly producing enormous amounts of experimental results. It is becoming almost impossible to formulate any hypotheses that are consistent with all these results depending solely on the human power. To overcome this situation, technology of computer science is strongly required.

There is a big gap between biology and computer science when they communicate to investigate biological phenomena. One of the problems is the difference between their expression systems. The expressions of biology clarify the components (e.g. gene, protein, and cells). Their transition and interaction (e.g. gene activation, protein phosphorylation and cell division) are diagramed. In many cases, mathematical models for these processes are not defined in biology. Thus, expressions of biology lack information essential for simulation analysis (e.g. the velocity of a process, the quantity of a component) and do not have background of mathematics (e.g. calculus and probabilistic arguments).

Expressions of computer science are designed to include information essential for simulation since simulation is one of the main analytical methods in this field. The expressions also have mathematical background since system analysis is another important method. Several expression systems have been proposed for expressing biological phenomena (e.g. differential equation system [5, 7], stochastic method [4], and stochastic differential equation system [1]). In these systems, components of biological systems are mingled with mathematical symbolism. Transition and interaction of the components are far less visualized than in the diagramed expressions of biology. This is a critical weak point for biology since the components and their transition/communication are the main targets of biology.

To make a bridge between biology and computer science, we have started to construct *bio-calculus*. The *bio-calculus* is an expression system that i) clarifies components of biological systems, ii) diagrams

their transition and communication, iii) includes information essential for simulation, and iv) possesses mathematical background. In this paper, we present this concept of *bio-calculus*. Then we define *bio-calculus* for molecular interaction. Finally, we discuss five major advantages of *bio-calculus* by expressing simple examples of molecular interaction.

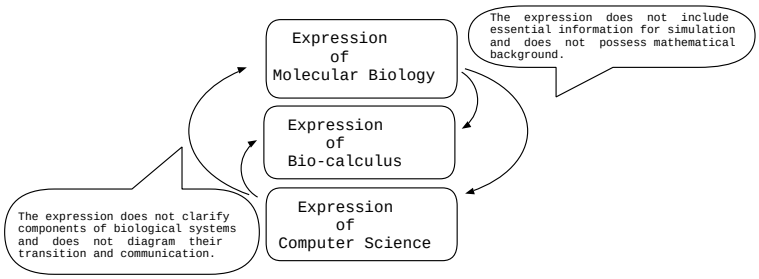


Figure 1: Problems of expression among biology and computer science. *Bio-calculus* solves those problems.

2 *Bio-calculus*

2.1 Requirements for *bio-calculus*

The expression of *bio-calculus* should satisfy the following requirements:

Req. 1 The expression clarifies components of biological systems.

Req. 2 The expression diagrams transition and communication of the components.

Req. 3 The expression includes information essential for simulation analysis.

Req. 4 The expression possesses background of mathematics.

2.2 *Multi-semantics system* for *bio-calculus*

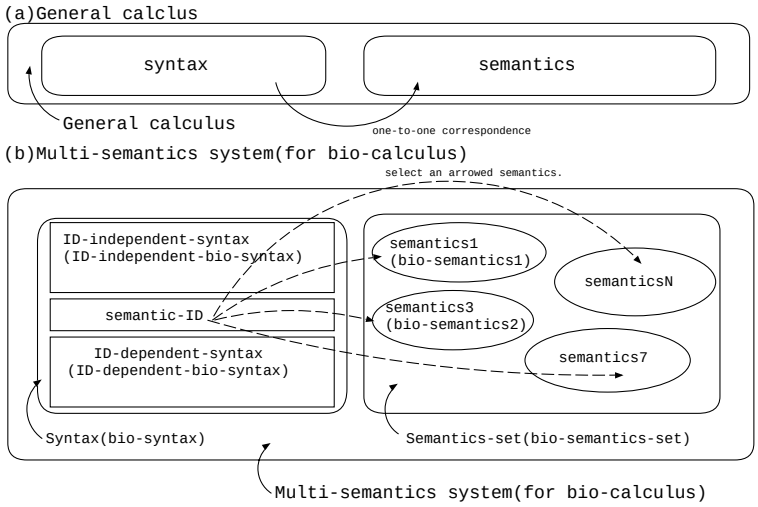


Figure 2: (a) A general calculus. (b) *Multi-semantics system* (for *bio-calculus*).

The distinguished feature of *bio-calculus* is *multi-semantics*.

To satisfy Req. 1 and Req. 2, the expression of *bio-calculus* is required to be quite similar to that of biology. Thus, we decided that in this calculus, each biological phenomenon is defined with the unique expression similar to that of biology. For Req. 3, we investigated several phenomena tring to define a single absolute simulation model for each phenominon. We then found that it is almost impossible to define such model for at least two reasons. Firstly, almost all phenomena are far from being completely understood. Thus, any explanation of a phenomenon will change. Secondly, there exist a huge number

of parameters - e.g. variety of molecule, pH and temperature. In many cases, some parameters are omitted for approximation. However, those parameters vary depending on what target is focused on. Accordingly, to fulfill Req. 3, each biological phenomenon should be defined by a unique expression with several explanations. This expression-explanation relationship is what *multi-semantics system* realizes in *bio-calculus*. To fulfill Req. 4, we define *bio-calculus* mathematically.

Bio-calculus is a calculus of *multi-semantics system*. Generally, a calculus consists of a syntax and a semantics corresponding to the syntax (Fig. 2(a)). Syntax is "the grammatical arrangement of words, showing their connection and relation". Semantics is "the evaluative rules of the language that is ruled by the syntax". In *multi-semantics system*, a syntax allows various kinds of semantics.

Bio-calculus consists of *bio-syntax* and *bio-semantics-set*. ***Bio-syntax*** is the syntax of this calculus, whose arrangement is designed to be quite similar to diagramed expressions of biology. ***Bio-semantics-set*** is a set of various kinds of *bio-semantics*. One of these *bio-semantics* is selected and applied to a *bio-syntax*. Thus, the semantics of a *bio-syntax* varies depending on the selected ***bio-syntax***. This syntax-semantics relationship is the special feature of *bio-calculus*. The relationship is quite analogous to the expression-explanation relationship in expressions for biological phenomena. Thus, it enables the calculus to fulfill Req. 1-4. We named a kind of system realizing this syntax-semantics relationship ***multi-semantics system***. We believe that this system is also quite useful for expressing other objects existing in the real world.

Selection of *bio-semantics* influences some parts of *bio-syntax*. So *bio-syntax* is divided into three parts, *ID-independent-bio-syntax, semantic-ID*, and *ID-dependent-bio-syntax*. ***ID-Independent-bio-syntax*** is the part of *bio-syntax* that is never influenced by selection of *semantic-ID*. This part contains ***template-rule***'s, which represent components of biological systems, and their transition and communication. The syntax of *template-rule* is designed to be quite similar to the expressions of biology. Thus, biological expressions for target phenomena are easily transported to the calculus and become available for computational analysis. The result of computational analysis under the calculus is quite familiar with biology. A ***semantic-ID*** specifies a *bio-semantics* for a *template-rule*. This is the special term that realizes *multi-semantics system* and makes the syntax of *template-rule* resemble the expressions of biology. ***ID-dependent-bio-syntax*** is the remaining part of *bio-syntax*. This part consists of constants for calculation. Contents of this part vary depending on the selection of *semantic-ID*.

Hence, all these parts are organized in this way and they function as a *multi-semantics system*.

2.3 Advantages of *bio-calculus*

Bio-calculus has many advantages. Some of them are directly related to Req. 1-4 and others are benefits of the *multi-semantics system*. In this section, we focus on five advantages that will strongly contribute to biology. The first two advantages are directly related to Req. 1-4. The others are the benefits of the *multi-semantics system*. In Section 4, each advantage is discussed with a simple example of molecular interaction.

Advantages

1. A biological expression is quite similar to a *template-rule* of *bio-calculus*. Thus the expression can be easily transported to *bio-calculus*. Then, the expression of *bio-calculus* is ready to be analyzed computationally.

2. An expression of *bio-calculus* includes information essential for computational analysis. Thus, the expression can be analyzed computationally. Results of computational analysis under this calculus are presented in the expressions quite familiar with biology.

3. A biological phenomenon can be analyzed in several models by simply changing the term *semantic-ID*. This is quite meaningful since simulation results sometimes differ among several models. The difference will give a hint for future investigation.

Table 1: Syntax of *bio-calculus* for molecular interaction.

S	:==	$def\ M_{set}$	*molecular-relation-set & initial-status-set*	
		$in\ A$		
M_{set}	:==	$M_1\ ,\ M_2$	*molecular-relation-set*	
	:==	M		
M	:==	$T\ k_v\ [semantic-ID]$	*molecular-relation*	
T	:==	$P_1 \Rightarrow P_2$	*template-relation*	*ID-independent-bio-syntax*
I	:==	$I_m\ ,\ I_e$	*initial-status-set*	
I_m	:==	B_1, B_2	*initial-molecular-status-set*	
	:==	B		
I_e	:==	E_1, E_2	*initial-environmental-status-set*	
	:==	E		
B	:==	$x(status)$	*initial-molecular-status*	
P	:==	$P_1 \mid P_2$	*molecule-set*	
	:==	R		
R	:==	x	*molecule*	
K	:==	$K_1\ ,\ K_2$	*coefficient-set*	
	:==	k_v		
$N \in \{x, y, u, v, ...\}$			*name*	*ID-independent-bio-syntax*
$semantic-ID$			*semantic-ID*	*ID-dependent-bio-syntax*
E			*initial-environmental-status*	*ID-dependent-bio-syntax*
$status$			*molecular-status*	*ID-dependent-bio-syntax*
k_v			*coefficient*	*ID-dependent-bio-syntax*

4. If the *information-level*[1] of *semanticsA* is higher than that of *semanticsB*, a model under *semanticsA* cannot be transformed into another one under *semanticsB* by merely changing the *semantic-ID*. This points out requirements for better understanding of the target biological phenomenon.

5. A model can be transformed into another model under a semantics whose *information-level* is lower than that of the original one. This enables us to select faster semantics at the cost of accuracy simply changing *semantic-ID*.

3 First Step of *Bio-calculus*: Definition of Molecular Interaction

For the first step of the development of *bio-calculus*, we discuss molecular interaction under this calculus. Molecular interaction is one of the basis of biological phenomena. At first, we define *bio-syntax* for molecular interaction. We then demonstrate several examples of *bio-semantics*. Finally, we discuss the advantages of *bio-calculus*, showing examples of molecular interaction.

3.1 *Bio-syntax* for molecular interaction

We define *bio-syntax* for molecular interaction as shown in Table 1. *Bio-syntax* for molecular interaction consists of two parts, *molecular-relation-set* and *initial-status-set*. *Molecular-relation-set* is defined between terms *def* and *in*, and describes relations among molecules (e.g. protein, mRNA,

[1] Information-level of semantics X($info_x$) is the amount of information (such as parameters) that is necessary for the expression of *bio-calculus* under semantics X. A relationship between two information-level is defined in partial order.

and amino acids). *Initial-status-set* describes initial status of those molecules (e.g. the number of molecules) and environment around them (e.g. pH and temperature). A ***molecular-relation-set*** is a set of ***molecular-relation***'s, each of which consists of a *template-relation*, a *semantic-ID* , and a *coefficient-set*.

A ***template-relation*** consists of two *molecule-set*'s, which are separated by term $\Rightarrow$. The left and right *molecule-set*'s are ***interactable-molecule-set*** and ***produced-molecule-set***, respectively. Each *molecule-set* consists of a *molecule* or *molecule*'s separated by term |. A *molecule* represents a biological molecule (e.g. protein, mRNA, and amino acids). A *template-relation* describes that *molecule*(s) in *interactable-molecule-set* react and produce *molecule*(s) in the *produced-molecule-set*. The ***semantic-ID*** of a *molecular-relation* determines the semantics under which the *molecular-relation* is executed. The ***coefficient-set*** of a *molecular-relation* is a set of ***coefficient***'s (e.g. reaction factor and reaction velocity) that are necessary to execute the *molecular-relation*.

An ***initial-status-set*** consists of an *initial-molecular-status-set* and an *initial-environmental-status-set*. An ***initial-molecular-status-set*** is a set of ***initial-molecular-status***'s. An *initial-molecular-status* consists of a *molecule* and a ***status***, and the *status* is described between parentheses (and). In an *initial-molecular-status*, the *status* defines the initial status (e.g. the number of molecule at the time point zero) of the *molecule*. An ***initial-environmental-status-set*** is a set of ***initial-environmental-status***'s, each of which represents the initial status of the environment (e.g. pH and temperature) for the molecular interaction defined by the *bio-syntax*.

Bio-calculus for molecular interaction is a *multi-semantics system*. Thus, its *bio-syntax* is divided into *ID-dependent-bio-syntax*, *semantic-ID*, and *ID-independent-bio-syntax*. *Initial-status-set* and *coefficient-set* are classified into *ID-dependent-bio-syntax*, whereas the others, except *semantic-ID*, are *ID-independent-bio-syntax*.

For the design of *template-rule*'s, we examined many expressions in biology. We also examined syntax of concurrency and communication calculi, π-calculus [6] and join-calculus [3] since molecular interaction is usually recognized as concurrency and communication. Thanks to the syntax of the *template-rule*'s, *bio-syntax* comes to be quite similar to expressions in biology and join-calculus [2].

3.2 Examples of *bio-semantics* for molecular interaction

We describe four *bio-semantics* for molecular interaction. Their mathematical backgrounds are confirmed by describing each of them in join-calculus.

3.2.1 Semantics1: stochastic *bio-semantics* (continuous version)

Under this *bio-semantics*, all biological molecules with the same name are represented by a *molecule* that holds a continuous value representing the number of the molecules. The reaction proceeds as follows: a *molecular-relation* is uniformly selected from executable *molecular-relation*'s. The selected *molecular-relation* is then executed. The details are as follows:

1. Initialize the 'current-molecule-set' to an empty set and the 'global-time' to zero. The value of each *molecule* is set to the value indicated in the term *status* in the *initial-molecular-status*. Add each *molecule* in the *initial-molecular-status-set* to the 'current-molecule-set'. If the value of some *molecule*'s is zero, remove the *molecule*'s from the 'current-molecule-set'.

2. Uniformly choose one *molecular-relation* from *molecular-relation*'s that are "executable"[3].

[2] Join-calculus and π-calculus has the same expressive power [3]. The syntax of join-calculus based on the chemical metaphor [2]. Thus, *bio-syntax* was designed to be similar to the syntax of join-calculus.

[3] We say that a *molecular-reaction* is "executable" if all molecules in the *interactable-molecule-set* exists in 'current-molecule-set'.

3. Execute the selected *molecular-relation*. As the result, the value of each *molecule* in the *interactable-molecule-set* and that in the *produced-molecule-set* decreases and increases, respectively. The variation corresponds to the product of i) the values each of which represents the original value of each *molecule* in the *interactable-molecule-set* and ii) the *coefficient* in the *molecular-relation*[4].

4. Increase the 'global-time'. The amount of increase is inversely proportional to the number of "executable" *molecular-reaction*'s.

5. If the value of a *molecule* in the *produced-molecule-set* changes from zero to a positive value, which means that a molecule appears, add the *molecule* to the 'current-molecule-set'. On the other hand, if the value of a *molecule* in the *interactable-molecule-set* becomes zero, delete the *molecule* from the 'current-molecule-set'. Then return to Step 2.

3.2.2 Semantics2: differential equation *bio-semantics*

Under *Semantics2*, the definition of *molecule* is the same as that under *Semantics1*. All *molecular-relation*'s are executed once at each time point. In short, under this *bio-semantics*, relations among molecules are executed as those defined under differential equations. The details are as follows:

1. Initialize the 'global-time' to zero. The value of each *molecule* is set to the value indicated in the term *status* in the *initial-molecular-status*.

2. All *molecular-relation*'s are executed. For each *molecular-relation*, the value of each *molecule* in the *interactable-molecule-set* and that in the *produced-molecule-set* decreases and increases, respectively. The variation corresponds to the product of i) the values each of which represents the value of each *molecule* in the *interactable-molecule-set* at the current 'global-time', and ii) the *coefficient* in the *molecular-relation*. Sum up all variations of each *molecule* in the *molecular-relation-set* and vary the value of the *molecule* at the current 'global-time' with the summed value.

3. Increase the 'global-time' by a constant value. Then return to Step 2.

3.2.3 Semantics3: Michaelis-Menten *bio-semantics*

Under this *bio-semantics*, the definition of *molecule* is the same as that under *Semantics1*. All *molecular-relation*'s are executed once at each time point. When a *molecular-relation* is executed, the change of the value of each *molecule* follows the Michaelis-Menten (M-M) equation [8]. The details are as follows:

1. The same as Step 1 in *Semantics2*.

2. All *molecular-relation*'s are executed. For each *molecular-relation*, the value of the *molecule*'s in the *interactable-molecule-set* and that in the *produced-molecule-set* decreases and increases, respectively. The variation follows the M-M equation applied to i) the values each of which represents the value of each *molecule* in the *interactable-molecule-set* at the current 'global-time' and ii) the two *coefficient*'s in the *molecular-relation*. Sum up all variations for each *molecule* in the *molecular-relation-set* and vary the value of the *molecule* at the current 'global-time' with the summed value.

3. The same as Step 3 in *Semantics2*.

[4]We do not use term "*coefficient-set*" but term "*coefficient*", because the semantics uses only one coefficient.

3.2.4 Semantics4: stochastic *bio-semantics* (discrete version)

Under *Semantics4*, all molecules are treated individually. The reaction proceeds as follows: From all existing molecules, uniformly choose two molecules. If the two molecules chosen can react, the reaction is executed with a predefined probability. Under *Semantics1* (Section 3.2.1), all molecules with the same name are represented by a *molecule*. Thus, the difference between *Semantics1* and *Semantics4* is the treatment of each molecule. The details are as follows:

1. Initialize the 'current-molecule-set' to an empty set and the 'global-time' to zero. For each *molecule*, generate the same number of elements as that indicated in the term *status* of its *initial-molecular-status*. Add those elements to the 'current-molecule-set'.
2. Uniformly choose two elements from the 'current-molecule-set'.
3. If the *molecule*(s) representing those elements are defined to react in a *molecular-relation*, select the *molecular-relation* and go to Step 4, otherwise go to Step 6.
4. Decide whether to execute the selected *molecular-relation*, depending on the *coefficient* of the *molecular-relation*. The *coefficient* represents the probability of the execution. With this probability, go to Step 5, otherwise go to Step 6.
5. Delete the two elements selected in Step 2 from the 'current-molecule-set'. Generate an element for each of all *molecule*'s in the *produced-molecule-set* and add all the generated element(s) to the 'current-molecule-set'.
6. Increase the 'global-time' by a constant value. Then return to Step 2.

3.3 Examples for molecular interaction

In Section 2.3, we have listed five major advantages of *bio-calculus*. Advantages 1 and 2 are directly related to the Req. 1-4, and Advantages 3-5 come from the *multi-semantics system*. In this section, we discuss those advantages of *bio-calculus*, showing simple examples of molecular interaction.

3.3.1 Advantage 1

As shown in Fig. 3, an expression of biology (Fig. 3(a)) is quite similar to a *template-rule* of *bio-calculus*. Thus the expression can be easily transported to *bio-calculus* (Fig. 3(b)). Then, the expression of *bio-calculus* can be analyzed by means of computer science (Fig. 3(c)).

3.3.2 Advantage 2

Fig. 4(a) shows an expression in computer science for a biological phenomenon. The same biological phenomenon can be expressed in *bio-calculus* (Fig. 4(b)). The *template-rule*'s of *bio-syntax* are easily transported to an expression of biology (Fig. 4(c)).

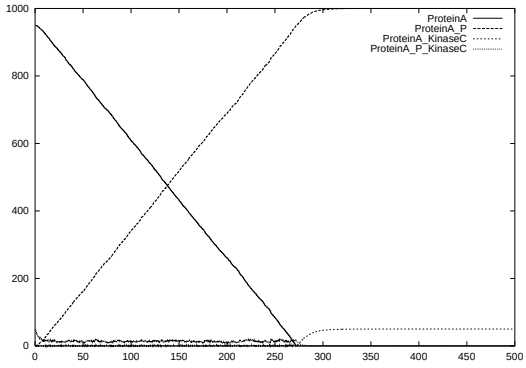
3.3.3 Advantages 3

We explain Advantage 3 using Fig. 5. Fig. 5(a) shows a *bio-calculus* expression of the mRNA transcription model described by Ko [4]. If the *semantic-ID* in this expression is **semantics2**, this expression is interpreted under *Semantics2* and if it is **semantics4**, the expression is interpreted under *Semantics4*. Thus, an expression of *bio-calculus* will be evaluated under several different semantics just by exchanging *semantic-ID*'s. It becomes an easy task to compare different semantics. If the executed results of different semantics are not the same, as in Fig. 5(b), the difference will give a hint for future investigation.

(a)

$$
\begin{array}{llcll}
\textit{let def} & & & & \\
 & ProteinA \,|\, KinaseC & \Rightarrow & ProteinA\_KinaseC & k_{A|C} \quad semantics1 \\
, & ProteinA\_KinaseC & \Rightarrow & ProteinA\_P \,|\, KinaseC & k_{A\_C\_1} \quad semantics1 \\
, & ProteinA\_KinaseC & \Rightarrow & ProteinA \,|\, KinaseC & k_{A\_C\_2} \quad semantics1 \\
\textit{in} & ProteinA\,(numA)\,,\, KinaseC\,(numC) & & &
\end{array}
$$

(b)



(c)

Figure 3: (a) A typical expression of biology for molecular interaction. KinaseC binds to proteinA and phosphorylate proteinA. (b) An expression of *bio-calculus* for the molecular interaction described in (a). (c) Executed results of the *bio-calculus* expression in (b). $k_{A|C} = 0.25$, $k_{A\_C\_1} = 0.10$, $k_{A\_C\_2} = 0.25$, $numA = 1000, numC = 50$

3.3.4 Advantage 4

The more detailed information your model contains, the more time the execution spends. To save the time, one can select other semantics under which the expression can be executed faster, at the cost of accuracy. *Semantics1* treats all molecules of the same name as a *molecule*. *Semantics4* treats those as individual elements. Thus, as shown in Fig. 6, *Semantics1* is faster than *Semantics4* by $O(n)$ (n is the number of all molecules in an expression of *bio-calculus*).

3.3.5 Advantage 5

Let us assume that the *information-level* of *SemanticsA* is higher than that of some *SemanticsB*. If one switches a *semantic-ID* in a *bio-calculus* expression from *SemanticsB* to *SemanticsA*, *bio-calculus* clearly show what information is necessary for the description under *SemanticsA*. An example is as follows. When one tries to translate an expression of *Semantics3* into that of *Semantics2*, each coefficient representing the M-M constant needs to be translated into two coefficients. This is because an M-M constant is the ratio of two coefficients for two simpler differential equations. Thus, when an expression under *Semantics3* is translated into that under *Semantics2*, the values of the two coefficients - new information - are required. This fact is easily recognized in *bio-calculus*. Owing to this advantage, we have already obtained an insightful biological result.[5]

[5]We have obtained an insightful biological result, by rewriting the paper [5]. This paper models the cell cycle of *Xenopus* embryo with differential equations which contain Michaelis-Mentens equation, which we translated to semantics1.

4 Conclusion

In this paper, we proposed *bio-calculus*, an expression system that bridges a gap between biology and computer science. At first, we proposed four requirements for such expression system. We proposed the design of *bio-calculus* that fulfills these requirements, then pointed out five major advantages of the calculus. For the first step of the development of *bio-calculus*, we presented *bio-syntax* for molecular interaction and presented four examples of *bio-semantics*. Finally, we discussed the five major advantages of *bio-calculus* showing some examples of molecular interaction.

The distinguished feature of *bio-calculus* is *multi-semantics*. In *bio-calculus*, the semantics of a *bio-syntax* varies depending on the selected *bio-semantics*. This syntax-semantics relationship is the most important point of *bio-calculus* since a biological phenomenon should be explained by several different explanations. We discussed that many advantages of *bio-calculus* come from this feature. We expected that other advantages will emerge from *multi-semantics system*. We believe that *multi-semantics system* can also be applied for expressing other phenomena existing in the real world.

Currently, we are developing a system that automatically generate the cell lineage from 4D microscope images of a *C. elegans* embryo [9]. We are planning to apply *bio-calculus* to the system and further extend the system in order to realize automatic gene interaction analysis using the enormous data produced by genomics.

For the first step of the development of *bio-calculus*, we defined the *bio-syntax* for molecular interaction since molecular interaction is one of the bases of biological phenomena. We are planning to extend the calculus for sub-cellular localization and cell-cell interaction to apply this calculus for the development of multi-cellular organisms.

References

[1] Arkin, A., Ross, J., and McAdams, H.H., Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells, *Genetics*, 149:1633–1648, 1998.

[2] Berry, G. and Boudol, G., The chemical abstract machine, *Theoretical Computer Science*, 96:217–248, 1992.

[3] Fournet, G. and Gonthier, G., The reflexive chemical abstract machine and the join-calculus, *Proceedings of the 23rd ACM Symposium on Principles of Programming Languages*, 372–385, 1996.

[4] Ko, M.S., A stochastic model for gene induction, *J. Theoretical Biology*, 153:181–194, 1991.

[5] Marlovits, G., Tyson, C.J., Novak, B., and Tyson, J.J., Modeling M-phase control in *Xenopus* oocyte extracts: the surveillance mechanism for unreplicated DNA, *Biophysical Chemistry*, 72:169–184, 1998.

[6] Milner, R., Parrow, J., and Walker, D., A calculus of mobile processes, parts I and II, *Information and Computation*, 100:1–40 & 44–77, 1992.

[7] Novak, B., Csikasz-Nagy, A., Gyorffy, B., Chen, K., and Tyson, J.J., Mathematical model of the *fission yeast* cell cycle with checkpoint controls at the G 1/S, G2/M and metaphase/anaphase transitions, *Biophysical Chemistry*, 72:185–200, 1998.

[8] Segel, I.H. (ed.), *Kinetics of Unireactant Enzymes in Enzyme Kinetics*, John Wiley & Sons Inc., New York, U. S. A., 1993.

[9] Yasuda, T., Bannai, H., Onami, S., Miyano, S., and Kitano, H., Towards automatic construction of cell-lineage of *C. elegans* from Nomarski DIC microscope images, *Genome Informatics 1999*, University Academy Press, 1999.

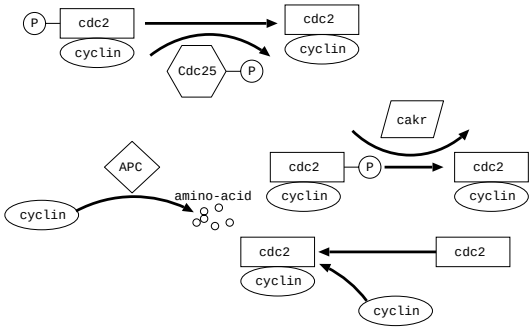
$$
\begin{aligned}
\tfrac{d}{dt}[cyclin] &= k_1 - [cyclin](V_2'[APC] + V_2''[APC^*] + k_3[cdc2]) \\
\tfrac{d}{dt}[cdc2\_cyclin] &= k_3[cyclin][cdc2] + k_{cakr}[cdc2\_p\_cyclin] + V_{25}'[cdc25][P\_cdc2\_Cyclin] + \\
&\quad V_{25}''[cdc25][P\_cdc2\_cyclin] + k_{ir}[CKI\_cdc2\_Cyclin] \\
&\quad -[cdc2\_cyclin]\{(V_2'[APC] + V_2''[APC^*]) + k_{cak} + \\
&\quad (V_{wee}'[Wee\_P] + V_{wee}''[Wee]) + k_i[cdc2]\} \\
&\ldots.
\end{aligned}
$$

(a)

let def

	$amino-acid$	$\Rightarrow$	$cyclin$	$k1$	$semantic-ID$
,	$cyclin \mid APC^*$	$\Rightarrow$	$amino-acid \mid APC^*$	V_2''	$semantic-ID$
,	$cdc2\_p\_cyclin \mid cakr$	$\Rightarrow$	$cdc2\_cyclin \mid cakr$	k_{cakr}	$semantic-ID$
,	$p\_cdc2\_cyclin \mid cdc25\_P$	$\Rightarrow$	$cdc2\_cyclin \mid cdc25\_P$	V_{25}''	$semantic-ID$
,	$cdc2\_cyclin \mid APC$	$\Rightarrow$	$amino-acid \mid APC$	V_2'	$semantic-ID$
,	$cdc2\_cyclin \mid cak$	$\Rightarrow$	$cdc2\_p\_cyclin \mid cak$	k_{cak}	$semantic-ID$
,	$cdc2\_cyclin \mid Wee$	$\Rightarrow$	$P\_cdc2\_cyclin \mid Wee$	V_{wee}''	$semantic-ID$
,	$cyclin \mid APC$	$\Rightarrow$	$amino-acid \mid APC$	V_2'	$semantic-ID$
,	$cyclin \mid cdc2$	$\Rightarrow$	$cdc2\_cyclin$	k_3	$semantic-ID$
,	$P\_cdc2\_cyclin \mid cdc25$	$\Rightarrow$	$cdc2\_cyclin \mid cdc25$	V_{25}'	$semantic-ID$
,	CKI_cdc2_cyclin	$\Rightarrow$	$CKI \mid cdc2\_cyclin$	k_{ir}	$semantic-ID$
,	$cdc2\_cyclin \mid APC^*$	$\Rightarrow$	$cdc2\_cyclin \mid APC^*$	V_2''	$semantic-ID$
,	$cdc2\_cyclin \mid WeeP$	$\Rightarrow$	$P\_cdc2\_cyclin \mid WeeP$	V_{wee}'	$semantic-ID$
,	$cdc2\_cyclin \mid CKI$	$\Rightarrow$	CKI_cdc2_cyclin	k_i	$semantic-ID$
,	...				

(b)

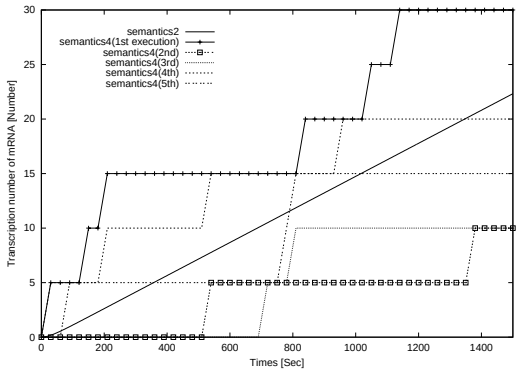


(c)

Figure 4: (a) Differential equations demonstrated in Novak *et al.* [5]. They used the differential equations for modeling the cell-cycle in *Xenopus* embryo. *cyclin*, APC, APC^*, *cdc2_cyclin*, *cdc2*, Wee, Wee_P, ... means proteins and k_1, k_{cak}, k_{cakr}, V_{25}', V_{25}'', V_2', V_2'', V_{wee}', V_{wee}'', k_i ... means coefficients. Only two equations are shown for saving the space. (b) A *bio-calculus* expression corresponding to the above simulation model. (a). Only several *molecular-relation*s are shown. (c) A typical expression of biology representing several molecular reactions expressed in the above *bio-calculus* expression (b).

$$
\begin{array}{l}
let\ def \\
\qquad \begin{array}{llll}
DNA \,|\, unbound & \Rightarrow & bound \,|\, mRNA & P1 \quad semantic-ID \\
DNA \,|\, unbound & \Rightarrow & DNA \,|\, unbound & 1-P1 \quad semantic-ID \\
Q \,|\, bound & \Rightarrow & DNA \,|\, Q \,|\, unbound & P2 \quad semantic-ID \\
polII \,|\, bound & \Rightarrow & polII \,|\, bound \,|\, mRNA & 1-P2 \quad semantic-ID
\end{array} \\
\qquad in\ DNA\,(1)\,,\ polII\,(1)\,,\ Q\,(1)\,,\ polII\,(1)\,,\ unbound\,(1)
\end{array}
$$

(a)



(b)

Figure 5: (a) An expression of *bio-calculus* representing the model of mRNA transcription described by Ko [4]. ($P1 = 0.90\,,\ P2 = 0.0005\,,\ step\,number = 51, mRNA\,transcription\,number\,for\,each\,step = 1$). (b) The execution results of the *bio-calculus* description in (a). The solid line is the result of the execution under semantics2. Each of other lines represents each result of five independent executions under semantics4.

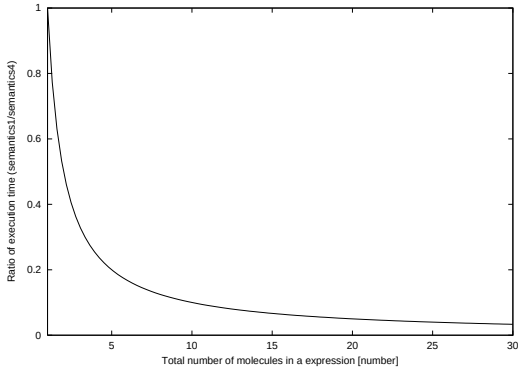


Figure 6: The ratio of the execution time under semantics1 to that under semantics4 for the same *bio-calculus* expression. The x-axis is the number of all molecules in an expression of *bio-calculus*. The y-axis is a relative ratio of execution time ($\frac{execution\ time\ under\ semantics1}{execution\ time\ under\ semantics4}$).