

Describing Metabolic Pathways Using an Artificial Chemistry Based on Pattern Matching and Recombination

Yoshikazu Suzuki

Kazuto Tominaga

School of Computer Science
Tokyo University of Technology
Katakura, Hachioji 192-0982 Japan

Abstract

In this paper, we apply our artificial chemistry to describing a real chemical system. The subject of description is a metabolic pathway: the oxidation of fatty acids. We modelled the pathway using straightforward design of molecules and reaction rules in the artificial chemistry. The description was executed by a simulator and correct behavior was observed. The qualitative modelling and simulation are a promising application of artificial chemistries.

Keywords: artificial chemistry, metabolic pathways, qualitative modelling, qualitative simulation.

1 Introduction

Artificial chemistry is becoming one of the most active research areas in artificial life field, and various artificial chemistry systems have been built [1]. An artificial chemistry is an abstract chemical model. It is usually used to construct virtual systems in order to study behavior of various living systems and/or necessary conditions for virtual systems to show such behavior. We proposed an artificial chemistry based on string pattern matching and recombination [2] and have constructed several virtual systems including an imaginary cell with cell cycle [3].

Recently, artificial chemistries are becoming applied to natural chemical systems. For example, Benkō et al. used an artificial chemistry based on graphs to produce chemically expected reaction pathways [4]. We applied our artificial chemistry to modelling DNA computing [5]. The capability of modelling natural chemical or biochemical systems will be a good property for artificial chemistries to broaden their applicable area.

Although DNA computing uses natural molecules (i.e., DNA) and their properties, modelling it does not contribute directly to understanding living systems, for it is performed in test tubes in laboratory. In this

paper, we attempt to model a system in natural organisms — the metabolic pathway of the oxidation of fatty acids — to explore possibilities of using the artificial chemistry as a tool to model such a system.

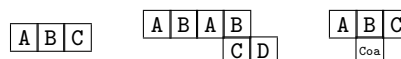
2 An Artificial Chemistry Based on Pattern Matching and Recombination

This section briefly explains a simple artificial chemistry based on pattern matching and recombination of stacks of strings [2], which we are going to use in the following discussion.

2.1 Elements and Objects

An *element* is denoted by an upper-case character, or a capitalized sequence of alphabetical characters. For example, A, B and Coa are elements. An element corresponds to an atom or a group of atoms in nature.

An *object* (corresponds to a molecule or a compound of molecules) is a character string or a stack of strings, such as those depicted below.



These objects are denoted by the string notations $0\#ABC/$, $0\#ABAB/3\#CD/$, $0\#ABC/1\#Coa/$ respectively; the numbers represent displacements of the line (in numbers of elements) relative to the first line. A displacement can be positive, zero or negative.

2.2 Patterns

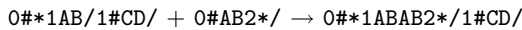
A *pattern* matches (or does not match) an object, and it can utilize two kinds of *wildcards*. An *element wildcard*, which is denoted by a number such as 1, matches any element. A *sequence wildcard*, denoted by a number and an asterisk such as 2^* and $*2$, matches

any sequence of zero or more elements; the position of an asterisk represents the direction in which the sequence can extend.

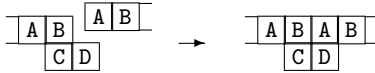
A pattern is denoted in a similar way to an object. For example, the pattern $0\#A1C/$ matches the object $0\#ABC/$; the pattern $0\#*1AB/1\#CD/$ matches the object $0\#ABAB/3\#CD/$. Note that the displacements are meaningful and that the length of a sequence wildcard is treated as zero in the notation for patterns.

2.3 Recombination Rules

A *recombination rule* transforms a group of objects into a group of objects, conserving elements just like a chemical reaction does. It is expressed in a manner similar to chemical formulae, but in terms of patterns. An example rule is



which is illustrated as follows.



If this rule is applied to the objects $0\#ABAB/3\#CD/$ and $0\#ABC/$, the object $0\#ABABABC/3\#CD/$ is produced and the reactants disappear.

2.4 Sources and Drains

Objects are kept in the *working multiset*. A *source* is defined as an object, and it supplies objects of the specified form to the multiset one at a time without limit. A *drain* is defined as a pattern, and it eliminates objects matched by the pattern, one at a time, from the multiset.

2.5 Dynamics

A *system* is a construct $\langle \Sigma, S, D, R, P_0 \rangle$ where Σ is a set of elements, S is a multiset of sources, D is a multiset of drains, R is a set of recombination rules, and P_0 is the *initial working multiset*, which specifies objects in the working multiset at the initial state. The system is interpreted nondeterministically as follows.

1. Initialize the working multiset to be P_0 .
2. Do one of the following operations.
 - Apply one recombination rule to a collection of objects.
 - Operate one source.
 - Operate one drain.
3. Go to Step 2.

3 Modelling a metabolic pathway

In this section, we are going to describe a metabolic pathway in natural organisms using the artificial chemistry explained in the previous section. The pathway to describe is the oxidation of fatty acids. Fat provides energy for organisms that do not perform photosynthesis. Fat is decomposed to glycerol and fatty acids, and fatty acids are degraded to acetyl CoA in mitochondria [6, Chapter 2].

3.1 The oxidation of fatty acyl CoA

The pathway of the oxidation of fatty acids to acetyl CoA is depicted in Figure 1. Shown at the top of the figure is a molecule of fatty acyl CoA, which is activated form of fatty acid. It has a long hydrocarbon chain; R in the figure stands for the tail of it.

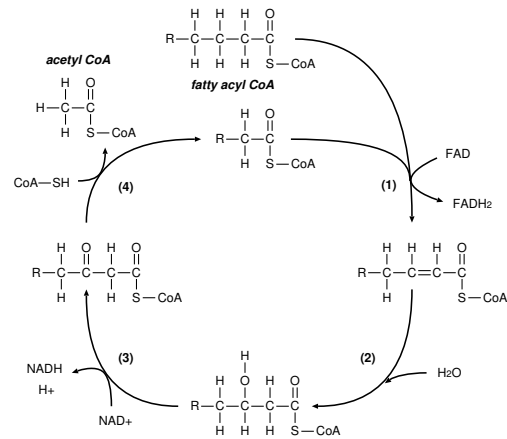


Figure 1: The oxidation cycle of fatty acids.

The pathway is cyclic. One cycle oxidizes the fatty acyl CoA and produces one molecule of acetyl CoA (at the top left of the figure). The acyl CoA loses two carbon atoms from its chain in one cycle and finally it becomes acetyl CoA itself.

The cycle comprises four steps of reactions ((1) to (4) in the figure); let us look at each step.

- Step 1: Acyl CoA reacts with FAD (flavin adenine dinucleotide; activated carrier). FAD gets two hydrogen atoms from the acyl CoA and becomes $FADH_2$.
- Step 2: Hydration. The product of Step 1 is hydrated.
- Step 3: Dehydrogenation. NAD^+ (nicotinamide adenine dinucleotide; activated carrier) removes two hydrogen atoms from the product of the previous step, making NADH and a proton (H^+).

- Step 4: Thiolysis. The product of Step 3 reacts with CoA-SH and becomes acyl CoA with shorter chain, producing one acetyl CoA. If no tail is left (i.e., R is H), two acetyl CoA are produced in this step, and the oxidation completes.

3.2 Atoms and Molecules

We design elements in our artificial chemistry as follows. For the atoms C, H, O and S, we provide corresponding elements C, H, O and S. The activated carriers CoA, FAD and NAD are represented by the elements Coa, Fad and Nad, respectively. Using these elements, NADH is represented by an object `0#NadH/`, for example. We use the element Po to denote positive charge so that NAD^+ and H^+ are expressed as `0#NadPo/` and `0#HPo/`. Furthermore, we introduce a dummy element, X, to represent vacancy.

Depicted below are example objects: fatty acyl CoA $\text{CH}_3\text{-CH}_2\text{-CH}_2\text{-CO-S-CoA}$ (left) and a product of Step 1 $\text{CH}_3\text{-CH=CH-CO-S-CoA}$ (right) assuming R is H in Figure 1. Because the latter has a double bond, two hydrogen atoms are missing compared to the former. The vacancies are denoted by the element X.



3.3 Recombination rules

The designed recombination rules are shown in Figure 2. The numbers given to the rules correspond to the numbers of steps in Figure 1.

The hydrocarbon chain of the acyl CoA changes its length during the process. In the rules, this part is represented by wildcards so that they can match a chain of any length.

Because we introduced the dummy element X to represent a vacancy, the rules do not conserve elements in the artificial chemistry (i.e., we relaxed a restriction imposed by the definition of our artificial chemistry given in Section 2.3); however, they are designed to conserve atoms in real reactions.

The oxidation cycle described in terms of our artificial chemistry is shown in Figure 3.

3.4 Initial working multiset

We give the following objects in the initial working multiset. A number in brackets is the number of objects of the form, determined empirically so that the simulator (discussed in the next section) effectively shows the process.

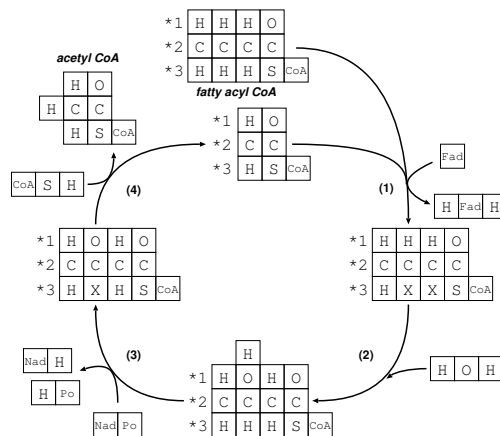


Figure 3: The oxidation cycle described by the artificial chemistry.

- Fatty acyl CoA [10]: below is an example object: `0#HHHHHHHHHO/-1#CCCCCCCCC/0#HHHHHHHHS Coa/`
- FAD [100]: objects of the form `0#Fad/`.
- NAD^+ [100]: objects `0#NadPo/`.
- CoA-SH [100]: objects `0:#CoaSH/`.
- H_2O [100]: objects `0#HOH/`.

4 Execution

We developed a simulator for our artificial chemistry. The current simulator is implemented in Objective-C and runs on the Cocoa framework of Mac OS X. Figure 4 shows the window of the simulator running the description given in the previous section. The table shows molecular species and their numbers, and the molecule window (bottom right) shows the form of specified species (by mouse click). In this figure, the product of the oxidation cycle, acetyl CoA, is shown. The execution of the above description consumed all the given acyl CoA molecules and produced acetyl CoA molecules, which is the correct behavior.

5 Discussion

As Figures 1 and 3 show, this artificial chemistry modelled the pathway in a straightforward manner. The elements correspond atoms and molecules naturally, and the recombination rules can be designed intuitively from the chemical formulae and the pathway graph. Describing chemical pathways using the artificial chemistry has significant benefits. First, a pathway is expressed as a mathematical system, so there is

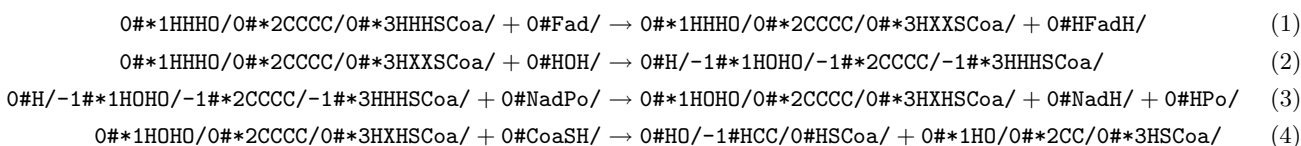


Figure 2: Recombination rules for the oxidation pathway.

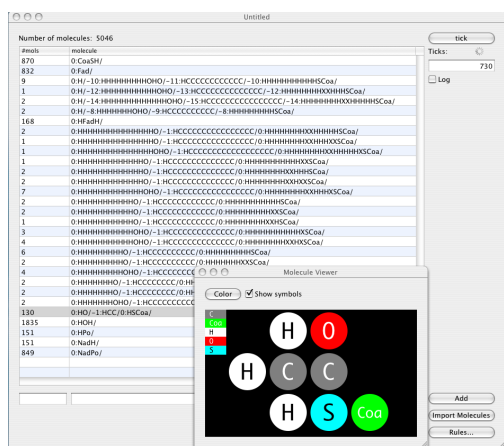


Figure 4: The simulator window.

no ambiguity. Second, it is executable on a simulator. Although the current simulator does not deal with concentrations or energy, it still can perform qualitative simulation. Furthermore, one can choose an arbitrary level of abstraction: for example, CoA can be represented by an element Coa like in the description given in this paper, or can be treated as composed of many atoms of C, H, O, etc.; the designer of the system can choose any level.

Other than the oxidation of fatty acids, we have described the following metabolisms: glycolysis, the citric acid cycle, the catabolism of amino acids, and the urea cycle. One more useful property of the artificial chemistry is that when the union of all the systems above (i.e., the union of the initial working multisets and the recombination rules) is given to the simulator, the execution shows the behavior of the composite system.

6 Concluding remarks

In this paper, we presented a description of a metabolic pathway described in terms of an artificial chemistry, and thereby showed possibility for the artificial chemistry to be applied to qualitative modelling

and simulation of chemical pathways. While there is room for improvement in the expressive power of the artificial chemistry, we think this is a promising application of artificial chemistries in general.

References

- [1] Peter Dittrich, Jens Ziegler, and Wolfgang Banzhaf. Artificial chemistries — a review. *Artificial Life*, 7:225–275, 2001.
- [2] Kazuto Tominaga. A formal model based on affinity among elements for describing behavior of complex systems. Technical Report UIUCDCS-R-2004-2413, Department of Computer Science, University of Illinois at Urbana-Champaign, March 2004.
- [3] Kazuto Tominaga. Describing protein synthesis and a cell cycle of an imaginary cell using a simple artificial chemistry. In *Proceedings of the Workshop on Artificial Chemistry and Its Applications, part of the 8th European Conference on Artificial Life (ECAL) 2005*, 2005.
- [4] Gil Benkö, Christoph Flamm, and Peter F. Stadler. Explicit collision simulation of chemical reactions in a graph based artificial chemistry. In *Advances in Artificial Life: Proceedings of the 8th European Conference on Artificial Life (ECAL) 2005*, pages 725–733. Springer, 2005.
- [5] Kazuto Tominaga. Modelling DNA computation by an artificial chemistry based on pattern matching and recombination. In *Proceedings of the Workshop on Artificial Chemistry and Its Applications, part of the 9th International Conference on the Simulation and Synthesis of Living Systems (ALIFE9)*, 2004.
- [6] Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. *Molecular Biology of the Cell*. Garland Science, New York, NY, 4th edition, 2002.