# Formal Description of Autopoiesis Based on the Theory of Category

Tatsuya Nomura
Faculty of Management Information, Hannan University,
5–4–33, Amamihigashi, Matsubara, Osaka 580–8502, Japan
URL: `http://www.hannan-u.ac.jp/~nomura/`

**Abstract**

Since the concept of autopoiesis was proposed as a model of minimal living systems by Maturana and Varela, there has been still few mathematically strict models to represent the characteristics of it because of its difficulty for interpretation. This paper proposes a formal description of autopoiesis based on the theory of category and Rosen's perspective of "closure under efficient cause", and discusses the effectiveness of autopoiesis in systems sciences through implication from the model.

## 1 Introduction

Autopoiesis gives a framework in which a system exists as an organism through physical and chemical processes, based on the assumption that organisms are machinary[6]. According to the original definition of it by Maturana and Varela, an autopoietic system is one that continuously produces the components that specify it, while at the same time realizing itself to be a concrete unity in space and time; this makes the network of production of components possible. An autopoietic system is organized as a network of processes of production of components, where these components:

1. continuously regenerate and realize the network that produces them, and
2. constitute the system as a distinguishable unity in the domain in which they exist.

The characteristics of autopoietic systems Maturana gives are as follows:

1. Autonomy:
   Autopoietic machinery integrates various changes into the maintenance of its organization. A car, the representative example of non–autopoietic systems, does not have any autonomy.
2. Individuality:
   Autopoietic machinery has its identity independent of mutual actions between it and external observers, by repeatedly reproducing and maintaining the organization. The identity of a non–autopoietic system is dependent on external observers and such a system does not have any individuality.
3. Self–Determination of the Boundary of the System:
   Autopoietic machinery determines its boundary through the self–reproduction processes. Since the boundaries of non–autopoietic systems are determined by external observers, self–determination of the boundaries does not apply to them.
4. Absence of Input and Output in the System:
   Even if a stimulus independent of an autopoietic machine causes continuous changes in the machine, these changes are subordinate to the maintenance of the organization which specifies the machine. Thus, the relation between the stimulus and the changes lies in the area of observation, and not in the organization.

Moreover, Kawamoto has continued his own development of autopoiesis. In his book[3], he designated the properties of autopoiesis by comparison with conventional system theories. In particular, he focuses on the fourth item among the above characteristics of autopoiesis, i.e., absence of input and output in the system. When we consider the "absence of input and output", important is the view where the system is understood based on the production processes. Kawamoto claims the following: the view of the relation between inputs and outputs in the system is one from external observers and it does not clarify the organization or the operation of the production in the system. A living cell only reproduces its components and does not produce the components while adjusting itself according to the relation between itself and oxygen in the air. Although the density of oxygen affects the production processes, external observers decide the influence and the cell does not. As long as the system is grasped from an internal view of the cell, the system does not have any "inputs and outputs". In addition, he gave the following gist in the concept of autopoietic systems[4]:

1. The set of components of a system is determined by the operation of the system.
2. The operation of the system precedes the initial condition.
3. The operation of the system is executed only to succeed itself and does not aim to produce by–products.
4. In the operation of the system, the things that happen in the system clearly differ from the things that external observers discriminate.

However, there has been still few mathematically strict models that represents autopoiesis. We discussed the difficulty of interpreting autopoiesis within conventional mathematical frameworks and problems of some models representing autopoiesis[7, 8]. The points are as follows:

**A Shift of Viewpoints:** How systems are grasped from the view of external observers is interpreted as separating the observers from the environment including the system, distinguishing between the system and the background in the environment, and verifying the relation between the system and the distinguished background, that is, the outside of the system. Autopoiesis forces us to give up this view, that is, to put our view in the system, not in the outside of the environment.
However, this shift of view is not easily acceptable in the contemporary situation where the view of external observers is still major in natural science. If a person bounded to this view observes an autopoietic system, the view shifts towards the outside of the environment and the system is grasped as a static map or dynamical system in a state space. Even if the view shifts towards the inside of the system, the production processes of the components themselves are grasped as the object of the observation and the view of external observers is not completely given up.

**Precedability of Operations to Elements and State Spaces:** As long as the view of external observers is not given up, the above gist of Kawamoto, in particular, the determination of the set of components by the operation and the precedability of the operation with the initial condition in the system cannot be understood. In the conventional system theories, state spaces where the operation is defined firstly exist, the initial condition is determined independent of the operation, and the properties in the state spaces by the operation such as time evolution are discussed.

**Lack of Some Characteristics in the Models of Autopoiesis:** In order to represent autopoiesis as mathematical or computational models, it is necessary to find machanism that a system creates the space where it exists and the boundary between it and the environments by itself. In the models previously proposed, however, some important characteristics are lost in the sense that the spaces that the systems exist are given in advance, some components are not reproduced in the systems, or a dynamical aspect of autopoiesis is not represented.

The aim in this paper is to clarify whether autopoiesis can really be represented within the conventional mathematical frameworks. For this aim, we introduce the theory of category[12], one of the most abstract algebraic structure representing relations between components. The focus is the concept of "closure under efficient cause" in "relational biology" by Rosen[9].

## 2 Closure under Entailment

In relational analysis, a system is regarded as a network that consists of components having functions. Rosen compared machine systems with living systems to clarify the difference between them, based on the relationship among components through entailment[9]. In other words, he focused his attention on where the function of each component results from in the sense of Aristotle's four causal categories, that is, material cause, efficient cause, formal cause, and final cause. As a result, Rosen claimed that a material system is an organism if and only if it is closed to efficient causation.

For exmaple, (M,R) systems[10] satisfy closure under efficient cause. This system model maintains its metabolic activity through inputs from environments and repair activity. The simplest (M,R) systems represent the above aspect in the following diagram and the left figure in Fig. 1:

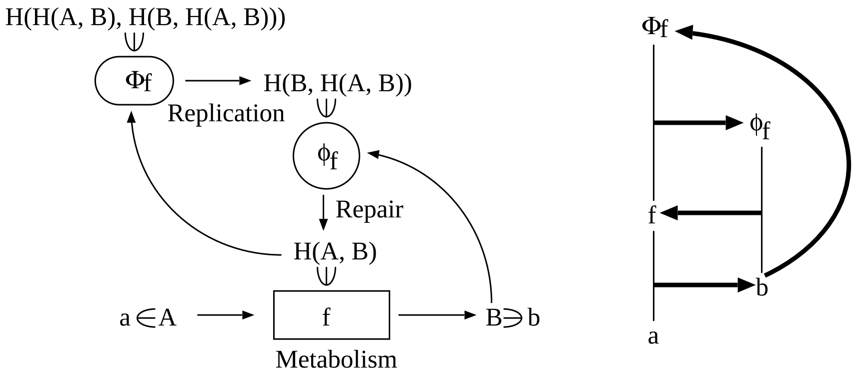


Figure 1: A (M,R) System and Its Hyperhigraph of Entailment

$$A \xrightarrow{f} B \xrightarrow{\phi_f} H(A,B) \xrightarrow{\Phi_f} H(B,H(A,B)) \tag{1}$$

Here, $A$ is a set of inputs from an environment to the system, $B$ is a set of outputs from the system to the environment, $f$ is a component of the system represented as a map from $A$ to $B$, and $\phi_f$ is the repair component of $f$ as a map from $B$ to $H(A,B)$ ($H(X,Y)$ is the set of all maps from a set $X$ to a set $Y$). In biological cells, $f$ corresponds to the metabolism, and $\phi_f$ to the repair. If $\phi_f(b) = f$ $(b = f(a))$ is satisfied for the input $a \in A$, we can say that the system self–maintains itself. In addition, $\Phi_f$ can be constructed by the preceding (M, R) system in the following way: For $a$ and $b$ such that $b = f(a)$ and $\phi_f(b) = f$, if $\hat{b} : H(B,H(A,B)) \to H(A,B)$ $(\hat{b}(\phi)(a') = \phi(b)(a')$ $(\phi \in H(B,H(A,B)), a' \in A))$ has the inverse map $\hat{b}^{-1}$, it is easily proved that $\hat{b}^{-1}(f) = \phi_f$. Thus, we can set $\Phi_f = \hat{b}^{-1}$. The right figure in Fig. 1 shows the aspect that the components except for $a$ are closed under entailment, by a hyperdigraph[1].

It is considered that closure under entailment or production is a necessary condition for a system to be autopoietic because the components reproduce themselves in the system. In order to clarify what system is closed under entailment in more general framework than the naive set theory, we use the theory of category.

## 3 Some Systems Closed under Entailment in a Category Theoretic Framework

In this paper, we assume that a category $\mathcal{C}$ has a final object 1 and product object $A \times B$ for any pair of object $A$ and $B$. The category of all sets is an example of this category. Moreover, we describe the set of morphisms from $A$ to $B$ as $H_{\mathcal{C}}(A,B)$ for any pair of objects $A$ and $B$. A element of $H_{\mathcal{C}}(1,X)$ is called a morphic point on $X$. For a morphism $f \in H_{\mathcal{C}}(X,X)$ and a morphic point $x$ on $X$, $x$ is called a fixed point of $f$ iff $f \circ x = x$ ($\circ$ means composition of morphisms)[11]. Morphic points and fixed points are respectively abstraction of elements of a set and fixed points of maps in the category of sets. As far as we do not explicitly note, we consider systems in this general category.

When there exists the power object $Y^X$ for objects $X$ and $Y$ (that is, the functor $\cdot \times X$ on $\mathcal{C}$ has the right adjoint functor $\cdot^X$ for $X$), note that there is a natural one–to–one correspondence between $H_{\mathcal{C}}(Z \times X, Y)$ and $H_{\mathcal{C}}(Z, Y^X)$ for any objects $X$, $Y$, $Z$ satisfying the diagram in the left figure of Fig. 2. Thus, there is a natural one–to–one corrspondence between morphic points on $Y^X$ and morphisms from $X$ to $Y$ satisfying the diagram in the right figure of Fig. 2.

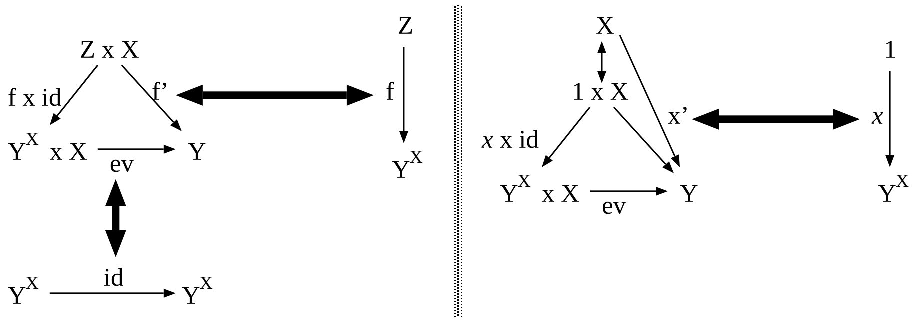


Figure 2: Natural One–To–One Correspondence between $H_{\mathcal{C}}(Z \times X, Y)$ and $H_{\mathcal{C}}(Z, Y^X)$

We propose some systems closed under entailment.

### 3.1 Completely Closed Systems

When components in a system are not only operands but also operators, the easist method for representing this aspect is the assumption of existence of a isomorphism from the space of operands to the space of operators[2].

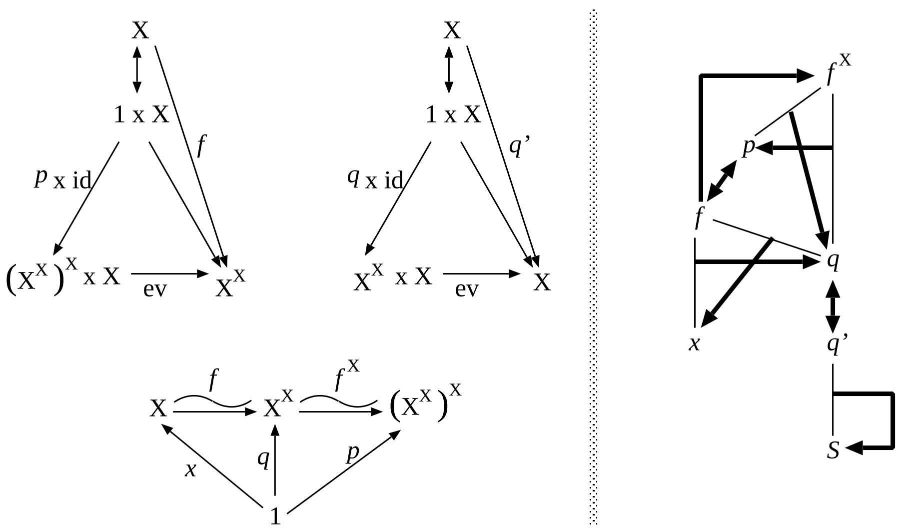


Figure 3: The Diagrams of a Completely Closed System and Its Hyperdigraph on Entailment

Now, we assume an object $X$ with powers and an isomorphism $f : X \simeq X^X$ in $\mathcal{C}$. Then, there uniquely exists a morphic point $p$ on $(X^X)^X$ corresponding to $f$ in the above sense, that is , $p' = f$. Since the morphism from $X^X$ to $(X^X)^X$ entailed by the functor $\cdot^X$, $f^X$, is also isomorphic, there uniquely exists a morphic point $q$ on $X^X$ such that $f^X \circ q = p$. We can consider that $p$ and $q$ entail each other by $f^X$. Furthermore, there uniquely exists a morphic point $x$ on $X$ such that $f \circ x = q$ because $f$ is isomorphic. Since we can consider that $x$ and $q$ entail each other by $f$, and $f$ and $p$ entail each other by the natural correspondence, the system consisting of $x$, $q$, $p$, $f$, and $f^X$ is completely closed under entailment. Moreover, if a set $S$ of morphic points on $X$ is fixed by $q' : X \to X$ naturally corresponding to $q$, that is, if $x' \in S$ then $q' \circ x' \in S$ and there exists $x'' \in S$ such that $q' \circ x'' = x'$, we can consider that $S$ entails itself by $f$ (the existence of these sets is guaranteed by Theorem 1 in [11], that is, the fact that $q'$ has fixed points by the labelling of $X^X$ by $X$, $f$).

Fig. 3 shows the diagrams of this completely closed system and its hyperdigraph on entailment. Thus, one isomorphism from $X$ to $X^X$ generates one completely closed system.

## 3.2 Generalized (M,R) Systems

As mentioned in section 2, (M,R) systems are closed under entailment except for the input $a$. We can generalize the closed part of (M,R) systems as follows.

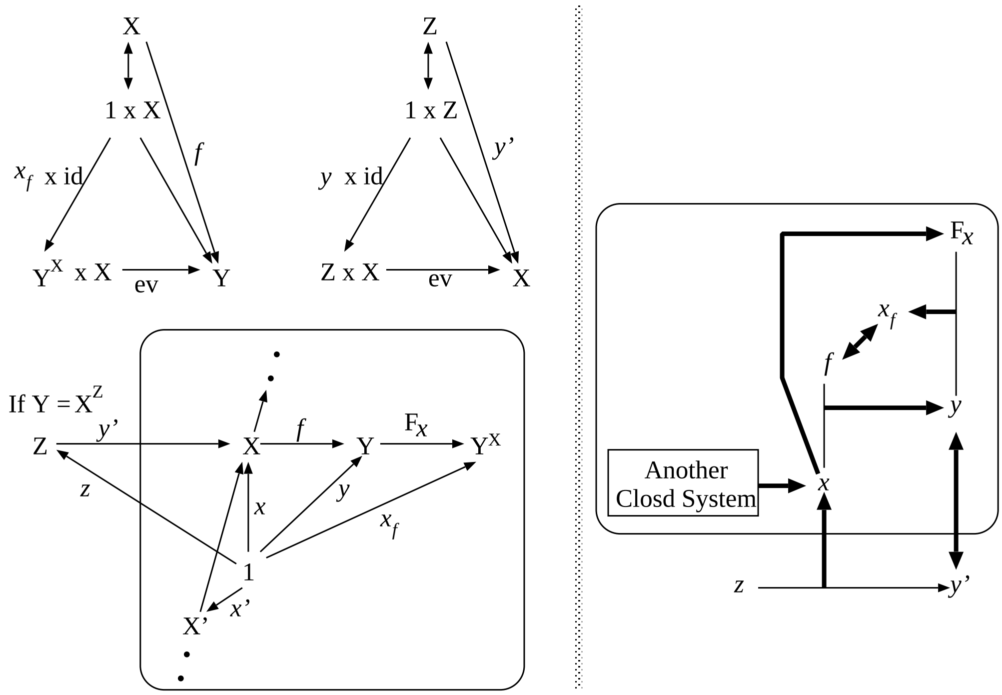


Figure 4: The Diagrams of a Generalized (M,R) System and Its Hyperdigraph on Entailment

For objects $X$ and $Y$ in $\mathcal{C}$, we assume that $X$ has powers. When a morphism $f : X \to Y$ and a morphic point $x$ on $X$ are given, we assume that $x$ satisfies the following conditions:

$$\exists\, G_x \in H_{\mathcal{C}}(Y^X, Y) \ s.t.,\ G_x \circ z = z' \circ x \text{ for any } z \in H_{\mathcal{C}}(1, Y^X) \quad (2)$$
$$\text{and } G_x \text{ has its inverse morphism } F_x \in H_{\mathcal{C}}(Y, Y^X)$$

here, $z'$ is the morphism from $X$ to $Y$ naturally corresponding to the morphic point $z$ on $Y^X$. When $y = f \circ x$ and $x_f$ is the morphic point on $Y^X$ naturally corresponding to $f$ ($(x_f)' = f$), we obtain $F_x \circ y = F_x \circ f \circ x = F_x \circ G_x \circ x_f = x_f$. Thus, $x_f$ is entailed by $y$ and $F_x$. If we regard $F_x$ as entailed by $x$, then $f$, $y$, $F_x$, and $x_f$ are entailed by themselves and $x$. Although $x$ is not entailed by $x$, $f$, $y$, $x_f$, or $F_x$, we can consider a larger system closed under entailment if $x$ is one of morphic points of another closed system (for example, a completely closed system in section 3.1). If there exist an object $Z$ with powers and morphic point $z$ on $Z$ such that $Y = X^Z$ and $y' \circ z = x$, the system including $y'$ and $z$ represents a generalized (M,R) system.

Fig. 4 shows the diagrams of this generalized (M,R) system and its hyperdigraph on entailment.

### 3.3 Infinte Regressive Systems

We consider a system like (M,R) systems including a kind of infinte regress. Now, we assume objects $X_i := {X_{i-1}}^{X_{i-2}}$, morphisms $f_i \in H_C(X_i, X_{i+1})$, and morphic points $x_i$ on $X_i$ such that $(x_{i+1})' = f_{i-1}$ and $f_i \circ x_i = x_{i+1}$ $(i = 0, \pm 1, \pm 2, \ldots)$. Although any morphic point and morphism are closely entailed each other in this system, its entailment cannot be reduced to any finite subset of components and represents a kind of infinite regress. Moreover, we assume that there exist the limit $X^\infty$ and colimit $X_\infty$ of $(X_i, f_i)$, and $X^\infty$ coincides with $X_\infty$. Furthremore, when we put $X = X^\infty = X_\infty$, and $p_i$ and $q_i$ are the projection from $X^\infty$ to $X_i$ and injection from $X_i$ to $X_\infty$ respectively, we assume $p_i \circ q_i = \mathrm{id}_{X_i}$ $(i = 0, \pm 1, \pm 2, \ldots)$.

Then, there uniquely exists a morphic point $x$ on $X$ such that $p_i \circ x = x_i$ for any $i$ since $X$ is the limit. On the other hand, the morphic point $y = q_i \circ x_i$ (determined independent on $i$ because of $q_i = q_{i+1} \circ f_i$) satisfies $p_i \circ y = p_i \circ q_i \circ x_i = x_i$. Thus, we obtain $x = y$ from the uniqueness of $x$. Moreover, when $\pi_2^{(i)}$ is the projection of $X_i \times X$ to $X$, the morphism $g_i$ naturally corresponding to $\pi_2^{(i)}$ satisfies $g_i = g_{i+1} \circ f_i$. Thus, there uniquely exists a morphism $f : X \to X^X$ such that $f \circ q_i = g_i$ since X is the colimit. Fig. 5 shows the diagrams of this system.

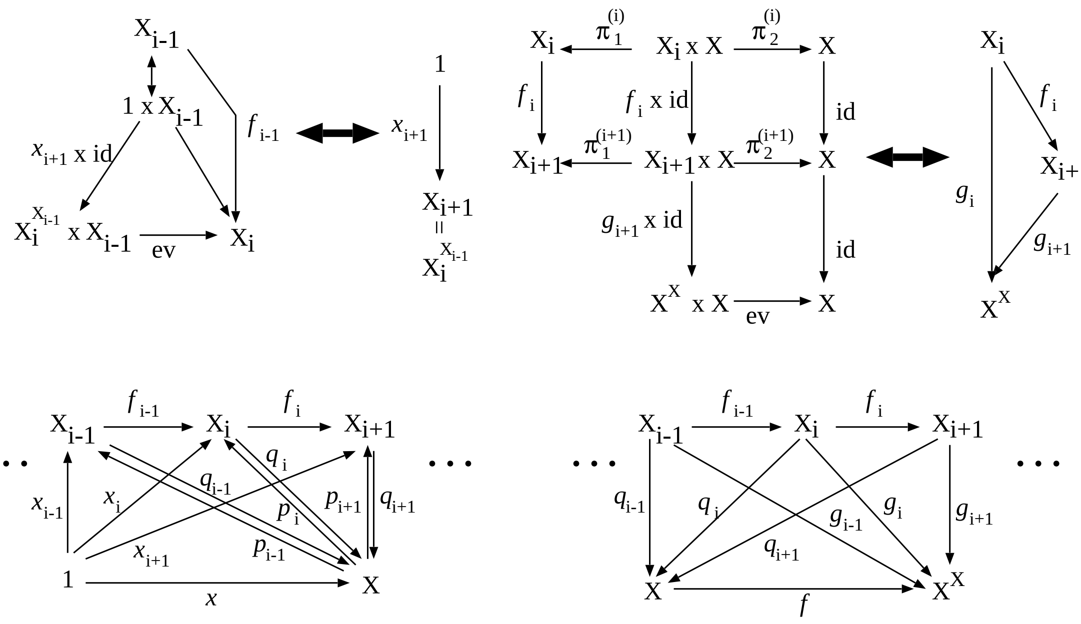


Figure 5: The Diagrams of an Infinite Regressive System

If the above $f$ is isomorphic, we can construct a completely closed system in section 3.1 including $f$ as one of its components. Then, if $x$ is a component in this closed system, $x$ is entailed in the system independent of $\{x_i, f_i\}$. Moreover, $x_i$ is entailed by $x$ and $p_i$, and $p_i$ is entailed by $p_{i-1}$ and $f_{i-1}$. Fig. 6 shows the hyperdigraph on entailment in the system. This represents a possibility that a system consisting of infinite regress construct a finite closed system and entailment from it by the system itself, that is, a kind of projection from the finite system to the infinite system.

# 4 Conclusion and Discussion

In this paper, we proposed some types of systems closed under entailment as autopoietic systems based on the theory of category and the assumption that closure under entailment is a necessary condition for

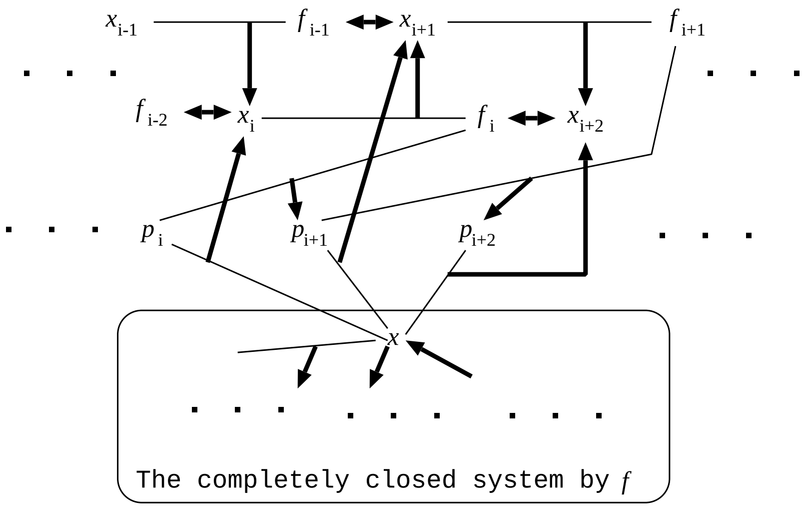


Figure 6: The Hyperdigraph on Entailment in an Infinite Regressive System

systems to be autopoietic. In particular, completely closed systems and generalized (M,R) systems are considered to represent the absence of input and output in the sense that there is no effect for entailment in the system even if there is a morphism from the outside of the system to the inside and a morphic point that entails one of components of the system; for a component $x$, another component $y$ entailing $x$ is entailed by $x$ and other components in the system, and it is independent whether $y$ is entailed from the outside of the system.

As one of future problems, we should consider coupling of these closed systems as far as we propose them as autopoietic systems. Although we do not deal with it in this paper, we can consider that a system $A$ is coupled with another system $B$ when a component of $A$ coincides with a component of $B$, that is, $A$ and $B$ are independently closed under entailment, and affect each other through the common component. Moreover, when we deal with a closed system in a category $\mathcal{C}$ and another closed system in another category $\mathcal{D}$, we can consider a functor $F$ from $\mathcal{C}$ to $\mathcal{D}$ such that maps $A$ to $B$, and a condition that there exists a closed system on $F$ as an object in the category of functors from $\mathcal{C}$ to $\mathcal{D}$. We are going to consider the possibility of this idea for contributing to mathematical implementation of Luhmann's communication system[5].

The most important problem is the conditions of the category used for constructing closed system. Although we required that operands coincide with operators ($X \simeq X^X$ in section 3.1 or $Y \simeq Y^X$ in section 3.2), this consition is difficult to be satisfied in the naive set theory. Although Soto–Andrade and Varela provided a category satisfying this condition (the category of partially ordered sets and continuous monotone maps with special conditions)[11], this category is very special. Furthermore, Rosen showed that systems closed under efficient cause cannot be described with their states because they lead to infinite regress[9]. If these closed systems can exist only in special categories not observable in the conventional sense, autopoiesis may be hard to be a general theory of a variety of systems.


**Acknowledgements:**

This research was supported by The Grant #B2002 "Studies on Social Complexity by Cognitive Reasoning Mechanisms and Artificial Life Algorithms" from the Institute of Industrial and Economic Research in Hannan University.


# References


[1] Higuchi, A., Matsuno, K., and Tsujishita, T.: Deductive Hyperdigraphs, A Method of Describing Diversity of Coherence. preprint (1997).

[2] Kampis, G: Self–Modifying Systems in Biology and Cognitive Science: A New Framework for Synamics, Information, and Complexity. Pergamon Press (1991).

[3] Kawamoto, H.: Autopoiesis: The Third Generation System. Seido–sha Publishers (1995). (Japanese)

[4] Kawamoto, H.: Autopoiesis and Mechanics of Cognition. Proc. of the 13–th Annal Meeting of Japanese Cognitive Science Society (1996) 12–13. (Japanese)

[5] Kneer, G. and Nassehi, A.: Niklas Luhmanns Theorie Sozialer Systeme. Wilhelm Fink Verlag (1993). (Japanese Edition: T. Tateno, et. al. Shinsensha (1995).)

[6] Maturana, H. R., and Varela, F. J.: Autopoiesis and Cognition: The Realization of the Living. D. Reidel Publishing (1980). (Japanese Edition: Kawamoto, H. Kokubun–sha Publishers. (1991).)

[7] Nomura, T: An Attempt for Description of Quasi–Autopoietic Systems Using Metabolism–Repair Systems. Proc. the Fourth European Conference on Artificial Life (1997) 48–56.

[8] Nomura, T: A Computational Aspect of Autopoiesis. Proc. the Fourth Asian Fuzzy Systems Symposium (2000) 1–6.

[9] Rosen, R: LIFE ITSELF. Columbia University Press (1991).

[10] Rosen, R: Some Relational Cell Models: The Metabolism–Repair Systems. In FOUNDATIONS OF MATHEMATICAL BIOLOGY. Academic Press (1972) 217–253.

[11] Soto–Andrade, J., and Varela, F. J.: Self–Reference and Fixed Points: A Discussion and an Extension of Lawvere's Theorem. Acta Applicandae Mathematicae **2** (1984) 1–19.

[12] Takeuchi, G: Sheaf, Category, and Topos. Nihon Hyoron–sha (1978).