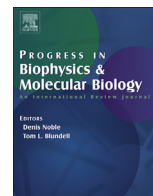




Contents lists available at ScienceDirect

## Progress in Biophysics and Molecular Biology

journal homepage: [www.elsevier.com/locate/pbiomolbio](http://www.elsevier.com/locate/pbiomolbio)

# Biologically inspired information theory: Adaptation through construction of external reality models by living systems



Toshiyuki Nakajima

Department of Biology, Ehime University, Bunkyo-cho 2-5, Matsuyama, Ehime 790-8577, Japan

## ARTICLE INFO

Article history:  
Available online 18 July 2015

Keywords:  
Information  
Inverse causality  
Entropy  
Sign  
Measurement  
Uncertainty

## ABSTRACT

Higher animals act in the world using their external reality models to cope with the uncertain environment. Organisms that have not developed such information-processing organs may also have external reality models built in the form of their biochemical, physiological, and behavioral structures, acquired by natural selection through successful models constructed internally. Organisms subject to illusions would fail to survive in the material universe. How can organisms, or living systems in general, determine the external reality from within? This paper starts with a phenomenological model, in which the self constitutes a reality model developed through the mental processing of phenomena. Then, the *it-from-bit* concept is formalized using a simple mathematical model. For this formalization, my previous work on an algorithmic process is employed to constitute symbols referring to the external reality, called the inverse causality, with additional improvements to the previous work. Finally, as an extension of this model, the cognizers system model is employed to describe the self as one of many material entities in a world, each of which acts as a subject by responding to the surrounding entities. This model is used to propose a conceptual framework of information theory that can deal with both the qualitative (semantic) and quantitative aspects of the information involved in biological processes.

© 2015 Elsevier Ltd. All rights reserved.

## 1. Introduction

Living systems have cognitive functions that help them cope with their uncertain environment and thereby survive. Through acting in the world and processing event data, the organisms of higher animals construct reality models for the environments with which they interact. Organisms that have not developed information-processing organs may also have external reality models built in the form of their biochemical, physiological, and behavioral structures, acquired by natural selection through successful models constructed internally. Organisms subject to illusions would fail to survive in the material universe. How can a subject, or a living system in general, understand the external reality through interaction from within?

In general, there are at least two different types of models in this context: (i) a world model for an external observer, which does not interact with the observer (called the external model) and is similar to an ordinary road map for a driver; and (ii) a model of the environment or external reality with which an observer interacts

(called the internal model in the sense that it is constructed internally). This is a perception-behavior system embodied in the organism, by which it can act in specific ways to maintain a particular relationship with the environment within the world (Conant and Ashby, 1970; Friston, 2010). This is similar to a car navigation system guiding a driver toward a desirable destination by voice and visual instructions.

A classic work by Jakob von Uexküll is useful (1926) for understanding an internal model for an organism. Uexküll developed the concept of the function circle, which forms a subject-based environment, called Umwelt: “every animal is a subject, which, in virtue of the structure peculiar to it, selects stimuli from the general influences of the outer world, and to these it responds in a certain way. These responses, in their turn, consist of certain effects on the outer world, and these again influence the stimuli. In this way there arises a self-contained periodic cycle, which we may call the function-circle of the animal” (Uexküll, 1926; italicized for emphasis by Nakajima). In this function-circle process, the focal subject and external entities interact with one another in specific ways, depending on their own properties in relation to how to move or change state. Through this process, the subject experiences a particular set of events at particular probabilities. Umwelt and an

E-mail addresses: [nakajima@sci.ehime-u.ac.jp](mailto:nakajima@sci.ehime-u.ac.jp), [cognizers@yahoo.co.jp](mailto:cognizers@yahoo.co.jp).

ecological niche can both be defined as a set of environmental conditions required for survival and reproduction, containing the implication that these conditions are shaped by direct and indirect interactions between the environment and a subject organism—both concepts involves what is required from and the impact on the environment (for niche concept, see e.g., Hutchinson, 1957; Chase and Leibold, 2003). However, these are different concepts in that Umwelt is viewed from (experienced by) a subject organism, while niche is viewed from an external perspective or by an external observer (i.e., ecologist).

Although Uexküll did not develop a mathematical model for this process, and it does not extend to molecular systems and those of lower entities, this process illuminates a general aspect of the information process between a subject and environmental entities within the world, which is also applicable to various levels of nested hierarchical organization, such as cells and molecules (e.g., Koshland, 1973; Loewenstein, 1999). Biologists usually use the term “information” to describe processes in which patterns (or configurations, forms) are processed within a cell, intercellular processes in development, organism–environment interactions, and so on. However, the information concept in biology is yet to be theoretically unified with a mathematical formalism.

Shannon’s mathematical theory of communication, usually called information theory, is widely recognized as a major mathematical theory that sheds light on a certain aspect of what we call information (Shannon and Weaver, 1949/1998). According to this theory, information is something that reduces the uncertainty (quantified as entropy) of an object (i.e., information source). As the authors themselves admit, the theory does not deal with the meaning of individual pieces of information, but instead, considers the amount of information as the degree of reduced uncertainty before and after receiving the information (a message). Despite its great importance in communication technology, there are several reasons why this theory has limitations in its application to biological processes. First, as previously stated, the theory cannot deal with semantic aspects of information. The reduction of uncertainty is important but is not the only necessity for living systems. They utilize goal-directed teleonomic properties to maintain *specific* relationships with others as a consequence of natural selection (Pittendrigh, 1958). Therefore, the semantics of information, which specifically relates the subject to the environment, is very important in biology. Second, this theory is not built upon an explicit causal model of entities interacting with one another. The theory may indeed have a material model such as an information source from which messages are sent from a transmitter, through a receiver, to a destination (noise may enter the path between the transmitter and receiver), but this is an abstract one-way symbol processing flow, with no explicit representations of causal interactions between component entities. In contrast, in biological systems, every material entity may serve as an information source, which interacts with others. Furthermore, the theory is based on the probability concepts of Bayesian probability and/or relative frequency, which both lack an explicit causal model of a material system. For example,  $p(a|b)$  implies the probability of  $a$  under the condition that  $b$  occurs, but no causal mechanism is explicit in this formulation.

From a study of quantum physics, Wheeler (1989) proposed another type of information concept, called *it from bit*. In contrast to Shannon’s model, which assumes that information sources exist out there, this concept does not assume the existence of real entities outside the subject. *It-from-bit* information is not a process from external reality to a subject, but a process (bit) used to construct external reality (*it*). This concept is important in focusing on the subject-dependent aspect of reality. However, like Shannon’s theory, the *it-from-bit* concept appears to be descriptive and

not explanatory for biological processes, and also lacks a semantic aspect of information. Therefore, we need to develop an information theory that can deal with both the qualitative (involving semantics and value) and quantitative aspects of the information involved in biological processes using a materialistic causal model.

This paper explores the information involved in living systems with the goal of developing a general framework of information theory to understand biological adaptation theoretically, beyond the limitation of verbal arguments in biology. We begin with information from the internal viewpoint, inspired by the philosophical concept of phenomenological reduction as a useful methodology, in which the external reality for the self is not supposed (but not denied) to exist out there (section 2). In this attempt, a simple mathematical model with an algorithmic process called inverse causality is introduced, focusing on the process used by the self to constitute symbols referring to the external reality or external objects from given phenomena. Using this model, it is shown that *it-from-bit* information can be formalized as a process based on the inverse causality (section 3). Then, information from the external viewpoint is considered, in which a focal subject receives information about external objects (section 4). In this analysis, a realistic world model is introduced, with additional assumptions to the first model, called the cognizers system model (Nakajima, 1999, 2013), in which material entities are described as the subjects of motion/action, interacting with one another within the world. Although this model has the external viewpoint of a meta-observer outside the world, it can also describe how a focal subject entity cognizes the environment and acts from an internal viewpoint. Here, the probability of events that a focal entity experiences (called the internal probability) is theorized in terms of the material properties of entities and their interactions (section 5). In this framework, information and its meaning are defined, and the uncertainty and order of the environment that a focal subject entity experiences in the environment during interaction are formalized, measured as internal entropy, which is based on the internal probability concept, rather than Bayesian probability. Finally, biological adaptations are analyzed in terms of the internal probability/entropy and the meaning of information (section 6).

## 2. Background: phenomenological view of reality as departure point

A realistic view presupposes the instinctive belief that objects that exist independently of the self cause percepts, which cause an inner, mental perception. In science, the subject gains information about the external reality by observation (i.e., a process from reality to phenomena). However, Descartes’ skepticism uncovered distinctions between the existence of percepts (Cogito) and the existence of the objects to which the percepts refer. A percept such as “There is an apple on the table,” is absolutely certain to the self because it is an immediate, indubitable experience. However, this fact does not guarantee that those things exist outside, independent of the self, despite its immediateness. That is, “this apple” would belong to the self, not to the outside realm. As revealed by skepticism, the semantic content cannot provide a secure base for proving the existence of such reality. Percepts in dreams do not reflect information about the external reality, raising a question: where is an “escape button” from solipsism (Rössler, 1993)?

This question concerns realization, which is a mental activity to constitute reality external to the self, based on phenomena or sense data occurring to the self (i.e., a process from phenomena to reality). Phenomena are generated from the reality by *observation* in the self, and the reality is constituted by *realization* in the self—“realization” is defined as the constitution of the reality external

to the self, based on phenomena or sense data internal to the self. Here, observation presupposes the existence of the reality, whereas the reality depends on realization (Maturana and Varela, 1980; Rössler, 1987; Matsuno, 1989; Varela et al., 1992; Nakajima, 2001; Saithe, 2001). There is an unavoidable circularity between phenomena and reality, which are related to each other by the mental activities of realization and observation.

In order to step out of this circularity, Husserl considered phenomena as mental processes as his starting point, and developed phenomenology as a rigorous science (Husserl, 1913/1983). Husserl analyzed the stream of mental processes to clarify the relation between the mind and world. He first methodologically rejected, i.e., bracketed, the existence of external reality or natural objects, and then examined how the mental process can “constitute” object things. He theorized that the directedness or aboutness of consciousness (i.e., intentionality) plays a pivotal role in singling out a particular object, e.g., a sheet of paper, in the background of other things such as books and pencils (Husserl, 1913/1983). Furthermore, he questioned what allows us (many subjects or selves) to share a common belief about an objective reality consisting of such objects derived through intentionality, starting from a solipsistic point. His answer is this: Under the assumption that, largely, the other subjects constitute objective reality in the same way as “I” (the self) do, “I” can believe that object things forming my own world exist independently of “my” subjective experiences, and are shared by other selves; this is the intersubjective constitution of the objective reality (Husserl, 1931/2012, 1936/1970).

We can extend (or naturalize) phenomenology by replacing the phenomenological self with observers in science, and further, with living systems in biology. The attempt to bridge the gap between the first-person view in phenomenology and a third-person view in science is particularly important in biology, because not only are we (humans) organisms, but also objects in biology can be viewed as the subjects of action in their environments (Uexküll, 1926, 2010). Sjolander (1997) asserted that a human’s epistemological activity may reflect the human’s evolutionary background, for which the construction of reality is an adaptation to the environment, not the truth about its outside. However, in Husserl’s phenomenology and that of other phenomenologists who seek to naturalize phenomenology, it remains unclear how the self can distinguish between material entities and imaginary or theoretical entities, because phenomenologists try to construct objective reality based on the semantic contents (e.g., “this book”) of percepts or phenomena.

In the quest for physical reality in quantum physics, Wheeler (1989) asserted that “every it—every particle, every field of force, even the space-time continuum itself—derives its function, its meaning, its very existence entirely—even if in some contexts indirectly—to yes or no questions, binary choices, bits.” He called this process of finding reality *it from bit*. For example, we watch an experimental device consisting of a polarizer over a distant source and an analyzer of polarization over a photo-detector. We then ask a yes-no question such as “Did the counter register a click during the specified second”? If yes, “a photon did it” (Wheeler, 1989). In this view, the reality is not assumed to exist out there independently of observation; instead it is derived from given sense data obtained from the yes-no questions posed by an experiment. That is, all things are information-theoretic in origin. Asking yes-no questions and registering answers always affect the object. In this reality construction, the semantic contents of percepts by observation are removed as “bits,” because yes-no answers only give the observer the distinction or determination of one of the two. Wheeler’s *it-from-bit* concept is a good departure point to develop a theory for a subject-dependent external-reality model based on an objective principle consistent with the ordinary scientific

methodology.

### 3. Constitution of reality model by inverse causality

We start with a solipsistic world, in which the self does not have any outside; that is, the self itself is the whole world. Specifically, we suppose a sequence of percepts or sense data, which are further abstracted into meaningless symbols. Then, a principle is derived by which the outside is constituted from percepts without recourse to perceptual contents.

#### 3.1. Perceptual sequence

Consider a temporal sequence of percepts, which occur during seeing, hearing, touching, thinking, imaging, etc. They are given as they are, as an immediate experience for the self. This is what Husserl called a stream of mental processes. We can then remove the semantic contents of each percept, such as meaning and qualia, and represent the sequence ( $M$ ) of symbolized percepts using meaningless symbols ( $mi$ ). For example,

$m0, m1, m1, m2, m3, m0, m2, m4, m5, \dots$ ; where  $mi = mj$ , if and only if  $i = j$ .

This sequence is a chronicle of perceptual symbolized elements arrayed in a time series, where only their order in time and distinctness are identified. This sequence is given as it is, not entailed by something else, and therefore there is no reason to consider that each of these elements was caused or entailed by its own predecessor. We have this sequence as an entire wholeness at a starting point. There is no entity outside  $M$ , although its outside can exist only as a form of meaning possessed by each percept.  $M$  is more primitive than a solipsistic world, because the former is deprived of the meaning and qualia possessed by mental states. Each percept may be divided into micro-timescale processes of elementary states, such as  $mi = mi_1, mi_2, \dots, mi_k$  and  $m_j = mj_1, mj_2, \dots, mj_k$ , where  $mi = mj$ , if and only if  $mi_h = mj_h$  ( $1 \leq h \leq k$ ).

#### 3.2. Principle of unique successor and inverse causality

Next, we introduce a rule, called *the unique-successor principle*, which stipulates that “Every element in the temporal sequence has a unique successor.” The above sequence  $M$  does not satisfy this principle, because  $m1$  is followed by  $m1$  and  $m2$ ; and  $m2$  is followed by  $m3$  and  $m4$ . Thus, we can modify the original sequence to fulfill the principle by introducing foreign elements to the original sequence. This modification is conducted by using the unique-successor principle inversely, which is called the inverse causality. Inverse causality is defined as follows: “If a given perceptual sequence satisfies the unique-successor principle, no operation is required. If not, new, foreign elements are introduced into the sequence in order that every perceptual element is a unique successor of an immediately previous element of the new sequence.”

We can demonstrate the operation of inverse causality using the simple example of  $M$ :

$m0, m1, m2, \dots, m0, m1, m3,$  (1)

where  $mi = mj$ , if and only if  $i = j$ ;  $\mathbf{M} = \{m0, m1, m2, m3, \dots\}$ . This sequence does not fulfill the unique-successor principle, because the first  $m1$  is followed by  $m2$ , and the second  $m1$  by  $m3$ ; i.e.,  $m1 \rightarrow m2$  or  $m3$ . Therefore, through an inverse causality operation, the first and second  $m1$  are differentiated by introducing the foreign symbols  $e0$  to the first  $m1$ , and  $e1$  to the second one—note

that  $m_2$  and  $m_3$  cannot be made identical to fulfill the principle because they are undoubtedly different, as directly experienced by the self. Therefore, the sequence now becomes

$$m_0, \underline{(m_1, e_0)}, m_2, \dots, m_0, \underline{(m_1, e_1)}, m_3, \tag{2}$$

where the underlined parts were modified. This sequence is generated in the downstream part of the  $M$  sequence.

From the meta-observer point of view, the operation of inverse causality on  $M$  can be interpreted as a measurement process for different states of the external reality,  $\mathbf{E}_0 = \{e_0, e_1\}$ . The perceptual changes occurring at  $m_1$ , such as  $(m_1, e_0) \rightarrow m_2$ ,  $(m_1, e_1) \rightarrow m_3$ , can be a process of perceptual discrimination about something by the self. If we use the state concept, the differences to be discriminated are differences in the state of something. Measurement is a causative discrimination process by a measurer between different states of an object (Rosen, 1977, 1978). Measurement requires the sameness of the subject's state as a reference state (e.g.,  $m_1$  in the example) to detect an object's state difference. As long as any measurement of objective differential states involves this discriminative response or cognition by the self, the derivation of  $\mathbf{E}_0$  by inverse causality is itself a measurement process.

The operation of inverse causality on  $M$  (denoted as ICM) is represented as follows: given a sequence  $M$ , if there are  $n$  different state changes from the same state ( $my$ ) to different states ( $mz_1, mz_2, \dots, mz_n$ ), where  $my$  may be identical to  $mz_i$ , implying no change, the ICM operation generates symbols that refer to the external reality's states,  $e_1, e_2, \dots, e_n$ , for each of these state changes ( $my \rightarrow mz_i$ ;  $1 \leq i \leq n$ ) at the time points when  $my$  occurs. From the meta-observer's viewpoint, if the ICM-operated sequence is replayed in the ordinary time direction, it represents the process in which the self in  $my$  discriminates between different states of the external reality ( $ei$ ) (Fig. 1). The inverse-causality operation requires several properties of the self that are not explicit in the above argument, including memory and function, to operate on the registered states and generate new states. This point is not discussed in this paper.

This derivation of the external reality is based on an objective principle involved in the measurement process in science, without recourse to the semantic contents of individual percepts. This process is based on a logic similar to that for the *it-from-bit* concept, in that perceptual differences (not semantic contents) generate symbols that refer to a reality that is foreign to the self ( $M$ ), in which the reality is not assumed to exist independently of the self, although the above process is not exactly the same as the *it-from-bit*

process discussed in section 3.4.2.

Bateson (1979) defined information as “any difference that makes a difference.” This definition appears to presuppose two entities as given, between which differences are related. As stated above, there is circularity between realization and observation. From a first-person viewpoint, we started from phenomena or data occurring to the self, constructed a reality model, and then understood the interaction between the phenomena and the reality. Once the self derived such symbols ( $\mathbf{E}_0 = \{e_0, e_1\}$ ), there were two entities, the self and its external reality (or the environment of the self), from the meta-observer's viewpoint, although the self can realize the reality only through its interior symbols.

### 3.3. Inverse causality at whole level (ICW)

We notice that seq. (2) does not fulfill the unique successor principle at the whole level, although it does for the determination of the succeeding elements of  $M$ . Consider “ $m_0 \rightarrow (m_1, e_0)$ ”, and “ $m_0 \rightarrow (m_1, e_1)$ ” in seq. (2). If we focus only on the elements of  $M$ , they are “ $m_0 \rightarrow m_1$ ” and “ $m_0 \rightarrow m_1$ ”, respectively; where is no violation of the unique-successor principle. However, if it is viewed from the holistic view, the sequence violates the principle. Therefore, inverse causality is used again for seq. (2) at the whole level (denoted as ICW). To fulfill the unique-successor principle, new symbols,  $e_2$  and  $e_3$ , must be introduced to the first and second  $m_0$ . Then, it gives

$$\underline{(m_0, e_2)}, (m_1, e_0), m_2, \dots, \underline{(m_0, e_3)}, (m_1, e_1), m_3, \tag{3}$$

where the underlined parts were modified.

By operating ICW, the symbols referring to the external reality  $E_1$  are found, whose state set  $\mathbf{E}_1 = \{e_0, e_1, e_2, e_3\}$ . Here,  $M$  in  $m_0$  changes to  $m_1$  whether  $E$  is in  $e_2$  or  $e_3$ . This implies that  $M$  in  $m_0$  does not discriminate between  $e_2$  or  $e_3$ . Introducing new symbols such as  $e_2$  and  $e_3$  to refer to the external reality's different states is a realization process by the self. After this process is completed, we, as a meta-observer, can see the resultant sequence as a process in which the self cannot discriminate between different states,  $e_2$  and  $e_3$ , because it changes from  $m_0$  to  $m_1$ , and then it discriminates between different states,  $e_0$  and  $e_1$ , changing to  $m_2$  or  $m_3$ , respectively. The inverse causality (if  $F(a) \neq F(b)$ , then  $a \neq b$ ) is the contraposition of the unique-successor principle (if  $a = b$ , then  $F(a) = F(b)$ ). (Note that this principle does not exclude the case where different elements are followed by the same element; that is,  $F(a) = F(b)$  when  $a \neq b$ .)

We still have the difficult issue of whether the existence of the external reality is justified by ICM/ICW. In a sense, it is hypothetical. However, the derivation is deduced from data using a logical rule used in scientific measurement. Therefore, it may be reasonable to claim that ICM or ICW discovers the external reality, as long as it is defined as what is directly or indirectly measured by the self (not as what is observed). If the self attempts to step out of a solipsistic world, the ICM assumption would lead it to a positivistic world, and then the ICW assumption would lead to a realistic deterministic world. Although it has been a significant issue in philosophy, as well as in science, to understand how the self can know (realize) the external reality, one possible approach to this issue, without necessarily being a resolution, is to make explicit assumptions in building the worldviews.

### 3.4. Interpretation of realization by ICW

Although ICW derives symbols referring to the external reality as a whole, it remains unclear how the self can differentiate the reality into individual objects situated in an interior space-time

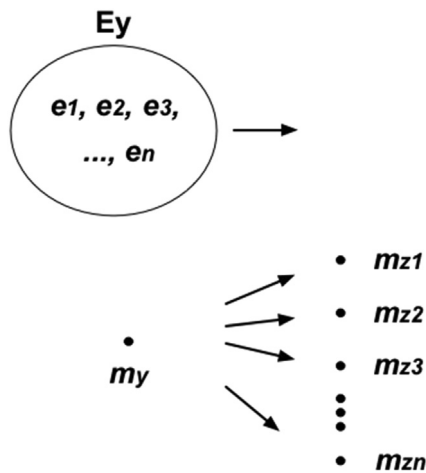
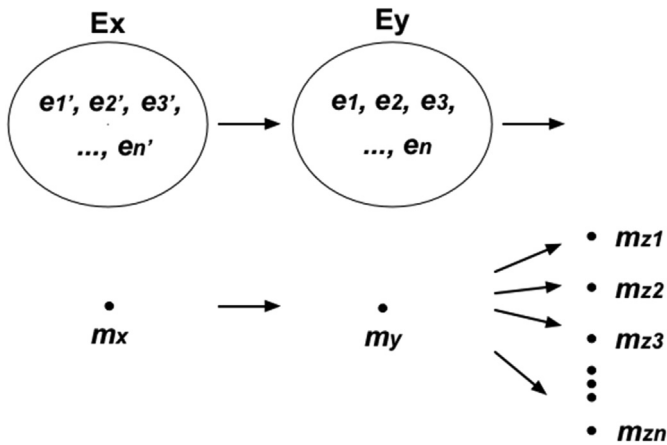


Fig. 1. Inverse causality on  $M$  sequence generates external reality's symbols ( $ei$ ). After the operation is completed, the process can be seen in the ordinary time direction (arrows) from the meta-observer's viewpoint.



**Fig. 2.** Inverse causality at whole level generates additional, external reality's symbols ( $e_i'$ ). After the operation is completed, the process can be seen in the ordinary time direction (arrows) from the meta-observer's viewpoint.

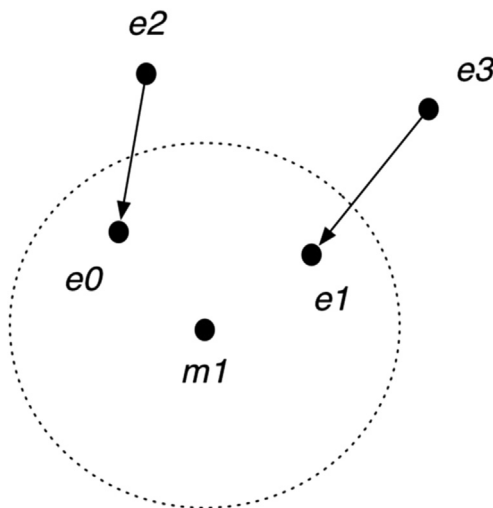
framework. Instead of addressing this difficult problem directly from within the self, the following discussion utilizes the strategy of giving possible interpretations of the ICW processes hypothetically.

We can represent the above case of reality construction by ICW (seq. 3) in a general form. Let us consider the following case, in which the self changes in state, such as  $mx \rightarrow my \rightarrow mz_1, mz_2, \dots$ , or  $mz_n$  (Fig. 2). As before, using inverse causality for  $M$  (ICM) generates the symbols  $e_i$  ( $\in E_y$ ;  $1 \leq i \leq n$ ) at time points  $my$ . By using inverse causality again on this sequence (ICW) to fulfill the unique successor principle at the whole level, additional symbols  $e_i'$  ( $\in E_x$ ;  $1 \leq i \leq n$ ) are generated at the time points when  $mx$  occurs.

How then can the self interpret the process when completed? In other words, what does this introduction of new symbols ( $E_x$ ) imply to the self? We can interpret how and why this process occurs in several ways from a meta-observer viewpoint. Using the example of seq. (3), let us discuss some major interpretations of this process.

3.4.1. Interpretation 1

Suppose that the external reality is a localized material entity. When the material object takes on the state  $e_0$  or  $e_1$ , it exists within the same locality in physical space with the self in  $m_1$  (Fig. 3).



**Fig. 3.** Different states ( $e_0, e_1$ ) of entities located within this locality can be discriminated by the self. Arrows show temporal changes. The dotted circle shows a local area for the self in  $m_1$ .

Therefore, the self can causally discriminate between them. However, the object positions  $e_2$  and  $e_3$  are located outside the locality, and therefore, the self in  $m_1$  cannot discriminate between them. When the object moves toward the local area of the self ( $e_2 \rightarrow e_0$ , or  $e_3 \rightarrow e_1$ ), the self can now discriminate between the positional differences,  $e_0$  and  $e_1$  (it is also possible that the self is moving toward the object).

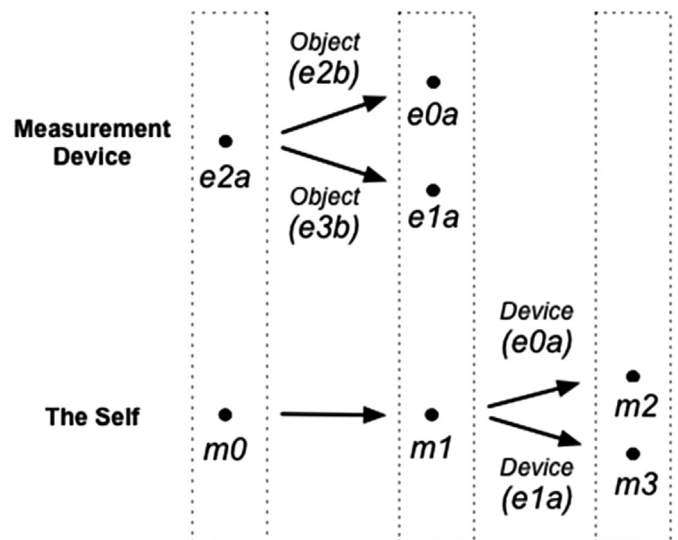
3.4.2. Interpretation 2

Suppose that the external reality is composed of two material entities. One is a measurement device, positioned in the locality of the self, and the other is an unknown object, whether in the locality or outside the locality. As illustrated in Fig. 4, let us rewrite seq. 3 by replacing “the external reality state” with a tuple of the states for a device and an unknown object, which gives the sequence of the whole states, “(self state, device state, unknown object state)”, such as

$$(m_0, e_{2a}, e_{2b}), (m_1, e_{0a}, e_{0b}), m_2, \dots, (m_0, e_{2a}, e_{3b}), (m_1, e_{1a}, e_{1b}), m_3, \dots,$$

where  $e_2 = (e_{2a}, e_{2b})$ ,  $e_3 = (e_{2a}, e_{3b})$ ,  $e_0 = (e_{0a}, e_{0b})$ ,  $e_1 = (e_{1a}, e_{1b})$ .

It should be noted that the device state is the same (i.e.,  $e_{2a}$ ) in both  $e_2$  and  $e_3$ , which is required for the measurement to work;  $m_0 \rightarrow m_1$  occurs for the device state of  $e_{2a}$ , implying that the self experiences that the external state is the same (such as “the device state is the same”), whereas the object takes on two different states ( $e_{2b}$  or  $e_{3b}$ ). The same cognition ( $m_0 \rightarrow m_1$  in this case) includes “the same experimental setup,” or “the same initial conditions in this experiment” for example. The device changes its state from  $e_{2a}$  to  $e_{0a}$  or  $e_{1a}$  by discriminating between different states of the object, i.e.,  $e_{2b}$  and  $e_{3b}$  (Fig. 4). The self can discriminate the states of the device ( $e_{0a}$  and  $e_{1a}$ ), although it cannot do so between the object states ( $e_{2b}$  and  $e_{3b}$ ). We can divide the device into more than two intermediate parts of a measurement system between which discriminations occur between different states of one part by another serially in time. To summarize, the self discriminates between different states of the unknown object indirectly by discriminating between the different states of the measurement device that directly interacts and discriminates the object's states.



**Fig. 4.** Detection (realization) of object by means of discrimination mediated by measurement device. Object state-changes are not shown in this diagram.

That is, the self obtains a *bit* about the object (e.g., the detection/non-detection of a photon) mediated through the device. As Wheeler stated, an elementary quantum phenomenon is “the elementary device-intermediated act of posing a yes-no physical question and eliciting an answer.” In the previous works (Nakajima, 2001, 2013), the *it-from-bit* process was described as equivalent to ICW, which is not correct. It is equivalent to ICW.

Many quantum physicists believe that our universe is indeterministic, although determinism and indeterminism are both metaphysical hypotheses that cannot be empirically determined (Nakajima, 2013). However, as far as we measure the states of something, quantum entities or anything else, using a measurement device, the “something” thus constituted is done under the assumption that the world fulfills the unique-successor principle. In other words, device-mediated measurement is equivalent to the ICW process, which is equivalent to the contraposition of determinism. This result seems to have a serious implication for science. It requires us to accept determinism as long as science constructs reality models of nature based on empirical supports from measurements using devices. This fact has not been clearly recognized so far in modeling natural phenomena.

However, we need to consider this implication carefully, because it does not imply that the future can be determined or predicted using data (percepts) from the past. This is because the deterministic state-sequence can be constructed for *past events* that have already happened. In other words, an *M* sequence will continue to generate new elements toward the future, which may possibly violate the unique-successor principle. In this sense, the world is open (i.e., indeterministic) to the future (Popper, 1982; Matsuno, 1996) or surrounded by an acausal environment (Rosen, 1991). Nevertheless, the fact that device-mediated measurement is logically equivalent to determinism poses the new problem of evaluating the role of determinism in modeling nature from data.

### 3.4.3. Interpretation 3

If we understand that mental states occur in a manner that depends on a particular sensory system, whose states occur in a manner that depends on the entities measured, then we can obtain another interpretation by slightly modifying the second one. It replaces the measurement device with a body that includes sense organs and neural systems, and replaces the unknown objects with the material objects around us, such as books, pens, and PCs. Without ICW, a brain cannot even realize these things surrounding us.

### 3.5. *It-from-bit* and Shannon's information in ICW/ICW process

In the realistic world model by Shannon, there are two entities, a subject that receives information about an object, which can take on various possible states in the form of a “message.” Shannon's information is not something conveyed by individual messages. It concerns the number of yes/no distinctions used to determine one message out of all of the possible ones. When a person receives a message that determines one among *N* possible states, the uncertainty of *N* possibilities (called entropy and formalized as  $\log_2 N$ , the logarithm of the number of yes/no distinctions) is then reduced to one by receiving the message. Here, the information relates to making a distinction, and the degree of uncertainty relates to the number of distinctions needed to determine the object. According to this interpretation, the amount of information is defined by “the degree of uncertainty before receiving information ( $\log_2 N$ )” minus “the degree of uncertainty after receiving information ( $\log_2 1$ ),” which is equal to  $\log_2 N$ , where *N* possibilities are reduced to one by the information (i.e., the distinction made by reading the message).

If the message is less discriminative, or less specific, containing *N'* different possibilities ( $N' \leq N$ ), then the amount of information becomes  $\log_2 N - \log_2 N' = \log_2 N/N' = -\log_2 p$ , where *p* is the probability of the subset of states that the message defines ( $N'/N$ ), under the assumption that each of *N* possibilities has an equal probability,  $1/N$ . In general, if state *i* occurs with a probability of  $p_i$  ( $1 \leq i \leq N$ ), then the entropy (*H*) is defined as the mean of the entropy for each state occurring at probability  $p_i$ ,  $H = -\sum_i p_i \log_2 p_i$ .

In the previous argument, ICW constitutes the symbols  $e1, e2, \dots, en$  by the perceptual state changes,  $my \rightarrow mz_1, mz_2, \dots, \text{or } mz_n$ , respectively (Fig. 1). For each change,  $my \rightarrow mz_i$ , the external state at time points of *my* is specified in a one-to-one correspondence. After the ICW operation is completed, the process can be replayed, or interpreted in a realistic model, in the ordinary time direction, which exhibits a process similar to the above process in Shannon's formalization. There are *n* possible states of an object, and an observer discriminates between them by a particular cognitive change such as  $my \rightarrow mz_i$ . By this cognition, the uncertainty ( $H = \log_2 n$ ) is reduced to 0 ( $H = \log_2 1$ ).

The ICW process constitutes an undetectable object (**Ex**) mediated by a constituted object (**Ey**); that is, the self in *my* constitutes **Ey** by the ICW process, and then constitutes **Ex** from **Ey** by ICW. If this process is replayed in the ordinary time direction (see the arrows in Fig. 2), it shows that the self in *mx* cannot discriminate between  $e1', e2', \dots, en'$  (i.e.,  $mx \rightarrow my$ ). In terms of Shannon's information, the self obtains no information about the external reality (in terms of ICW/ICW, “ $mx \rightarrow my$ ” cannot constitute **Ex**). There is a one-to-*n* correspondence in state between the self and the environment at the time points of *my*, implying that the uncertainty (*H*) is  $\log_2 n$ . This one-to-*n* correspondence then changes to a one-to-one correspondence (i.e.,  $H = 0$ ).

This process of *nondiscrimination followed by discrimination* in a replayed ICW process can be related to chance events for the observer. An observer in state *mx* observes an object, and changes to *my*. This state change (“ $mx \rightarrow my$ ”) indicates a cognitive “event” about the object occurring to the self. If the same cognitive event ( $mx \rightarrow my$ ) occurs repeatedly, and if the following events vary, such as  $my \rightarrow mz_1, mz_2, \dots, \text{or } mz_n$ , this one-to-many correspondence in event occurrence appears to the self as uncertainty about the object.

## 4. Realistic world model from meta-observer viewpoint

The replayed process of the ICW-derived sequence (Fig. 2) can occur within the self, which does not describe a process in a realistic world composed of many objects. The ICW model is modest and limited in its power of explanation (section 3.4), because it is constructed based on a given perceptual sequence of the self; it removes meaningful contents from perceptual states; and it does not explicate how symbols for the derived external reality can differentiate into those for component objects. However, the ICW model suggests a general aspect of a material entity required for building a realistic model of the information process. Namely, any entity, including the self, has a state at every moment of time; it can change its own state, and can change the state-relation to other entities by discriminating between their states. A better approach to a comprehensive theory of information requires modeling a subject-based world (system) model axiomatically by standing out of the self, based on the general aspect of the material entity, in which material entities are described as the subjects of motion or action.

### 4.1. Material entities and observers

The relationship between an observer and an observed system

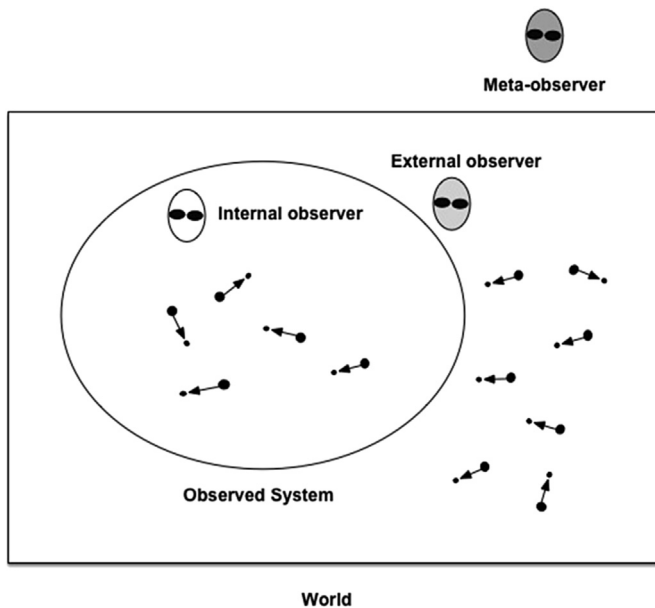


Fig. 5. Three types of observers. Dots indicate material entities, including living systems, whose arrows show their temporal changes in state (not necessarily velocity).

may include various situations (Fig. 5). First, given a world, which includes every material entity, living or non-living, in the universe, the meta-observer as a theoretical construct describes this world from outside. Scientists exist within the world, and observe objective systems of interest. There are at least two kinds of observers: external and internal observers. The external observer observes a set of entities from outside the boundary of an observed system. The boundary is made based on the degree of interactions. Although by definition the external observer does not causally affect entities within the boundary, entities located outside the boundary may affect the behavior of entities within the boundary, like noises, which cause indeterministic behaviors of the observed entities, even if the whole world is completely deterministic. An internal observer, usually a scientific observer, is located within the observed system, in which the observer interacts internally with the entities comprising the system. For example, quantum physical observations usually affect the observed system, in which observation is a material process within the observed system. From this meta-observer viewpoint, the phenomenological process can occur within the internal observer. Although, from a phenomenological viewpoint, this whole scheme (Fig. 5), or any other kind of observer-world model, can occur in a mental process of the self.

#### 4.2. Working definition of information

According to Young (1987), for hundreds of years prior to its entry into modern science, the word information had come to mean knowledge. However, after the early 1950s, in many scientific fields, “information” began to mean a certain kind of material process, in addition to meaning knowledge. Since then, various meanings have been attached to this word, such as imbuing form or shape to an object; putting in proper form or order; and arranging, composing, and giving form to the mind. The words related to the concept of information range between the epistemic and ontic (materialistic) realms. They are all important in describing the natural processes of physical, chemical, and biological entities, as well as realizing the reality in science.

Although we have not yet reached a common definition of information, we can find a common aspect of nature underling a

variety of these information-related words. Here, a working definition of information is given as follows: information is the state change of an entity that occurs in relation to the state of another entity, i.e., a relational state change between two entities, including particles and fields, in which the state can include internal states such as the pattern, form, or structure of an entity (see section 5.1). A pattern (form, structure) itself is not information. Similarly, stored memory and knowledge are not. Their change in relation to another entity (such as relational pattern formation, memorizing, knowing) is information. Here, the terms “knowledge” (or “knowing”) and “observation” (or “observing”) are used interchangeably—the use of knowledge does not necessarily connote its truth or justification, which includes incomplete knowledge about objects, although philosophers clearly distinguish knowledge from belief (e.g., Salmon, 1967).

Information has both qualitative and quantitative aspects. The qualitative aspect of information describes the specificity of a consequent relation by a relational change in response to the other's state; and the quantitative aspect describes the degree to which it is specified. Barbieri (2013) emphasized that organic (biological) information, observed in DNA replication or translation processes, is specificity, which is nominable, but cannot be measured, unlike Shannon's physical entropy. Although this characterization is suggestive and important, we need to incorporate both aspects of information into a single theory to effectively comprehend biological processes.

Relational state changes can occur between the following regions: (i) between material entities within an observer system (e.g., DNA replication occurs between DNA, deoxyribonucleotides, and enzyme proteins; from DNA through RNA to proteins in translation; from sense organs through neural networks to a brain); (ii) between an internal observer (scientist) and material entities within an observed system (e.g., between a researcher, measurement device, and quantum object); and (iii) from an observed system to an external observer (scientist) (e.g., from a DNA sequence to a researcher who reads it; from interacting material objects to physicists in a classical physical experiment) (Fig. 5). Finally, we add *it-from-bit* information as a unique case of the above definition, which is an inverse process of constituting a reality (or a reality state) as relational to a given subject's state change.

Information processes occur at various spatiotemporal levels of natural hierarchy; in biology, these levels include molecular or lower entities, cells, organisms, groups of organisms, ecosystems, evolving populations forming lineages, and phylogenetic patterns (Salthe, 1985; Nakajima, 2004). The following argument focuses on the material information process at a particular level of a nested biological hierarchy, not on the processes that cross between different levels of the hierarchical organization.

### 5. Information and probabilities of events

This section describes a general mathematical framework that is applicable to the process of producing relational changes in state or a state-pattern between entities (i.e., information), with a particular focus on (i) how a relational state change can occur between material entities and (ii) how such a relational state change can determine the probabilities of events occurring to a subject entity in interaction with the environment.

#### 5.1. Cognizers-system model

As previously demonstrated (section 3), the ICM/ICW process constitutes symbols referring to the external reality (set  $E$ ), given a temporal sequence of percepts ( $M$  with set  $M$ ). The consequent constituted sequences are generated downstream of  $M$ , i.e., within

the self. Extending this model hypothetically, a world that includes the self and external reality can be modeled from a meta-observer stance. As such a model, the cognizers-system model (CS model) may be useful to describe the information processes within a material world, including the processes between material entities within an observed system and between an internal or external observer and an observed system (Fig. 5) (Nakajima, 1999, 2007, 2013). In the CS model, the world is composed of material entities, called cognizers  $C_i$  (Fig. 6). The meta-observer describes the model, which does not interact materially with the world. Internal and external observers are also cognizers within the world. An internal observer is a cognizer that can interact with cognizers within the observed system. An external observer is a special case of a cognizer that does not affect the observed system, at least during the period of observation. However, an external observer can interact with the observed system because of its materiality when setting up experimental systems and giving them a particular initial condition.

The CS model is summarized as follows. Each cognizer takes on a particular state at every moment, and its state changes according to the current states of the cognizer itself and all other cognizers in the world (“all other cognizers” are also called the “cognizers in the environment” or “environmental cognizers”), based on the cognizer’s properties (formalized with motion function  $f_{ci}$ ), which uniquely determine the succeeding state in time. Consider a focal cognizer ( $C_1$ ) with motion function  $f_{c1}$  and state space  $\mathbf{C1}$ , and the environment ( $E$ ) with motion function  $f_E$  and state space  $\mathbf{E}$  (Fig. 7).  $C_1$  in state  $cx$  changes to  $cy$ , i.e.,  $f_{c1}(cx, ex) = cy$ , and  $E$  in state  $ex$  changes to  $ey$ , i.e.,  $f_E(cx, ex) = ey$ , where  $f_{c1}: \mathbf{C1} \times \mathbf{E} \rightarrow \mathbf{C1}$  and  $f_E: \mathbf{C1} \times \mathbf{E} \rightarrow \mathbf{E}$ . (In this paper, italicized capitals are used to denote cognizers, and bold-faced capitals denote their state sets. A space means a set with a particular topological structure such as distance.)

In this formalization, a cognizer’s state-change from  $ci$  to  $cj$  (including  $i = j$ , implying no change) is called “cognition.” Like a

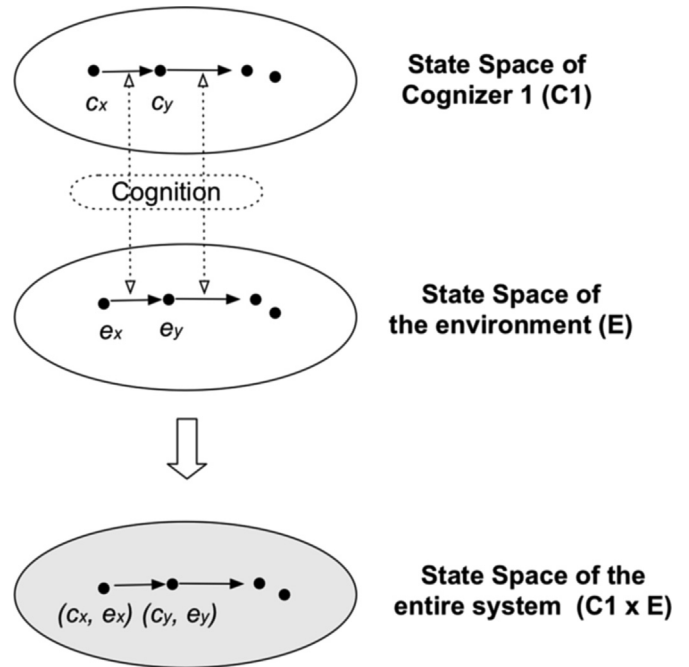


Fig. 7. Schematic diagram of cognizers system, composed of focal cognizer,  $C_1$ , and its environment,  $E$ . The environment is a cognizer, which may be composed of many cognizers such as  $C_2, C_3, \dots$ . Arrows indicate temporal state-changes of component cognizers by cognition, which is formalized as  $f_{ci}(ci, ei) = cj$ ;  $f_E(ci, ei) = ej$  ( $i = x, j = y$  in the figure). The state transition of the entire system is given as  $\dots, (ci, ei), (f_{ci}, f_E)(ci, ei), (f_{ci}, f_E)^2(ci, ei), \dots$ , where “ $(f_{ci}, f_E)$ ” is defined as  $(f_{ci}(ci, ei), f_E(ci, ei))$ .

perceptual sequence of  $M$  (section 3.1), cognition  $cx \rightarrow cy$  may be divided into micro-scale processes such as  $cx = cx_1, cx_2, \dots, cx_k$  and  $cy = cy_1, cy_2, \dots, cy_k$ , where  $ci = cj$ , if and only if  $ci_h = cj_h$  ( $1 \leq h \leq k$ ). The CS model is a deterministic system in order to clarify the mechanism of how a chance event can occur to internal entities (Nakajima, 2013). However, if we want to model an observed system that interacts with the outside unspecified entities (e.g., noises) within the whole world (Fig. 5), it is possible to construct an indeterministic CS model by incorporating an unspecified entity into the model, which affects component cognizers indeterministically from outside the system.

If the environment is composed of  $n-1$  cognizers, then  $ex$  is represented as a  $[n-1]$ -tuple of the states of these  $n-1$  cognizers ( $C_2, C_3, \dots, C_n$ ) existing in the environment, such as  $ei = (c2i, c3i, \dots, cni)$ . Here, “ $ei$ ” indicates “the state of cognizers in the environment” (or simply, the “environmental state”). The relation between two or more cognizers ( $C_1, C_2, \dots$ ) can be defined in terms of subsets of the direct product ( $\prod \mathbf{Ci}$ ) of their state spaces ( $\mathbf{Ci}$ ).

Furthermore, each cognizer can be modeled as made up of cognizers at a lower level. For example, an animal is modeled as composed of sub-cognizers such as a brain, sensors, neural systems, and effectors. By doing so, a focal cognizer can be realized as consisting of component cognizers that inter-cognize one another. In this case, the internal states of a cognizer can be defined in terms of subsets of the direct product of the state spaces of the component cognizers (Appendix 1). Such internal states include memories and organizational structures. In this paper, living systems (or living entities) are viewed as cognizers because they are composed of many parts, although the use of “system” might be terminologically confusing in the scheme of Fig. 5.

The CS model focuses on two general aspects of the properties of a motion function: selectivity and discriminability. Selectivity is an aspect of a cognizer’s property regarding which state a cognizer

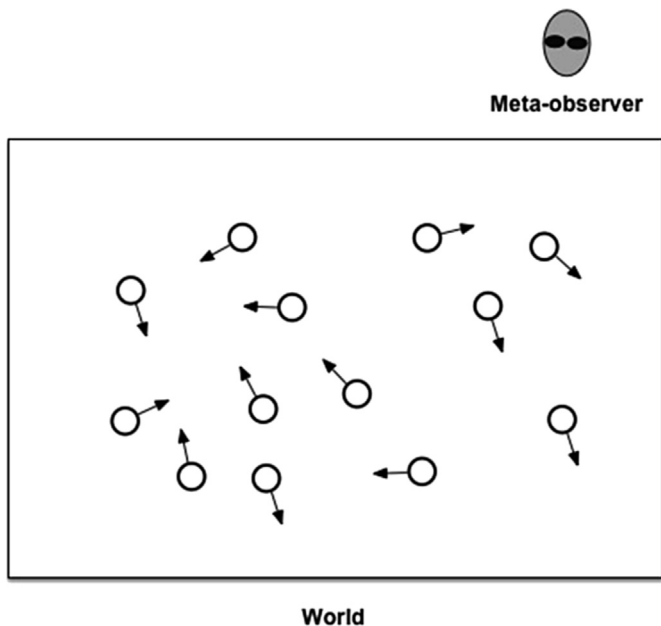


Fig. 6. Material entities are formalized as cognizers, which include internal and external observers, shown in Fig. 5, as special cases. Circles represent cognizers, and arrows indicate their state changes in an abstract sense, not velocity. State includes internal states (e.g., memory, structure) for cognizers that are composed of many cognizers at a lower level. The whole world is described as a deterministic system by the meta-observer.



selects (determines) as a succeeding state under a given environmental situation. When  $C1$  in state  $cx$  changes to  $cy$  when the environmental state is  $ex$ ,  $C1$  in  $cx$  selects  $cy$  to  $E$  in  $ex$ . Similarly,  $E$  in  $ex$  selects  $ey$  to  $C1$  in  $cx$ . Therefore, selectivity focuses on the aspect of the motion function that affects the relationship between a subject cognizer and its environment through interaction (inter-cognition). Selectivity relates to the value and meaning in the framework of the CS model, which will be focused on later (section 6.3).

Discriminability is another aspect of a cognizer's property and relates to the ability to discriminate between different states of the environment. When  $C1$  in state  $cx$  changes to different states against different environmental states,  $ex$  or  $ex'$ ,  $C1$  discriminates between  $ex$  and  $ex'$  (i.e.,  $fc1(cx, ex) \neq fc1(cx, ex')$ ). It should be noted that cognition in the CS model is not identical with discrimination; the former includes both discriminative and non-discriminative state changes against different environmental states. In terms of selectivity, discrimination can be expressed as *selecting different succeeding states in response to different environmental states*. Therefore, selection and discrimination can be understood to be different aspects of the same process of cognition.

When viewed from the ICM/ICW model, cognizers and their motion functions are hypothetical constructs that may not be derived uniquely from a finite set of perceptual sequences obtained in the past, and therefore, may not be justified as true (Hume, 1777/1902; Salmon, 1967). However, such a formal model seems to be essential to give a mechanistic explanation for observations in science (Cassirer, 1923). From the viewpoint of the CS model, the ICM/ICW model can be understood to be a special case of the former model in that the ICM/ICW model describes only a single cognizer or its part, whose state-change process can be situated in the entire system of cognizers, as interpreted in section 3.4.

## 5.2. Discrimination between patterns

In the above definition, discrimination concerns the state differences of all of the environmental cognizers. Generally, there are many cognizers comprising the environment, such as  $C2, C3, \dots, Cn$  as above. Therefore, we need to consider the application of the discrimination concept to individual cognizers in the environment. Let us consider a simple example of three cognizers  $\{C2, C3, C4\}$  comprising the environment, each of which can take on two states  $\{0, 1\}$ . There are  $2^3$  different states ( $\{e1, e2, \dots, e8\}$ ) for the environmental cognizers for a focal cognizer  $C1$  in state  $cx$  (Fig. 8). Here, the focal cognizer's cognitions (responses) to the environmental

states are classified using different colors for the table rows in Fig. 8; in other words, the states ( $ei$  and  $ej$ ) with the same color are not discriminated between by the focal cognizer  $C1$ . If the focal cognizer discriminates every environmental state ( $ei$ ) from the others ( $ej, i \neq j$ ), this discrimination occurs for all three cognizers in the environment (Fig. 8A). However, as in the case of Fig. 8C, discrimination occurs for only  $C2$ , because the cognizer's responses are determined only by whether  $C2$ 's state is 0 or 1, irrespective of the others' states. Fig. 8B shows an intermediate case, in which the focal cognizer discriminates between different states of only  $C2$  and  $C3$ . Fig. 8D shows that the discrimination occurs for all three cognizers, but only  $e8$  produces a different cognition, while the other seven states are not discriminated.

Notably, the environmental states can be described as *patterns* of the states assumed by the component cognizers. This is also true for the focal cognizer's states. A cognizer composed of more than two component cognizers has internal states such as memories of experiences that have occurred. Therefore, the cognition of a focal cognizer ( $cx \rightarrow cy; fc1(cx, ex) = cy$ ) involves a correlated change in an internal pattern/structure with the environment, which is an information process, as previously defined (section 4.2.).

As before, internal and external observers are also cognizers within the world (Fig. 6). Since an internal observer interacts (or inter-cognizes) with entities within the observed system, the formalization is the same as previously given. For the external observer, the environment is the observed system, and component entities within the environment do not discriminate between different states of the observer. Information flows in one direction, from the observed system (the environment) to the observer. Therefore, the motion function of the observed system in this case can be modified as follows:  $f_E(\bullet, ei) = ej$ , where “ $\bullet$ ” indicates a constant, or simply as  $f_E(ei) = ej$ , where  $f_E: E \rightarrow E$ .

## 5.3. Local causality and choice of objects for discriminative actions

The patterns involved in discrimination (Fig. 8) can be related to the local causality. Discriminability is related to the locality of the state space of the cognizers, under the assumption that causation can occur between entities only in the physical locality, known as local causality, which has been widely accepted in physics. In the case of Fig. 8C, it can be interpreted that the focal cognizer shares the same locality with cognizer  $C2$ , but might not with others,  $C3$  and  $C4$  (although the determination of the latter point requires the discriminability of  $C3$  or  $C4$  regarding  $C1$  states). Note that “discrimination” in the language of the CS model refers to

A	B	C	D
	C2	C3	C4
e1	0	0	0
e2	0	0	1
e3	0	1	0
e4	0	1	1
e5	1	0	0
e6	1	0	1
e7	1	1	0
e8	1	1	1

**Fig. 8.** There are  $2^3$  different states ( $\{e1, e2, \dots, e8\}$ ) of environment composed of  $C2, C3$ , and  $C4$ , with each taking on two states (0 or 1), for focal cognizer  $C1$  in state  $cx$ . The focal cognizer in  $cx$  cognizes each of these eight environmental states, and it changes to a succeeding state. According to the succeeding states, the cognizer's discriminative cognitions in response to the environmental states are classified using the different colors of the table rows; the states ( $ei$  and  $ej$ ) shown using the same color are not discriminated between by the focal cognizer  $C1$ . Only four cases are shown (A–D).

discrimination between different states of a given entity, not between the presence and absence of a given entity within the system. However, the latter case of discrimination can be represented as discrimination between an entity's state within the same locality and states outside the locality.

The patterns (Fig. 8) can be related to another factor, i.e., the choice of objects for discriminative actions by a focal cognizer. Suppose that C2, C3, and C4 all exist in the same locality of a focal cognizer's position. A comparison between Fig. 8A, B, and C suggests that the focal cognizer (C1) chooses the object cognizers (C2, C3, and C4) for taking discriminative actions: C1 chooses C2, C3, and C4 in 8A; C2 and C3 in 8B; and only C2 in 8C. Here, "choose" is used instead of "select" because the latter is used to mean the determination of a succeeding state in the CS model (section 5.1). As previously quoted, Uexküll emphasized that every animal selects (chooses) stimuli from the general influences of the outer world. In other words, an animal chooses a particular subset of entities in the environment, to which it acts appropriately (i.e., selectively) by discriminating between their states. This choice implies that the subject responds to a particular pattern of the environment in a particular way, which is an information process within the world, constituting an internal model of the external reality for the subject organism.

5.4. Probability of events in CS model

Although it has been widely recognized that the events occurring to a given actor such as a living system are affected by the interaction between the actor and the environment, none of the traditional theories of probability can effectively deal with the probabilities of the events occurring to the actor (Nakajima, 2013). In the deterministic CS model, the probabilities of the events that occur to a focal entity are explained as follows. As shown in Fig. 9, a focal cognizer C1 in  $cx$  does not discriminate between different  $n$  environmental states ( $\{e1, e2, \dots, en\}$ ) and changes to  $cy$  (i.e.,  $cx \rightarrow cy$  for  $e1, e2, \dots, en$ ). Likewise, the  $E$  in each state ( $\in Ex$ ) changes to the succeeding states of the environment ( $\in \{e1', e2', \dots, en'\} = Ey$ ). The focal cognizer's cognition ( $cx \rightarrow cy$ ) is a conditional event, which is repeated in the entire cognizers system; under this conditional event, a variety of events occur subsequently: C1 in  $cy$  encounters  $n$  different environmental states ( $\in Ey$ ), then it changes to  $cz1$  for  $e1', e2', e3'$ , or to  $cz2$  for  $e4', e5', \dots, en'$ ; in other words, C1 discriminates between states belonging to two subsets,  $Ey1 = \{e1', e2', e3'\}$  and  $Ey2 = \{e4', e5', \dots, en'\}$ , but not between elements within the subsets. C1 experiences two events as

cognitions,  $cy \rightarrow cz1$  at a probability of  $3/n$  and  $cy \rightarrow cz2$  at a probability of  $(n-3)/n$ ; the former cognition corresponds to  $Ey1$ , and the latter cognition corresponds to  $Ey2$ . This process defines the probabilities of events that occur to a material entity in interaction with the environment, which is conceptualized as the internal probability in the sense that the probability of an event occurring to an entity is internal to the system (Nakajima, 2013). The internal probability can be measured in terms of the relative frequency of an event that follows a particular cognition in question ( $cx \rightarrow cy$ ) during a given state sequence of the system.

In this process, if C1 has a different ability to discriminate between the states ( $\in Ex$ ) (for example,  $cx \rightarrow cy1$  for  $e1, e2, e4$ , and  $cx \rightarrow cy2$  for others), then the consequent environmental states that C1 encounters under each conditional event would become different; an increase in discriminability reduces the size of the state set, i.e., uncertainty. Therefore, the discrimination ability of entities is an important property that affects the probabilities of the events occurring to them. The role of the discrimination ability of entities in organization and adaptation is widely recognized in biological systems, ranging from molecules through cells to organisms (e.g., Wolfenden, 1972; Koshland, 1973; Fersht, 1987; Stewart and Dahlquist, 1987; Lerner and Tramontano, 1988; Dusenbery, 1992). The above formulation of internal probability can also be used when considering the internal states of a cognizer (Appendix 1).

An "event" should be defined in the framework of the CS model. As before,  $ci \rightarrow cj$  is called cognition. However, cognition occurs at the same time as an event to the cognizer, because the cognition ( $cx \rightarrow cy$ ) corresponds to the environmental states of  $Ex$ , which implies that  $cx \rightarrow cy$  refers to the environmental states ( $\in Ex$ ) that occur to C1. Therefore, a particular cognition defines a particular event in relation to the environment for the subject. For example, when tossing a coin repeatedly, we observe two events, heads or tails, but each of these events includes many states of the environment (including the coin and a table). We observers recognize these in terms of two subsets, "heads" and "tails." In other words, cognitions partition the environmental states (e.g.,  $Ey$ ) into equivalence sets (e.g.,  $Ey1$  and  $Ey2$ ), with each corresponding to a single event occurring to the subject. This partitioning can vary according to the subjects. This definition of an event is compatible with the definition of events in the mathematical theory of probability.

Here, it is useful to compare the event concept with a measurement by ICM (Fig. 1), in order to understand the difference between a phenomenological view and a meta-observer view in the CS model. In an ICM process, the inverse causality on a given  $M$  sequence (interpreted as measurement) generates the external reality's symbols  $ei$  at the time points of  $my$ , from  $my \rightarrow mz_i$  ( $1 \leq i \leq n$ ) (Fig. 1).  $my \rightarrow mz_i$  are phenomenological events for the self, referring to events concerning the external reality as  $ei$ . In the CS model, an event is a subset of the environmental states whose elements cannot be discriminated between by the focal cognizer (who is ignorant about them) in that state. Therefore,  $ei$  in the ICM process corresponds to a subset of the environmental states such as  $Ex$ ,  $Ey1$ , and  $Ey2$  in the CS model.

Finally, the internal probability (the measure of certainty of an event occurring to a material entity) can be extended to an overall collection of cognitions for a focal cognizer, including different conditional cognitions, in order to indicate how often the cognizer experiences an event in question in the total population of events occurring during a given state sequence of the system (Nakajima, 1999: p88-89; Nakajima, 2013: p76). To avoid conceptual confusion, let us call this measure as the overall internal probability ( $P_{overall}$ ), and the measure relative to a conditional cognition (such as  $cx \rightarrow cy$  in Fig. 9) as the cognition-based internal probability ( $P_{cog}$ ).

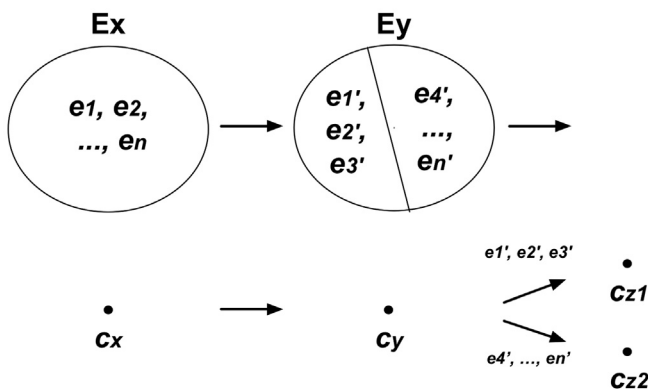


Fig. 9. Mechanism for generating uncertainty of event occurrence to focal cognizer in CS model. Non-discrimination ( $cx \rightarrow cy$ ) between differences ( $e1, e2, \dots$ ) brings about uncertainty of events ( $cy \rightarrow cz1$  or  $cz2$ ) to a focal cognizer in  $cx$ .

### 5.5. Increase in probabilities of favorable events by semiosis

Semiosis is an adaptive mechanism of living systems used to maintain an appropriate relationship with the environment (Uexküll, 1926, 2010; Salthe, 2001, 2014; Hofmeyer, 2007; Nakajima, 2007). However, we have not reached a comprehensive mathematical model for adaptation by semiosis in biological systems. Here, an attempt is made to represent some aspects of semiotic processes by focusing on the increase in the probabilities of favorable events by semiosis using the CS model. The general scheme of internal probability formalized in Fig. 9 (section 5.4) does not make explicit the topological structure (i.e., distance) of the state spaces of cognizers. However, the discriminability of cognition is fundamentally affected by the distance between the positions of focal and target entities. Semiosis connects local and nonlocal processes as follows.

Consider a focal cognizer  $C1$ , and assume that  $C1$  needs to have a certain state relationship with cognizer  $Co$ . Under the assumption of local causality,  $C1$  cannot discriminate between different states of  $Co$  when its position is outside the locality of  $C1$ . However, if a certain cognizer ( $Cs$ ) is within the locality, whose states are correlated with  $Co$ 's state by discriminative and selective actions (cognitions) between them,  $C1$  can discriminate between the different states of  $Co$  mediated by the discrimination between the different states of  $Cs$  as a sign. In the discussion using Fig. 8, we can consider another case in which there is a correlation between the states of environmental cognizers ( $C2$ ,  $C3$ , and  $C4$ ) (Fig. 10). In this case, the states of  $C2$  and  $C4$  are correlated. Therefore, if the focal cognizer cognizes  $C2$  states, these uniquely determine the  $C4$  states, whereas the relationship is random between  $C2$  (or  $C4$ ) and  $C3$ . Therefore, if  $C2$  (or  $C4$ ) is an important object for  $C1$  to survive,  $C4$  (or  $C2$ ) plays an important role in the survival of  $C1$  within a world of local causality.

Generally, there are many entities within and without the locality of a focal subject, and there may potentially be many entities in a given locality of the subject that are significantly correlated with non-local entities. Semiotic action involves a choice of a particular local entity (or a subset of local entities) that is correlated to non-local entities (such as natural enemies, resources, or mates) relevant to the subject entity's activity of system maintenance—these relevant entities vary according to the subject living system. The semiotic action enables living systems to discriminate between state differences of their non-local surroundings by local causality, which is one of the most important biological properties of living systems as teleonomic systems. The signal and tag information in molecular biology may be a similar

	$C2$	$C3$	$C4$
$e1$	0	0	0
$e2$	0	1	0
$e3$	1	0	1
$e4$	1	1	1

**Fig. 10.** Environment of focal cognizer  $C1$  is composed of  $C2$ ,  $C3$ , and  $C4$ , taking on 0 or 1.  $e1$ ,  $e2$ ,  $e3$ , and  $e4$  are the states of the environment that actually occur in the system, in which the states of  $C2$  and  $C4$  occur dependently.  $C1$  can use  $C2$  as a sign for  $C4$ , and vice versa.

mechanism (Holland, 1995, 2012).

The “relational state (or state-pattern) change between two entities” given as a working definition of information (section 4.2) requires at least two entities. However, in the above process, there are three entities: a subject ( $C1$ ), second entity in the locality ( $Cs$ ), and third entity outside the locality ( $Co$ ). Peirce highlighted this triadic relation between entities (including mental concepts) in his semiotics. In Peircean semiotics (Chandler, 2007), sign is the unity of the following three concepts: (i) the representamen (how the object is represented), sometimes called the sign vehicle (or simply, sign; which is similar to Saussure's signifier, e.g., the sounds of a word) (Saussure, 1916/2013); (ii) an interpretant (how the object is represented), which is similar to Saussure's signified (e.g., the concept of a word) and is related to how the focal cognizer cognizes other cognizers (i.e., how to act/move); and (iii) an object (what is represented, a referent). The interaction between the representamen, interpretant, and object is called semiosis. Although Peirce's object can include abstract concepts and functional entities, the Peircean model explicitly allocates a place for materiality and reality outside the sign system, which Saussure's model did not directly feature (Chandler, 2007). Hoffmeyer (2007) called such a sign-mediated adaptive process “semiotic causation,” which operates through the mechanisms of material efficient causation.

## 6. Information and entropy

In this section, the framework of the CS model is used to explore (i) how a relational state change can reduce the uncertainty of the environment for a subject entity by introducing the concept of internal entropy, (ii) whether internal entropy is always decreased by an increase in adaptation, and (iii) in what sense a relational state change of entities (i.e., information) can have meaning.

### 6.1. Internal entropy: uncertainty and disorder of Umwelt

Uncertainty is based on the probability concept, but they are not identical concepts. Probability focuses on a particular event, for which the degree of certainty is measured under a particular condition (Nakajima, 2013)—the condition varies according to the probability interpretation, which includes knowledge or observation in subjective (epistemological) theory, a collection of data or events in frequency theory, an experimental setup in propensity theory, a conditional cognition or action ( $cx \rightarrow cy$ ) for a material entity in internal probability theory ( $P_{cog}$ ), and a collection of cognitions (events) for a material entity in internal probability theory ( $P_{overall}$ ). Uncertainty focuses on the diversity of events, with their probability distribution occurring under a particular condition. Shannon provided a measure for uncertainty known as entropy  $H = -\sum_i p_i \log_2 p_i$ , where  $p_i$  is the probability of event  $i$ .

The meaning of  $H$  may vary according to the interpretation of probability. If  $p_i$  is interpreted as the subjective probability,  $H$  measures the uncertainty of events occurring to an external subject, whose values may change according to the degree of knowledge. If  $p_i$  is interpreted as the relative frequency,  $H$  measures the disorder of the system behavior. There is a controversy in interpreting what entropy means in physics, concerning whether it is epistemic or ontic (Denbigh and Denbigh, 1985; Ben-Naim, 2012). This problem arises from the dual meaning of entropy in statistical mechanics and information theory, which appears to have been inherited from the duality of the probability concept, called the Janus-faced character of probability (Gillies, 2000; Hacking, 2006). Its epistemic aspect concerns uncertainty, while the ontic aspect concerns disorder; these may correspond to regarding information as knowing, which decreases uncertainty, and information as a material process that decreases disorder of a system, respectively.

Now, what if  $p_i$  is interpreted as the internal probability of event  $i$ ? If it is the cognition-based internal probability ( $P_{cog}$ , section 5.4), entropy measures uncertainty of events relative to a given cognition. This is denoted as *the cognition-based internal entropy*,  $H_{cog}$ .  $H_{cog}$  can measure the uncertainty of the environment that a subject experiences *under a given cognition* ( $cx \rightarrow cy$ ). In Fig. 9,  $H_{cog}$  is calculated as  $(3/n)\log_2(n/3) + ((n-3)/n)\log_2(n/(n-3))$ . This value may differ according to the material properties of the subject and the environmental entities. However, it cannot be affected by the external observer's knowledge (or observation). In this sense, it is objective. We can consider another type of internal entropy, based on the overall internal probability ( $P_{overall}$ , section 5.4), which is denoted as *the overall internal entropy*,  $H_{overall}$ . If a cognizer experiences only a single event following each conditional cognition during a given period of time, no uncertainty exists of an event occurring after each cognition (i.e.,  $H_{cog} = 0$ ). However, even during this period, the cognizer may experience different events at different frequencies, yielding a positive  $H_{overall}$ . For example, an omniscient Laplace's demon that observes coin tosses will experience no uncertainty of consequent events under any given observation ( $H_{cog} = 0$ ). However, the demon experiences roughly 50% heads and 50% tails in the coin-toss sequence, yielding  $H_{overall} = (1/2)\log_2 2 + (1/2)\log_2 2$ . Unlike the demon, cognizers can affect event probabilities by discriminative and selective cognitions (actions). For example, a skillful tosser can manipulate the coin toss to bias the probabilities (Clark and Westerberg, 2009), in which  $H_{overall}$  can vary according to the tosser's preference. Like  $H_{cog}$ ,  $H_{overall}$  is also determined by material properties of the subject, as well as environmental entities and their interactions, although they generally yield different values.  $H_{cog}$  and  $H_{overall}$  can measure uncertainty and disorder aspects of Umwelt, respectively, that a subject organism experiences.

## 6.2. Internal entropy does not necessarily decrease by adaptation

In the CS model, knowing (observing) and action (movement) are unified as a state change of the entity, called cognition, within the observed system (Fig. 6). Therefore, what decreases the uncertainty and what decreases the disorder of a system both arise from the relational state change between two entities by the discriminative and selective cognition/action of entities. Maxwell's demon may illustrate this point (Maxwell, 1871/2001). This imaginary entity is described as an internal material observer, which *discriminates* between different states of particles and acts on them *selectively* to produce an ordered organization of the system. Whether such a demon is physically possible or not is an empirical problem, not a logical one. From a logical point of view, we can learn from this thought experiment that epistemic information (i.e., the reduction of uncertainty by observation) is connected through action with ontic information (i.e., the reduction of disorder by pattern transmission), if knowing is a material process within the system.

The cognitive ability of Maxwell's demon decreases both  $H_{cog}$  and  $H_{overall}$ , suggesting that biological adaptation decreases both types of internal entropy because of discriminative and selective action that increases the internal probabilities (cognition-based or overall) of favorable events, while decreasing those of unfavorable events (section 5.4). However, is it true? To answer this question, let us consider an experiment involving taking a ball from a opaque box. There are 10 balls, including seven red and three yellow balls, in the box. A person repeatedly removes one ball and returns it to the box after recording the color. Perceiving the ball's color is an event occurring to the person, whose probability is the internal probability since there are material interactions, though very simple, between the person and the balls (we could also consider a

complex case using ball-like robots equipped with visual sensors and motors, or living animals such as mice, which move in a specific way in response to the person's hand in the box). If the person takes out a ball randomly because the box is *opaque*, there is 10 degree of uncertainty to determine one. Therefore, the cognition-based internal probability of red and that of yellow are 7/10 and 3/10, respectively. Therefore  $H_{cog}$  is given as  $(7/10)\log_2(10/7) + (3/10)\log_2(10/3) \approx 0.881$ . Their overall internal probabilities ( $P_{overall}$ ) are expected to be 7/10 and 3/10, respectively, because each cognition in the repeated trials occurs in a similar manner from the same box. Therefore,  $H_{overall} \approx 0.881$ .

Now, consider another experiment in which the person can discriminate with some degree of precision between different positions of the red and yellow balls using a *semi-transparent* box (i.e., increased discriminability), and where the person likes yellow balls (i.e., non-neutral selectivity) because they are valuable to him or her. In this case, the person can discriminate between different positions of differently colored balls with some degree of precision. We can interpret this experiment biologically. These balls represent environmental entities that a focal living system (a person in the experiment) encounters. Red balls may be predators and yellow ones are prey or mates; or, either color may be prey, but with different dietary qualities (values). Therefore, in the second experiment, if the person is adaptive, the internal probability ( $P_{cog}$ ) of red would be less than 7/10, say 2/5, and that of yellow would be 3/5. Here, the increased discriminability reduces  $|Ex|$  in Fig. 9 from 10 to 5, and the selectivity for yellow is reduced from 7 to 2. The internal entropy ( $H_{cog}$ ) is then given as  $(2/5)\log_2(5/2) + (3/5)\log_2(5/3) \approx 0.971$ . The  $P_{overall}$  values are expected to be the same as the  $P_{cog}$  values due to the reason mentioned in the first experiment. Therefore  $H_{overall}$  is also expected to be the same as the  $H_{cog}$  value. In the CS model, a person can be viewed as a cognizer composed of component cognizers, such as sense organs, sensory neurons, the brain, motor neurons, effectors (e.g., muscles), and others, with a particular topological structure. A given cognition ( $ci \rightarrow cj$ ) involves micro-scale interactions among these component cognizers at a micro-timescale through which sensory data are used to produce an appropriate action of effectors.

The comparison between the first and second experiments suggests that a discriminative and adaptively selective action does not necessarily reduce the entropy of events that the person experiences, compared to when the action is random. Of course, if the discriminative and selective ability to obtain yellow is much higher, and if the probability of taking yellow ones becomes higher than 7/10, then the entropy will become lower than that in the random-taking case. However, it is important to learn from this case that an increase in the degree of adaptation does not necessarily decrease the internal entropy. This is because events can have different effects on the performance of a living system. Adaptive evolution leads a living system's properties to increase the internal probabilities ( $P_{cog}$ ,  $P_{overall}$ ) of a particular set of events favorable (valuable) to the system maintenance (survival) and reproduction. Therefore, if a starting stage is a low internal entropy ( $H_{overall}$ ) with a particular probability distribution of events, and if moving to the most favorable probability distribution requires crossing the highest entropy stage, the entropy would then increase toward the peak.

The relationship between adaptation and the reduction of uncertainty is controversial (Friston et al., 2012). Friston (2010, 2012) proposed that any self-organizing system that is at equilibrium with its environment must minimize its free energy. The free-energy principle is used in statistical estimation (e.g., variational Bayesian methods), derived from statistical mechanics, based on Kullback–Leibler divergence, which is a measure of the difference between two probability distributions. Approximate Bayesian

inference is performed, given some data and a model of how the data were generated. Free energy is an upper bound on surprise (prediction errors), in which the minimization of free energy implicitly indicates the minimizing of surprise (equivalently, the maximization of expectation. This theory asserts that the principle of free energy unifies all adaptive autopoietic and self-organizing behavior under one simple imperative, “avoid surprises and you will last longer.”

It seems reasonable that living systems act to increase the degree to which predictions match outcomes (i.e., reduction in  $H_{cog}$ ), as long as the outcomes benefit system persistence and reproduction. However, the overall-internal probability distribution of beneficial events does not necessarily yield lower entropy in the sense of  $H_{overall}$ : From an ecological point of view, various means/solutions exist for living systems to relate them to environmental entities to maintain their own organization effectively and reproduce more in a given ecological community. For example, organisms of generalist species are expected to have higher  $H_{overall}$  values than those of specialists in a given local community, because generalists have a wider range of habitat choice and diet preference than specialists as an adaptive strategy (Pianka, 1999). This suggests they should have a good internal probability distribution ( $P_{overall}$ ) of events, which is understandable by considering a case in which red and yellow balls are both essential resources in the aforementioned experiment. Therefore, we can conclude that adaptation does not necessarily reduce  $H_{overall}$ .

We must next critically examine another issue of whether it is imperative that *all* living systems decrease uncertainty as prediction errors ( $H_{cog}$ ). It is noteworthy that fitness is not identical to survival ability, but includes the reproduction rate as another component. A genotype with a high reproduction rate but relatively low survivability can be competitive in evolutionary time; and generally trade-offs exist between survivability and the reproduction rate (e.g., Mayr, 1982; Begon et al., 1996; Pianka, 1999). For example, even in the same ecological niche, an ancestral clone of a bacterium (*Escherichia coli*) differentiated adaptively into two distinctive types in a long-term experimental culture of the bacterium and its predator (*Tetrahymena thermophila*), one with a high reproduction rate that was vulnerable to the predator, and another with a low reproduction rate that was less vulnerable to the predator (Nakajima and Kurihara, 1994). To conclude, the reduction of uncertainty ( $H_{cog}$ ) is not necessarily the goal, but is somehow required for achieving a better internal probability distribution ( $P_{overall}$ ) of events to persist *over generations*.

### 6.3. Mutual information, meaning, and adaptation

Shannon's information theory focuses on the quantitative aspect of information in terms of the degree of reduction in uncertainty (i.e., the amount of information). However, the amount of information is usually simply called “information,” which causes much confusion in biology (Barbieri, 2013); and the role of information theory in biology is also controversial (Battail, 2013). The amount of information is described as the difference between the quantities of uncertainty or disorder (i.e., entropy,  $H$ ) before and after receiving a certain piece of information (or a message). As described in section 5.5, a subject can reduce the uncertainty about an object by experiencing events about a given sign vehicle in the locality. Let us describe this reduction in uncertainty using Shannon's information theory. Suppose that two entities ( $X$  and  $Y$ ) exist in the environment of a focal entity, in which  $X$  is an object, and  $Y$  is a sign vehicle. The event space concerning  $Y$  is represented as  $\mathbf{Y} = \{y_j\}$  and  $\sum p(y_j) = 1$ ; for example,  $\mathbf{Y} = \{y_0, y_1\}$ , where  $y_1$ : “ $Y$  is here”;  $y_0$ : “ $Y$  is not here.” Similarly, the event space concerning  $X$  is represented as  $\mathbf{X} = \{x_i\}$  and  $\sum p(x_i) = 1$ ; for example,  $\mathbf{X} = \{x_0, x_1\}$ ,

where  $x_1$ : “the entity  $X$  is near here”;  $x_0$ : “ $X$  is not near (far from) here”; where “near” means outside the locality but near it. The whole event space of the environment can be given as  $\mathbf{X} \times \mathbf{Y} = \{(x_0, y_0), (x_1, y_0), (x_0, y_1), (x_1, y_1)\}$ .

If the subject does not know about  $Y$  (sign vehicle), then the probability distribution of the object ( $X$ ) events is given as  $p(x_i)$  with entropy  $H(X)$ ;  $H(X) = -\sum_i p(x_i) \log_2 p(x_i)$ . However, if the subject experienced an event about  $Y$  (i.e., it determined whether it was  $y_0$  or  $y_1$ ), the probability distribution of the object ( $X$ ) events changes to  $p(x_i|y_j)$  with entropy  $H(X|Y)$ , where  $p(x_i|y_j)$  is the conditional probability of event  $x_i$  under the occurrence of  $y_j$ ;  $H(X|Y) = -\sum_j p(y_j) H(x_i|y_j) = -\sum_{i,j} p(y_j) p(x_i|y_j) \log_2 p(x_i|y_j)$ . The amount of information about  $X$  (object) found by knowing which event occurred about  $Y$  (sign) is given as  $I(X; Y) = H(X) - H(X|Y)$ , i.e., the difference between the two entropies before and after knowing about  $Y$ , which is called the mutual information between  $X$  and  $Y$ .

$$I(X; Y) = \sum_{i,j} p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i)p(y_j)}.$$

As an extreme case, if  $P(x_i|y_j) = 0$  or 1, implying that knowing  $Y$  completely determines whether  $x_i$  occurs or not,  $H(X|Y) = 0$ , and therefore,  $I(X; Y) = H(X)$ , which indicates that there is no uncertainty about  $X$  for the subject after knowing about  $Y$ . As another extreme in which  $X$  and  $Y$  occur independently,  $P(x_i|y_j) = P(x_i)$ ; therefore,  $H(X|Y) = H(X)$ , and mutual information  $I(X; Y) = 0$  (i.e., no information). It is evident that  $I(X; Y) = I(Y; X)$  from the above equation (Shannon and Weaver, 1949/1998; Cover and Thomas, 2006; see Bergstrom and Rosvall [2009] and Adami [2004] for a discussion in biology).

If  $I(X; Y)$  is relatively large, then it could be understood that  $Y$  means  $X$ . However, in this case,  $X$  also means  $Y$  because  $I(X; Y) = I(Y; X)$ . This thought lacks the presence of interpretation in the Peircean scheme, in which a sign vehicle ( $Y$ ) and an object ( $X$ ) are related mediated through interpretation by the subject (interpreter). For example, molecules ( $Y$ ) are closely correlated in state with the organism ( $X$ ) that released them. However, the meaning of the molecules to an organism eaten by this organism must be different to its predator. For the former, it is something to flee from, whereas it is something to approach for the latter. Therefore, understanding meaning as the correlation between the sign and object seems limited.

In the ICM/ICW model, as well as the CS model, the meaningful contents of states and their temporal changes are apparently removed, leaving only meaningless symbols and functions. How then can we incorporate meaning into this mathematical formalism? In Uexküll's function circle, the meaning of an event (or a cognition) is closely related to the relation between a subject and entities in the environment. MacKay (1969) defined the meaning of message as “its selective function on the range of the recipient's states of conditional readiness for goal-directed activity.” This aspect of the subject's properties has something to do with the selectivity of a motion function in the CS model (section 5.1). Based on these ideas, let us define the meaning of an event (including receiving a message or signal) to a subject entity as the effect of the event on the subject's action (or state change) that specifically affects the consequent subject–environment relation. Mathematically, an event can be defined as a subset of the state-space of the environment (section 5.4); similarly, a state relation between a subject and environment entities can be defined in terms of a subset of the direct product of their state-spaces (section 5.1). Therefore, the meaning of an event is something used to relate this event subset to a subject–environment relation subset. “Something” is the property of the subject entity, which determines a

specific action/movement in response to the event, affecting the consequent relationship with the environment. This property is nothing more than the *selectivity* of a motion function of a cognizer in the CS model. It may be that meaningful contents are implicitly embedded in the selectivity of the motion function in the CS model. This selectivity property is at the core of biological adaptation to the environment, which exists as properties resulting from its effect on the success of the possessors in persistence through natural selection. In this sense, the specific meaning systems embodied in living systems can be understood as their own internal models for environments, in which actions/movements are guided at molecular, cellular, and organismic levels.

This definition of meaning may also be applied to the relationship between the component entities within a living system, although molecular entities do not possess highly organized cognitive properties like brains. For example, in the signal transduction pathways within a living cell or a multicellular organism, a metabolic subsystem (A) receives a signal molecule released from another subsystem (B), which causes an effect on system A in a particular way, thus leading to a particular relationship between them. In this process, cognizing the molecule's state by contact (i.e., an event occurring to A) produces A's specific change as the effect (i.e., the meaning of receiving the molecule), by which A correlates with B's state in a specific way.

Lastly, it is noteworthy that physical entities also have selectivity in response to electromagnetic or gravitational forces according to their own material properties. However, the above definition of meaning should not be applied to these entities, because physical entities do not possess these selectivity properties because of their positive effects on the persistence of the entities.

## 7. Conclusions

It is frequently stated in the literature that living systems act to reduce the uncertainty of events to ensure survival by gathering information from their environments and using it. However, such a statement would be circular or a mere description of what we already know about biological adaptation, if information is defined using concepts about the reduction of uncertainty and/or the generation of order among entities. In the quest for a general framework of information applicable to living systems, a CS model (including the ICM/ICW model as a special case) was proposed to construct an information theory for living systems. In the CS model, knowing (or observing) and action (or movement) are unified as a single concept of a state change of entity within an observed system. Focal and environmental entities change their state at every moment according to their material properties of discriminating between different environmental states and selecting one particular succeeding state among many possibilities (section 5). Based on this process, the probabilities of the events that a subject entity experiences within the world (i.e., internal probabilities) can be derived without assuming a probability concept a priori. Based on this, the internal entropy is derived as a measure of the uncertainty or order of the environment that a subject entity experiences in interaction with the environment (section 6.1).

In this scheme, information is defined as a state change of an entity that occurs in relation to the state of another entity, i.e., a relational state change between two entities. Within a given observed system, an event (e.g., receiving a signal) occurring to a subject entity can have a particular effect on the subject's action (or state change), which will then affect the subject–environment relation. The semiosis involved in adaptation links an event to a particular action, which can lead to a favorable relation with the environment. In other words, the link between an event and

a subsequent action is the *meaning* of the event to the subject living system. The whole set of such links constitutes an internal model of the external reality, which is embodied or built in the organization of the living system. The qualitative aspect of information describes the specificity of a consequent relation by a relational change in response to the other's state, while the quantitative aspect describes the degree to which it is specified. It was suggested that the activity to achieve a better probability distribution ( $P_{overall}$ ) of events for survival and reproduction through an embodied internal model for external reality, rather than the reduction of uncertainty ( $H_{cog}$ ) is an essential biological property to characterize adaptation. This distribution may vary according to the subject living entity and the ecosystem organization that it inhabits. However, the analysis of adaptive evolution from the internal view is still in its infancy, requiring further investigations to incorporate evolutionary ecology into the theoretical scheme.

## Acknowledgements

I thank Pedro C. Marijuán, Koichiro Matsuno, and Plamen L. Shimeonov for suggesting valuable comments and criticisms for improving an early version of the paper.

## Appendix 1. Internal states of cognizer

The state space ( $\mathbf{C}$ ) of a focal cognizer  $C$  can be represented as the direct product of the position space ( $\mathbf{X}$ ) and the space of the other qualities of states ( $\mathbf{Y}$ ), i.e.,  $\mathbf{C} = \mathbf{X} \times \mathbf{Y}$ . Consider a simple case in which the internal states of a cognizer are defined in terms of the relative positions of component cognizers. If a cognizer is composed of  $n$  component cognizers, a cognizer's position ( $x \in \mathbf{X}$ ) is represented as an  $n$ -tuple of the components' positions. We can introduce a local coordinate system (LCS) such as a barycentric coordinate system to define the internal positional states, with the center of mass being the origin. Different  $n$ -tuples in this LCS imply different internal states characterizing specific patterns or structures, and the origin's position determines the positional relation to environmental cognizers. The same internal state with different positions for the origin relative to the environment is identified with an equivalence class (subset) of  $\mathbf{C}$  in the direct product of the position spaces of the component cognizers.

In general, the internal states (or patterns) of a cognizer are defined in terms of the relative states of the component cognizers, not necessarily confined to their positional relation.  $\mathbf{C} = \mathbf{C}_1 \times \mathbf{C}_2 \dots \mathbf{C}_n$ , where  $\mathbf{C}_j$  ( $1 \leq j \leq n$ ) is the state space of a component cognizer of cognizer  $C$ . Therefore, the cognizer's states fulfilling a particular state-relation among the component cognizers constitute a subset of  $\mathbf{C}_1 \times \mathbf{C}_2 \dots \mathbf{C}_n$ . The inter-component relation of cognizer  $C$  partitions the state space ( $\mathbf{C}$ ) into  $m$  equivalence classes,  $\mathbf{C}^{(a)}$ ,  $\mathbf{C}^{(b)}$ , ..., and  $\mathbf{C}^{(m)}$ , with each corresponding to an internal state of cognizer  $C$ . We can thus consider the set  $\mathbf{C}[\mathbf{in}] = \{a, b, \dots, m\}$ , which constitutes  $C$ 's internal state space.  $C$  also has external states that relate it to the environmental cognizers' states, like the positions of the origin of LCS in a global coordinate system as above, and denote the external state set as  $\mathbf{C}[\mathbf{ex}]$ . Therefore,  $\mathbf{C} = \mathbf{C}[\mathbf{in}] \times \mathbf{C}[\mathbf{ex}]$ . Different states of the same internal state ( $\in \mathbf{C}^{(k)}$ ;  $k \in \mathbf{C}[\mathbf{in}]$ ) have different states relative to the environment.

The calculation of internal probability given in section 5.4 and Fig. 9 can also be applied to a case where a cognizer's states ( $cx, cy, czi$ ) are its internal states. In this case, we need to take  $\mathbf{C}[\mathbf{in}]$  as the focal cognizer's state space, and  $\mathbf{C}[\mathbf{ex}]$  is incorporated into the environmental cognizers to convert their states relative to  $C$ 's states.

## References

- Adami, C., 2004. Information theory in molecular biology. *Phys. Life Rev.* 1, 3–22.
- Barbieri, M., 2013. The paradigms of biology. *Biosemiotics* 6, 33–59.
- Battail, G., 2013. Biology needs information theory. *Biosemiotics* 6, 77–103.
- Bateson, G., 1979. *Mind and Nature*. E. P. Dutton, New York.
- Begon, M., Harper, J.L., Townsend, C.R., 1996. *Ecology: Individuals, Populations and Communities*. Wiley-Blackwell.
- Ben-Naim, A., 2012. *Entropy and the Second Law: Interpretation and Misinterpretation*. World Scientific Pub Co Inc.
- Bergstrom, C.T., Rosvall, M., 2009. The transmission sense of information. *Biol. Philos.* 26, 159–176.
- Cassirer, E., 1923. *Substance and Function, and Einstein's Theory of Relativity*. Primary Source Edition. The Open Court Publishing Company, Chicago. Translated by W. C. Swabey and M. C. Swabey.
- Chase, J.M., Leibold, M.A., 2003. *Ecological Niches: Linking Classical and Contemporary Approaches*, second ed. Univ of Chicago Press.
- Chandler, D., 2007. *Semiotics, the Basics*, second ed. Routledge, London.
- Clark, M.P.A., Westerberg, B.D., 2009. How random is the toss of a coin? *CMAJ* 181, E306–E308. <http://dx.doi.org/10.1503/cmaj.091733>.
- Conant, R.C., Ashby, W.R., 1970. Every good regulator of a system must be a model of that system. *Int. J. Syst. Sci.* 1, 89–97.
- Cover, T.M., Thomas, J.A., 2006. *Elements of Information Theory*, second ed. Wiley-Interscience.
- Denbigh, K.G., Denbigh, J.S., 1985. *Entropy in Relation to Incomplete Knowledge*. Cambridge University Press, Cambridge.
- Dusenbery, D.B., 1992. *Sensory Ecology*. W. H. Freeman and Company, NY.
- Fersht, A.R., 1987. The hydrogen bond in molecular recognition. *Trends Biochem. Sci.* 12, 301–304.
- Friston, K., 2010. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138.
- Friston, K., 2012. A free energy principle for biological systems. *Entropy* 14, 2100–2121. <http://dx.doi.org/10.3390/e14112100>.
- Friston, K., Thornton, C., Andy Clark, A., 2012. Free-energy minimization and the dark-room problem. *Front. Psychol.* 3, 1–7. <http://dx.doi.org/10.3389/fpsyg.2012.00130>. Article 130.
- Gillies, D., 2000. *Philosophical Theories of Probability*. Routledge, New York.
- Hacking, I., 2006. *The Emergence of Probability: a Philosophical Study of Early Ideas About Probability, Induction and Statistical Inference*, second ed. Cambridge University Press.
- Hutchinson, G.E., 1957. Concluding remarks. *Cold Spring Harb. Symp. Quant. Biol.* 22, 415–427.
- Husserl, E., 1913/1983. *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, First Book, General Introduction to a Pure Phenomenology*. Springer (translated by F. Kersten).
- Husserl, E., 1931/2012. *Cartesian Meditations*. Springer.
- Husserl, E., 1936/1970. *The Crisis of European Sciences and Transcendental Philosophy*. Northwestern University Press, Evanston. Carr, D., trans.
- Hoffmeyer, J., 2007. Semiotic scaffolding of living systems. In: Barbieri, M. (Ed.), *Introduction to Biosemiotics*. Springer, Netherlands.
- Holland, J.H., 2012. *Signals and Boundaries*. The MIT Press, MA.
- Holland, J.H., 1995. *Hidden Order: How Adaptation Builds Complexity*. Addison-Wesley Publishing Company, Inc., Reading, MA.
- Hume, D., 1777/1902. *An Enquiry Concerning Human Understanding*, second ed. (Kindle Edition).
- Koshland, D.E., 1973. Protein shape and biological Control. *Sci. Am.* 229 (4), 52–64.
- Lerner, R.A., Tramontano, A., 1988. Catalytic anti-bodies. *Sci. Am.* 258, 58–70.
- Loewenstein, W.R., 1999. *The Touchstone of Life: Molecular Information, Cell Communication, and the Foundations of Life*. Oxford Univ Press.
- Mackey, D.M., 1969. *Information, Mechanism and Meaning*. The MIT Press.
- Matsuno, K., 1996. Internalist stance and the physics of information. *BioSystems* 38, 111–118.
- Matsuno, K., 1989. *Protobiology: Physical Basis of Biology*. CRC Press, FL, Boca Raton.
- Maturana, H.R., Varela, F.J., 1980. *Autopoiesis and Cognition—the Realization of the Living*. Reidel, Dordrecht.
- Mayr, E., 1982. *The Growth of Biological Thought*. Harvard University Press, Cambridge, MA.
- Maxwell, J.C., 1871/2001. *Theory of Heat*. Longmans, Green, and Co., London. Reprinted in *Theory of Heat*. Dover, Publications, Inc., New York.
- Nakajima, T., Kurihara, Y., 1994. Evolutionary changes of ecological traits of bacterial populations through predator-mediated competition 1. Experimental analysis. *Oikos* 71, 24–34.
- Nakajima, T., 1999. Biological probability: cognitive processes of generating probabilities of events in biological systems. *J. Theor. Biol.* 200, 77–95.
- Nakajima, T., 2001. Is the world deterministic?: mental algorithmic process generating external reality by inverse causality. *Int. J. Gen. Sys* 30, 681–702.
- Nakajima, T., 2004. Synchronic and diachronic hierarchies of living systems. *Int. J. General Syst.* 33, 505–526.
- Nakajima, T., 2007. How living systems manage the uncertainty of events: adaptation by semiotic cognition. In: Barbieri, M. (Ed.), *Biosemiotic Research Trends*. Nova Science Publishers, Inc., New York, pp. 157–172.
- Nakajima, T., 2013. Probability in biology: overview of a comprehensive theory of probability in living systems. *Prog. Biophys. Mol. Biol.* 113, 67–79.
- Pianka, E.R., 1999. *Evolutionary Ecology*, sixth ed. Benjamin Cummings.
- Pittendrigh, C.S., 1958. Adaptation, natural selection, and behavior. In: Roe, A., Simpson, George Gaylord (Eds.), *Behavior and Evolution*. Yale University Press, New Haven, pp. 390–416.
- Popper, K.R., 1982. *The Open Universe: an Argument for Indeterminism*. Routledge, London.
- Rosen, R., 1977. Observation and biological systems. *Bull. Math. Biol.* 39, 663–678.
- Rosen, R., 1978. *Fundamentals of Measurement and Representation of Natural Systems*. Elsevier North-Holland, Inc., New York.
- Rosen, R., 1991. *Life Itself*. Columbia University Press, New York.
- Rössler, O., 1987. Endophysics. In: Casti, J., Karlqvist, A. (Eds.), *Real Brains-artificial Minds*. North-Holland, New York, pp. 25–46.
- Rössler, O., 1993. Endophysics—Descartes taken seriously. In: Atmanspacher, H., Dallenort, G.D. (Eds.), *Inside Versus Outside*. Springer-Verlag, Berlin, pp. 153–161.
- Salmon, W., 1967. *The Foundations of Scientific Inference*. University of Pittsburgh Press.
- Salthé, S.N., 1985. *Evolving Hierarchical Systems: Their Structure and Representation*. Columbia University Press, New York.
- Salthé, S.N., 2001. Theoretical biology as an anticipatory text: the relevance of Uexküll to current issues in evolutionary systems. *Semiotica* 134, 1–22.
- Salthé, S.N., 2014. Creating the Umwelt: from chance to choice. *Biosemiotics* 7, 351–359.
- Saussure, F., 1916/2013. *Course in General Linguistics*. Bloomsbury Academic, London.
- Shannon, C.E., Weaver, W., 1949/1998. *The Mathematical Theory of Communication*. University of Illinois Press.
- Sjolander, S., 1997. On the evolution of reality—some biological prerequisites and evolutionary stages. *J. Theor. Biol.* 187, 595–600.
- Stewart, R.C., Dahlquist, F.W., 1987. Molecular components of bacterial chemotaxis. *Chem. Rev.* 87, 997–1025.
- Uexküll, J. von, 1926. *Theoretical Biology*. Harcourt, Brace, New York.
- Uexküll, J. von, 2010. *A Foray into the Worlds of Animals and Humans: with a Theory of Meaning*. Univ of Minnesota Press.
- Varela, F., Thompson, E., Rosch, E., 1992. *The Embodied Mind*. The MIT Press, Cambridge, MA.
- Wolfenden, R., 1972. Analog approaches to the structure of the transition state in enzyme reactions. *Acc. Chem. Res.* 5, 10–18.
- Wheeler, J.A., 1989. Information, physics, quantum: the search for links. In: Proc. 3rd Int. Symp. Foundations of Quantum Mechanics, Tokyo, pp. 354–368. Reprinted In: Feynman and Computation, edited by A. J. G. Hey (1999), Perseus Books, Cambridge, pp. 309–336.
- Young, P., 1987. *The Nature of Information*. Praeger.