組織化された系の複雑さを測る方法

岡本龍明
NTT 情報通信処理研究所

組織化された系の複雑さを測る一方法を提案する。そのため、情報源に対し、システムコンプレキシティと呼ぶ概念を定義する。また、条件付きシステムコンプレキシティ、及び資源制約システムコンプレキシティを導入する。特に、多項式時間制約条件付きシステムコンプレキシティが、Goldwasserらによって導入された知識コンプレキシティと密接に関連することを示す。さらに、情報源の持つ情報量のみでなく受信者の能力や知識にも依存して決まるような転送情報量を測る一方法を示す。また、複雑度を決定する不完全性並びに実用的測定方法について述べる。最後に、システムコンプレキシティの応用として、計算量理論、通信、及び、生物学的オートマトン理論への応用例を示す。

How to Measure the Complexity of Organized Systems

(Preliminary Report)

Tatsuaki Okamoto

NTT Laboratories
Nippon Telegragh and Telephone Corporation
1-2356, Take, Yokosuka-shi, Kanagawa-ken, 238-03 Japan

May 1988

Abstract

The main purpose of this paper is to present a new way to measure the degree of the complexity of organized systems. For this purpose, we define a new notion of complexity, called *system complexity*. Conditional system complexity and a resource bounded variant of system complexity are also presented. In particular, we show that the polynomial-time bounded conditional system complexity is closely related to the knowledge complexity introduced by Goldwasser et al. In addition, we show a way to measure the amount of transmitted information from sources to destinations and inference processes in the light of system complexity. Incompleteness and practical measurement are also discussed. Finally, a number of applications to computational complexity, communications, and biological automata theory are presented.

1. Introduction

In 1948, Weaver classified scientific problems into three classes: problems of *simplicity*, problems of *disorganized complexity*, and problems of *organized complexity*[W]. The classical dynamics can be used to analyze and predict the motion of a few ivory balls as they move about on a billiard table. This is a typical problem of simplicity. Imagine, then, a large billiard table with millions of balls rolling over its surface, colliding with one another and with the side rails. Although to be sure the detailed history of one specific ball can not be traced, statistical mechanics can analyze and predict the average motions. This is a typical problem of disorganized complexity. Problems of organized complexity, however, deal with features of organizations such as living things and artificial machines. Here, cells in a living thing are interrelated into an organic whole, whereas the balls in the illustration of disorganized complexity problems are distributed, in their positions and motions, in a helter-skelter manner.

The degree of complexity is an essential notion in problems of *disorganized complexity*, and is measured by entropy as defined in statistical mechanics. In a similar manner, disorganized complexities of sources and strings are measured by Shannon entropy[SW] and Kolmogorov complexity[K, C1]. Although these three approaches are obviously different, they have been shown to share a number of similar properties[SW, C2].

In problems of *organized complexity*, the degree of complexity is also an essential notion. Imagine a situation in which we are receiving a signal from deep space. If the signal is perfectly random (e.g., a random noise radio wave) or perfectly regular (e.g., a regular radio wave pulse from a pulsar), then we do not suppose it is being sent by an intelligent creature. On the contrary, if we notice that the signal is constructed according to a complicated rule (e.g., the radio wave pulse represents a prime numbers sequence), then we would know that intelligent creatures exist in deep space. In this situation, the degree of organized complexity of the signal plays an essential role in deciding whether the signal is sent by intelligent creatures or not.

How, then, can we measure the degree of *organized complexity*? The following two examples show that the notions of entropy and Kolmogorov complexity cannot be used.

First, if a supercomputer is broken up into random pieces, then the degree of organized complexity of the randomly broken supercomputer is at the minimum level, while its entropy is at the maximum level. Second, a random alphabet sequence which a chimpanzee has typed has a considerably higher value of entropy or Kolmogorov complexity than the plays of Shakespeare. Obviously, the degree of organized complexity of the former sequence should be much less than that of the latter.

On the other hand, the notion of negative entropy (*negentropy* [Sc]) cannot be a measure of organized complexity either, because the perfectly regular thing, or the most simply organized thing, has the maximum negentropy value.

Thus, a new notion for measuring the degree of organized complexity is needed. The measure of organized complexity should meet at least the following conditions.

(1) Perfectly random things have the *minimum* organized complexity value. (They have the *maximum* entropy or Kolmogorov complexity value.)

(2) Perfectly regular things have the *minimum* organized complexity value. (They have the *maximum* negentropy value.)

Then, what is the common property of perfectly random and perfectly regular things in light of organized complexity? Also, what is the common property of things with a high organized complexity value? We can describe the characteristics both of perfectly random and perfectly regular things in the simplest manner. On the contrary, we must create a complicated description to explain the characteristics of a supercomputer or the plays of Shakespeare. This fact leads to a thought regarding the measure of organized complexity: Something like the quantity of description required to characterize an organized system can be the measure of its organized complexity.

In this paper, we present a new way to measure the complexity of organized systems on the basis of the above concept. To do this, a new notion of complexity of organized *sources* is defined, because the complexity of many organized systems can be approximately represented as the complexity of the outputs, or *sources*, from these systems (Remember the above-mentioned example of signals from deep space, in which we estimate the complexity of a system through observing a signal from the system). The complexity is also the measure of the quantity of information that a source contains in relation to some kind of semantics (Remember the above-mentioned example of the plays of Shakespeare and the sequence typed by a chimpanzee).

The new definition has its origin in Kolmogorov complexity. Roughly speaking, this new complexity, which we will refer to as *system complexity*, is defined based on *probabilistic* Turing machines, whereas Kolmogorov complexity is based on *deterministic* Turing machines. For example, the system complexity of a perfectly random source is the lowest because the probabilistic Turing machine can generate a perfectly random source by the shortest program. Furthermore, the system complexity of a perfectly regular source is also the lowest because the probabilistic Turing machine encompasses the capability of the deterministic Turing machine, and the deterministic Turing machine can generate the perfectly regular source by the shortest program. Then, Kolmogorov complexity corresponds to system complexity, when a source is deterministic.

In section 2, the system complexity and conditional system complexity are defined. System complexity in this section is defined under the conditions of a *resource unbounded* Turing machine. For Kolmogorov complexity and Shannon entropy, some resource bounded variants have been proposed [H, Si, BB, Y]. In section 3, we present a *resource bounded* variant of system complexity. In particular, the *polynomial-time bounded conditional* system complexity is shown to be closely related to the *knowledge complexity* defined by Goldwasser et al.[GMR]. In section 4, we show a way to measure the amount of information transmitted from a source to a destination by means of the conditional system complexity. This amount depends not only on the source's properties but also on the *capability* and *knowledge* which are held at the destination. For this definition of the information amount, the notion of *inference process* and the *capability bounded* system complexity are also introduced. In section 5, we offer some discussions regarding incompleteness and practical measurement. Finally, in section 6, we show system complexity can be applied to many areas including computational complexity, communications, and biological automata theory.

2. System complexity of sources

Definition 1. $\Sigma = \{0,1\}$ and $\Sigma^* = \{\Lambda, 0, 1, 00, 01 \ldots\}$ is a set of finite binary strings. Λ is the empty string. $l(x)$ is the length of a string $x \in \Sigma^*$. $\Sigma^n = \{x \mid l(x) = n, x \in \Sigma^*\}$. N is the set of natural numbers.

Definition 2. A *Probabilistic Turing machine* (PTM) is a Turing machine with a read-only input tape, a work tape, a write-only output tape, and a read-only random tape. The random tape contains an infinite sequence of random bits. The random tape can be scanned only from left to right. When we say that the PTM flips a coin, we mean that it reads the next bit of its own random tape.

Definition 3. A *source* X_n is a probability distribution P_{X_n} over Σ^n, where $\sum_{x\in\Sigma^n} P_{X_n}(x) = 1$. A source set $\mathcal{X}$ is a set of sources $\{X_n \mid n \in N\}$.

Definition 4. The partial functions $\phi_i : \Sigma^* \times \Sigma^* \to \Sigma^*$ and $M_i : \Sigma^* \to \mathcal{X}$ computed by a PTM M_i is defined by

$$\phi_i(p; \alpha) = x,$$

$$P_{M_i(p)}(x) = \sum_{\alpha} \Pr\{\phi_i(p; \alpha) = x\},$$

where p is an input, and α is the string of coin flips it takes the PTM to compute x from p. If no such x exists, $\phi_i(p; \alpha)$ is not defined. Note that this definition differs from that in Gill[Gi].

The next theorem implies that i is an acceptable Gödel number of the partial recursive functions.

Theorem 1. There exists a *universal* PTM M_u such that $\phi_u(< i, p >; \alpha) = \phi_i(p; \alpha)$ for every binary string α. In particular, $M_u(< i, p >) = M_i(p)$ for every i and p. Here, we assume a standard computable pairing function $< i, p >$.

The proof of this theorem uses simulations essentially identical to that used for deterministic Turing machines. Then, we define the system complexity of a source $X_n \in \mathcal{X}$ with respect to the partial function M_i computed by a PTM M_i as the length of the shortest program which generates the source.

Definition 5. Let $M_i : \Sigma^* \to \mathcal{X}$ be a partial function, and $X_n \in \mathcal{X}$.

$$S_i(X_n) = \min\{l(p) \mid M_i(p) = X_n\}$$

If no such p exists, then $S_i(X_n) = \infty$.

This complexity measure depends in an essential way on the function M_i. We almost get rid of this dependence by means of the following theorem, in the same manner as for Kolmogorov complexity.

Theorem 2. There exists a universal function M_u such that for any function M_i

$$S_u(X_n) \leq S_i(X_n) + c,$$

where $X_n \in \mathcal{X}$, and c is a constant (dependent on M_u and M_i but not on X_n).

In an analogous way we can introduce the concept of *conditional* system complexity.

Definition 6. Let $M_i : \Sigma^* \times \Sigma^* \rightarrow \mathcal{X}$ be a partial function.

$$S_i(X_n \mid t) = \min\{l(p) \mid M_i(p,t) = X_n\},$$

where $X_n \in \mathcal{X}$ and $p, t \in \Sigma^*$. If no such p exists, then $S_i(X_n \mid t) = \infty$.

Theorem 3. There exists a universal function M_u such that for any function M_i

$$S_u(X_n \mid t) \leq S_i(X_n \mid t) + c,$$

where $X_n \in \mathcal{X}$, $t \in \Sigma^*$, and c is a constant.

Definition 7. The *system complexity of a source* $X_n \in \mathcal{X}$ is defined by $S_u(X_n)$ for a universal PTM M_u, and is simply written by $S(X_n)$. The *conditional system complexity of a source* $X_n \in \mathcal{X}$ with respect to a *string* $t \in \Sigma^*$ is defined by $S_u(X_n \mid t)$ for a universal PTM M_u, and is simply written by $S(X_n \mid t)$.

Example 1. Let source $X_n \in \mathcal{X}$ such that $P_{X_n}(x) = 2^{-n}$, where $x \in \Sigma^n$ (that is, X_n is a perfectly random source). Then,

$$S(X_n \mid n) = O(1),$$

where $O(1)$ denotes a function whose absolute value is less than or equal to a constant.

Note that the Shannon entropy of this source is n.

Example 2. Let source $X_n \in \mathcal{X}$ such that $P_{X_n}(x) = 1$ if $x = 1^n \in \Sigma^n$ and 0 otherwise (that is, X_n is a perfectly regular source). Then,

$$S(X_n \mid n) = O(1).$$

Note that the Shannon entropy of this source is 0.

3. Resource bounded variant of system complexity

In this section, we show only a time bounded variant of the system complexity. The space bounded variant of the system complexity can be shown in a similar manner. First, we define a variant of the indistinguishability introduced by Yao[Y] and Goldwasser et al.[GM].

Definition 8. Let $X_n \in \mathcal{X}$, $Y_n \in \mathcal{X}$ be two sources. Let a *distinguisher* D be a PTM that on input strings from a source outputs either 0 or 1 and halts within $f(n)$ steps. X_n and Y_n are at most g-$f(n)$-*distinguishable*, if for any D

$$\mid \Pr\{D(X_n) = 1\} - \Pr\{D(Y_n) = 1\} \mid < g(n) + 1/f(n),$$

with a sufficiently large n. We say that X_n and Y_n are $f(n)$-*indistinguishable*, when they are at most 0-$f(n)$-*distingishable* .

Definition 9. Let $M_i : \Sigma^* \rightarrow \mathcal{X}$ be a partial function, and $X_n \in \mathcal{X}$.

$S_i^f(X_n) = \min\{l(p) \mid M_i(p)$ and X_n are $f(n)$-indistinguishable; $M_i(p)$ halts within $f(n)$ steps with large enough n $\}$.

If no such p exists, then $S_i^f(X_n) = \infty$.

Theorem 4. There exists a universal function M_u such that for any function M_i

$$S_u^{(f)}(X_n) \leq S_i^f(X_n) + c,$$

where c is a constant, and

$S_u^{(f)}(X_n) = \min\{l(p) \mid M_u(p)$ and X_n are $f(n)$-indistinguishable; $M_u(p)$ halts within $(cf(n)\log f(n) + c)$ steps with large enough n $\}$.

The proof of this theorem is obtained by the Hennie-Stearns simulation theorem [HS], in a manner similar to that shown in [H].

Definition 10. The *$f(n)$-time bounded system complexity of a source $X_n \in \mathcal{X}$* is defined by $S_u^{(f)}(X_n)$ and is simply written by $S^{(f)}(X_n)$.

The $f(n)$-time bounded *conditional* system complexity of a source $X_n \in \mathcal{X}$ with respect to a string $x \in \Sigma^*$ is similarly defined by $S_u^{(f)}(X_n \mid x)$, simply written as $S^{(f)}(X_n \mid x)$. $S^{(f)}(X_n \mid x)$ represents the amounts of information which the sequences of source X_n contain when the destination includes the knowledge x. Here, the resource is limited to within $f(n)$. This notion is shown to be closely related to the knowledge complexity introduced by Goldwasser et al.[GMR]. First, we define a non-interactive variant of the knowledge complexity.

Defintion 11. Let X_n be a source and x be a string of length n. We say that (X_n, x) has a *non-interactive knowledge complexity* no greater than $r(n)$, if there exist a polynomial time PTM A such that $A(x)$ and X_n are at most *g-$poly(n)$*-distinguishable, where $r(n) = \lfloor -\log(1-g(n)) \rfloor$, and *poly* denotes any polynomial function. We denote this fact by $(X_n, x) \in NKC(r(n))$.

Theorem 5. Let $X_n \in \mathcal{X}$ and $x \in \Sigma^n$. If $(X_n, x) \in NKC(r(n))$, then

$$S^{(poly)}(X_n \mid x) \leq r(n) + O(1).$$

(An informal proof sketch)

Let $\alpha(A)$ be a coin flips subset of $A(x)$, and $\alpha_D(A, X_n)$ be a coin flips subset of $A(x)$ that corresponds to the same output from each D as X_n gives. $\#\{Z\}$ denotes the number of elements in a set Z. Let $A^*(x)$ be $A(x)$ which minimizes the value of $g(n)$ and the minimum value be $g^*(n)$, where X_n and $A(x)$ are at most *g-$poly(n)$-distinguishable*. Then, for any distinguisher D, the probability that $A^*(x)$ and X_n give the same output from D is at least $(1-g^*(n))$, because $\Pr\{D(X_n) = 1\}$ is asymptotical to 1 or 0. Then, we will show that a subset whose elements number ratio to $\#\{\alpha(A^*)\}$ is at least $(1 - g^*(n))$ is common among $\alpha_D(A^*, X_n)$, for all D's. If we can show this, we can represent the common subset with information whose size is at most $\lfloor -\log(1-g^*(n)) \rfloor$, and we can generate X_n by the coalition of the most powerful D and $A^*(x)$, and using at most $\lfloor -\log(1-g^*(n)) \rfloor$ bit information. Then, the program lengths of D and A^* are bounded.

$\alpha_D(A^*, X_n)$ differs for each D. If there are subsets $\alpha_{D_1}(A^*, X_n)$ for D_1 and $\alpha_{D_2}(A^*, X_n)$ for D_2 such that the ratio of $\#\{\alpha_{D_1}(A^*, X_n) \cap \alpha_{D_2}(A^*, X_n)\}$ to $\#\{\alpha(A^*)\}$ is less than $(1-g^*(n))$, then the coalition of D_1 and D_2 can distinguish A^* from X_n with a probability no less than $g^*(n)$. This is a contradiction. Thus a subset whose elements number ratio to $\#\{\alpha(A^*)\}$ is at least $(1 - g^*(n))$ is common among $\alpha_D(A^*, X_n)$, for all D's.

4. Transmitted information and inference processes

In this section, we show the amount of information transmitted from a source to a destination with respect to inference processes. First, we define the *capability bounded* system complexity, the inference process and the transmitted information amount. Then, we show the results regarding the relationship between the transmitted information amount and the capability bounded conditional system complexity. First, we define another variant of the indistinguishability.

Definition 12. Let $X_n \in \mathcal{X}$, $Y_n \in \mathcal{X}$ be two sources. Let $\Omega \subset \{M \mid M \text{ is a PTM}\}$, and $D_\Omega \in \Omega$ be a *distinguisher* that on input strings from a source outputs either 0 or 1. X_n and Y_n are *Ω-$f(n)$-indistinguishable*, if for any $D_\Omega \in \Omega$

$$\mid \Pr\{D_\Omega(X_n) = 1\} - \Pr\{D_\Omega(Y_n) = 1\} \mid < 1/f(n),$$

with a sufficiently large n.

In a manner similar to that shown in Definition.9, 10, Theorem.4 and [H], we can define the *capability bounded* system complexity, $S_i^{\Omega,f}(X_n)$, and $S_u^{(\Omega,f)}(X_n)$, simply written as $S^{(\Omega,f)}(X_n)$. Its *conditional* version, $S^{(\Omega,f)}(X_n \mid t)$, can be also defined.

Definition 13. Suppose that the strings x_i ($i = 1, 2, \ldots; x_i \in \Sigma^n$) generated by a source X_n is transmitted to a destination Y. (Ω, f) is the *capability* of Y, and a string t_i is the i-th *knowledge* which is held at Y after receiving the strings $\{x_1, \ldots x_i\}$. *Inference process* $Ip(X_n, \Omega, f, t)$ is a set of strings$\{t_i \mid i = 0, 1, 2, \ldots\}$ such that

$$S^{(\Omega,f)}(X_n \mid t_i) < S^{(\Omega,f)}(X_n \mid t_{i-1}) + O(1),$$

$$\lim_{i \to \infty} S^{(\Omega,f)}(X_n \mid t_i) = O(1),$$

$$t_0 = t.$$

Definition 14. In the inference process $Ip(X_n, \Omega, f, t)$, the *transmitted information amounts* T_i (by sending the string x_i) and T (by sending the strings $x_1, x_2, \ldots$), are defined as

$$T_i = S^{(\Omega,f)}(X_n \mid t_{i-1}) - S^{(\Omega,f)}(X_n \mid t_i),$$

$$T = \sum_{i=1}^{\infty} T_i.$$

Definition 15. A PTM is a *prefix-free* PTM, if programs on the input tape satisfy the extention of the Kraft inequality condition[C2]. *Prefix-free* system complexity is defined on a *prefix-free universal* PTM.

The following theorem shows the relationship between the transmitted information amount and the prefix-free conditional system complexity.

Theorem 6. Let T be the transmitted information amounts in the inference process $Ip(X_n, \Omega, f, t)$. Then,

$$T = \widetilde{S}^{(\Omega,f)}(X_n \mid t) + O(1),$$

where $\widetilde{S}^{(\Omega,f)}(X_n \mid t)$ denote the prefix-free variant of $S^{(\Omega,f)}(X_n \mid t)$.

Thus, $\widetilde{S}^{(\Omega,f)}(X_n \mid t)$ means the amount of information which is incorporated in any sequence generated by a source X_n when the destination has the capability (Ω, f) and the knowledge t.

5. Incompleteness and practical measurement

Chaitin [C3] has shown that it is impossible to prove that a particular string x is of a Kolmogorov complexity greater than $l(x) + c$, where c is a constant. From a practical viewpoint, however, we can measure approximately the complexity (randomness) of a finite string by measuring the length of a universal coding of the string[ZL, R].

In a similar manner, we will be able to show that it is impossible to prove that a particular source X_n is of a system complexity greater than $f(n) + c$, where f is a function. From a practical viewpoint, however, we can roughly measure the system complexity of a source through strings generated by the source, when the structure of the source is relatively simple and can be approximately presumed. Then, its system complexity can be measured by measuring the length of the source parameters which are determined by criteria such as the Minimum Description Length Criteria[R] or the Akaike Information Criteria[A].

6. Applications

This section suggests a few representative applications of this notion of system complexity to the areas of computational complexity, communications, and biological automata theory.

6.1 Computational complexity

In computational complexity theory, it has been shown that the expected running times of instances randomly selected from among several NP-hard problems can be polynomial under certain conditions[AV, DF]. By contrast, the expected running times of instances maliciously selected from among these NP-hard problems can be much longer than those of the randomly selected instances. A set of instances selected from a problem with a distribution is considered to be a source. Therefore, the system complexity of these probabilistic problems can be defined, and the system complexity of the randomly selected problem is the lowest. In contrast, the system complexity of the maliciously selected problem is much higher than that of the randomly selected problem. Consequently, studying the relationship between the expected running time of a probabilistic problem and its system complexity is expected to yield many interesting and practical insights.

6.2 Communications

Recent advances in the communications and computations technology have led to the proliferation of multimedia communications. There are many kinds of media: voice (e.g., telephone), character (e.g., data communication), and graphic (e.g., facsimile), among others. When we make a media-conversion in multi-media communications (e.g., voice to character), can we measure the amount of information lost through the conversion from one medium to another? In this section, we show a way to measure the information amount lost through the conversions. First, we define the media.

Definition 16. A partial function $\Phi : \Sigma^* \to \mathcal{X}$ is a *medium*. We say that a source $X \in \mathcal{X}$ is *on a medium* Φ, if there is a program $p \in \Sigma^*$ such that $X = \Phi(p)$. p_X is a *minimum* program of a source X on Φ, if

$$\Phi(p_X) = X \text{ and } l(p_X) = \min\{l(p) \mid \Phi(p) = X\}.$$

$\widehat{p}_X$ is *transmitted information* of a source X on Φ, if

$$p_X = <\widehat{p}_X, q> \text{ and } l(q) = O(1) \text{ and}$$

$$l(\widehat{p}_X) = \min\{l(\widehat{p}) \mid p_X = <\widehat{p}, q> \text{ and } l(q) = O(1)\}.$$

Definition 17. Let $X \in \mathcal{X}$ and $Y \in \mathcal{X}$ be sources on media Φ and Ψ, respectively. X and Y are *equivalent*, if $\widehat{p}_X$ and $\widehat{p}_Y$ are equivalent, where $\widehat{p}_X$ and $\widehat{p}_Y$ are transmitted information of sources X and Y, respectively.

Definition 18. Let Φ and Ψ be media, and a subset $\mathcal{X}_\Phi \subset \{X \mid X \text{ is a source on } \Phi\}$. A function $\mathcal{C}_{\Phi,\Psi} : \mathcal{X}_\Phi \to \mathcal{X}_\Psi$ is a *two-media conversion*. A two-media conversion $\mathcal{C}_{\Phi,\Psi}$ is *complete*, if $\mathcal{C}_{\Phi,\Psi}$ is bijective, and $X \in \mathcal{X}_\Phi$ and $Y = \mathcal{C}_{\Phi,\Psi}(X)$ are equivalent for any $X \in \mathcal{X}_\Phi$.

Theorem 7. If $\mathcal{C}_{\Phi,\Psi}$ is a complete two-media conversion, then, for any $X \in \mathcal{X}_\Phi$ and $Y = \mathcal{C}_{\Phi,\Psi}(X)$,

$$S(X) = S(Y) + O(1).$$

We can say that two-media conversion from Φ to Ψ loses no information, when the condition of Theorem 7 is satisfied. On the other hand, we can measure the lost information amount by $\max\{S(X) - S(Y) \mid X \in \mathcal{X}_\Phi \text{ and } Y = \mathcal{C}_{\Phi,\Psi}(X)\}$.

6.3 Biological automata theory

A more challenging application to biological automata theory is discussed. Von Neumann has asked basic conceptual questions regarding biology such as: "How is self-reproduction possible?", "What is an organism?", "What is its degree of organization?", "How probable is evolution?" [N ,C4]. He answered the first question in [N]. The notion of system complexity can apparently be used to answer the third question. In addition, we show one of the directions which might be taken to answer the fourth question. This is represented as a challenging open problem: is there a *principle of increasing system complexity* in some formally defined biological automata, similar to the *principle of increasing entropy* in thermodynamics.

Here, in considering this problem, we treat the biological automata as a system which generates a source X_n.

Definition 19. A PTM B is a *system with increasing knowledge*, if B satisfies the following conditions.

(1) The information amount written on the B's input tape, B's knowledge, increases as time passes. Moreover, information once written on the input tape is never erased. B_t is the state of the system B at the time t. $B_{t_1}(B_{t_2})$ is the state of B at the time t_2, when the state of B at the time t_1 is B_{t_1}, where $t_2 > t_1$. p_t is the program on the input tape of B_t, and $p_{t_2}(p_{t_1})$ is that of $B_{t_2}(B_{t_1})$. $l(p_{t_2}(p_{t_1})) = f(t_2 - t_1)$, where $f(\cdot)$ is increasing. $\mathcal{P}_{t_2}(p_{t_1})$ is a set $\{p_{t_2}(p_{t_1})\}$. Given t_1, t_2, and p_{t_1}, for any program $p \in \mathcal{P}_{t_2}(p_{t_1})$, $\Pr\{p_{t_2}(p_{t_1}) = p\} = 1/\#\{\mathcal{P}_{t_2}(p_{t_1})\}$.

(2) B_t generates a source $X_{(t)}$.

(3) The *complexity* of B_t, $S(B_t)$, is defined by $S(X_{(t)})$.

In the above-mentioned definition, the fact of increasing knowledge is essential. On the other hand, we measure the complexity of a system by means of system complexity of the source generated by the system. This method is based on the idea that the complexity of the system can be measured by measuring that of the system output.

Informal examples of systems with increasing knowledge.

(1) *Living things:* The genes of living things correspond to input tapes. Principally, information written on a gene which affects the representation of the living thing increases. The behavior and actions of living things correspond to the sources generated by these systems.

(2) *Social systems (e.g., society, company):* The important document files, books and data bases possesed by social systems correspond to input tapes. Principally, the accumulated information on them for each social system increase as long as these systems remain active. Their products and behavior correspond to the sources generated by these systems.

Proposition 1. Let B be a system with increasing knowledge, and $t_2 > t_1$.

$$S(B_{t_1}) \le E(S(B_{t_2}(B_{t_1}))) + O(1),$$

where $E(\cdot)$ denotes the expected value.

Thus, this proposition is considered to be a kind of general evolutionary principle. Remember that living things were born from the sea of chaos, or the minimum system complexity state, and evolved in degree of organization, or in the level of system complexity.

7. Conclusion

A new notion of complexity, called system complexity, has been presented to measure the degree of the organized complexity of a source. We have also shown the time bounded system complexity and the relationship with the knowledge complexity. In addition, we have shown a way to measure the amount of information transmitted from a source to a destination by means of conditional system complexity. Incompleteness and practical measurement have also been discussed. Finally, applications to computational complexity, communications, and biological automata theory have been presented.

Acknowledgements

The author would like to thank Kenji Naemura, Sadami Kurihara, Kenji Koyama and Kazuo Ohta for their valuable comments on this paper. He would also like to thank Tetsuro Kamae for his valuable suggestions regarding Kolmogorov complexity.

References

[A] H. Akaike: A New Look at the Statistical Model Identification, *IEEE Trans. Automat. Contr., AC-19* (1974), 716-723.

[AV] D. Angluin and L.G. Valiant: Fast Probabilistic Algorithms for Hamiltonian Circuits and Matching, *Proc. of 9th STOC* (1977), 30-41.

[BB] J.L. Balcázar and R.V. Book: Sets with Small Generalized Kolmogorov Complexity, *Acta Informatica, 23* (1986), 679-688.

[C1] G. J. Chaitin: On the Length of Programs for computing finite binary sequences, *J. ACM, 13* (1966), 547-569.

[C2] G. J. Chaitin: A Theory of Program Size Formally Identical to Information Theory, *J. ACM, 22* (1975), 329-340.

[C3] G. J. Chaitin: Information Theoretic Limitations of Formal Systems, *J. ACM, 21* (1974), 403-424.

[C4] G. J. Chaitin: Algorithmic Information Theory, *IBM J. Res. and Develop.* (July 1977), 350-359.

[DF] M.E. Dyer and A.M. Frieze: Fast Solution of Some Random NP-Hard Problems, *Proc. of 27th FOCS* (1986), 331-336.

[Gi] J. Gill: Computational Complexity of Probabilistic Turing Machines, *SIAM J. Comput., 6, 4* (1977), 675-695.

[GM] S. Goldwasser, S. Micali: Probabilistic Encryption, *JCSS, 28, 2* (1984), 270-299.

[GMR] S. Goldwasser, S. Micali, and C. Rackoff: The Knowledge Complexity of Interactive Proof-Systems, *Proc. of 17th STOC* (1985), 291-304.

[H] J. Hartmanis: Generalized Kolmogorov Complexity and the Structure of Feasible Computations, *Proc. of 24th FOCS* (1983), 439-445.

[HS] F.C. Hennie, and R.E. Stearns: Two-Tape Simulation of Multitape Turing Machines, *J. ACM, 13, 4* (1966), 533-546.

[K] A.N. Kolmogorov: Three Approaches to the Quantitative Definitions of Information, *Prob. Info. Transmission, 1, 1* (1965), 1-7.

[M] 森毅編：数学近未来，倍風館，1986

[N] J. von Neumann: *Theory of Self-reproducing Automata.* University of Illinois Press, Urbana, Ill, 1966.

[R] J. Rissanen: A Universal Prior for Integers and Estimation by Minimum Description Length, *Annals of Statistics, 11, 2* (1983), 416-431.

[Sc] E. Schrödinger: *What is Life?* Cambridge Univ. Press, England, 1944.

[Si] M. Sipser: A Complexity Theoretic Approach to Randomness, *Proc. of 15th STOC* (1983), 330-335.

[SW] C.E. Shannon and W. Weaver: *The Mathematical Theory of Communication.* University of Illinois Press, Urbana, Ill, 1949.

[W] W. Weaver: Science and Complexity, *American Scientist, 36* (1948), 536-544.

[Y] A.C. Yao: Theory and Applications of Trapdoor Functions, *Proc. of 23rd FOCS* (1982), 80-91.

[ZL] J. Ziv and A. Lempel: Compression of Individual Sequences via Variable-rate Coding, *IEEE Trans. Inf. Theory, IT-24, 5* (1978), 530-536.