Proc. IMEKO-TC7 Symposium (Cracow, Poland, June 25-27, 2002)

Modified concept of indiscernibility and a formal basis of classification

Hiroshi WATANABE

Higashi-mukohjima 6–19–10, Sumida-ku, Tokyo, 131–0032, Japan

hwatanab@hkg.odn.ne.jp

ABSTRACT

Establishing the notion of equivalence, or indiscernibility, with the concept of attribute, precedes every measurement practice. The present paper proposes a formal theory of classification, and sheds a new light upon the notion of equivalence underlying nominal level of measurement. Assuming that each object is characterized by a set of predicates as a binary vector, whose components are 1 or 0 according as the object affirms a predicate or not, we first examine the relationship among objects in reference to the choice of predicates, or classifier, and formulate a formal theory of classification of objects by predicates. The formalization is perfectly symmetrical between the objects and the predicates, and allows us to consider a classification of predicates by objects. Based on this reciprocal view of classification we introduce a modified concept of equality between two groups of objects, propose a new approach to clustering, and discuss its merits by showing example.

Keywords: Classification, Clustering, Measurement science, Nominal scale, Pattern recognition.

1. INTRODUCTION

In his 1887 paper [13] Hermann von Helmholtz considered the basis of the procedure of measurement as means to acquire objective knowledge of the external world, and asked two questions. First, on what ground we can represent various relationships among real objects by numerals, and second, under what condition it is possible? His conclusion is summerized as: in order to deal with empirical objects and relations quantitatively we should know first the method of comparison to determine ⟫die Gleichheit⟪, or the equivalence, of the attribute under consideration, and secondly, should have either of the operation of physical addition formally similar to the operation of numerical addition or the natural law which has the quantity of numerical addition as its coefficient. Taking over this second respect, from the end of the nineteenth century to the first half of the twentieth century, the notion of extensive measurement was established. This course of development was also a process in which the idea that only physical variables are measureble and have their meaning prevailed.

The measurement scales are for physical, non-human objects by nature, while nowadays there are increasing many occasions where we need more than numbers to express what we want, feel and judge. One possible way to meet this new situation is to recapiturate the role of the nominal scale, and to take advantage of classification of objects by some predicates, or attributes, as an additional, more basic approach to our numerical scaling practice.

Helmholtz began his consideration with the assumption that the methods of comparison are known, and he himself did not explicitly discuss a problem of how to establish the concept of the equivalence, nor how to determine ⟫die Methode der Vergleichung⟪, yet he had noticed the importance of forming the concept of a *class*. As Stevens pointed out [12], forming a class of similar objects, or events, is the first step of measurement, while the nature of classification procedure rarely had a theoretical consideration so far in the context of measurement theory on the grounds that it concerns the qualitative aspects of objects rather than their quantitative aspects, and there have been arguments that classification procedure should not be regarded as measurement [1][2]. Determing the equivalence of objects, first of all, is the function of the nominal scale, yet the question as to whether classification is really measurement in the nominal scale is not completely resolved [6].

Recently, however, in an effort to establish a new framework of measurement science, nominal scaling underlying classification procedure came to be reconsidered, and be received a new treatment as basis of testing, diagnosis, identification and pattern recognition [3]. As early as 1976, Finkelstein analysed the concepts of pattern and pattern recognition from the point of view of measurement theory, and pointed out that the procedure of pattern recognition is closely analogous to that of nominal scale measurement [5]. He also proposed regarding the description of objects by using non-numeral symbols as an extension of the concept of measurement [4][6]. As for Stevens, it is interesting to note, from our point of view, that in his 1956 paper he had regarded emerging information theory as an example which upholds the importance of the nominal scale: a transmission of a sequence of symbols is nothing but 'the treatment of data at the nominal level of measurement' [12]

Establishing the notion of equivalence, or indiscernibility, with the concept of attribute, precedes every measurement practice, and extending the notion of measurement entails reexamination of the concept of attribute, physical or psychological. It requires new view of the notion of equivalence. The present paper proposes a formal theory of classification, and, through which, sheds a new light upon the notion of indiscernibility underlying the nominal level of measurement. In the following consideration, assuming that each object is characterized by a set of predicates as a binary vector, whose components are 1 or 0 according as the object affirms a predicate or not, we examine the relationship among objects in reference to the choice of predicates, or classifier, and show some essential features of classification procedure, regarding it as a form of nominal measurement. Then we introduce a modified concept of indiscernibility with the help of the reciprocal relationship between the objects and the predicates, and discuss a formal basis of the equivalence among groups of objects.

Recently, Mari examined carefully the representational view of measurement [7], and proposed a new operational, or functional, view of measurement [8][9][10]. He discernes two components of measuring system, called *acquisition component* and *presentation component*, and points out that the former involves a function of *classifier* [9]. Such functional components play an important role in the actual measurement practice, while they have no real role to play in constructing idealized homomorphism [10]. Thus Mari's theory explaines cleary why the nominal scale, which was originally introduced from the operational point of view, came to be ignored in the development of axiomatic approach. The following consideration will also include some insights into this new trend of measurement science.

2. A FORMAL THEORY OF CLASSIFICATION

Classification of objects by predicates

Let $X = \{x_1, x_2, \cdots, x_n\}$ be a finite set of objects which we are going to classify, and $Y = \{y_1, y_2, \cdots, y_m\}$ be a set of predicates, or attributes, which are relevant and meaningful to describe the members of X. That is, we assume for all pairs of x_i and y_j the proposition 'the object x_i satisfies, or affirms, the predicate y_j' is meaningful whether it is true or false. We further assume that each predicate is applied to each object either affirmatively or negatively. We denote this object-predicate relationship τ, and call the triplet $\langle X, Y, \tau\rangle$ the *object-predicate system.* Such an object-predicate relationship can be identified with a table T, which is called the *object-predicate table*, with n rows and m columns, whose (i, j)-component is 1 or 0 according to whether an object x_i affirms or negates a predicate a_j.

Let us take a subset A of predicates from Y. Since our object-predicate system is two-valued, an object, say x_i, divides a set A of predicates to two parts; one is a subset whose elements are satisfied by x_i, and the other is a subset whose elements are negated by it. We call the former an *affirmative subset of A with respect to* x_i, denoting $A\,(x_i)$:

$$A\,(x_i) \equiv \{a \in A \mid x_i \text{ affirms } a\ \}. \qquad (1)$$

For two members of X, say x_i and x_j, we say these two are *indiscernible with respect to A*, denoting $x_i \overset{A}{\sim} x_j$, if corresponding affirmative subsets are the same:

$$x_i \overset{A}{\sim} x_j \iff A\,(x_i) = A\,(x_j). \qquad (2)$$

This definition can be paraphrased by using the object-predicate table T as follows: Take a subtable of T, say T_A, by picking up columns corresponding to the elements of A. Then x_i and x_j are indiscernible under A if two rows of T_A which correspond to x_i and to x_j are identical.

The equivalence relation $\overset{A}{\sim}$ induces a partition on X in a natural way. Regarding this partition as a classification, we call it a *classification of X with respect to A*, or *under A* for short, and denote this family of equivalence classes $\mathscr{C}(A)$, instead of the usual mathematical notation $X/\overset{A}{\sim}$, to emphasize the role of the set A of predicates as classifier. If A has only a single element, $A = \{a\}$, then we write $\mathscr{C}(a)$ instead of $\mathscr{C}(\{a\})$ to simplify our notation.

Two subsets of predicates taken from Y, say A and B, will give two ways of classification, $\mathscr{C}(A)$ and $\mathscr{C}(B)$, of the same set X of objects. When the indiscernibility under B implies that of under A, that is, when all indiscernible pairs under B are also indiscernible under A, we say $\mathscr{C}(B)$ is *finer* than $\mathscr{C}(A)$, or the former is a *refinement* of the latter, denoting $\mathscr{C}(A) \prec \mathscr{C}(B)$. If both of $\mathscr{C}(A) \prec \mathscr{C}(B)$ and $\mathscr{C}(B) \prec \mathscr{C}(A)$ hold then we say $\mathscr{C}(A)$ and $\mathscr{C}(B)$ are *same* classification, and denote $\mathscr{C}(A) \sim \mathscr{C}(B)$. If neither $\mathscr{C}(A) \prec \mathscr{C}(B)$ nor $\mathscr{C}(B) \prec \mathscr{C}(A)$ holds then we say these two classifications are *independent.*

Careful examination shows various, interesting properties of operations of classifier [14][15]. Here are excerpts of them given without proofs. In so far as we obey the logical rules faithfully the operation of classifier is not so simple as it might appear. It needs careful treatment.

Proposition 2.1 (*Absorptive law*) A classification given by A is less fine than a classification given by B if and only if $A \cup B$ and B give the same classification:

$$\mathscr{C}(A) \prec \mathscr{C}(B) \iff \mathscr{C}(A \cup B) \sim \mathscr{C}(B). \qquad (3)$$

Proposition 2.2 (*Addition of classifier*) The inequality of the finess of classification is preserved for the addition of any set of classifier:

$$\mathscr{C}(A_1) \prec \mathscr{C}(A_2) \Rightarrow \mathscr{C}(A_1 \cup B) \prec \mathscr{C}(A_2 \cup B). \qquad (4)$$

Corollary 2.1 (*Replacement of classifier*) If a subset A_1 of A and another set B of predicates give the same classification then A_1 can be replaced by B preserving the result of classification $\mathscr{C}(A)$.

Proposition 2.3 (*Subtraction of classifier*) Let A_1 and A_2 share a certain non-empty subset B in common. Even if $\mathscr{C}(A_1)$ is less fine than $\mathscr{C}(A_2)$, $\mathscr{C}(A_1 \backslash B)$ is not always less fine than $\mathscr{C}(A_2 \backslash B)$:

$$B \subset A_1 \cap A_2,\ \ \mathscr{C}(A_1) \prec \mathscr{C}(A_2) \not\Rightarrow \mathscr{C}(A_1 \backslash B) \prec \mathscr{C}(A_2 \backslash B). \qquad (5)$$

Multiplicity of classifier selection

One may conceive that a choice of predicates is determined anyhow by the purpose of classification, and that once the purpose is given then we can choose suitable set of predicates as classifier for the classification uniquely. However, this is not always true. From the formal point of view, same classification can be formed via different set of predicates, and in so far as we take various predicates with the same importance, we cannot select suitable set of predicates for the classification uniquely [14][15]. To decide which is suitable, or which is the best, we must bring in some additional factor, which may depend on our subjective preference or aesthetic taste. The existence of multiplicity of this kind was first discussed by Satosi Watanabe, the physicist, in his information-theoretical consideration on the basis of pattern recognition and clustering [19][20]. He showed that in so far as we attach same importance to the various predicates there exists no such thing as a class of similar objects in the world, and argued that if we acknowledge the empirical existence of such classes then it means we are bringing in non-uniform weighing of some extra-logical origin from the outside of our framework. This assertion is known as *Theorem of the ugly-duckling.*

	y_1	y_2	y_3
x_1	1	1	1
x_2	1	0	0
x_3	0	1	0
x_4	0	0	1

Table 1 Object-predicate table ($n=4$, $m=3$). In this table any pair of predicates discerns all four objects, and gives the same classification of $X = \{x_1, x_2, x_3, x_4\}$.

3. A FORMAL BASIS OF THE EQUIVALENCE

Classification of predicates by objects

So far we have discussed a classification of objects by predicates. However, since our formalization is completely symmetrical between the objects and the predicates, we can consider also a *classification of predicates by objects* by interchanging role of objects with that of predicates [14][15].

Let $A = \{a_1, a_2, \cdots, a_k\}$ be a set of predicates. An *affirmative subset of* X *with respect to* a_i is given by

$$X\,(a_i) \;\equiv\; \{x \in X \mid x \text{ affirms } a_i\,\}\,, \qquad (1')$$

as a counterpart of $A\,(x_i)$. Then the indiscernibility between two predicates, say a_i and a_j, is defined by

$$a_i \overset{X}{\sim} a_j \quad\Longleftrightarrow\quad X\,(a_i) \;=\; X\,(a_j). \qquad (2')$$

The equivalence relation $\overset{X}{\sim}$ induces a partition on A as before, which we call a *classification of A with respect to* X, denoting $\mathscr{C}(X)$. For two subsets of X, say X_I and X_J, we can consider relations such as $\mathscr{C}(X_I) \prec \mathscr{C}(X_J)$ and $\mathscr{C}(X_I) \sim \mathscr{C}(X_J)$, and obtain similar results as we stated above. Thus we can consider not only a classification of objects by predicates but also a classification of predicates by objects. Combining these two views, we have

Proposition 3.4 If two predicates a and b are indiscernible with respect to X then these two predicates give the same classification of X:

$$a \overset{X}{\sim} b \quad\Longrightarrow\quad \mathscr{C}(a) \sim \mathscr{C}(b). \qquad (6)$$

The converse, however, is false in general. Even if two predicates give the same classification, they are not necessarily indiscernible with respect to X.

Principle of object-predicate reciprocity

Considering the formal symmmetry between the objects and the predicates, this 'oneway-ness' seems rather peculiar. One may conceive that if two predicates give the same classification of X then these two should be identified with respect to X. Such a consideration leads us to the notion of reciprocity between the predicates as classifier and the objects to be classified [14][15]: *The indiscernibility of classifier and the coincidence of corresponding classifications should be equivalent.* We take this view of reciprocity for our heuristic guiding principle, and modify the notion of indiscernibility in order for the reciprocity to be maintained as follows: For two predicates a and b, we say these two predicates are *weakly indiscernible with respect to* X, denoting $a \underset{\mathrm{w}}{\overset{X}{\sim}} b$, if a and b are indiscernible with respect to X *or* if $\bar{a}$ and b are indiscernible with respect to X:

$$a \underset{\mathrm{w}}{\overset{X}{\sim}} b \quad\Longleftrightarrow\quad (a \overset{X}{\sim} b) \vee (\bar{a} \overset{X}{\sim} b), \qquad (7)$$

where $\bar{a}$ is the *negation* of a whose affirmative subset is given by $X(\bar{a}) = X \backslash X(a)$. Then two predicates give the same classification if and only if these two are weakly indiscernible with respect to X:

$$a \underset{\mathrm{w}}{\overset{X}{\sim}} b \quad\Longleftrightarrow\quad \mathscr{C}(a) \sim \mathscr{C}(b). \qquad (6')$$

Weak indiscernibility thus defined is also an equivalence relation, and induces a partition on A, which we call *weak classification of A with respect to* X, denoting $\widetilde{\mathscr{C}}(X)$ [14].

Indiscernibility between two groups

Now we come to a problem of interest. We already have defined the equality of two classifications $\mathscr{C}(A) \sim \mathscr{C}(B)$. Then how should we define the indiscernibility between those two sets of predicates, which is equivalent with the relation $\mathscr{C}(A) \sim \mathscr{C}(B)$, and deserves the notation $A \underset{\mathrm{w}}{\overset{X}{\sim}} B$ as a generalization of weak indiscernibility?

Let us take a subset A' of A, and extend the notion of affirmative subset $X(a)$, defined by $(1')$, to

$$X\,(A') \equiv \{\, x \in X \mid x \text{ affirms any } a' \in A'\,\}. \qquad (8)$$

We make a collection of such affirmative subsets $X\,(A')$'s for all subsets A' of A, and call it an *affirmative family of* X *with respect to* A, denoting $\mathscr{X}_A$:

$$\mathscr{X}_A \equiv \{\, X\,(A')\,\}_{A' \subset A}. \qquad (9)$$

One may conceive that $\mathscr{X}_A$ reflects all possible relations of X to A, and expect to obtain the desired definition by matching of such families of subsets. Careful examination shows, however, that we should take not only affirmative aspects but also negative aspects into account: Let us form a collection of subsets of X by augmenting the members of $\mathscr{X}_A$, by the use of $\cup$, $\cap$ and $\backslash$, until finally we can produce nothing new. We call this collection the *Boolean completion* of $\mathscr{X}_A$, denoting $\hat{\mathscr{X}}_A$. Then we have [14]

Theorem 3.1 The Boolean completion of the affirmative family $\mathscr{X}_A$ coincides with the Boolean completion of the affirmative family $\mathscr{X}_B$ if and only if A and B give the same classification of X:

$$\hat{\mathscr{X}}_A = \hat{\mathscr{X}}_B \quad\Longleftrightarrow\quad \mathscr{C}(A) \sim \mathscr{C}(B). \qquad (10)$$

We are now prepared to answer to our problem: We say two sets of predicates A and B are *indiscernible with respect to* X if the Boolean completion $\hat{\mathscr{X}}_A$ coincides with the Boolean completion $\hat{\mathscr{X}}_B$, denoting $A \underset{\mathrm{w}}{\overset{X}{\sim}} B$:

$$A \underset{\mathrm{w}}{\overset{X}{\sim}} B \quad\Longleftrightarrow\quad \hat{\mathscr{X}}_A = \hat{\mathscr{X}}_B. \qquad (11)$$

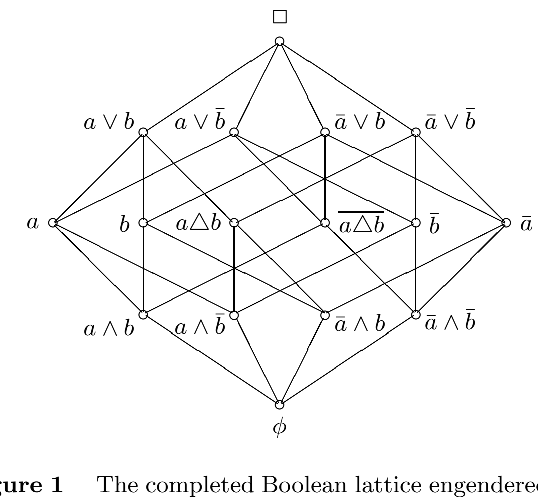

Figure 1 The completed Boolean lattice engendered by two predicates a and b. The symbol $\triangle$ stands for the exclusive or: $a \triangle b \equiv (a \wedge \bar{b}) \vee (\bar{a} \wedge b)$. The upward path between two nodes denotes implication relation. Note that if $a = y_1$ and $b = y_2$ of Table 1 then we have $y_3 = \overline{a \triangle b}$.

Then we obtain the desired relation

$$A \underset{\mathrm{w}}{\overset{X}{\sim}} B \quad \Longleftrightarrow \quad \mathscr{C}(A) \sim \mathscr{C}(B). \tag{12}$$

in accordance with (6′).

Our result can be paraphrased by using the logical operations, $\wedge$, $\vee$, and $^{-}$, defined on the object-predicate table: Starting from the members of A, we can form a collection $\hat{\mathscr{A}}_X$ of all possible predicates which are given as the *conjunction*, the *disjunction* and the *negation* of the members of A. In mathematical terminology $\hat{\mathscr{A}}_X$ thus defined is the *smallest completed Boolean lattice* of predicates that contains all members of A. Then we obtain [15]

Theorem 3.2 The completed Boolean lattice $\hat{\mathscr{A}}_X$ coincides with $\hat{\mathscr{B}}_X$ if and only if A and B give the same classification of X :

$$\hat{\mathscr{A}}_X = \hat{\mathscr{B}}_X \quad \Longleftrightarrow \quad \mathscr{C}(A) \sim \mathscr{C}(B). \tag{13}$$

Hence we can define the indiscernibility between two sets of predicates also by the coincidence of the corresponding completed Boolean lattices:

$$A \underset{\mathrm{w}}{\overset{X}{\sim}} B \quad \Longleftrightarrow \quad \hat{\mathscr{A}}_X = \hat{\mathscr{B}}_X. \tag{11′}$$

The result obtained above suggests that if we have no prior knowledge, and hence, if we have to take various predicates with same importance then we should start not from an arbitrarily chosen set of predicates but from the completed Boolean lattice of predicates. There are many set of predicates that engenders same completed Boolean lattice, and if we single out a certain set of predicates as 'elementary' predicate set without any legitimate reason then it means that we are bringing in some extra-evidential factor from the outside of our formal framework [19][20].

Basis of the equivalence

By interchanging objects with predicates once again, we obtain a definition of indiscernibility between two *groups* of objects. Then we have

$$X_I \underset{\mathrm{w}}{\overset{A}{\sim}} X_J \quad \Longleftrightarrow \quad \mathscr{C}(X_I) \sim \mathscr{C}(X_J). \tag{14}$$

In order to compare two groups of objects, therefore, we only have to consider a classification of *predicates by objects* instead of the completed Boolean lattice.

The results we obtained above suggest that our modified notion of indiscernibility between two groups of objects is based upon a structure of the completed Boolean lattice. This implies that if we reinterpret nominal scaling as a description of objects by predicates then we can regard the completed Boolean lattice as a basis of the nominal scale. In the current understanding of measurement theory, traditional four scale types are linearly ordered according to the algebraic structure, and the nominal scale is regarded as the most simple, trivial scale type. The reinterpreted nominal scale is released from the tight numerical structure, and can be distinguished from other three scale types.

Extension and intension

In the context of traditional philosophy the notion of concept is supposed to have two aspects, called *extension* and *intension*. Extension is a range of individuals to which the concept is applicable, while intension is the entire set of properties which specifies those individuals. If a predicate a corresponds to some concept then we can interpret our affirmative subset $X(a)$ as the extension of the concept. Let us denote a collection of such extensions for all members of A by $\mathscr{E}_A$: $\mathscr{E}_A \equiv \{X(a)\}_{a\in A}$. $\mathscr{X}_A$ includes $\mathscr{E}_A$, while the Boolean completion of $\mathscr{E}_A$ coincides with that of $\mathscr{X}_A$: $\hat{\mathscr{X}}_A = \hat{\mathscr{E}}_A$. Thus we can paraphrase the definition (11) of the indiscernibility between two sets of predicates once again:

$$A \underset{\mathrm{w}}{\overset{X}{\sim}} B \quad \Longleftrightarrow \quad \hat{\mathscr{E}}_A = \hat{\mathscr{E}}_B. \tag{11″}$$

It should be remembered that we defined the indiscernibility between two predicates by the coincidence of their extensions, (2′). This result suggests that in our formal framework $\hat{\mathscr{E}}_A$ plays more important role than $\mathscr{E}_A$ per se.

Besides, with the help of an affirmative subset $A(X')$, which is introduced as a counterpart of (8), we can define also the intension of the concept a by $A(X(a))$, which allows us to examine various formal relationship between two aspects of the concept. For simplicity, let us denote our formal extension and formal intension

$$\mathrm{ext}\,(a_i) \equiv X(a_i), \tag{15}$$

$$\mathrm{int}\,(a_i) \equiv A(\mathrm{ext}\,(a_i)) = A(X(a_i)), \tag{16}$$

respectively. In the object-predicate table, $\mathrm{ext}\,(a_i)$ corresponds to the rows whose ith component is 1, while $\mathrm{int}\,(a_i)$ corresponds to the columns whose components corresponding to the members of $\mathrm{ext}\,(a_i)$ are all 1. We have

$$\mathrm{int}\,(a_j) \subset \mathrm{int}\,(a_i) \Longleftrightarrow \mathrm{ext}\,(a_i) \subset \mathrm{ext}\,(a_j), \tag{17}$$

which shows the inversely proportional relation between the extension and the intension. That is, the more extension contains individuals, the smaller intension becomes. This result agrees nicely with the traditional notion of extension and intension. In particular, $\mathrm{ext}\,(a_j)$ includes $\mathrm{ext}\,(a_i)$ if and only if a_j belongs to $\mathrm{int}\,(a_i)$. Then we have

$$\mathrm{ext}\,(a_i) = \bigcap_{a_j \in \mathrm{int}\,(a_i)} \mathrm{ext}\,(a_j), \tag{18}$$

$$\mathrm{ext}\,(a_i) = \bigcup_{a_i \in \mathrm{int}\,(a_j)} \mathrm{ext}\,(a_j). \tag{19}$$

The first formula (18) shows another aspect of the inversely proportional relationship. These formulae will uphold our interpretation that $\mathrm{int}\,(a_i)$ is the set of predicates, or attributes, that characterize the group $\mathrm{ext}\,(a_i)$ of objects.

By virtue of the formal symmetry, we can consider also formal extension and intension of *object*: $\mathrm{ext}^*(x_i) \equiv A(x_i)$ and $\mathrm{int}^*(x_i) \equiv X(\mathrm{ext}^*(x_i)) = X(A(x_i))$. Regarding the predicates as attributes, we can interpret $\mathrm{ext}^*(x_i)$ as a collection of attributes that the object x_i has, and $\mathrm{int}^*(x_i)$ as a group of objects whose member has all attributes that x_i has. Those member may have further another attributes. The intension of predicate and the extension of object are related by

$$\mathrm{int}\,(a_i) = \bigcap_{x_j \in \mathrm{ext}\,(a_i)} \mathrm{ext}^*(x_j), \tag{20}$$

which shows, in accordance with the traditional terminology once again, that $\mathrm{int}\,(a_i)$ is a collection of attributes which are shared by all the members of $\mathrm{ext}\,(a_i)$. (20) gives another expression of the inversely proportional relationship between the extension and the intension of predicate.

Extension measures the *richness* of a concept in terms of participation of individuals in the class corresponding

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}
x_1	0	1	1	1	1	1	0	1	1	1	1	1	1	1
x_2	1	0	1	1	1	1	1	1	1	0	1	1	1	1
x_3	1	0	0	1	1	1	1	1	1	0	1	1	1	1
x_4	1	1	1	0	1	0	1	1	0	1	1	0	0	1
x_5	0	1	1	1	0	1	0	1	1	1	1	1	1	1
x_6	1	1	1	1	1	0	1	1	1	1	1	0	1	1
x_7	1	1	1	1	1	1	0	1	1	1	1	1	1	1
x_8	1	1	1	0	1	0	1	0	0	1	1	0	0	1
x_9	1	1	1	1	1	0	1	1	0	1	1	0	0	1
x_{10}	1	1	1	1	1	1	1	1	1	0	1	1	1	1
x_{11}	1	1	1	1	1	0	1	1	0	1	0	0	0	1
x_{12}	1	1	1	1	1	1	1	1	1	1	1	0	1	1
x_{13}	1	1	1	1	1	0	1	1	1	1	1	0	0	1
x_{14}	1	1	1	1	1	0	1	1	0	1	0	0	0	0

Table 2 Gene expression pattern [11][17]. These data are fictitious for explanation. x_i is a gene or gene product. a_j is a predicate 'is expressed under the condition that j th gene is disrupted' (1: expressed, 0: not expressed).

to the concept [20]. Now let us consider how to estimate the degree of the richness of a concept. Let n be the number of the members of X, d_i be the number of the members of $\mathrm{ext}\,(a_i)$, and we denote the number of the column types corresponding to $\mathrm{ext}\,(a_i)$ in the object-predicate table $|\ \mathrm{ext}\,(a_i)\,|$. If a set A of predicates itself is completed then we have $|\ \mathrm{ext}\,(a_i)\,| = 2^{d_i}$. In general we can regard $|\ \mathrm{ext}\,(a_i)\,|$ as the degree of the variety of the members of $\mathrm{ext}\,(a_i)$. Indeed, even if d_i is large, if $|\ \mathrm{ext}\,(a_i)\,|$ is small then, in so far as the set A is concerned, the members of $\mathrm{ext}\,(a_i)$ are more or less of similar type, and we should say that the content of the concept a_i is not so fruitful. Hence we may safely take $|\ \mathrm{ext}\,(a_i)\,|$, or an increasing function of it, as a measure of the degree of richness of the concept a_i.

In a similar way we denote the number of the column types corresponding to $\mathrm{int}\,(a_i)$ in the object-predicate table $|\ \mathrm{int}\,(a_i)\,|$. If A itself is completed then $|\ \mathrm{int}\,(a_i)\,| = 2^{n-d_i}$, and we have $|\ \mathrm{ext}\,(a_i)\,||\ \mathrm{int}\,(a_i)\,| = 2^n$. That is,

$$\log\ |\ \mathrm{ext}\,(a_i)\,|\ +\ \log\ |\ \mathrm{int}\,(a_i)\,|\ =\ \mathrm{const.} \qquad (21)$$

If A is not completed then the object-predicate table lacks some column types, and we have to replace the equality of (21) with the inequality '$\leq$' in general. This relation, a counterpart of (7.214) in [20], gives another expression of

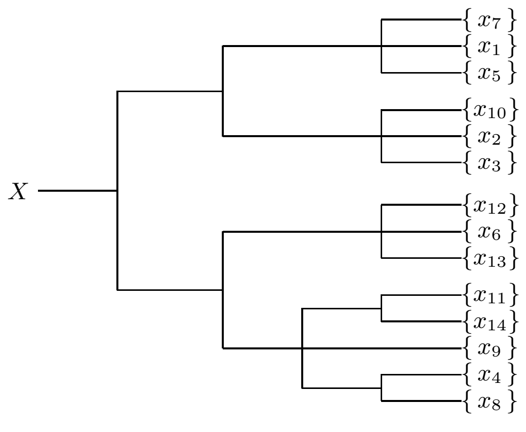

Figure 2 Clustering of fourteen fictitious genes according to the gene expression pattern described in Table 2.

	a_7	a_1	a_5	a_{10}	a_2	a_3	a_{12}	a_6	a_{13}	a_9	a_{11}	a_{14}	a_4	a_8
x_7	0	1	1	1	1	1	1	1	1	1	1	1	1	1
x_1	0	0	1	1	1	1	1	1	1	1	1	1	1	1
x_5	0	0	0	1	1	1	1	1	1	1	1	1	1	1
x_{10}	1	1	1	0	1	1	1	1	1	1	1	1	1	1
x_2	1	1	1	0	0	1	1	1	1	1	1	1	1	1
x_3	1	1	1	0	0	0	1	1	1	1	1	1	1	1
x_{12}	1	1	1	1	1	1	0	1	1	1	1	1	1	1
x_6	1	1	1	1	1	1	0	0	1	1	1	1	1	1
x_{13}	1	1	1	1	1	1	0	0	0	1	1	1	1	1
x_9	1	1	1	1	1	1	0	0	0	0	1	1	1	1
x_{11}	1	1	1	1	1	1	0	0	0	0	0	1	1	1
x_{14}	1	1	1	1	1	1	0	0	0	0	0	0	1	1
x_4	1	1	1	1	1	1	0	0	0	0	1	1	0	1
x_8	1	1	1	1	1	1	0	0	0	0	1	1	0	0

Table 3 Rearranged expression pattern corresponding to the result of clustering. It is easy to see that $\{x_7, x_1, x_5\}$ and $\{x_{10}, x_2, x_3\}$ respectively form classes. Remaining 8 members are classified also into two groups. See Figure 2.

the inversely proportional relationship between the extension and the intension defined in our formal framework.

Furthermore, we can improve estimation of the degree of the richness of concept by replacing $\log|\ \mathrm{ext}\,(a_i)\,|$ with

$$S\,(\,\mathrm{ext}\,(a_i))\ \equiv\ -\sum_j p_j \log p_j\ , \qquad (22)$$

where p_j is a relative frequency of the appearance of the j th column type in the subtable of the object-predicate table corresponding to $\mathrm{ext}\,(a_i)$. From the point of view of statistical physics, the relationship between $\log|\ \mathrm{ext}\,(a_i)\,|$ and $S\,(\,\mathrm{ext}\,(a_i))$ thus defined is analogous to the relationship between Boltzmannian entropy and Gibbsian entropy.

4. ANALYSIS OF INTERDEPENDENCE AMONG GROUPS OF OBJECTS

The strength of cohesion among groups of objects
The notion of indiscernibility between two groups enables us to examine a complicated relationship among groups of objects. In order to clarify our issue, let us consider a partitioning of a set of objects: Suppose that we are given a set X of objects, and asked to classify its members into two groups based on the relations reflected on the corresponding object-predicate table. How should we divide it?

There are obviously many different ways of thinking about how we look at such a partitioning, but it is natural to conceive that the entire set should be divided so that intragroup dependence is strong and that intergroup dependence is weak as far as possible. On the other hand, a subset of X that gives fine classification of predicates can be regarded as a group whose member has a full of variety as to which predicate satisfies and which negates, while a subset that gives less fine classification of predicates is understood as a group whose member satisfies and negates more or less same predicates, which suggests the exsitence of some binding force or coherence among the members. Hence we can consider that the finess of classification of predicates by objects reflects the degree of strength of cohesion, or interdependence, among those objects.

Let us take a subset X_I of X, and suppose that we divide X into X_I and $X \backslash X_I$. If $\mathscr{C}(X_I) \prec \mathscr{C}(X \backslash X_I)$ holds then we have $\mathscr{C}(X \backslash X_I) \sim \mathscr{C}(X)$ by Proposition 2.1, and we have to consider that $X \backslash X_I$ and X are essentially the same. When we divide the entire set into two subgroups, therefore, we should divide it so that those subgroups give mutually independent classification of the predicates.

Interdependence analysis among groups of objects Now let us newly denote our $S(\mathrm{ext}\,(a_i))$ of (22) $S(X_I)$, where $X_I \equiv \mathrm{ext}\,(a_i)$, as the quantity that characterizes the group X_I. This formal entropy $S(X_I)$ and the fineness of classification of predicates by objects are related by

Theorem 4.1 If a classification of a set A of predicates with respect to X_I is less fine than a classification with respect to X_J then the formal entropy of X_I is never larger than that of X_J :

$$\mathscr{C}(X_I) \prec \mathscr{C}(X_J) \quad \Longrightarrow \quad S\,(X_I) \leq S\,(X_J). \tag{23}$$

Then, in our terminology, if two groups are indiscernible, these two groups have the same formal entropy [16][17].

This quantity can be used for various useful purposes, one of which is analysis of interdependence among groups of objects. Suppose that we are given a set X of objects, and divide it into two groups X_1 and $X_2 \equiv X \setminus X_1$. Then the average entropy corresponding to this partitioning is

$$\langle\, S\,\rangle_{\{X_1,X_2\}} \;\equiv\; \frac{n_1 S(X_1) + n_2 S(X_2)}{n_1 + n_2}\,, \tag{24}$$

where n_i is the cardinality of X_i $(i = 1, 2)$. Close examination shows that the average entropy $\langle\, S\,\rangle$ thus defiend can be a good measure of the degree of interdependence between two groups X_1 and X_2, and that we can take this quantity as a criterion for choosing suitable partitioning of a set of objects: when we divide a given set of objects we should divide it so that this entropy is minimized [16][17].

This idea works fairly reasonable. The data, called gene expression pattern, as shown in Table 2 were taken from a survey paper on current topics of genome analysis. These are not the actual experimental data but the fictitious ones for introductory exposition. x_i $(i = 1, 2, \cdots, 14)$ is a fictitious gene mutually related in a certain regulation process [11], and a_j $(j = 1, 2, \cdots, 14)$ is a predicate 'appeared, or is expressed, under the condition that j th gene is disrupted'. Applying our strategy proposed above to these fourteen fictitious genes we obtain the polychotomic tree as shown in Figure 2, and, according to the result of clustering, the rows and the columns of Table 2 are rearranged as shown in Table 3, which allows us to infer, for instance, a causal relationship such as: $x_7 \to x_1 \to x_5$, and $x_{10} \to x_2 \to x_3$. That is, the proposed method helps us to decompose a complex regulation process into its elementary, constituent components, though we have no knowledge of how to understand their biochemical implications for the present [17]. As for details of algorithm, see [16][17]. The proposed method is very general in principle. Recently this method was extended, and applied successfully to structure analysis of quantum-mechanical systems [18].

5. CONCLUSION

The notion of quantity requires clear definition of equivalence, and reexamination of the concept of equivalence sheds a new light upon the basis of measurement. In the present paper we have examined the nature of classification procedure, and introduced a modified concept of indiscernibility among groups of objects. Measurement should be objective, while our result suggests that measurement practice inevitably requires some subjective factor such as intention of measurement. Our modified concept of indiscernibility may seem rather counter-intuitive: if we apply this concept to estimate the distance between two binary sequences, it tells us that a binary sequence and its 'negation' are essentially identical. This seems somewhat quire, considering the usual information measure such as Hamming distance. On the other hand, in analysis of complex systems, our modified concept helps us to understand complicated structure of complex system, identfying its constituent components. Such a formal framework will open up a new vista on newly emerging measurement science.

REFERENCES

[1] Berka, K.: *Measurement; its concepts, theories and problems.* D.Reidel, Dordrecht (1983)

[2] Berka, K.: Scales of measurement; a critical analysis of the concept of scales and of their function it the theory of measurement. In *Language, Logic and Method.* R.S.Cohen and M.W.Wartofsky (eds.), pp.1–73, D.Reidel (1983)

[3] Van Biesen, L. : Advances of measurement science. *IMEKO TC7 International Workshop on Advances of Measurement Science.* 13–19, Kyoto, Japan (1999)

[4] Finkelstein, L.: Representation by symbols as an extension of the concept of measurement. *Kybernetes*, 4, 215–223. (1975)

[5] Finkelstein, L.: The relation between the formal theory of measurement and pattern recognition. *IMEKO VII CONGRESS 1976*, CTH/301/1–9. (1976)

[6] Finkelstein, L. and Leaning, M.S.: A review of the fundamental concepts of measurement. *Measurement*, 2, 1, 25–34. (1984)

[7] Krantz, D.H., Luce, R.D., Suppes, P. and Tversky, A.: *Foundations of measurement.* 3 volumes. Academic Press, San Diego (1971, 1989, 1990)

[8] Mari, L.: The meaning of "quantity" in measurement. *Measurement*, 17, 127–138. (1996)

[9] Mari, L.: The role of determination and assignment in measurement. *Measurement*, 21, 79–90. (1997)

[10] Mari, L.: Beyond the representational viewpoint: a new formalization of measurement. *Measurement*, 27, 71–84. (2000)

[11] Somogyi, R., Sniegoski, C.A.: Modeling the complexity of genetic networks: understanding multigenic and pleiotropic regulation. *Complexity*, 1, 45–63 (1996)

[12] Stevens, S.S.: Measurement, psychophysics and utility. In *Measurement: definitions and theories.* C.W.Churchman and P.Ratoosh (eds.) pp.18–63, Wiley, New York (1959)

[13] von Helmholtz, H. : Zählen und Messen—erkenntnistheoretish betrachtet, Philosophische Aufsätze Eduard Zeller Gewidmet (1887). In Gesammelte Abhandlungen, vol.3, pp.356–391 (1895). Numbering and Measuring from Epistemological Viewpoint. (English translation by M.F.Love) In *H.v.Helmholtz, Epistemological Writings*, Boston Studies in the Philosophy of Science, vol.37 (1977)

[14] Watanabe, H.: Theory of classification: a new aspect of the nominal scale. *Proceedings of IMEKO-XV*, II, pp.147–154. Osaka, Japan (1999)

[15] Watanabe, H.: Theory of classification and its application to the interdependence analysis. *Proceedings of IEEE SMC'99*, VI-1031–1038, Tokyo, Japan (1999)

[16] Watanabe, H.: A new view of the formal entropy as a measure of interdependence and its application to pattern recognition. *Proceedings of IEEE SMC 2000*, 2827–2832, Nashville, USA (2000)

[17] H. Watanabe: Clustering as average entropy minimization and its application to structure analysis of complex systems. *Proceedings of IEEE SMC 2001*, 2408–2414, Tucson, Arizona, USA (2001)

[18] H. Watanabe: Interdependence analysis in a composite quantum system and its application to partitioning of quantum dot register, *Superlattices and Microstructures* (in printing)

[19] Watanabé, S.: Une explication mathématique du classement d'objets. In *Information and Prediction in Science.* S. Dockx and P. Bernays (eds.), pp.39–76, Academic Press (1965)

[20] Watanabe, S.: *Knowing and guessing: a formal and quantitative study.* Wiley, New York (1969)