CLASSIFICATION

par L. APOSTEL
Gewoon hoogleraar R.U.G.

1. *INTRODUCTION*

The universal significance of the operation of classification lies in the fact that we receive too much information about the external world to be able to grasp the whole of it, to preserve the whole of it and to retrieve useful bits of information if we are compelled to register and preserve everything. We must diminish the quantity of information received, and in order to do this we must introduce redundancies (this means we must organise the information).

But on the other hand, these redundancies, this organisation, facilitating reception, preservation and retrieval, must be such that not too much information is lost. These double aims, organisation of information, and preservation of information cannot be simultaneously achieved in an optimal fashion. Some compromise is necessary. The concept of a "natural classification" is precisely the instrument we use to obtain this result. Indeed, we can say informally that *a natural classification is a classification that represents as well as possible as many as possible other classifications of the same region of reality.* To put it otherwise : if we know that two objects a and b belong to the same class in the natural classification N, then we know for a certain number of other classifications of the same domain that a and b equally belong to the same classes in these other ones, or to not too different classes. But it can not be true that a classification is completely natural. (In the only non trivial case : where the *n* partitions to be represented in *one* classification are *not* refinements of each other). For this reason, we must define a degree of naturality of a classification, and this degree of naturality must be defined with reference to a set of other classifications of the same region. It is the aim of this article to point out the necessary distinctions that have to be made in order to define this concept. This first paragraph had the intention to show that

a) for a given definition of "natural classification" or of "degree of naturality of a classification", there will not in general be a uniquely defined "most natural classification" of the region,
b) there will not be a unique definition of "degree of naturality of classifications".

We start with the most general formal account of classification we know of; we give a brief exposition of it, and we point out various generalizations of it and various specifications and additions.

Extrait de la Revue (R.E.L.O.)
VII, 1 à 4, 1971. C.I.P.L. - Université de Liège -

2. *THE CONCEPT OF "CLASSIFICATION"*

Let S be a region to be classified. Let M be a matrix of classes with n rows. The first row contains S itself, and is indicated as $S_{1,1}$ (first element of first row). The second row contains (for instance) four classes : S_{21}, S_{22}, S_{23}, S_{24}. Such a matrix has n rows (n can be finite or infinite), each row contains s classes (s can be finite or infinite); each class contains t elements (t can be finite or infinite). In the simplest cases, s and t are all finite. Every row is a partition; this implies

1. No element of the partition is empty;
2. All intersections of members of the same row are empty;
3. For all rows, the union of the elements of that row is equal to S.

But in order to have a classification of S, it is not sufficient to have at our disposal an ordered sequence of partitions of S. The order of the set of partitions must be intrinsically determined by that set. It is well known that the only semi-order intrinsically determined by a sequence of classes is class inclusion. (K_1 is included in K_2, if and only if every element of K_1 is also an element of K_2). In this intrinsic semi-order on a sequence of partitions. One partition intrinsically *precedes* another if P_2 is a refinement of P_1 . P_2 is a refinement of P_1 if every class of P_2 is either identical to a class of P_1 , or is included (as a subclass) in a class of P_1 . A refinement P_1 of P_2 is progressive if at least one class of the refining partition is a real subclass of a class of the partition it is a refinement of. A refinement is strictly progressive if all classes of the refining partition are real subclasses. As postulate 4, we stipulate that the ordered sequence of partitions must be progressive at all levels (strict progressiveness is desirable but not imposed).

3. *GENERALIZED CLASSIFICATIONS*

The earlier concept of classification is the one that has been used in the two publications that have studied the concept of natural classification in the most general fashion (Luscewska-Romahnowa, 1961, Apostel, 1963). We claim however that the concept is too stringent and that, in order to conform to classificatory practice, we have to generalize it in various directions; suggested by Dobrowolski, Hillman, and Apostel.

1) Dobrowolski stresses that not every branch of a classification tree has a prolongation upon every level. This obliges us to weaken postulate 3. as follows : for all rows, the union of the sets of the row is included in the set S (and not : is identical to the set S). This is a very large departure from our starting point, and in order to stay somewhat closer, it might be advisable to add : at least some unions of rows are identical to S, while all unions of rows are included in S.

2) Dobrowolski stresses also that in order to allow the classification to grow, some empty classes should be allowed. This naturally introduces an *intensional* element (2). We have to substitute for postulate 1, the following stipulation : no class of any row is *necessarily* empty, and in any row there are at least some non empty classes (if we want some row to have more than one empty class, then we must develop a non extensional logic of classes. Only in modal classification theory can this be formalized).

3) A third weakening of the postulates would be that we do not ask that all partitions are progressive refinements. We simply ask that some partitions are progressive refinements (and, if we want to generalize in this direction, without getting too far away from our starting point, we can ask moreover that no non-progressive refinement be the last refinement in the branch in which it occurs).

4) The classical idea of partition presupposes the fact that every row is a Boolean algebra : for every two elements of the row it is possible to form the meet and the join, and every element of the row has the union of all other elements of the row as its complement. Hillman proposes to use not only Boolean algebras but also Brouwerian algebras in a given row of the classification matrix (and next to Brouwerian algebras, subtractive lattices). In order to see what changes would be produced in the postulate set for classifications, let us rewrite Postulate 2 (page 3) as follows (Post 2') : the union of all members of a row, different from this member, is the complement of it. This postulate implies both postulates 2 and 3 (page 3). Let us now weaken it, by splitting the concept of "complement" in two

a. for a given class K, (of S) let -K be the largest set of elements of S that are with certainty not in K (if there are undecided or undecidable cases, even though they might belong outside of K, they will not be found here) : this is the pseudo-complement of K.
b. let K be the smallest set of elements containing all elements not in K (here the undecided cases are included, but some elements of S may be also included - The Brouwerian complement).

Consider then two substitutes for $P_2{}'$: let the union of non members of K_1 be the pseudo-complement of K_1, or let it be the Brouwerian complement of K_1 (both relative, not to S but to the union of all the members of the row in question).

5) Finally, to remain close to practice, it is necessary to consider the partitioning of fuzzy sets (Zadeh), and to consider fuzzy partitions. This has already been stressed in Apostel, 1963, but the theory of fuzzy sets was not at our disposal at that moment. Let us consider first a classical set S, but let us consider for every class K_i, subclass of S a kernel of elements, (KK_i) certainly contained in it, and an indetermination region the elements of which either are both in and out the set, or neither in nor out the set (IK_i). We say that we have a fuzzy classification of a set S if :

I. All sets in any row are either non empty, or are included in the indetermination class of the empty class.

II. Any row is a sequence of pairs, the first element of a pair being the kernel of a class, the second element being the indetermination class of It. For any pair (K_i-IK_i, K_j-Ik_j) we can consider the following intersections : K_iK_j, IK_iIK_j, K_iIk_j, K_jIk_i and stipulate
   a) the first intersection is empty;
   b) none of the terms of any intersection are identical to the intersection;
   c) the three last intersections are for all pairs of a given row of the same order of extension (to put it strongly : of same cardinality);
   d) all the three last intersections are small (to be precise : have less than n elements).

III. In as far as members of a given row have a progressive prolongation, on each new level some members of I-classes of the earlier level become members of K-classes (the indeterminacy is monotonically decreasing somewhere or everywhere).

4. *TWO CONCEPTS OF NATURAL CLASSIFICATION : LUSCEWSKA AND APOSTEL*

Having, in order to come closer to scientific practice, defined various directions in which the theory of ordered sequences of partitions has to be generalized, we are now going to return to the classical point of view, embodied in Luscewska-Romahnowa's article, showing how the definition of "naturality" present in it can be generalized (in the direction of Apostel 1963, but giving various much needed definitions lacking there). The basic idea of Luscewska is simple :

a) given a classification matrix, one defines a distance between elements, in function of this matrix,
b) one defines *independently* a distance among elements of the set S,
c) one calls the classification natural, if *both* distances have *the same values.*

Our own theory differs from hers on the following points

a) we want to study *in general* functions measuring the degree of adaptation of classification-parameters, to set parameters. The identity function is a much too simple special case of such an adaptation measure. And in general, we do not think it advisable to define only the classificatory concept " the classification M is natural with respect to a distance function on S" but we need an ordinal or even a quantitative measure of adaptation, defining "degrees of naturality". We want to consider as large as possible a class of measures of adaptation;
b) we want to study the degree of adaptation of one classification parameter to a multiplicity of set parameters and not only to one set parameter;
c) we want to define classes, orders, and distances (numerical or non numerical ones) both upon the classification matrix and upon the set to be classified and we have to define the meaning of the expressions "a classification with given degree of naturality, from the point of view of relations between classes, orders or distances; the first ones defined upon the matrix and the second one upon the set".

5. *THE DISTANCE FUNCTIONS*

A function d of two arguments is a distance in the classical sense of the word if

1. it maps the arguments on the set of natural or real numbers;
2. if $d(xx) = o$;
3. $d(xy) = d(yx)$;
4. $|d(xy) + d(yz)| \geqslant$ to $d(xz)$.

A generalized distance (defined in Luscewska-Batog, 1965) is a mapping of the set of pairs of arguments upon a semi lattice (with the union and the inclusion as lattice operations), where the last postulate becomes

$$d(xz) \subseteq |d(xy) \cup (dyz)|$$

A rather natural method to define distance (a,b) in function of the classification tree is the following one : we define as the line of an element of S, in the matrix M, the sequence of all classes of which a is an element, provided

a) that it is a strictly increasing sequence under inclusion (its earlier elements are all included in its later elements) and

b) that in the sequence is to be found one and only one member of every row of the Clf matrix.

The lines of any two elements of S meet in a given row of the matrix (eventually only in the first row : the smallest class in which both elements are included is S itself). It is clear that the closer the two elements are according to M, the earlier their lines must meet. If we now map the n rows of M upon the numbers (1, 2 .... n), and if w is the number of the first row in which the lines of a and b meet (have a common member), then the M distance will be n-w (it is clear that this number will be large for w small, and that w will be small when the lines meet closer to the top, as intuition demands). Even for transfinite matrices, Luscewska-Batog use a transfinite sequence of indexing numbers attributed to the various rows, in order to be able to preserve this intuition. It seems to us important that *it is possible to define in at least three different ways other distances in function of the classification*

a) we can take as distance the ordered or unordered pair of the two segments of the lines (IL), that have no common members. It is easy to verify for the unordered pair that the distance of an element to itself is the empty class (any line having all elements common with itself), that the distance is symmetrical (the union being commutative) and that the distance has the triangular property (the union of SLx with SLz, is included into the union of SLx with SLy, and of SLy with SLz) (3). If we take the ordered pair we lose the symmetry.

b) We can also call "distance" the complement (with reference to the set of all classes of M) of the common part of the lines of x and of y. This set varies also inversely with the length of this common part (and thus directly with the closeness in the classification of the elements x and y). Here we have the following properties :

   1. the complement of the common part of two identical lines is the complement of each of this lines itself;
   2. the complement of the intersection of the final segments of two sequences has the symmetry property, intersection being commutative;

c) finally (and this suggestion is to be preferred for all classifications in which the final rows are not unit classes) we can define an order upon the classification tree in any conventional manner (for instance starting with the top, going to the extreme left of the tree, and then coming back inserting all diverging branches until their base point). The distance of any two members of the matrix is the interval, in the order so defined, between the two given points of the tree (4). The distance properties are verified. This distance defined upon the classification can be used to define a distance between elements of S with reference to the classification, as follows : d(ab) is the union of the distances of all elements of the lines of a and b to each other.

The three generalised distance definitions are non numerical; they are intuitively satisfying; they can, being non numerical, be applied to infinite classifications without introducing transfinite ordinals or cardinals (a fragment of set theory suffices). The last definition has the drawback of being dependent upon a conventionally chosen ordering of the M matrix. One could overcome this obstacle by considering a multiplicity of orderings and taking as distance in the classificationtree the union of distances in these various orderings. We do not pursue here this topic further.

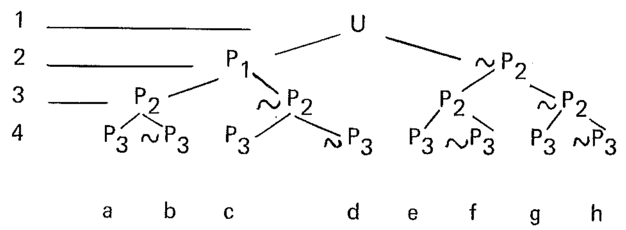
6. *SET DISTANCE SHOULD NOT ALWAYS BE IDENTICAL TO CLASSIFICATION DISTANCE*

As explained before according to Luscewska-Romahnowa, a distance function independently defined upon S will determine an n level classification, of S, if and only if there is a classification M with n rows of the same set S and if the distances of the elements with reference to the classification are identical to the distances in S itself.

Using two examples, we want to demonstrate that this stipulation is acceptable for given ways of defining the independent distances, and for given methods of classification but inacceptable for others.

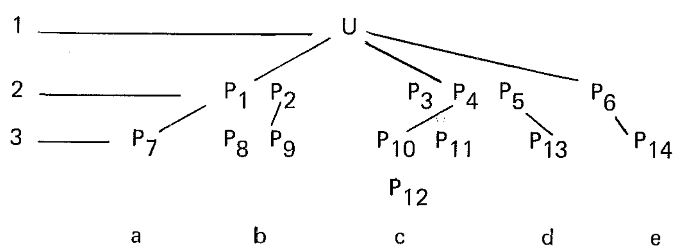
We consider in both cases a set S upon which only one argument predicates and no relations are defined. We consider in both cases two methods to define the independent distance. Both methods can only be applied when the list of primitive predicates characterizing the objects of the given set is determined and when it is finite. DI(ab) is the number of non shared primitive predicates. D2(ab) is the number of non shared primitive predicates, divided by the total number of primitive predicates (set dependent). LRD is Luscewska's concept (classification dependent).

Example I.



a has : $P_1\ P_2\ P_3$
b has : $P_1 \sim P_2 \sim P_3$
$DI(a,d) = 2$
$DII(a,d) = 2/3$
$LRD(a,d) = 4-2 = 2$
So : $DI(a,d) = LRD(a,d)$
$DII(a,d) \neq LRD(a,d)$

Example II.



a has : $P_1\ P_2\ P_7$
b has : $P_1\ P_2\ P_8\ P_9$
$DI(ab) = 3$
$DII(ab) = 3/5$
$LRD(ab) = 3-2 = 1$
So : $LRD(ab) \neq DI(ab)$
$\neq DII(ab)$

Our first example has been widely discussed : it gives all possible combinations of the primitive predicates; characterizes every row by means of one unique predicate and allows to preserve its strictly dichotomic structure at the expense of the definition of classes by means of negative predicates. The second classification is much less regular but much more realistic :

a) the various rows have different numbers of members;
b) no negative predicates are allowed;
c) the various classes are defined by different numbers of primitive predicates occurr (this last feature mirrors the existence of law-like regularities in the field).

Both definitions of distance can be defended as natural. The existence of these two acceptable distance definitions and the frequent occurrence of classifications, considered natural and being of the second type, shows convincingly that even if we want to define the naturality of a classification with reference to a comparison of distances, defined in the set and defined in the classification, *we cannot simply ask for identity.* Let us only *ask that the classification distance be a function of the set-distance, and moreover a monotonic increasing function* : whenever the set distance between a and b is larger than the set distance between cand d, then the classification distance between a and b should also be larger than the set distance between c and d.

7. *GENERALIZED CORRESPONDENCES BETWEEN SET PARAMETERS AND CLASSIFICATION PARAMETERS*

For this generalization of the requirement of correspondence of the two distances, we can also define an order of naturality. For a given set distance (SD), one classification distance (KD_1) is more natural than another classification distance (KD_2) if there are more refinements of KD_2 than of KD_1 that are also function of SD. We call a classification distance a refinement of another classification distance if it is the distance defined upon a classification matrix, but moreover some additional rows either added between existing rows in such a way that the classification postulates remain true, or added at the base.

We consider to be the aim of the theory of classification, the discovery of the following facts :

a) a measure (or an infinite set of measures) of the degree of naturality of a classification;
b) an algorithm allowing to classify a given set according to a classification of any given degree of naturality (whenever such an algorithm exists).

In order to reach these aims we can now leave our starting point (the generalization of Luscewska-Romahnowa's conception of classification, of distance and of naturality), and define various other standards in comparison with which a classification is judged to be more or less natural.

These standards must in general be dependent upon weaker functions than upon distances defined upon the set to be classified. A distance is nothing more than numerical order. So instead of characterizing the natural similarities and dissimilarities between sets and classifications by means of distances, we can take as a standard

a) an already existent classification;
b) an already existent order;
c) or a multiplicity of classifications, or of orders, or a mixture of classifications and orders.

There are at least two reasons to make this move

a) the examination of existent classifications, considered to be natural, shows that they conform to preexistent classifications or orders;
b) if it is possible to take a predefined distance as standard, it is necessarily also possible to take either predefined classifications or orders as standards (in as much as there cannot be a distance without there being necessarily a classification and an order, the converse of this statement being not true).

This weakening of the standards of naturality is not however for us an aim in itself. It is only a preparation for the more important task : write general axioms for the adaptation measure that should define the degree of naturality.

8. *DEGREE OF ADAPTATION OF PARTITIONS TO PARTITIONS*

Classifications being ordered sequences of partitions, let us start by defining the meaning of the sentence "Partition Pl is closer to partition P_2 than P_3 is to P_2". The definition will be the following one : we superimpose P_1 upon P_2. Let the greatest common divisor of the two partitions be a partition $Pr_{1,2}$ satisfying the following two conditions

a) every set of $Pr_{1,2}$ is included in only one set of P_1 and is also included in only one set of P_2 ;
b) there is no partition Px of the same set satisfying condition a, and such that every set of $Pr_{1,2}$ is included in only one set of Px.

We define equally the greatest common divisor of P_3 and P_2 ; $Pr_{3,2}$. It is intuitively clear that the closer two partitions are, the more nearly identical sets they contain. Two sets are nearly identical if their greatest common divisor is either identical to one of the sets (and then the sets are completely identical) or if their greatest common divisor is the union of a very large set, and a very small set. The sentence under analysis can now be defined as follows : "$Pr_{1,2}$ contains more pairs such as described than $Pr_{3,2}$". This definition can only be applied in the finite case but, every pair of partitions having a greatest common divisor (Ore) it can be applied in all such cases. In the infinite case we can only have comparability as to closeness among partitions if $Pr_{3,2}$ is a subpartition of $Pr_{1,2}$. In this way we can order partitions as to their closeness. A more general idea is to define a comparative distance between 2 partitions P_1 and P_2. An algorithm for defining such a distance would be : take all pairs of elements of the set S. Determine the ratio of the number of pairs that share a P_2 class when they share a P_1 class to the number of pairs that share a P_3 class when they share a P_2 class. Let these numbers be $NC(P_1, P_2)$ and $NC(P_1, P_3)$. If $NC(P_1 P_2) > NC(P_1 P_3)$, the first distance is larger than the second distance.

9. *DEGREE OF ADAPTATION OF A CLASSIFICATION TO CLASSIFICATIONS*

Classification Clf_1 is closer to Clf_2 than Clf_3 is to Clf_2 if and only if taken in the order they have in the ordered sequence of rows, more rows of Clf_1 are closer to the corresponding rows of Clf_2 than corresponding rows of Clf_3 to those of Clf_2. We have to take two independent features into account because of the following difficulty : we can have a two-row classification, the two rows of which are completely identical to two initial or intermediate or final rows of another classification, and yet this classification is not as close as another one that has no identical rows but that has a more equal number of rows, many of which are more close to their correspondents than any other classification of comparable length. We can also meet the case of having to compare close partitions, that however have not the same rank in the development of their respective classification. We should perhaps either have to put new restrictions on the definition to express the fact that closeness of partitions has not the same significance at different ranks : closeness at later ranks should perhaps determine more strongly the closeness of classifications, distinguishing the comparison of classifications of equal lengths from the comparison of classifications of different lengths.

10. *DEGREE OF ADAPTATION OF A CLASSIFICATION TO A SET OF CLASSIFICATIONS*

A) If we have a workable definition of the closeness order of classifications to each other, we can proceed to define the degree of closeness of a classification to a set of other classifications. Various definitions could be proposed. We could say that Clf_1 is closer to a set of classifications K, than Clf_2 is to K if and only if Clf_1 is closer to the greatest common divisor of classifications in K, than Clf_2 is. But we did not yet define the greatest common divisor of a class of classifications. We define it as follows : taking the ordered sequence of partitions of all classifications of the set, we define the greatest common divisor of the partitions of corresponding rank. We thus obtain a sequence of partitions and we can not be sure in general that this sequence satisfies again the postulates of a classification. Let us then consider among all the classifications of the same set, any classification such that there is no other classification closer to the sequence of partitions just obtained as divisor (we can not in general be certain that there is only one classification that is closest; for this reason we do prefer to give our definition in this negative terminology). This classification will be called the greatest common divisor of the set of classifications. It will be clear that, even though this definition is the most natural one, it is quite complex (and in general an algorithm applying it will take a lot of time).

B) For this reason it is perhaps advisable to consider some other possible definitions of the closeness of a classification to a set of other classifications; in order to decrease complexity we shall let the relation depend upon a unique classification chosen among the members of the set : and this selection can be made in various ways. We can select the classification in the set K, Clf_1 is closest to, and also the one lying at maximal distance. The same pair can be determined for Clf_2. If Clf_2 is closer to the member of the set K, Clf_1 is closest to, than Clf_1 itself or if Clf_1 is more distant from the member of K at maximal distance, than Clf_2 (and certainly if both these conditions are satisfied), then Clf_2 is closer to the set than Clf_1.

We can say something in favour and something against every possibility mentioned (*but only a general axiom system about closeness measures could make us select, in general or for specific cases, one of these measures*). We can combine the various definitions in various ways. For instance, within the class K we can try to define a "representative". This means the following : we construct for every classification of the set, the subset of classifications of the set, the first classification is closer to than any other. The classification having the largest associated set (being closer to more other classifications of the set than any other) is then called the representative and Clf_1 is closer than Clf_2 to K if and only if Clf_1 is closer to the representative of K.

Comparing those various proposals, one realizes that much has to be done, starting with the comparison of sets (whose symmetrical differences give an immediately satisfactory definition for their degree of closeness), studying the comparison of partitions, and then of sequences of partitions, to come to *the comparison of sequences of partitions with sets of sequences of partitions.* We consider however this rather difficult concept to be the basic definition of a natural classification.

11. *DEGREE OF ADAPTATION OF A PARTITION TO A RELATION*

A) We are going to repeat our whole inquiry all over again, this time concentrating upon the comparison of classifications with relations (and more specifically with orders and semi orders). We do not have to study the degree of adaptation of classes to equivalence relations, because this topic is formally identical to the one analysed before. Before starting this inquiry, let us however make still two remarks about the topic we are now leaving

a) if we have at our disposal ways to order the degree of closeness of classifications to sets of classifications, we also have at our disposal ways of comparing the degrees of closeness of sets of classifications to other sets of classifications. Indeed all our techniques consist in reducing the comparison of a Clf with a set of Clf's to the comparison of one Clf with one other Clf. The same methods of reduction can be twice applied.

b) Luscewska-Romahnowa defines the degree of naturality by means of the comparison of two distances : one depending upon the intrinsic properties of the set, the other depending upon the properties of the classification to be evaluated. We can combine her idea with ours in the following fashion : let us define the distance of any two elements of the set to be classified, in all the classifications we are comparing.

This gives us, in the simplest finite case for every pair of elements a sequence of numbers nl......nr., representing their classification dependent distances in the various classifications to be compared. We want to utilize the matrix of these vectors to define the concept of "distance between two classifications". Indeed, let there be 2 pairs of objects in S : it is natural to say that the difference between the distances of the same pairs in different Clf's measures the distance between the classifications. For a pair of classifications we can take as distance measure the sum or the average of the difference between the distances of the same pairs. For a set K' of classifications let us now compute the distances we just defined. Let us take for any classification the average of its distances to all the others. If there is a classification such, that the average of its distance to all the others is minimal, then we shall call this classification a natural classification with reference to the set K'. It is not necessary that there is only one such classification, but in the finite case we can be certain that there always will be at least one such classification at minimal average distance (or at minimal maximum distance or at minimal minimal distance). This concept of distance between classifications can perhaps equally be generalized in a non metrical fashion, as we have indicated in an earlier paragraph. It can also be used to define classes of related classifications (being defined as it is put by Fernandez de la Vega, either monothetically or polythetically, either by the fact that the distance within one unique class does not grow larger than a given number or by the fact that the average distance within one class is smaller than the average distances of class members to non class members). We mention this last possibility because we can use this classification of classifications to see if there is or is not a real chance to find a natural classification.

B) Having made these remarks, we are now, as we have announced, taking the next step in order to study the relations between classifications and orders.
Let us start with the simplest case : let an anti reflexive, antisymmetrical, transitive and connex relation be defined upon the set S. We wish to define the meaning of the statement "The Clf_1 of S is closer to the relation R than the Clf_2 of S". This problem however is only the first of an infinite series of problems of the same type : instead of adapting a classification to an order of the elements of the set, we can also try to adapt a classification to an order on the n the Cartesian product of S with itself (supposing that the pairs, or triads or quandruplets asf of S are completely ordered). The method of Benzécri is one solution for one of this infinite series of problems (the second one). Its significance can only rightly be understood in the wider context we give it here.

A classification being, as we know an ordered sequence of partitions, it is natural to start with a definition of "the partition P_1 is closer to the order R than the partition P_2". An example taken from a small universe U will suggest some answers. Let U contain (abcde), and let the lexicographical order mirror the relation R upon U. The following bipartitions are possible : (a) (bcde), (ab) (cde), (abc) (de), (abcd) (e) and the following tripartitions (a) (b) (cde); (abc) (d) (e); we can also have quadripartitions. All the partitions mentioned *preserve the order* in this sense that when two elements belong, according to the partition to the same class, then all elements who ly in the order, between them, equally belong to that class (this was the definition used by Apostel 1963 for the adaptation of a partition to an order). We now are of the opinion however

a) that this condition is neither sufficient nor
b) necessary to define the degree of closeness of a partition to an order.

Suppose that we have a classification (ae) (bcd) and a classification (a) (bd) (ce). It would be natural to say that the second partition is closer to the order than the first. The reasons for this impression can be easily given : the order abcde can be transformed into the order aebcd, and also into the order abdce by means of a certain number of one place permutations. Let the shortest permutation having this effect be called the canonical permutation. We shall say that a partition, not satisfying the interval condition, lies closer to an order than another partition not satisfying the interval condition if the canonical permutation necessary to transform the given order into an order with reference to which both partitions satisfy the interval condition, has a smaller number of elements in the first than in the second case. This definition implies naturally that all partitions satisfying the interval condition are closer to the order than any partition not satisfying the interval condition. Among the partitions to be compared in our miniature universe, we claim that the bipartitions are preferable to the tripartitions, and that among the bipartitions the following two (ab) (cde) and (abc) (de) are preferable for the following reasons :

a) they only do not contain unit classes (a classification partition that has as many classes as there are elements is certainly trivial);

b) the classes of the preferred partitions are as close as possible to be one-one projectible upon each other (the degree of one-one projectibility could be measured by the following measure : "K_1 is more 1-1 projectible upon K_2 than is K_3" means "The symmetric difference between the subsets of K_1 1-1 projectible upon subsets of K_2 and K_1 or K_2 themselves is smaller than the symmetric difference between their analogues in the K_2 - K_3.1 comparison).

This allows us to give the following definition : Partition P_1 is closer to R than Partition P_2

a) if the canonical permutation of R necessary to make P_1 satisfy the interval condition is shorter than the canonical permutation necessary to make P_2 satisfy the interval condition

b) if P_1 and P_2 have equal values as to the associated canonical permutation, then P_1 is closer to P_2 if P_1 has less unit classes and

c) finally if P_1 and P_2 have equal values for the two first indices, P_1 is closer to R than P_2 if the vector having as first elements the number of one-one projectible sets, as second elements the number of nearly 1-1 projectible sets to degree d (d being the measure of the symmetric differences mentioned before), as third element the number of 1-1 projectible sets to degree d′ (d precedes immediately d′) dominates until the r th rank (r can be arbitrarily chosen) the similar vector for the pair P_2 .

This definition implies an evaluation as to the importance of the three indices, and is heavily dependent upon the finiteness of order and classifications. It is imperiously necessary to generalize these definitions in order to have wider applicability. But it is obvious that the problem is already so complex in the finite case that we should wait with this generalization.

12. *DEGREE OF ADAPTATION OF A CLASSIFICATION TO ONE RELATION*

A) Having thus studied an ordering of partitions with reference to a given relation, we now have to study the ordering of classifications with reference to a given relation R. We can see at least two mechanical methods to generalize the earlier definition for a sequence of partitions : 1. The first method is the following one : we order the rows of the classifications to be compared in some order of importance (either we consider the basic rows as being more important, or the superior rows, or the intermediate rows) and then we compare the most important row of Clf_1 with the most important row of Clf_2.

The degree of closeness of Clf_1 to R is larger than, equal to or smaller than the degree of closeness of Clf_2 (with reference to the order R) according to the relation between the degrees of closeness of the most important rows. If two classifications are equally close in the most important row, then we go to the row immediately following in importance in both and we order again with reference to this row. We continue the procedure until we have found a pair of rows that are unequal in closeness or until we do not have any rows left in one or both of the classifications. The method is again only applicable in the finite case and is artificial in this sense that we have to define an order of preference upon the rows. We can eliminate this last source of arbitrariness, and give all rows equal weight by saying that Clf_1 is closer to R than Clf_2 if the number of rows of equal rank of Clf_1 that are closer to R than their correspondents in Clf_2 is greater than the number of rows of Clf_2 that are closer to R than their correspondents of equal rank.

B) The second mechanical method supposes that we introduce orders upon the classes of a row of a classification. We then take any partition of this set of subsets of S as the following row (going up from top to bottom) if there is no other partition of this new set S' (the ordered sets of the first partition of S) that is closer to the order R than itself and we make this partition the immediately subsequent row of the classification. As our formulation shows clearly enough this procedure does not give unique results. The classifications constructed by means of this procedure have the same degree of closeness to the underlying order R as has the most basic row (in order to give a meaning to this statement we do not need to have numerical values for the degrees of closeness : if dc is the degree of closeness order, then dc(x,R) =dc(yR) if for all y and for all z, dc(x,R) greater than dc(w,R) implies dc(y,R) smaller than dc(z,R)). This second mechanical method again has somewhere an arbitrary feature in as much as we take some row as basic, measure its degree of closeness and then construct in some way the classification that is in some sense optical with reference to an order defined upon that basic row.

13. *WHEN ARE CLASSIFICATIONS POSSIBLE OR DESIRABLE ?*

We say that both methods are mechanical because none of them can allow us to decide for what type of orders a simple partition is optimal, and for what type of orders a classification is preferable. Intuitively, it seems that the Benzécri method has an immediate appeal : we shall say that a partition is closer to a relation of order upon the pairs of elements of S than another partition if the number of classes that contain more pairs of elements preceding, in the S^2 (= SxS) order, all pairs of elements having their first element within the class, the other outside the class, is larger for the first partition than for the second. This idea gives a natural criterion for the desirability of an n row classification. Let us consider a partition. There is need for a second row if, in the S^2 order there are at least three classes in the partition satisfying the following condition : there are more pairs the first element of which is chosen in K_1, the second in K_2 (or inversely) who precede in the S^2 order such pairs whose first element is chosen in K_1 and their second in K_3, than pairs satisfying the opposite relation. In that case we may say that, with reference to R, K_1 and K_2 ly closer to each other than K_1 and K_3 and we are justified to add to the initial row a second row, classifying K_1 and K_2 together as subclasses. We shall call classi-

fications justified. A classification having more justified classes in more justified rows is closer to a relation on S^2 than a classification having less justified classes in less justified rows. The degree of justification of a class K_1 is greater than the degree of justification of another class K_2 if the number of preceding pairs (in the definition of the justified superclass) is larger in K_1 than in K_2.
The reader will already have remarked that any order (abcde) defines a natural order upon S^2 (for instance : ab, ac, ad, ae; bc, bd, be; asf). This order is however not complete, ab and bc can not be compared. Inversely when we have an order upon S^2, it is possible to derive from it an order upon S under certain conditions. This shows us that the existence of a classification with reference to a relation upon S is not independent from the existence of a classification with reference to a relation upon S^2. We can now use higher powers of the Cartesian product and state in general : a classification Clf_1 is closer to a relation upon S and to a sequence of relations upon S^2, S^3, ... S^n than a classification Clf_1 if Clf_2 is closer to more orders R′, R″ .... R‴...... upon all these cartesian powers. Let us stress that for any order upon any n th cartesian power we can redefine the concept of justified rows of higher rank. The Clf_1 will be strongly closer to the sequence of relations on the cartesian powers if
a) it will be closer according to all relations of all order and
b) if moreover it will have more justified classes according to the order of majority of Cartesian products.
Let us not leave this subject without stating that there are certainly situations in which no natural classification can be achieved : if all initial partitions, or classifications, or orders are at large (1) and equal (2) distance of each other, the task should not be undertaken to construct at great cost a classification system. There are cases, as we have remarked, trivializing the problem; other cases making it unsolvable; many intermediary ones making it possible, with non unique solutions and some rare, important situations where the non trivial problem is uniquely solvable. By no means do we have at our disposal exact criteria to separate these cases from each other.

14. *THE DEGREE OF CLOSENESS OF A PARTITION TO A SET OF RELATIONS*

One should not be astonished to encounter here exactly the same situation as when we have to define the degree of naturality of a classification with respect to a set of classifications. We must try to replace the complex comparison by a simpler one. Again the simplest method will be to start saying that partition P_1 is closer to the set (R) of relations than P_2 if and only if P_1 is closer to the intersection relation of this set than P_2. This definition again will not do all the work because of the fact that very often this intersection relation will be the empty relation or, if not so drastically reduced, will still contain a very small number of pairs (and thus be too undiscriminating). In these cases we have to give other definitions. We can select among the set the relation that implies the largest set of other relations of the set (this selection is not necessarily unique). Let us call these relations the fundamental relations of the set. We then can say that P_1 is closer to the set if it is closer to a larger number of fundamental relations of the set. This definition can be used, even when the intersection relation is empty or too poor.
It is perhaps the place here to refer to incomplete orders (ordering relations that leave certain comparisons of elements undetermined) and to say that in determining the closeness to a set of relations some of which are incomplete, we can try to order the relations as to their degree of incompleteness, and prefer closeness to the most complete relations. Another method to overcome the emptiness or poverty of

the intersection, is to compare initial or final segments of relations (the n first or r last elements of the order) and consider their intersections. Let us take the largest initial or final segments having non trivial intersections and say that a partition is closer to a set of relations if and only if it is closer to the intersection of the longest segments having non trivial intersections. These definitions give some of the methods one might use to compare the closeness of partitions to sets of relations.

15. *THE CLOSENESS OF CLASSIFICATIONS WITH RESPECT TO SETS OF RELATIONS*

A) We shall perhaps find a suitable measure by considering an ideal case, using it to define approximations to this ideal case.

Let us consider a set of relations that is strictly increasing (in this sense that R_1 ab implies R_2 ab but not inversely for at least one pair ab, and the same asymmetrical implication holds for any R_i and R_j in the order mentioned). This is simply defining an order upon the set of relations. Let us now also define an order upon the set of rows of a classification and let us say that Clf_1 is more closely adapted to the ordered set of relations than Clf_2 if the first partition of Clf_1 is more closely adapted to the first relation in the order, and the second of Clf_1 is more closely adapted than the second of Clf_2 to the intersection of the first and second relation, and the third row of Clf_1 is more closely adapted than the third of Clf_2 to the intersection of the three first relations. This case is an ideal case in this sense that every later row utilizes a new ordering relation in its definition. We do not think that a more regular case of adaptation to a set of relations is possible.

B) If we want to generalize this concept then we come to the following conclusion : a Clf_1 is closer than a Clf_2 to a given set of relations (R) absolutely (without reference to an order defined on Clf_1 or Clf_2, and without reference to an order defined on the set of orders), if, when we compare one partition of Clf_1 to the totality of orders in (R) the Clf_1 wins if the first partition of Clf_1 is closer *more often* than the second. We execute this comparisons for all pairs of partitions of Clf_1 and Clf_2 and the classification that is closer is the one that has more winning partitions. This method avoids an artificial ordering on the sets to be compared, but loses the impact of the order of the ordered sequences of partitions typical for any classification. A more natural (but partly artificial) procedure is the following one : we order the set of relations and the set of partitions, and we compare Clf_1 to Clf_2 using first the comparative closeness of the first partition in each classification to the first order in r; then we relate the second partition in both to the second relation and we continue that way, making the final decision by counting as in the earlier case. Going one step further we can make the final decision more strongly dependent upon both orderings of sets by continuing the procedure only when the earlier comparisons have led to the equality of closeness result. These procedures do not reach the ideal we started out with because the feature of increasing content that is typical for the first comparison procedure is absent in the three last attempts. But they can still very well be applied in all those cases where the first would lead to insignificant results (empty or poor intersections).

16. *ADAPTATION OF CLASSIFICATION TO OTHER MEASURES*

We shall only very briefly mention the problem of measuring the degree of adaptation of classifications either to similarity measures or to distances. The adaptation to distances upon the set S has been thoroughly studied by Luscewska; the adaptation to similarity measures has been studied more practically and less generally by Fernandez de la Vega. A similarity measure is simply a weak distance : it has the first two properties of a distance (any object has maximal similarity to itself, just as it has minimal distance, and the degree of similarity is symmetrical - but the triangular property, denoting transitivity is not present).

We only wish to suggest that the problem of classification with respect to similarity indices reduces to the problem of decomposition of graphs. Indeed we can consider the elements of the set S as nodes on a graph. The nodes are linked by means of vertices if their degree of similarity is not zero. The degree of similarity can have various values so that we have valued graphs. Moreover the similarity index can be calculated with reference to a multiplicity of points of view : so that we have multi-graphs, where two nodes can be connected by many different types of vertices. In the book "Structural Methods" (Harary) methods of decomposition of bivalued simple graphs into connected subgraphs are studied. These methods when generalised for many valued multi-graphs give the solution of the problem of the classification of a set with reference to similarity indices.
We cannot leave this topic without reminding the reader that, in our introduction, we justified the necessity of classification by referring to the needs of information storage and retrieval.

Computer development has shown that the best organisation of programs is the one closest to the working of the human mind. This implies an important consequence : *the problem of the organisation of an information storing and retrieval system is the same as the problem of the organisation of memory.* Here library scientists, psychologists, neurologists and logicians should meet.

We consider our earlier paragraphs as a study of potential use to those who wish to analyse human memory.

If this is the case however, we have to mention the following possibility : it could be that our *classification systems are adapted to associative memories.* To be more precise : let the information units be stored as references (not of one, but of n types) to each other. Formally : every unit would be a sequence of n types of pairs (or of n-adds). We could study the degree of adaptation of a partition to such associative structures (taking into account that the link is neither reflexive, nor symmetrical, nor transitive, nor necessarily anti-reflexive, anti-symmetrical or anti-transitive). We could then say that P_1 is closer to the association set A than P_2 if the points referred to by members of the same classes in P_1 are more often the same than in P_2 (again averages, maxima or minima, or functions of all these parameters could be chosen to measure).

A final remark should be added opening up a new field of research. We owe it to a discussion intervention of Dr. Louchard.

We should not only study the adaptation of a classification tree to n relations, partitions or other pre-classifications of the same field.

The information stored has to be used. Search-programs have to look for a given type of information stored away in certain classes. A program is an algorithm. One should have *trees optimally adapted to algorithms.* The concept here introduced is of the same type of the ones studied in this paper, but it needs new solutions. We can only consider a proposal : a tree T_1 is more adapted to an algorithm A than a T_2 if the average length of the path needed to reach the end of A on T_1 is less. This length will in general depend upon the number of cyclic motions needed after a useless move in the search. Minimalizing the length of these cycles would imply Yngre's lop-sided trees.

The following drawing would make this clear. Let the information be stored in G, and let exploration always begin to the left.

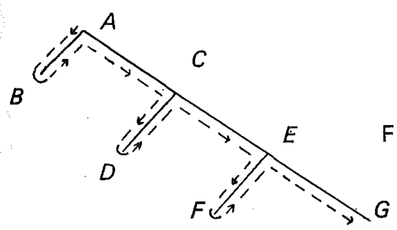


Fig. 1

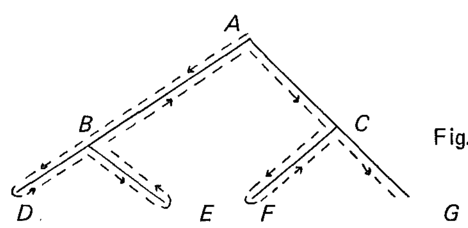


Fig. 2

The path is much longer in (2). One could naturally have the same result with left-branching. The machine realizing the search must be a pushdown because every motion must be controlled both by the earlier position and the present position. This remark is however only a first introduction to the study of the adaptation of classifications to algorithms. *It shows however why irregular trees (where not all branches reach the bottom row) are preferred to regular trees.*

The various problems discussed in the earlier paragraphs have the property of being genuine classification problems, in this sense that the standards of reference determining the degree of naturality can be based upon quantitative or qualitative data, and that the number of classes both in the classifications and in the standards of reference is free and is in general unknown for the classification sought for. The problems of multivariate statistics are specifications and special cases of the ones we had to study here.

17. *GENERAL CONDITIONS OF ADEQUACY FOR ADAPTATION MEASURES*

And now finally we come to our most important and also most difficult task : the attempt to determine certain very general axioms for the function that has in its domain the set of possible classifications of S, and that has in its codomain either the set of possible classifications, or the set of orders, or the set of similarity indices, or the set of distances on S or the set of association link, or the set of algorithms on S. The axioms should be such that intuitively the function F is a measure of the degree of closeness of the elements of the domain to the elements of the codomain. As far as we know this task has not yet been undertaken before.

We propose the following certainly very incomplete and insufficient list of axioms and add certain suggestions as to the direction in which they can be enlarged.

A) Let the binary function F be defined upon the set of all sets, and upon the set of all classifications (it is obvious that we need some theory of types or some version of category theory to avoid the paradoxes).

1. Let for any pair, the first element of which is a classification and the second element of which is a set, the function F project this pair upon an element of an incompletely ordered set (a stronger postulate would be : let it project the pair upon a completely ordered set).
2. Let the function F be on all such pairs at least partially recursive (other postulates could be : everywhere recursive, or primitive recursive).
3. Let the function be univalued : giving for every argument only one value.
4. Let there be as many pairs $<$ Clf, S $>$ as possible such that the function has a unique maximum (minimally : at least one pair, or at least an infinite number of pairs on which the function has a unique maximum).
5. If we take Clf as invariant and if we enlarge successively the set S upon which the Clf is defined, let a) F(Clf, S_1) remain equal to the function value for the classification upon the same set considered as subset of larger sets, and
   let b) there be refinements of Clf for which F gives the same values on the successive enlargements of S as Clf did on S.
6. If we take S as invariant then there are for given values of F, classifications that are refinements of the given classification, that have higher F values upon S.
   Comment : while the first four postulates are very general and rather innocent, the 5th and 6th postulates are extremely conservative : they presuppose that given classifications can, without radical upheaval be preserved for larger regions or be improved by simple refinements. This conservative stand-point can only be defended by a general postulate on the relative continuity of scientific research. The continuity can however not be absolute but only relative. For this very reason we must add to 5 and 6, another also *very specific* postulate about classifications.
7. Definition : any given set S defines a set of associated sets. These associated sets are of the following type :

a) the $sets_1$ of structures of the elements of S, decomposed into parts according to a certain decomposition procedure;
b) the $sets_2$ of structures of wholes, the parts of which are the elements of S;
c) the set of structures of wholes, the parts of which are sequences of elements taken in S, S_1, or S_2, or S^n;
d) the set of sets of S that stand to each other in a genetic relation (n elements of one subset S_1, producing one element of a subset S_2). Our postulate 7 will be of the following type : for any Clf and any S if F(Clf, S) cannot rise higher than a certain value v, then there is at least one Clf of at least one associated set, such that the value of F for that associated set is higher than v.

The justification of this last postulate is both historical and systematic. Historically we see that in very different domains morphological classifications (classifications upon associated sets 1 and 2), ecological classifications (on associated set 3) and genetic classifications (on associated set 4) appear regularly when natural classifications are sought. Systematically we remark that if the degree of naturality of the classification of a given set cannot rise high enough to satisfy our requirements, we must substitute for S other sets S′ that, if the step is not completely arbitrary, are functions of S. The modifications of S that are closest to S are, in the order to be given

a) the subdivision of the elements of S into parts
b) the fusion of elements of S into wholes
c) quotient sets upon S that stand in relationships that have some degree of homomorphism to the structure of a classification tree
d) cartesian products of S with arbitrary other sets.

B) In Apostel, 1963 morphological and genetic classifications have been studied formally and the relevant litterature has been mentioned. Here we introduce them in a manner less completely ad hoc by means of natural postulates upon the measure of the degree of naturality. A complete treatment of the subject would however demand that we show that, for a given set S and a given set of orders or classifications of S, if a certain degree of naturality cannot be reached for a summarizing classification, the best method to use to reach the aims of classification is to replace the set S by its associated sets in the order mentioned (and the meaning of "best method" would be here : the simplest algorithm giving the highest degree of naturality is to be found by applying the existent naturality measures to associated sets). We cannot pursue this topic here because we should have to introduce measures for the complexities of algorithms, a topic that is now under intensive study but that cannot be referred to in the present context.
We do not claim that the axioms for F are in any sense complete. We think that we should show in the future that the measures proposed satisfy the axioms mentioned. This also is not yet done. It is the main problem in this field.

18. *CLASSIFICATIONS ADAPTED TO LAWS*

Finally we have to remark that a definition of natural classification we did not yet mention is the following one "a classification is more natural than another one if more important laws true for the classes defined by the classification". The importance of a law is measured by the number of other laws of given degree of importance it implies. To come closer to practice, we should perhaps prefer to say "A Clf_1 is more natural than a Clf_2 if more laws implying with higher probability other laws of given degrees of confirmation are more highly confirmed for the classes of the first Clf than for those of the second". If the number of laws is finite, if their formulation and their degrees of confirmation are known we can apply immediately this definition, to obtain a partition (if we have resolved the difficulty of comparing the justification of a class figuring in a large quantity of laws either as antecedent or consequent, when these laws having few other laws either as their inductive or as their deductive consequents, with the justification of a class present in a smaller number of laws either as antecedent or consequent, with these laws being however more important in the sense of implying more consequences). In order to use a system of laws to come to a given classification, we have to define the order of an ordered and embedded sequence of partitions starting from this set of laws. To define this order, we can precisely use the difficulty mentioned. Indeed this difficulty derives simply from the fact that three independent parameters interfere : the number of laws referring to given classes; their degree of confirmation, and their deductive or inductive order of implication. We can use in order to define partitions some of these parameters, and use for defining the order of the partitions other parameters (for instance, it would seem natural that in the later rows of a Clf matrix, strongly confirmed laws with a smaller number of implicate laws are used and in the higher rows of the matrix more weakly confirmed laws with a large number of implicates). We must however leave the topic with these rather schematic remarks, the concepts of law and confirmation are too complex to be dealt with here.

19. *INTENSIONAL PERSPECTIVES*

Until now we did adopt a completely extensional point of view. We did not consider the classification criteria, their definitions or their meanings. This purely extensional theory of classifications is a natural outgrowth of the purely extensional theory of classes. The observer of present day logic will however immediately remark that intensional logics (and also intensional logics of classes) are more and more systematically introduced. If we consider the literature of library classifications, (for instance the facet classification of Vickery), it is certain that a general theory of Ranganathan and Vickery's efforts can only be formalized by means of a study of the intensional formal relations between the meanings of the class definitions. It is also certain that a more practical definition of distances or similarity indices on sets has to refer to the form and logical type of the predicates observed in the region (for instance : if we measure similarity by the number of shared one place predicates of lowest logical type, the problem is much simpler to solve than if we measure similarity by means of the number of shared predicates having various different logical types and various different numbers of arguments. The general problem of natural similarity measures for this more complex logical situation, has to be studied in interrelation with the form of definition of sets, the criteria of the partitions, the intensional and formal relations between the meanings of these criteria. This intensional aspect, much studied in the tradition and of great importance for library classification, has been informally studied on the most general level in Apostel 1963.

20. *THE GROWTH OF CLASSIFICATIONS*

A last topic, essential in the study of classifications, only to be mentioned here, is the genetic and historic study of classifications. Classifications can grow from the bottom (the individuals are known, the total set is reached through successive rows of partitions) or from the top (the superset is known, and various partitions reach the bottom). Most frequently both bottom and top will be partially known and by growing toward each other will receive their various final delimitations. Classifications can grow from one row to n row classifications, from low to high degree of naturality, from adaptation to other classification sets (larger and larger ones). Piaget and Bruner have studied in the adult and in the child the growth of classifications. The formal instrument for their undertaking is the theory we are trying to develop here.

21. *LIBRARY-SCIENCE AS A SPECIAL CASE OF OUR ADAPTATION PROBLEM*

The problem of library classification is fundamentally similar to the problem of memory storage.
Books or articles are complex information bearers that have to be stored at certain places (similar to addresses) of a 3 dimensional structure (location).
Let us call B the set of complex information units, a finite set of finite sets. Upon this set B a multidimensional order is defined (store-structure).
Let us call U the set of potential users of B. Each member of U is characterized by a set of questions ? $R(X_iX_j)$. The questions may be n-adic. These questions may be said to presuppose n classifications of human knowledge. All types of cataloguing and indexing (automatic or non automatic) can be seen as attempts to represent the information stored in such a way that :

a. it either intersects with all or many user-classifications, or
b. that operations performed on user, or on supplyer classifications (for instance : forming unions, or intersections, or complements of the classes in one or both) will make both classifications meet.

How can we optimalize the classification system ?
1. An estimate has to be made of the set U of potential users, and 2. of the set of their questions expressed in given terms. 3. of the content of the messages, by means of most frequent terms, or n-ads of terms or of privileged terms or termsequences. 4. A translation procedure has to be provided for either B-descriptions in terms of QU-descriptions, or of QU-descriptions in terms of B-descriptions. 5. the most frequently used information units must be a) present in many replications b) at short distances (analogy to Zipf's law for bibliotheconomy). 6. The problem is fundamentally statistical and reduces to constructing n partitions of the B-units in such a way that the probability that any UQ-partition has a

non zero intersection with the B-partition, is not too low (before or after the application of the search transformations).

The author hopes that this remark will make it clear that his apparently very abstract developments are simply giving the algebraical core of the library problem (neglecting on purpose the essential statistical and feedback features present here).

In conclusion, we can point out that a new chapter of logic, the formal study of the relations of classification systems, orders and measures has now reached independent existence but must still receive its intensional and dynamic counterpart.. The very important study of information retrieval, rightly but severely criticised by Bar-Hillel is only possible as an outgrowth and application of the more general discipline we tried to sketch here. Controversies between genetical and morphological classifications can only be solved in its framework.

BIBLIOGRAPHY

L. APOSTEL, *Le problème formel des classifications empiriques*, in La Classification dans les sciences, Gembloux, J. Duculot - Centre National Belge de Recherches de Logique, pp. 157-230.

Y. BAR-HILLEL, *A logician's reaction to Recent theorizing on Information Search Systems,* American Documentation, 8, pp. 103-122.

J.P. BENZECRI, *Sur les algorithmes de classification,* Paris, Cours I.S.U.P., 1965.
J.P. BENZECRI, *Problèmes et Méthodes de la taxinomie,* Paris, Cours I.S.U.P., 1965-1966.

J. BRUNER, *A study of Thinking* (in collaboration with J.S. Goodnow and G.A. Austin), New York, Wiley, 1957, pp. 330.

Zygmunt DOBROWOLSKI, *Etude sur la construction des systèmes de classification,* Paris, Gauthier-Villars, 1964.

D.J. HILLMAN, *Mathematical Classification Techniques for Non-static Document collections, with Particular Reference to the Problem of Relevance*, in Classification Research, Copenhagen, Munksgaard, P. Atherton, 1965, pp. 177-210.

S. LUSCEWSKA-ROMAHNOWA, *Classification as a Kind of distance Function*, Studia Logica, XII, 1961, pp. 41-81.
S. LUSCEWSKA-ROMAHNOWA and TADEUSZ BATOG, *A generalized Theory of classifications*, I, Studia Logica, XVI, 1965, pp. 53-70, II, ibidem XVII, 1965, pp. 7-24. The second series of papers is a generalization for infinite trees of the distance approach developed in the first paper. The problem is *not* trivial; but perhaps the application to actual classification trees is not as immediate.

A.F. PARKER-RHODES and R.M. NEEDHAM, *The theory of clumps,* Cambridge Language research Unit, Report M.L. 126, Febr. 1960.

J. PIAGET (avec BARBEL-INHELDER), *La genèse de structures logiques élémentaires : classifications et sériations,* Neuchatel, Delachaux et Niestle, 1959, pp. 295.

R.R. SOKOEL and P.A.H. SNEATH, *Principles of Numerical Taxonomy,* San Francisco and London, W.H. Freeman and C˘, 1963, pp. 359.

W. FERNANDEZ de la VEGA, *Quelques Aspects du problème de la détermination automatique des classifications,* Quality and Quantity, Bologna, IV, 1, June, 1970.

B.C. VICKERY, *Classification and Indexing in science,* 2d ed. London, Butterworth, New York, Academic Press.

L.A. ZADEH : 1) *Fuzzy Sets,* Information and Control, 8, 1965, pp. 338-353.
2) *Fuzzy Sets and Systems,* Proceedings of a symposium on systems theory, Polytechnic Institute of Brooklyn, pp. 29-37.
3) *Shadows of Fuzzy Sets,* Problems Trans. Information, 2, pp. 37-44.
4) *Fuzzy Algorithms,* Information and Control, 12, 1968, pp. 94-102.

SUMMARY

The aim of this paper is to show the truth of a few assertions :

1. Whenever we are seriously considering classifications, we plan to have "natural classifications" (i.e. : classifications such that when two objects *a* and *b* belong to the same class, we know that a and b have many other things in common).

2. A *natural* definition of *"natural"* classification (in purely extensional terms) can be given. A successful attempt due to Mrs Luscewska-Romahnowa can be generalised, using ideas present in Apostel, 1963).

3. Algorithms can be constructed finding for a given domain the most natural classifications. But these algorithms are not in general able to find a *unique* solution : various sub-optimal solutions can be in many cases the only answer.

4. a) Often the most efficient classification is an irregular one (see for the definition of this word the following text).
   b) Often the most efficient classification is a serie of different irregular classifications of the same domains to be switched on by given signals.

NOTES

(1) All names of authors refer to the bibliography at the end of the paper.

(2) The reader will remember the familiar distinction between the extension of a term and its intension or comprehension. The second concept is identical to the meaning, the first refers to objects name-able by that term.

(3) SL means : segment of a line.

(4) We call here "interval" the union of all sets lying between 2 endpoints, including the one, included in the other.