WADD-TN-61-29 March 1961

Hierarchical Grouping to Maximize Payoff

By
Joe H. Ward, Jr.

PERSONNEL LABORATORY
WRIGHT AIR DEVELOPMENT DIVISION
AIR RESEARCH and DEVELOPMENT COMMAND
UNITED STATES AIR FORCE
LACKLAND AIR FORCE BASE, TEXAS

Fred E. Boldrege, Col USAF
Chief

A. Carp
Technical Director

Personnel Laboratory
Wright Air Development Division (ARDC)

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever: and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

.....

Qualified requestors may obtain copies of this report from the Armed Services Technical Information Agency, Documents Serivce Center, Arlington 12, Virginia. Department of Defense contractors may obtain unclassified documents from ASTIA on request by stating their official need and citing the Defense contract involved.

.....

Copies of WADD Technical Reports and Technical Notes should not be returned to the Wright Air Development Division unless such return is required by security considerations, contractual obligations, or notice on a specific document.

Since the ordering of elements is arbitrary and does not affect the definition of the set, the union of two sets is commutative. We may state, for example,

$$[S(1, n)] \cup [S(2, n)] = [S(2, n)] \cup [S(1, n)]$$

$$|e_1, e_2| = |e_2, e_1|.$$

We also can use such statements to express the union of all possible pairs of the n mutually exclusive sets $S(i, n)$, $i = 1, 2, \ldots, n$. These $n(n-1)/2$ possible unions may be designated as follows:

$$[S(1, n)] \cup [S(2, n)] = |e_1, e_2|$$

$$[S(1, n)] \cup [S(3, n)] = |e_1, e_3|$$

$$\vdots$$

$$[S(i, n)] \cup [S(j, n)] = |e_i, e_j|$$

$$\vdots$$

$$[S(n-1, n)] \cup [S(n, n)] = |e_{n-1}, e_n|.$$

These $n(n-1)/2$ possible unions also may be designated more simply by

$$[S(i, n)] \cup [S(j, n)],$$

$$(i = 1, 2, \ldots, n-1; j = i+1, \ldots, n).$$

Value of objective function associated with new set. The values of the objectives functions associated with each of the $n(n-1)/2$ possible unions of sets may be designated as follows:

Value of Objective Function	Union with which Value of Objective Function Is Associated
$Z[1, 2, n-1]$	$[S(1, n)] \cup [S(2, n)]$
$Z[1, 3, n-1]$	$[S(1, n)] \cup [S(3, n)]$
$\vdots$	$\vdots$
$Z[i, j, n-1]$	$[S(i, n)] \cup [S(j, n)]$
$\vdots$	$\vdots$
$Z[n-1, n, n-1]$	$[S(n-1, n)] \cup [S(n, n)]$

These $n(n-1)/2$ values of objective functions also may be designated more simply as

$Z[i, j, n-1]$ associated with $[S(i, n)] \cup [S(j, n)]$,

$(i = 1, 2, \ldots, n-1; j = i+1, \ldots, n)$.

ERRATA

Ward, J.H., Jr. *Hierarchical grouping to maximize payoff.* Personnel Laboratory, Wright Air Development Division, March 1961. (WADD-TN-61-29)

p. 5 This page is underprinted and many of the symbols are illegible. Please substitute the reprinting on the reverse of this sheet.

p. 9 Last line within brackets, near bottom of page:

For $i = q_{n-1}, q_{n-2}$

Read $i \neq q_{n-1}, q_{n-2}$

p. 11 Middle of page, right member of equation:

For $= |e_{m_1}, \ldots, e_{m_s}, \ldots, e_{m_t}|,$

Read $= |e_{m_1}, \ldots, e_{m_s}, \ldots, e_{m_t}|,$

Div. 15/2, 23/1, 28/4

Wright Air Development Division. Personnel Laboratory, Lackland Air Force Base, Texas. HIERARCHICAL GROUPING TO MAXIMIZE PAY-OFF, by Joe H. Ward, Jr. March 1961. v + 18 p. (Project 7734; Task 17016) (WADD-TN-61-29)
Unclassified report

This report describes mathematically a general procedure for forming hierarchical groups of mutually exclusive sets in a manner which yields an optimum value for the functional relation, or objective function, that reflects the criterion chosen by the investigator. The number of groups to be formed need not be specified in advance. Given k sets, this technique permits their reduction to $k-1$ mutually exclusive sets by considering the union of all possible pairs that can

(over)

UNCLASSIFIED

Div. 15/2, 23/1, 28/4

Wright Air Development Division. Personnel Laboratory, Lackland Air Force Base, Texas. HIERARCHICAL GROUPING TO MAXIMIZE PAY-OFF, by Joe H. Ward, Jr. March 1961. v + 18 p. (Project 7734; Task 17016) (WADD-TN-61-29)
Unclassified report

This report describes mathematically a general procedure for forming hierarchical groups of mutually exclusive sets in a manner which yields an optimum value for the functional relation, or objective function, that reflects the criterion chosen by the investigator. The number of groups to be formed need not be specified in advance. Given k sets, this technique permits their reduction to $k-1$ mutually exclusive sets by considering the union of all possible pairs that can

(over)

UNCLASSIFIED

UNCLASSIFIED

be formed and the selection of that union which has the highest payoff value with respect to the criterion chosen. This procedure can be repeated until only one set remains. Hence decisions on the number of groups to be used can be based on a knowledge of the "costs" of grouping at each stage in the entire hierarchical structure. A computer flowchart and a numerical example of the grouping procedure are provided. An Appendix shows how to determine the number of possible ways of forming groups and the number of distinguishable unions possible.

be formed and the selection of that union which has the highest payoff value with respect to the criterion chosen. This procedure can be repeated until only one set remains. Hence decisions on the number of groups to be used can be based on a knowledge of the "costs" of grouping at each stage in the entire hierarchical structure. A computer flowchart and a numerical example of the grouping procedure are provided. An Appendix shows how to determine the number of possible ways of forming groups and the number of distinguishable unions possible.

UNCLASSIFIED

UNCLASSIFIED

Div. 15/2, 23/1, 28/4

Wright Air Development Division. Personnel Laboratory, Lackland Air Force Base, Texas. HIERARCHICAL GROUPING TO MAXIMIZE PAY-OFF, by Joe H. Ward, Jr. March 1961. v + 18 p. (Project 7734; Task 17016) (WADD-TN-61-29)
Unclassified report

This report describes mathematically a general procedure for forming hierarchical groups of mutually exclusive sets in a manner which yields an optimum value for the functional relation, or objective function, that reflects the criterion chosen by the investigator. The number of groups to be formed need not be specified in advance. Given k sets, this technique permits their reduction to $k-1$ mutually exclusive sets by considering the union of all possible pairs that can

(over)

UNCLASSIFIED

UNCLASSIFIED

be formed and the selection of that union which has the highest payoff value with respect to the criterion chosen. This procedure can be repeated until only one set remains. Hence decisions on the number of groups to be used can be based on a knowledge of the "costs" of grouping at each stage in the entire hierarchical structure. A computer flowchart and a numerical example of the grouping procedure are provided. An Appendix shows how to determine the number of possible ways of forming groups and the number of distinguishable unions possible.

UNCLASSIFIED

Div. 15/2, 23/1, 28/4

Wright Air Development Division. Personnel Laboratory, Lackland Air Force Base, Texas. HIERARCHICAL GROUPING TO MAXIMIZE PAY-OFF, by Joe H. Ward, Jr. March 1961. v + 18 p. (Project 7734; Task 17016) (WADD-TN-61-29)
Unclassified report

This report describes mathematically a general procedure for forming hierarchical groups of mutually exclusive sets in a manner which yields an optimum value for the functional relation, or objective function, that reflects the criterion chosen by the investigator. The number of groups to be formed need not be specified in advance. Given k sets, this technique permits their reduction to $k-1$ mutually exclusive sets by considering the union of all possible pairs that can

(over)

UNCLASSIFIED

UNCLASSIFIED

be formed and the selection of that union which has the highest payoff value with respect to the criterion chosen. This procedure can be repeated until only one set remains. Hence decisions on the number of groups to be used can be based on a knowledge of the "costs" of grouping at each stage in the entire hierarchical structure. A computer flowchart and a numerical example of the grouping procedure are provided. An Appendix shows how to determine the number of possible ways of forming groups and the number of distinguishable unions possible.

UNCLASSIFIED

WADD-TN-61-29
March 1961

HIERARCHICAL GROUPING TO MAXIMIZE PAYOFF

By
Joe H. Ward, Jr.

Project 7734; Task 17016

Personnel Laboratory
WRIGHT AIR DEVELOPMENT DIVISION
AIR RESEARCH AND DEVELOPMENT COMMAND
UNITED STATES AIR FORCE
Lackland Air Force Base, Texas

ABSTRACT

This report describes mathematically a general procedure for forming hierarchical groups of mutually exclusive sets in a manner which yields an optimum value for the functional relation, or objective function, that reflects the criterion chosen by the investigator. The number of groups to be formed need not be specified in advance. Given k sets, this technique permits their reduction to $k-1$ mutually exclusive sets by considering the union of all possible pairs that can be formed and the selection of that union which has the highest payoff value with respect to the criterion chosen. This procedure can be repeated until only one set remains. Hence decisions on the number of groups to be used can be based on a knowledge of the "costs" of grouping at each stage in the entire hierarchical structure. A computer flowchart and a numerical example of the grouping procedure are provided. An Appendix shows how to determine the number of possible ways of forming groups and the number of distinguishable unions possible.

PREFACE

In many situations it is desirable to group large numbers of persons, jobs, or objects into smaller numbers of mutually exclusive classes in which the members are as much alike as possible with respect to some criterion. When the grouping is done in a manner that establishes a taxonomy, or system, of mutually exclusive clusters wherein each larger unit is a combination of subgroups, these clusters are called "hierarchical groups." Hierarchical grouping is particularly useful for classification purposes. In the past it has been used to classify plants and animals with respect to genetic background; also to organize and catalog materials, such as library holdings, so as to facilitate the storage and retrieval of information. Similarly, hierarchical groupings of persons and of jobs make it easier to consider all the information available for purposes of personnel administration. Until now, however, hierarchical grouping has usually been accomplished by armchair rather than computer techniques. Hence its use has been limited, optimally homogeneous groups have not been formed, and the loss resulting from the grouping has not been quantified.

This report describes mathematically a general procedure for forming hierarchical groups of mutually exclusive sets in a manner which yields an optimum value for the functional relation, or objective function, that reflects the criterion chosen by the investigator. The number of groups to be formed need not be specified in advance. Given k sets, this technique permits their reduction to $k-1$ mutually exclusive sets by considering the union of all possible $k(k-1)/2$ pairs that can be formed and the selection of that union which has the highest payoff value with respect to the criterion chosen. This procedure can be repeated until only one set remains. Hence decisions on the number of groups to be used can be based on a knowledge of the costs of grouping at each stage in the entire hierarchical structure. A computer flowchart and a numerical example of the grouping procedure are provided. The Appendix shows how to determine the number of possible ways of forming groups and the number of distinguishable unions possible.

TABLE OF CONTENTS

LIST OF FIGURES

HIERARCHICAL GROUPING TO MAXIMIZE PAYOFF[1]

INTRODUCTION

Situations frequently arise in which it is desired to combine or group mutually exclusive sets, or collections of well-defined objects, so as to increase the efficiency with which they can be considered. The grouping, or union, of collections of objects makes it easier to comprehend large collections and often increases the efficiency of practical operations. However, grouping ordinarily results in some loss that may be quantified in a "value-reflecting" number associated with the grouping.

Suppose, for example, a collection, or set, of ratings has been obtained for 10 individuals. Let us say these values are {2, 6, 5, 6, 2, 2, 2, 0, 0, 0}. A common practice is to use the mean value of the observations to represent all scores rather than to consider the individual scores. Here, for instance, the mean, 2.5, would represent the 10 scores. The loss in information resulting from treating the individual scores as one group is indicated by a value-reflecting number called the "error sum of squares" (ESS).

OBJECTIVE FUNCTION

The error sum of squares is given by the functional relation,

$$\text{ESS} = \sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2,$$

where x_i is the score of the ith individual. The error sum of squares for the example is given by

$$\text{ESS (one group)} = \sum_{i=1}^{10} x_i^2 - \frac{1}{10}\left(\sum_{i=1}^{10} x_i\right)^2 = 113 - 62.5 = 50.5.$$

The functional relation used to obtain a value-reflecting number will be referred to as an "objective function." In this example, the objective function yields a value that represents the information lost when the 10 individual scores are treated as a single group. In general, the objective function is any type of value-reflecting number. As we have seen, it may be the cost associated with a particular grouping. It might refer, for example, to the crosstraining time expected to result from grouping jobs into mutually exclusive categories. Again, in information-theory terminology, the objective function might refer to the amount of information still available after grouping.

HIERARCHICAL GROUPING

If we are willing to classify the 10 individuals on the basis of their scores into four groups,

$$(0, 0, 0),\ (2, 2, 2, 2),\ (5),\ (6, 6),$$

we can use a different value – the mean of each group – to represent each of the four sets of scores

---

[1] Manuscript released by the author for publication as a WADD Technical Note in March 1961.

without any loss of information. In the example, the objective function ESS (four groups) would be computed as:

$$\text{ESS (four groups)} = \text{ESS (Group 1)} + \text{ESS (Group 2)} + \text{ESS (Group 3)} + \text{ESS (Group 4)}$$

$$\text{ESS (four groups)} = \left[\sum_{i=1}^{3} x_i^2 - \frac{1}{3}\left(\sum_{i=1}^{3} x_i \right)^2 \right] \quad \text{(Group 1 scores)}$$

$$+ \left[\sum_{i=1}^{4} x_i^2 - \frac{1}{4}\left(\sum_{i=1}^{4} x_i \right)^2 \right] \quad \text{(Group 2 scores)}$$

$$+ \left[\sum_{i=1}^{1} x_i^2 - \frac{1}{1}\left(\sum_{i=1}^{1} x_i \right)^2 \right] \quad \text{(Group 3 scores)}$$

$$+ \left[\sum_{i=1}^{2} x_i^2 - \frac{1}{2}\left(\sum_{i=1}^{2} x_i \right)^2 \right] \quad \text{(Group 4 scores)}$$

$$\text{ESS (four groups)} = 0 + 0 + 0 + 0 = 0.$$

From this we see that when all the scores are combined in one group, considerable information is lost; but when the scores are considered in four mutually exclusive groups no information is lost, as indicated by the value of the objective function.

APPROACH TO HIERARCHICAL GROUPING

The grouping procedure described here is based on the premise that the greatest amount of information, as indicated by the objective function, is available when the array of n objects is ungrouped. Hence the grouping process starts with these n objects, which are termed groups or sets although they contain only one element. The first step in grouping is to select which two of these n sets should be combined in order to reduce by one the number of groups, while producing the least impairment of the optimum value of the objective function. The $n-1$ resulting sets then are examined to identify which pair of groups should be merged in order to secure the optimum value of the objective function for $n-2$ groups. This procedure can be repeated, if desired, until all the original n objects are in a single group. Since the number of sets is systematically reduced $[n, n-1, \ldots, 1]$, the process will be referred to as "hierarchical grouping." The resulting mutually exclusive groups, or sets, are called "hierarchical groups" inasmuch as each is a combination, or union, of subgroups.

Hierarchical groups formed in this way are particularly useful for classification purposes (*1*, *2*, *3*, *4*, 5). Such groups may be used, for example, to establish taxonomies of plants and animals with respect to genetic background. They also permit the organization and cataloging of materials, e.g., library documents, in a manner that facilitates the storage and retrieval of

information. Similarly, hierarchical groupings of jobs may be helpful in identifying job "types" and "subtypes."

We shall not attempt to catalog the various applications of the hierarchical grouping process. However, the usefulness of this method is not restricted to the familiar types of classification problems. For example, in practical prediction problems it may be desirable to form hierarchical groups of prediction equations so as to select a smaller number of equations for use in place of a large number of different equations. A current study[2] is determining loss against several cost criteria when a smaller number of predictive equations, so selected, are substituted for 60 different regression equations developed for predicting success in Air Force Technical Schools. Although it may be desirable sometimes to disregard the hierarchical arrangement in finding the "best" partitioning into groups for predictive purposes, a hierarchical grouping is preferred in many practical situations. Even when a nonhierarchical grouping is sought, the hierarchical approach described in this article will yield a good solution although it may be one that does not optimize the objective function.

In some situations the desired number of groups can be specified in advance; in others, this is difficult to do.[3] In either event, it may be feasible to reflect in the objective function whatever "value" is associated with the number of groups. If this is done, the optimum value of the objective function may be attained at some stage when the number of groups is greater than one. This situation arises, for example, when it is desirable to obtain a hierarchical grouping of regression equations and the objective function is the probability associated with the test of the hypothesis that the regression equations within the groups are homogeneous. In this case the goal is to identify and use that particular grouping for which the test statistic has the largest probability of occurring by chance under the hypothesis of equal regression coefficients. The number of groups that would satisfy this objective function is quite likely to be greater than one group.

A problem often linked with the grouping problem is that of assigning or classifying new objects within accepted groupings. This problem may be resolved by developing discriminant predictive equations from the data available on the objects that have been grouped. However, this problem, as well as those involved in selecting objective functions and deciding upon the appropriate number of groups to be utilized, will not be discussed here. The purpose of this article is to describe a hierarchical grouping procedure which is believed to be of general interest since it has many applications.

HIERARCHICAL GROUPING PROCEDURE

UNIVERSAL SET (U) AND ITS SUBSETS

Universal set, U. Consider a collection of n well-defined objects, symbols, or persons. This collection will be called a *set.* The n members of the set will be referred to as *elements* and designated $\{e_1, e_2, e_3, \ldots, e_n\}$. The order in which elements appear in a set is arbitrary;

---

[2] Directed by Dr. Raymond E. Christal and Dr. Ernest C. Tupes, Personnel Laboratory.

[3] When it is desired to assign objects to k, a specified number of groups — without regard for the subdivisions of these groups — the goal is only to identify the k groups that optimize the objective function. If the objective function is a linear form, the problem can be formulated as a linear programming problem. If the objective function is nonlinear in form, the problem can be formulated as a nonlinear programming problem (2); however, certain computational difficulties must be considered.

they are numbered in sequence only as a means of convenient identification. The set of n elements to which we shall restrict our attention will be referred to as the "universal" set, which will be denoted

$$U = \{e_1, e_2, \ldots, e_n\}.$$

Subsets of U. If each element of a set A is also an element of a set B, we shall say that "A is a *subset* of B."[4]

Now let us define the following n sets, each of which consists of a single element and each of which is a subset of the universal set, U:

$$S(1, n) = \{e_1\}$$
$$S(2, n) = \{e_2\}$$
$$\vdots$$
$$S(i, n) = \{e_i\}$$
$$\vdots$$
$$S(n, n) = \{e_n\}.$$

Within the parentheses, the left-hand term is the number that identifies the set, and the right-hand term – n in these sets – shows the number of sets under consideration. It is important to observe that each of these n subsets of U has only one element, e.g., the set designated $S(i, n)$ contains a single element, e_i.

Note also that no two of these n subsets of U have the same element. Whenever two sets, designated C and D respectively, have no element in common, we shall say "the two sets, C and D, are *mutually exclusive*." Similarly, when a collection of sets, such as the n sets $S(1, n)$, $S(2, n)$, . . . , $S(n, n)$ defined here, have no elements in common, we shall describe them as mutually exclusive sets.

UNION OF SETS

Formation of new set by union. Let us now consider the formation of a new set by combining or grouping the elements of a pair of mutually exclusive sets. The resulting new set will be termed the *union* of the two sets.[5] For example, the set resulting from the union of $S(1, n) = \{e_1\}$ and $S(2, n) = \{e_2\}$ is a new set with two elements that will be denoted $\{e_1, e_2\}$. The union of the two sets will be represented as

$$[S(1, n)] \cup [S(2, n)] = \{e_1, e_2\}.$$

The symbol "∪" is read as "union," i.e., $S(1, n)$ *union* $S(2, n)$.

---

[4] If A is a subset of B and A does not include the entire set B, then A is called a "proper subset of B."

[5] The concept of the "union" of two sets is applicable to all sets, not just mutually exclusive sets.

Since the ordering of elements is arbitrary and does not affect the definition of the set, the union of two sets is commutative. We may state, for example,

$$|S(1, n)| \cup |S(2, n)| = |S(2, n)| \cup |S(1, n)|$$

$$|e_1, e_2| = |e_2, e_1|.$$

We also can use such statements to express the union of all possible pairs of the n mutually exclusive sets $S(i, n)$, $i = 1, 2, \ldots, n$. These $n(n-1)/2$ possible unions may be designated as follows:

$$|S(1, n)| \cup |S(2, n)| \quad |e_1, e_2|$$

$$|S(1, n)| \cup |S(3, n)| \quad |e_1, e_3|$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$|S(i, n)| \cup |S(j, n)| \quad |e_i, e_j|$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$|S(n-1, n)| \cup |S(n, n)| \quad |e_{n-1}, e_n|.$$

These $n(n-1)/2$ possible unions also may be designated more simply by

$$|S(i, n)| \cup |S(j, n)|,$$
$$(i = 1, 2, \ldots, n-1; \; j = i+1, \ldots, n).$$

Value of objective function associated with new set. The values of the objectives functions associated with each of the $n(n-1)/2$ possible unions of sets may be designated as follows:

Value of Objective Function	Union with which Value of Objective Function Is Associated
$Z\|1, 2, n-1\|$	$\|S(1, n)\| \cup \|S(2, n)\|$
$Z\|1, 3, n-1\|$	$\|S(1, n)\| \cup \|S(3, n)\|$
$\vdots$	$\vdots$
$Z\|i, j, n-1\|$	$\|S(i, n)\| \cup \|S(j, n)\|$
$\vdots$	$\vdots$
$Z\|n-1, n, n-1\|$	$\|S(n-1, n)\| \cup \|S(n, n)\|$

These $n(n-1)/2$ values of objective functions also may be designated more simply as

$$Z|i, j, n-1| \text{ associated with } |S(i, n)| \cup |S(j, n)|,$$
$$(i = 1, 2, \ldots, n-1; \; j = i+1, \ldots, n).$$

Within the bracketed expression for the value of an objective function, the left-hand and center terms identify the first and the second sets, respectively, of the pair combined in the union with which the value is associated. The term on the right – $n-1$ in this instance – denotes the number of groups that will result from the union of these two of the n existing sets.

HIERARCHICAL GROUPING CYCLE

SUBSETS $S(i, n)$

Selection of optimum union of subsets $S(i, n)$. As each of the $n(n-1)/2$ possible unions is considered in turn and the corresponding objective function evaluated, it can be hypothesized that the value of the objective function that results from this particular union of two sets is "equal to or better than" that for any preceding union that has been considered. If the identity of the "optimum" or "best" union is maintained throughout the sequence of comparisons, it is possible to select, from the $n(n-1)/2$ possible unions, the one union for which the value of the objective function is "equal to or better than" that of the other possible unions. This union, then, will be accepted as an optimal grouping when the number of sets is reduced from n to $n-1$.

Designation of optimum union and associated objective function. Once a union of two sets has been selected as the optimum union for going from n to $n-1$ sets because it results in the optimum value of the objective function, special designations will be given to the sets in this union. The set with the smaller identification number will be designated p_{n-1}; the set with the larger identification number will be designated q_{n-1}.

The new set resulting from the optimum union will be identified as p_{n-1}. This new set will be represented as

$$S(p_{n-1}, n-1) \quad |S(p_{n-1}, n)| \cup |S(q_{n-1}, n)|$$

$$|e_{p_{n-1}}, e_{q_{n-1}}|.$$

Within the parentheses, the left-hand term, p_{n-1}, is the identification of the new set resulting from the union; the right-hand term, $n-1$, indicates that $n-1$ sets now exist as a result of this union.

The value of the objective function corresponding to the optimum union will be designated

$$Z|p_{n-1}, q_{n-1}, n-1|.$$

As before, the left-hand and center terms refer to the identification numbers of the united sets. The term at the right, $n-1$, shows the number of sets remaining for consideration after the union of these two sets.

$S(i, n)$ has been selected, the resulting $n-1$ new subsets will be defined as follows:

$$
\begin{array}{lll}
S(1, n-1) = S(1, n) = |e_1| \\
S(2, n-1) = S(2, n) = |e_2| \\
\vdots \quad \vdots \quad \vdots \\
S(p_{n-1}, n-1) = |S(p_{n-1}, n)| \cup |S(q_{n-1}, n)| = |e_{p_{n-1}}, e_{q_{n-1}}| \\
\vdots \quad \vdots \quad \vdots \\
S(q_{n-1}-1, n-1) \quad S(q_{n-1}-1, n) \quad |e_{q_{n-1}}-1| \\
S(q_{n-1}+1, n-1) \quad S(q_{n-1}+1, n) \quad |e_{q_{n-1}}+1| \\
\vdots \quad \vdots \quad \vdots \\
S(n, n-1) \quad S(n, n) \quad |e_n|
\end{array}
$$

Or, in simpler notation, we have

$$
S(i, n-1) \quad S(i, n) \left[\begin{array}{ll} i & 1, 2, \ldots, n; \\ & i \neq p_{n-1} \\ & i \neq q_{n-1} \end{array} \right]
$$

Further, when $i = p_{n-1}$,

$$
S(p_{n-1}, n-1) \quad |S(p_{n-1}, n)| \cup |S(q_{n-1}, n)|.
$$

Here, within the parentheses, the left-hand term identifies the new subsets; and the right-hand term indicates the number of sets resulting from the union. Notice that the identification number q_{n-1} is no longer used in the identification of the $n-1$ new mutually exclusive sets. When an identification number, such as q_{n-1}, is not used, it will be referred to as "inactive."

Selection of optimum union of subsets S(i, n-1). A procedure analogous to that used to consider the $n(n-1)/2$ possible unions of the subsets $S(i, n)$, $i = 1, 2, \ldots, n$, and the values of their corresponding objective functions will be used to consider the $(n-1)(n-2)/2$ possible unions of the subsets $S(i, n-1)$, $i = 1, 2, \ldots, q_{n-1} - 1; q_{n-1} + 1, \ldots, n$. The values of the objective functions and the unions with which they are associated will be denoted as follows:

Value of Objective Function	Union with which Value of Objective Function Is Associated
$Z[1, 2, n-2]$	$[S(1, n-1)] \cup [S(2, n-1)]$
$Z[1, 3, n-2]$	$[S(1, n-1)] \cup [S(3, n-1)]$
.	.
.	.
.	.
$Z[i, j, n-2]$	$[S(i, n-1)] \cup [S(j, n-1)]$
.	.
.	.
.	.
$Z[n-1, n, n-2]$	$[S(n-1, n-1)] \cup [S(n, n-1)]$

The values of objective functions associated with these $(n-1)(n-2)/2$ possible unions of subsets may be given in a more explicit notation in this fashion:

$$Z[i, j, n-2] \text{ associated with } [S(i, n-1)] \cup [S(j, n-1)],$$

$$\begin{bmatrix} i = 1, 2, \ldots, q_{n-1} - 1, q_{n-1} + 1, \ldots, n-1; \\ j = i+1, i+2, \ldots, q_{n-1} - 1, q_{n-1} + 1, \ldots, n \end{bmatrix}.$$

As before, we shall hypothesize that each of the $(n-1)(n-2)/2$ unions of subsets $S(i, n-1)$ yields an objective function having an optimal value. When all comparisons have been made, an optimum union and its associated objective function will be selected and designated in a fashion analogous to that employed when the n sets were reduced to $n-1$ sets:

$$S(p_{n-2}, n-2) \quad [S(p_{n-2}, n-1)] \cup [S(q_{n-2}, n-1)]$$

and

$$Z[p_{n-2}, q_{n-2}, n-2], \ (p_{n-2} < q_{n-2}).$$

The identifications are similar to those used before, i.e., p_{n-2} is the identification number with the smaller numerical value, and q_{n-2} is that with the larger numerical value.

Identification of n-2 new, mutually exclusive subsets. **After the optimum union has been selected for subsets S(*i*, *n*-1), the resulting *n*-2 new, mutually exclusive subsets will be defined in the same manner as the *n*-1 subsets were defined[6]:**

$$
\begin{aligned}
S(1, n-2) &= S(1, n-1) &&= \{e_1\} \\
S(2, n-2) &= S(2, n-1) &&= \{e_2\} \\
&\vdots &&\vdots \\
S(p_{n-1}, n-2) &= S(p_{n-1}, n-1) &&= \{e_{p_{n-1}}, e_{q_{n-1}}\} \\
&\vdots &&\vdots \\
S(q_{n-1}-1, n-2) &= S(q_{n-1}-1, n-1) &&= \{e_{q_{n-1}-1}\} \\
S(q_{n-1}+1, n-2) &= S(q_{n-1}+1, n-1) &&= \{e_{q_{n-1}+1}\} \\
&\vdots &&\vdots \\
S(p_{n-2}, n-2) &= [S(p_{n-2}, n-1)] \cup [S(q_{n-2}, n-1)] &&= \{e_{p_{n-2}}, e_{q_{n-2}}\} \\
&\vdots &&\vdots \\
S(q_{n-2}-1, n-2) &= S(q_{n-2}-1, n-1) &&= \{e_{q_{n-2}-1}\} \\
S(q_{n-2}+1, n-2) &= S(q_{n-2}+1, n-1) &&= \{e_{q_{n-2}+1}\} \\
&\vdots &&\vdots \\
S(n, n-2) &= S(n, n-1) &&= \{e_n\}
\end{aligned}
$$

We may abbreviate these definitions, and state:

$$S(i, n-2) = S(i, n-1), \begin{bmatrix} i = 1, 2, \ldots, n; \\ i \neq p_{n-2} \\ i = q_{n-1}, q_{n-2} \end{bmatrix}$$

Further, when $i = p_{n-2}$,

$$S(p_{n-2}, n-2) = [S(p_{n-2}, n-1)] \cup [S(q_{n-2}, n-1)].$$

Notice that, at this stage in the grouping cycle, the two identification numbers, q_{n-1} and q_{n-2}, are no longer used. They are said to be inactive.

---

[6] The definition indicated here represents the situation in which the *n*-2 subsets contain either one or two elements. The grouping at this stage could result, of course, in a set which contains three elements.

SUBSETS $S(i, k)$

Notation for general case. **The grouping cycle described can be continued until all subsets have been united to form the universal set, U. At any phase in which k mutually exclusive subsets are under consideration, the objective function and the union with which it is associated would be expressed as**

$$Z[i, j, k-1] \text{ associated with } [S(i, k)] \cup [S(j, k)]$$

where

$$i = 1, 2, \ldots, n-1$$

$$i \neq q_{n-1}, q_{n-2}, \ldots, q_k$$

$$j = i+1, i+2, \ldots, n$$

$$j \neq q_{n-1}, q_{n-2}, \ldots, q_k.$$

The optimum union and the value of the corresponding objective function will be designated

$$S(p_{k-1}, k-1) = [S(p_{k-1}, k)] \cup [S(q_{k-1}, k)]$$

and

$$Z[p_{k-1}, q_{k-1}, k-1], \quad (p_{k-1} < q_{k-1}).$$

Furthermore, the elements of any subset, $S(i, k)$, will be designated

$$S(i, k) = [e_{m_1}, \ldots, e_{m_a}, \ldots, e_{m_t}],$$

where

t = number of elements in the subset,

and m_a = identification number of ath element in the subset.

FLOW CHART

The procedure for hierarchical grouping is described explicitly in the flow chart in Figure 1. Note that the optimum union of two subsets $S(i, k)$ is found, the identification numbers of the subsets are specified, and the value for the objective function associated with this optimum union is determined.

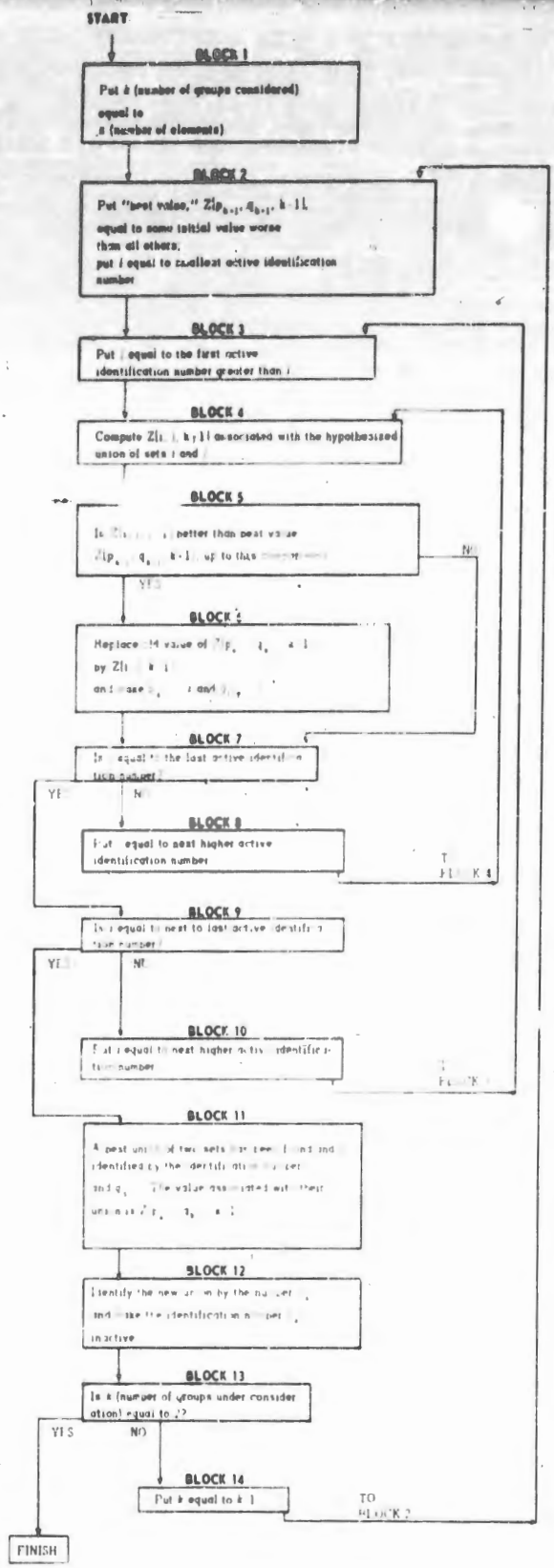

Fig. 1. Flow chart for hierarchical grouping procedure.

NUMERICAL EXAMPLE

To illustrate the grouping procedure, let us consider a problem in which five individuals are to be grouped on the basis of ratings they have given to some object. The data are:

Person	Rating of Object
1	1
2	7
3	2
4	9
5	12

The universal set, U, will have one element corresponding to each person,

$$U = \{e_1, e_2, e_3, e_4, e_5\},$$

where e_i refers to person i. The rating given to the object by person i will be denoted as x_i.

The objective function to be optimized during the hierarchical grouping process will be defined in the following manner. First, consider the hypothesized union,

$$[S(i, k)] \cup [S(j, k)] = \{e_{m_1}, \ldots, e_{m_a}, \ldots, e_{m_t}\},$$

where

t = number of elements in the new set resulting from this hypothesized union,
and m_a = identification number of ath element in this set.

The rating given by the person identified as m_a will be x_{m_a}. Accordingly the objective function associated with the hypothesized union will be defined as

$$Z[i, j, k-1] = \sum_{a=1}^{t} (x_{m_a})^2 - \frac{1}{t}\left(\sum_{a=1}^{t} x_{m_a}\right)^2.$$

This is recognized as the sum of squares of the deviations about the mean of the t observations in the hypothesized union. Hence it will be desired to select that particular union which yields the *smallest* value for the objective function $Z[i, j, k-1]$.

The computational operations to accomplish the hierarchical grouping of the five persons follow this pattern (block numbers here correspond to those in Figure 1):

Block Number	Time Through	Operation
1	1	$k = 5$
2	1	$Z[p_4, q_4, 4]$ = a high value, say, 100
2	1	$i = 1$
3	1	$j = 2$
4	1	$Z[1, 2, 4] = [(1)^2+(7)^2]-.50[(1+7)^2] = 50-32 = 18$
5	1	Is $Z[1, 2, 4] < Z[p_4, q_4, 4]$? i.e., Is $18<100$? Yes; go to Block 6
6	1	$Z[1, 2, 4]$ replaces $Z[p_4, q_4, 4]$ as temporary best
6	1	18 replaces 100 as temporary best
6	1	$(i = 1)$ replaces p_4
6	1	$(j = 2)$ replaces q_4.
7	1	Is $j = 5$? i.e., is $2 = 5$? No; go to Block 8
8	1	Set j = next higher active number = 3. Go to Block 4
4	2	$Z[1, 3, 4] = [(1)^2+(2)^2]-.50[(1+2)^2] = 5-4.5 = .50$
5	2	Is $Z[1, 3, 4] < Z[p_4, q_4, 4]$? i.e., Is $.50<18$? Yes; go to Block 6
6	2	$Z[1, 3, 4]$ replaces $Z[p_4, q_4, 4]$ as temporary best
6	2	.50 replaces 18
6	2	$(i = 1)$ replaces p_4
6	2	$(j = 3)$ replaces q_4
7	2	Is $j = 5$? i.e., is $3 = 5$? No; go to Block 8

The procedure can be continued until all possible unions have been compared. We find the optimum union is one combining $p_4 = 1$ and $q_4 = 3$ for which the objective function is

$$Z[1, 3, 4] = .50.$$

As a result of this union, we have four new sets:

$$S(1, 4) = \{1, 3\}$$
$$S(2, 4) = \{2\}$$
$$S(4, 4) = \{4\}$$
$$S(5, 4) = \{5\}.$$

Notice that the value $q = 3$ is no longer used to identify a set.

When the grouping procedure is repeated with these four sets, we find the best union resulting in three groups is that in which $p_3 = 2$ is combined with $q_3 = 4$, for which the objective function is

$$Z[2, 4, 3] = 2.00$$

As a result of this union, we have three new sets:

$$S(1, 3) = \{1, 3\}$$
$$S(2, 3) = \{2, 4\}$$
$$S(5, 3) = \{5\}.$$

Notice that the values $q_4 = 3$ and $q_3 = 4$ are inactive.

When the procedure is repeated, we find the optimum union leading to two groups is $p_2 = 2$ and $q_2 = 5$, for which the objective function is

$$S[2, 5, 2] = 12.67.$$

The two new sets resulting from this union are:

$$S(1, 2) = \{1, 3\}$$
$$S(2, 2) = \{2, 4, 5\}.$$

Notice that the values $q_4 = 3$, $q_3 = 4$, and $q_2 = 5$ are now inactive identifiers.

Finally, when these two sets are united in one group, $p_1 = 1$ and $q_1 = 2$, the objective function now is

$$Z[1, 2, 1] = 86.80.$$

The final set, then, is expressed as

$$S(1, 1) = \{1, 3, 2, 4, 5\}.$$

Notice that the sequence of the elements reflects the order in which they entered into unions during the grouping process. (See Figure 2.)

Element Identity	Number of Groups				
	1	2	3	4	5
1	1	1	1	1	1
3	3	3	3	3	3
2	2	2	2	2	2
4	4	4	4	4	4
5	5	5	5	5	5
Value of Objective Function:	86.80	12.67	2.00	.50	

Fig. 2. Summary of results of hierarchical grouping for numerical example.

If we prepare the final listing of elements so as to identify the sequence in which elements were united during the grouping operations, it will be extremely useful in describing the results of the hierarchical grouping procedure. Take, for example, Figure 2, which gives a hierarchical listing for the numerical example described. This display has the added advantage of showing at which phase the element groupings took place and the values of the objective functions for each union. In general, when the final ordering of elements is such that each pair of mutually exclusive sets that was united in the grouping process occupies adjacent positions in a list, this list will be called a "list of hierarchical groups," or a "hierarchical listing."

There obviously is more than one list that will satisfy this condition. In the 5-person problem discussed, three different listings of the elements could be called hierarchical listings:

1, 3, 2, 4, 5
or 3, 1, 2, 4, 5
or 5, 4, 2, 3, 1.

However, of all the possible hierarchical listings, only one has its items arranged so as to minimize the number of interchanges needed to place the numbers identifying these items in an ascending series. A hierarchical listing that has this property will be called a "proper list." For the example, the proper list would read 1, 3, 2, 4, 5. The grouping procedure described in this article yields such proper lists as a result of the way in which the identification numbers are ordered during the grouping operations.

SUMMARY

A procedure has been described for forming hierarchical groups of mutually exclusive sets in a manner which yields an optimum value of an objective function. Given k sets, this method permits their reduction to $k-1$ mutually exclusive sets by considering the union of all possible $k(k-1)/2$ pairs that can be formed from these sets and selecting the union with which the optimal value of the objective function is associated. The procedure can be repeated until only one set remains. The method also yields lists of hierarchical groups showing the sequence in which sets have been grouped. A flow chart and numerical example are provided.

REFERENCES

1. **Cox, D.R.** Note on grouping. *J. Amer. statis. Ass.*, 1957, **52**, 543-7.
2. **Fisher, W.D.** On grouping for maximum homogeneity. *J. Amer. statis. Ass.*, 1958, **53**, 789-798.
3. **Tanimoto, T.T., & Loomis, R.G.** *A taxonomy program for the IBM 704.* New York: International Business Machines Corporation (Data Systems Division, Mathematics & Applications Department), 1960. (M&A-6, *The IBM Taxonomy Application*)
4. **Thorndike, R.L.** Who belongs in the family? *Psychometrika*, 1953, **18**, 267-276.
5. **Thorndike, R.L., Hagen, Elizabeth P., Orr, D.B., et al.** *An empirical approach to the determination of Air Force job families.* Lackland Air Force Base, Texas: Air Force Personnel and Training Research Center, August 1957. (*Technical Report* AFPTRC-TR-57-5, ASTIA Document No. AD-134 239)

APPENDIX

DETERMINING THE NUMBER OF POSSIBLE WAYS OF FORMING GROUPS AND THE NUMBER OF DISTINGUISHABLE UNIONS POSSIBLE

NUMBER OF POSSIBLE WAYS OF FORMING GROUPS

To determine $N(P_n)$, the number of possible ways of forming the groups that could result from applying the hierarchical grouping procedure, let n be the number of elements to be grouped. Then we have

$$N(P_n) = \binom{n}{2}\binom{n-1}{2}\binom{n-2}{2}\cdots\binom{2}{2}$$

$$= \left[\frac{n(n-1)}{2}\right]\left[\frac{(n-1)(n-2)}{2}\right]\cdots\left[\frac{2\cdot 1}{2}\right]$$

$$= \frac{n[(n-1)!]^2}{2^{n-1}}.$$

Also observe that

$$N(P_{n+1}) = N(P_n)\left[\frac{n(n+1)}{2}\right]$$

For example, $N(P_n)$ for ns 2 through 8 is

n	$N(P_n)$
2	1
3	3
4	18
5	180
6	2,700
7	56,700
8	1,587,600

NUMBER OF DISTINGUISHABLE UNIONS POSSIBLE

The investigator may wish to determine the number of distinguishable unions that are possible when the hierarchical grouping procedure is used. The way in which this is done is best observed in specific examples that illustrate several independent situations. Therefore, for each value of the number of elements $n = 2$ through 6, we shall consider the possible occupancy numbers that could arise.

APPENDIX (*Continued*)

Possible Occupancy Number	Number of Distinguishable Unions	
For $n = 2$		
1 1	$(\frac{2!}{1!1!})/2!$	= 1
2 0	$(\frac{2!}{2!0!})$	= 1
	Total	2
For $n = 3$		
1 1 1	$(\frac{3!}{1!1!1!})/3!$	= 1
2 1 0	$(\frac{3!}{2!1!})$	= 3
3 0 0	$(\frac{3!}{3!})$	= 1
	Total	5
For $n = 4$		
1 1 1 1	$(\frac{4!}{1!1!1!1!})/4!$	= 1
2 1 1 0	$(\frac{4!}{2!1!1!})/2!$	= 6
2 2 0 0	$(\frac{4!}{2!2!})/2!$	= 3
3 1 0 0	$(\frac{4!}{3!1!})$	= 4
4 0 0 0	$(\frac{4!}{4!})$	= 1
	Total	15

APPENDIX (*Continued*)

Possible Occupancy Number	Number of Distinguishable Unions
For $n = 5$	
1 1 1 1 1	$\left(\frac{5!}{1!1!1!1!1!}\right)/5! = 1$
2 1 1 1 0	$\left(\frac{5!}{2!1!1!1!}\right)/3! = 10$
2 2 1 0 0	$\left(\frac{5!}{2!2!1!}\right)/2! = 15$
3 1 1 0 0	$\left(\frac{5!}{3!1!1!}\right)/2! = 10$
3 2 0 0 0	$\left(\frac{5!}{3!2!}\right) = 10$
4 1 0 0 0	$\left(\frac{5!}{4!1!}\right) = 5$
5 0 0 0 0	$\left(\frac{5!}{5!}\right) = 1$
	Total 52
For $n = 6$	
1 1 1 1 1 1	$\left(\frac{6!}{1!1!1!1!1!1!}\right)/6! = 1$
2 1 1 1 1 0	$\left(\frac{6!}{2!1!1!1!1!}\right)/4! = 15$
2 2 1 1 0 0	$\left(\frac{6!}{2!2!1!1!}\right)/2!2! = 45$
3 1 1 1 0 0	$\left(\frac{6!}{3!1!1!1!}\right)/3! = 20$
2 2 2 0 0 0	$\left(\frac{6!}{2!2!2!}\right)/3! = 15$
3 2 1 0 0 0	$\left(\frac{6!}{3!2!1!}\right) = 60$

(Continued on next page)

Possible Occupancy Number	Number of Distinguishable Unions	
For $n = 6$ (Cont.)		
411000	$(\frac{6!}{4!1!1!})/2!$	= 15
330000	$(\frac{6!}{3!3!})/2!$	= 10
420000	$(\frac{6!}{4!2!})$	= 15
510000	$(\frac{6!}{5!1!})$	= 6
600000	$(\frac{6!}{6!})$	= 1
	Total	203

Div. 15/2, 23/1, 28/4

Wright Air Development Division. Personnel Laboratory, Lackland Air Force Base, Texas. HIERARCHICAL GROUPING TO MAXIMIZE PAY-OFF, by Joe H. Ward, Jr. March 1961. v + 18 p. (Project 7734; Task 17016) (WADD-TN-61-29)
Unclassified report

This report describes mathematically a general procedure for forming hierarchical groups of mutually exclusive sets in a manner which yields an optimum value for the functional relation, or objective function, that reflects the criterion chosen by the investigator. The number of groups to be formed need not be specified in advance. Given k sets, this technique permits their reduction to $k-1$ mutually exclusive sets by considering the union of all possible pairs that can

(over)

UNCLASSIFIED

UNCLASSIFIED

be formed and the selection of that union which has the highest payoff value with respect to the criterion chosen. This procedure can be repeated until only one set remains. Hence decisions on the number of groups to be used can be based on a knowledge of the "costs" of grouping at each stage in the entire hierarchical structure. A computer flowchart and a numerical example of the grouping procedure are provided. An Appendix shows how to determine the number of possible ways of forming groups and the number of distinguishable unions possible.

UNCLASSIFIED

Div. 15/2, 23/1, 28/4

Wright Air Development Division. Personnel Laboratory, Lackland Air Force Base, Texas. HIERARCHICAL GROUPING TO MAXIMIZE PAY-OFF, by Joe H. Ward, Jr. March 1961. v + 18 p. (Project 7734; Task 17016) (WADD-TN-61-29)
Unclassified report

This report describes mathematically a general procedure for forming hierarchical groups of mutually exclusive sets in a manner which yields an optimum value for the functional relation, or objective function, that reflects the criterion chosen by the investigator. The number of groups to be formed need not be specified in advance. Given k sets, this technique permits their reduction to $k-1$ mutually exclusive sets by considering the union of all possible pairs that can

(over)

UNCLASSIFIED

UNCLASSIFIED

be formed and the selection of that union which has the highest payoff value with respect to the criterion chosen. This procedure can be repeated until only one set remains. Hence decisions on the number of groups to be used can be based on a knowledge of the "costs" of grouping at each stage in the entire hierarchical structure. A computer flowchart and a numerical example of the grouping procedure are provided. An Appendix shows how to determine the number of possible ways of forming groups and the number of distinguishable unions possible.

UNCLASSIFIED

Div. 15/2, 23/1, 28/4

Wright Air Development Division. Personnel Laboratory, Lackland Air Force Base, Texas. HIERARCHICAL GROUPING TO MAXIMIZE PAY-OFF, by Joe H. Ward, Jr. March 1961. v + 18 p. (Project 7734; Task 17016) (WADD-TN-61-29) Unclassified report

This report describes mathematically a general procedure for forming hierarchical groups of mutually exclusive sets in a manner which yields an optimum value for the functional relation, or objective function, that reflects the criterion chosen by the investigator. The number of groups to be formed need not be specified in advance. Given k sets, this technique permits their reduction to $k-1$ mutually exclusive sets by considering the union of all possible pairs that can

(over)

UNCLASSIFIED

UNCLASSIFIED

be formed and the selection of that union which has the highest payoff value with respect to the criterion chosen. This procedure can be repeated until only one set remains. Hence decisions on the number of groups to be used can be based on a knowledge of the "costs" of grouping at each stage in the entire hierarchical structure. A computer flowchart and a numerical example of the grouping procedure are provided. An Appendix shows how to determine the number of possible ways of forming groups and the number of distinguishable unions possible.

UNCLASSIFIED

Div. 15/2, 23/1, 28/4

Wright Air Development Division. Personnel Laboratory, Lackland Air Force Base, Texas. HIERARCHICAL GROUPING TO MAXIMIZE PAY-OFF, by Joe H. Ward, Jr. March 1961. v + 18 p. (Project 7734; Task 17016) (WADD-TN-61-29) Unclassified report

This report describes mathematically a general procedure for forming hierarchical groups of mutually exclusive sets in a manner which yields an optimum value for the functional relation, or objective function, that reflects the criterion chosen by the investigator. The number of groups to be formed need not be specified in advance. Given k sets, this technique permits their reduction to $k-1$ mutually exclusive sets by considering the union of all possible pairs that can

(over)

UNCLASSIFIED

UNCLASSIFIED

be formed and the selection of that union which has the highest payoff value with respect to the criterion chosen. This procedure can be repeated until only one set remains. Hence decisions on the number of groups to be used can be based on a knowledge of the "costs" of grouping at each stage in the entire hierarchical structure. A computer flowchart and a numerical example of the grouping procedure are provided. An Appendix shows how to determine the number of possible ways of forming groups and the number of distinguishable unions possible.

UNCLASSIFIED