AD-767 199

COMPREHENSIVE OCCUPATIONAL DATA ANALYSIS PROGRAMS (CODAP): GROUP MEMBERSHIP (GRMBRS/GRPMBR) AND AUTOMATED DIAGRAMMING (DIAGRM) PROGRAMS

William J. Phalen, et al

Air Force Human Resources Laboratory
Brooks Air Force Base, Texas

April 1973

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

AIR FORCE

HUMAN RESOURCES

AD 767199

AFHRL-TR-73-5

COMPREHENSIVE OCCUPATIONAL DATA ANALYSIS PROGRAMS (CODAP)

GROUP MEMBERSHIP (GRMBRS/GRPMBR) AND AUTOMATED DIAGRAMMING (DIAGRM) PROGRAMS

By

William J. Phalen
Raymond E. Christal

PERSONNEL RESEARCH DIVISION
Lackland Air Force Base, Texas 78236

DDC
RECEIVED
OCT 3 1973
B

April 1973

Approved for public release; distribution unlimited.

LABORATORY

18

AIR FORCE SYSTEMS COMMAND

BROOKS AIR FORCE BASE, TEXAS

NOTICE

Unclassified
Security Classification

AD-767199

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*
Personnel Research Division
Air Force Human Resources Laboratory
Lackland Air Force Base, Texas 78236

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP

3. REPORT TITLE
COMPREHENSIVE OCCUPATIONAL DATA ANALYSIS PROGRAMS: GROUP MEMBERSHIP (GRMBRS/GRPMBR) AND AUTOMATED DIAGRAMMING (DIAGRM) PROGRAMS

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

5. AUTHOR(S) *(First name, middle initial, last name)*
William J. Phalen
Raymond E. Christal

6. REPORT DATE
April 1973

7a. TOTAL NO. OF PAGES
18

7b. NO. OF REFS
10

8a. CONTRACT OR GRANT NO.

b. PROJECT NO. 7734

c. Task No. 773403

d. Work Unit No. 77340304

9a. ORIGINATOR'S REPORT NUMBER(S)
AFHRL-TR-73-5

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT
Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY
Personnel Research Division
Air Force Human Resources Laboratory
Lackland Air Force Base, Texas 78236

13. ABSTRACT

This technical report describes two Comprehensive Occupational Data Analysis Programs (CODAP): (*a*) the "Group Membership" (GRMBRS/GRPMBR) programs, which produce a report that identifies the two groups combining at each stage of a hierarchical grouping process, and (*b*) the "Automated Diagramming Cluster Merger" (DIAGRM) program, which uses GRMBRS information as input and displays the hierarchical grouping actions in diagrammatic form.

The detailed descriptions of the GRMBRS/GRPMBR and DIAGRM programs contained in this report should be of particular interest to agencies and organizations using the CODAP system.

DD FORM 1 NOV 65 1473

Unclassified
Security Classification

ib

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
automated diagramming program (DIAGRM)						
clustering procedures						
CODAP programs						
Comprehensive Occupational Data Analysis Programs (CODAP)						
computerized diagramming						
computerized job analysis techniques						
computerized clustering procedures						
computer programs used in job analysis						
data analysis computer programs						
diagramming by computer						
DIAGRM program						
GRMBRS program						
GRPMBR program						
grouping procedures						
group membership programs						
group membership reports						
hierarchical grouping						
inventory of tasks						
job analysis						
job grouping						
job inventory						
job typing						
occupational analysis						
occupational survey						
task-level job analysis						
task ratings						

FOREWORD

The Personnel Research Division has been working continuously for the past 10 years on a comprehensive set of data analysis programs called CODAP. These programs are designed for analyzing and retrieving occupational survey information collected by means of job inventories. The CODAP system was originally written for execution on an IBM 7040 computer. Many of the time-consuming routines were coded in machine language (MAP), thus making it difficult for agencies not having access to this specific piece of equipment to make use of these powerful analysis programs. The United States Marine Corps had the CODAP system rewritten under contract for execution on an IBM 360-65 computer for application by their Service; and the Office of the Assistant Secretary of Defense for Manpower and Reserve Affairs has arranged to get the system operational on an IBM 370-155 computer for possible use by all military services, as well as the United States Coast Guard. As the CODAP system becomes available to agencies outside the Air Force, there is a recognized need for the various programs to be described in some detail. It is not feasible to describe the entire system in a single document. Instead, a series of technical reports is being prepared, each of which will describe one or more of the CODAP programs. These reports are written for the analyst or investigator who is charged with making occupational data meaningful to personnel managers. Other manuals and documents of a more technical nature will be published for use by programmers, systems analysts, computer operators, and control card specialists who service the analyst.

The technical development in job analysis methodology described in this particular report was accomplished under Project 7734, Development of Methods for Describing, Evaluating, and Structuring Air Force Occupations; Task 773403, Development and Evaluation of Methods for Grouping Work Activities into Positions, Specialties, and Career Fields. Specifications were supplied to personnel of the Computer Sciences Corporation, who wrote the programs. This particular report describes two CODAP programs: (*a*) the "Group Membership" (GRMBRS/GRPMBR) programs, which produce a report that identifies the two groups combining at each stage of a hierarchical grouping process, and (*b*) the Automated Diagramming Cluster Merger" (DIAGRM) program, which uses GRMBRS information as input and displays the hierarchical grouping actions in diagrammatic form.

Technical assistance and suggestions by SSgt Philip Aitken-Cade appreciably improved the format of the DIAGRM output.

This report has been reviewed and is approved.

Harold E. Fischer, Colonel, USAF
Commander

ABSTRACT

This technical report describes two Comprehensive Occupational Data Analysis Programs (CODAP): (*a*) the "Group Membership" (GRMBRS/GRPMBR) programs, which produce a report that identifies the two groups combining at each stage of a hierarchical grouping process, and (*b*) the "Automated Diagramming Cluster Merger" (DIAGRM) program, which uses GRMBRS information as input and displays the hierarchical grouping actions in diagrammatic form.

The detailed descriptions of the GRMBRS/GRPMBR and DIAGRM programs contained in this report should be of particular interest to agencies and organizations using the CODAP system.

TABLE OF CONTENTS

Page
I. Introduction . . . 1
II. The CODAP Group Membership (GRMBRS/GRPMBR) Programs . . . 1
Group Membership – Short Listing (GRMBRS) . . . 1
Group Membership – Full Listing (GRPMBR) . . . 2
III. The CODAP Automated Diagramming Program (DIAGRM) . . . 4
IV. Anticipated Developments . . . 10
References . . . 11

LIST OF FIGURES

Figure Page
1 Group membership – short listing (GRMBRS) . . . 2
2 Group membership – full listing (GRPMBR) . . . 3
3 Segment of Column 1 of a cluster diagram . . . 6
4 Selected portions of a group membership printout . . . 7
5 DIAGRM program flow chart . . . 8
6 Diagram of clusters and subclusters in the Inventory Management and Materiel Facilities Career Ladder – 645X0/647X0 . . . 9

Preceding page blank

COMPREHENSIVE OCCUPATIONAL DATA ANALYSIS PROGRAMS: GROUP MEMBERSHIP (GRMBRS/GRPMBR) AND AUTOMATED DIAGRAMMING (DIAGRM) PROGRAMS

I. INTRODUCTION

The Personnel Research Division of the Air Force Human Resources Laboratory conducts occupational surveys as authorized by Air Force Manual 35-2. The detailed procedures for developing and administering surveys are spelled out by Archer and Fruchter (1963) and by Morsh and Archer (1967).

A system of Comprehensive Occupational Data Analysis Programs (CODAP)[1] has been developed to aid in the analysis of occupational surveys. Explanations and practical applications of the CODAP package are stated by Morsh and Christal (1966) and by Archer (1966). The mathematical processes underlying the grouping of jobs on an index of similarity are discussed in detail by Ward (1961), Morsh (1965), and by Christal and Ward (1967). Detailed program descriptions are spelled out in the *Program Description Manual–CODAP* (1969) and in the *Program Description Manual–DIAGRM* (1969), which were authored by the Computer Sciences Corporation.

The purpose of this report is to describe two CODAP programs: (*a*) the "Group Membership" (GRMBRS/GRPMBR) programs, which produce a report that identifies the two groups combining at each stage of a hierarchical grouping process, and (*b*) the "Automated Diagramming Cluster Merger" program (DIAGRM), which uses GRMBRS information as input and displays the hierarchical grouping actions[2] in diagrammatic form.

II. THE CODAP GROUP MEMBERSHIP (GRMBRS/GRPMBR) PROGRAMS

The GRMBRS program is referred to in the CODAP Manual as "Group Membership – Short Listing," because there is also a program designated GRPMBR, which is referred to as "Group Membership – Full Listing." The GRPMBR Program (full listing) is seldom used, because the report it generates is too bulky and repetitive to satisfy job analysis needs as well as GRMBRS (short listing). However, GRPMBR will be described in terms of its unique features which may, at times, serve a useful purpose and these features should be documented.

Group Membership – Short Listing (GRMBRS)

GRMBRS is a report that identifies the two groups combining at each stage of the hierarchical grouping process and displays information describing each group. Figure 1 shows the final 12 stages of a GRMBRS printout. The information contained in GRMBRS includes: the stage number; the number of members in the combined group; the average percentage of overlap "between" the merging groups and "within" the combined group; the number of members within the merging groups; a computer-assigned group number which is used by the computer but is of no consequence to the job analyst; the stages at which the subgroups of the combined group were formed; and the sequences of special case identification numbers (called KPATH sequence numbers) assigned to members of the merging groups.

The average percentage of overlap "between" the merging groups and "within" the combined group are similarity indices computed during the "Grouping" (GROUP) program, although the initial matrix of pairwise overlaps between *individual* cases is computed during the "Overlap" (OVRLAP) program, which precedes GROUP. By way of a brief definition, "between" = the average percentage of overlap between all possible pairings of members in Group A with members in Group B; "within" = the average percentage of overlap between all possible pairs of individuals within the combined A and B Groups. A more thorough explanation of overlap "between" and "within" is presented by Archer (1966). "KPATH sequence number" and "KPATH order" refer to the sequencing of cases and their related data after the cases have been hierarchically grouped. Upon completion of a hierarchical grouping, all cases are sequenced in such a way that those cases which

---

[1] Formerly called "Computerized Occupational Data Analysis Programs (CODAP)."

[2] Grouping (or clustering) procedures are designed to group larger numbers of persons, objects, jobs, etc., into smaller numbers of mutually exclusive classes in which the members have similar characteristics. When the grouping is done in a manner that establishes a taxonomy of mutually exclusive clusters wherein each larger unit is a unique combination of the next-subordinate units, the clusters are called "hierarchical groups."

Membership at Each Stage of Hierarchical Grouping for Data Systems Inventory (DAFSC 68XX0)

Stage	No. Members	Average Overlaps Between	Average Overlaps Within	No. Mbrs. in Merging Groups	Group	Stage Formed	KPATH From	KPATH To
1	1621	0.02	13.20	1620	1	2	1	1620
				1	259	1	1621	1621
2	1620	0.11	13.22	1619	1	3	1	1619
				1	305	2	1620	1620
3	1619	0.21	13.24	1618	1	4	1	1618
				1	827	3	1619	1619
4	1618	0.39	13.25	1617	1	5	1	1617
				1	1068	4	1618	1618
5	1617	0.46	13.27	1615	1	6	1	1615
				2	554	1620	1616	1617
6	1615	0.57	13.30	1614	1	7	1	1614
				1	802	6	1615	1615
7	1614	0.62	13.32	1520	1	8	1	1520
				94	4	17	1521	1614
8	1520	1.65	14.82	1512	1	9	1	1512
				8	359	12	1513	1520
9	1512	1.71	14.96	1509	1	10	1	1509
				3	187	62	1510	1512
10	1509	2.33	15.01	1497	1	11	1	1497
				12	55	18	1498	1509
11	1497	2.65	15.21	1496	1	13	1	1496
				1	945	11	1497	1497
12	8	4.17	25.11	3	359	678	1513	1515
				5	754	21	1516	1520
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Fig. 1. **Group membership – short listing (GRMBRS).**

merged at each stage of the grouping process are positioned adjacently. The sequential numbering of cases so positioned produces what are known as "KPATH sequence numbers." The topic of KPATH ordering will be thoroughly discussed in a future report.

Because GRMBRS orders groups in stage number sequence, that is, from the pair of combined groups with the lowest "between" value to the pair with the highest "between" value, the hierarchical sequence of groups is not directly discernible. Therefore, the column entitled "Stage Formed" (Figure 1) was added to the initial version of the GRMBRS program to enable the job analyst to identify, without lengthy searching, the prior stages at which two merged subgroups had been formed in the hierarchical grouping process. Thus, one can see that the group of 1520 members formed at Stage 8 (Figure 1) is composed of a group of 1512 members formed at Stage 9 (KPATH range = 1 to 1512) and a group of 8 members formed at Stage 12 (KPATH range = 1513 to 1520.

Group Membership – Full Listing (GRPMBR)

The GRPMBR program lists for each grouping stage not only the group formed at a particular stage, but also all other groups that have already been formed and are still in existence as of that stage. At the requester's option, a specified stage, a range of stages, or several ranges of stages will be reported. Figure 2 is an example of a GRPMBR printout for Stage 100 of the hierarchical grouping of the Medical Laboratory job inventory data. The group formed at Stage 100 is preceded by two

KPATH Order – Group Membership for Medical Laboratory Inventory (DAFSC 904X0)
Membership at Each Stage of Hierarchical Grouping

Stage Formed	No. Members	Average Overlaps		No. Mbrs. in Merging Groups	Group	KPATH	
		Between	Within			From	To
103	227	51.07	58.28		1	1	227
136	16	54.21	60.07		85	228	243
* 100	49	50.61	54.46	21	110	245	265
*				28	145	266	293
118	2	52.53	76.27		558	295	296
116	2	52.46	76.22		379	298	299
122	2	52.98	76.50		152	300	301
132	2	53.71	76.85		156	302	303
107	5	51.53	62.26		35	304	308
245	3	61.03	75.57		5	310	312
197	2	58.32	79.17		21	313	314
109	148	51.71	60.53		24	315	462
175	2	56.84	78.42		86	466	467
121	2	52.92	76.46		600	471	472
105	3	51.23	68.93		6	475	477
140	2	54.53	77.27		62	478	479
108	6	51.53	65.13		28	480	485
126	7	53.15	62.67		34	487	493
208	2	58.98	79.48		111	494	495
185	2	57.61	78.81		248	496	497
102	2	50.94	75.47		277	498	499
120	3	52.90	69.14		313	500	502
187	2	57.67	78.84		607	503	504
134	3	54.17	70.07		133	506	508
106	19	51.46	60.28		20	515	533
171	3	56.63	71.90		92	535	537
130	5	53.53	66.22		23	538	542
101	2	50.88	75.44		27	543	544
138	2	54.44	77.22		59	556	557
300	2	63.30	81.64		11	563	564
332	3	64.65	80.43		38	570	572
232	2	60.51	80.26		369	573	574
110	3	51.78	69.01		182	575	577
145	2	54.90	77.44		66	581	582
221	2	60.15	80.07		30	585	586
262	2	61.75	80.86		141	591	592
131	7	53.67	64.68		2	594	600
111	7	51.84	61.90		50	601	607
133	2	54.13	77.06		41	610	611
	62 Single Members						

***

Fig. 2. **Group membership – full listing (GRPMBR).**

asterisks. All groups existing at Stage 100 which contain two or more members are listed in KPATH order. As might be expected, all other stage numbers in the "Stage Formed" column are higher than 100. At the bottom of the report for Stage 100 is the statement "62 Single Members," which means that 62 individuals have not yet been grouped. The values in the "No. Members" column indicate how many members there are in each group, and if the values in this column are summed, the total would equal the number of cases in the study, *i.e.*, 619.

III. THE CODAP AUTOMATED DIAGRAMMING PROGRAM (DIAGRM)

Before the development of the DIAGRM program, job analysts produced handmade diagrams of hierarchical grouping actions from the "Group Membership" (GRMBRS) printout. The DIAGRM program was designed to fill two needs: (*a*) to mechanize a routine and time-consuming task which was a prerequisite to all job analysis studies, *i.e.*, to produce a diagram of the hierarchical grouping actions reported in GRMBRS, and (*b*) to standardize the diagramming procedure so that task scientists could read and interpret each other's diagrams of hierarchically grouped data. Both objectives were achieved. The diagramming procedure has been reduced from a 40 man-hour chore of hand drafting to a 15-minute computer routine. The computerized diagram, once it has been explained to users, enables them to read and interpret other job analysts' diagrams without having to review a maze of subjective, and often confusing, configurations and embellishments.

The DIAGRM program uses the data from the GRMBRS program to generate a tree-like diagram that graphically displays which groups were merged during the hierarchical grouping process. Each group is represented by a rectangular block of data containing the most important information reported in GRMBRS. Rows and columns of asterisks show the branches leading back from a group to its subgroups, and leading forward to the group with which it merges. A small portion of a DIAGRM printout is shown below.

```
Group 634 →  0634 0016  0716 0004  ← Group 716
             0549-0564  0565-0568
             58.6  64.4 60.6  71.9
                  *           *
                  *           *
Group 574 →  0574 0020        *
             0549-0568******
             57.0 - 62.3
                  *
                  *
```

Using Group 574 as an example, the numbers in the data block are interpreted as follows:

0574 = the stage in the clustering process at which this group was formed by the merging of two smaller groups. The stage number is also used to identify the group formed at this stage, *i.e.*, Stage 574 = Group 574.

0020 = the number of members in the group.

0549 -
0568 = the range of numbers that is used to identify the location of the 20 group members on the KPATH history tape or on any KPATH printout of case data. This range is referred to as the "KPATH range."

57.0 = the average percentage of overlap "between" members of the two subgroups that merged to form Group 574, namely, Group 634 and Group 716.

62.3 = the average percentage of overlap "within" or among the members of Group 574.

Vertical and horizontal lines of asterisks are used to show mergers of two groups into a larger group. Thus, Group 634 and Group 716 merged to form Group 574, which has 20 members grouped in the KPATH range from 549 to 568. Group 574 has an average "between" value of 57.0 ("between" = the average percentage of overlap between all possible pairings of members of Group 634 with members of Group 716), and an average "within" value of 62.3 ("within" = the average percentage of overlap between all possible pairs of individuals within the newly formed Group 574). The stage numbers of the combining groups are higher than those of the combined group, because stage numbers decrease iteratively from "N-1" to "1" as the grouping process progresses. At Stage 0001, all cases have been merged into a single group. The stage number also indicates the number of groups existing at that stage, including one-member groups. In this context, the number "0574" in the combined group states that there are *574 groups remaining* at the time *Group 574* is formed at *Stage 574*.

Although DIAGRM can show all grouping stages, the program has features which may be used to limit the number of stages printed. Through judicious application of a set of constants, it is possible to vary the point at which a group is established as a "starter." A starter group is the furthest breakdown of a cluster to be shown in the diagram; thus, all groups appearing in the top row of the diagram are starters. The constants and their functions are defined as follows:

N/V = the minimum number of stages to be diagrammed.

W = an integer value that defines the minimum number of members in a starter group.

X = an "average between group overlap" value used to establish a beginning point for selecting starter groups. The X-value is subject to change by the program, depending on the value of N/V. If the average "between" value at stage N/V is less than or equal to the input X value, then output X = input X, and N/V is changed to the stage number corresponding to the stage number at the value of input X. If the average "between" value at stage N/V is greater than the input X value, then the output X becomes the observed value of X at stage N/V.

Y = an "average between group overlap" value used as the lower limit cutoff in selecting starter groups.

An example of the way in which these constants operate is displayed by jointly viewing Figures 3, 4, and 5. Figure 3 is an excerpt of a DIAGRM printout; Figure 4 shows segments from a GRMBRS printout; Figure 5 is a flow chart of how the DIAGRM program works.

Assume a set of input values for the four constants as: X = 50.00, Y = 25.00, W = 3, and N/V = 810. Compared with the output values shown in the heading of Figure 3, note that only X is different.[3] The explanation can be traced in Figure 4. The defined value of N/V = 810 instructed the program to go first to Stage 810; the value of X ("average overlap between") at that stage is 53.87. Since this quantity is greater than the *input* value of X = 50.00, *output* X became 53.87. Had we used N/V as 677 rather than 810, then the program would have read the "between" value at that stage as X = 49.98 (Figure 4). This value is *less* than *input* X = 50.00, and, in that case, N/V would have been changed to Stage 678, where the "average overlap between" was 50.00. In this manner, N/V would have been changed to correspond to the *input* value of X.

Consider now the function of W in the system. With Stage 810 as the defined beginning point and moving in the direction of Stage 1620, the program selects starter groups; that is, those groups which will form the top row of the diagram. These are the smallest groupings that would be of interest to the job analyst. Limiting the number of starter groups holds the printout to manageable size, and the W constant is one way to establish such a limit. In the assumed set of values stated above, W = 3. The first stage after 810 to meet this minimum is 812, followed by 813, 815, and 816 (Figure 4).

To assure selection of nonoverlapping groups (or to preclude the selection of subgroups of groups already selected), the KPATH range of each group is used as a control. The KPATH range identifies precisely which cases are in the group. In Figure 4, the KPATH range is shown in the last two columns. For example, Group 816 contains 96 cases (second column) and includes all cases with KPATH numbers from 1115 to 1210. The DIAGRM program records this range, will select no other groups which overlap that range, and thereby precludes the selection of any subgroups within the selected group.

When the program completes the selection of starter groups between Stages 810 and 1620 (NSTAGE), it goes to Stage 809 and works toward Stage 1, selecting groups of size W or larger. Groups added in this process will have no more than four cases, since any group larger than four must contain a subgroup of at least three cases, which would have been selected at a higher stage.

The constant Y, to which we gave a value of 25.00, establishes the lower limit cutoff for "average between group overlap." In the present example (Figure 4), Stage 110 met the criterion of W = 3, but the "between" value was only 24.82; therefore, Stage 110 would not have been selected as one of the starter groups, because we had set 25.00 as the Y value.

Upon completion of the selection process, the starter groups are sequenced from low to high on

---

[3] The statement "NSTAGE = 1620" at the top of Figure 3 tells how many hierarchical grouping stages were required to complete the grouping. NSTAGE = the number of cases in the sample minus one. The "minus one" is a result of two individuals being merged by the first grouping action, thus reducing the total number of single-case groups by one. Therefore, after the first grouping action, the number of groups remaining is 1619 single-case groups and one two-case group: a total of 1620 groups.

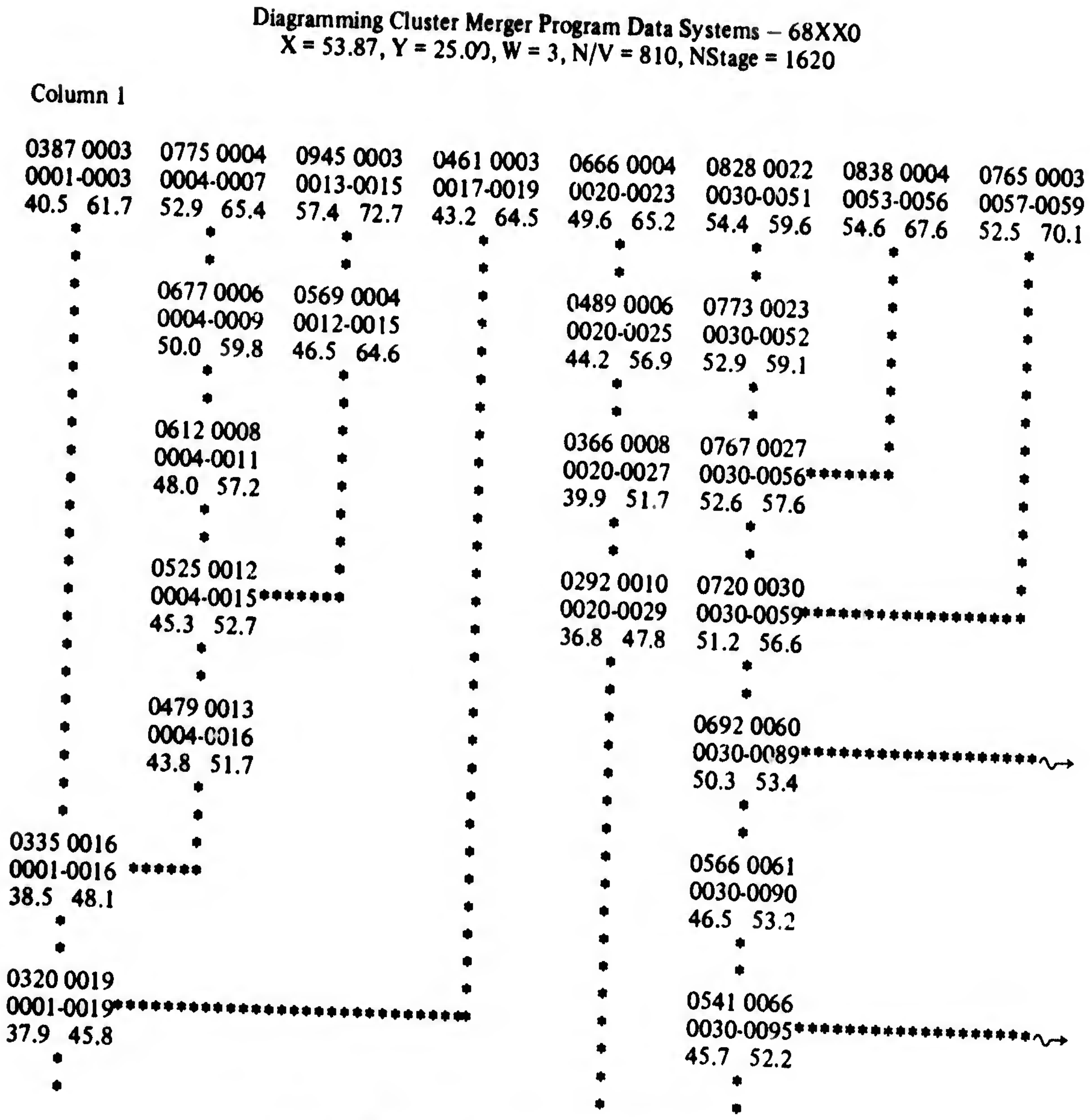

Fig. 3. **Segment of Column 1 of a cluster diagram.**

KPATH range across the top row of the printout. Each print sheet accommodates 11 groups across. The DIAGRM excerpt in Figure 3 shows the first eight starter groups from the Data Systems job inventory analysis. Following the layout of the starter groups, all subsequent mergers of groups are displayed on the diagram. If a group accumulates additional cases or adds on groups that failed to be selected as starter groups, these additions are shown as part of an enlarged group directly below the groups to which they are added. For example, the second starter group in Figure 3 (Group 0775) has four members and runs from KPATH 0004 to 0007. Directly below this group (separated by two asterisks) is Group 0677, which contains six members and includes KPATH 0004 to 0009. Here we can see that Group 0677 includes Group 0775 plus two additional members.

Wherever two groups that are starter groups or enlargements of starter groups come together, two sets of asterisks are used to show into what group they are merged. Thus, Group 0612 merges with Group 0569 to form Group 0525, which contains the 12 members previously contained in Groups 0612 and 0569.

Membership at Each Stage of Hierarchical Grouping

		Average Overlaps			KPATH	
Stage	No. Members	Between	Within	No. Mbrs. in Merging Groups	From	To
1	1621	0.02	13.20	1620	1	1620
				1	1621	1621
2	1620	0.11	13.22	1619	1	1619
				1	1620	1620
⋮	⋮	⋮	⋮	⋮	⋮	⋮
110	3	24.82	53.12	2	490	491
				1	492	492
111	2	25.00	62.50	1	1508	1508
				1	1509	1509
112	2	25.00	62.50	1	1518	1518
				1	1519	1519
113	14	25.01	43.28	13	1428	1440
				1	1441	1441
114	349	25.22	30.89	208	1	208
				141	209	349
115	3	25.23	52.56	2	1604	1605
				1	1606	1606
116	6	25.26	45.76	2	1414	1415
				4	1416	1419
⋮	⋮	⋮	⋮	⋮	⋮	⋮
677	6	49.98	59.78	4	4	7
				2	8	9
678	3	50.00	72.22	1	1513	1513
				2	1514	1515
⋮	⋮	⋮	⋮	⋮	⋮	⋮
807	10	53.84	60.50	7	1029	1035
				3	1036	1038
808	2	53.84	76.92	1	1363	1363
				1	1364	1364
809	2	53.87	76.93	1	479	479
				1	480	480
810	2	53.87	76.94	1	597	597
				1	598	598
811	2	53.87	76.94	1	18	18
				1	19	19
812	3	53.90	72.75	1	1416	1416
				2	1417	1418
813	7	53.94	62.82	6	147	152
				1	153	153
814	2	53.97	76.98	1	91	91
				1	92	92
815	4	53.98	67.36	2	1245	1246
				2	1247	1248
816	96	54.01	60.45	91	1115	1205
				5	1206	1210
⋮	⋮	⋮	⋮	⋮	⋮	⋮
1619	2	99.99	99.99	1	1348	1348
				1	1349	1349
1620	2	100.00	100.00	1	1616	1616
				1	1617	1617

Fig. 4. **Selected portions of a group membership printout.**

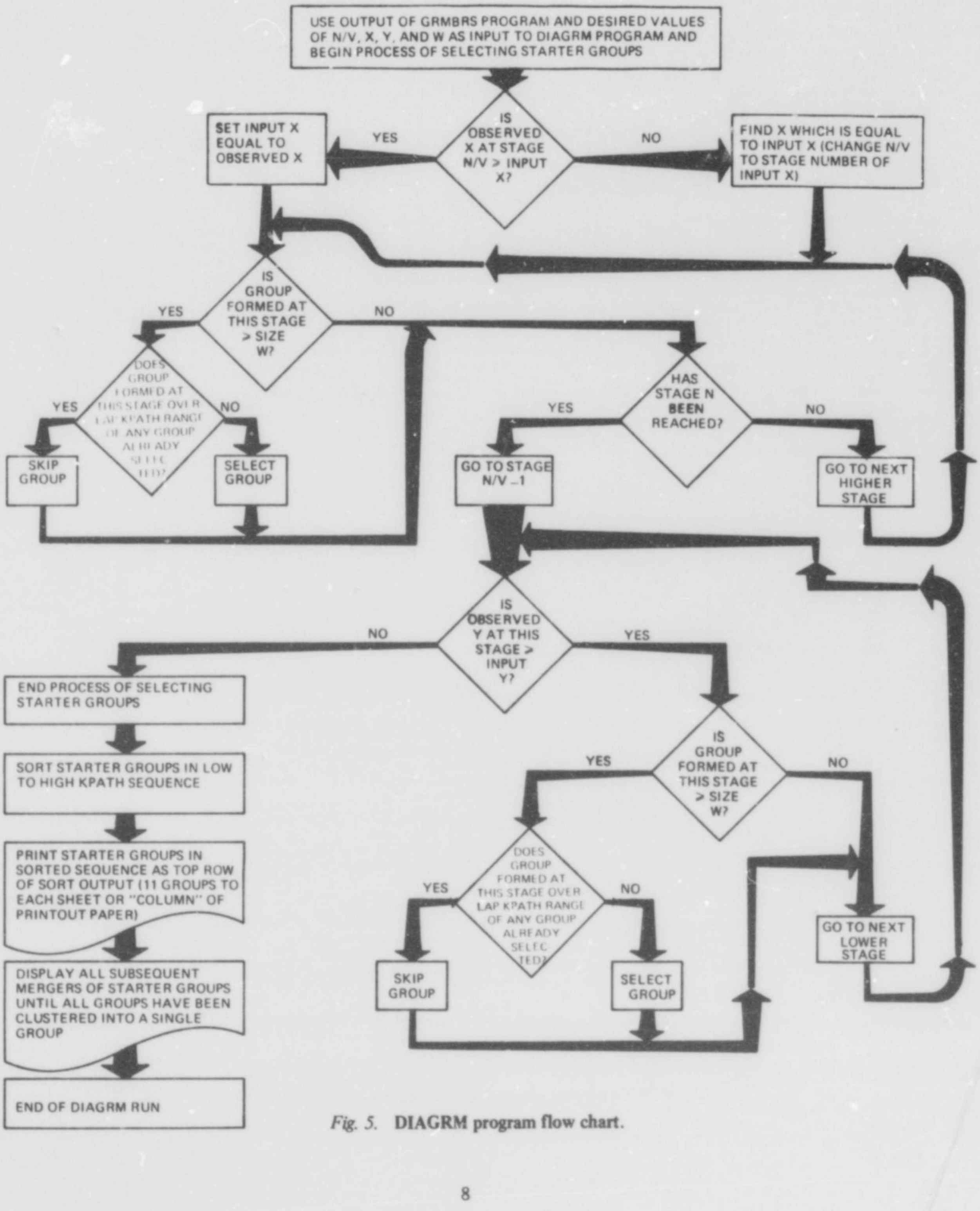

Fig. 5. **DIAGRM program flow chart.**

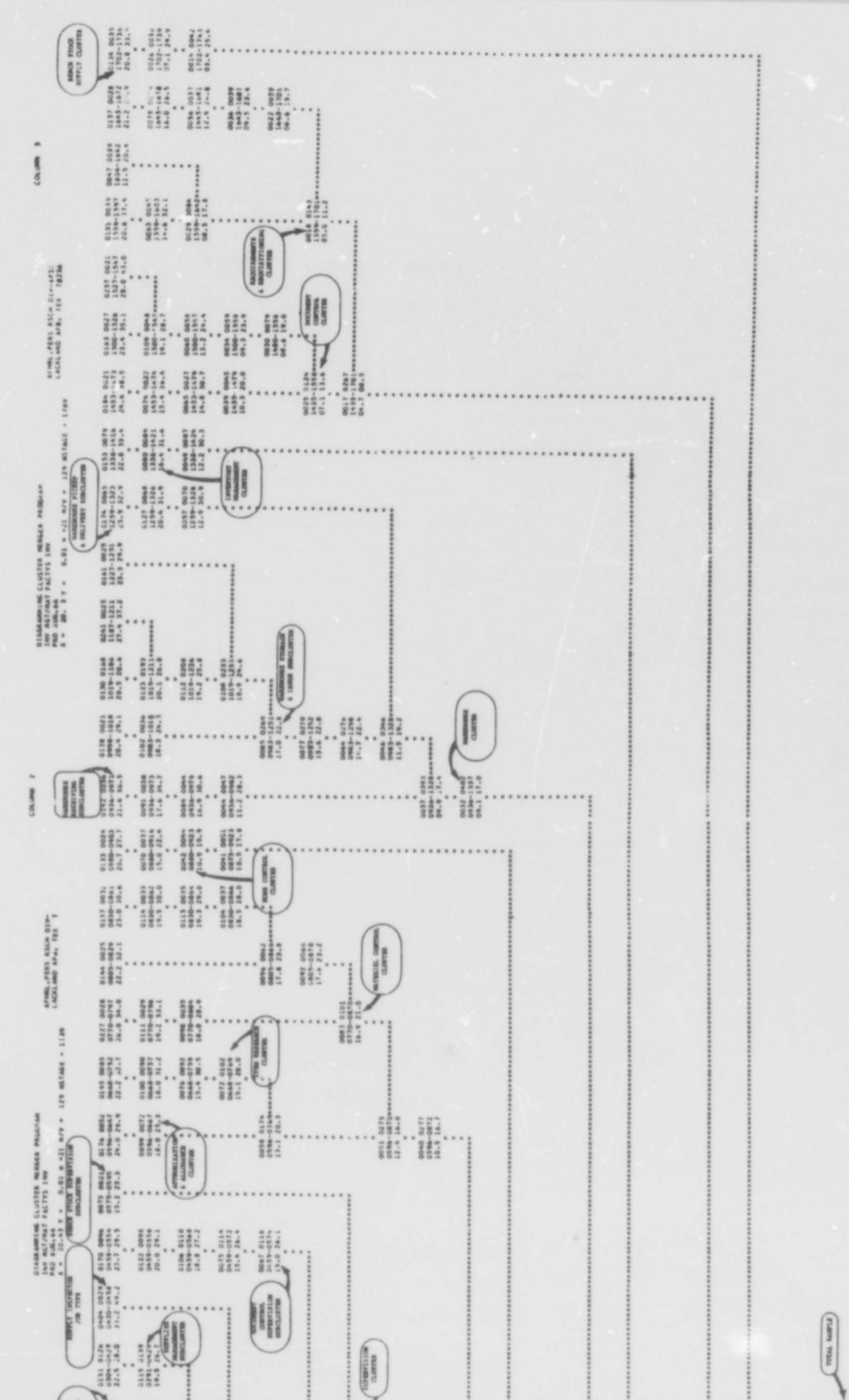

Fig. 6. **Diagram of clusters and subclusters in the Inventory Management and Materiel Facilities Career Ladder - 645X0/647X0.**

Figure 6 is a complete diagram that has been photographically reduced to give a visual impression of how a complete diagram appears. Because only the largest and most important groups were desired in this diagram, the X, Y, W, and N/V values were adjusted so as to limit the displayed groups primarily to those desired, plus the remaining mergers leading to Group 0001. The principal groups, called clusters or subclusters, are identified on the diagram by typed-in names supplied by a job analyst. Specialized diagrams, such as Figure 6, are easy to request once the groups desired for display have been decided upon. In the case of Figure 6, X was set at 20.43, so as to obtain as a starter group the "Warehouse Pickup and Delivery Subcluster" rather than Group 0127 below it (which had a "between" value of 20.4); Y was set at 0.01, so as to avoid leaving out starter groups with low "between" values, such as the "Bench Stock Supervision Subcluster," which had a "between" value of 15.2; W was set at 21, so that no group smaller than the "Bench Stock Supervision Subcluster," which had exactly 21 members, would be selected as a starter group; N/V was set at Stage 0129, because it was already known by looking at the GRMBRS printout, which is not shown, that the desired value of X (X = 20.43) occurred at Stage 0129.

In normal job analysis situations, where groups much smaller than clusters are of interest, the following input constants are routinely requested: X = 50.00, Y = 25.00, W = 4 or 5, and N/V = the median stage. These constants seem to provide an adequate amount of diagram without prior scrutiny of the GRMBRS printout. In fact, GRMBRS and DIAGRM are usually requested simultaneously.

IV. ANTICIPATED DEVELOPMENTS

As the programs in the CODAP package grow and become more sophisticated in the types of information they provide, it is quite likely that additional or different kinds of information will be made available in the DIAGRM program.[4]

The method DIAGRM employs in selecting starter groups also provides a basis for determining at what point in the clustering process one should start comparing groups, if a totally automated job type selection program were implemented. Steps are already being taken in this direction.

Recently, researchers have shown interest in using DIAGRM for displaying grouping processes other than those encountered in job analysis, *e.g.*, biological taxonomies. It is hoped that this publication will be instrumental in eliciting as yet unanticipated applications of the DIAGRM program and of the CODAP system as a whole.

---

[4] The DIAGRM program has recently been adapted for use in conjunction with the CODAP Profile Analysis Program, which uses raw difference scores rather than percent similarity values. The GRMBRS program, which has been used in CODAP Profile Analysis from the beginning, was also modified so that all overlap "between" and "within" raw difference scores could be expressed as percentages of the Stage 1 "between" score. This change was necessary because DIAGRM is limited to a maximum value of ±99.9. The percentage value 100.00, which will occur at Stage 1, is automatically converted to 99.9. All other values will be less than 100.00. The maximum limit for GRMBRS is ±99999.99.

REFERENCES

AF Manual 35-2. *Occupational analysis: Procedures for surveying and evaluating Air Force Specialties.* Washington: Department of the Air Force, 9 July 1968.

Archer, W.B., & Fruchter, D.A. *The construction, review, and administration of Air Force job inventories.* PRL-TDR-63-21, AD-426 755. Lackland AFB, Tex.: Personnel Research Laboratory, Aerospace Medical Division, August 1963.

Archer, W.B. *Computation of group job descriptions from occupational survey data.* PRL-TR-66-12, AD-653 543. Lackland AFB, Tex.: Personnel Research Laboratory, Aerospace Medical Division, December 1966.

Christal, R.E., & Ward, J.H., Jr. The MAXOF clustering model. In M. Lorr & S. B. Lyerly (Eds.), *Proceedings of the conference on cluster analysis of multivariate data.* New Orleans: Catholic University of America, June 1967, 11.02-11.45.

Computer Sciences Corporation. *Program description manual – CODAP – Computerized Occupational data analysis programs.* (5th ed.) El Segundo, Calif.: Computer Sciences Corporation, September 1969.

Computer Sciences Corporation. *Program description manual – DIAGRM – Computerized diagram of hierarchical grouping actions.* (supplement to *Program description manual – CODAP – Computerized occupational data analysis programs* (5th ed.) El Segundo, Calif.: Computer Sciences Corporation, September 1969.

Morsh, J.E. *Identification of job types in the personnel careerfield.* PRL-TR-65-9, AD-622 433. Lackland AFB, Tex.: Personnel Research Laboratory, Aerospace Medical Division, April 1965.

Morsh, J.E., & Christal, R.E. *Impact of the computer on job analysis in the United States Air Force.* PRL-TR-66-19, AD-656 304. Lackland AFB, Tex.: Personnel Research Laboratory, Aerospace Medical Division, October 1966.

Morsh, J.E., & Archer, W.B. *Procedural guide for conducting occupational surveys in the United States Air Force.* PRL-TR-67-11, AD-664 036. Lackland AFB, Tex.: Personnel Research Laboratory, Aerospace Medical Division, September 1967.

Ward, J.H., Jr. *Hierarchical grouping to maximize payoff.* WADD-TN-61-29, AD-261 750. Lackland AFB, Tex.: Personnel Laboratory, Wright Air Development Division, March 1961.