AD-A057 337 MINNESOTA UNIV MINNEAPOLIS MANAGEMENT INFORMATION SY--ETC F/G 9/2
COMPUTERIZED FILE ORGANIZATION DESIGN SYSTEM (VERSION 2): USER'--ETC(U)
JUN 77 S T MARCH, D G SEVERANCE N00014-76-C-1063
UNCLASSIFIED MISRC-TR-77-12 NL

| OF |
AD
A057337

END
DATE
FILMED
9-78
DDC

LEVEL II

Report Number: MISRC-TR-77-12

AD A057337

COMPUTERIZED FILE ORGANIZATION DESIGN SYSTEM (VERSION 2):

USER'S MANUAL

Management Information Systems Research Center
Graduate School of Business Administration
University of Minnesota
Minneapolis, MN 55455

1977 June

AU NO.
DDC FILE COPY

UNCLASSIFIED; DISTRIBUTION UNLIMITED

DDC
RECEIVED
AUG 10 1978
D

Prepared for and delivered to:

David W. Taylor Naval Ship Research and Development Center
ATTN: Code 1821, Dr. David K. Jefferson
U.S. Department of the Navy
Bethesda, MD 20084

The research reported here was funded by the David W. Taylor Naval Ship Research and Development Center under contract N00014-76-C-1063.

78 08 09 069

LEVEL II

Report Number: MISRC-TR-77-12

COMPUTERIZED FILE ORGANIZATION DESIGN SYSTEM (VERSION 2):

USER'S MANUAL

Salvatore T. March, Research Assistant
Dennis G. Severance, Principal Investigator

Management Information Systems Research Center
Graduate School of Business Administration
University of Minnesota
Minneapolis, MN 55455

1977 June

UNCLASSIFIED; DISTRIBUTION UNLIMITED

DDC
RECEIVED
AUG 10 1978
D

Prepared for and delivered to:

David W. Taylor Naval Ship Research and Development Center
ATTN: Code 1821, Dr. David K. Jefferson
U.S. Department of the Navy
Bethesda, MD 20084

The research reported here was funded by the David W. Taylor Naval Ship Research and Development Center under contract N00014-76-C-1063.

78 08 09 069

UNCLASSIFIED, DISTRIBUTION UNLIMITED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(9) Final rept. 1 Sep 76-30 Sep 77

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

(14) 1. REPORT NUMBER: MISRC-TR-77-12

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

(6) 4. TITLE (and Subtitle): Computerized File Organization Design System (Version 2): User's Manual

5. TYPE OF REPORT & PERIOD COVERED: FINAL REPORT 9/1/76 - 9/30/77

6. PERFORMING ORG. REPORT NUMBER

(10) 7. AUTHOR(s): Salvatore T. March, Research Assistant; Dennis G. Severance, Principal Investigator

(15) 8. CONTRACT OR GRANT NUMBER(s): N0014-76-C-1063

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Management Information Systems Research Center, 269 19th Ave. So.; University of Minnesota, Minneapolis, MN 55455

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

(11)

11. CONTROLLING OFFICE NAME AND ADDRESS: David W. Taylor Ship Research & Development Center; ATTN: Code 1821, Dr. David K. Jefferson, U.S. Department of the Navy; Bethesda, MD 20084

12. REPORT DATE: 1977 June 77

13. NUMBER OF PAGES: 15

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report): Unclassified (12) 13p.

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution Unlimited

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Automatic Database Design, File Structure Evaluation, Analytic Model, Segmented Record

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This The Computerized File Organization Design System is one of a set of database design aids developed at the Management Information Systems Research Center at the University of Minnesota. This program is used to automatically generate an efficient database design for a single record type in a multi-user environment. Given a quantitative problem specification, an efficient combination of access paths is determined for a collection of segmented records. This document describes the use of the Design System.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-014-6601

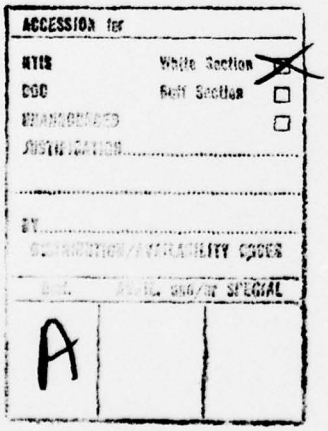
UNCLASSIFIED, DISTRIBUTION UNLIMITED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

409 153

TABLE OF CONTENTS

FIGURES



1. INTRODUCTION

The computerized database design system contains two optimizing design procedures:

(1) Record Segmentation Optimization, and

(2) File Organization Optimization.

Duhne [1, 2, 3] developed a file organization design procedure which is capable of automatically selecting an efficient file organization for database design problems in which multiple database users retrieve subsets of entire data records from a single set of stored data records. Data records, however, are commonly composed of both "high usage" and "low usage" data items (it is generally believed that eighty percent of all database retrieval is directed at only twenty percent of the stored data items). For operating efficiency, therefore, it is desirable to partition the data items stored within each record into a primary record segment retrieval by all users and a secondary record segment retrieval at extra cost only when required. Eisner and Severance [4] developed a procedure to determine a record segmentation which minimizes the total system operating cost.

In the computerized database design system, Duhne's file organization design procedure is modified to solve database design problems with a given record segmentation. This record segmentation is either specified by an analyst or determined by the record segmentation procedure.

The operation of the database design system is illustrated in Figure 1. Initially a problem is defined for both the record segmentation procedure (see [5]) and the file organization design procedure (see [2]). Care must be taken to insure that the problem definition's are consistent.* The record segmentation procedure is then invoked and an analyst may either define a record segmentation or use the algorithm to determine an efficient record segmentation for the defined problem.

---

*In a later version of the database design system the problem definitions for each procedure will be obtained from a single problem specification.

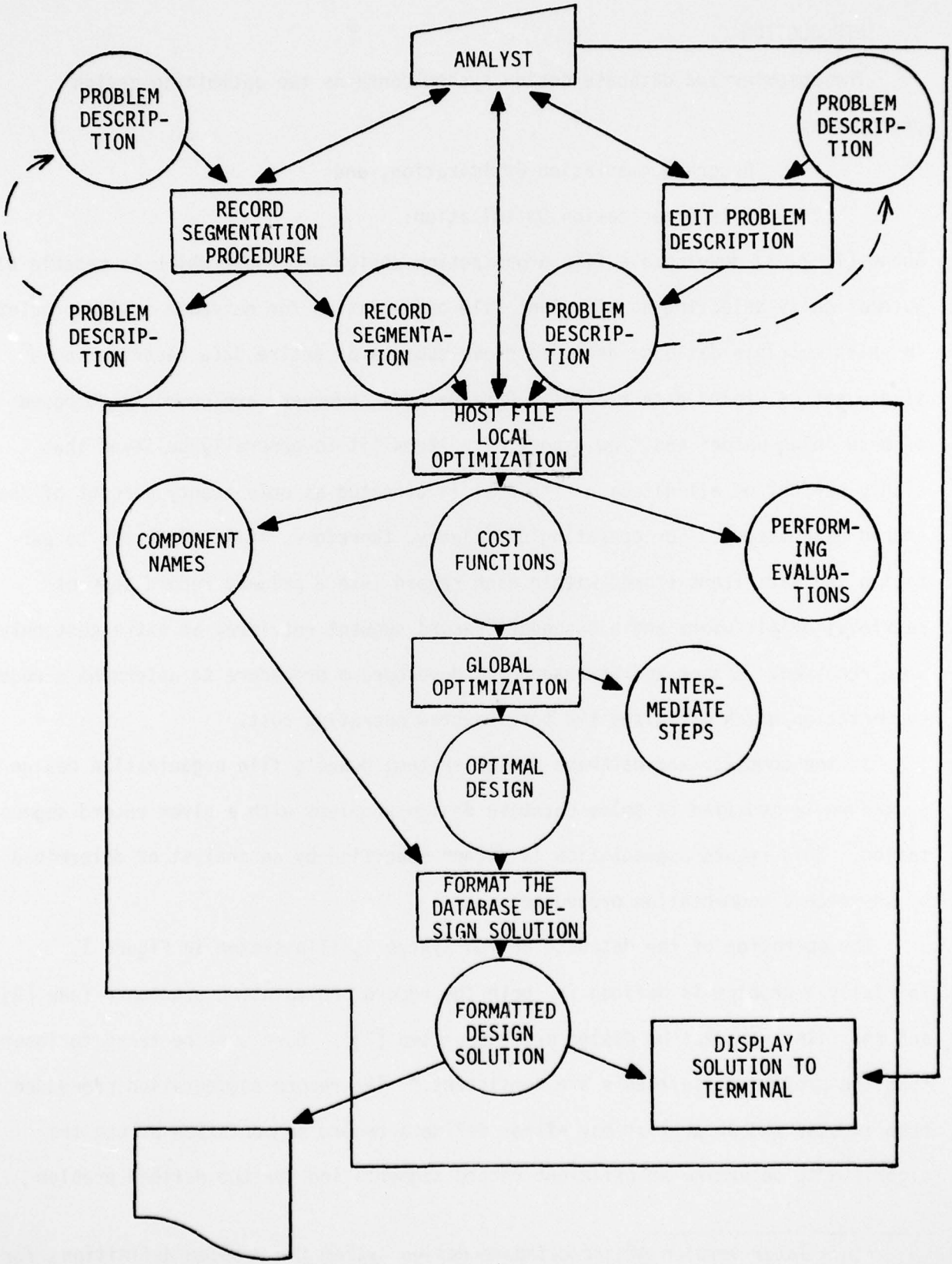


FIGURE 1. COMPUTERIZED DATABASE DESIGN SYSTEM

The selected segmentation is passed to the file organization design procedure (via a file called DBSEG) which is invoked next. This procedure consists of four subprocedures and six data files as illustrated in Figure 1. The selected database design may be displayed at a terminal or output to a line printer. After the results have been examined, the analyst may select a different record segmentation or modify the problem definitions and reinvoke the design procedure. The following section describes the user interaction with the design procedure.

2. AN EXAMPLE OF USER DIALOG.

This section presents an annotated user dialog to illustrate the use of the database design system. In the example a user first retrieves a previously stored record segmentation problem definition (see [5]) and evaluates a user defined segmentation. The record segmentation algorithm is then invoked to determine an optimal record segmentation for the design problem. This segmentation is saved for the host file optimization procedure and the record segmentation procedure is ended.

The user then invokes the file organization design procedure, again with a previously stored problem definition (corresponding to the problem analyzed by the record segmentation algorithm). An optimal file organization design is determined given the record segmentation determined by the record segmentation procedure.

In the annotated user dialog presented below capital letters designate user-computer dialog. Lines typed by the user begin with either a question mark or a slash (supplied as a prompting character by the system).* The user terminates an input line with a carriage return (or ETX, depending on the computer system).

The present dialog corresponds to a CDC CYBER 74, KRONOS operating system.

In order to use the design system, a user initially logs on through an operating system and enters procedure call mode (BATCH mode in the KRONOS operating system); the user types BATCH, the system responds

```
$RFL,20000.                (The system acknowledges the command.)
/-DBSYST                   (The procedure -DBSYST readies the computer system
                            to accept other database system commands.)
/-DBDSGN                   (The user calls the database design procedure.)
```

```
 RECORD SEGMENTATION ALGORITHM
 BEGIN ALL RESPONSES IN COLUMN 1                    (The record segmentation
 REQUEST BRIEF MODE (YES / NO)                       procedure is entered)
? YES
 REQUEST PREVIOUSLY STORED PROBLEM (YES / NO)
? YES

 REQUEST DISPLAY OF INPUT   (YES / NO)
? NO

 REQUEST CORRECTIONS TO INPUT (YES / NO)
? NO
 REQUEST PROBLEM STORED (YES / NO)
? NO
 REQUEST BLOCKED PRIMARY FILE (YES / NO)
? YES
 ENTER BUFFER SIZE (CHARACTERS)                     (NOTE this buffer size must be
? 317Ø                                               the same as the length of a memory
                                                     block for the data area in the
                                                     file organization problem description.)
```

```
 ENTER SEGMENTATION OPTION  OPTIMAL / PARAMETRIC / BOTH / DEFINE / END
? DEFINE                              (A segmentation is defined by the user.)
  ENTER RECORD SEGMENTATION VECTOR     (1 =  PRIMARY, 2 =  SECONDARY)
? 1 1 1 1 2 2 2 2 1 2 1 2 1 2 1 2 1  1 1 2 2 2 2 1 2 1 1 2 1 1


 FOR C   .GE.         Ø                                  (The user defined segmentation
                                                         is automatically evaluated and
 DATA ELEMENTS IN PRIMARY                                displayed.)
          1- 4, 9,11,13,15,17-19,24,26-27,29-3Ø

 SUBSET OF DATA ITEMS DELETED
          5- 8,1Ø,12,14,16,2Ø-23,25,28

 USERS WITH POSITIVE PROBABILITY OF BEING DISSATISFIED
          1-18,2Ø
 REQUEST DISPLAY OF PROBABILITIES FOR ABOVE
  DISSATISFIED USERS (YES / NO)
? NO
 COST =       1532.96
 ENTER SEGMENTATION OPTION  OPTIMAL / PARAMETRIC / BOTH / DEFINE / END
? OPTIMAL                                               (An optimal segmentation
                                                         is determined.)

 OPTIMAL SOLUTION FOUND


 FOR C   .GE.     .Ø143                                  (The optimal segmentation
                                                         is evaluated and displayed.)
 DATA ELEMENTS IN PRIMARY
          1- 4, 9,11,13,15,24,26-27

 SUBSET OF DATA ITEMS DELETED
         17-19,29-3Ø

 USERS WITH POSITIVE PROBABILITY OF BEING DISSATISFIED
          1- 2, 4,12,17-18,2Ø
 REQUEST DISPLAY OF PROBABILITIES FOR ABOVE
  DISSATISTIED USERS (YES / NO)
? NO
 COST =       336.38

 REQUEST SENSITIVITY ANALYSIS (YES / NO)
? NO
  SEGMENTATION OPTION FOR HOST FILE OPTIMIZATION    CURRENT / DEFINE
? CURRENT                                           (The current segmentation, i.e.,
                                                    the optimal, is saved for the file
                                                    organization evaluation.)

  REQUEST RECORD SEGMENTATION ALGORITHM FOR A NEW
  PROBLEM ( YES / NO)
 ? NO
  RECORD SEGMENTATION ALGORITHM ENDED
  STOP
```

(The file organization design procedure is automatically entered.)

REQUEST AUTOMATIC GENERATIONS OF PARAMETERS FOR ALL HOST FILES (YES/NO)

?YES (The user wants automatic optimization)

ENTER RECORD SEGMENTATION OPTION
UNSEGMENTED / FROM SEGMENTATION MODULE

?FROM SEGMENTATION MODULE (The record segmentation previously saved is used in the design optimization. Optionally the user could have selected UNSEGMENTED records for the design.)

(The system then outputs some information as the optimization procedure progresses. First the host file names, then minimum cost in each interaction, finally, when the solution is reached it is stored in file DBSOLN, and a display program is called to permit the user to analyze the solution. The system outputs page zero:)

ENTER PAGE NUMBER YOU WANT / (CARRIAGE-RETURN) FOR NEXT PAGE / (Q) TO QUIT

PAGE 0.NEXT:

? (Carriage-Return) (The user wants to see the first page) (The first part of the solution summary is displayed. The output is divided into "pages" which are not larger than the size of a CRT display. The bottom two lines read:)

PAGE L. NEXT: (The system indicates that the displayed page is number one and asks for the next page to be displayed)

? (Carriage-return) (After analyzing the displayed page, the user wants the next page)

(Page two is displayed. The user continues analyzing the next pages. The user is allowed to jump among pages. Let's say, for example, that the current page is page 5, and the user wants to go back to page 1, he types:)

PAGE 5. NEXT:

?1 (The user wants page 1 to be displayed next)

(Page 1 is then displayed again. When the user finishes analyzing solution he types:)

PAGE 1. NEXT:

?Q (Quit: end display)

WANT TO QUIT (YES/NO) (The system asks again)

?YES

ENDISPLAY (The system leaves the display program).

See Duhne [2] for a detailed description of the output produced and additional details of the file organization design procedure. See [5] and Eisner and Severance [4] for details of the record segmentation procedure. System documentation and complete listings of all programs used in this database design procedure are found in [6].

REFERENCES

[1] Duhne, Ricardo A. Optimal Design of a Generalized File Organization, Ph.D. Dissertation, Cornell University, 1977.

[2] Duhne, Ricardo A. Computerized File Organization Design System: User's Manual, Management Information Systems Research Center Technical Report 77-02, University of Minnesota, 1976.

[3] Duhne, Ricardo A. Computerized File Organization Design System: System Documentation, Management Information Systems Research Center Technical Report 77-03, University of Minnesota, 1976.

[4] Eisner, M.J. and Severance, D.G. "Mathematical Techniques for Efficient Record Segmentation in Large Shared Databases," JACM 23,4 October 1976.

[5] Record Segmentation Analyzer: User's Manual, Management Information Systems Research Center Technical Report 77-08, University of Minnesota, 1977.

[6] March, Sal Computerized File Organization Design System (Version 2): System Documentation, Management Information Systems Research Center Technical Report 77-13, University of Minnesota, 1977.