A Relational Data Base Implemented Using MBASIC

R. M. Smith
DSN Facility Operations Office

The implementation of a relational data base in MBASIC is described. Five relational operations are implemented, and their use on two data bases is described.

I. Introduction

The relational model of data is presently state-of-the-art in data base technology. The relational approach offers a simply understood, generalized data model that allows a relatively easy attainment of security, integrity, and privacy controls as well as implementation of the usual query and update transaction capability. The model also allows relatively easy expansion and reorganization. For these reasons, the relational model is an attractive candidate for the DSN Facility Operations data base. This article describes initial efforts at exploring the features of the relational model. Included in the scope of this effort was the construction of a demonstration relational data base and the implementation in MBASIC of several relational data base manipulation operations. The capability to manipulate the data base with the operations was demonstrated.

II. Relational Structure

In the relational model, data are organized into arrays (called relations) with fields (called domains), so that each record entry (called a tuple) is essentially a set of attributes describing one or more characteristics of a real-world entity. For instance, the relation "equipment identification" might be described by the following set of domain names:

CONTROL NUMBER	MANUFACTURER CODE	MODEL	USAGE CATEGORY	PROPERTY NUMBER

Adding domain values would record the identification of equipment:

CONTROL NUMBER	MANUFACTURER CODE	MODEL	USAGE CATEGORY	PROPERTY NUMBER
AA1A23	HEA	5245L	Ø3	J270(F)12345
AB5CØ1	TEA	547	Ø7	LA546785
AA2B47	EPI	100	Ø2	NONE

Other relations may be constructed to describe other relationships such as equipment location, service history, etc.

III. Relational Operations

There are several specific relational operations (described in Refs. 1 –3) that are used to manipulate data in the data model:

(1) Join

(2) Restriction

(3) Division

(4) Projection

(5) Permutation

The reader may refer to the references for detailed descriptions of the relational operations; the following brief description may suffice for this reading:

(1) *Join.* Two relations that have at least one domain in common may be combined in a way that preserves all the information in both relations but does not repeat the duplicate domain values in the resultant tuples. The join is equivalent in action to a union of sets.

(2) *Restriction.* Tuples may be selected from a relation by testing each tuple for the presence of a specified attribute.

(3) *Division.* Tuples may be selected from a relation by testing each tuple for the presence of a specified constant value. Tuples so selected are listed or stored, minus the domain containing the constant value, as though the selected tuples were remainders in a division process.

(4) *Projection.* Specific domains in a relation may be selected (striking out the others) and duplicate tuples removed. The resultant array is a projection of the original relation.

(5) *Permutation.* Domain positions in a relation may be interchanged, resulting in a permutation of the relation.

IV. Examples of the Relational Operations

Figures 1, 2 and 3 use flow charts derived from MBASIC programs described in this article (see *MBASIC, Vol. 1: Fundamentals,* and *MBASIC, Vol. 2: Appendices,* Jet Propulsion Laboratory, 1973) to illustrate the relational operations: join, restriction, and division. The projection operation is inherent in each of the relational operations when either "PRINT" or "WRITE ON" is used. Permutation is not illustrated separately but occurs in the program 'REFERENCE' when the record domains are reordered in the "PRINT" operation.

V. Description of the Data Base

The data base consists of two parts. The first part is purely for demonstration and consists of five MBASIC files containing equipment data to be manipulated and one file that serves as a directory. There is one program that may be used to demonstrate relational operations on the files. Figure 4 illustrates the domain assignments for each demonstration relation. Figure 5 contains a flow diagram of the demonstration program 'RELATIONAL'. Figure 6 illustrates the domain assignment for the directory relation.

The program 'RELATIONAL' performs a relational operation using relations, domains, and domain values selected by the user. In performing the selected operation, the program uses the relational operators as intermediate steps in manipulating the selected relations and domains.

The second part is a working data base that records a collection of technical articles and the subjects and key words contained in the articles. There are three MBASIC programs written especially for using the data base:

(1) 'REFERENCE': Locates and prints the author and title of all articles that reference a subject or key word selected by the user.

(2) 'INPUTREF': Maps into storage new articles and subjects or key words.

(3) 'TERNARY': Prints a list of subjects and key words without printing duplications.

Figure 7 illustrates the domain assignment for the working part of the data base. Figures 8, 9, and 10 are flow charts for the special MBASIC programs.

VI. Details of the Implementation and Functions

There are several "traits" that are desirable in a relational data base, and the following were specifically chosen for inclusion in the described data base:

(1) Record order independence. Application programs should generally not be dependent upon stored order of the records.

(2) Use of domain names rather than domain positions. The user should not have to know the positions of domains.

(3) Deletion and addition of relations. These should not affect application programs.

For this implementation, no effort was made to produce a data sublanguage (Ref. 3); therefore, all transactions are carried on in a full prompting mode, and the user deals with individual programs.

Figures 11 and 12 contain, respectively, the content of the demonstration files and the content of the working files. Figure 13 illustrates a session at a terminal using the demonstration program 'RELATIONAL' to produce the join of two relations over a common domain and the restriction of a relation. Figure 14 illustrates a session at a terminal using the working program 'REFERENCE' to produce a list of articles on a selected subject.

VII. Conclusions

The capability for exclusive use of domain names was fully realized in the program entitled 'RELATIONAL'. Record order independence is observed in all relations but one: 'SUBJECTS'. For practical reasons, it was decided to store 'SUBJECTS' in a particular, sorted order rather than to write a program containing a sort routine. The ability to add and delete relations was fully realized in the demonstration program ('RELATIONAL') and associated directory file ('RREL'). All of the previously listed relational operations take place during the operation of the special MBASIC programs referenced in this article and are annotated on the 'RELATIONAL' flow diagram.

References

1. Codd, E. F., "A Relational Model of Data for Large Shared Data Banks," *Communications of ACM,* Vol. 13, No. 6, June 1970.
2. Codd, E. F., "Normalized Data Base Structure: A Brief Tutorial," *Proceedings of 1971 ACM SIGFIDET Workshop on Data Description, Access and Control,* ACM, New York, 1971.
3. Date, C. J., *An Introduction to Database Systems,* Addison-Wesley, 1975.

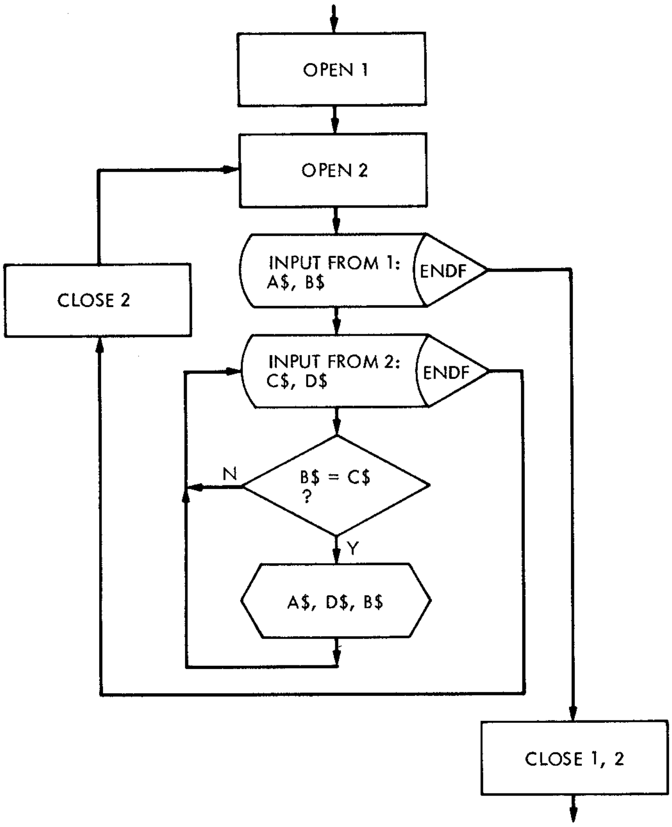


Fig. 1. Flow chart for the relational operation "Join"

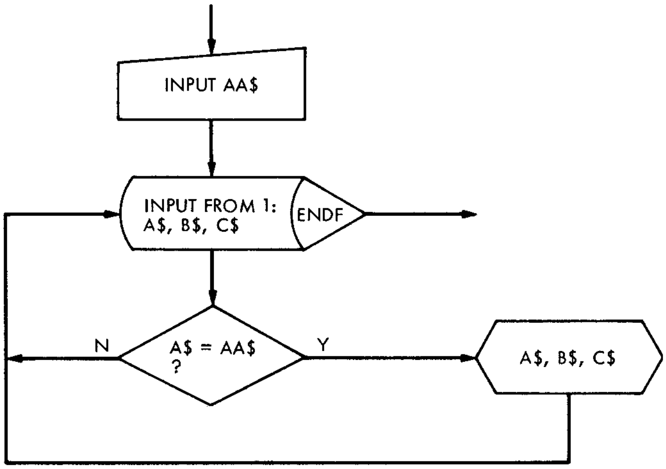


Fig. 2. Flow chart for the relational operation "Restriction"

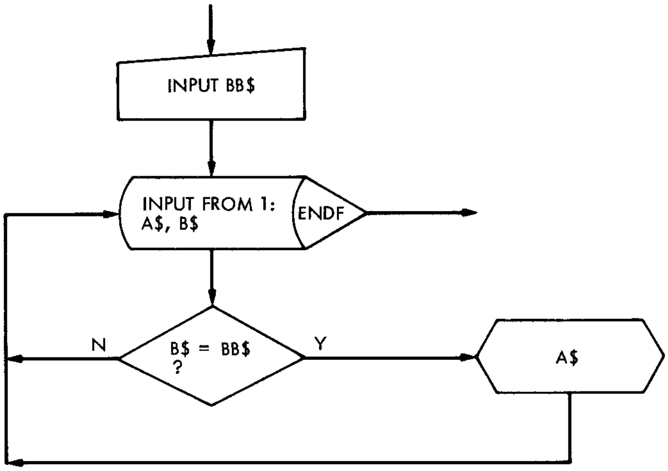
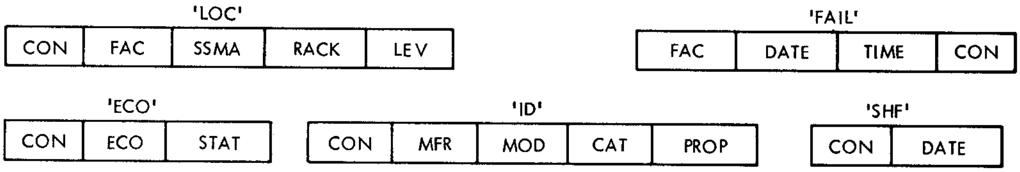


Fig. 3. Flow chart for the relational operation "Division"



'LOC': CON | FAC | SSMA | RACK | LEV

'FAIL': FAC | DATE | TIME | CON

'ECO': CON | ECO | STAT

'ID': CON | MFR | MOD | CAT | PROP

'SHF': CON | DATE

CON	=	DSN CONTROL NUMBER
FAC	=	FACILITY
SSMA	=	SUBSYSTEM AND MAJOR ASSEMBLY
RACK	=	RACK
LEV	=	LEVEL
ECO	=	ENGINEERING CHANGE ORDER
STAT	=	STATUS
DATE	=	DATE
TIME	=	TIME
MFR	=	MANUFACTURER THREE LETTER CODE
MOD	=	MODEL
CAT	=	CATEGORY
PROP	=	PROPERTY NUMBER

'LOC':	EQUIPMENT LOCATION DATA
'FAIL':	EQUIPMENT FAILURE HISTORY DATA
'ECO':	EQUIPMENT DESIGN DATA
'ID':	EQUIPMENT IDENTIFICATION DATA
'SHF':	EQUIPMENT SERVICE HISTORY DATA

Fig. 4. Demonstration relations and domains

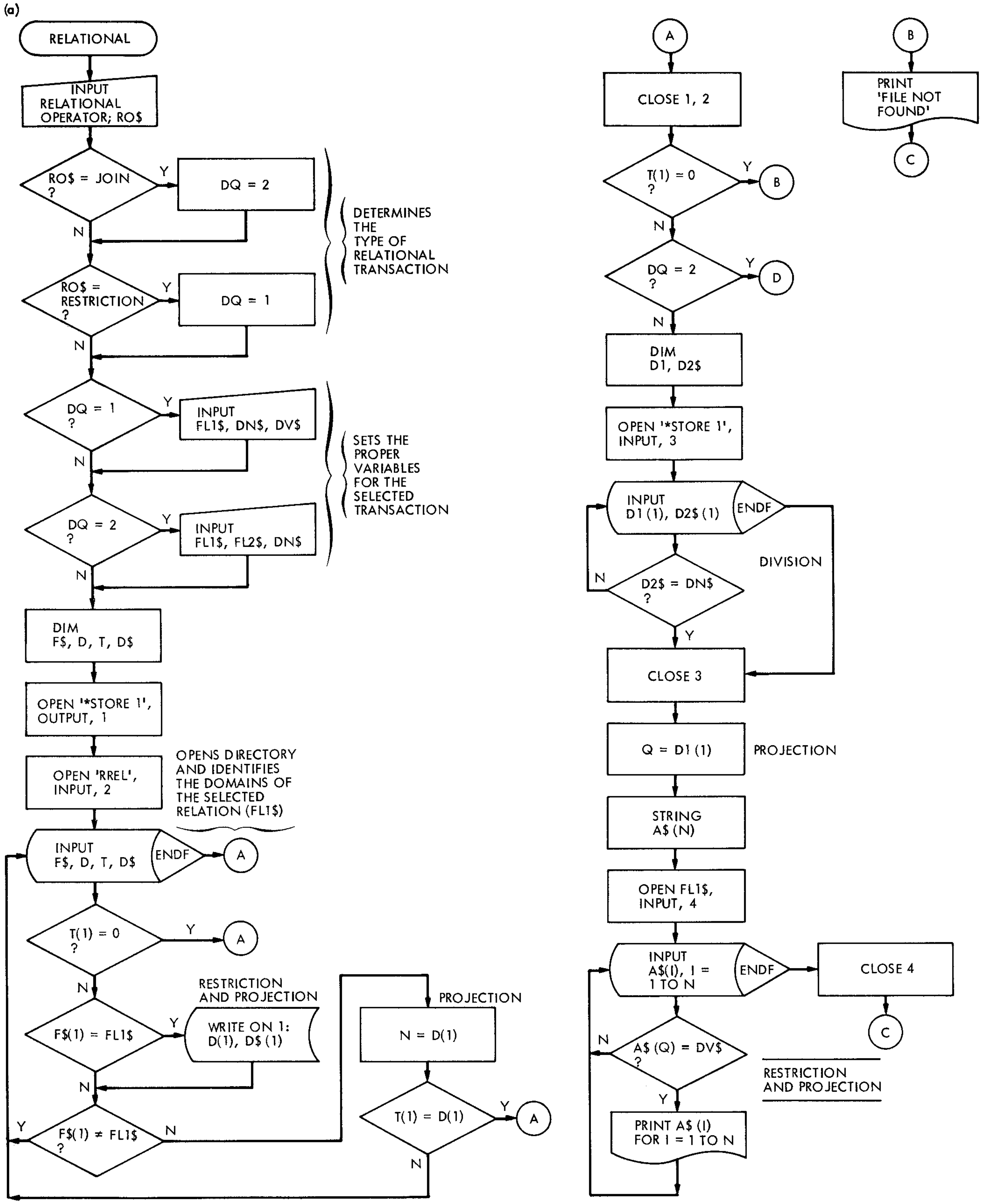


Fig. 5. Flow diagram for the MBASIC demonstration program 'RELATIONAL'

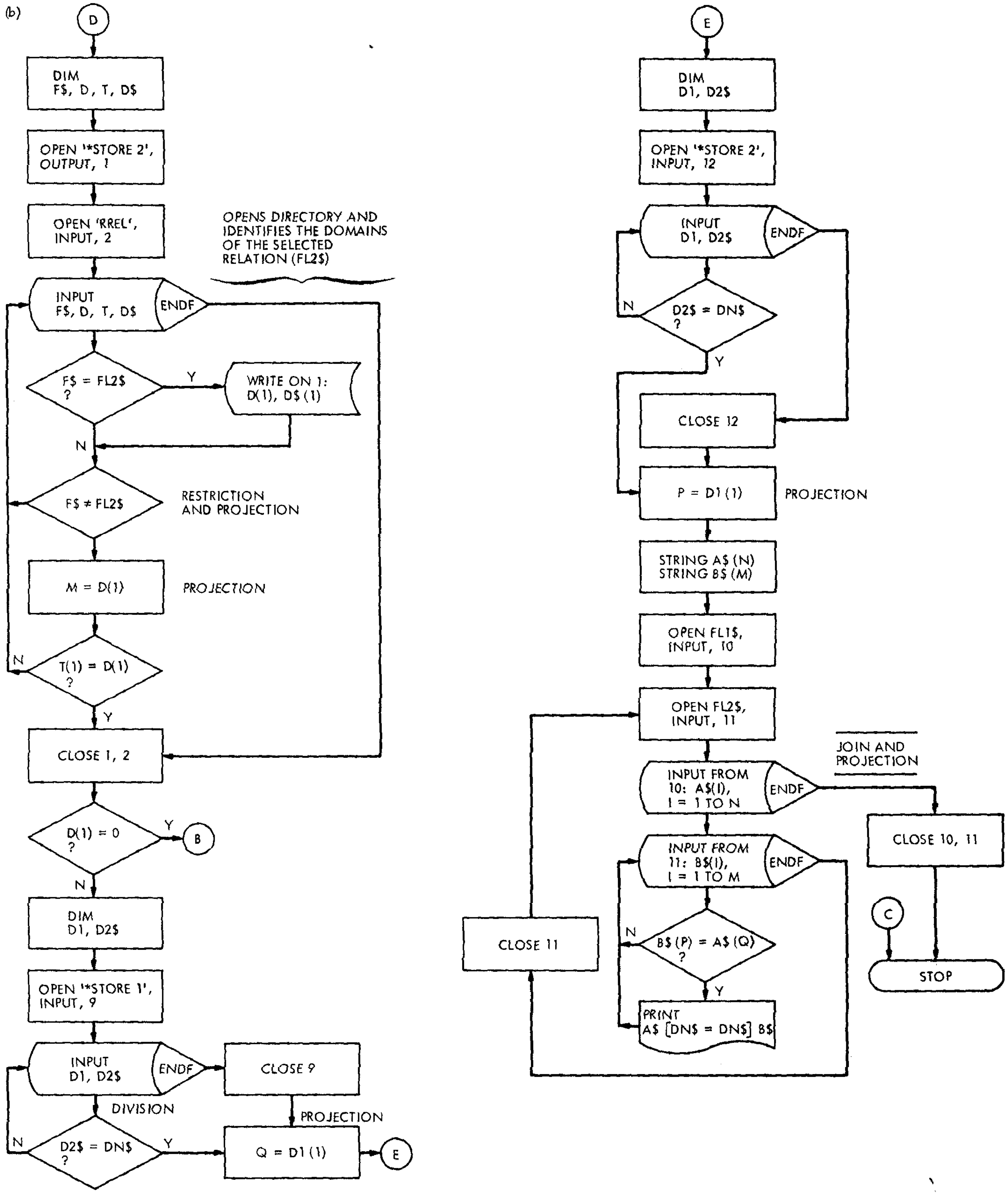

(b)
D
DIM
F$, D, T, D$
OPEN '*STORE 2',
OUTPUT, 1
OPEN 'RREL',
INPUT, 2
OPENS DIRECTORY AND
IDENTIFIES THE DOMAINS
OF THE SELECTED
RELATION (FL2$)
INPUT
F$, D, T, D$
ENDF
F$ = FL2$
?
Y
WRITE ON 1:
D(1), D$ (1)
N
F$ ≠ FL2$
RESTRICTION
AND PROJECTION
M = D(1)
PROJECTION
N
T(1) = D(1)
?
Y
CLOSE 1, 2
D(1) = 0
?
Y
B
N
DIM
D1, D2$
OPEN '*STORE 1',
INPUT, 9
INPUT
D1, D2$
ENDF
CLOSE 9
DIVISION
PROJECTION
N
D2$ = DN$
?
Y
Q = D1 (1)
E
E
DIM
D1, D2$
OPEN '*STORE 2',
INPUT, 12
INPUT
D1, D2$
ENDF
N
D2$ = DN$
?
Y
CLOSE 12
P = D1 (1)
PROJECTION
STRING A$ (N)
STRING B$ (M)
OPEN FL1$,
INPUT, 10
OPEN FL2$,
INPUT, 11
JOIN AND
PROJECTION
INPUT FROM
10: A$(I),
I = 1 TO N
ENDF
CLOSE 10, 11
INPUT FROM
11: B$(I),
I = 1 TO M
ENDF
C
CLOSE 11
N
B$ (P) = A$ (Q)
?
Y
STOP
PRINT
A$ [DN$ = DN$] B$

Fig. 5. (contd)

'RREL'

FILENAME	DOMAIN NUMBER	QUANTITY OF DOMAINS	DOMAIN NAME

FILENAME = MBASIC NAME OF A RELATION FILE

DOMAIN NUMBER = THE DOMAIN POSITION IN THE RELATION

QUANTITY OF DOMAINS = TOTAL QUANTITY OF DOMAINS IN THE RELATION

DOMAIN NAME = THE NAME ASSIGNED TO THE SPECIFIC DOMAIN

Fig. 6. Domain assignment for the directory relation 'RREL'

'SUBJECTS'

REF #	SUBJECTS

'ARTICLES'

REF #	TITLE	AUTHOR

REF # = A UNIQUE FILE NUMBER ASSIGNED TO AN ARTICLE

SUBJECT = A SUBJECT NAME OR A KEY WORD

TITLE = THE TITLE OF AN ARTICLE

AUTHOR = THE NAME OF THE ARTICLE'S AUTHOR

Fig. 7. Working relations and domains

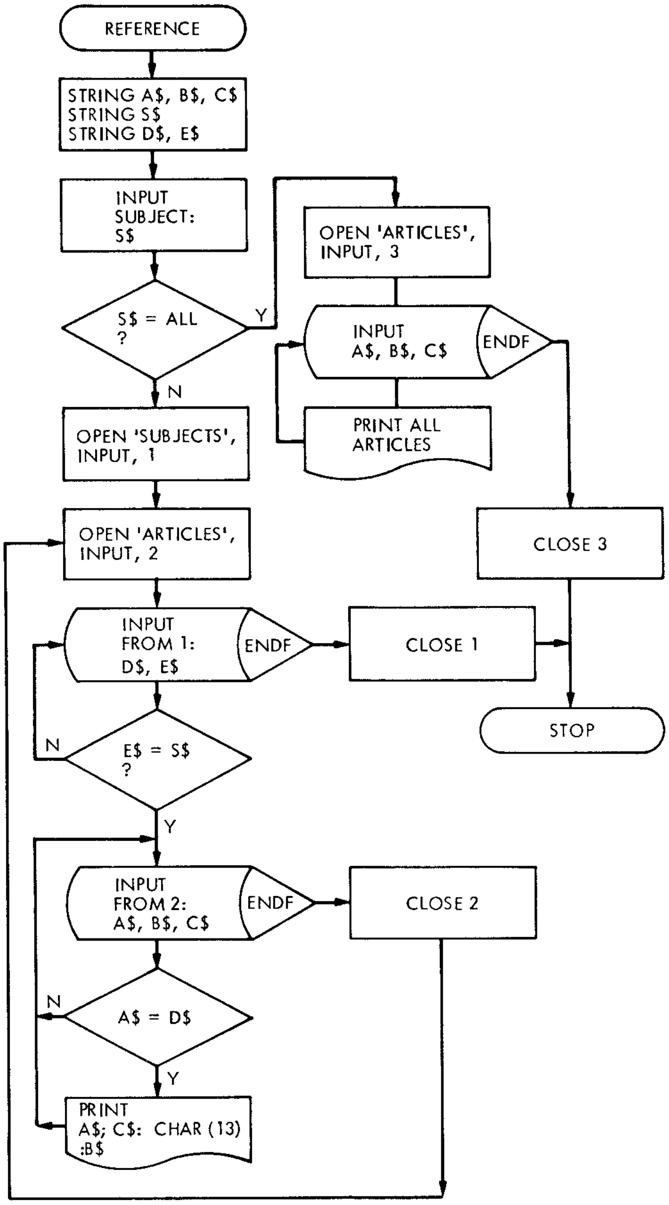


Fig. 8. Flow diagram for the MBASIC working program 'REFERENCE'

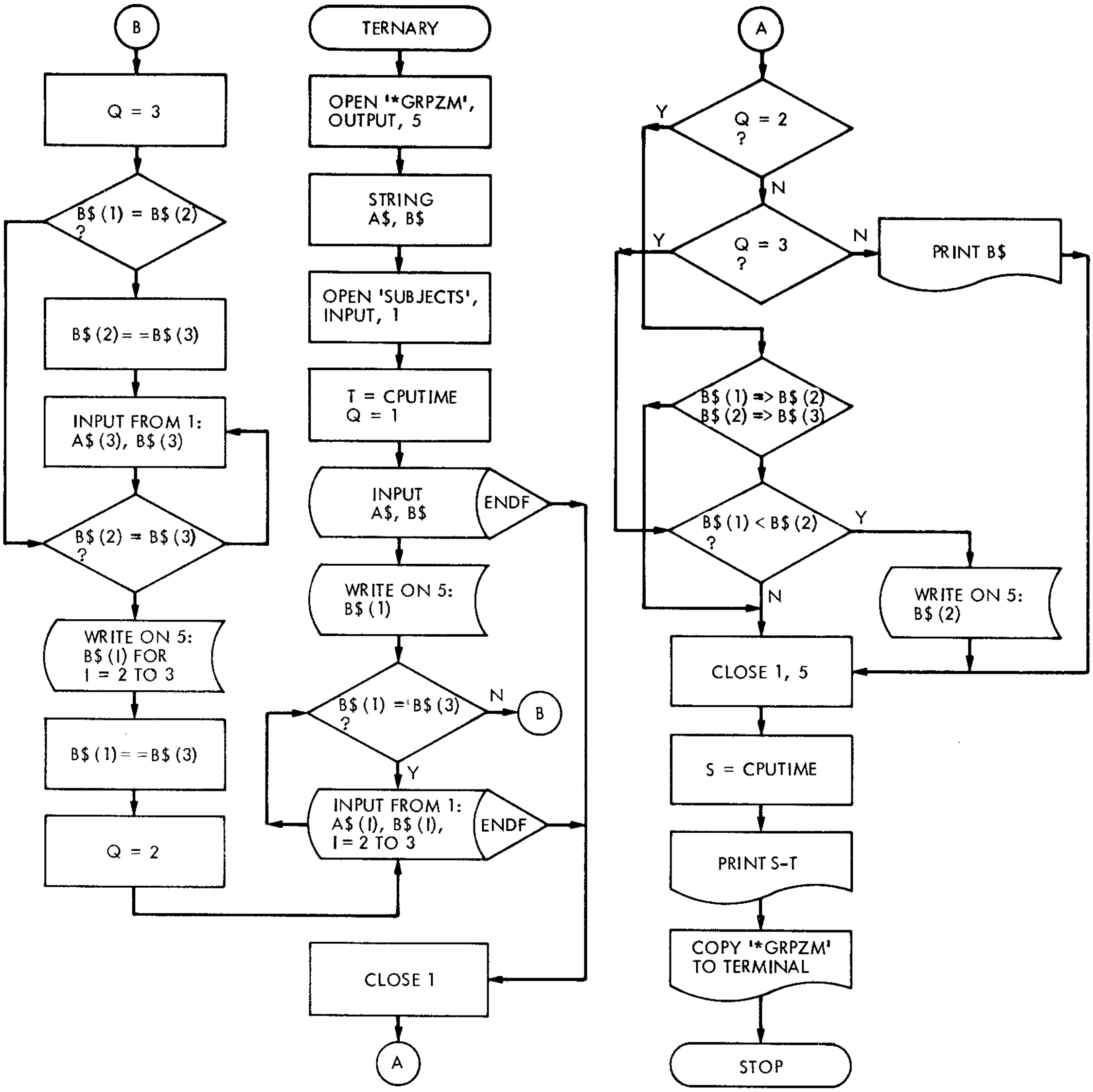


Fig. 9. Flow diagram for the MBASIC working program 'TERNARY'

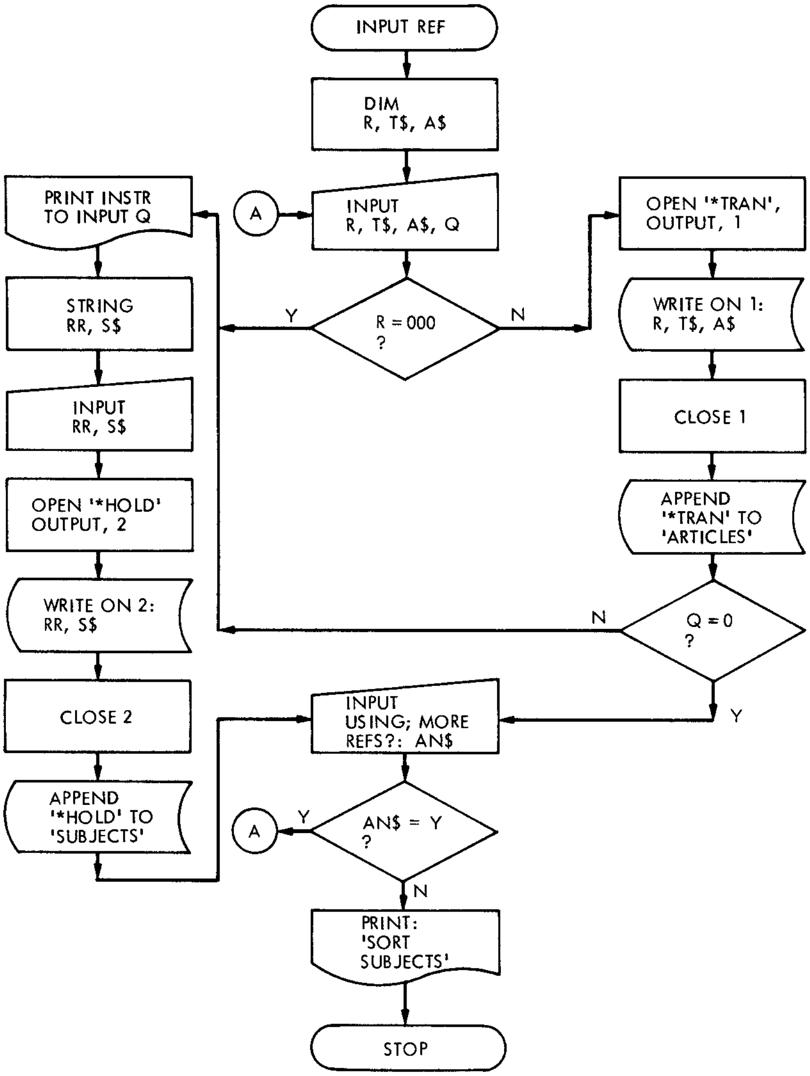
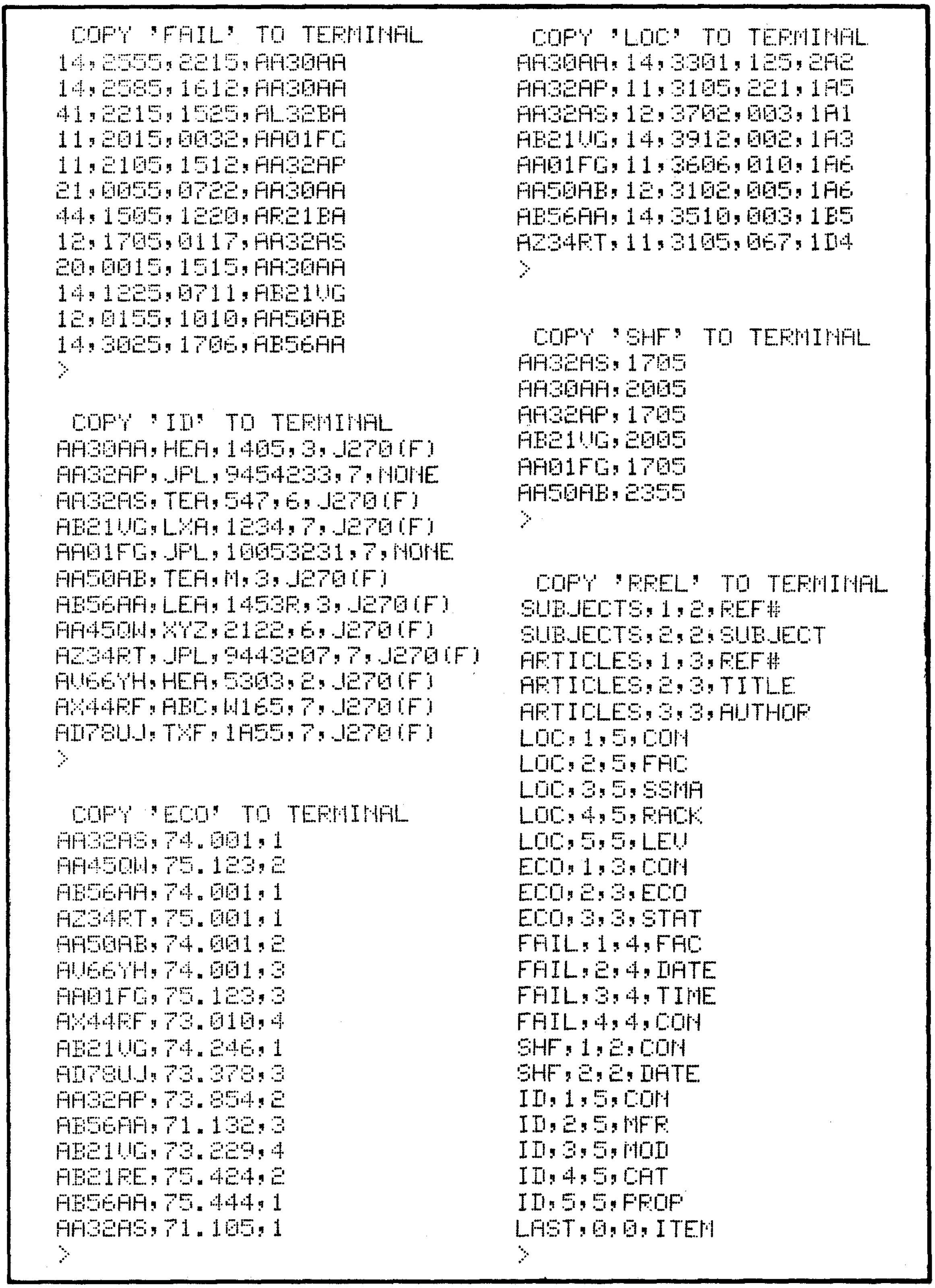


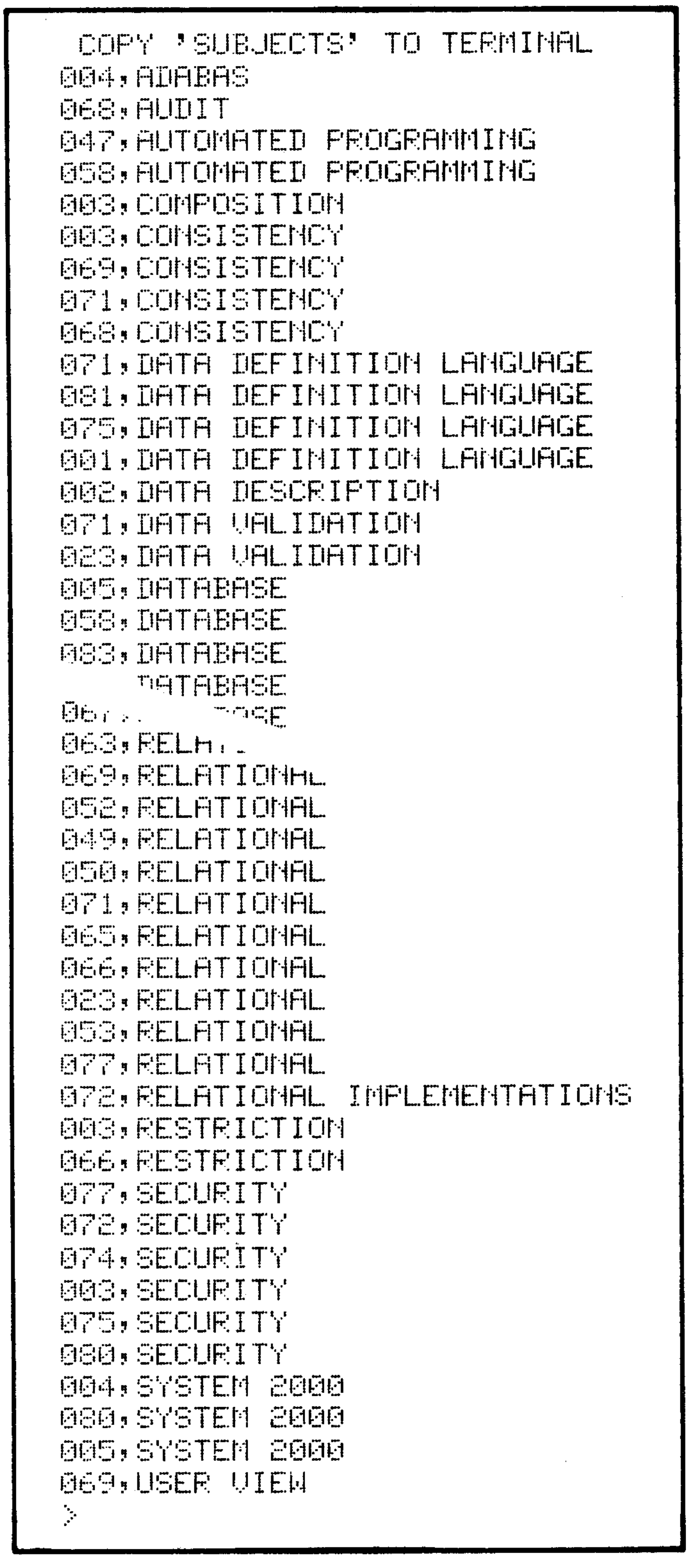
Fig. 10. Flow diagram for the MBASIC working program 'INPUTREF'



```
 COPY 'FAIL' TO TERMINAL
14,2555,2215,AA30AA
14,2585,1612,AA30AA
41,2215,1525,AL32BA
11,2015,0032,AA01FG
11,2105,1512,AA32AP
21,0055,0722,AA30AA
44,1505,1220,AR21BA
12,1705,0117,AA32AS
20,0015,1515,AA30AA
14,1225,0711,AB21VG
12,0155,1010,AA50AB
14,3025,1706,AB56AA
>

 COPY 'ID' TO TERMINAL
AA30AA,HEA,1405,3,J270(F)
AA32AP,JPL,9454233,7,NONE
AA32AS,TEA,547,6,J270(F)
AB21VG,LXA,1234,7,J270(F)
AA01FG,JPL,10053231,7,NONE
AA50AB,TEA,M,3,J270(F)
AB56AA,LEA,1453R,3,J270(F)
AA450W,XYZ,2122,6,J270(F)
AZ34RT,JPL,9443207,7,J270(F)
AV66YH,HEA,5303,2,J270(F)
AX44RF,ABC,W165,7,J270(F)
AD78UJ,TXF,1A55,7,J270(F)
>

 COPY 'ECO' TO TERMINAL
AA32AS,74.001,1
AA450W,75.123,2
AB56AA,74.001,1
AZ34RT,75.001,1
AA50AB,74.001,2
AV66YH,74.001,3
AA01FG,75.123,3
AX44RF,73.010,4
AB21VG,74.246,1
AD78UJ,73.378,3
AA32AP,73.854,2
AB56AA,71.132,3
AB21VG,73.229,4
AB21RE,75.424,2
AB56AA,75.444,1
AA32AS,71.105,1
>
```

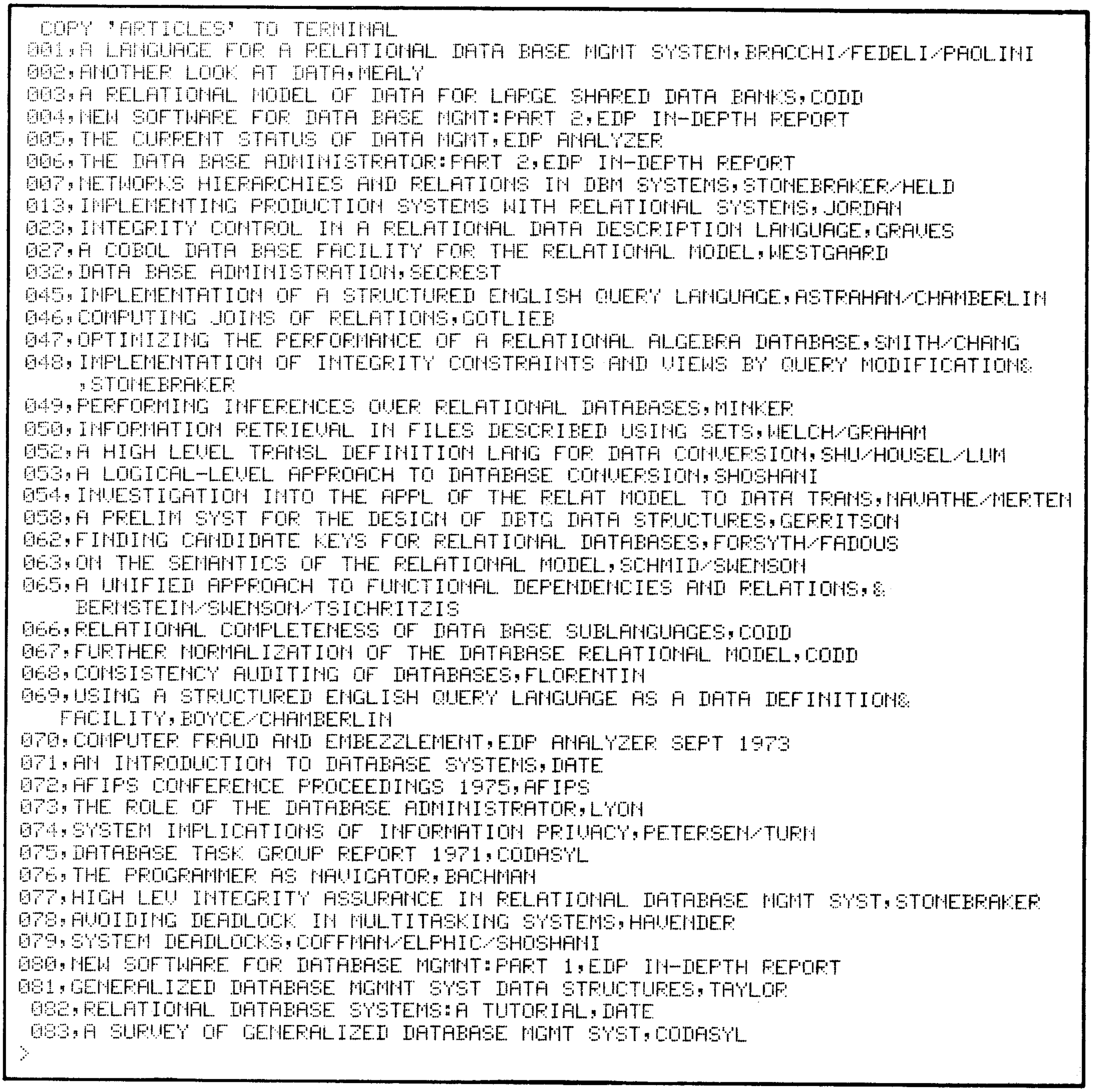
```
 COPY 'LOC' TO TERMINAL
AA30AA,14,3301,125,2A2
AA32AP,11,3105,221,1A5
AA32AS,12,3702,003,1A1
AB21VG,14,3912,002,1A3
AA01FG,11,3606,010,1A6
AA50AB,12,3102,005,1A6
AB56AA,14,3510,003,1B5
AZ34RT,11,3105,067,1D4
>

 COPY 'SHF' TO TERMINAL
AA32AS,1705
AA30AA,2005
AA32AP,1705
AB21VG,2005
AA01FG,1705
AA50AB,2355
>

 COPY 'RREL' TO TERMINAL
SUBJECTS,1,2,REF#
SUBJECTS,2,2,SUBJECT
ARTICLES,1,3,REF#
ARTICLES,2,3,TITLE
ARTICLES,3,3,AUTHOR
LOC,1,5,CON
LOC,2,5,FAC
LOC,3,5,SSMA
LOC,4,5,RACK
LOC,5,5,LEV
ECO,1,3,CON
ECO,2,3,ECO
ECO,3,3,STAT
FAIL,1,4,FAC
FAIL,2,4,DATE
FAIL,3,4,TIME
FAIL,4,4,CON
SHF,1,2,CON
SHF,2,2,DATE
ID,1,5,CON
ID,2,5,MFR
ID,3,5,MOD
ID,4,5,CAT
ID,5,5,PROP
LAST,0,0,ITEM
>
```

Fig. 11. Content of demonstration relations



```
 COPY 'SUBJECTS' TO TERMINAL
004,ADABAS
068,AUDIT
047,AUTOMATED PROGRAMMING
058,AUTOMATED PROGRAMMING
003,COMPOSITION
003,CONSISTENCY
069,CONSISTENCY
071,CONSISTENCY
068,CONSISTENCY
071,DATA DEFINITION LANGUAGE
081,DATA DEFINITION LANGUAGE
075,DATA DEFINITION LANGUAGE
001,DATA DEFINITION LANGUAGE
002,DATA DESCRIPTION
071,DATA VALIDATION
023,DATA VALIDATION
005,DATABASE
058,DATABASE
083,DATABASE
    [illegible]ATABASE
06[illegible]ASE
063,RELA[illegible]
069,RELATIONAL
052,RELATIONAL
049,RELATIONAL
050,RELATIONAL
071,RELATIONAL
065,RELATIONAL
066,RELATIONAL
023,RELATIONAL
053,RELATIONAL
077,RELATIONAL
072,RELATIONAL IMPLEMENTATIONS
003,RESTRICTION
066,RESTRICTION
077,SECURITY
072,SECURITY
074,SECURITY
003,SECURITY
075,SECURITY
080,SECURITY
004,SYSTEM 2000
080,SYSTEM 2000
005,SYSTEM 2000
069,USER VIEW
>
```

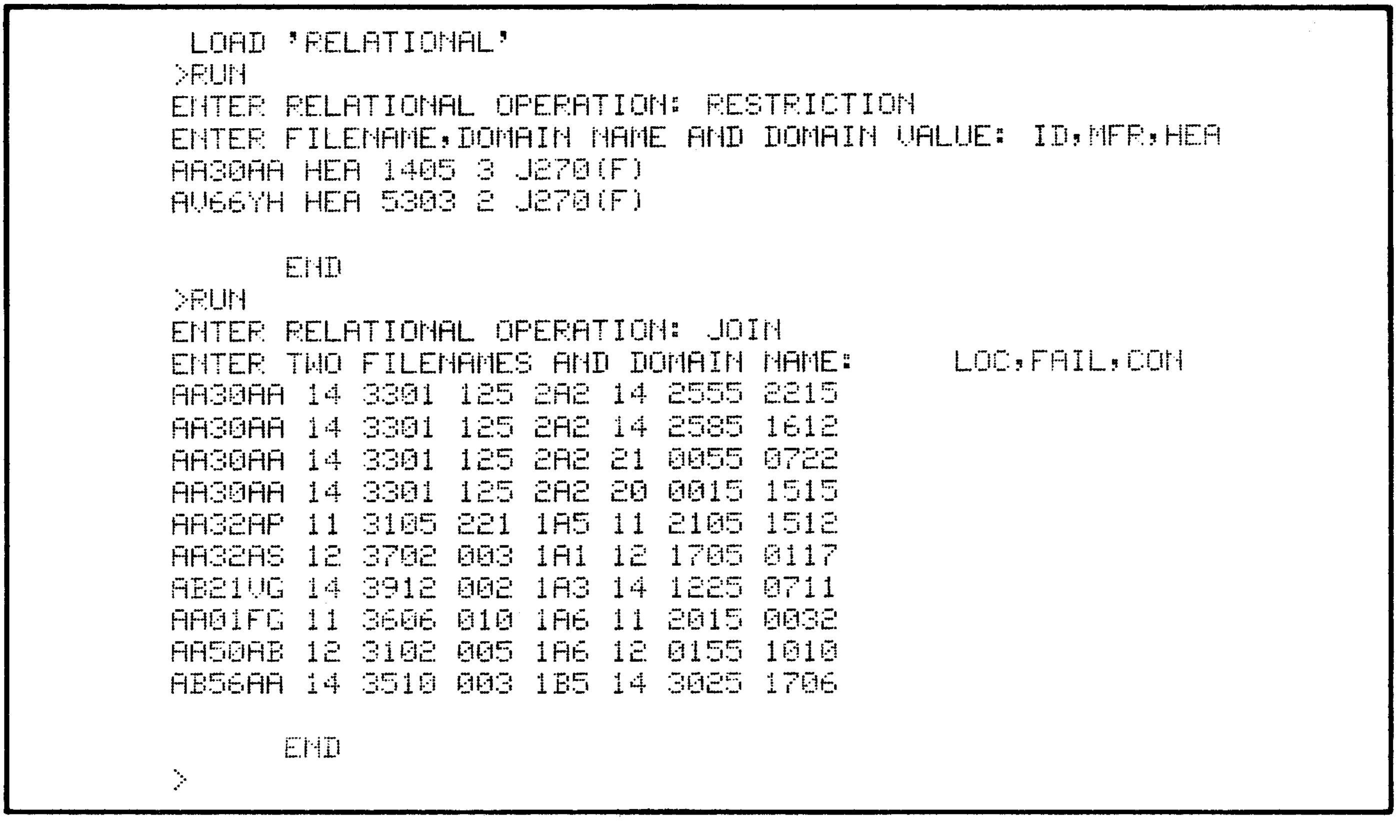
Fig. 12. Content of working relations



```
 COPY 'ARTICLES' TO TERMINAL
001,A LANGUAGE FOR A RELATIONAL DATA BASE MGMT SYSTEM,BRACCHI/FEDELI/PAOLINI
002,ANOTHER LOOK AT DATA,MEALY
003,A RELATIONAL MODEL OF DATA FOR LARGE SHARED DATA BANKS,CODD
004,NEW SOFTWARE FOR DATA BASE MGMT:PART 2,EDP IN-DEPTH REPORT
005,THE CURRENT STATUS OF DATA MGMT,EDP ANALYZER
006,THE DATA BASE ADMINISTRATOR:PART 2,EDP IN-DEPTH REPORT
007,NETWORKS HIERARCHIES AND RELATIONS IN DBM SYSTEMS,STONEBRAKER/HELD
013,IMPLEMENTING PRODUCTION SYSTEMS WITH RELATIONAL SYSTEMS,JORDAN
023,INTEGRITY CONTROL IN A RELATIONAL DATA DESCRIPTION LANGUAGE,GRAVES
027,A COBOL DATA BASE FACILITY FOR THE RELATIONAL MODEL,WESTGAARD
032,DATA BASE ADMINISTRATION,SECREST
045,IMPLEMENTATION OF A STRUCTURED ENGLISH QUERY LANGUAGE,ASTRAHAN/CHAMBERLIN
046,COMPUTING JOINS OF RELATIONS,GOTLIEB
047,OPTIMIZING THE PERFORMANCE OF A RELATIONAL ALGEBRA DATABASE,SMITH/CHANG
048,IMPLEMENTATION OF INTEGRITY CONSTRAINTS AND VIEWS BY QUERY MODIFICATION&
    ,STONEBRAKER
049,PERFORMING INFERENCES OVER RELATIONAL DATABASES,MINKER
050,INFORMATION RETRIEVAL IN FILES DESCRIBED USING SETS,WELCH/GRAHAM
052,A HIGH LEVEL TRANSL DEFINITION LANG FOR DATA CONVERSION,SHU/HOUSEL/LUM
053,A LOGICAL-LEVEL APPROACH TO DATABASE CONVERSION,SHOSHANI
054,INVESTIGATION INTO THE APPL OF THE RELAT MODEL TO DATA TRANS,NAVATHE/MERTEN
058,A PRELIM SYST FOR THE DESIGN OF DBTG DATA STRUCTURES,GERRITSON
062,FINDING CANDIDATE KEYS FOR RELATIONAL DATABASES,FORSYTH/FADOUS
063,ON THE SEMANTICS OF THE RELATIONAL MODEL,SCHMID/SWENSON
065,A UNIFIED APPROACH TO FUNCTIONAL DEPENDENCIES AND RELATIONS,&
    BERNSTEIN/SWENSON/TSICHRITZIS
066,RELATIONAL COMPLETENESS OF DATA BASE SUBLANGUAGES,CODD
067,FURTHER NORMALIZATION OF THE DATABASE RELATIONAL MODEL,CODD
068,CONSISTENCY AUDITING OF DATABASES,FLORENTIN
069,USING A STRUCTURED ENGLISH QUERY LANGUAGE AS A DATA DEFINITION&
   FACILITY,BOYCE/CHAMBERLIN
070,COMPUTER FRAUD AND EMBEZZLEMENT,EDP ANALYZER SEPT 1973
071,AN INTRODUCTION TO DATABASE SYSTEMS,DATE
072,AFIPS CONFERENCE PROCEEDINGS 1975,AFIPS
073,THE ROLE OF THE DATABASE ADMINISTRATOR,LYON
074,SYSTEM IMPLICATIONS OF INFORMATION PRIVACY,PETERSEN/TURN
075,DATABASE TASK GROUP REPORT 1971,CODASYL
076,THE PROGRAMMER AS NAVIGATOR,BACHMAN
077,HIGH LEV INTEGRITY ASSURANCE IN RELATIONAL DATABASE MGMT SYST,STONEBRAKER
078,AVOIDING DEADLOCK IN MULTITASKING SYSTEMS,HAVENDER
079,SYSTEM DEADLOCKS,COFFMAN/ELPHIC/SHOSHANI
080,NEW SOFTWARE FOR DATABASE MGMNT:PART 1,EDP IN-DEPTH REPORT
081,GENERALIZED DATABASE MGMNT SYST DATA STRUCTURES,TAYLOR
 082,RELATIONAL DATABASE SYSTEMS:A TUTORIAL,DATE
 083,A SURVEY OF GENERALIZED DATABASE MGMT SYST,CODASYL
>
```

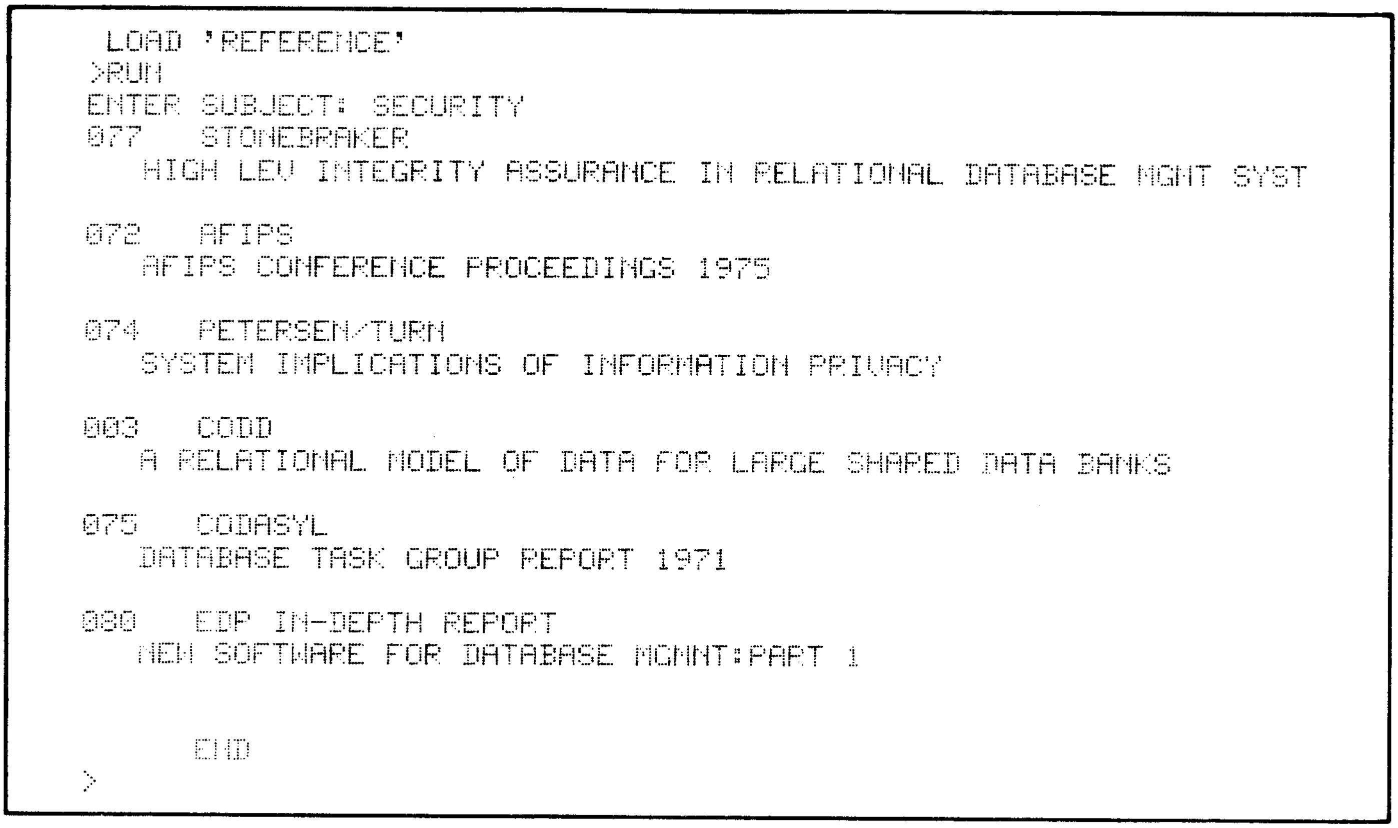
Fig. 12. (contd)

]



```
 LOAD 'RELATIONAL'
>RUN
ENTER RELATIONAL OPERATION: RESTRICTION
ENTER FILENAME,DOMAIN NAME AND DOMAIN VALUE: ID,MFR,HEA
AA30AA HEA 1405 3 J270(F)
AV66YH HEA 5303 2 J270(F)

      END
>RUN
ENTER RELATIONAL OPERATION: JOIN
ENTER TWO FILENAMES AND DOMAIN NAME:      LOC,FAIL,CON
AA30AA 14 3301 125 2A2 14 2555 2215
AA30AA 14 3301 125 2A2 14 2585 1612
AA30AA 14 3301 125 2A2 21 0055 0722
AA30AA 14 3301 125 2A2 20 0015 1515
AA32AP 11 3105 221 1A5 11 2105 1512
AA32AS 12 3702 003 1A1 12 1705 0117
AB21VG 14 3912 002 1A3 14 1225 0711
AA01FG 11 3606 010 1A6 11 2015 0032
AA50AB 12 3102 005 1A6 12 0155 1010
AB56AA 14 3510 003 1B5 14 3025 1706

      END
>
```

Fig. 13. A session at a terminal using the MBASIC program 'RELATIONAL'



```
 LOAD 'REFERENCE'
>RUN
ENTER SUBJECT: SECURITY
077   STONEBRAKER
   HIGH LEV INTEGRITY ASSURANCE IN RELATIONAL DATABASE MGNT SYST

072   AFIPS
   AFIPS CONFERENCE PROCEEDINGS 1975

074   PETERSEN/TURN
   SYSTEM IMPLICATIONS OF INFORMATION PRIVACY

003   CODD
   A RELATIONAL MODEL OF DATA FOR LARGE SHARED DATA BANKS

075   CODASYL
   DATABASE TASK GROUP REPORT 1971

080   EDP IN-DEPTH REPORT
   NEW SOFTWARE FOR DATABASE MGMNT:PART 1


      END
>
```

Fig. 14. A session at a terminal using the MBASIC program 'REFERENCE'

NASA — JPL — Coml., L.A., Calif.