A Preliminary Design of a Data Retrieval Language to Handle a Generalized Data Base: DRL
The National Bureau of Standards\(^1\) was established by an act of Congress March 3, 1901. The Bureau’s overall goal is to strengthen and advance the Nation’s science and technology and facilitate their effective application for public benefit. To this end, the Bureau conducts research and provides: (1) a basis for the Nation’s physical measurement system, (2) scientific and technological services for industry and government, (3) a technical basis for equity in trade, and (4) technical services to promote public safety.

The Bureau consists of the Institute for Basic Standards, the Institute for Materials Research, the Institute for Applied Technology, the Center for Computer Sciences and Technology, and the Office for Information Programs.

**THE INSTITUTE FOR BASIC STANDARDS** provides the central basis within the United States of a complete and consistent system of physical measurement; coordinates that system with measurement systems of other nations; and furnishes essential services leading to accurate and uniform physical measurements throughout the Nation’s scientific community, industry, and commerce. The Institute consists of a Center for Radiation Research, an Office of Measurement Services and the following divisions:

- Applied Mathematics
- Electricity
- Heat
- Mechanics
- Optical Physics
- Linac Radiation
- Nuclear Radiation
- Applied Radiation
- Quantum Electronics
- Electromagnetics
- Time and Frequency
- Laboratory Astrophysics
- Cryogenics.

**THE INSTITUTE FOR MATERIALS RESEARCH** conducts materials research leading to improved methods of measurement, standards, and data on the properties of well-characterized materials needed by industry, commerce, educational institutions, and Government; provides advisory and research services to other Government agencies; and develops, produces, and distributes standard reference materials. The Institute consists of the Office of Standard Reference Materials and the following divisions:

- Analytical Chemistry
- Polymers
- Metallurgy
- Inorganic Materials
- Reactor Radiation
- Physical Chemistry.

**THE INSTITUTE FOR APPLIED TECHNOLOGY** provides technical services to promote the use of available technology and to facilitate technological innovation in industry and Government; cooperates with public and private organizations leading to the development of technological standards (including mandatory safety standards), codes and methods of test; and provides technical advice and services to Government agencies upon request. The Institute also monitors NBS engineering standards activities and provides liaison between NBS and national and international engineering standards bodies. The Institute consists of the following technical divisions and offices:

- Engineering Standards Services
- Weights and Measures
- Flammable Fabrics
- Invention and Innovation
- Vehicle Systems Research
- Product Evaluation Technology
- Building Research
- Electronic Technology
- Technical Analysis
- Measurement Engineering.

**THE CENTER FOR COMPUTER SCIENCES AND TECHNOLOGY** conducts research and provides technical services designed to aid Government agencies in improving cost effectiveness in the conduct of their programs through the selection, acquisition, and effective utilization of automatic data processing equipment; and serves as the principal focus within the executive branch for the development of Federal standards for automatic data processing equipment, techniques, and computer languages. The Center consists of the following offices and divisions:

- Information Processing Standards
- Computer Information
- Computer Services
- Systems Development
- Information Processing Technology.

**THE OFFICE FOR INFORMATION PROGRAMS** promotes optimum dissemination and accessibility of scientific information generated within NBS and other agencies of the Federal Government; promotes the development of the National Standard Reference Data System and a system of information analysis centers dealing with the broader aspects of the National Measurement System; provides appropriate services to ensure that the NBS staff has optimum accessibility to the scientific information of the world, and directs the public information activities of the Bureau. The Office consists of the following organizational units:

- Office of Standard Reference Data
- Office of Technical Information and Publications
- Library
- Office of Public Information
- Office of International Relations.

---

\(^1\) Headquarters and Laboratories at Gaithersburg, Maryland, unless otherwise noted; mailing address Washington, D.C. 20234.

\(^2\) Part of the Center for Radiation Research.

\(^3\) Located at Boulder, Colorado 80302.
TECHNICAL NOTE 590

ISSUED JULY 1971

Nat. Bur. Stand. (U.S.), Tech. Note 590, 26 pages (July 1971)
CODEN: NBTNA

A Preliminary Design of a Data Retrieval Language to Handle a Generalized Data Base: DRL

Elizabeth Fong

Systems Development Division
Center for Computer Sciences and Technology
National Bureau of Standards
Washington, D.C. 20234

NBS Technical Notes are designed to supplement the Bureau's regular publications program. They provide a means for making available scientific data that are of transient or limited interest. Technical Notes may be listed or referred to in the open literature.
TABLE OF CONTENTS

I. INTRODUCTION ........................................... 2

II. BRIEF DESCRIPTION OF LANGUAGE. ...................... 3
   II.1 Input/Output Commands. ............................. 4
   II.2 Data Description Commands ......................... 7
   II.3 Data Maintenance and Manipulation Commands ....... 9
   II.4 Data Retrieval Commands ............................ 11
   II.5 Higher Order Functions ............................. 13
   II.6 Built-In String Manipulation Functions .......... 14

III. METHOD OF IMPLEMENTATION ............................ 14

IV. OPERATION ............................................... 18

V. REFERENCES ............................................... 18

Appendix A
   A Sample Run of DRL ..................................... A-1
A PRELIMINARY DESIGN OF A DATA RETRIEVAL LANGUAGE TO HANDLE A GENERALIZED DATA BASE: DRL

ELIZABETH FONG

DRL (Data Retrieval Language) is a high-level programming language for information retrieval. The language includes a data description language which can describe fixed-length hierarchical data structures, and DRL includes a data retrieval statement whereby a user can retrieve data by specifying conditions on to the data value. DRL also has an environment declaration statement in which the user can indicate specific peripheral devices by unit number for files. The rest of the language consists of an operation repertory of input-output functions and other data manipulations.

DRL is implemented as a preprocessor to FORTRAN V on the UNIVAC 1108, under EXEC II Operating system. Key-words act as triggers and are replaced by blocks of FORTRAN code.

The purpose of this project is to investigate the design of an information retrieval language to handle a generalized data base. The DRL system consists of a set of primitives utilizing both compile-time macros and run-time subroutines. These primitives are embedded in a high-level procedure-oriented programming language--the "host language" -- FORTRAN in this case. These primitives form a base upon which a class of languages can be defined.

Key words: Data base; data retrieval; data structure; information storage and retrieval; language extension; preprocessor; programming language.
I. Introduction

In implementing an information storage and retrieval system, one is faced with the problem of choosing a suitable programming language. Requirements of that programming language are:

- ability to define data structures
- ability to describe an environment
- ability to manipulate data structures
- ability to address data by content
- ability to offer computational power comparable to FORTRAN.

It is possible to implement information storage and retrieval system in procedure-oriented language such as ALGOL, FORTRAN, COBOL, or PL/1, but it is unnatural and requires indirectness in the use of primitives of the language. For example, let us consider writing a program to do the following simple task:

Fine the first element of the array A, of length N, whose value is equal to 3 and set I equal to the value of the index. If none, set I equal to zero.

In a procedural-oriented language such as FORTRAN, one would say:

```
DO 10 I = 1,N
IF (A(I) .EQ. 3) GO TO 20
10 CONTINUE
I = 0
20 . . .
```
Using languages such as LISP, SNOBOL, L6, etc., in doing information retrieval seems even more awkward. COBOL has file manipulation capability but lacks the ability of addressing data by content.[5] This leads to the notion of addressing data by content, which is more natural to a user, who is not a professional programmer. Hence the design of an information retrieval language should be convenient and natural for the user, yet the language should be powerful enough to handle complicated types of data structure. Such a language must have primitives at the level of what is to be done rather than how it is to be done.

An information retrieval language called DRL (Data Retrieval Language) has been designed and is partially implemented at the National Bureau of Standards to meet these objectives on an experimental basis. The DRL language is designed as an extension to FORTRAN explicitly to include the primitives necessary for an information retrieval language. The language will enable us to investigate the benefits of this approach to retrieval problems.

II. Brief Description of the Language

The DRL language is embedded in the FORTRAN V language system on the UNIVAC 1108 computer. It will allow all the usual FORTRAN capabilities plus the following four classes:

1. Input/Output statements
2. Data description statements
3. Data maintenance and manipulation statements
4. Data retrieval statements
II.1 Input/Output Commands

The input/output commands include the peripheral device declaration and also record accessing commands. The input/output devices permitted at present are card reader, card punch, printer, magnetic tape and drum. The logical unit number is the same as FORTRAN V standard table as set up at National Bureau of Standards, Gaithersburg.

a. ENVIRO

The ENVIRONMENT declaration provides information about the physical location of the data set associated with a file. This information allows the preprocessor to determine the method of accessing the data set, and causes the tape or drum to position at the beginning.

FORMAT

ENVIRO ( <filename >, <logical unit no> )

<filename> ::= A FORTRAN variable.

<logical unit no> ::= An integer variable/integer.

Allowable values for the logical unit numbers and their assignments are listed as follows:

<table>
<thead>
<tr>
<th>LOGICAL UNIT</th>
<th>ASSIGNMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Reread</td>
</tr>
<tr>
<td>1</td>
<td>Card reader</td>
</tr>
<tr>
<td>2</td>
<td>Printer</td>
</tr>
<tr>
<td>3</td>
<td>Card punch</td>
</tr>
<tr>
<td>4</td>
<td>Console</td>
</tr>
</tbody>
</table>
Card reader

Printer

Tapes A - Z

Tape (  

Tape -

Entire Drum 1300000 to 3777777 (octal)

Lower Half 1300000 to 2537777 (octal)

Upper Half 2540000 to 3777777 (octal)

Lower third 1300000 to 2177777 (octal)

Middle third 2200000 to 3077777 (octal)

Upper third 3100000 to 3777777 (octal)

EXAMPLE

ENVIRO (PAYROL, 35)

The above means declaring the entire drum as a mass storage device for a file called PAYROL.

b. PUTOUT

PUTOUT will write a record currently set up in core onto the indicated peripheral device. Unless a FORTRAN format label is supplied, the output is assumed to be binary.

FORMAT

PUTOUT ( <where> , {<label> ,} <filename> )

<where>::= peripheral device unit number. It is assumed that the device is positioned to be written.

<label>::= FORTRAN format label. This field is optional
EXAMPLE

PUTOUT (35, PAYROL)

or

PUTOUT (6,100, PAYROL)

100 FORMAT (1X, 22A6)

The first PUTOUT example means write out a record which is in main storage array PAYROL onto the previously positioned drum. The second PUTOUT example means write on the printer the record according to the FORTRAN format statement labeled 100.

c. GETIN

GETIN reads in a record from the indicated peripheral device onto the record image space in core. Unless a FORTRAN format label is supplied, the input is assumed to be binary.

FORMAT

GETIN ( <where> , {<label> ,}<filename> )

<where>::= peripheral device unit number. It is assumed that the device is positioned to be read.

<filename>::= A FORTRAN variable.

<label>::= FORTRAN format label. This field is optional

EXAMPLE

GETIN (35, PAYROL)

or

GETIN (5, 100, PAYROL)

100 FORMAT (80A1)
The first GETIN example reads one record from the previously positioned drum into the main storage array PAYROL. The second GETIN example reads from the card reader according to FORTRAN format statement labeled 100 into the array PAYROL.

II.2 Data Description Commands

Facility for declaring hierarchically structured data is provided. The declaration format is patterned after PL/1 where the level of hierarchy is indicated by the level number in front of the variables in the declaration. When the variable occurs without the level number in front, it is assumed that the declaration is merely a single data item or an array. To facilitate character manipulation, the string declaration is added to the FORTRAN type statements.

a. DECLAR CHARAC or BITS

DECLAR with the data type CHARAC means the variable being declared is a string of n six-bit Fielddata code as defined for the UNIVAC 1108. DECLAR with the data type BITS means the variable being declared is of n bits taking on values one or zero.

FORMAT

DECLAR ( <variable> CHARAC ( <n> ))

DECLAR ( <variable> BITS ( <n> ))

<variable> ::= A FORTRAN Variable.

<n> ::= An integer greater than 0.

EXAMPLES

DECLAR ( NAME CHARAC (10))

DECLAR ( MATRIX BITS (8))
b. DECLAR hierarchically

DECLAR with a hierarchical data list declares a data structure consisting of elementary data items (terminal nodes) and composite data items (non-terminal nodes or meta syntactical variable). The composite data items must have one or more subordinates and the elementary data items must have data type and size specifications associated with them.

**FORMAT**

DECLAR ( <hierarchical data list> )

<hierarchical data list>::= <hierarchical data>/

<hierarchical data list>, <hierarchical data>

<hierarchical data>::= <n> <variable>/

<n> <variable> <type> <dimension>

<type>::= all legal FORTRAN data types (only first six characters) plus CHARAC and BITS

<dimension>::= ( <n> )

<variable>::= A FORTRAN variable

<n>::= An integer greater than 0

**EXAMPLE**

DECLAR ( 0 PAYROL,
    1 NAME,
    2 FIRST CHARAC (10),
    2 MIDDLE CHARAC (10),
    2 LAST CHARAC (10),
    1 SALARY,
    2 REGU INTEGE (1),
    2 OVER INTEGE (1),
    1 OCC CHARAC (20) )

8
II.3 Data Maintenance and Manipulation Commands

a. PUT

PUT assumes a data structure into which values are to be stored. If the attribute name happens to be an elementary data item, then the value is simply put in. If the attribute is not an elementary data item, then the lists of values to fill the attributes subordinate to it must be given.

**FORMAT**

\[
\text{PUT (} \text{<attribute>, <value list> )}
\]

\[
\text{<attribute>:= elementary data item variable or composite data item variable as declared in DECLAR statement}
\]

\[
\text{<value list> := <value>/<value list>, <value>}
\]

\[
\text{<value>:= any expression (The present version can only handle constants, literals and variables)}
\]

**EXAMPLE**

The following example assumes the declaration which appears in the DECLAR hierarchically example given above.

\[
\text{PUT (LAST, 'FONG')}
\]

\[
\text{PUT (NAME, 'LIZ', 'NEE', 'FONG')}
\]

The first PUT expression whose first argument is the elementary data item LAST and therefore the value is immediately assigned. In the second PUT expression the first argument is the composite data item NAME consisting of three subordinates and therefore the three values 'LIZ', 'NEE', and 'FONG' are assigned to FIRST, MIDDLE and LAST respectively.
b. LOCATE

Records within a file have an ordinal number according to their position within the file. LOCATE positions the file with respect to this ordinal number of the record.

**FORMAT**

```
LOCATE ( <filename>, <index> )
```

- `<filename>` := A FORTRAN variable defined in ENVIRO and DECLAR statements
- `<index>` := An integer / An integer variable. If index = 0 the file is position to the beginning of the file.

**EXAMPLE**

```
LOCATE (PAYROL, 5)
LOCATE (PAYROL, IXGET)
```

c. DELETE AND DELIX

DELETE deletes the first encountered record which satisfies the given condition list.

DELIX deletes the Nth record in the file where N is given.

**FORMAT**

```
DELETE ( <filename> , <condition list> )
DELIX ( <filename >, <index> )
```

- `<condition list>` := This is the same as described under the Get command
- `<filename>` := A FORTRAN variable defined in ENVIRO and DECLAR statements
- `<index>` := An integer / An integer variable
EXAMPLE

The following example assumes the declaration which appears in the DECLAR (hierarchically) example.

DELETE (PAYROL, LAST EQ 'SMITH')
DELIX (PAYROL, 5)

II.4 Data Retrieval Commands

a. GET

GET is a retrieval function which returns the value of the specified attribute in the first encountered record which satisfies the given conditions.

FORMAT

A = GET (<filename> , <attribute> , <condition list> )

<filename> ::= A FORTRAN variable declared in ENVIRO and DECLAR statements. It is assumed that the peripheral device is positioned to start searching.

<attribute> ::= Elementary data item variable or complex data item variable as declared in DECLAR statement.

<condition list> ::= <wff>

<wff> ::= <proposition> / <wff><connectives><wff>

<proposition> ::= <attribute><rel><value>

<connectives> ::= AND / OR

<rel> ::= EQ / NE / GT / GE / LT / LE

[value] ::= Any expression
EXAMPLE

The following example assumes the declaration which appears in the DECLAR (hierarchically) example.

\[ A = \text{GET ( PAYROL, NAME, REGU EQ 400 AND OCC EQ 'MATH') } \]

This GET command will search the PAYROL file. If the field \text{REGU} equal 400 and the field \text{OCC} equals MATH, then the field \text{NAME} will be retrieved and stored in \text{A}. \text{A} must be properly dimensioned.

DEFAULT CONDITIONS

If an error occurs, RUNERR routine is executed. RUNERR is a routine which may be supplied by the user to handle error recovery. If the user does not supply a RUNERR routine, the DRL system will execute the UNIVAC EXEC II error routine which will just stop execution.

REMARK

If there is no second argument of the GET command, i.e., two commas with nothing in between, then it is assumed that the index value is required. In any case an internally defined variable called IXGET will always contain the index value after each GET function. The IXGET value will be destroyed upon initiation of the next GET function.

b. GETALL

GETALL is the same as GET except instead of retrieving a single item, a whole set is retrieved. The variable occurring on the left of the GETALL statement must be pre-declared as one dimensional array with an estimate of maximum size. After retrieving, the first entry of that array will contain the count of the number of items retrieved followed by the values. The user must also DIMENSION the IXGET to be one dimensional array of maximum size. After each execution of GETALL, IXGET (1)
contains the count of the number of items retrieved followed by the list of pointer to the retrieved records.

II.5 Higher Order Functions

These functions could be defined by combining appropriate previously defined primitive functions.

a. REPLACE

REPLACE may be defined by combining the following four DRL primitives.

GET - get a whole record that meets given conditions
PUT - change value as desired
LOCATE - position back
PUTOUT - put back the modified record in the file.

b. SUBSET

SUBSET may be defined by the following pseudo statements.

ENVIRO - declared a different unit number for a subfile

GET - get a record that meets the given conditions

IF end-of-file THEN stop

PUTOUT - putout on the new unit

GO TO 10

An alternative way of defining assumes the existence of the primitive MOVE. This may be defined as a user's subroutine.

ENVIRO - declare a different unit number for a subfile
GETALL - get all records that meets the given conditions
MOVE - move to individual buffers for a unit record
PUTOUT - putout on the new unit.
c. COUNT

COUNT may be defined with the GETALL command and reading out the first word of the IXGET array or the user-defined array containing the answers.

II.6 Built-in String Manipulation Functions

There exists in the DRL system a group of run-time subroutines which are accessible to the user. The following is a set of string i.e., characters manipulation functions which are patterned after PL/1. The string variables must be declared as characters in a DECLAR CHARAC statement.

CONC (A,B) is the concatenation of string A and B.
LENGTH (A) is the length of string A.
INDEX (A, 'B') returns the starting location for the first occurrence of 'B' in string A or zero if not found.
SUBSTR (A, I, J) extracts the substring starting at position I of length J. If J equals 0 the rest of String A is returned.
MATCH (A,B,N) compares the first N characters of A to B. If they are identical, the value of MATCH is true, otherwise the value of MATCH is false.

III. Method of Implementation

The DRL language translator is a preprocessor to FORTRAN V on the UNIVAC 1108. It consists of two major phases:

Phase 1 - A scanner reads the input stream and traps all the DRL keywords and replaces them with appropriate blocks of FORTRAN code.
The declarative statements generate FORTRAN dimensioning statements and tables containing the data descriptions.

Phase 2 - A collection of predefined run-time subroutines to perform all of the above described tasks.

Both phase 1 and phase 2 work. All of the primitives defined above have been implemented. The higher order functions have not been implemented.

The main routine is the lexical scanner called LSCAN. This routine reads the DRL statements and branches according to the keywords scanned. The DRL syntax conforms to the FORTRAN statement format. If a given statement does not contain a DRL keyword, then it is assumed to be a FORTRAN statement, and the line is carried over to the generated program. The DRL keyword analysis is described individually as follows:

a. DROP - This should be the first statement of the DRL program. The operand of this statement defines a name for the output FORTRAN source.

b. ENVIRO - This statement generates the equivalent of an open file statement by positioning the peripheral device indicated by the unit number to the beginning. The file name is entered into the FILNAM table.
c. DECLAR - This generates FORTRAN DIMENSION statements. If it is hierarchical data declaration, then it generates DIMENSION statements and EQUIVALENCE statements for all the complex data item names. Also the hierarchical data information is stored in the 8-column matrix J. The meaning of these 8 columns are as follows: J(I,1) contains the hierarchical level number J(I,2) contains the BCD symbol table index number J(I,3) contains the degree or number of subtrees of that node J(I,4) contains the index no. of the parent node J(I,5) contains the index no. of the sister node J(I,6) contains the next occupant of the same name J(I,7) contains the size in characters J(I,8) contains the starting character position from the beginning.

The details of the J-matrix can be found in Lawson [2].

The variable names are also entered into the SYMBOL table.

Layout of SYMBOL table

| SYMBOL | SPOINT | Pointer to next available entry |
|--------|--------|---------------------------------
|        | SYMBOL in BCD | |
|        | Type | Size | Index to J |

Type = 1 means character
Type = 2 means BITS
Type = 0 means other
SIZE = no. of characters or bits.
d. **PUTOUT** - This generates an appropriate output statements including calls to **NTRAN** if the output is to be binary. **NTRAN** is a FORTRAN callable subroutine which provides buffered input/output routines for tape and drum. Detailed description of **NTRAN** can be found in section 7.5 of the UNIVAC 1107 FORTRAN manual [3].

e. **GETIN** - This is the same as **PUTOUT** except it generates appropriate output statements.

f. **GET** - This will first generate a **NTRAN** call to read in a record. A boolean function containing the conditions given is next generated. On the 'true' branch the extraction of the specified attribute code is supplied. On the 'false' branch, a **GOTO** statement back to the **NTRAN** call to read in the next record is generated.

g. **GETALL** - Same as **GET** except a loop is set up to continuing retrieving until an end of file is reached.

h. **PUT** - This will generate a DO Loop containing a call to a run-time routine called **MOVECH**. **MOVECH** will move a character from the indicated source to the indicated destination.

i. **LOCATE** - This will generate a **NTRAN** call to position the peripheral device **N** records from the beginning. If the device is drum, the drum address is positioned from the beginning by the amount **N** times the length of the record.

j. **DELETE** - (This command is to be used primarily with drum) This statement will first call **GET** to determine which record meets the conditions given. Then it calls **LOCATE** to position the file to this record. The record is then zeroed. Garbage collection
is not yet coded.

k. DELIX - (This command is to be used primarily with drum). This
first calls LOCATE to position the file to the Nth record, and
the record is zeroed.

IV. Operation

The input to the DRL translator is the program text written in the
drl language. The output is a FORTRAN program automatically residing on
the drum, linked-edited, and ready to be executed. This output FORTRAN
program, together with the predefined run-time action routines and block
data, will be the final executable program capable of manipulating data,
accessing the peripheral storage and performing any kind of retrieval
tasks. This present version consists of approximately 1900 lines of
FORTRAN and approximately 300 lines of UNIVAC assembly code.

The source listing of the entire DRL system is available from the
author upon request.

The author is deeply indebted to Mr. Charles T. Meadow for first
suggesting the topic and for his interest throughout the implementation.

V. References

[2] Lawson, Harold W., Jr., "The Use of Chain List Matrices for the
Analysis of COBOL Data Structures," presented at the ACM National
C 28-8201-1, IBM 1968.
Systems - A Survey", IEEE Transaction on Electronic Computers,
APPENDIX A -- A Sample Run of DRL

In this appendix, the following three outputs are presented:

(1) DRL sample program. These DRL statements are translated into FORTRAN codes which appear indented to the right. The data description table, symbol table, and file name table are generated as FORTRAN assignment statements.

(2) FORTRAN compilation of the DRL generated program.

(3) Execution of the sample program. The sample program reads in from the card reader a file of personnel records. It sets up the data base on drum. Two retrieval commands are executed.
BEFORE CALLING DRUM WRITE
DRUM WRITE CALLED

| J | 5 | 6 |
| J | 2 | 1 |
| J | 3 | 2 |
| J | 5 | 1 |
| J | 2 | 0 |
| J | 3 | 1 |
| J | 4 | 2 |
| J | 5 | 3 |
| J | 6 | 2 |
| J | 7 | 3 |
| J | 8 | 2 |
| J | 9 | 3 |
| J | 10 | 2 |
| J | 11 | 3 |
| J | 12 | 2 |
| J | 13 | 3 |
| J | 14 | 2 |
| J | 15 | 3 |
| J | 16 | 2 |
| J | 17 | 3 |
| J | 18 | 2 |
| J | 19 | 3 |
| J | 20 | 2 |
| J | 21 | 3 |
| J | 22 | 2 |

SYMBOLI 1) = ANDNAME
SYMBOLII 2) = ANDPATTERN
SYMBOLIII 3) = ANDNAME
SYMBOLIV 4) = ANDNAME
SYMBOLV 5) = ANDNAME
SYMBOLVI 6) = ANDNAME
SYMBOLVII 7) = ANDNAME
SYMBOLVIII 8) = ANDNAME
SYMBOLIX 9) = ANDNAME
SYMBOLX 10) = ANDNAME
SYMBOLXI 11) = ANDNAME
SYMBOLXII 12) = ANDNAME
SYMBOLXIII 13) = ANDNAME
SYMBOLXIV 14) = ANDNAME
SYMBOLXV 15) = ANDNAME
SYMBOLXVI 16) = ANDNAME
SYMBOLXVII 17) = ANDNAME
SYMBOLXVIII 18) = ANDNAME
SYMBOLXIX 19) = ANDNAME
SYMBOLXX 20) = ANDNAME
SYMBOLXXI 21) = ANDNAME
SYMBOLXXII 22) = ANDNAME
SYMBOLXXIII 23) = ANDNAME
SYMBOLXXIV 24) = ANDNAME

GO TO 2
END

A-2
END OF UNIVAC 1108 FORTRAN V COMPIILATION, 1 "DIAGNOSTIC" MESSAGE(S)

PHASE 1 TIME = 1 SEC.
PHASE 2 TIME = 0 SEC.
PHASE 3 TIME = 0 SEC.
PHASE 4 TIME = 0 SEC.
PHASE 5 TIME = 0 SEC.
PHASE 6 TIME = 0 SEC.

TOTAL COMPIILATION TIME = 2 SEC

NAME-PAUL REG$100000 OVER*120000
NAME=BEGIN NEW PRZBY,SKJ REG$100000 OVER*100000
NAME=TRUMAN E, TURNIPSEED REG$ 60000 OVER* 7000
NAME=PEACE HAPINESS REG$500000 OVER*460000
NAME=CRYSTAL SHAANDA LEAR REG$400000 OVER* 0
NAME=PRISE GOD BAREBONES REG$400000 OVER*460000
NAME=IMA A, HOGG REG$425000 OVER*800000
NAME WHERE LAST CO LEAR =CRYSTAL SHAANDA LEAR
NAME WHERE REG GT 700=PAUL KRC$
1. **Title and Subtitle**

   A PRELIMINARY DESIGN OF A DATA RETRIEVAL LANGUAGE TO HANDLE A GENERALIZED DATA BASE: DRL

7. **Author(s)**

   Elizabeth Fong

9. **Performing Organization Name and Address**

   NATIONAL BUREAU OF STANDARDS
   DEPARTMENT OF COMMERCE
   WASHINGTON, D.C. 20234

16. **Abstract**

   DRL (Data Retrieval Language) is a high-level programming language for information retrieval. The language includes a data description language which can describe fixed-length hierarchical data structures, and DRL includes a data retrieval statement whereby a user can retrieve data by specifying conditions on to the data value. DRL also has an environment declaration statement in which the user can indicate specific peripheral devices by unit number for files. The rest of the language consists of an operation repertory of input-output functions and other data manipulations.

   DRL is implemented as a preprocessor to FORTRAN V on the UNIVAC 1108. Keywords act as triggers and are replaced by blocks of FORTRAN code.

   The purpose of this project is to investigate the design of an information retrieval language to handle a generalized data base. The DRL system consists of a set of primitives utilizing both compile-time macros and run-time subroutines. These primitives are embedded in a high-level procedure-oriented programming language—the "host language"—FORTRAN in this case. These primitives form a base upon which a class of languages can be defined.

17. **Key Words**

   Data base; data retrieval; data structure; information storage and retrieval; language extension; preprocessor; programming language.

18. **Availability Statement**

   [✓] UNLIMITED.

   [ ] FOR OFFICIAL DISTRIBUTION. DO NOT RELEASE TO NTIS.
PERIODICALS

JOURNAL OF RESEARCH reports National Bureau of Standards research and development in physics, mathematics, chemistry, and engineering. Comprehensive scientific papers give complete details of the work, including laboratory data, experimental procedures, and theoretical and mathematical analyses. Illustrated with photographs, drawings, and charts.

Published in three sections, available separately:

- Physics and Chemistry
  Papers of interest primarily to scientists working in these fields. This section covers a broad range of physical and chemical research, with major emphasis on standards of physical measurement, fundamental constants, and properties of matter. Issued six times a year. Annual subscription: Domestic, $9.50; foreign, $11.75*.

- Mathematical Sciences
  Studies and compilations designed mainly for the mathematician and theoretical physicist. Topics in mathematical statistics, theory of experiment design, numerical analysis, theoretical physics and chemistry, logical design and programming of computers and computer systems. Short numerical tables. Issued quarterly. Annual subscription: Domestic, $5.00; foreign, $6.25*.

- Engineering and Instrumentation
  Reporting results of interest chiefly to the engineer and the applied scientist. This section includes many of the new developments in instrumentation resulting from the Bureau's work in physical measurement, data processing, and development of test methods. It will also cover some of the work in acoustics, applied mechanics, building research, and cryogenic engineering. Issued quarterly. Annual subscription: Domestic, $5.00; foreign, $6.25*.

TECHNICAL NEWS BULLETIN

The best single source of information concerning the Bureau's research, developmental, cooperative and publication activities, this monthly publication is designed for the industry-oriented individual whose daily work involves intimate contact with science and technology—for engineers, chemists, physicists, research managers, product-development managers, and company executives. Annual subscription: Domestic, $3.00; foreign, $4.00*.

* Difference in price is due to extra cost of foreign mailing.

Order NBS publications from:
Superintendent of Documents
Government Printing Office
Washington, D.C. 20402

NONPERIODICALS

Applied Mathematics Series. Mathematical tables, manuals, and studies.


Handbooks. Recommended codes of engineering and industrial practice (including safety codes) developed in cooperation with interested industries, professional organizations, and regulatory bodies.

Special Publications. Proceedings of NBS conferences, bibliographies, annual reports, wall charts, pamphlets, etc.

Monographs. Major contributions to the technical literature on various subjects related to the Bureau's scientific and technical activities.

National Standard Reference Data Series. NSRDS provides quantitative data on the physical and chemical properties of materials, compiled from the world's literature and critically evaluated.

Product Standards. Provide requirements for sizes, types, quality and methods for testing various industrial products. These standards are developed cooperatively with interested Government and industry groups and provide the basis for common understanding of product characteristics for both buyers and sellers. Their use is voluntary.

Technical Notes. This series consists of communications and reports (covering both other agency and NBS-sponsored work) of limited or transitory interest.

Federal Information Processing Standards Publications. This series is the official publication within the Federal Government for information on standards adopted and promulgated under the Public Law 89–306, and Bureau of the Budget Circular A–86 entitled, Standardization of Data Elements and Codes in Data Systems.

Consumer Information Series. Practical information, based on NBS research and experience, covering areas of interest to the consumer. Easily understandable language and illustrations provide useful background knowledge for shopping in today's technological marketplace.
