Author(s)	Prokop, Jan Stuart.
Title	Remote information retrieval and display.
Publisher	Monterey, California: U.S. Naval Postgraduate School
Issue Date	1964
URL	

REMOTE INFORMATION RETRIEVAL AND DISPLAY

JAN S. PROKOP

REMOTE INFORMATION RETRIEVAL AND DISPLAY

by

Jan S. Prokop

Lieutenant, Supply Corps, United States Navy

Submitted in partial fulfillment of
the requirements for the degree of

MASTER OF SCIENCE
IN
MANAGEMENT/DATA PROCESSING

United States Naval Postgraduate School
Monterey, California

1 9 6 4

REMOTE INFORMATION RETRIEVAL AND DISPLAY

by

Jan S. Prokop

This work is accepted as fulfilling

the thesis requirements for the degree of

MASTER OF SCIENCE

IN

MANAGEMENT/DATA PROCESSING

from the

United States Naval Postgraduate School

ABSTRACT

This paper describes an experiment in the remote retrieval of information. A program was written to permit interrogation of the CDC 1604 computer from the Data Display dd65 display unit, and to provide a visual display of the result of the interrogation on the Cathode Ray Tube displays of the dd65 console. The interrogator also has an option of requesting a hard-copy printout of a display within a time limit of thirty seconds after the commencement of each new display presentation. Operator instructions are shown on the dd65 scope to assist the operator in making the interrogation. In addition, certain dd65 console lights are activated to indicate that the CDC 1604 has accepted each new input segment. This experimental program is integrated with and based upon the SEMI-AUTOMATIC BIBLIOGRAPHIC INFORMATION RETRIEVAL SYSTEM (SABIRS), developed by the Computer Facility, and used by the Technical Reports Library, U. S. Naval Postgraduate School.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

ABBREVIATIONS

ASA .... American Standards Association

B .... When used as a suffix to a number, indicates that the number is expressed in octal notation.

BCD .... Binary Coded Data

BYTE .... Six bits representing one BCD character

CDC .... Control Data Corporation

CRT .... Cathode Ray Tube, or 'scope' (Oscilliscope)

IBM .... International Business Machines, Inc.

I/O .... Input/Output

IR .... Information Retrieval

MCS .... Master Control System

SABIRS........................ SEMI-AUTOMATIC BIBLIOGRAPHIC INFORMATION RETRIEVAL SYSTEM

usec .... Micro-second; one millionth of a second

I. INTRODUCTION

A. General Criteria

One of the most critical problem areas in any automated information processing system is the time required to extract and present selected information when needed. The increasing trend toward higher processing speed in the arithmetic unit of computers has served to underline this problem. It is now possible to process a unit of data in nano-seconds, but the time required to obtain a copy of this information and present it to a requestor still varies from minutes to hours (and sometimes days) in a job-shop computer application. This delay in the physical receipt of requested information is a function of the administrative process, and while one particular item may be individually handled for rapid processing, a mass of such requests imposes an undue, and in many cases impossible, burden on the administration of the computer center. A solution to the problem has been to provide remote on-line interrogation units and printers for selected uses. These units provide a hard-copy of the requested information on demand, with minimal interruption of the computer center work.

This solution, while certainly far better than waiting for a regular printout to be run and distributed, has certain disadvantages. While the on-line printer and interrogation unit may work well in some locations, it is not suitable for executive office use, or in a location where the noise of a printer would be disruptive, or where rapid updating of the information (such as in command and control problems) is the rule. In these and other instances, display units have become popular by combining the least number of disadvantages with the most flexibility. While visual display in some few instances of information retrieval is not as desirable as a hard-copy, it does present the information immediately. Meanwhile, a hard-copy may be in process if requested, for further study or dissemination. In most applications, however, only a one-time inspection of a specific subset of data is desired. In these situations, an on-line display unit presents unique advantages over other on-line equipment.

This paper will discuss one particular display program, and its application within the general context of Information Retrieval. Section II gives a description of the general hardware and program environment. Particular emphasis is given to the Data Display dd65 unit, and the SABIR2 version of the SEMI-AUTOMATIC BIBLIOGRAPHIC INFORMATION RETRIEVAL SYSTEM (SABIRS) [2]. Section III discusses the demonstration program in detail, describing the peculiarities of programming the display and interrogation for the dd65. In this section also will be found suggestions for further improvements and modifications of the program. Section IV gives the sequence of operating instructions, starting with the activating of the dd65 logic unit, the bootstrap of the SABIR2, and ending with the display console operation. Section V discusses Information Retrieval in general, and provides correlation between the demonstration program and its usefulness in a real world environment. Section V is followed by References and Bibliography, then the Appendices, in which are detailed the flow-charts of the system and of the program, and then the program listing. Throughout the paper, illustrations are provided as necessary to clarify the discussion.

B. Definition and Scope

This paper presents one method of remote on-line interrogation and display, with requestor option of hard-copy printout. The system is integrated with the present SABIR2 [5] modification of SABIRS [2] , a search and retrieval of bibliographic information from the Technical Reports Section of the U. S. Naval Postgraduate School Library. This search and retrieval is based on the selection of identifiers, specifically Uniterms (see Section III), and the search for the logical intersection of these identifiers among the bibliography listings of the Library's Technical Reports Section. The demonstration program was written in the SCRAP [3] language, making use of I/O Channel #7, allowing direct core-to-core transfer between the CDC 1604 computer and the Data Display dd65 display unit. (See Figure 1).

In this program, the 1604 senses for a display request from the dd65. If a display request has been made, the 1604 displays the operating instructions on the dd65 for the operator's convenience, and then processes the request information as it is keyed in from the dd65 in the required format. A search of the Document Tape (Figure 7, Page 35) is then

BLOCK DIAGRAM OF CDC 1604 - dd65 SYSTEM

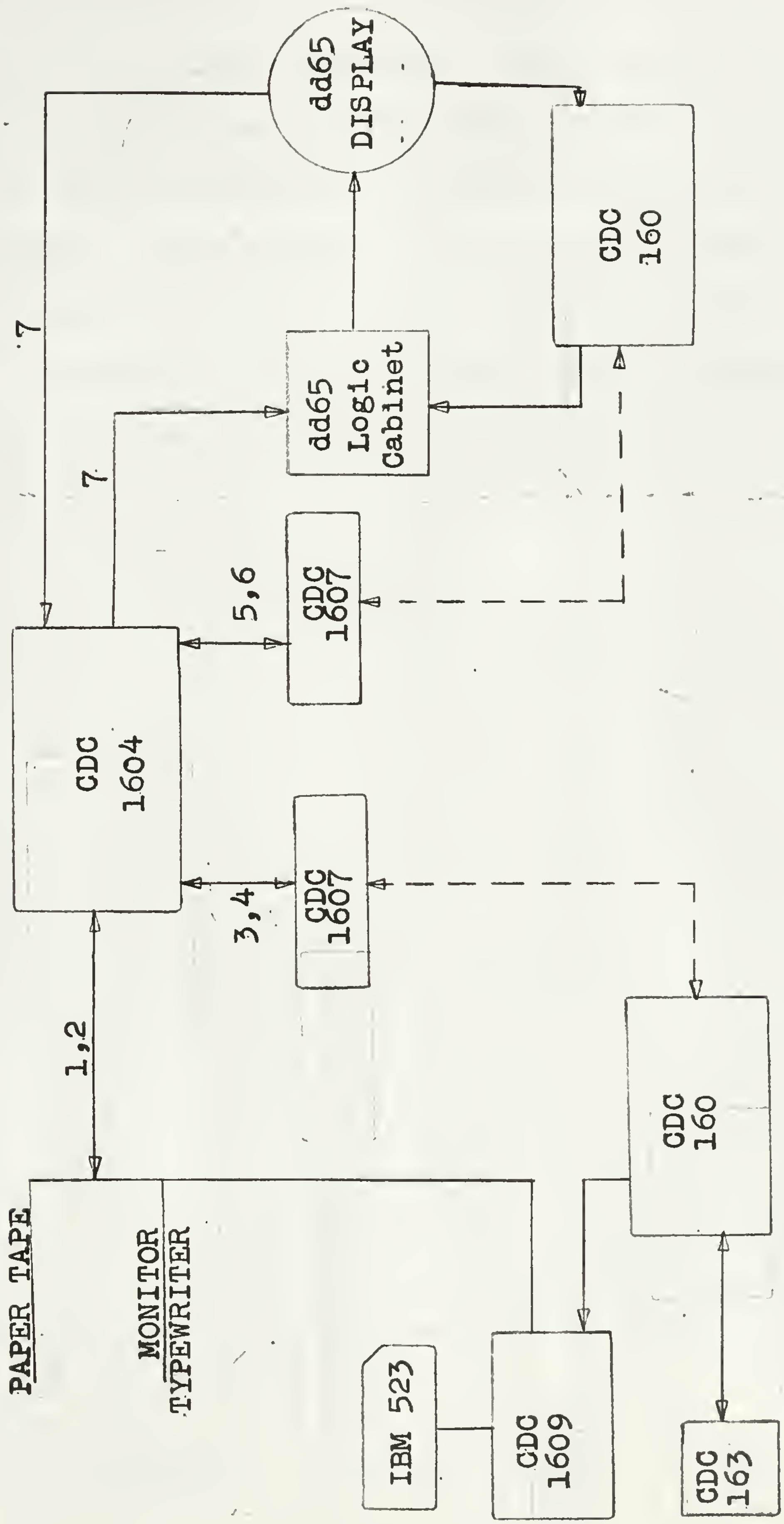


FIGURE 1

begun. If a document fulfilling the request is located, the English text of the 84 word abstract, if available, is presented on the display unit. The requestor then has the option of requesting a hard-copy printout of the abstract, and the search continues for another document which fulfills the request. Each display is available to the requestor for an interval of thirty seconds. If no decision has been made by the requestor during this time limit, a printout of the display is made, and the next available abstract is presented.

II. ENVIRONMENT DESCRIPTION

A. CDC 1604

The CDC 1604 is a stored-program general-purpose digital computer of solid-state design. It incorporates a storage unit of 32,768 48-bit words (or 262,144 6-bit bytes, or characters) located in two independent 16,384 word banks. These banks are phased alternately with overlapping cycles to provide a 4.8 usec effective cycle time for random addresses. The mode of operation is parallel; that is, arithmetic operations are performed 48 bits at a time.

Input/output is conducted over three buffered input channels, (#1, #3, #5), three buffered output channels (#2, #4, #6) and one high speed channel (#7). (See Figure 1.) Data transfer is accomplished in 48-bit words. Computation continues during I/O except when a word is brought in from memory to the disassembly register or from the assembly register to memory. A peak transfer rate on any channel of 4.8 usec is available, however, the actual transfer rate is governed by the rate of the external equipment. The peak transfer rate of 4.8 usec for a 48-bit word is achieved on the core-to-core transfer of Channel #7. Channel #7 is used exclusively in the demonstration program for communication with the display unit [6].

B. dd65

The dd65 consists of two basic units, the display console and the logic unit. The display console has two 12-inch

CRTs, each with a usable $8\frac{1}{2}$ inch square display area, and two keyboards. The keyboards are input devices for the associated computer, either a CDC 160 or the 1604 (Figure 1), and are basically a typewriter keyboard (keyboard #1) and a special purpose keyboard (keyboard #2). This keyboard combination is shown in Figure 5, Page 23. The logic unit contains the I/O cable connections, the logical circuitry of the system, and a core memory of 512 words of 48 bits each.

The facilities at the console provide a man-machine interface with the main computer (the 1604 in this case). Keyboard hits are decoded, one at a time, by the program written for the CDC 1604, and the console operator is able to examine the displayed results of his keyboard interrogation. Symbols cannot be entered directly on the CRT display areas from the keyboard. As indicated in Figure 1, the circuitry is one-way for the display. Hence, if a keyboard hit is to be displayed, it must go from the keyboard to the computer, then to the display scope via the dd65 logic unit, and must be accompanied by a designator word. This designator word is discussed below in Section III.

The logic unit receives the output from the computer in BCD format, as two 24-bit words, and stores them sequentially in its own core memory. Similarly, the computer cannot read out of the logic unit memory directly. One further feature of the logic unit is controlled by a miniature toggle switch located on the inside edge of the cabinet. This switch, when in the "up" position, disables the upper 256 words of core memory. Despite the advice of instruction manual [7], it is found to be much more useful in the "down" position, thus enabling the entire memory. Unfortunately this decreases the symbol generation frequency, causing a slight flickering on the right tube. However, this was not found to be particularly annoying, and was in fact quite acceptable in view of the desirability of having all of the 512 words of core memory available for use.

The I/O cable connections at the bottom of the dd65 logic cabinet should be examined before each use. The original cable markings as described in the dd65 maintenance manual are no longer on the cables. The markings indicated in Figure 8, should be used as a guide. This same figure shows the proper cable connections in the CDC 1604 main frame for Channel #7. Only the connections shown need concern the user.

C. SABIR2

SABIR2 is a modification of the SABIRS [2] information retrieval routine written for the Technical Reports Section of the U. S. Naval Postgraduate School Library [5]. This system was written to accept a paper-tape input to request technical report bibliographies which fulfill certain requirements of subject, source, or dates. The paper tape is read into the CDC 1604, and a search is initiated on the Document Tape (See Figure 7, Page 35.) for documents which have all the desired qualifications. If a match is found, an English Title File Tape is searched for an English abstract which corresponds to the document number. The abstract is printed out if it is available, otherwise only the accession number of the document is printed for each match. Up to fifty requests can be accepted at one time. SABIR2 also has the facilities for updating the Document and English abstract tapes; however, this paper will be concerned only with the search portion of the program, and its resultant display.

The necessary input to SABIR2 is a request record. The discussion in Section III (C) includes sample data of the various parts of a request record. Briefly, the request record is prepared according to the format described in Reference 5, and consists of:

1. An indentification-8 characters which identify the requestor.

2. Dates-16 characters specifying the inclusive first and last dates desired.

3. Sources-zero to twelve 8-character source codes.

4. Uniterms-one to twelve 8-character Uniterms, which specify subject areas of interest. Section V of this paper gives a discussion of the Uniterm system of classification; its uses and limitations.

The SABIR2 system will print error messages on the output magnetic tape if the specifications for input are not met in any one of the request records. The particular request in error will then be deleted.

The unmodified SABIR2 search system basically consists of the MASTER CONTROL SYSTEM (1604 memory cells 00030B -

MAP OF STORAGE ALLOCATION

dd65 STORAGE

Start	dd65 STORAGE	End
000B	"IDENTIFICATION" INSTRUCTION	031B
032B	"DATES" INSTRUCTION	127B
130B	"NO. OF SOURCES" INSTRUCTION	167B
170B	"SOURCES" INSTRUCTION	257B
260B	"NO. OF UNITERMS" INSTRUCTION	315B
316B	"UNITERMS" INSTRUCTION	371B
372B	"CARRIAGE RETURN" INSTRUCTION	404B
416B	"BIBLIOGRAPHY" DISPLAY	716B

MODIFIED SABIR2 IN THE CDC 1604 MEMORY

Start	MODIFIED SABIR2 IN THE CDC 1604 MEMORY	End
00030B	MCS	00235B
00300B	COMMON PORTIONS OF SEARCH PART 1 AND SEARCH PART 2	03713B
04000B	SEARCH PART 1	04672B
04000B	SEARCH PART 2	05103B
05200B	OUTPUT BUFFER	77767B

FIGURE 2

00222B), SEARCH PART 1 (cells 00300B - 04271B), and SEARCH PART 2 (cells 00300B - 05066B). Cells 00300B - 03563B are common to SEARCH PART 1 and SEARCH PART 2. About 1450 of these common cells are taken up with the blocks of request records, the rest being program and constants. The beginning of the output block is at 05200B and the ending is at 77767B (See Figure 2).

After the MCS has been bootstrapped, the operator then types the date and the SEARCH option on the CDC 1604 console typewriter. The MCS then reads in SEARCH PART 1 and passes control to it, by jumping to START 1 at cell 04000B. SEARCH PART 1 reads in the request records, sorts them and stores them in the reserved block area in the COMMON STORAGE. It then reads in SEARCH PART 2, which overlaps that portion of SEARCH PART 1 from cell 04000B to 05200B. This overlap is referred to as SEARCH COMMON. SEARCH PART 2 conducts the search on the Document and English Tapes, and passes control back to the MCS when the search is terminated, and an output tape is printed.

This system is limited, in that a quick retrieval of technical sources is not feasible due to the inherent administrative problems of printout and distribution, and particularly in view of the present off-line printer. The on-line interrogation and display capabilities of the demonstration program as incorporated into the SABIR2 system will provide this needed link with the real-time problem of retrieving information exactly when it is needed.

The response time of the system to retrieve a specific record is, of course, a function of the location of the record on the Document Tape. These records are stored in ascending order of accession number, i.e. order of receipt, on the Document Tape, with the most recent addition at the end of the tape. The longest time to retrieve any one record would be a function of the tape speed and the tape length. Assuming that a full Document Tape was on the tape unit, 2400 feet would be searched at a speed of 150 inches per second. This would give a maximum search time of 2.5 minutes, with no stops between records. However, stopping between each record, for the comparison, introduces a factor of 1/150 seconds for each record. Assuming 15,000 records on a full reel of tape, the maximum search time would be about 4.2 minutes. The average time would then be 2.1 minutes to retrieve any record requested. It should be noted that the English Title File Tape is paced approximately with the Document Tape, to minimize the actual retrieval time of an abstract once the match has been found. This relatively slow method of search and retrieval, necessitated by the use of tapes, is another limitation of the system. The limitation will become more apparent and more severe as the Technical Reports bibliography is increased and additional tapes are added.

III. PROGRAMMING

A. General

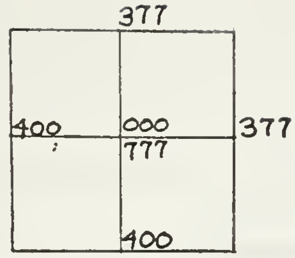
The keyboard input and CRT display program were written in SCRAP, and in a manner compatible with the SABIR2 MASTER CONTROL SYSTEM and SEARCH routines. Advantage was taken of common symbols and storage locations in order to reduce the memory allocation required for the demonstration program. The various parts of the program were initially written as subroutines to speed the check-out on the CDC 1604. When these subroutines were debugged on their own, the dummy constants were removed, and the entire program was assembled to run with the dd65 on-line. The initial uncertainty about Channel #7 transfer prompted this method of attack, in order to more clearly isolate the system errors.

The space constraint in the dd65 core caused less of a problem than that imposed by the SABIR2 program. Since the SABIR2 routines are read in sequentially and overlap one another in some locations, absolute addressing was used for many of the constants and block addresses. Of course, the insertion of the demonstration program changed most of these absolute addresses, and each change in the final debugging phase once again involved changes to these addresses. Some slack area was available, for instance between the MCS and the SEARCH COMMON portion, and between the end of SEARCH COMMON and cell 04000B. This slack area was used to advantage as explained in Section IV (C). See also Appendix I.

B. Hardware Constraints

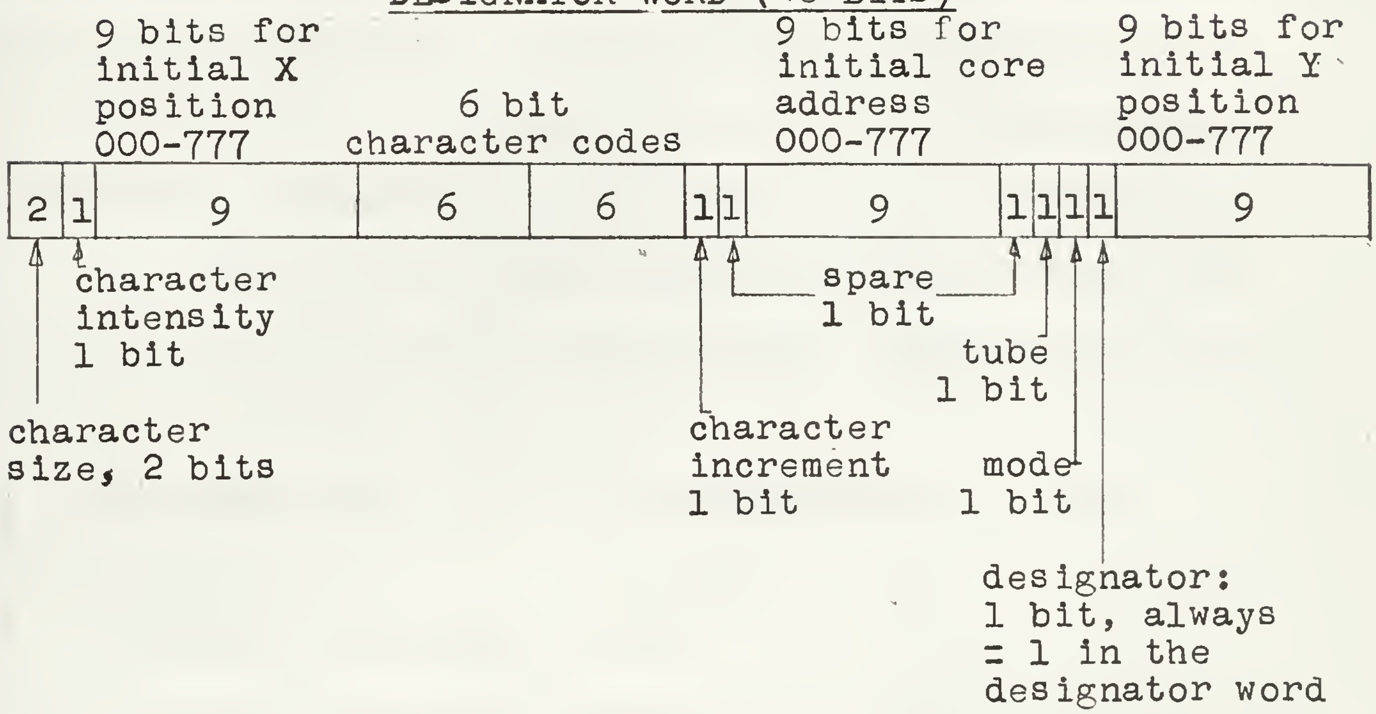
The dd65 must be programmed entirely through a computer, i.e. the memory of the dd65 is neither directly addressable nor can it be programmed to perform any logical operation. A stored program receives input into memory in a specified format and displays this input on one of the two CRTs. Prior to each new transfer of data, the dd65 must be selected. (See Figure 4, Page 21). In the case of this program, the SELECT MEMORY UPDATE from the 1604 on Channel #7 was the select code used. Once selected, Channel #7 stays available for high speed transfer to the dd65 until the select is removed. One 48-bit word is transferred at a time and disassembled into two 24-bit words by the dd65 logic. The first word to be transferred must be a designator word, and each succeeding word must contain a control byte as indicated in Figure 3.

The designator word triggers a sequence of instructions which selects the scope to be used, the size of the characters, the mode of writing (character or vector), the method of incrementing (horizontal or vertical), intensity, and the dd65 memory location. As can be seen from Figure 3, this implies bit-manipulative programming. The only parameter not directly evident in the designator word is the dd65 memory location. In order to examine the designator word to determine where in the dd65 memory an item is located, that portion of the designator word must be broken down into binary and reassembled in octal format in the following



CRT X,Y COORDINATE POSITION CODES

DESIGNATOR WORD (48 BITS)



Character size: 00=128, 01=64, 10=32 characters per line

Character intensity: 1=bright, 0=normal

Character increment: 0=horizontal, 1=vertical

Mode: 0=character, 1=vector

Tube: 0=left, 1=right

SUBSEQUENT 48-BIT STRING WORDS

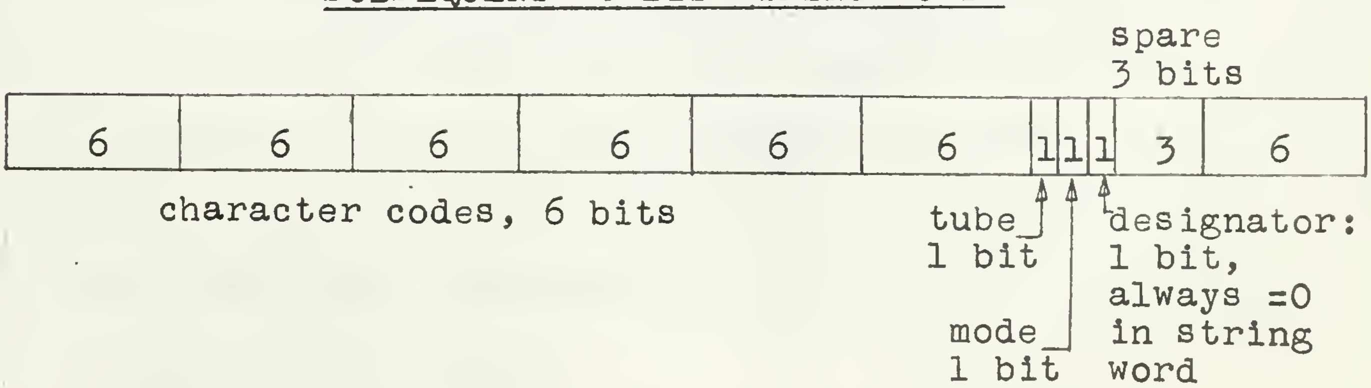


FIGURE 3

manner, e.g.

Designator word	23770000<u>0560</u>1777
Location: Binary format	00<u>0101110000</u>
Reassembled memory location	<u>270</u>

It should be noted that only the even numbered locations in memory are addressable through the designator word, due to the 24-bit size of a word in memory. If an odd location is inadvertantly addressed, the proper display will be lost.

Referring to the above sample designator word, the following can be noted by inspection, reading from left to right:

Character Size	-	64 characters per line
Intensity	-	Normal
Initial X position	-	377

There are no characters transmitted in the designator word.

Character increment	-	horizontal
Tube	-	left
Mode	-	character
Initial Y position	-	777

The X, Y position codes (See Figure 3) locate position 000 as the center of the scope. The code increases from 000 to 377 along the positive X and Y axes, until the top of the scope and the right hand edge are reached. At this point the display is cyclic, and "wraps around" the scope. The code continues to increase from 400 to 777 = 000 along the negative X and Y axes.

After the designator word has been transmitted and located in the dd65 core memory, the subsequent string words follow sequentially in storage, and are displayed and incremented on the scopes as prescribed by the designator word. The designator bits in the string words are a source of difficulty in programming, and involve a sequence of byte shifts and several indexing operations. Thus, the string words evolve with seven display characters, out of the eight which are transmitted.

The provision for two display characters within the designator word was utilized for spacing, rather than display transmission. Since the display characters are incremented horizontally, but display lines are not incremented vertically by the dd65 logic, only one line of display can be transmitted before the designator word must be changed to reorient the Y position. Using 64 characters per line, the transmission of 9 string words (each with 7 display characters) plus two characters in the designator word would give a "wrap-around" display. For this reason, the designator word characters were left blank, and the starting point for each display line was given as X = 367, which, in effect, displayed one blank before wrapping around, displayed another blank at position 400, and then started the actual display line of 63 characters with an indentation of one character space on each line.

The Y axis incrementing could be done by either setting up a table of designator words, or by incrementing the initial designator word in the last nine bit-positions. This second method was chosen to save the space that the table would have taken up in memory. This solution involves one difficulty, however, namely the changeover at Y = 777 = 000. The bit position immediately to the left of the Y-position bits is not a spare bit, as is the case with the initial memory address portion of the designator word. Thus, the last nine bits could not be incremented beyond 777, but rather the word had to be re-oriented to step over this problem.

An additional difficulty occurs since the CDC 1604 "63" instruction (OUT) transmits words from the bottom of a list to the top, in a manner similar to the working of the "64" instruction, Equality Search. The "OUT" instruction is one of the two high-speed data transfer instructions (the other being INT), for Channel #7 [6]. This instruction transfers a number of words, as specified in an index register, from a sequence of addresses which begin at the location specified by the execution address. The transfer takes the first output word from the last address and reduces the index register by one for each word transferred. The transfer time is (4.0 + 4.8n) usec, where n is the number of words to be transferred. If, for instance, the following phrase were translated and prepared for output, its appearance in the CDC 1604 memory would be as shown below. (A space is indicated by "_").

TEST. THIS IS A TEST PROGRAM UTILIZING A CHANNEL 7 TRANSFER OF INFORMATION BETWEEN THE CDC 1604 AND THE DD65 DISPLAY UNIT.

1604 Format:

IT.___6_

PLAY_U6N

D65_DI6S

D_THE_6D

1604_A6N

THE_CD6C

ETWEEN6_

ATION_6B

_INFOR6M

2367000014245357 Octal Designator for second line

SFER_06F

_7_TRA6N

CHANNE6L

ING_A_6_

_UTILI6Z

PROGRA6M

A_TEST6_

HIS_IS6_

TEST._6T

2367000014005377 Octal designator word for first line

The numeral 6 which appears in each string word is a programming device to get the proper control bits inserted into the string word. For display on the left tube, the control bits must all be zero, as indicated in Figure 3. Since the numeral 6 is translated as the BCD character 06, this places the zero bits in the proper string word location.

Any octal number could have been used, since any "1" bits fall in the spare bit positions, and do not affect the display.

A further example of this may be seen in the instructions to the operator which are to be displayed on the left scope of the dd65. This display was hand-coded, as above, since it is not variable, and was located in SEARCH PART 1 of SABIR2. The translation program for the display of the abstracts on the right scope produces the same effect from any input the size of the abstracts encountered in the library search routine.

A complete series of sense-select operations is required to allow one input to the CDC 1604 via the dd65 keyboard. (See Figure 4). When a keyboard #1 hit is made, the keyboard remains mechanically locked until a transfer from this keyboard is made to the 1604. In this program, the 1604 first senses for a hit, selects an input from the sensed keyboard, and then goes into a wait loop until the selection is made. For each key hit this cycle is repeated, this in marked contrast to the single select necessary for the display routine. One other peculiarity of the keyboard "select-sense" cycle is that before the keyboard can be selected again, another different select must be made. (See Figure 4). In this case, SELECT RADAR TO AUXILIARY EQUIPMENT was used. This sequence is required because the keyboard must be positively de-selected before it can again be successfully selected. One method of de-selection is to consciously

PARTIAL LIST OF EXTERNAL FUNCTION CODES (dd65)

Function	Code	Input word format
Select Kbd 1 for input	77140	0...0FFFFFF
" Kbd 2 for input	77120	
" Memory update from 1604	77010	
" Radar target data to Auxiliary Equipment	77001	
" Interrupt on Kbd 1 hit	77105	
" Interrupt on Kbd 2 hit	77103	
Release interrupt request	77111	
Remove all interrupt selects	77121	
Sense Kbd 1 hit (full exit)	77172	
" Kbd 2 hit "	77175	
" Kbd 1 not hit "	77173	
" Kbd 2 not hit "	77174	
" Tab not hit "	77157	
" Carriage return not hit (full exit)	77166	
" Kbd 1 not selected (full exit)	77167	
" Kbd 2 not selected "	77037	
" dd65 interrupt "	77156	
" dd65 from 1604 selected "	77010	
" Radar to Auxiliary Equipment not selected (full exit)	77001	

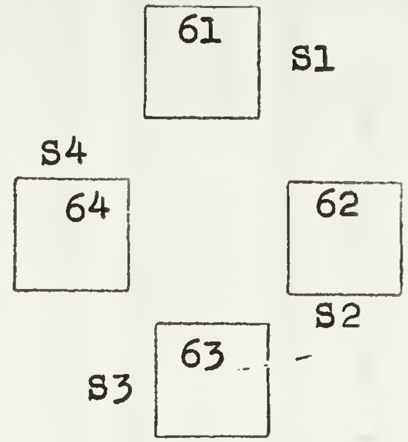
FIGURE 4

select another mode of operation. The most harmless mode to select in this process turned out to be SELECT RADAR TO AUXILIARY EQUIPMENT, since it is a legal select code, but has no meaning within the present hardware configuration.

C. Integration with SABIR2

The interrogation and display portions of the demonstration program were set into the SABIR2 program by some slight rearrangement of the basic SABIR2 program, and by utilizing two-way switches to indicate whether the program was operating in the remote-display mode or not. A sense function (See Figure 4) was located in the MASTER CONTROL SYSTEM, which senses for a keyboard #1 hit each time control is passed to the MASTER CONTROL SYSTEM. Any keyboard #1 hit indicates to the program that a dd65 console operator wishes to interrogate the system. If this hit is sensed, a switch, or flag, is set. This switch subsequently jumps the program around any CDC 1604 operator options or paper tape input, and puts the program into the SEARCH routine. After the SEARCH PART 1 (See Figure 2) is read in, the operator instructions are displayed on the left scope of the dd65. At this point the program goes into a wait loop until the interrogator begins the input on keyboard #1. (See Figure 5). This is a wasted time-period in terms of computer utilization, and is necessitated by the character transmission format from the dd65. A single character is transmitted as part of a 48-bit word, right justified and with leading zeros. (See Figure 4). The 1604 must receive and pack each individual keyboard

	1	2	3	4	
	01 7202 7203	11 7302 7303	21 7402 7403	31 7502 7503	A
	02 7204 7205	12 7304 7305	22 7404 7405	32 7504 7505	B
	03 7210 7211	13 7310 7311	23 7410 7411	33 7510 7511	C
	04 7220 7221	14 7320 7321	24 7420 7421	34 7520 7521	D
	05 7240 7241	15 7340 7341	25 7440 7441	35 7540 7541	E



	A	B	C	D	E
5	41 7602 7603	42 7604 7605	43 7610 7611	44 7620 7621	45 7640 7641
6	51 7702 7703	52 7704 7705	53 7710 7711	54 7720 7721	55 7740 7741

Each square represents a switch on Keyboard #2. Each key (switch) is located by the alpha-numeric coordinates shown along the border of the squares. Inside the squares, the two-digit number indicates the code which is generated by depressing the switch. The 4-digit numbers control the indicator lights associated with each switch. The even codes turn the light on, and the odd codes turn the light off. Switches S1-S4 have no light codes.

dd65 KEYBOARD #2 KEY LOCATIONS AND LIGHT CODES

FIGURE 5

character into a compact word format. Because this handling of the individual hits has to be done by the computer rather than the dd65, the capability of the 1604 is not fully utilized during this time. Of course, a solution to this would be to have the logic of the dd65 arranged so that the key hits are assembled into 8-character words before a transmission is made. In this way, the computer would be available between word transmissions for computation. As presently programmed, this keyboard hit translation routine is located in SEARCH PART 1 of SABIR2.

As each segment of input is accepted, a dd65 console light is activated to inform the operator. Below is the text of the instructions displayed, along with the lights (See Figure 5.) which are activated as each instruction is accomplished and accepted. The program messages and the resulting actions are in capital and small letters respectively.

FIRST ENTER AN 8 CHARACTER (INCLUDING SPACES) IDENTIFICATION. EXAMPLE (JONES JP)

(Activate light 5A)

NEXT ENTER THE DATE SPECIFICATION OF 16 CHARACTERS IN THE FORM DATEYYMMTHRUYYMM WHERE Y = YEAR AND M = MONTH FOLLOWED BY A PERIOD. IF NO DATE, ENTER PERIOD ONLY. EXAMPLE (DATE6302THRU6401.)

(Activate light 6A)

NEXT ENTER THE NUMBER OF SOURCES, (FROM 0 THRU 12), FOLLOWED BY A PERIOD. EXAMPLE (12.) OR (0.)

(Activate light 5B)

NEXT ENTER 1 THRU 12 SOURCES (FOLLOWED BY A PERIOD), OF THE FORM OOXNNNNN, WHERE X = 1, 3, 5, 7, and 9 AND N IS ANY NUMBER. IF NO SOURCES, OMIT THIS INSTRUCTION. EXAMPLE

(00391208.)

(Activate light 6B)

NEXT ENTER THE NUMBER OF UNITERMS (FROM 1 THRU 12), FOLLOWED BY A PERIOD. EXAMPLE (7.)

(Activate light 6C)

NEXT ENTER 0 THRU 12 UNITERMS (FOLLOWED BY A PERIOD) OF THE FORM 00YNNNNN WHERE Y = 0, 2, 4, 6, 8, AND N IS ANY NUMBER. EXAMPLE (00609123.)

(Turn off lights)

SIGNAL ERROR WITH CARRIAGE RETURN THEN A KBD 1 HIT. BEGIN AGAIN.

Hitting the "Carriage Return" at any time during the interrogation phase will cause the program to abandon the input and return control to the Master Control System. If an error is made in input, the Carriage Return key (CR), should, in fact, be hit. The operator then has ten seconds within which to hit any of the keyboard #1 keys to signify that the interrogation is to begin again. The input should then be started over again, from the IDENTIFICATION through to completion. If the keyboard #1 key is not hit within 10 seconds after a carriage return hit, the MCS passes control to the 1604 operator. There is no time limit between hits once the program has entered the remote-display mode.

When all of the required input information is accepted by the CDC 1604, the 1604 starts the search of the Document Tapes. The information required for this search has been previously sorted and stored in the normal input areas by the demonstration program input routine. The usual SABIR2 SORT routine is bypassed completely, since it deals only with paper-tape input format.

Once a match has been found on the Document Tape, the English Title File is checked for an entry which matches the document number. If none is found, the accession number is moved to the output buffer for printing. However, if an English abstract is available, it is immediately displayed on the right scope, and the display program goes into a thirty-second wait loop. The interrogator at this point has four options:

1. Hit the "Y" key on the #1 keyboard to signify "Yes" - a hard-copy printout is desired. The abstract is moved to the output buffer and the search continues.

2. Hit the "N" key on the #1 keyboard to signify "No" - hard-copy printout not desired. The search continues and the abstract on display is destroyed in core memory.

3. Do nothing. After thirty seconds of display, the abstract is moved to the output buffer for a hard-copy printout, and the search continues.

4. Hit the Carriage Return to abandon the search. Control will then be returned to the MCS. The interrogator has the option of hitting any key at random from keyboard #1 within 10 seconds after this to make another search request. This option is signaled by the activating of light #6E on keyboard #2. When this "request for interrogation" hit is accepted, light #6E is turned off, and light #6D is activated. (See Figure 5)

When the end of the Document Tape has been reached, the

CHARACTER CODES

The following are the octal equivalents of the character codes available for use from the dd65 keyboard, or for display on the console CRTs.

Character	Octal Code	Character	Octal Code	Character	Octal Code
A	61	W	26	}	52
B	62	X	27	{	75
C	63	Y	30	[	17
D	64	Z	31	]	32
E	65	Ø	12	(	34
F	66	1	01	)	74
G	67	2	02	:	15
H	70	3	03	||	56
I	71	4	04	;	77
J	41	5	05	.	73
K	42	6	06	↑	55
L	43	7	07	,	33
M	44	8	10	→	35
N	45	9	11	<	72
O	46	CR	76	≥	37
P	47	Tab	36	≠	14
Q	50	Space	20	=	13
R	51	Blank	00	/	21
S	22	$	53	*	54
T	23	∧	57	+	60
U	24			-	40
V	25			'	16

FIGURE 6

scopes are cleared and an END OF SEARCH message is displayed. If another interrogation is pending, keyboard #1 must be hit (on any key). Light #6E will be activated as a reminder of this for the operator. In this way, the program will immediately return to the remote-display mode. Otherwise, the MCS program may either accept a paper tape input or return to CDC 1604 operator control before another dd65 input can be sensed. This provides more efficient use of the computer, but allows less leeway for the interrogator in the event that there is a series of requests to be made.

In order to preserve as much of the space as possible in the output buffer area, advantage was taken of the spaces available within the SABIR2 program proper. Thus, much of the programming was done in short subroutines, and these subroutines were packed into the available spaces with the intention of preserving output space rather than preserving the sequential order of use of the subroutines. This way the entire capacity of the output buffer was preserved. However, the available space between SABIR2 program segments was almost exhausted, and any significant additions to the program in the future will have to use part of the present output buffer.

D. Proposed Modifications

The system, as written, works in a very straightforward manner, with a minimum of "bells and whistles". However, in order to make it more user-oriented, some modifications might be made. The following paragraphs present some of these suggested improvements, with a short discussion of each.

An interrupt capability might be the most useful, and yet, at the same time, the most difficult change to make. An interrupt capability would make the system appear to respond more in the manner of real-time interrogation and display network than it now does. One approach would be to read out the paper tape requests now in the core memory onto a Scratch tape, destroy the bibliographies presently in the output buffer, and accept the keyboard input. At the conclusion of the interrupt the requests could be read in again from the Scratch Tape and the search on them started over. A second approach might be to reserve the first request location in each input block (IDENT, UNITERMS, FRSTDATE, LASTDATE, SOURCES) for interrupts only. Under these conditions the present requests and bibliographies could remain in memory. However, a record count from the Document Tape would have to be programmed, in order that the normal search could pick up where it left off when the interrupt was satisfied.

To help the operator catch his own input errors, his keyboard hits could be displayed on the left scope just below each operator instruction. There is presently enough space left in the dd65 core memory to display this, and also room on the scope. Each left-scope display line may have to be incremented for a full nine words along the X axis, in order to allow more space vertically for this display addition. The left display is now seven words per line vice the allowable nine words, but for ease of programming only.

A reminder of the thirty second time limit might be well worth the programming effort. On the right scope just below the displayed bibliography, the time in seconds could be shown in a count-down display manner.

The thirty second wait loop could be integrated with the search for the next match, rather than as presently arranged. This would involve some rather delicate manipulation of SEARCH PART 2. Before proceeding with a modification of this type, several other questions should be investigated, e.g. what is to be done when (1) the time limit runs out and a search is still in progress, and (2) when another match has been found but the time limit has not run out.

The thirty second wait loop may prove to be too long, in practice. The constant TWOTIMER controls this time limit, and is presently set equal to 1800. Since the real time clock is of the 60 cycle variety, the number of seconds desired in the wait loop, multiplied by 60, will be the new value of TWOTIMER.

The legibility of the operator instructions could be improved by using a 32 character per line mode, and splitting the instructions between both scopes. Since these instructions are not needed after the bibliography display begins, this display may be destroyed after the last input character is accepted, and recovery instructions concerning the use of the Carriage Return displayed on the left scope.

The display routine can be made more sophisticated by programming a scanner to look for a space near the end of each display line. This would then be the break point of the line, and the next line of display would start with a new word. This will make the display longer along the Y axis, since presently the display line is filled for the complete 64 characters.

A routine to search the input for illegal characters or illegal combinations would be a desirable feature. This routine could display the error diagnosis and permit that character or input record to be keyed in again, rather than the present return to the beginning of the interrogation.

One item falls in the category of revision rather than modification. Another mode of operation can be programmed in addition to the present mode. This would provide for a moving display, showing a continuous list of requested items. Perhaps the title, author and accession number would be more appropriate than the entire abstract in this case. One method of solution would be to set up a table of designator words to be continually substituted in display lines in order to make each line move up, and finally disappear. The moving display mode could be signaled by a hit or a sequence of hits on the keyboards, e.g. a keyboard #2 hit while the program is in a wait-loop. This would take up some of the output buffer for the display list, but the buffer would not be needed for any extensive printout capability in this mode. A HOLD signal would need to be provided, in order that the display be kept stationary for the user while a HOLD key on keyboard #2 is depressed. This mode would be most useful when examining long lists of items, such as all the publications from a given source, or the matches of only one or two commonly used Uniterms.

IV. OPERATING INSTRUCTIONS

A. Sequential List of Operations

a. Turn on 400 cycle current for dd65. The power switch is located in Room 501, Spanagel Hall.

b. Check dd65 cable connections, located at bottom right side of the dd65 logic cabinet against the diagram of Figure 8.

c. Hold the Master CLEAR-RUN Logic Switch down (in the CLEAR position) momentarily, press the start button, and place the switch into the RUN position. This Logic switch is located on the front panel of the dd65 logic cabinet.

d. Check the HALF MEMORY Switch, which is located on the top inside left edge of the logic cabinet. It should be in the "Down" position.

e. Set the rotary Mode Switch, located on the front of the dd65 logic cabinet, to "1604 Only".

f. Check the cables inside the CDC 1604 Main frame for the Channel #7 connections. (Figure 8)

g. Mount tapes on the CDC 1607 tape units, according to Figure 7.

h. Bootstrap the system tape as follows:

```
Clear up and down then step.
ENTER 270         in  A Register
ENTER 20000003    in  Function Code and Execution
                      Address
                    Step
ENTER 74032011    in  FC and EA
                    Step
ENTER 74300030    in  FC and EA
                    Step
ENTER 75000030    in  FC and EA
               Step and Start
```

i. When bootstrapped, the program will type "DATE" on the CDC 1604 console typewriter. In response the operator must type today's date in the following format:

mm/dd/yy

mm is a two-digit representation of the month, e.g. 08.

dd is a two-digit representation of the day, e.g. 30.

yy is a two-digit representation of the year, e.g. 64.

j. Wait for light #6E (See Figure 5.) to be activated on the dd65 console, then hit any key from keyboard #1. This will put the program into the remote-display mode.

k. Follow the instructions which appear on the left scope of the dd65 console. Light #6E will be turned off and light #6D will be activated to indicate that the remote mode has been entered.

BLOCK DIAGRAM OF CHANNEL AND UNIT

ASSIGNMENTS FOR SABIR2

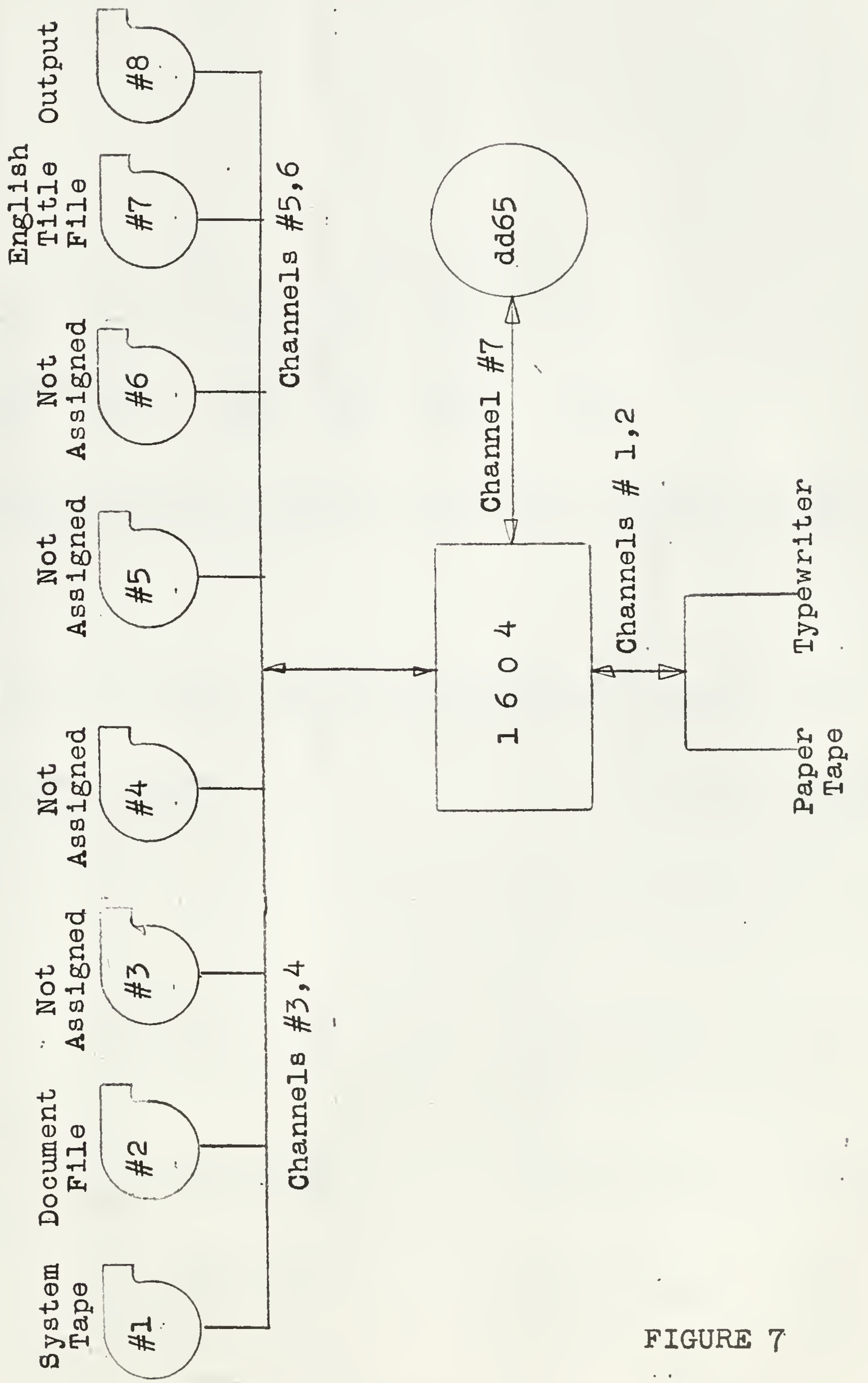
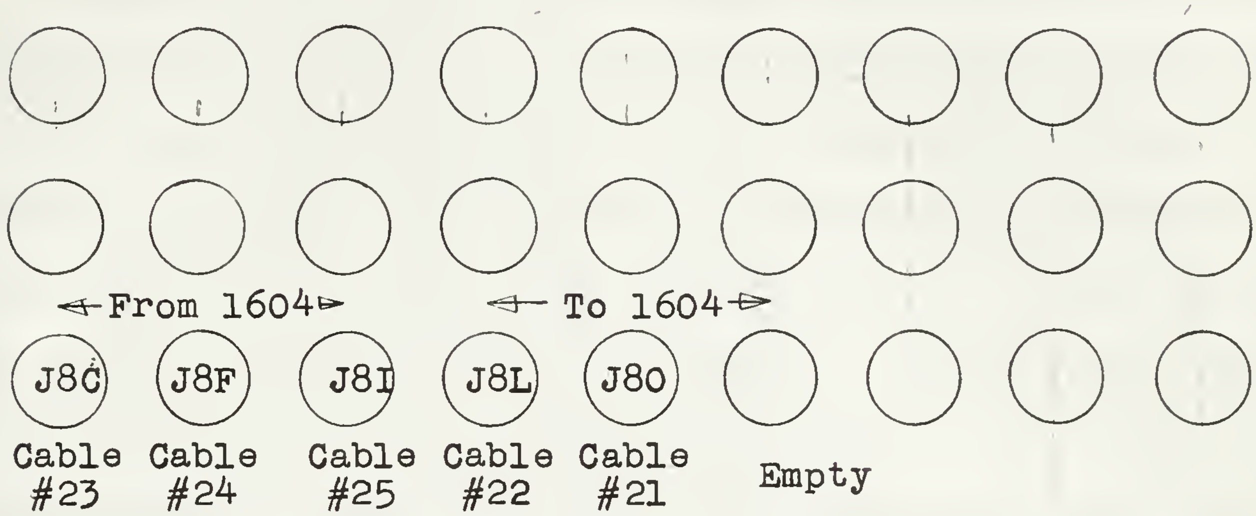


FIGURE 7

CHANNEL #7 dd65 LOGIC CABINET CABLE CONNECTIONS*



*Connections not marked either do no concern Channel #7, or are already installed in their proper sockets.

Channel #7 CDC 1604 Main Frame Cable Connections*

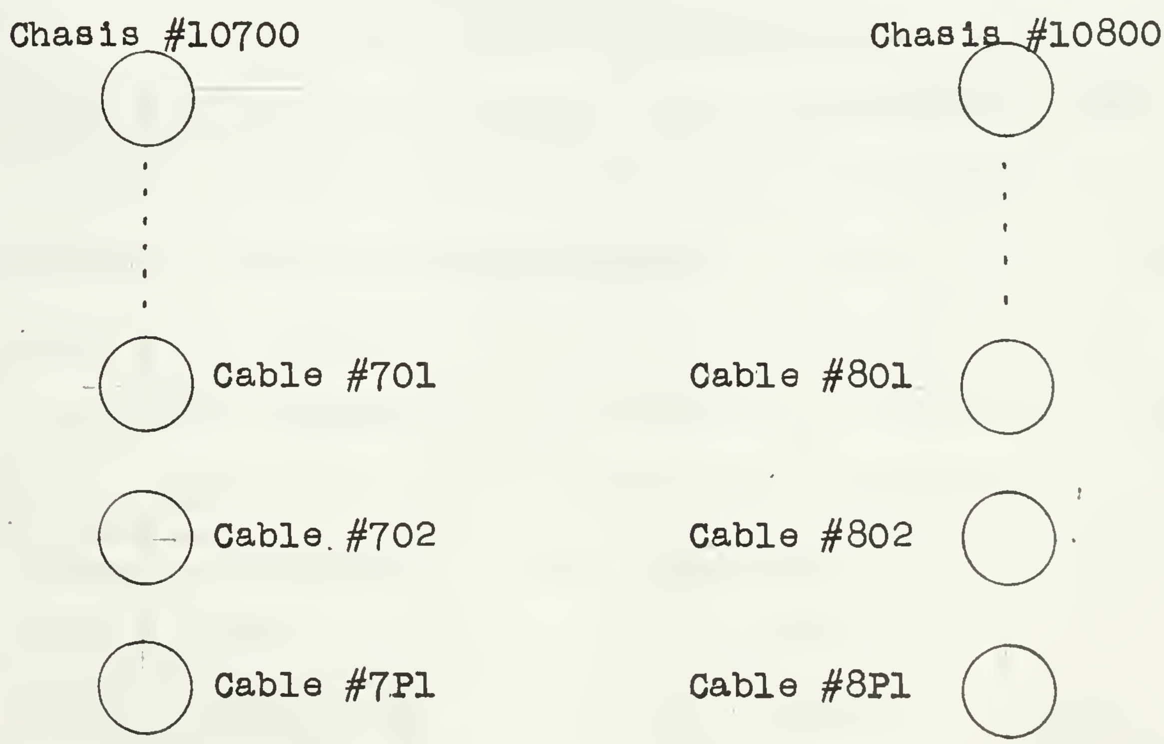


FIGURE 8

V. DISCUSSION

A. Generalized Information Retrieval

Automated Information Retrieval is still very much in its infancy when judged by any yardstick which equates progress in this field against developments in computer technology. Although computers and Information Retrieval are intimately associated, it appears that the separation of the two is at the point where the ability to program a sophisticated IR routine does not begin to match the ability of the machines for logical accuracy and speed. The design and production of visual display units to accompany an IR system are themselves not very much further along in technology than as experimental models developed for specific applications. A survey of the many surveys of Information Retrieval reveals a substantial number of examples of the very real problems facing the professional person who wants to find the latest information on a given subject. It is pointed out that for any research costing less than $100,000 it is cheaper to duplicate the desired research than it is to try to discover if it has been done before, and if so, what the results were. Another frequently mentioned fact is that the daily output of technical papers in the world would fill an Encyclopaedia Britannica; there presently being about 100,000 technical journals published in 60 languages [1].

While the accuracy of the above figures may be open to challenge, the situation which they describe is not. Information Retrieval at a local level has been accomplished successfully in many instances, both with and without computers. However, IR on a large scale is a very complicated problem, and cannot rely on local procedures for success. The very item which allows local successes in specialized IR routines, at the same time complicates a generalized IR routine. This item is the question of what the information to be retrieved should be called [8]. While a seemingly innocent question, this is the very same problem of indexing which leads to the field of the Boolean Algebras, cybernetics and linguistics in search of solutions; and which determines the method of storage and of retrieval of the information in question. In reality the problem has, then, two aspects, i.e. document storage, and retrieval; and the solution of these two very important problems pivots on the vital question of the method of indexing.

Basically, indexing involves assigning a fixed identifying tag or tags to an item in order that the item may be quickly located at some future time. While this may be solved within the restricted vocabulary of a command and control system, or an inventory control problem, finding such an indexing system of general applicability has proved to be considerably more complicated than it appears at first glance.

Almost all of the present systems of indexing in Information Retrieval involve coordinate indexing. This is a structure similar to a decision table [4], where the particular items required are represented by the logical intersection of two or more descriptors. This is a desirable feature, since only one search - for the intersection - need be made. Other systems can produce multiple searches. If a restricted list of tags is used, so that the searcher knows that he is using an admissable tag to search on, a reasonably reliable system can be developed. It is evident that the weak link in this system is the person who assigns tags to the items. If he mis-interprets the article or assigns an erroneous tag for some other reason, neither computer technology nor IR technology will be of any great assistance to the searcher.

In order to avoid as many of the pitfalls of indexing as possible, the SABIR2 system uses Uniterms as its coordinate index system. This is a free-list system (the admissable list is unrestricted) which uses the significant words, or keywords, of the document as the tags. This permits a great depth of indexing and allows for the documents which deal with uncommon subjects. It has been suggested that the author of the article himself choose the Uniterms to avoid the obvious difficulty of indexing by another party. The use of Uniterms or similar devices is, of course, suggested only where a more obvious choice is not available, such as stock number or service number. Whatever method is used, a workable index organization is a prerequisite to success in the Information Retrieval field.

Information Retrieval is of vital concern to any manager simply because correct decisions can only be made with up-to-date information. Immediate access to any item of information within a military commander's purview is now a requirement of higher command in the Armed Services. This is particularly true in view of the speed at which developments take place in the present military environment. Anything less than immediate evaluation of all pertinent information would be generally useless. While not as strikingly obvious, the retrieval of information in a logistics system is none the less important. Command decision in the management review of provisioning, the audit of financial posture, and the control of inventory are some of the areas now available for the application of on-line retrieval of information on a real-time basis.

Projecting the needs for a remote interrogation and display capability into the future, shows the necessity for thinking of systems design in terms of future capabilities rather than past performance. Starting from the hardware, it can be readily seen that using magnetic tape as the storage medium soon will be unacceptable from a time and maintenance standpoint. For a file of any practical size, the time to accomplish a search of the tapes becomes intolerable. The experimental program discussed in this paper treats time as though it were an expandable quantity, when in practice, time may well be one of the strictest constraints. The tape search and the limitation of the first generation display unit acted together to produce a situation which served well for experiment, but which is not recommended in a crucial working environment. It is evident that a random access storage medium must be employed for maximum efficiency, and that a display and interrogation unit must be of a non-tempermental, and logically efficient design to coordinate with the computer. The random access requirement is not a condition imposed only by IR needs; rather IR is but another in a long list of applications which depend heavily on a random access capability.

The question of the interrogation and display unit itself should be considered from the viewpoint of the problem-oriented user rather than the machine-oriented programmer. However, a waste of main computer time by an inefficient interrogation or display system cannot be considered a justification under any circumstances for a system which purports to be user-oriented. The most significant results of an IR capability are going to be achieved, not by a technician checking, for instance, to see if a correct update was made, but by a manager who is asking for information to finance his decision. It is for this executive that the equipment application must be considered. Small size, silence of operation, and ease of interrogation will be the key features looked for by this user.

Interest in the retrieval and display of information in military activities is passing from the real-time operational phase characterized by the target acquisition and assignment problem, to the management decision area. It is in this new area that advances are slow to be realized. With a working system (and this implies indexing, programming and hardware), a military commander should have the capability for the retrieval and display of intelligence, politico-military, or administrative information upon his request. Many of the staff research assignments which necessarily precede a delicate decision, can be shortened in both time and effort by having the request information requested and displayed for the commander at his desk. Predicted developments in the field of large screen display units should bring about increased use of displays in staff operation centers, war rooms and similar information gathering, intelligence briefing, or decision-making environments.

In non-operational commands, management decisions are also predicated on the information available, and the more timely the information, the more appropriate the decision should be. With an IR system, management does not have to manage only by exception, since for any organization information is on immediate recall. The Stock Control Officer, for instance, has instantaneous access to the stock position, demand history or any other recorded history, on any stock item. The information need be no older than the time between the interrogation and the display, in contrast to the age of the information on a printed report. The display need not be a stationary seek-and-display operation, but could provide a rolling presentation of a sequence of items for review. This is the counterpart to the periodic written report.

In a generalized IR system, the auditor recalls and pursues an audit trail, the Legal Officer recalls the totality of references for a given circumstance, and the Personnel Officer reviews the personnel jackets - all using the same hardware equipments, and all within the working area of the officials. IR techniques are promising as useful tools for all levels of management; and as the terms of reference of the executive become larger, the IR applications become broader in scope. With the properly programmed security checks on the user; and security classification, sub-paragraphing or tagging of information, this executive application can be extended in a manner similar to the command and control application.

The implication of top management use of IR is that in the utilization of basic IR equipment and techniques to further the executive decision process, management by exception will now be the exception rather than the rule. Complete control of executive responsibilities must include the capacity to retrieve a detail upon demand, and not just when something goes wrong. These executive officials can be freed from dependence on periodic or specially written reports by a remote display IR. They can have instant access to the files of any department, with the computer acting the role of file clerk. This means that a check can be made on any project or item of recorded information by this official without necessity for a briefing or a written report. Such use can have profound effects on the methods of organization management. This independent assessment of the work of any individual, or the progress of any undertaking could be one of the most important management uses of Information Retrieval.

END

REFERENCES

1. B. Nanus, The Use of Electronic Computers for Information Retrieval, Bulletin of the Medical Library Association Vol. 48, No. 3, July 1960.

2. A. M. Wildberger, Information Retrieval, U. S. Naval Postgraduate School M. S. Thesis, 1961.

3. E. N. Ward, Symbolic Coded Relocatable Assembly Program (SCRAP), September 1961.

4. IBM Manual #F20-8102, Decision Tables, A Systems Analysis and Documentation Technique, 1962.

5. C. A. Haworth, Technical Report #44, SABIR2, 1964.

6. CDC Manual #167, 1604 Computer Programming Manual

7. Data Display, Model dd65 Instruction Manual, Volume I.

8. IBM Manual #C20-8062, Index Organization for Information Retrieval.

BIBLIOGRAPHY

1. C. N Mooers, The Next Twenty Years in Information Retrieval: Some Goals and Predictions, Proceedings of the Western Joint Computer Conference, 1959.

2. J. C. R. Licklider, Man-Computer Symbiosis, IRE Transactions on Human Factors in Electronics, HFE-1, 1960.

3. IBM Manual #E20-8044, An Introduction to Information Retrieval, 1960.

4. IBM Manual #E20-8034, Inventory Control and Material Accounting, 1961.

5. L. H. Hattery and E. M. McCormick, Editors, Information Retrieval Management, American Data Processing, Inc., 1962.

6. C. G. Lawson, An Integrated Display and Control System for Man-Machine Communication, U. S. Naval Postgraduate School M. S. Thesis, 1962.

7. R. P. Warrick, Design of an Automatic Muli-Processing System, U. S. Naval Postgraduate School M. S. Thesis, 1962.

8. IBM Manual #E20-0100, An Information Retrieval Experiment by Saint-Gobain Company, Paris, France, 1962.

9. IBM Manual #E20-8095, Storage, Retrieval and Dissemination of Information, 1962.

10. R. L. Hogg/D. C. Glover, Control System Programming, Remote Computing and Data Display, U. S. Naval Postgraduate School M. S. Thesis, 1963.

11. G. W. Barnard, Information Storage and Retrieval: A Survey, Air Force Systems Command Technical Documentary Report No. AMRL - TDR-63-8, January 1963.

12. W. Griffith, Programming A Remote Monitor and Display System for a Time Sharing Computer, U. S. Naval Postgraduate School M. S. Thesis, 1963.

13. A. J. Owens, The Uniform Automatic Data Processing System: Its Origin and Impact on the Navy Supply System, U. S. Naval Postgraduate School M. S. Thesis, 1963.

14. IBM Manual #E20-8094, Mechanized Library Procedures.

15. CDC Manual #187, Programming Manual for Control Data Satellite Computer System.

APPENDIX I

FLOW DIAGRAMS:

System Integration

Display Section

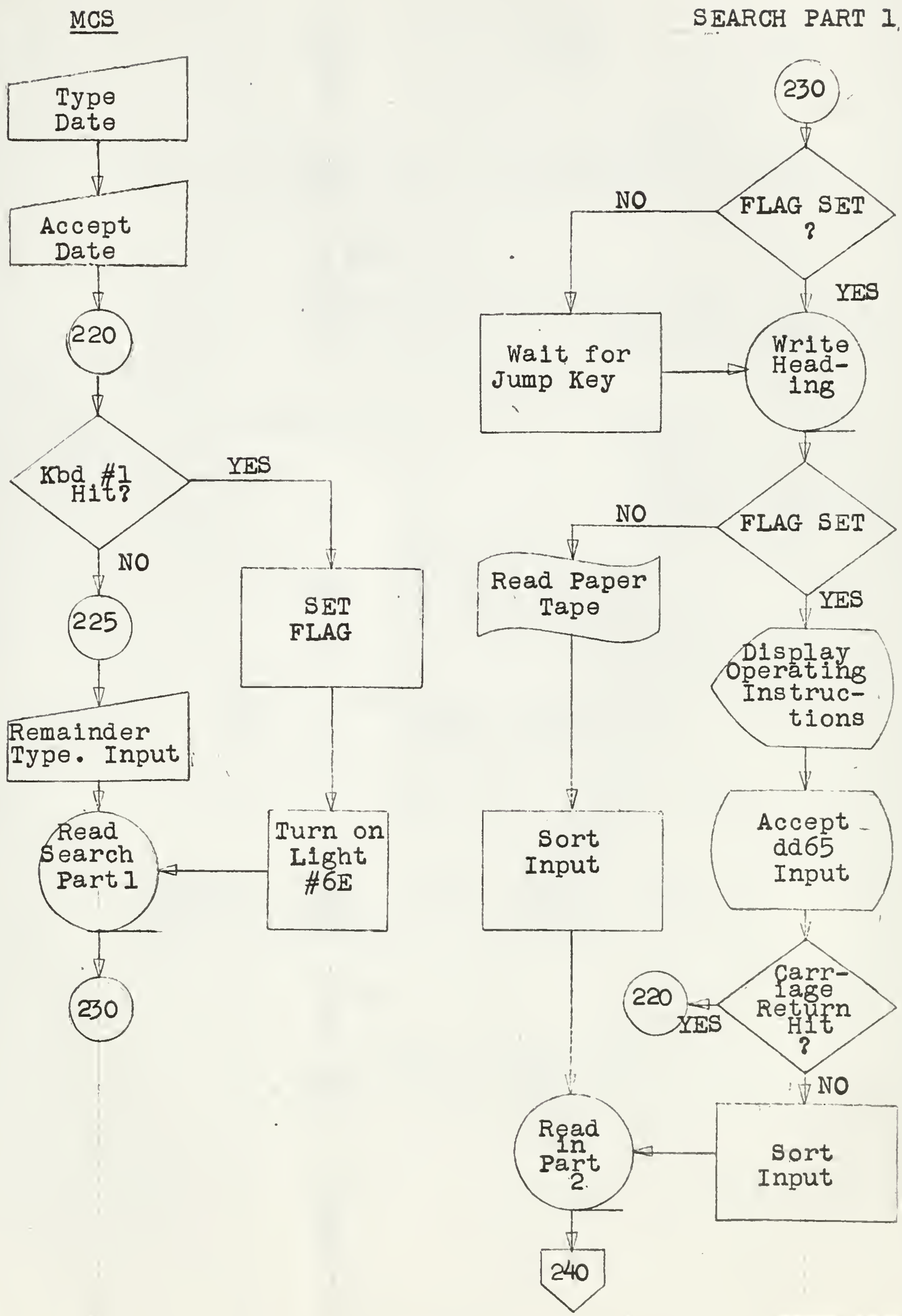
Interrogation

Notes:

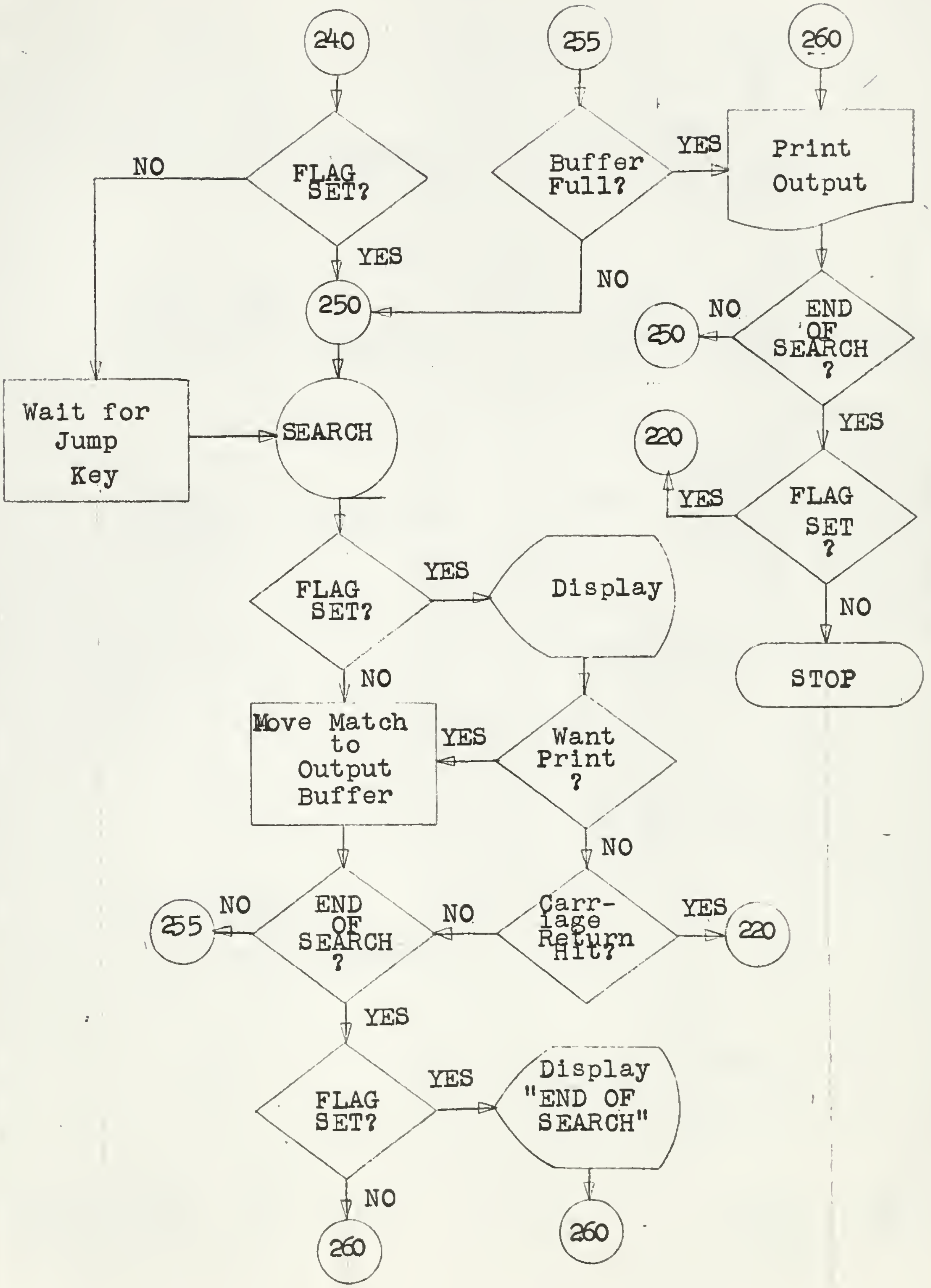
1. Flow diagrams are in accordance with ASA X3.6/12 standards.

2. The detailed coding is not included here, but the program card deck and a listing have been placed on file in the Computer Facility, USNPGS.

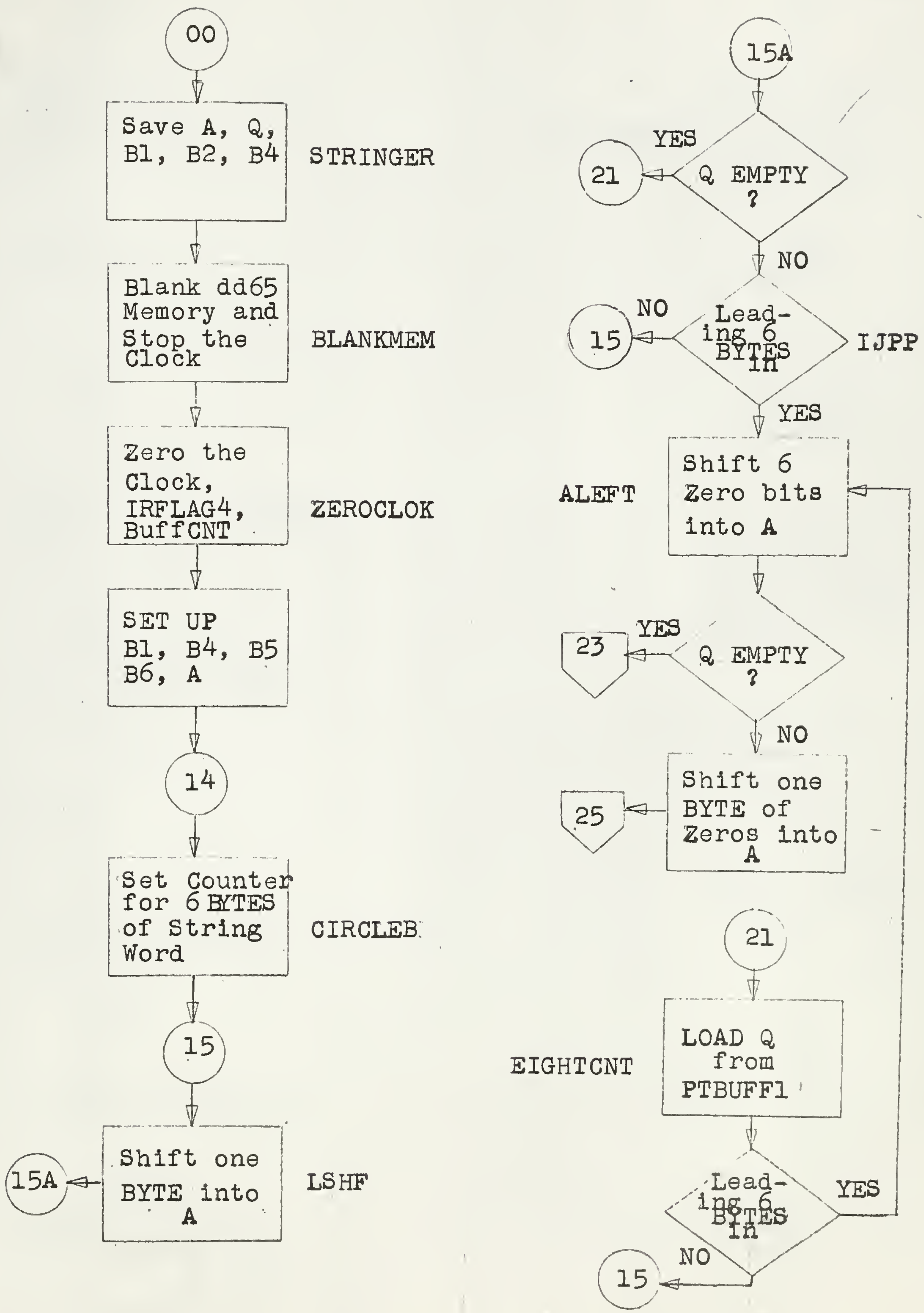
SYSTEM INTEGRATION

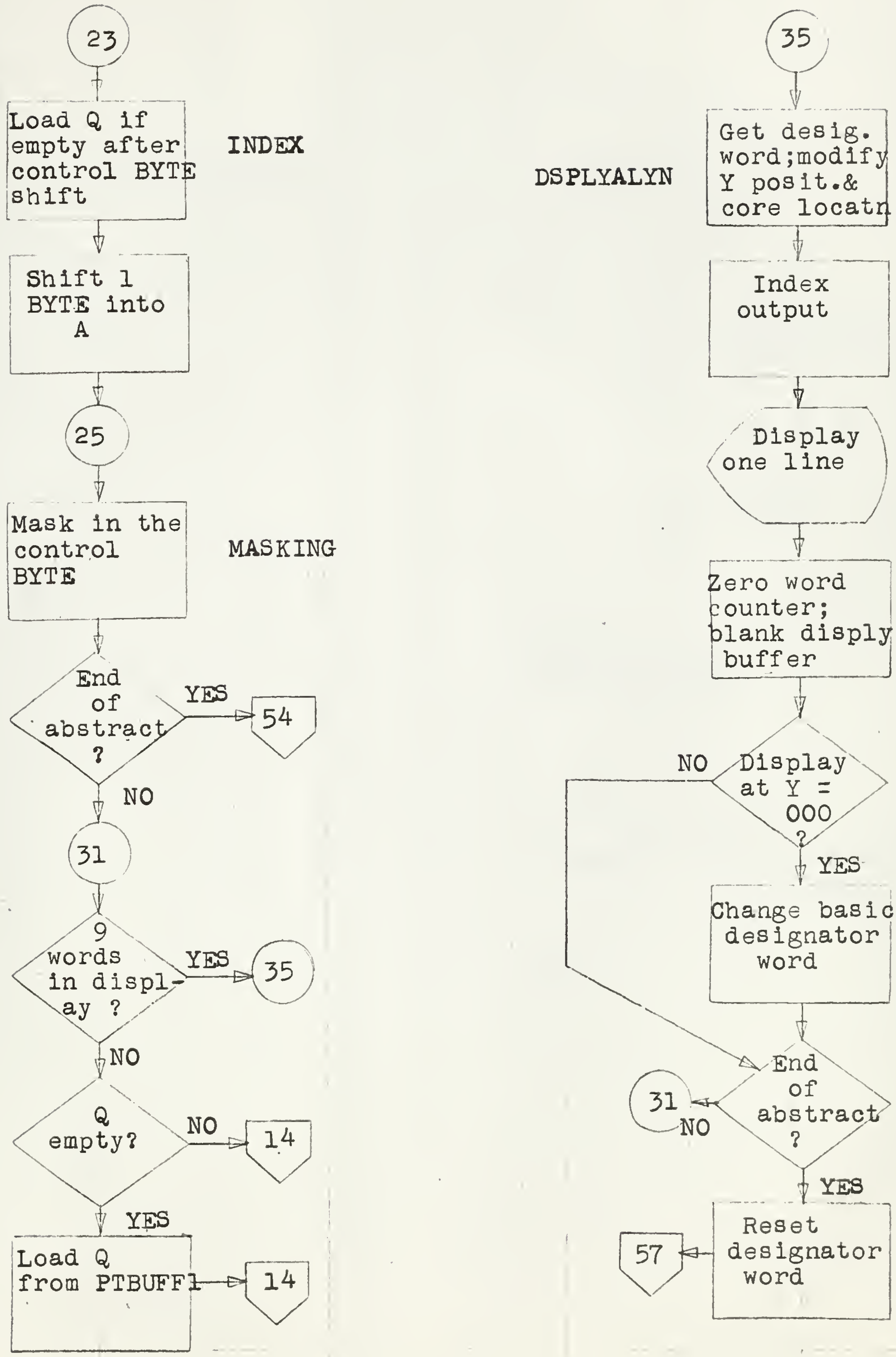


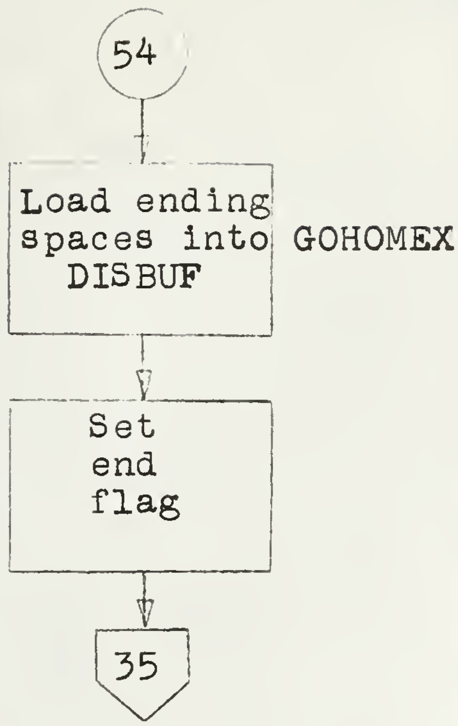
SEARCH PART 2

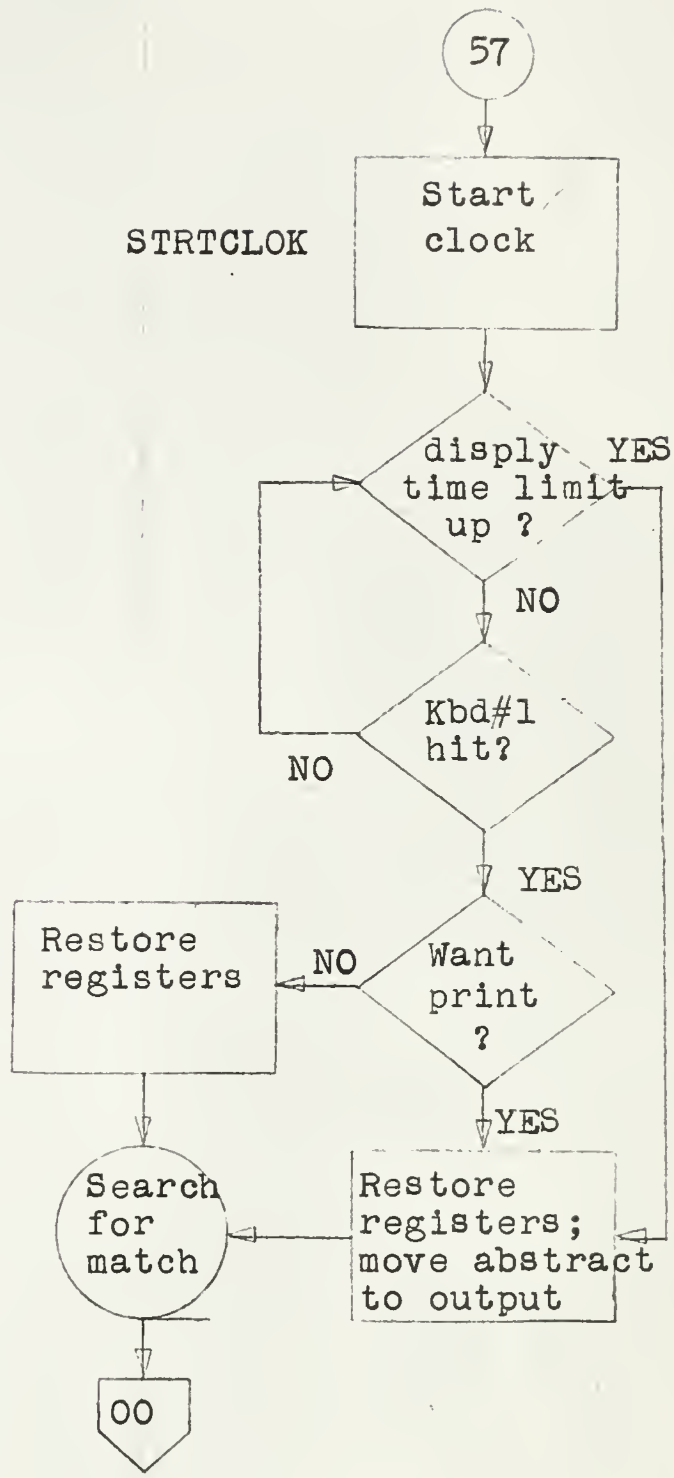


DISPLAY SECTION



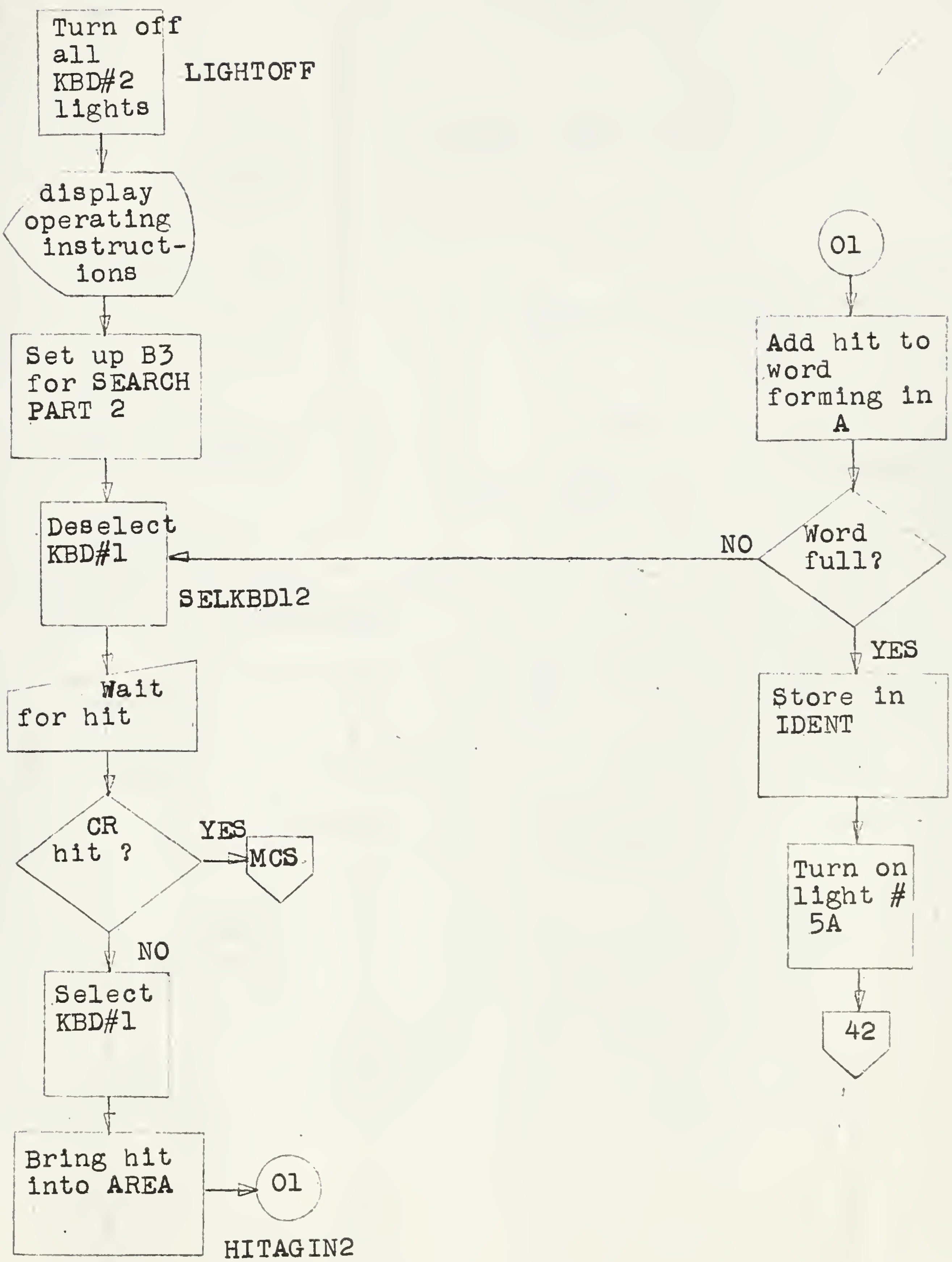

23
Load Q if empty after control BYTE shift
INDEX
Shift 1 BYTE into A
25
Mask in the control BYTE
MASKING
End of abstract ?
YES
54
NO
31
9 words in display ?
YES
35
NO
Q empty?
NO
14
YES
Load Q from PTBUFF1
14
35
DSPLYALYN
Get desig. word;modify Y posit.& core locatn
Index output
Display one line
Zero word counter; blank disply buffer
Display at Y = 000 ?
NO
YES
Change basic designator word
End of abstract ?
31
NO
YES
Reset designator word
57


54
Load ending spaces into DISBUF
GOHOMEX
Set end flag
35

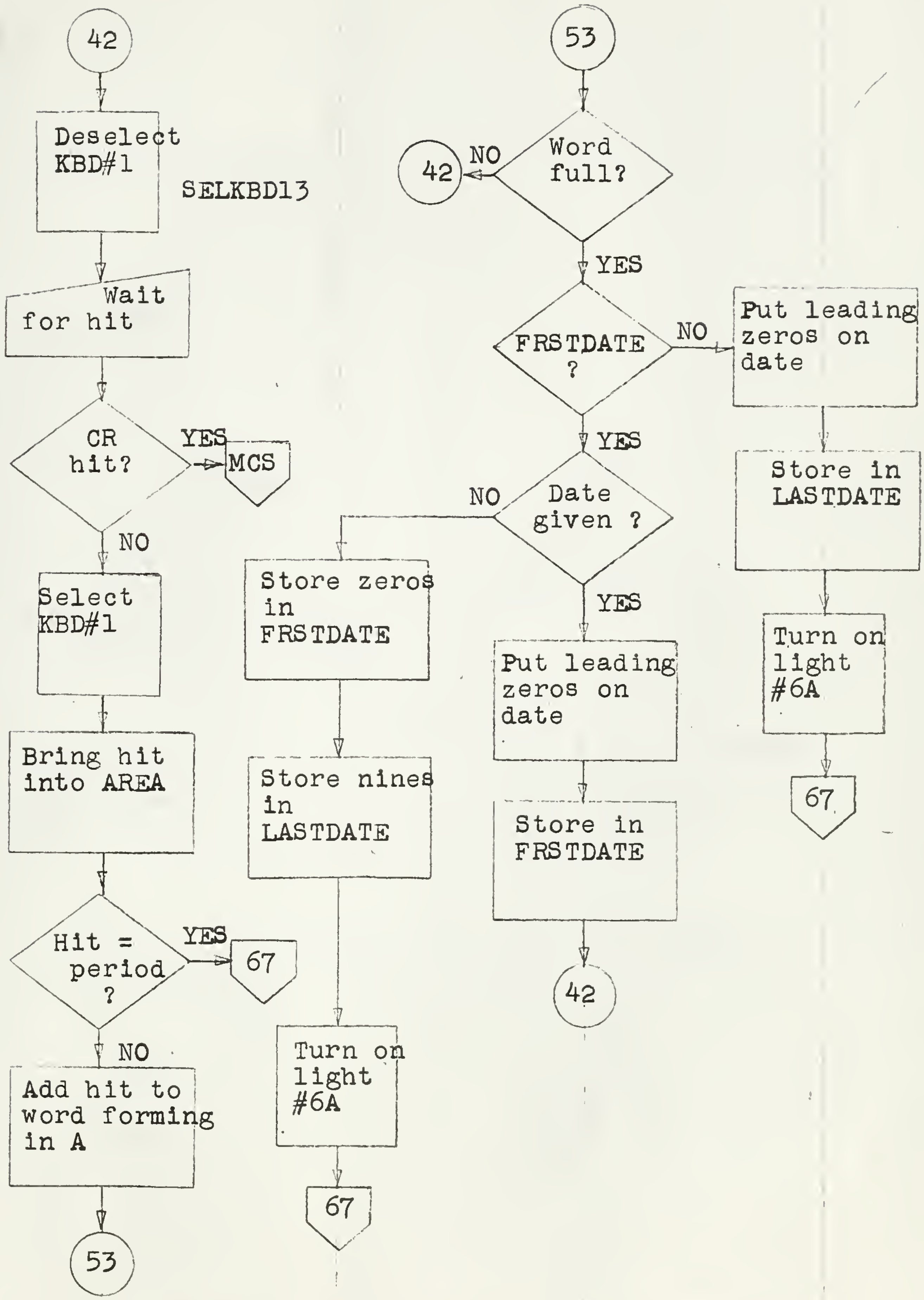

57
STRTCLOK
Start clock
disply time limit up ?
YES
NO
Kbd#1 hit?
NO
YES
Want print ?
NO
Restore registers
YES
Restore registers; move abstract to output
Search for match
00

INTERROGATION SECTION

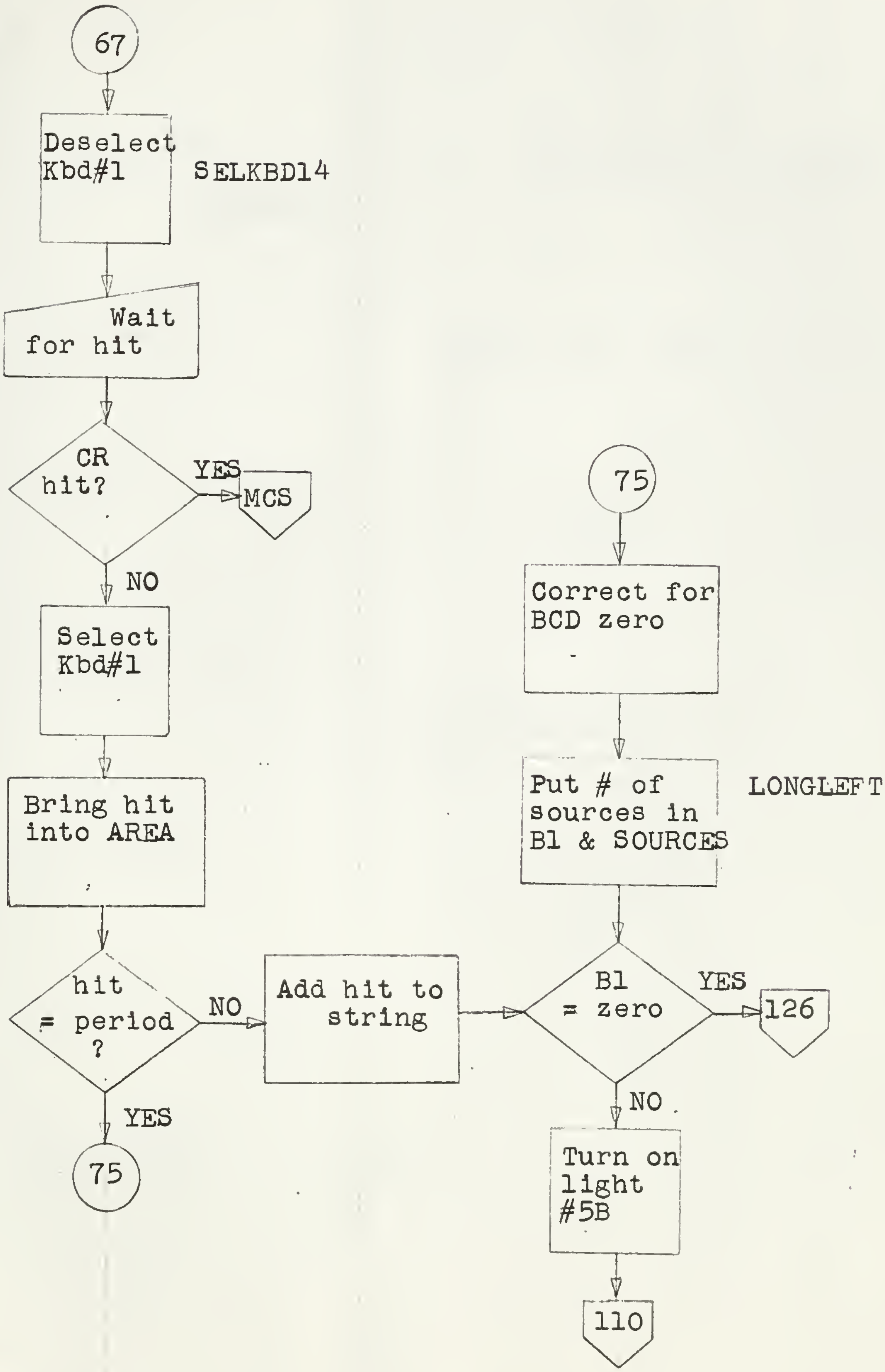
IDENT INPUT



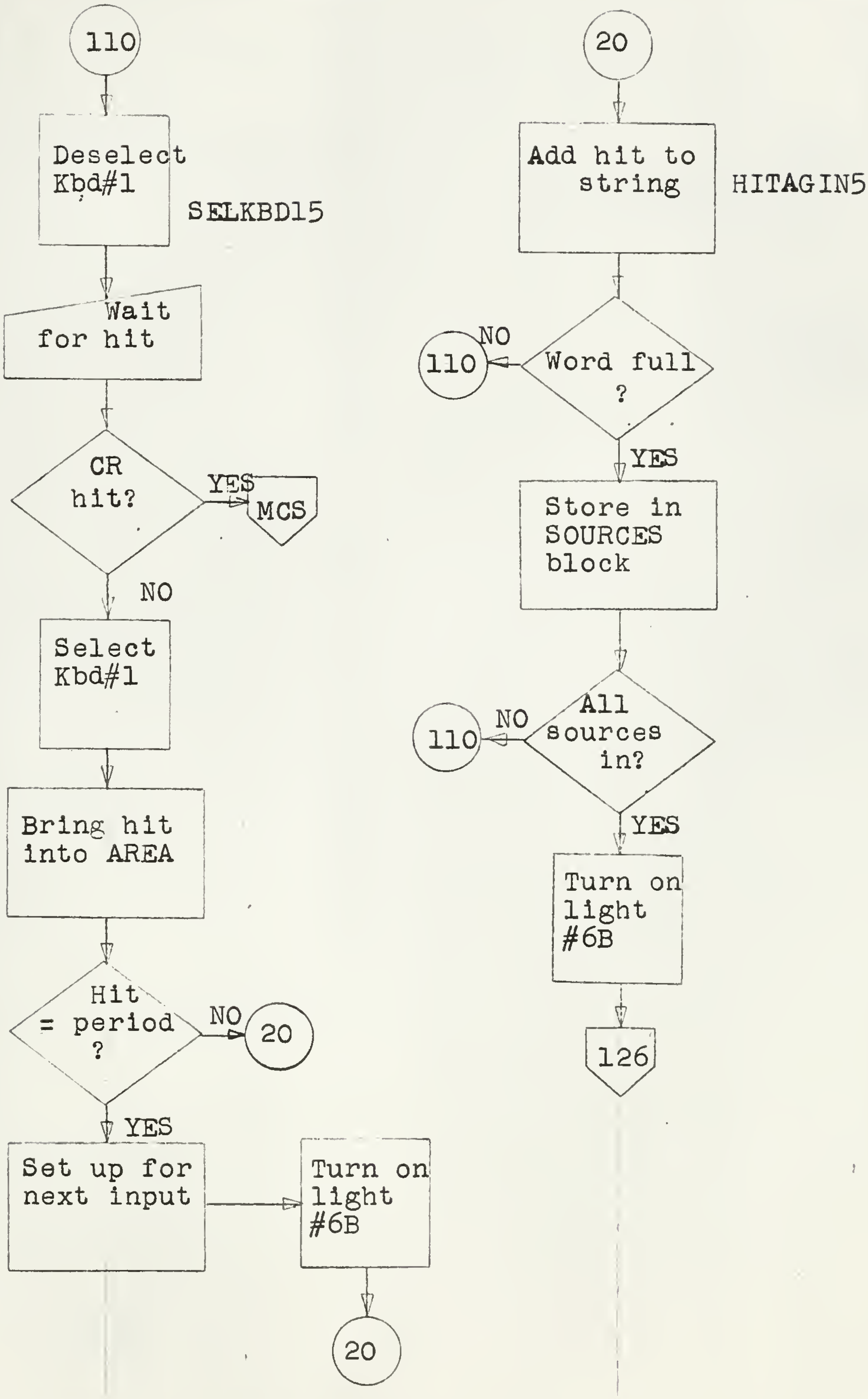
DATE INPUT


42
Deselect KBD#1
SELKBD13
Wait for hit
CR hit?
YES
MCS
NO
Select KBD#1
Bring hit into AREA
Hit = period ?
YES
67
NO
Add hit to word forming in A
53
53
Word full?
NO
42
YES
FRSTDATE ?
NO
Put leading zeros on date
Store in LASTDATE
Turn on light #6A
67
YES
Date given ?
NO
Store zeros in FRSTDATE
Store nines in LASTDATE
Turn on light #6A
67
YES
Put leading zeros on date
Store in FRSTDATE
42

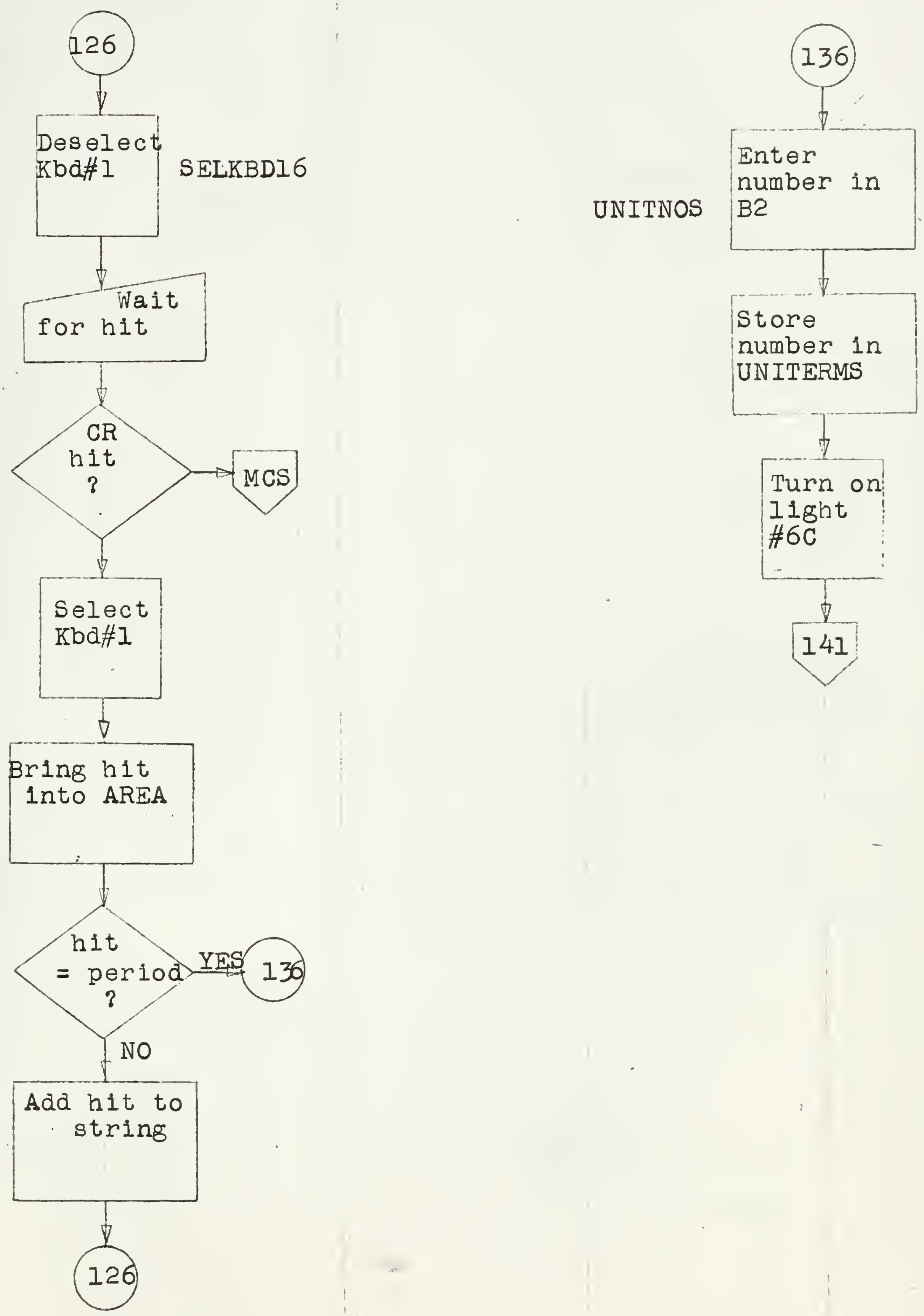
NUMBER OF SOURCES INPUT



SOURCE INPUT



NUMBER OF UNITERMS INPUT



UNITERMS INPUT