POLISH, A Fortran Program to Edit Fortran Programs

by

John Dorrenbacher, David Paddock
David Wisneski, and Lloyd D. Fosdick

Department of Computer Science
University of Colorado
Boulder, Colorado 80309

Report #CU-CS-050-76(Revised) May, 1976

* This work supported in part by National Science Foundation Grant DCR 75-09972.
ABSTRACT

POLISH is a program which will read ANS Fortran programs and rewrite them in a stylized format designed to be easy for the reader and conservative in its use of space. Although some editing features, such as spacing conventions about tokens are fixed, many are subject to user control. POLISH itself is written in ANS Fortran.

Keywords: Fortran text editor.
1. INTRODUCTION

The text editor POLISH reformats a FORTRAN program to give it systematic spacing, indentation, labelling, and other stylistic features desirable for the reader. The edited program has a different physical appearance but it is identical to the original program with respect to its execution characteristics. Figures 1, 2, and 3 illustrate the kind of editing done by POLISH. Style is a matter of taste, but almost any reasonable stylistic conventions will produce a program listing more pleasing to the eye and easier for the human reader to follow. Other FORTRAN text editors are TODY [1], FORDOC [2], FORTREDIT [2]. Particular features of POLISH which distinguish it from others are that it is written in ANS FORTRAN (cf. Appendix D), no manual insertion of editing cues in the unedited text is necessary but such cues may be used, line widths may be controlled, and spacing between tokens is controlled. Since early in 1974 POLISH has been used to prepare the FORTRAN listings which appear in the Algorithms Department of the Communications of the ACM and more recently in ACM Transactions on Mathematical Software; see for example [3].

POLISH is a subroutine subprogram. Parameters which control certain editing features are passed to POLISH through COMMON storage. When called into execution it will read the unedited FORTRAN text from a file and write the edited FORTRAN text onto another file; both files are specified by the user. Aside from certain exceptions noted in Section 2, it is assumed that the text to be edited is a syntactically correct ANS FORTRAN program. There are some circumstances in which it is desirable to inhibit editing of a text segment; most typically in
C EXAMPLE 1.
C THIS IS A NONSENSE PROGRAM ILLUSTRATING SOME FEATURES OF EDITING
C DONE BY POLISH.
C
COMPLEX  Z
EQUIVALENCE  (C,D),(E,F) , (G , H )
DATA C,A/3.0, 2.0/
3  A=1.0
7  B=3.0*C+D*E/F
CD(I , J )=AB(I,J)-EF(I+1)
DO 10  I=1,50
   X=X+Y
   C =C +D
   DO10K=1,40
   E=E+1.0
   DO 10  L=1,100
10  M=M+1
   IF(A.LT. C)GOTO3
   IF(A .LE.D)GO TO 3
   IF(A .EQ. C)GOTO 3
   IF(A .LT. B .AND. C.LE.D.OR.X.EQ.Y.OR.A+B.NE.C*D.AND. 6.0.LE.A-B)
+GOTO3
   X=X ** 5+A
   X=- 3.0
   X=+ 3.0
   IF(.NOT. C .OR..NOT.D)GOTO7
   Z=(3.0,2.0)+(2.0, 4.0)
STOP
END

Figure 1.a: Program before processing by POLISH.
C EXAMPLE 1.
C THIS IS A NONSENSE PROGRAM ILLUSTRATING SOME FEATURES OF EDITING
C DONE BY POLISH.
COMPLEX Z
EQUIVALENCE (C,D), (E,F), (G,H)
DATA C, A /3.0,2.0/
   10 A = 1.0
   20 B = 3.0*C + D*E/F
CD(I,J) = AB(I,J) + EF(I+1)
DO 50 I=1,50
   X = X + Y
   C = C + D
   DO 40 K=1,40
   E = E + 1.0
   DO 30 L=1,100
      M = M + 1
   CONTINUE
   40 CONTINUE
   50 CONTINUE
IF (A.LT.C) GO TO 10
IF (A.LE.D) GO TO 10
IF (A.EQ.C) GO TO 10
IF (A.LT.B .AND. C.LE.D .OR. X.EQ.Y .OR. A+B.NE.C*D .AND. 6.0.LE.A-B) GO TO 10
   X = X**5 + A
   X = -3.0
   X = +3.0
IF (.NOT.C .OR. .NOT.D) GO TO 20
Z = (3.0,2.0) + (2.0,4.0)
STOP
END

Figure 1.b: Program after processing by POLISH.
Example 2.
This is another nonsense program illustrating some features of editing done by POLISH.

X = 1.0
Observe how this block of comments is left adjusted, preserving the relative indentation.

Y = 2.0
This block has a comment line which is too long (cf. KEYS(1)) so break and append procedures are applied. This may not always happen and depends on the KEYS parameter values.

Z = 3.0
This block also has a comment line which is too long, however break and append procedures are not applied in this case because this comment block consists of more than three lines. Note that the default value for KEYS(10) is 3.

Z = 1.0
WRITE(6,10)
10 FORMAT(*NOTE REPOSITIONING OF THIS STATEMENT*)
WRITE(6,20)
20 FORMAT(*NOTE STRING DELIMITER REPLACEMENT*)
WRITE(6,30) X,Y,Z
30 FORMAT(3H X = ,E20.10,3H Y = ,E20.10,3H Z = ,E20.10/*OBSERVE THE SPLIT +WHICH IS DONE HERE*)
STOP
END

---

Figure 2.a: Program before processing by POLISH.
C EXAMPLE 2.
C THIS IS ANOTHER NONSENSE PROGRAM ILLUSTRATING SOME FEATURES OF
C EDITING DONE BY POLISH.
    X = 1.0
C OBSERVE HOW THIS BLOCK OF COMMENTS
C IS LEFT ADJUSTED, PRESERVING THE
C RELATIVE INDENTATION.
    Y = 2.0
C THIS BLOCK HAS A COMMENT LINE WHICH IS TOO LONG (CF.
C KEYS(1)), BREAK AND APPEND PROCEDURES ARE APPLIED. THIS MAY
C NOT ALWAYS HAPPEN AND DEPENDS ON THE KEYS PARAMETER VALUES.
    Z = 3.0
C THIS BLOCK ALSO HAS A COMMENT LINE WHICH IS TOO LONG, HOWEVER BREAK
C AND APPEND PROCEDURES ARE NOT APPLIED IN THIS CASE BECAUSE THIS
C COMMENT BLOCK CONSISTS OF MORE THAN THREE LINES. NOTE THAT THE
C DEFAULT VALUE FOR KEYS(10) IS 3.
    Z = 1.0
    WRITE (6,99999)
    WRITE (6,99998)
    WRITE (6,99997) X, Y, Z
STOP
99999 FORMAT (36HNOTE REPOSITIONING OF THIS STATEMENT)
99998 FORMAT (33HNOTE STRING DELIMITER REPLACEMENT)
99997 FORMAT (3H X=, E20.10, 3H Y=, E20.10, 3H Z=,
* E20.10/37HOBSERVE THE SPLIT WHICH IS DONE HERE)
END

Figure 2.b: Program after processing by POLISH.
C EXAMPLE 3.
C THIS EXAMPLE ILLUSTRATES ALPHABETIZATION OF SUBPROGRAMS.
X=2.0
READ(5,10)X
10 FORMAT(F10.0)
CALL ASUB(X)
STOP
END
SUBROUTINE BSUB(U,V)
V=U+1.
X=Y
RETURN
END
SUBROUTINE ASUB(Y)
Y=4.0
CALL BSUB(Y,R)
RETURN
END

Figure 3.a: Program before processing by POLISH.

C EXAMPLE 3.
C THIS EXAMPLE ILLUSTRATES ALPHABETIZATION OF SUBPROGRAMS.
X = 2.0
READ (5,99999) X
CALL ASUB(X)
STOP
99999 FORMAT (F10.0)
END
SUBROUTINE ASUB(Y)
Y = 4.0
CALL BSUB(Y, R)
RETURN
END
SUBROUTINE BSUB(U, V)
V = U + 1.
X = Y
RETURN
END

MAN  10
MAN  20
MAN  30
MAN  40
MAN  50
MAN  60
MAN  70
MAN  80
ASU  10
ASU  20
ASU  30
ASU  40
ASU  50
BSU  10
BSU  20
BSU  30
BSU  40
BSU  50

Figure 3.b: Program after processing by POLISH.
COMMENT blocks that have already been formatted in a special way. The
editing will be suspended by the appearance of a card with an asterisk
in column 1 and will resume with the second appearance of such a card.
Also, parameter specifications and delimiters can be used to suspend
all editing of COMMENT statements or limit it to indicated areas.

POLISH is a descendant of a program called STYLE written by
Dorothy Lang Wedel [4]. While POLISH contains many of the design goals
of STYLE, it is an entirely new program. POLISH uses Sale's algorithm
to identify Fortran statement types [5]. Most of the production work
on POLISH was done by a group of twelve undergraduate students at the
University of Colorado as a class project, supervised by one of the
authors [LDF]. The coauthors of this report were the principal contribu-
tors in this project.

The general operating requirements and run time data for POLISH on
a CDC 6400 operating under Kronos 2.1 without subprogram alphabetization
are*:

(1) Field length to run--57000 (octal)
(2) Files on secondary storage--8
(3) Processing time--30.3 statements/CPU sec
(4) Job costs--$1.50 (100 statements)

    --$2.50 (200 statements)
    --$3.20 (300 statements)

* This data depends on the environment. It is correct for the systems
environment at the Computing Center of the University of Colorado,
September 14, 1976.
2. EDITING FEATURES

Fixed editing features. There are three editing features, not subject to user control:

(1) Spacing conventions between tokens;
(2) Renumbering of statement labels and removal of unreferenced labels;
(3) Insertion of CONTINUE statements to force termination of every DO-block on a unique CONTINUE statement.

The goal of the spacing conventions is to produce a text pattern which is easy for the reader of FORTRAN text to follow, and at the same time is economical in its use of blanks. An important principal in this connection is that the token sequence <identifier> <operator> <identifier> is grouped closer together for operations executed first (in the execution hierarchy) than for others, except inside parentheses. A list of spacing conventions used in POLISH appears below:

(1) No space before or after operators within parentheses except before and after logical operators .AND. and .OR.;
(2) No space before or after the operators *, **, /, .LT., .LE., .EQ., .NE., .GE., .GT.;
(3) Single space before and after the operators +, -, =, .AND., .OR. except in DO statement and within parentheses (cf.1);
(4) No space after the unary operators +, -, .NOT.;
(5) One space to the right of the comma, except in complex constants, subscripts, within solidus delimiters of a DATA statement, and within parenthesis delimiters of an EQUVALENCE statement;
(6) No space to the right of a left parenthesis or a solidus as a left delimiter;
(7) No space to the left of a right parenthesis or a solidus as a right delimiter;
(8) No space between nested parentheses;
(9) No space within any subscript;
(10) Normal English spacing is used elsewhere, except no space to the right of a solidus used as a field separator in a FORMAT statement.

Certain commonly used, but non-standard, FORTRAN constructions are accepted by POLISH. These are:

(a) Hollerith strings in FORMAT statements delimited by a special symbol (e.g. *SAMPLE* instead of 6HSAMPLE);
(b) FORTRAN II READ, PRINT, and PUNCH statements (e.g. READ 99, X instead of READ (5,99) X);
(c) DATA statement for array elements without explicit naming of the elements (e.g. DATA A/5.0, 3.0, 1.0/ instead of DATA A(1), A(2), A(3) / 5.0, 3.0, 1.0 /)

POLISH will replace the Hollerith string with delimiters in a FORMAT statement by the standard form; thus *SAMPLE* would be replaced by 6HSAMPLE.

Labels are renumbered so that they increase by a fixed amount in order of their appearance on statements. The fixed increment can be set by the user (cf. KEYS(11) description). While unreferenced labels do not violate the ANSI standard, they are a source of confusion and may not be accepted by a compiler. POLISH removes unreferenced labels.
User controlled editing features. There are a number of editing features under control of the user. Each of these features is designated by a parameter value passed to POLISH in a thirteen element array in labelled COMMON; viz.

COMMON/PARMS/KEYS(13)

A description of each element of the KEYS array follows. Within POLISH each element of KEYS is initialized by a DATA statement; these initial values thus serve as default values. It should be noted that once these values are changed by a user in the course of execution of the program, the "default" value is permanently overwritten for that execution run.
<table>
<thead>
<tr>
<th>KEYS</th>
<th>DESCRIPTION</th>
<th>DEFAULT</th>
<th>RANGE&lt;sup&gt;+&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Rightmost card column in which FORTRAN code can appear; subsequent code is put on continuation cards with an * in card column 6.</td>
<td>64</td>
<td>[32,72]</td>
</tr>
<tr>
<td>2.</td>
<td>Left margin for comment statements.</td>
<td>3</td>
<td>[3, KEYS(1)-29]</td>
</tr>
<tr>
<td>3.</td>
<td>Card column for right justification of statement labels.</td>
<td>5</td>
<td>[1,5]</td>
</tr>
<tr>
<td>4.</td>
<td>Increment for statement labels.</td>
<td>10</td>
<td>&gt;0</td>
</tr>
<tr>
<td>5.</td>
<td>Number of columns indented on continuation cards.</td>
<td>10</td>
<td>≥0</td>
</tr>
<tr>
<td>6.</td>
<td>Increment for sequencing in card columns 73-80. If 0 then no sequencing. Sequencing consists of the first three letters of the name of the subprogram and a sequence number. POLISH assigns the name MANPRG to a main program and BLKDAT to all BLOCK DATA subprograms.</td>
<td>10</td>
<td>≥0</td>
</tr>
<tr>
<td>7.</td>
<td>Number of columns indented in DO loops. POLISH indents all code within DO loops according to this specification. Note: Indentation of DO loops or continuation cards cannot exceed 20 spaces to the left of the right margin. If this occurs, then indentation is set to KEYS(1)-20.</td>
<td>2</td>
<td>≥0</td>
</tr>
<tr>
<td>8.</td>
<td>Blank comment cards removed. 0---- they are removed, anything else they are not.</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td>9.</td>
<td>Authorization to break and append comments. If 0---- as permitted by KEYS(10) and + delimiters, anything else --- break and append comments to conform to KEYS(1) and KEYS(2).</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td>10.</td>
<td>Maximum size of non-delimited comment block in which break and append procedures may be used when KEYS(9)=0.</td>
<td>3</td>
<td>≥0</td>
</tr>
<tr>
<td>11.</td>
<td>Optional Hollerith delimiting character Effect is to replace say, <em>SAMPLE</em> by 6HSAMPLE in FORMAT statements only.</td>
<td>*</td>
<td>N/A</td>
</tr>
<tr>
<td>12.</td>
<td>Edit comments. 0--- yes, anything else--- no. If no then KEYS(2), KEYS(8), KEYS(9) and KEYS(10) are ignored, except to check if KEYS(2) is in range. If it is not in range an error message is printed and execution is terminated.</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td>13.</td>
<td>Alphabetize subprograms. 0--- yes, anything else--- no.</td>
<td>0</td>
<td>N/A</td>
</tr>
</tbody>
</table>

<sup>+[a,b] means all integers in this interval including a and b.</sup>
3. INPUT AND OUTPUT

Seven files are used by POLISH for input, output, and intermediate storage. Unit numbers for these files may be passed to POLISH through labeled COMMON, otherwise default values are used. The specification of the unit numbers in labeled COMMON, the file function, and the default value is given below.

<table>
<thead>
<tr>
<th>File</th>
<th>Function</th>
<th>Default Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>K1</td>
<td>Input file. Contains program to be edited. End of file flag is card with $ in column 1. This file must be positioned before POLISH is called, POLISH does not rewind this file. File read in 80A1 format.</td>
<td>1</td>
</tr>
<tr>
<td>K2</td>
<td>Scratch file for COMMENT statements. File read and written in 80A1 format.</td>
<td>2</td>
</tr>
<tr>
<td>K3</td>
<td>Scratch file. File read and written in unformatted binary.</td>
<td>3</td>
</tr>
<tr>
<td>K4</td>
<td>Scratch file for FORMAT statements. File read and written in unformatted binary.</td>
<td>4</td>
</tr>
<tr>
<td>K5</td>
<td>Not used in present implementation.</td>
<td>5</td>
</tr>
<tr>
<td>K6</td>
<td>Scratch file for subprogram alphabetization. File read and written in 80A1 format.</td>
<td>6</td>
</tr>
<tr>
<td>K7</td>
<td>Output file. The edited program is written on this file. POLISH does not rewind this file. No carriage control characters are written on this file. File is written in 80A1 format.</td>
<td>7</td>
</tr>
<tr>
<td>K8</td>
<td>Error message file. Error messages generated by POLISH are written on this file. POLISH does not rewind this file. Carriage control characters are written on this file. File is written in A1 format.</td>
<td>8</td>
</tr>
</tbody>
</table>

* Default values are set by BLOCK DATA
FORMAT statements in the program text may be moved to the end of the subprogram block, immediately before the END statement, or left in place. Movement to the end of the subprogram block is achieved by assigning different numbers to files K3 and K4; explanation of file specification appears in the next section. FORMAT statements are left in place if the same number is assigned to files K3 and K4.

FORMAT statements are renumbered starting at 10**KEYS(3)-1, and decreasing by 1. Thus for KEYS(3)=5 the FORMAT statements are numbered 99999, 99998, 99997, ... Unreferenced FORMAT statements are given a blank label and an error message is printed.

A card with an asterisk in column 1 will suspend all editing of the cards which follow except for the generation of sequence numbers; thus columns [1, 72] will appear in the output identically as input. The card with the asterisk in column 1 is not reproduced on the output. Editing is thus suspended until the next appearance of a card with an asterisk in column 1. Suspension of editing can be repeated. This feature is useful for preserving the layout of COMMENT statements.

If editing is suspended as just described on a segment of code containing statement labels, unreferenced labels or duplicate labels may result since labels are ignored while editing is suspended.
4. COMMENT STATEMENT EDITING

COMMENT statements are edited in blocks of 1 to 120 statements. A COMMENT block begins with the first COMMENT after a non-comment statement and is terminated by a non-comment statement or by one of the editing delimiter cards. The 120th COMMENT will terminate a block and the 121st COMMENT start another block. A delimited COMMENT table is processed as a block regardless of length.

COMMENT statement editing is controlled by six KEYS parameters (cf. Section 2, KEYS description) and three sets of delimiters. In decreasing order of precedence the editing controls are:

<table>
<thead>
<tr>
<th>Editing Control</th>
<th>Use</th>
<th>Result of Editing Control Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>* in card column one</td>
<td>Do not edit delimiter</td>
<td>Columns (1,72) of COMMENTS and other statements appearing between * delimiter cards are copied directly with no editing. Sequence numbers are generated and added in columns (73,80) in accordance with KEYS(6).</td>
</tr>
<tr>
<td>KEYS(12)</td>
<td>Indicates COMMENT editing is permitted/not permitted</td>
<td>For a non-zero value, each COMMENT block is treated as if it were delimited by do not edit delimiters. For a zero value COMMENT blocks are edited as specified by other COMMENT editing controls.</td>
</tr>
<tr>
<td>KEYS(1)</td>
<td>Indicates right margin for edited COMMENT statements</td>
<td>Point at which COMMENT is to be broken if required and permitted. Error message is issued if edited COMMENT exceeds KEYS(1) value.</td>
</tr>
<tr>
<td>KEYS(8)</td>
<td>Indicates blank COMMENTS are to be retained/deleted</td>
<td>For a zero value, blank COMMENTS are deleted. For a non-zero value blank COMMENTS are retained.</td>
</tr>
<tr>
<td>++ in card columns one and two</td>
<td>Table delimiter</td>
<td>COMMENTS appearing between ++ delimiter cards are processed as a table in accordance with the data card immediately following the initial ++ delimiter card (cf. Appendix C, COMMENT Table Processing).</td>
</tr>
<tr>
<td>KEYS(2)</td>
<td>Indicates left margin for edited COMMENT statements</td>
<td>Indented COMMENTS are indented from this value.</td>
</tr>
</tbody>
</table>
KEYS(9) Indicates specially formatted COMMENT statements may be/are not present

For a non-zero value specially formatted COMMENTS are not present, COMMENTS may be broken and appended to conform to KEYS(1) and KEYS(2). For a zero value, specially formatted COMMENTS may be present. COMMENT blocks may be broken and appended only if permitted by + delimiters or KEYS(10).

+ in card column one COMMENT edit delimiter

COMMENTS appearing between + delimiter cards may be broken and appended to conform to KEYS(1) and KEYS(2). FORTRAN code, delimited COMMENT tables and do not edit delimited statements may be nested, in the sense of DO loop nesting, between the + delimiter cards.

KEYS(10) Indicates maximum size of non-delimited COMMENT block in which break and append procedures may be used when KEYS(9)=0

COMMENT blocks of length ≤ KEYS(10) may be broken and appended to conform to KEYS(1) and KEYS(2). Primary purpose of this is to permit editing of small COMMENT blocks, which presumably do not have a tabular structure which needs to be preserved.

For non-delimited and + delimited COMMENT blocks, editing is divided into minimum, advanced and blank COMMENT editing phases. Minimum editing consists of shifting COMMENT blocks en bloc to conform to KEYS(2). If the COMMENT block does not then conform to KEYS(1), an attempt is made to achieve conformance by reducing relative indentations. If the attempt is successful, indentations are changed and there is no need for advanced editing. If the attempt is not successful the COMMENT lines are left in the shifted position as required by KEYS(2), indentations are not changed, and if advanced editing is not permitted an error message is issued; in the latter instance, the right end of COMMENT statements may be lost. Advanced editing consists of breaking and appending COMMENT statements so that the block will conform to KEYS(1) (cf. Appendix B). Blank COMMENTS are edited in accordance with KEYS(8).
5. CALLING PROCEDURE

If default values for KEYS and file numbers are used, then the main program is simply

CALL POLISH
STOP
END

If FORMAT statements are to be left in place but otherwise default options are used, then the main program is

COMMON/INOUT/K1, K2, K3, K4, K5, K6, K7, K8
K3=K4
CALL POLISH
STOP
END

If no editing of COMMENT statements is desired, FORMAT statements are to be left in place and subprograms are not to be alphabetized but otherwise default options are used, then the main program is

COMMON/PARMS/KEYS(13)
COMMON/INOUT/K1, K2, K3, K4, K5, K6, K7, K8
K3=K4
KEYS(12)=1
KEYS(13)=1
CALL POLISH
STOP
END
POLISH can be used to edit more than one program, with different editing features selected. For example:

```
COMMON/PARMS/KEYS(13)
COMMON/INOUT/K1, K2, K3, K4, K5, K6, K7, K8
CALL POLISH
K3=K4
CALL POLISH
KEYS(12)=1
CALL POLISH
STOP
END
```

This main program will cause the editing of three programs on the input file: The first program will be edited with all default options; the second program will have FORMAT statements left in place; the third program will have FORMAT statements left in place and COMMENT statements will not be edited.

**Input file.** The input file must contain a card with a $ in column 1 at the end of the program to be edited by a call to POLISH. In the third example above three programs are processed by POLISH and each would have to be followed by a card with a $ in column 1.
6. ERROR MESSAGES

Two types of error messages are generated by POLISH.

**TYPE 1 FORMAT**

ERROR NN WAS DETECTED IN XXX.
ERRSUB WAS CALLED BY YYY.

Where:

NN is the error number.
XXX is user Routine name
YYY is name of POLISH routine that detected the error.

**TYPE 2 FORMAT**

****** ERROR TYPE NN HAS BEEN DETECTED IN THE ABOVE STATEMENT.
THIS STATEMENT IS IN ROUTINE XXX STATEMENT NUMBER III.

Where:

NN is the error number.

XXX is the name of the users routine.

III is the sequence number of the bad statement.
# ERROR NUMBERS FOR TYPE 1 ERRORS.

<table>
<thead>
<tr>
<th>Routine in which detected</th>
<th>Error number</th>
<th>Meaning</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>LABEL</td>
<td>-1</td>
<td>too many statement labels in one subroutine</td>
<td>length of DEF array must be increased with MAXDEF changed to reflect this change. Length is now 100.</td>
</tr>
<tr>
<td></td>
<td>-2</td>
<td>too many reference labels in one subroutine</td>
<td>length of REF array must be increased with MAXREF changed to reflect this change. Length is now 200.</td>
</tr>
<tr>
<td></td>
<td>-3</td>
<td>DO statements nested too deep</td>
<td>length of DOSTCK array must be increased with MAXSTK changes to reflect this increase. Length is now 10.</td>
</tr>
<tr>
<td></td>
<td>-4</td>
<td>too many DO statements in one subprogram</td>
<td>length of DOARRY array must be increased with MAXRRY changed to reflect this increase. Length is now 50.</td>
</tr>
<tr>
<td>ENLDO</td>
<td>-1</td>
<td>Same as LABEL -2</td>
<td>Should never occur.</td>
</tr>
<tr>
<td></td>
<td>-2</td>
<td>Same as LABEL -1</td>
<td>Should never occur.</td>
</tr>
<tr>
<td>BCKSCH</td>
<td>±1</td>
<td>Program error</td>
<td>Should never occur.</td>
</tr>
<tr>
<td>DEFPRO</td>
<td>-1</td>
<td>Same as LABEL -1</td>
<td>Should never occur.</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>renumbering of FORMAT statements overlaps with renumbering of other statements. Non-fatal.</td>
<td>Adjust KEYS(3) and KEYS(4) and rerun.</td>
</tr>
<tr>
<td>DSORT</td>
<td>-1</td>
<td>Program error</td>
<td>Should never occur.</td>
</tr>
<tr>
<td></td>
<td>-2</td>
<td>Same as LABEL -2</td>
<td>Should never occur.</td>
</tr>
<tr>
<td>KSETA</td>
<td>-1</td>
<td>Program error</td>
<td>Should never occur.</td>
</tr>
<tr>
<td>KSETB</td>
<td>-1</td>
<td>Program error</td>
<td>Should never occur.</td>
</tr>
<tr>
<td>KDECNA</td>
<td>0,-1</td>
<td>Program error</td>
<td>Should never occur.</td>
</tr>
<tr>
<td>KADVBA</td>
<td>-1</td>
<td>Program error</td>
<td>Should never occur.</td>
</tr>
<tr>
<td>KADVBB</td>
<td>-1</td>
<td>Program error, or B array overflow.</td>
<td>Should never occur.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>See note below.</td>
</tr>
<tr>
<td>KDECBA</td>
<td>-1</td>
<td>Program error</td>
<td>Should never occur.</td>
</tr>
<tr>
<td>KADVNA</td>
<td>-1</td>
<td>Program error</td>
<td>Should never occur.</td>
</tr>
</tbody>
</table>

Some of the program errors which normally should never occur, may occur if the input to POLISH is not recognizable by POLISH.
NOTE: POLISH can input up to 19 continuation cards with no errors. This is equivalent to 1326 characters (72 characters from the first card and 66 from each continuation card). During processing POLISH inserts and deletes blanks and renumbers statement labels. If the number of characters ever increases over 1326 then the B array overflows. To correct this the statements should be broken up into two or more statements, or the A and B arrays must be increased with MAX (located in labeled common DEMMAX) also changed to reflect this increase.

All above errors are fatal except as noted.

ERROR NUMBERS FOR TYPE 2 ERRORS.

All errors non-fatal except 25.

<table>
<thead>
<tr>
<th>Error Number</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>non-ANS use of a letter.</td>
</tr>
<tr>
<td>2</td>
<td>non-ANS use of a number.</td>
</tr>
<tr>
<td>3</td>
<td>This is a program error and should never occur.</td>
</tr>
<tr>
<td>4</td>
<td>non-ANS use of a left parenthesis.</td>
</tr>
<tr>
<td>5</td>
<td>non-ANS use of a right parenthesis.</td>
</tr>
<tr>
<td>6</td>
<td>non-ANS use of a comma.</td>
</tr>
<tr>
<td>7</td>
<td>non-ANS use of an equals sign.</td>
</tr>
<tr>
<td>8</td>
<td>non-ANS use of a plus or minus sign or non-ANS use of a character not included in errors 1-7.</td>
</tr>
<tr>
<td>9</td>
<td>non-ANS use of an asterisk or solidus.</td>
</tr>
<tr>
<td>10</td>
<td>non-ANS use of a period.</td>
</tr>
<tr>
<td>11</td>
<td>This is a program error and should never occur.</td>
</tr>
<tr>
<td>12</td>
<td>non-ANS use of a character not included in errors 1-11.</td>
</tr>
<tr>
<td>13</td>
<td>This is a program error and should never occur.</td>
</tr>
<tr>
<td>14-16</td>
<td>Not used.</td>
</tr>
<tr>
<td>17</td>
<td>non-ANS useage may have induced error into statement containing Hollerith string.</td>
</tr>
<tr>
<td>18-19</td>
<td>Program error. Possible non-ANS statement.</td>
</tr>
<tr>
<td>20</td>
<td>non-ANS statement, or second portion of LOGICAL IF statement is non-ANS, or program error.</td>
</tr>
<tr>
<td>21</td>
<td>Second delimiter missing in FORMAT Hollerith string.</td>
</tr>
</tbody>
</table>
22-23 Program error. Possible non-ANS statement.
24 Too many continuation cards on output. Adjust KEYS(1).
25 Error in values of KEYS. FATAL.
26 Non-digit entry in statement label position.
27 Program error. Possible non-ANS statement.
28 More than 19 continuation cards on input.
29 Hollerith string too long to fit on one card image. Adjust KEYS(1) to make card image bigger.
30 Possible improper nesting of DO statements.
31 Missing statement label. Slashes are inserted wherever this label is referenced.
32 Statement label exceeds 5 characters after renumbering, or statement label exceeds number permitted by KEYS(3). Asterisks are inserted.
33 A FORMAT statement is unreferenced.
34 Same as 31 but applies only to FORMAT statements.
35 Same as 32 but applies only to FORMAT statements.
36 Statement is unrecognizable by POLISH.
37 Not used.
38-39 Program error. Possible non-ANS statement.
40 Sequence number greater than 9999. Restart at KEYS(6).
41 Zero-length Hollerith string.
42 This comment block contains at least one COMMENT with length exceeding KEYS(1).
43 Unable to break COMMENT within maximum authorized text length.
44 Unable to break last column of COMMENT table where indicated.
45 Length of table line exceeds KEYS(1).
46 Length of table line exceeds KEYS(1). Line not broken because corrected column number for continuation exceeds column 72.
47 Table processing terminated by a non-comment statement not located between do not process delimiters. If terminating table delimiter now encountered errors including loss of next card image will be introduced possibly without warning.

48 Label not referenced - blanks are inserted.
WARNING:

When an error is detected in one statement, the user is cautioned that POLISH has a tendency to spread this error to other statements without giving error messages. Therefore, when an error is detected, the user should fix the offending statement and re-run the whole sub-program through POLISH. An example:

```
FUNCTION WRONG(A)
C THE FOLLOWING SHOULD BE GO TO
  GTO 10
  RETURN
10  WRONG = 2.0*A
  RETURN
END
```

POLISH will detect an error in the GO TO statement and hence not process it. Therefore, statement label 10 appears to be unreferenced, and POLISH will delete it.

A NOTE ABOUT HOLLERITH STRINGS:

A Hollerith string will not be detected unless the string follows a , ( /* All ANS FORTRAN input will have this condition satisfied. Generally if an error is detected in a statement containing a Hollerith string, the output will not be correct in that the string length is missing and the letter H is shifted to the left.
7. MACHINE INDEPENDENCE

SUBROUTINE TYPE is written to be completely machine independent and is therefore very inefficient requiring an average of thirteen comparisons to classify a character as a letter, thirty-one comparisons to classify as a digit, and forty-one to classify as a blank or special character. A listing of SUBROUTINE TYPE is provided in Appendix A to assist users in preparing an efficient machine dependent version of this intensively used routine which classifies each character in a non-comment statement.

Alphanumeric character storage is one character per word for data and program constants with the following exceptions:

a. End of statement mark, 3HEOS
b. Statement label flag, 2HFG
c. Hollerith string flag, 2HH*
d. Error messages contained in DATA statements in the various routines and printed by SUBROUTINE ERRSUB have 4Hxxxx Hollerith specifications.

The entire program is input/output bound with four mass storage accesses per statement and two additional mass storage accesses for each COMMENT or FORMAT statement if they are edited in any way. Mass storage accesses are nearly continuous during subprogram alphabetization. Installations where mass storage access are relatively expensive may desire to use machine dependent input/output buffering routines.
REFERENCES:

[1] TIDY, University of Colorado Computer Center program Q-TIDY-237. (Adapted from program by Harry M. Murphy, Air Force Weapons Laboratory, Kirtland Air Force Base, New Mexico 1966).


SUBROUTINE TYPE(IA1, MM, JTYPE)

C IDENTIFICATION
C PROGRAM NAME TYPE
C SOURCE LANGUAGE ANSI FORTRAN
C DATE 12/01/73
C PROGRAMMER M. BOLKE
C READER J.S. DORRENBACHER
C UPDATE 1 OF JULY 1974

C ABSTRACT
C THIS PROGRAM CLASSIFIES A CHARACTER OR CHARACTER STRING AND ASSIGNS A TYPE CODE TO IT.
C METHOD OF USE
C THIS SUBROUTINE IS CALLED BY
C CALL TYPE(IA1, MM, JTYPE)
C WHERE
C IA1 -- VARIABLE CONTAINING THE CHARACTER OR CHARCTERS TO BE CLASSIFIED. (INPUT)
C MM -- FLAG INDICATING HOW IA1 SHOULD BE CLASSIFIED.  MM=0 OR -1. (INPUT)
C JTYPE-- INTEGER REPRESENTING THE TYPE CODE. (OUTPUT)
C JTYPE CHARACTER(S) IN IA1
C 1 LETTER
C 2 DIGIT
C 3 END OF STATEMENT MARK (3EOS)
C 4 LEFT PARENTHESIS
C 5 RIGHT PARENTHESIS
C 6 COMMA
C 7 EQUAL SIGN
C 8 IF MM=0, THEN JTYPE=8 IF IA1 IS NOT ONE OF THE ABOVE.
C 9 IF MM=-1, THEN JTYPE=8 IF IA1 IS A PLUS OR MINUS SIGN.
C 10 PERIOD
C 11 BLANK
C 12 OTHER
C 13 HOLLERITH FLAG (2HH*)
C THE CHARACTER TO BE CLASSIFIED IS REPRESENTED IN IA1 IN A1
C FORMAT EXCEPT FOR THE END OF STATEMENT MARK AND THE HOLLERITH
C EXAMPLE
C IF IA=1H5 AND MM=-1 THEN CALL TYPE(IA,MM,JTYPE) WOULD RETURN
C JTYPE=2.
C IF IA=1H$ AND MM=0 THEN JTYPE=8 UPON RETURN.
C PROGRAM CHARACTERISTICS
C IA -- IA1 REPRESENTED AS A LOCAL VARIABLE
C KLTTTR -- ARRAY CONTAINING ALPHABET IN NORMAL ORDER. EACH ELEMENT CONTAINS ONE LETTER IN 1H FORMAT
C KDIG -- ARRAY CONTAINING DIGITS FROM 0 TO 9. EACH ELEMENT CONTAINS ONE DIGIT IN 1H FORMAT.
C KSPEC -- ARRAY CONTAINING SPECIAL CHARACTERS OR SYMBOLS IN 1H FORMAT. ORDER OF CHARACTERS IS +, -, *, /, (, ), $, +, BLANK, COMMA, PERIOD, END OF STATEMENT MARK.
C IHOLFG -- HOLLERITH FLAG
C THERE ARE NO MACHINE DEPENDENT VARIABLES AND NO MACHINE
C Dependent code in this subroutine.
C More efficient machine dependent code should be available at
C most installations for determining if a given character is a
C letter or a number.
C
COMMON /CHRS/ KLETTR(26), KDIG(10), KSPEC(12)
DATA IHOLFG /2HH*/
IA = IA1
C CHECK FOR LETTER
DO 10 I=1,26
   IF (IA.EQ.KLETTR(I)) GO TO 30
10 CONTINUE
C CHECK FOR NUMBER
DO 20 I=1,10
   IF (IA.EQ.KDIG(I)) GO TO 40
20 CONTINUE
C CHECK FOR END-OF-STATEMENT MARK
   IF (IA.NE.KSPEC(12)) GO TO 50
   JTYPE = 3
   RETURN
C CHARACTER IS A LETTER
30 JTYPE = 1
   RETURN
C CHARACTER IS A NUMBER
40 JTYPE = 2
   RETURN
C CHECK FOR LEFT PARENTHESIS
50 IF (IA.NE.KSPEC(5)) GO TO 60
   JTYPE = 4
   RETURN
C CHECK FOR RIGHT PARENTHESIS
60 IF (IA.NE.KSPEC(6)) GO TO 70
   JTYPE = 5
   RETURN
C CHECK FOR COMMA
70 IF (IA.NE.KSPEC(10)) GO TO 80
   JTYPE = 6
   RETURN
C CHECK FOR EQUAL SIGN
80 IF (IA.NE.KSPEC(8)) GO TO 90
   JTYPE = 7
   RETURN
C CHECK FOR ABRIDGED TABLE
90 IF (MM.NE.0) GO TO 100
   JTYPE = 8
   RETURN
C CHECK FOR + OR - SIGN
100 IF (IA.NE.KSPEC(1) .AND. IA.NE.KSPEC(2)) GO TO 110
    JTYPE = 8
    RETURN
C CHECK FOR * OR / SIGN
110 IF (IA.NE.KSPEC(3) .AND. IA.NE.KSPEC(4)) GO TO 120
    JTYPE = 9
    RETURN
C CHECK FOR PERIOD
120 IF (IA.NE.KSPEC(11)) GO TO 130
JTYP = 10
RETURN

C CHECK FOR BLANK
130 IF (IA.NE.KSPEC(9)) GO TO 140
    JTYP = 11
RETURN

C CHECK FOR HOLLERITH FLAG.
140 IF (IA.NE.IHOLFG) GO TO 150
    JTYP = 13
RETURN

C CHARACTER IS OF TYPE OTHER
150 JTYP = 12
RETURN
END
Appendix B
BREAK AND APPEND PROCEDURES

When the text of a COMMENT statement extends past KEYS(1) it is broken at the first blank space located at or to the left of KEYS(1) + 1. Decisions concerning the continuation indentation and whether or not to append the next COMMENT line are made by examining as many as necessary, or possible, of the two preceding and two following COMMENT lines. Appending of additional COMMENT lines is terminated when any of the following occur: 1) end of COMMENT block, 2) occurrence of blank COMMENT, 3) next COMMENT has greater indentation than current COMMENT, or 4) less than eight characters of the next COMMENT line could be appended to the current COMMENT line. Terminal condition 3) precludes appending the second line of a subparagraph to the first when that form of indentation is used with subparagraphs.

Hints to users:

1. Liberal use of blank COMMENTS is recommended to improve readability and processing of COMMENTS. The blank COMMENTS are removed by the KEYS(8) default value.

2. Generally it is best to separate consecutive subparagraphs with a blank COMMENT. This precludes incorrectly appending subparagraphs together.

3. A single, one line subparagraph should be followed by a blank COMMENT.

4. Use a long text line for COMMENT text, i.e., keep KEYS(2) small.

5. Blank COMMENT statements or an empty COMMENT block delimited by * or + delimiters may be used to separate COMMENT statements. The
former keeps all the COMMENTS in the same block while the later
separates the COMMENTS into two blocks.

6. Break and append procedures are not recommended for a COMMENT
block that contains an outline format.

The following example contains most of the problem cases and shows
the use and misuse of delimiters, blank COMMENTS, and text line length.

Comment block ready to edit.

EXAMPLE COMMENT BLOCK

THE BLANK COMMENT KEEPS A LONG TITLE LINE FROM BEING
CONSIDERED AS THE INDENTED FIRST LINE OF THE FOLLOWING
PARAGRAPH. THE BLANK COMMENT FOLLOWING THE SUBPARAGRAPH
SERVES A SIMILAR PURPOSE.
A. AN ERROR COULD RESULT IF THIS SUBPARAGRAPH BROKEN.

A DIFFERENT PROBLEM IS ILLUSTRATED BY THESE SUBPARAGRAPHS.
A. AN ERROR COULD RESULT IF THIS SUBPARAGRAPH BROKEN.

B. THE BLANK COMMENT SAVES THE SUBPARAGRAPH FORMAT
WHERE THE DELIMITED EMPTY COMMENT BLOCK DOES NOT AS
SHOWN HERE.

C. THE SUBPARAGRAPH FORMAT WILL BE LOST.

HERE THE USE OF A * OR + EMPTY COMMENT BLOCK
WOULD HAVE THE SAME RESULT IF KEYS(9).NE.0 BUT
WOULD DEPEND ON KEYS(10) IF KEYS(9).EQ.0.

Let KEYS(1)=50, KEYS(2)=10, and KEYS(9)=1. Assume other KEYS are
at their default values.
Results of editing are:

EXAMPLE COMMENT BLOCK
THE BLANK COMMENT KEEPS A LONG TITLE LINE
FROM BEING CONSIDERED AS THE INDENTED
FIRST LINE OF THE FOLLOWING PARAGRAPH.
THE BLANK COMMENT FOLLOWING THE
SUBPARAGRAPH SERVES A SIMILAR PURPOSE.
   A. AN ERROR COULD RESULT IF THIS
      SUBPARAGRAPH BROKEN.
A DIFFERENT PROBLEM IS ILLUSTRATED BY
THESE SUBPARAGRAPHS.
   A. AN ERROR COULD RESULT IF THIS
      SUBPARAGRAPH BROKEN.
   B. THE BLANK COMMENT SAVES THE
      SUBPARAGRAPH FORMAT WHERE THE
      DELIMITED EMPTY COMMENT BLOCK DOES
      NOT AS SHOWN HERE.
C. THE SUBPARAGRAPH FORMAT WILL BE LOST.
HERE THE USE OF A * OR + EMPTY COMMENT
BLOCK WOULD HAVE THE SAME RESULT IF
KEYS(9).NE.0 BUT WOULD DEPEND ON KEYS(10)
IF KEYS(9).EQ.0.
Appendix C
COMMENT TABLE EDITING

COMMENT table editing includes the deletion of indicated card columns, the insertion of a space following an indicated card column, and the breaking and continuation with possible appending of the last table column of the respaced COMMENT table if it exceeds KEYS(1). The broken last column is continued at the indicated card column corrected for revised table spacing. If the following COMMENT table line starts at the indicated continuation indentation it is available to be appended to the broken last column. KEYS(8) is used to edit blank COMMENTS.

COMMENT tables are delimited by the ++ delimiter cards. Immediately following the initial ++ delimiter card must be a data card containing processing information for the COMMENT table. The data card has the following format: columns 1-72 contain +, -, or blank; columns 73-78 are blank; columns 79-80 are blank or contain a right justified card column number corresponding to the card column, uncorrected for added spaces or deleted characters, in which the continuation of the last table column is to be started. When columns 79-80 are blank the last table column is not broken. The last table column is broken at the first blank space located at or to the left of KEYS(1) + 1. The meaning of the characters in columns 1-72 are: -
- delete card column from table,
+ insert space after card column is copied,
blank copy card column as is.

Tables of COMMENTS may contain blocks of code and/or COMMENTS between do not edit delimiters. A non-comment statement not located
between do not edit delimiters will terminate table editing with an error message. If a terminal ++ delimiter card is then encountered, the next card is lost and other errors may be introduced without warning or error messages. The other errors should be restricted to COMMENT statements and COMMENT tables. An error message is provided if a table line exceeds KEYS(1) after editing. KEYS(2) is not used in COMMENT table editing.
Example using table processing to insert and delete blanks.

Input table:

++
+++++ +++  ------
C TIME SPEED DISTANCE
C 10.3 54.7
C 15.2 52.3
++
$

Edited table:

C TIME SPEED DISTANCE
C 10.3 54.7 107.1
C 15.2 52.3 153.9
Example using multiple table editing within one table to delete non-blank character, to align table sections, and to break, continue and append the last column of the table. KEYS(1)=45.

Input table:

++

C  IPTR--POINTER TO A-ARRAY.
C  JPTR--POINTER TO B-ARRAY.
++
++

C  A - ARRAY CONTAINING INPUT DATA TO BE PROCESSED.
C  ARRAY IS IN A1 FORMAT.
C  B - ARRAY CONTAINING PROCESSED DATA TO BE OUTPUT.
C  ARRAY IS IN A1 FORMAT.
++
$ 

Edited table:

C  IPTR--POINTER TO A-ARRAY.
C  JPTR--POINTER TO B-ARRAY.
C  A - ARRAY CONTAINING INPUT DATA TO BE PROCESSED.
C  ARRAY IS IN A1 FORMAT.
C  B - ARRAY CONTAINING PROCESSED DATA TO BE OUTPUT.
C  ARRAY IS IN A1 FORMAT.
Appendix D
SECOND LEVEL DEFINITION

The second level definition restrictions on array subscripts of ANSI FORTRAN Standard X3.9-1966, Sections 10.2.8 and B8.1 have not been observed in subprograms COPY, DRICON, DRIVER, LABEL, and LBOCSC. To date, no user has reported any problems in this area. The following changes, line replacements except as noted, will cause second level definition restrictions to be observed:

SUBROUTINE COPY
   OUTBUF(I) = A(IP)

SUBROUTINE DRICON
   IF (KSETB(JP,6).NE.0) GO TO 330
   IF (A(IP).NE.KLITTR(6)) GO TO 190
   190 IF (A(IP).NE.KLITTR(18)) GO TO 200
   200 IF (A(IP).NE.KLITTR(4)) GO TO 210
   B(JP) = A(IP)
   IF (KADVBB(JP).NE.0) GO TO 330
   B(JP) = IBLANK
   IF (KADVBB(JP).NE.0) GO TO 330
   IF (KSETB(JP,5).NE.-1) GO TO 330
   B(JP) = LDSTCK
   IF (K.EQ.10) B(JP) = B(JP) - 1

SUBROUTINE DRIVER
   B(J) = A(IP)
   B(JP) = IBLANK
   60 CALL TYPE(A(IP), MM, JTYP)
   A(IP) = INST
   90 B(JP) = IBLANK
   100 B(JP) = A(IP)
   120 B(JP) = A(IP)
   B(JP) = IBLANK
170 B(JP) = A(IP)
    IF (KSETB(JP,1).NE.-1) GO TO 470
    B(JP) = JJ
200 B(JP) = IBLANK
    B(JP) = A(IP)
240 NPOINT = A(IP)
    CALL TYPE(A(IP), MM, JTYP)
    IF (A(IP).EQ.KLTTR(1) .OR. A(IP).EQ.KLTTR(15))
* GO TO 260
250 B(JP) = NPOIN
    B(JP) = IBLANK
    B(JP) = NPOINT
    B(JP) = A(IP)
    B(JP) = IBLANK
320 B(JP) = IBLANK
    B(JP) = A(IP)
    NAMES(LL) = A(IP)
    IF (A(IP).EQ.KSPEC(5)) GO TO 60
380 IF (A(IP).NE.KSPEC(4)) GO TO 100
    NCHAR = A(IP)
410 B(JP) = A(IP)
    B(JP) = A(IP)
    B(JP) = KSPEC(9)
    B(JP) = A(IP)
    IF (A(IP).NE.IBLANK) GO TO 60
Replace DRI 1030, 1430, 1460, 1610, 1630, 2070, 2090, 2330, 2370, 2390,
2490, 2510, 2800, 2930, 3320, 3350, 3380, and 3440 by
    IF (KADVBB(JP).NE.0) GO TO 470
SUBROUTINE LABEL

* IP, JP

20 IF (KADVBA(IPNTRA).GT.0) GO TO 210
   IF (KADVBB(IPNTRB).GT.0) GO TO 210
   IF (KADVNA(IPNTRA).LT.0) GO TO 210
60 IF (KDECBA(IPNTRA).NE.0) GO TO 210
   IF (KSETB(IPNTRB,3).GT.0) GO TO 210

Insert after LAB 1060
IPNTRA = IP
IPNTRB = JP

FUNCTION LBOCSC

* IP, JP

   IF (KSETA(IPNTRA,ISTART).NE.0) GO TO 50
   IF (KDECBA(IPNTRA).NE.0) GO TO 50

LAB  890
LAB 1210
LAB 1250
LAB 1310
LAB 1360
LAB 1410
LAB  130
LAB  190
LAB  330