Survey of Coded Character Representation

R. W. BEMER, *IBM Corp., White Plains, N. Y.*

Technical Committee 97 of the International Standards Organization (ISO) is concerned with standards in data processing. The American Standards Association holds the secretariat of this committee. Sectional Committee X3 of A.S.A. is responsible for national data processing standards in the U. S.

One of the most important areas of standardization (and one of the most pressing) is that of the logical representation of the character sets. These representations may be by punched holes in paper tape, pulses on a communications line, bits stored in memory, marks on paper, etc. The existing standards work in this area has been done by MEE 149 (now DPE) of the British Standards Institution and by TR 24.4 (now TR 27.6) of the Electronic Industries Association. This particular problem is now under the cognizance of A.S.A. Sub-committee X3.2 on character sets and data format.

The chart (pp. 640–1) is presented as staff work for the deliberations of X3.2. It is the most complete information we have been able to assemble to date, but obviously there may be errors and omissions. The primary aims in publishing this chart are:

(1) To indicate to the information processing industry why standardization is vital in this area.

(2) To request further information from the various experts who possess it.

The chart is presented in 64-character modules (sufficient for a 6-bit set) and the positions are given designations in the octal number system from 00 to 77. For the benefit of anyone not familiar with this notation, a conversion table is given:

Octal	Binary
0	000
1	001
2	010
3	011
4	100
5	101
6	110
7	111

"1" represents an ON or a punched hole condition. "0" represents OFF or an unpunched condition. Obviously, the conversion between physical and logical representation may be made in two ways (for example, depending upon which edge of the paper tape is on your right hand).

Five-track paper tape is shown as though one of either the letter or figure shifts had a sixth track. In most cases this is theoretical. The letter shift-figure shift relationship is indicated on the chart together with the tape orientation. If a "3" appears in the column marked units, then the "3" hole side is octal 01. If a "2" appears in the units column the reverse is true. In most British tapes the letter shift is shown in the high position because the collating sequence adopted by the British for six-track codes puts the digits low to the alphabet. The converse is mostly true in the U. S.

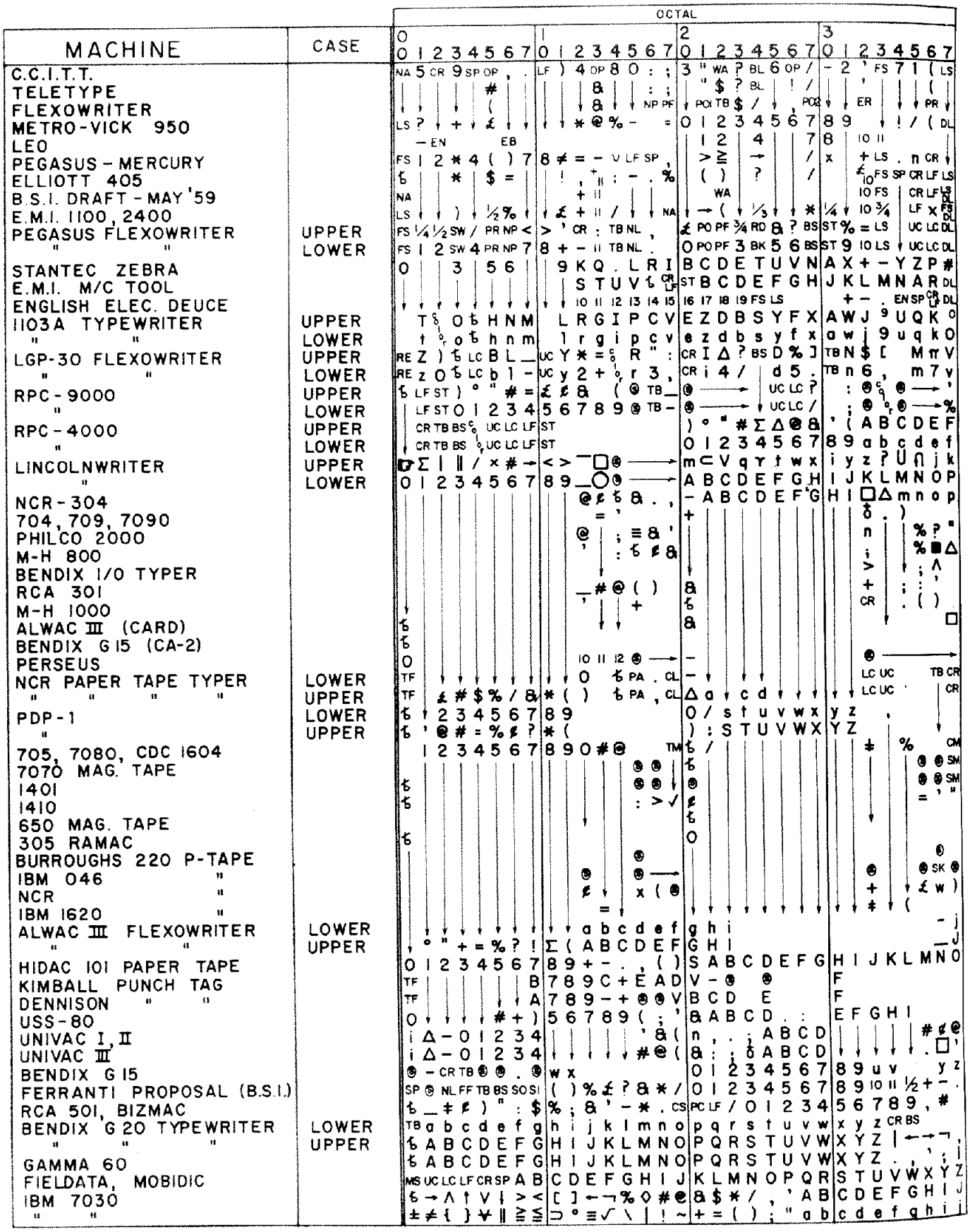
NOTE: The 704 and 705 codes, for example, apparently violate this statement since the digits have lower octal representations than the alphabet. *However*, ordering of files is controlled by a collating sequence in which the digits are higher than the alphabet. This is accomplished either by a comparison matrix in hardware or by programmed replacement of the keys in the records to be ordered.

Some seven element codes are shown on two lines. In most cases these are accomplished by an upper and lower case shift on the input key board. These cases are indicated. The 7030 actually uses the seven (and eight) bit set internally.

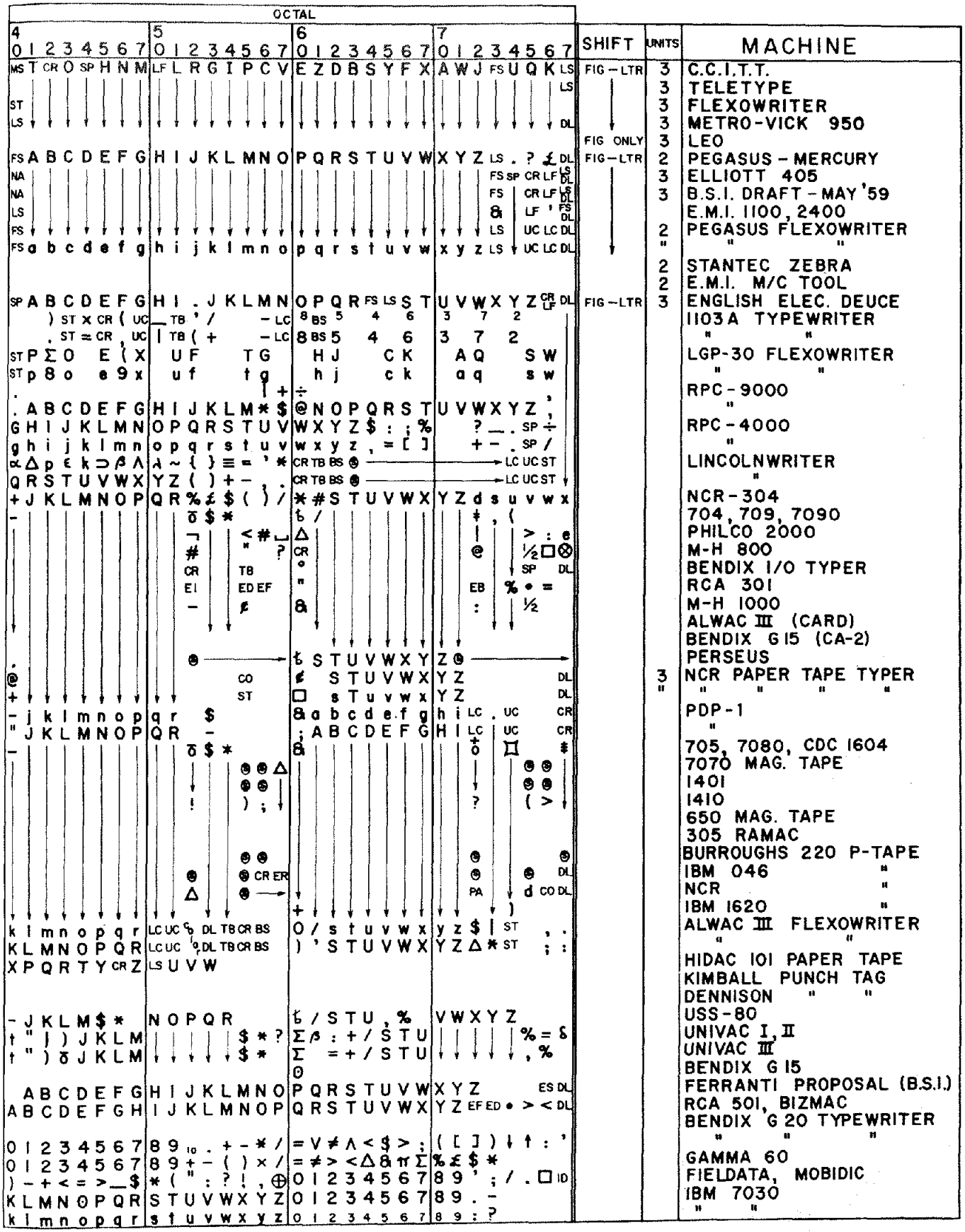
The elements of the code sets may possess either informational or control characteristics. In my personal opinion they should not possess both. Informational characters are shown by their single graphics. Control and functional characters are coded with two-letter mnemonics according to the following table, except for blank (which is shown as a lower case *b* with a slash through the stem) and special (Ⓢ).

BK—black
BL—bell
BS—backspace
CL—clear
CM—card mark
CO—compute
CR—carriage return
CS—carriage shift
DL—delete (erase)
EB—end block
ED—end data
EF—end file
EI—end information
EN—end number
ER—error
ES—escape
FF—form feed (paper throw)
FS—figure shift
ID—idle
LC—lower case
LF—line feed
LS—letter shift
MS—master space
NA—no action
NL—new line (CR + LF)
NP—non-print
OP—optional
PA—put away
PC—page change
PF—punch-off
PO—punch-on
PR—print restore
RD—red
RE—read
SI—shift in
SK—skip
SM—segment mark
SO—shift out
SP—space
ST—stop
TB—tabulate
TF—tape feed
TM—tape mark
UC—upper case
WA—who are you?

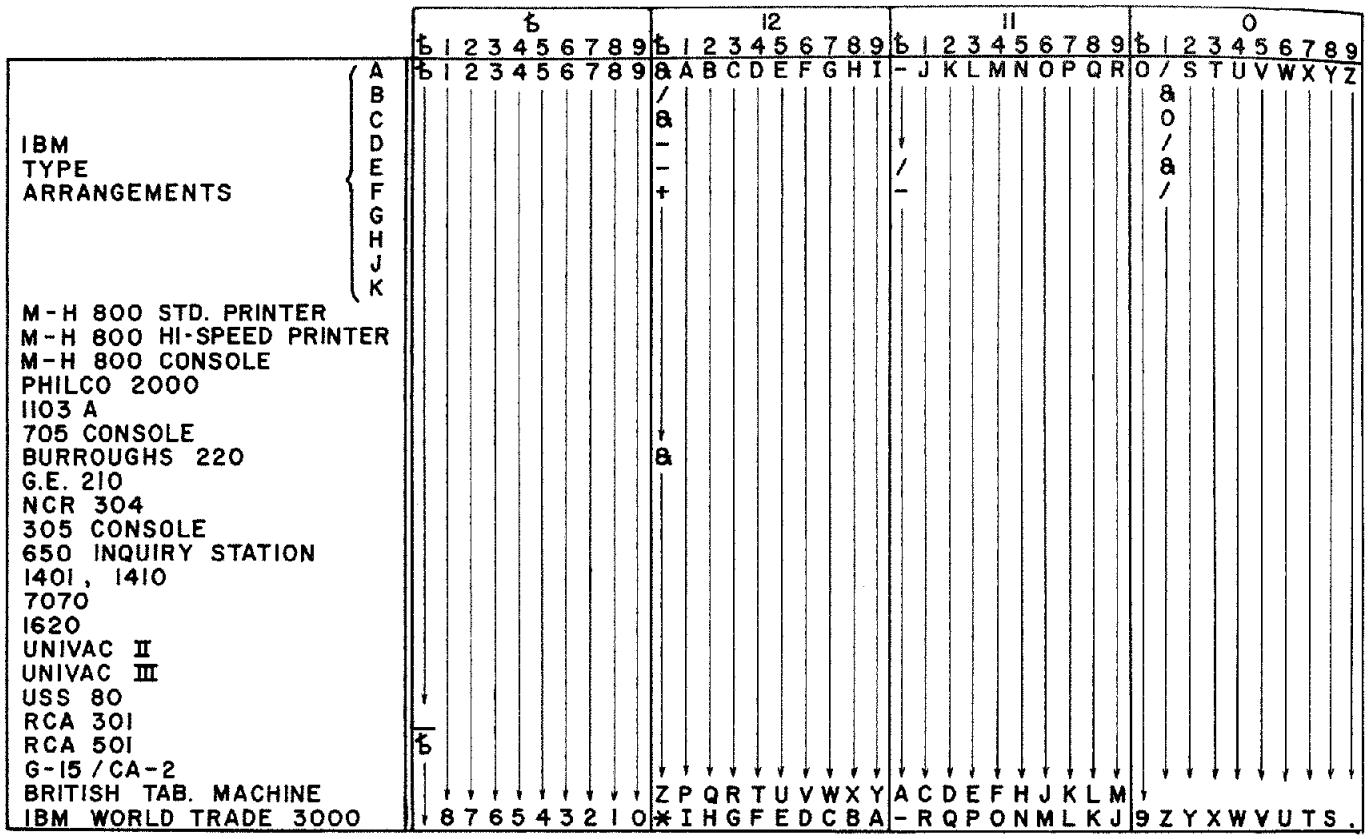
Survey of Coded Character Representation



R.W.BEMER, F.A.WILLIAMS REV. 9 OCT 1960


OCTAL
SHIFT
UNITS
MACHINE
C.C.I.T.T.
TELETYPE
FLEXOWRITER
METRO-VICK 950
LEO
PEGASUS - MERCURY
ELLIOTT 405
B.S.I. DRAFT - MAY '59
E.M.I. 1100, 2400
PEGASUS FLEXOWRITER
STANTEC ZEBRA
E.M.I. M/C TOOL
ENGLISH ELEC. DEUCE
1103A TYPEWRITER
LGP-30 FLEXOWRITER
RPC - 9000
RPC - 4000
LINCOLNWRITER
NCR - 304
704, 709, 7090
PHILCO 2000
M-H 800
BENDIX I/O TYPER
RCA 301
M-H 1000
ALWAC III (CARD)
BENDIX G 15 (CA-2)
PERSEUS
NCR PAPER TAPE TYPER
PDP - 1
705, 7080, CDC 1604
7070 MAG. TAPE
1401
1410
650 MAG. TAPE
305 RAMAC
BURROUGHS 220 P-TAPE
IBM 046
NCR
IBM 1620
ALWAC III FLEXOWRITER
HIDAC 101 PAPER TAPE
KIMBALL PUNCH TAG
DENNISON
USS-80
UNIVAC I, II
UNIVAC III
BENDIX G 15
FERRANTI PROPOSAL (B.S.I.)
RCA 501, BIZMAC
BENDIX G 20 TYPEWRITER
GAMMA 60
FIELDATA, MOBIDIC
IBM 7030

Survey of Punched Card Codes



F.A. WILLIAMS, H.J.SMITH 28 OCT 60