                The EBCDIC Codes and Their Mapping to ASCII

Abstract

   The uniquely map the ASCII codes into corresponding EBCDIC codes in a
   consistent manner throughout the ARPA Network, this RFC describes and
   defines the IBM Standard Extended BCD Interchanged Code.

Introduction

   The IBM Corporate Systems Standard, Extended BCD Interchanged Code
   (EBCDIC) defines 8-bit graphic and control codes (See Figure 1). The
   basic EBCDIC code consists of 54 controls (including space) and 88
   graphics. This set is extended to include 10 special graphics and 1
   special control (EO). These special graphics originate from the 7-
   bit hollerith code and include 6 ASCII graphics. The EBCDIC code is
   further extended to include the publishing and printing graphics
   option which specifics 52 graphics. Of these graphics, 32 appear on
   the IBM TN print chain. Four of these graphics are duals with
   graphics not on the TN print chain, and one graphic (degree) is dual
   with a graphic in the special graphics set of the basic code (tilde).

   It is desirable to uniquely map the ASCII codes into corresponding
   EBCDIC codes in a consistent manner throughout the ARPA network.

   For each of the 34 ASCII controls (including space and delete) there
   is a corresponding BDCDIC control (assigning ASCII control DC3 to the
   EBCDIC code X'13'). For 85 of the 94 ASCII graphics, there is a
   corresponding graphic in the basic EBCDIC set. Three different
   correspondences can be made for the other 9 ASCII graphics.

I.   IBM Correspondence

   a) IBM recommends the following ASCII duals with the basic EBCDIC
      graphics.

             ASCII   EBCDIC   Code
             [[cent sign]   X'4A'
             ]   !   X'5A'
             !   l   X'4F'
             [carrot sign]   [upper right corner]   X'5F'

   Note that the EBCDIC graphic for exclamation point (!) is not chosen
   to correspond to the ASCII for exclamation point (!), though this
   would be a sensible choice, and thus another code must be used to
   represent this graphic.

   b) Special EBCDIC graphics would be used to represent the other ASCII
      graphics.

             Graphic   Code
             l   X'6A'
             l   X'79'
             [diagonal slash]   X'A1'
             ~   X'E0'
             [diagonal slash]   XC0'
             {   XD0'
             }

II. Publishing Correspondence

   a) Associate the following special EBCDIC graphics with the
      corresponding ASCII graphics.

             Graphic   Code
             [carrot]   X'71'
             [X'AD'
             ]   X'BD'
             {   X'8B'
             }   X'9B'

   The codes for open bracket and close bracket are chosen since these
   graphics appear on the TN print chain. The codes for left brace and
   right brace are chosen rather than the codes in the special graphics
   set for opening brace and closing brace, respectively, since these
   graphics are similar and also appear on the TN print chain.

III. Graphical Correspondence

   a) Associate the following basic EBCDIC graphics with the indicated
      ASCII graphics because of their graphic similarity.

                  | X'4F'   with |
                                                           |

              [upper right corner]   X'5F'   with ~

   b) Associate the basic EBCDIC graphic for cent with the ASCII graphic
      for reverse slash.

              [cent]   X'4A'   with [diagonal slash]

   This choice is made since the cent graphic is not an ASCII graphic
   and is the only graphic in the basic EBCDIC set which would not
   otherwise be associated with any ACII graphic.

   c) Associate the special EBCDIC graphic grave accent.

          ` X'79'

      with the corresponding ASCII graphic.

   d) Associate the following publishing EBCDIC graphics with the
   corresponding ASCII graphics.

              [carrot]   X'71'
              [X'AD'
              ]   X'BD'
              {   X'8B'
              }   X'9B'

   The codes for open bracket and close bracket are chosen since these
   graphic appear on the TN print chain. The codes for left brace and
   right brace are chosen rather than the codes in the special graphics
   set for opening brace and closing brace, respectively, since these
   graphics are similar and also appear on the TN print chain.

   Standards:

   In order that the mapping from ASCII into EBCDIC and vice versa could
   become standardized, I would appreciate comments on the above from
   each site whose operating system uses EBCDIC as the internal code.

Telnet Codes:

For those sites who may wish to provide our use TELNET services that
communicate using an EBCDIC code, a standard code must be specified.
The codes given in Figure 1 can form the basis for a standard.
Specific codes must also be specified for the TELNET control codes.
The following are suggested:

	Hex Code
sync	38
break	39
NOP	3A
Return to ASCII	FF
No echo	14
Echo	23
Hide input	24

To eliminate using one code for two graphics, I propose that the TN
graphics be associated with their corresponding code. The graphic
tilde (~) might be assigned to the code X'E1' rather than keeping the
dual with the graphic for degree. This would have no effect if the
Graphical Correspondence were chosen for the EBCDIC to ASCII mapping
with the code X'5F' for logical not associated with tilde. The other
graphics of the publishing and printing option (Double Acute,
Inferior Hook, Macron, and Inferior Comma) which are not on the TN
print chain but have the same codes as graphics on the TN print chain
would not be considered to be part of the standard EBCDIC code.

EBCDIC Questionnaire

   1. For ASCII to EBCDIC mapping of the 9 special ASCII graphics do you
      prefer:

      a)   The IBM correspondence   _____
      b)   The Publishing correspondence   _____
      c)   The Graphical correspondence   _____
      d)   Another correspondence (describe) _____

   2. Do you concur with the definition of the standard EBCDIC code,
      including TELNET control codes?

              YES _____   NO _____

      Comments:
      _____
      _____
      _____

   3. Please list for your operating system:

      a)   graphics not included in the complete EBCDIC code.
      b)   Graphics given a different code.
      c)   Controls given one of the graphic codes.
      d)   Controls given one of the control codes but defined to be a
            different control
      e)   All the controls which have meaning with your operating
            system (i.e., for which special action is taken) and state
            the action.

   Reply from: Name   _____
                    Telephone   _____
                    Site   _____
                    Host Computer   _____

   Send to: Joel M. Winett
                    M.I.T. Lincoln Laboratory
                    Room C-151
                    Lexington, Mass. 02173

   Or call: (617) 862-5500 ext. 7474

   Figure 1. [Please view the PDF version of this RFC.]

   Figure 2. [Please view the PDF version of this RFC.]

Hex Code	Category	Control	Name
00	CC	NUL	Null
01	CC	SOH	Start of Heading
02	CC	STX	Start of Text
03	CC	ETX	End of Text
04	DC	PF	Punch off
05	FE	HT	Horizontal Tab
06	GR	LC	Lower Case
07	GR	DEL	Delete
08	GR	GE	Graphic Escape
09	FE	RLF	Reverse Line Feed
0A	CC	SMM	Start of Manual Message
0B	FE	VT	Vertical Tab
0C	FE	FF	Form Feed
0D	FE	CR	Carriage Return
0E	GR	SO	Shift Out
0F	GR	SI	Shift In
10	CC	DLE	Data Line Escape
11	DC	DC1	Device Control 1
12	DC	DC2	Device Control 2
13	DC	TM/DC3	Tape Mark/Device Control 3
14	DC	RES	Restore
15	FE	NL	New Line
16	FE	BS	Backspace
17	DC	IL	Idle
18	GR	CAN	Cancel
19	DC	EM	End of Medium
1A	DC	CC	Cursor Control
1B	CU	CUI	Customer Use 1
1C	IS	IFS	Info. Field Separator
1D	IS	IGS	Info. Group Separator
1E	IS	IRS	Info. Record Separator
1F	IS	IUS	Info. Unit Separator
20	ED	DS	Digit Select
21	ED	SOS	Start of Significance
22	ED	FS	Field Separator
23			(Reserved)
24	DC	BYP	Bypass
25	FE	LF	Line Feed
26	CC	ETB	End of Text Block
27	GR	ESC	Escape
28			(Reserved)
29			(Reserved)
2A	DC	SM	Set Mode
2B	CU	CU2	Customer Use 2
2C			(Reserved)

```
2D          CC            ENQ            Enquiry
2E          CC            ACK            Acknowledge
2F          DC            BEL            Bell

30                                       (Reserved)
31                                       (Reserved)
32          CC            SYN            Synchronous Idle
33                                       (Reversed)
34          DC            PN             Punch On
35          DC            RS             Reader Stop
36          GR            UC             Upper Case
37          CC            EOT            End of Transmission
38                                       (Reserved)
39                                       (Reserved)
3A                                       (Reserved)
3B          CU            CU3            Customer Use 3
3C          DC            DC4            Device Control 4
3D          CC            NAK            Negative Acknowledge
3E                                       (Reserved)
3F          GR            SUB            Substitute
```

Figure 3: EBCDIC Control Functions

   CC   ([illegible] Control). A functional character [illegible] to
           control or facilitate transmission of introducing [illegible]
           communication networks.

   FB   (Format Bisector). A functional character which controls the
           layout of positioning or information in printing or display
           devices.

   IS   (Information Separator). A character which is used to separate
           and qualify information in a logical sense. There is a group of
           four such characters, which are to be used in a hierarchical
           order.

   DC   (Device Control). A functional character used for the control
           of ancillary devices associated with data processing of
           telecommunication systems, more especially switching devices
           "on" and "off".

   ED   (Edit and Mark). A control character used by the
           System/[illegible]...and Mark ([illegible]) instruction for the
           formatting of alphanumeric fields.

   GH   (Graphic Control). A control character indicating that the core
           combinations which follow are to be [illegible] in a particular
           code table, depending upon the particular control character.

   CU   (Customer Use). A character excluded from future assignment by
           IBM. These "protected" codes are intended for use by customer
           systems so that their use will not conflict with a possible
           future IBM use.

                            Figure 4
                  Categories of Control Functions

	Hex Code	Graphic	Name
*	6A	|	Vertical Line
		|	
*	79	[diagonal slash]	Grave Accent
*	A1	~	Tilde
*	C0	{	Opening Brace
	CC	[hook]	Hook
	CE	[fork]	Fork
*	D0	}	Closing Brace
*	E0	/	Reverse Slant
	EC	[chair]	Chair
	FA	|	Long Vertical Line
	FF	EO	Eight Ones

Figure 5: Special EBCDIC Graphics

*ASCII Graphic

```
        Hex Code     Graphic                     Name
  **       A0         –                           Superscript Minus
  *        A1         [degree]                    Degree
  *        B0         [superscript 0]             Superscript Zero
  *        B1         [superscript 1]             Superscript One
  *        B2         [superscript 2]             Superscript Two
  *        B3         [superscript 3]             Superscript Three
           B4         [superscript 4]             Superscript Four
           B5         [superscript 5]             Superscript Five
           B6         [superscript 6]             Superscript Six
           B7         [superscript 7]             Superscript Seven
           B8         [superscript 8]             Superscript Eight
           B9         [superscript 9]             Superscript Nine

           SB         {                           Left Brace
           SC         [equal or less than]        Equal or Less Than
           SD         [superscript (]             Superscript Left Parenthesis
           SE         [superscript +]             Superscript Plus Sign
           SF         +                           Plotting Cross
           9B         }                           Right Brace
           9C         [lozenge]                   Lozenge
           9D         [superscript )]             Superscript Right Parenthesis
           9E         [plus or minus]             Plus or Minus
           9F         [histogram]                 Histogram
           AB         [lower left corner]         Lower Left Corner
           AC         [upper left corner]         Upper Left Corner
           AD         [                           Open Square Bracket
           AE         [= or >]                    Equal or Greater Than
           AF         [bullet]                    Bullet (Plotting Circle)
           EB         [lower right corner]        Lower Right Corner
           EC         [upper right corner]        Upper Right Corner
           ED         ]                           Close Square Bracket
           EE         [not equal]                 Not equal
           EF         ---                         Entended Dash
```

            Figure 6: Publishing and Printing Graphics
                    Also on the TN Print Chain

```
  **       Dual with the special EBCDIC graph c tilde
  *        Dual with another graphic which is not on the TN print chain
```

```
        Hex Code        Graphic                 Name
            70          [Scandinavian accent]   Scandinavian Accent
            71          [carrot]                Circumflex
            72          [diaeresis]             Diaeresis
            73          /                       Diacritical Virgule
            74          '                       Acute Accent
            75          [superior .]            Superior Dot
            76          ,                       Cedilla
            77          [breve]                 Breve
            78          [caron]                 Caron
            8A          [up arrow]              Up Arrow
            9A          [dagger]                Dagger

  *         B0          "                       Double Acute
  *         B1          '                       Inferior Hook
  *         B2          –                       Macron
  *         B3          ,                       Inferior Comma

            CD          '                       Open Quote
            DB          [pound sign]            Pound Sign
            DC          [section sign]          Section Sign
            DD          [paragraph sign]        Paragraph Sign
            ED          '                       Close Quote
```

              Figure 7: Publishing and Printing Graphics
                      not on the TN Print Chain

  *   Dual with another graphic which is on the TN print chain

Name	Graphic	Hex Code	Graphic	Name
Tilde	~	A1	[degree]	Degree
Double Acute	"	B0	[superscript 0]	Superscript Zero
Inferior Hook	'	B1	[superscript 1]	Superscript One
Macron	-	B2	[superscript 2]	Superscript Two
Inferior Comma	,	B3	[superscript 3]	Superscript Three

Figure 8: Graphic Duals

_____

Codes	Graphics	Name
AF--75	.	Bullet--Superior Dot
8B--C0	{	Left Brace--Opening Brace
9B--D0	}	Right Brace--Closing Brace
61--73	/	Slash--Diacritical Virgule
A1--7.0	[degree]	Degree--Scandinavian Accent
4F--FA	|	Logical Or--Long Vertical Mark
6B--76--B3	,	Comma--Cedilla-Inferior Comma
60--B2	-	Dash--Macron

Figure 9: Similar Graphics

_____

Name	Control	Hex Code	Graphic	Name
Group Mark	GM	4F	|	Logical or
Mode Change	MC	5F	[upper right corner]	Logical Not
Plus Zero	PZ	C0	{	Opening Brace
Minus Zero	MZ	D0	}	Closing Brace
Record Mark	RM	E0	\	Reverse Slant
Word Separator	WS	6D	__	Underscore
Segment Mark	SM	6F	?	Question Mark
Substitute Blank	SB	7A	:	Colon
Tape Mark	TM	7F	"	Quotation Marks

Figure 10: Graphic Control Duals

_____

This material has not been reviewed for public release and is
intended only for use with the ARPA network. It should not be quoted
or cited in any publication not related to the ARPA network.