UDC 681.3.042 : 003.62

Ref. No.: ISO / R 646 - 1967 (E)

ISO

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION

ISO RECOMMENDATION

R 646

6 AND 7 BIT CODED CHARACTER SETS FOR INFORMATION PROCESSING INTERCHANGE

1st EDITION
December 1967

For each individual country the only valid standard is the national standard of that country.

Printed in Switzerland

Also issued in French and Russian. Copies to be obtained through the national standards organizations.

BRIEF HISTORY

The ISO Recommendation R 646, *6 and 7 bit coded character sets for information processing interchange*, was drawn up by Technical Committee ISO/TC 97, *Computers and information processing*, the Secretariat of which is held by the United States of America Standards Institute (USASI).

Work on this question by the Technical Committee began in 1962 and led, in 1966, to the adoption of a Draft ISO Recommendation.

In October 1966, this Draft ISO Recommendation (No. 1052) was circulated to all the ISO Member Bodies for enquiry. It was approved, subject to a few modifications of an editorial nature, by the following Member Bodies:

Argentina	Germany	Turkey
Australia	Israel	U.A.R.
Belgium	Italy	United Kingdom
Brazil	Netherlands	U.S.A.
Canada	Sweden	U.S.S.R.
France	Switzerland	

Four Member Bodies opposed the approval of the Draft:

Austria
Japan
Poland
Spain

The Draft ISO Recommendation was then submitted by correspondence to the ISO Council, which decided, in December 1967, to accept it as an ISO RECOMMENDATION.

ISO Recommendation R 646 December 1967

6 AND 7 BIT CODED CHARACTER SETS FOR INFORMATION PROCESSING INTERCHANGE

1. GENERAL

1.1 Scope

1.1.1 This ISO Recommendation describes two character sets for general usage, including letters, figures, punctuation marks, other symbols and controls, with their coded representation.

1.1.2 The requirements for graphics and controls in data handling and programming, in accordance with computer and ancillary equipment characteristics, have been taken into account in determining these character sets.

1.1.3 This ISO Recommendation consists of two tables with notes, a key, explanatory notes and the definitions of functional characteristics related to control characters.

1.1.4 These character sets are primarily intended for the interchange of information among data processing systems and associated equipment. They may be regarded as basic alphabets in an abstract sense.

1.1.5 These character sets are applicable to all Latin alphabets.

1.1.6 One character set contains 6-bit characters and provides 64 coded characters, the other contains 7-bit characters and provides 128 coded characters. The two sets are compatible in that they are logically related. This facilitates the conversion between the two sets.

1.2 Implementation

1.2.1 These character sets may be implemented in various media with logically related codes, as appropriate. For example, this could include punched tapes, punched cards, magnetic tapes and transmission channels, thus permitting interchange of data to take place either indirectly by means of an intermediate recording in a physical medium, or by local electrical connection of various units (such as input and output devices and computers) or by means of data transmission equipment.

1.2.2 Other ISO Recommendations describe the means of recording these character sets in physical media and the means of encoding them for transmission while taking into account the need for error checking *.

* For the implementation of codes
- on punched tape, see ISO Recommendation R ..., *Representation of 6 and 7 bit coded character sets on punched tape* (at present at the stage of draft proposal);
- on punched cards, see ISO Recommendation R ..., *Implementation of ISO 7 bit coded character set on 12 row punched card* (at present at the stage of draft proposal);
- on 7 and 9 track magnetic tape, see ISO Recommendation R ..., *Implementation of the 6 and 7 bit coded character sets on 7 track 1/2 in (12.7 mm) magnetic tape* (at present Draft ISO Recommendation No. 1320) and ISO Recommendation R ..., *Implementation of the 7 bit coded character set on 9 track 1/2 in (12.7 mm) magnetic tape* (at present Draft ISO Recommendation No. 1321).

2.1 6-Bit set table

Bits b_6	b_5	b_4	b_3	b_2	b_1	Row \ Column	0 0 0	0 1 1	1 0 2	1 1 3
0	0	0	0	0	0	0	SP	0	NUL	P
0	0	0	1	0	1	1	F_1 (HT)	1	A	Q
0	0	1	0	1	0	2	F_2 (LF) ①	2	B	R
0	0	1	1	1	1	3	F_3 (VT)	3	C	S
0	1	0	0	2	0	4	F_4 (FF)	4	D	T
0	1	0	1	2	1	5	F_5 (CR) ①	5	E	U
0	1	1	0	3	0	6	SO	6	F	V
0	1	1	1	3	1	7	SI	7	G	W
1	0	0	0	4	0	8	(	8	H	X
1	0	0	1	4	1	9	)	9	I	Y
1	0	1	0	5	0	10	*	: ⑧	J	Z
1	0	1	1	5	1	11	+	; ⑧	K	([) ③
1	1	0	0	6	0	12	,	< ⑨ $ ②	L	(£) ② ③
1	1	0	1	6	1	13	–	= ⑨ %	M	(]) ③
1	1	1	0	7	0	14	.	> ⑨ &	N	ESC
1	1	1	1	7	1	15	/	'	O	DEL

NOTES ABOUT

① The controls CR and LF are intended for printer equipment which requires separate combinations to return the carriage and to feed a line.

For equipment which uses a single control for a combined carriage return and line feed operation,

— in the 6-bit set table, the function F_2 will have the meaning of " New Line " (NL). F_5 will then have the meaning of " Back-space " (BS);

— in the 7-bit set table, the function FE_2 will have the meaning of " New Line " (NL).

These substitutions require agreement between the sender and the recipient of the data.

The use of this function " NL " is not allowed for international transmission on general telecommunication networks (Telex and Telephone networks).

② For international information interchange, $ and £ symbols d not designate the currency of a given country. The use of the symbols combined with other graphic symbols to designate natio al currencies may be the subject of other ISO Recommendation

③ Reserved for National Use. These positions are primarily intend for alphabetic extensions. If they are not required for that pu pose, they may be used for symbols and a recommended choi is shown in parenthesis in some cases.

Some restrictions are placed on the use of these characters on t general telecommunication networks for international tran mission.

④ Positions 5/14, 6/0 and 7/14 of the 7-bit set table are normal provided for the diacritical signs " circumflex ", " grave accent and " overline ". However, these positions may be used for oth

NOTES

2.2 7-Bit set table

b_7	b_6	b_5	b_4	b_3	b_2	b_1	Row \ Column	0	0	0	0	1	1	1	1
								0	0	1	1	0	0	1	1
								0	1	0	1	0	1	0	1
								0	1	2	3	4	5	6	7
			0	0	0	0	0	NUL	(TC_7)DLE	SP	0	(@) ③	P	`④	p
			0	0	0	1	1	(TC_1) SOH	DC_1	!	1	A	Q	a	q
			0	0	1	0	2	(TC_2) STX	DC_2	" ⑥	2	B	R	b	r
			0	0	1	1	3	(TC_3) ETX	DC_3	£ ② ⑦	3	C	S	c	s
			0	1	0	0	4	(TC_4) EOT	DC_4	$ ②	4	D	T	d	t
			0	1	0	1	5	(TC_5) ENQ	(TC_8)NAK	%	5	E	U	e	u
			0	1	1	0	6	(TC_6) ACK	(TC_9)SYN	&	6	F	V	f	v
			0	1	1	1	7	BEL	(TC_{10})ETB	' ⑥	7	G	W	g	w
			1	0	0	0	8	FE_0 (BS)	CAN	(	8	H	X	h	x
			1	0	0	1	9	FE_1 (HT)	EM	)	9	I	Y	i	y
			1	0	1	0	10	FE_2 (LF) ①	SUB	*	: ⑧	J	Z	j	z
			1	0	1	1	11	FE_3 (VT)	ESC	+	; ⑧	K	([) ③	k	③
			1	1	0	0	12	FE_4 (FF)	IS_4 (FS)	,	<	L	③	l	③
			1	1	0	1	13	FE_5 (CR) ①	IS_3 (GS)	–	=	M	(]) ③	m	③
			1	1	1	0	14	SO	IS_2 (RS)	.	>	N	^ ④ ⑥	n	¯ ④ ⑤
			1	1	1	1	15	SI	IS_1 (US)	/	?	O	_	o	DEL

7 BIT SET TABLE

aphical symbols when it is necessary to have 8, 9 or 10 positions r national use.

or international information interchange, position 7/14 of the bit set table is used for the graphical symbol ¯ (overline), the aphical representation of which may vary according to national e to represent ~ (tilde) or another diacritical sign provided that ere is no risk of confusion with another graphical symbol cluded in the table.

he graphics in positions 2/2, 2/7, 5/14 of the 7-bit set table have spectively the significance of " quotation mark ", " apostrophe " d " upwards arrow "; however, these characters take on the gnificance of the diacritical signs " diæresis ", " acute accent " and circumflex accent " when they precede or follow the " back-ace " character.

⑦ For international information interchange position 2/3 of the 7-bit set table has the significance of the symbol £. Within a country where there is no requirement for the symbol £, the symbol # may be used in that position.

⑧ If 10 and 11 as single characters are needed (for example, for Sterling currency subdivision), they should take the place of " colon " (:) and " semi-colon " (;) respectively. These substitutions require agreement between the sender and the recipient of the data.

On the general telecommunication networks, the characters " colon " and " semi-colon " are the only ones authorized for international transmission.

⑨ Either of the two sets of three symbols shown in these positions in the table may be chosen; the interpretation of the corresponding combinations requires agreement between the sender and the recipient of the data.

3. KEY

3.1 Controls

ACK	ACKNOWLEDGE
BEL	BELL
BS	BACKSPACE
CAN	CANCEL
CR	CARRIAGE RETURN
DC	DEVICE CONTROL
DEL	DELETE
DLE	DATA LINK ESCAPE
EM	END OF MEDIUM
ENQ	ENQUIRY
EOT	END OF TRANSMISSION
ESC	ESCAPE
ETB	END OF TRANSMISSION BLOCK
ETX	END OF TEXT
F	FUNCTION
FE	FORMAT EFFECTOR
FF	FORM FEED
FS	FILE SEPARATOR
GS	GROUP SEPARATOR
HT	HORIZONTAL TABULATION
IS	INFORMATION SEPARATOR
LF	LINE FEED
NAK	NEGATIVE ACKNOWLEDGE
NL	NEW LINE
NUL	NULL
RS	RECORD SEPARATOR
SI	SHIFT-IN
SO	SHIFT-OUT
SOH	START OF HEADING
SP	SPACE
STX	START OF TEXT
SUB	SUBSTITUTE
SYN	SYNCHRONOUS IDLE
TC	TRANSMISSION CONTROL
US	UNIT SEPARATOR
VT	VERTICAL TABULATION

3.2 Graphical Symbols

Graphical Representation	Name	Position in the code tables †	
		6-bit	7-bit
(Space)	A normally non-printing graphic character	0/0	2/0
!	Exclamation mark	—	2/1
"	Quotation mark. Diaeresis (note ⑥)	—	2/2
£	Currency symbol £ (note ②)	3/12	2/3
$	Currency symbol $ (note ②)	1/12	2/4
%	Percent	1/13	2/5
&	Ampersand	1/14	2/6
'	Apostrophe, Acute accent (note ⑥)	1/15	2/7
(	Left parenthesis	0/8	2/8
)	Right parenthesis	0/9	2/9
*	Asterisk	0/10	2/10
+	Plus sign	0/11	2/11
,	Comma	0/12	2/12
-	Hyphen, minus sign	0/13	2/13
.	Full stop (period)	0/14	2/14
/	Solidus	0/15	2/15
:	Colon	1/10	3/10
;	Semi-colon	1/11	3/11
<	Less than	1/12	3/12
=	Equals	1/13	3/13
>	Greater than	1/14	3/14
?	Question mark	—	3/15
@	Commercial at	—	4/0
[	Left square bracket	3/11	5/11
]	Right square bracket	3/13	5/13
^	Upwards arrow, Circumflex accent (note ⑥)	—	5/14
_	Underline	—	5/15
`	Grave accent	—	6/0
—	Overline (note ⑤)	—	7/14

Notes ② ⑤ ⑥ see pages 6-7.

† See clause 4.1 page 10.

4. EXPLANATORY NOTES

4.1 Numbering of the positions in the code tables

Within any one character the bits are identified by b_7, b_6... b_1, where b_7 or b_6 is the highest order, or most significant bit, and b_1 is the lowest order, or least significant bit.

If desired these may be given a numerical significance in the binary system, thus:

Bit identification:	b_7	b_6	b_5	b_4	b_3	b_2	b_1
Significance:	64	32	16	8	4	2	1

In the code tables the columns and rows are identified by numbers written in binary and decimal notations.

Any one position in a code table may be identified either by its bit pattern, or by its column and row numbers. For instance, the position containing the figure 1 in the 7-bit set table may be identified:

— by its bit-pattern, e.g. 011 0001

— by its column and row numbers, e.g. 3/1.

4.2 Diacritical signs

In the 7 bit character set, some printing symbols may be designed to permit their use for the composition of accented letters when necessary for general interchange of information. A sequence of three characters, comprising a letter, " backspace " and one of these symbols, is needed for this composition, and the symbol is then regarded as a diacritical sign. It should be noted that these symbols take on their diacritical significance only when they are preceded or followed by the " backspace " character; for example, the symbol corresponding to the code combination 2/7 (') normally has the significance of " apostrophe ", but becomes the diacritical sign " acute accent " when it precedes or follows the " backspace " character.

In order to increase efficiency, it is possible to introduce accented letters (as single characters) in the positions marked by note ③ in the code tables. According to national requirements, these positions may contain special diacritical signs.

4.3 Interpretation of graphics

The meaning of the graphics is not defined by this ISO Recommendation. It will be necessary to reach agreement on the meaning and this will depend upon the particular application except in cases where other ISO Recommendations already exist. However no interpretation may be chosen which is contradictory to the customary meaning. A graphical symbol can have more than one meaning, e.g. the graphical symbol—(minus) also can have the meaning of hyphen or separation mark. The font design of the symbols is not part of this ISO Recommendation.

4.4 Dual allocations

A character allocated to a position in the code table may not be placed elsewhere in the table. In the case of positions having two characters allocated to them (1/10 and 1/11 of the 6-bit set table and 2/3, 3/10 and 3/11 of the 7-bit set table), the character not used cannot be placed elsewhere. This applies also to positions showing a preferred graphic. If such a position is filled with a national character the preferred graphic cannot be placed in another code position.

5. FUNCTIONAL CHARACTERISTICS RELATED TO CONTROL CHARACTERS

Some definitions in this section are stated in general terms and more explicit definitions of use may be needed for specific implementation of the code tables on recording media or on transmission channels. These more explicit definitions may become the subject of other ISO Recommendations.

A control character may be identified by a general designation, by a specific designation or by a combination of both.

5.1 General designations of control characters

The general designations of control characters involve a specific name followed by a subscript number.

They are defined as follows:

TC TRANSMISSION CONTROL

A functional character intended to control or facilitate transmission of information over telecommunication networks.

The use of the TC characters on the general telecommunication networks may be the subject of future ISO Recommendations.

FE FORMAT EFFECTOR

A functional character which controls the layout or positioning of information in Input/Output media.

DC DEVICE CONTROL

A functional character for the control of an ancillary device associated with a data processing or telecommunication system, for example an " On " or " Off " switching device.

The following is an example of how device control characters could be used in a specific tape system employing two ancillary tape punches and one ancillary tape reader:

DC_1 — First Punch " On "
DC_2 — Second Punch " On "
DC_3 — Tape reader " On "

The DC_4 character has a specific function, defined in clause 5.2.

IS INFORMATION SEPARATOR

A functional character which is used to separate and qualify information blocks logically. There is a group of four such characters, which are intended to be used in a hierarchical order.

F FUNCTION

A functional character in the 6-bit set, which may have the combined or alternative meanings of IS or FE.

Control characters with the general designations FE, IS or F have specific designations. These appear next to them, in parenthesis, in the code tables. These specific designations, the definitions of which are given in clause 5.2, have been chosen from the terminology used for some types of equipment. In other types of equipment, different but related meanings may be associated with certain of these control characters, but this requires agreement between the sender and the recipient of the data.

5.2 Specific designations of control characters

These are defined as follows:

ACK ACKNOWLEDGE

A Transmission Control character transmitted by a receiver as an affirmative response to the sender.

BEL BELL

A character for use when there is a need to call for human attention; it may control alarm or attention devices.

BS BACKSPACE

A layout character which controls the movement of the printing position one printing space backward on the same printing line.

CAN CANCEL

A character used to indicate that the information it accompanies is in error.

CR CARRIAGE RETURN

A layout character which controls the movement of the printing position to the first printing position on the same printing line.

DC_4 A Device Control character used to interrupt or turn off ancillary devices (STOP).

DEL DELETE

This character is used primarily to erase or obliterate erroneous or unwanted characters in punched tape. DEL characters may be inserted into or removed from a stream of data without affecting the information content of that stream. DEL characters may serve to accomplish media-fill or time-fill but then the addition or removal of these characters may affect the information layout and/or the control of equipment.

DLE DATA LINK ESCAPE

A Transmission Control character which will change the meaning of a limited number of contiguously following characters. It is used exclusively to provide supplementary data transmission control functions. Only graphics and Transmission Control characters can be used in DLE sequences.

EM END OF MEDIUM

A control character which may be used to identify the physical end of the medium, or the end of the used, or wanted, portion of information recorded on a medium. The position of this character does not necessarily correspond to the physical end of the medium.