INTERNATIONAL STANDARD ISO 646

7-bit coded character set for information processing interchange

First edition — 1973-07-01

UDC 681.3.042 : 003.62

Ref. No. ISO 646-1973 (E)

Descriptors : data processing, data transmission, character sets, coding, information interchange.

Price based on 12 pages

FOREWORD

ISO (the International Organization for Standardization) is a worldwide federation of national standards institutes (ISO Member Bodies). The work of developing International Standards is carried out through ISO Technical Committees. Every Member Body interested in a subject for which a Technical Committee has been set up has the right to be represented on that Committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work.

Draft International Standards adopted by the Technical Committees are circulated to the Member Bodies for approval before their acceptance as International Standards by the ISO Council.

International Standard ISO 646 was drawn up by Technical Committee ISO/TC 97, *Computers and information processing*, and circulated to the Member Bodies in April 1972.

It has been approved by the Member Bodies of the following countries :

Belgium	Ireland	Sweden
Brazil	Italy	Switzerland
Canada	Japan	Thailand
Czechoslovakia	Netherlands	United Kingdom
Denmark	Portugal	U.S.A.
Egypt, Arab Rep. of	Romania	U.S.S.R.
France	South Africa, Rep. of	
Germany	Spain	

No Member Body expressed disapproval of the document.

7-bit coded character set for information processing interchange

1 SCOPE AND FIELD OF APPLICATION

1.1 This International Standard contains a set of 128 characters (control characters and graphic characters such as letters, digits and symbols) with their coded representation. Most of these characters are mandatory and unchangeable, but provision is made for some flexibility to accommodate special national and other requirements.

1.2 The need for graphics and controls in data processing and in data transmission has been taken into account in determining this character set.

1.3 This International Standard consists of a general table with a number of options, notes, a legend and explanatory notes. It also contains a specific international reference version, guidance on the exercise of the options to define specific national versions and application oriented versions.

1.4 This character set is primarily intended for the interchange of information among data processing systems and associated equipment, and within message transmission systems.

1.5 This character set is applicable to all latin alphabets.

1.6 This character set includes facilities for extension where its 128 characters are insufficient for particular applications.

1.7 The definitions of some control characters in this International Standard assume that data associated with them is to be processed serially in a forward direction. Their effect when included in strings of data which are processed other than serially in a forward direction or included in data formatted for fixed record processing may have undesirable effects or may require additional special treatment to ensure that the control characters have their desired effect.

2 IMPLEMENTATION

2.1 This character set should be regarded as a basic alphabet in an abstract sense. Its practical use requires definitions of its implementation in various media. 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.

2.2 The implementation of this coded character set in physical media and for transmission, taking into account the need for error checking, is the subject of other ISO publications. (See Appendix Y.)

3 BASIC CODE TABLE

TABLE 1 — Basic code table

b7	b6	b5	column \ row	0	1	2	3	4	5	6	7
				0	0	0	0	1	1	1	1
				0	0	1	1	0	0	1	1
				0	1	0	1	0	1	0	1
b4	b3	b2	b1								
0	0	0	0	NUL	TC7 (DLE)	SP	0	③	P	` ④	p
0	0	0	1	TC1 (SOH)	DC1	!	1	A	Q	a	q
0	0	1	0	TC2 (STX)	DC2	" ⑥	2	B	R	b	r
0	0	1	1	TC3 (ETX)	DC3	£(#) ②	3	C	S	c	s
0	1	0	0	TC4 (EOT)	DC4	$(¤) ②	4	D	T	d	t
0	1	0	1	TC5 (ENQ)	TC8 (NAK)	%	5	E	U	e	u
0	1	1	0	TC6 (ACK)	TC9 (SYN)	&	6	F	V	f	v
0	1	1	1	BEL	TC10 (ETB)	' ⑥	7	G	W	g	w
1	0	0	0	FE0 (BS)	CAN	(	8	H	X	h	x
1	0	0	1	FE1 (HT)	EM	)	9	I	Y	i	y
1	0	1	0	FE2 (LF)①	SUB	*	:	J	Z	j	z
1	0	1	1	FE3 (VT)①	ESC	+	;	K	⑨	k	⑨
1	1	0	0	FE4 (FF)①	IS4 (FS)	, ⑥	<	L	③	l	③
1	1	0	1	FE5 (CR)①	IS3 (GS)	-	=	M	③	m	③
1	1	1	0	SO	IS2 (RS)	.	>	N	^ ④⑤	n	‾ ④⑤
1	1	1	1	SI	IS1 (US)	/	?	O	_	o	DEL

NOTES ABOUT TABLE 1

① The format effectors are intended for equipment in which horizontal and vertical movements are effected separately. If equipment requires the action of CARRIAGE RETURN to be combined with a vertical movement, the format effector for that vertical movement may be used to effect the combined movement. For example, if NEW LINE (symbol NL, equivalent to CR + LF) is required, FE_2 shall be used to represent it. This substitution requires agreement between the sender and the recipient of the data.

The use of these combined functions may be restricted for international transmission on general switched telecommunication networks (telegraph and telephone networks).

② The symbol £ is assigned to position 2/3 and the symbol $ is assigned to position 2/4. In a situation where there is no requirement for the symbol £ the symbol # (number sign) may be used in position 2/3. Where there is no requirement for the symbol $ the symbol ¤ (currency sign) may be used in position 2/4. The chosen allocations of symbols to these positions for international information interchange shall be agreed between the interested parties. It should be noted that, unless otherwise agreed between sender and recipient, the symbols £, $ or ¤ do not designate the currency of a specific country.

③ National use positions. The allocations of characters to these positions lies within the responsibility of national standardization bodies. These positions are primarily intended for alphabet extensions. If they are not required for that purpose, they may be used for symbols.

④ Positions 5/14, 6/0 and 7/14 are provided for the symbols UPWARD ARROW HEAD, GRAVE ACCENT and OVERLINE. However, these positions may be used for other graphical characters when it is necessary to have 8, 9 or 10 positions for national use.

⑤ Position 7/14 is used for the graphic character ‾ (OVERLINE), the graphical representation of which may vary according to national use to represent ˜(TILDE) or another diacritical sign provided that there is no risk of confusion with another graphic character included in the table.

⑥ The graphic characters in positions 2/2, 2/7, 2/12 and 5/14 have respectively the significance of QUOTATION MARK, APOSTROPHE, COMMA and UPWARD ARROW HEAD; however, these characters take on the significance of the diacritical signs DIAERESIS, ACUTE ACCENT, CEDILLA and CIRCUMFLEX ACCENT when they are preceded or followed by the BACKSPACE character (0/8).

4 LEGEND

4.1 Control characters

Abbrevi-ation	Note	Meaning	Position in the code table
ACK		Acknowledge	0/6
BEL		Bell	0/7
BS		Backspace	0/8
CAN		Cancel	1/8
CR	1	Carriage Return	0/13
DC		Device control	–
DEL		Delete	7/15
DLE		Data Link Escape	1/0
EM		End of Medium	1/9
ENQ		Enquiry	0/5
EOT		End of Transmission	0/4
ESC		Escape	1/11
ETB		End of Transmission Block	1/7
ETX		End of Text	0/3
FE		Format Effector	–
FF	1	Form Feed	0/12
FS		File Separator	1/12
GS		Group Separator	1/13
HT		Horizontal Tabulation	0/9
IS		Information Separator	–
LF	1	Line Feed	0/10
NAK		Negative Acknowledge	1/5
NUL		Null	0/0
RS		Record Separator	1/14
SI		Shift-In	0/15
SO		Shift-Out	0/14
SOH		Start of Heading	0/1
SP		Space (see 7.2)	2/0
STX		Start of Text	0/2
SUB		Substitute Character	1/10
SYN		Synchronous Idle	1/6
TC		Transmission Control	–
US		Unit Separator	1/15
VT	1	Vertical Tabulation	0/11

4.2 Graphic characters

Graphic	Note	Name	Position in the code table
(space)		Space (see 7.2)	2/0
!		Exclamation mark	2/1
"	6	Quotation mark, Diaeresis	2/2
£	2	Pound sign	2/3
#	2	Number sign	2/3
$	2	Dollar sign	2/4
¤	2	Currency sign	2/4
%		Percent sign	2/5
&		Ampersand	2/6
'	6	Apostrophe, acute accent	2/7
(		Left parenthesis	2/8
)		Right parenthesis	2/9
*		Asterisk	2/10
+		Plus sign	2/11
,	6	Comma, Cedilla	2/12
–		Hyphen, Minus sign	2/13
.		Full stop (period)	2/14
/		Solidus	2/15
:		Colon	3/10
;		Semi-colon	3/11
<		Less than sign	3/12
=		Equals sign	3/13
>		Greater than sign	3/14
?		Question mark	3/15
^	4, 6	Upward arrow head, Circumflex accent	5/14
_		Underline	5/15
`	4	Grave accent	6/0
~	4, 5	Overline, Tilde	7/14

5 EXPLANATORY NOTES

5.1 Numbering of the positions in Table 1

Within any one character the bits are identified by b_7, b_6 . . . b_1, where b_7 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 table the columns and rows are identified by numbers written in binary and decimal notations.

Any one position in the table may be identified either by its bit pattern, or by its column and row numbers. For instance, the position containing the digit 1 may be identified :

— by its bit pattern in order of decreasing significance, i.e. 011 0001 ;

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

The column number is derived from bits b_7, b_6 and b_5 giving them weights of 4, 2 and 1 respectively. The row number is derived from bits b_4, b_3, b_2 and b_1 giving them weights of 8, 4, 2 and 1 respectively.

5.2 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; the symbol is then regarded as a diacritical sign. It should be noted that these symbols take on their diacritical significance when they are preceded or followed by one "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 a "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 table. According to national requirements, these positions may contain special diacritical signs.

5.3 Names, meanings and fonts of graphic characters

This International Standard assigns at least one name to denote each of the graphic characters displayed in Tables 1 and 2. The names chosen to denote graphic characters are intended to reflect their customary meanings. However, this International Standard does not define and does not restrict the meanings of graphic characters. In addition, it does not specify a particular style or font design for the graphic characters.

Under the provision of Note ③ of Table 1, graphic characters which are different from the characters of the international reference version may be assigned to the national use positions. When such assignments are made, the graphic characters shall have distinct forms and be given distinctive names which are not in conflict with any of the forms or the names of any of the graphic characters in the international reference version.

5.4 Uniqueness of character allocation

A character allocated to a position in Table 1 may not be placed elsewhere in the table. For example, in the case of position 2/3 the character not used cannot be placed elsewhere. In particular the POUND sign (£) can never be represented by the bit combination of position 2/4.

6 VERSIONS OF TABLE 1

6.1 General

6.1.1 In order to use Table 1 for information interchange, it is necessary to exercise the options left open, i.e. those affected by Notes ② to ⑤. A single character must be allocated to each of the positions for which this freedom exists or it must be declared to be unused. A code table completed in this way is called a "version".

6.1.2 The notes to Table 1, the explanatory notes and the legend apply in full to any version.

6.2 International reference version

This version is available for use when there is no requirement to use a national or an application-oriented version. In international information processing interchange the international reference version (Table 2) is assumed unless a particular agreement exists between sender and recipient of the data. The following characters are allocated to the optional positions of Table 1 :

#	Number sign	2/3
¤	Currency sign	2/4
@	Commercial at	4/0
[	Left square bracket	5/11
\	Reverse solidus	5/12
]	Right square bracket	5/13
{	Left curly bracket	7/11
|	Vertical line	7/12
}	Right curly bracket	7/13

It should be noted that no substitution is allowed when using the international reference version.

TABLE 2 — International reference version

b7 →		0	0	0	0	1	1	1	1
b6 →		0	0	1	1	0	0	1	1
b5 →		0	1	0	1	0	1	0	1
b4 b3 b2 b1	column / row	0	1	2	3	4	5	6	7
0 0 0 0	0	NUL	TC7 (DLE)	SP	0	@	P	`	p
0 0 0 1	1	TC1 (SOH)	DC1	!	1	A	Q	a	q
0 0 1 0	2	TC2 (STX)	DC2	"	2	B	R	b	r
0 0 1 1	3	TC3 (ETX)	DC3	#	3	C	S	c	s
0 1 0 0	4	TC4 (EOT)	DC4	¤	4	D	T	d	t
0 1 0 1	5	TC5 (ENQ)	TC8 (NAK)	%	5	E	U	e	u
0 1 1 0	6	TC6 (ACK)	TC9 (SYN)	&	6	F	V	f	v
0 1 1 1	7	BEL	TC10 (ETB)	'	7	G	W	g	w
1 0 0 0	8	FE0 (BS)	CAN	(	8	H	X	h	x
1 0 0 1	9	FE1 (HT)	EM	)	9	I	Y	i	y
1 0 1 0	10	FE2 (LF)	SUB	*	:	J	Z	j	z
1 0 1 1	11	FE3 (VT)	ESC	+	;	K	[	k	{
1 1 0 0	12	FE4 (FF)	IS4 (FS)	,	<	L	\	l		
1 1 0 1	13	FE5 (CR)	IS3 (GS)	−	=	M	]	m	}
1 1 1 0	14	SO	IS2 (RS)	.	>	N	^	n	‾
1 1 1 1	15	SI	IS1 (US)	/	?	O	_	o	DEL

6.3 National versions

6.3.1 The responsibility for defining national versions lies with the national standardization bodies. These bodies shall exercise the options available and make the required selection.

6.3.2 If so required, more than one national version can be defined within a country. The different versions shall be separately identified. In particular when for a given national use position, for example, 5/12 or 6/0, alternative characters are required, two different versions shall be identified, even if they differ only by this single character.

6.3.3 If there is in a country no special demand for specific characters, it is strongly recommended that the characters of the international reference version be allocated to the same national use positions.

6.4 Application-oriented versions

Within national or international industries, organizations or professional groups, application-oriented versions can be used. They require precise agreement among the interested parties, who will have to exercise the options available and to make the required selection.

7 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 table on recording media or on transmission channels. These more explicit definitions and the use of these characters are the subject of other ISO Publications.

7.1 General designations of control characters

The general designation of control characters involves a specific class name followed by a subscript number.

They are defined as follows :

TC   **Transmission control characters**

Control characters intended to control or facilitate transmission of information over telecommunication networks.

The use of the TC characters on the general telecommunication networks is the subject of other ISO publications.

The transmission control characters are :

ACK, DLE, ENQ, EOT, ETB, ETX, NAK, SOH, STX and SYN.

FE   **Format effectors**

Control characters mainly intended for the control of the layout and positioning of information on printing and/or display devices. In the definitions of specific format effectors, any reference to printing devices should be interpreted as including display devices.

The definitions of format effectors use the following concept :

    a)   a page is composed of a number of lines of characters;

    b)   the characters forming a line occupy a number of positions called character positions;

    c)   the active position is that character position in which the character about to be processed would appear, if it were to be printed. The active position normally advances one character position at a time.

The format effector characters are : BS, CR, FF, HT, LF and VT (see also Note 1).

DC   **Device control characters**

Control characters for the control of a local or remote ancillary device (or devices) connected to a data processing and/or telecommunication system. These control characters are not intended to control telecommunication systems; this should be achieved by the use of TCs.

Certain preferred uses of the individual DCs are given in 7.2.

IS   **Information separators**

Control characters that are used to separate and qualify data logically. There are four such characters. They may be used either in hierarchical order or non-hierarchically; in the latter case their specific meanings depend on their applications.

When they are used hierarchically, the ascending order is :

US, RS, GS, FS.

In this case, data normally delimited by a particular separator cannot be split by a higher order separator but will be considered as delimited by any higher order separator.

7.2 Specific control characters

Individual members of the classes of controls are sometimes referred to by their abbreviated class name and a subscript number (for example, TC_5) and sometimes by a specific name indicative of their use (for example, ENQ).

Different but related meanings may be associated with some of the control characters but in an interchange of data this normally requires agreement between the sender and the recipient.

ACK Acknowledge

A transmission control character transmitted by a receiver as an affirmative response to the sender.

BEL Bell

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

BS Backspace

A format effector which moves the active position one character position backwards on the same line.

CAN Cancel

A character, or the first character of a sequence, indicating that the data preceding it is in error. As a result, this data is to be ignored. The specific meaning of this character must be defined for each application and/or between sender and recipient.

CR Carriage return

A format effector which moves the active position to the first character position on the same line.

Device controls

DC₁ A device control character which is primarily intended for turning on or starting an ancillary device. If it is not required for this purpose, it may be used to restore a device to the basic mode of operation (see also DC₂ and DC₃), or for any other device control function not provided by other DCs.

DC₂ A device control character which is primarily intended for turning on or starting an ancillary device. If it is not required for this purpose, it may be used to set a device to a special mode of operation (in which case DC₁ is used to restore the device to the basic mode), or for any other device control function not provided by other DCs.

DC₃ A device control character which is primarily intended for turning off or stopping an ancillary device. This function may be a secondary level stop, for example wait, pause, stand-by or halt (in which case DC₁ is used to restore normal operation). If it is not required for this purpose, it may be used for any other ancillary device control function not provided by other DCs.

DC₄ A device control character which is primarily intended for turning off, stopping or interrupting an ancillary device. If it is not required for this purpose, it may be used for any other device control function not provided by other DCs.

Examples of use of the device controls :

1) One switching

| | on−DC₂ | off−DC₄ |

2) Two independent switchings

| first one | on−DC₂ | off−DC₄ |
| second one | on−DC₁ | off−DC₃ |

3) Two dependent switchings

| general | on−DC₂ | off−DC₄ |
| particular | on−DC₁ | off−DC₃ |

4) Input and output switching

| output | on−DC₂ | off−DC₄ |
| input | on−DC₁ | off−DC₃ |

DEL Delete

A character used primarily to erase or obliterate an erroneous or unwanted character in punched tape. DEL characters may also serve to accomplish media-fill or time-fill. They may be inserted into or removed from a stream of data without affecting the information content of that stream 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 graphic characters and transmission control characters can be used in DLE sequences.

EM End of medium

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

ENQ Enquiry

A transmission control character used as a request for a response from a remote station — the response may include station identification and/or station status. When a "Who are you" function is required on the general switched transmission network, the first use of ENQ after the connection is established shall have the meaning "Who are you" (station identification). Subsequent use of ENQ may, or may not, include the function "Who are you", as determined by agreement.

EOT End of transmission

A transmission control character used to indicate the conclusion of the transmission of one or more texts.

ESC Escape

A control character which is used to provide additional control functions. It alters the meaning of a limited number of contiguously following bit combinations. The use of this character is specified in ISO 2022.

ETB End of transmission block

A transmission control character used to indicate the end of a transmission block of data where data is divided into such blocks for transmission purposes.

ETX End of text

A transmission control character which terminates a text.

FF Form feed

A format effector which advances the active position to the same character position on a pre-determined line of the next form or page.

HT Horizontal tabulation

A format effector which advances the active position to the next pre-determined character position on the same line.

Information separators

IS₁ (US) A control character used to separate and qualify data logically; its specific meaning has to be defined for each application. If this character is used in hierarchical order as specified in the general definition of IS, it delimits a data item called a UNIT.

IS₂ (RS) A control character used to separate and qualify data logically; its specific meaning has to be defined for each application. If this character is used in hierarchical order as specified in the general definition of IS, it delimits a data item called a RECORD.

IS₃ (GS) A control character used to separate and qualify data logically; its specific meaning has to be defined for each application. If this character is used in hierarchical order as specified in the general definition of IS, it delimits a data item called a GROUP.

IS₄ (FS) A control character used to separate and qualify data logically; its specific meaning has to be defined for each application. If this character is used in hierarchical order as specified in the general definition of IS, it delimits a data item called a FILE.

LF Line feed

A format effector which advances the active position to the same character position of the next line.

NAK Negative acknowledge

A transmission control character transmitted by a receiver as a negative response to the sender.

NUL Null

A control character used to accomplish media-fill or time-fill. NUL characters may be inserted into or removed from a stream of data without affecting the information content of that stream; but then the addition or removal of these characters may affect the information layout and/or the control of equipment.

SI Shift-in

A control character which is used in conjunction with SHIFT-OUT and ESCAPE to extend the graphic character set of the code. It may reinstate the standard meanings of the bit combinations which follow it. The effect of this character when using code extension techniques is described in ISO 2022.

SO Shift-out

A control character which is used in conjunction with SHIFT-IN and ESCAPE to extend the graphic character set of the code. It may alter the meaning of the bit combinations of columns 2 to 7 which follow it until a SHIFT-IN character is reached. However, the characters SPACE (2/0) and DELETE (7/15) are unaffected by SHIFT-OUT. The effect of this character when using code extension techniques is described in ISO 2022.

SOH Start of heading

A transmission control character used as the first character of a heading of an information message.

SP Space

A character which advances the active position one character position on the same line.

This character is also regarded as a non-printing graphic.

STX Start of text

A transmission control character which precedes a text and which is used to terminate a heading.

SUB Substitute character

A control character used in the place of a character that has been found to be invalid or in error. SUB is intended to be introduced by automatic means.

SYN Synchronous idle

A transmission control character used by a synchronous transmission system in the absence of any other character (idle condition) to provide a signal from which synchronism may be achieved or retained between data terminal equipment.

VT Vertical tabulation

A format effector which advances the active position to the same character position on the next pre-determined line.