INTERNATIONAL STANDARD ISO 4873

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION•МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ•ORGANISATION INTERNATIONALE DE NORMALISATION

Information processing — 8-bit coded character set for information interchange

Traitement de l'information — Jeu de caractères codés à 8 éléments pour l'échange d'information

First edition — 1979-02-15

STANDARDSISO.COM : Click to view the full PDF of ISO 4873:1979

UDC 681.3.04 : 003.62 **Ref. No. ISO 4873-1979 (E)**

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

Price based on 4 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 4873 was developed by Technical Committee ISO/TC 97, *Computers and information processing*, and was circulated to the member bodies in December 1977.

It has been approved by the member bodies of the following countries :

Australia	Japan	Sweden
Belgium	Mexico	Switzerland
Brazil	Netherlands	United Kingdom
Czechoslovakia	New Zealand	U.S.A.
Egypt, Arab Rep. of	Poland	U.S.S.R.
France	Romania	Yugoslavia
Germany, F.R.	South Africa, Rep. of	
Italy	Spain	

The member body of the following country expressed disapproval of the document on technical grounds :

Canada

STANDARDSISO.COM : Click to view the full PDF of ISO 4873:1979

© International Organization for Standardization, 1979 •

Printed in Switzerland

Information processing — 8-bit coded character set for information interchange

STANDARDSISO.COM : Click to view the full PDF of ISO 4873:1979

1 SCOPE AND FIELD OF APPLICATION

1.1 This International Standard specifies an 8-bit coded character set derived from, and compatible with, the 7-bit coded character set defined in ISO 646.

The characteristics of this set are also in conformance with the code extension techniques described in ISO 2022.

1.2 ISO 2022 defines, among other things, a family of 8-bit coded character sets, each of which accommodates a basic set of 128 characters together with 32 additional control characters and 94 additional graphic characters. The present International Standard is limited to the members of that family which have the characters of a version of the 7-bit coded character set of ISO 646 as their basic set.

1.3 While the 7-bit coded character set of ISO 646 is the standard coded character set for general information interchange, the 8-bit coded character set specified in this International Standard is intended for general information interchange within an 8-bit environment.

1.4 This International Standard consists of a basic table defining the character set and its coded representation, complemented by notes, legend and guidance on the exercise of the options to define specific versions.

2 REFERENCES

ISO 646, *7-bit coded character set for information processing interchange.*

ISO 2022, *Code extension techniques for use with the ISO 7-bit coded character set.*

3 IMPLEMENTATION

The implementation on physical media and in transmission, taking into account the need for error checking, will be the subject of other International Standards.

4 BASIC CODE TABLE

4.1 Table 1 shows a 16 by 16 array of columns numbered 00 to 15 and rows numbered 0 to 15 containing 256 code positions. Each position contains a character and/or a reference to notes. This table together with the notes constitutes the basic 8-bit coded character set. In order to use it, versions according to clause 5 must be defined. The requirements, definitions and explanatory notes given in ISO 646 are equally applicable to this International Standard, although they are not repeated here.

4.2 Columns 00 to 07 of the array contain 128 characters which are in one-to-one correspondence with the 128 characters of the basic set defined in ISO 646, except for positions 00/14 and 00/15. Their coded representation is the same as in the 7-bit code with the addition of an 8th most significant bit which is ZERO.

In particular, a national or an application-oriented version derived in accordance with ISO 646, or the International Reference Version of the 7-bit code, is an appropriate candidate for utilization in this part of the table.

4.3 Columns 08 to 15 of the array contain a further 128 code positions; the 8th bit of their coded representation is ONE.

4.3.1 Columns 08 and 09 are reserved for a set of 32 control characters. If this set has to be designated, the rules relevant to C1 sets defined in ISO 2022 apply. This set shall not include transmission control characters.

4.3.2 Columns 10 to 15, with the exception of positions 10/0 and 15/15, are reserved for a set of 94 graphic characters. If this set has to be designated, the rules relevant to G1 sets defined in ISO 2022 apply.

4.3.3 Positions 10/0 and 15/15 have no meaning assigned to them by this International Standard. The use of these bit-combinations requires agreement on their meaning and, when required, on their representation in a 7-bit environment.

TABLE 1 – **Basic code table**

b_4 b_3 b_2 b_1		b_8 0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
		b_7 0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
		b_6 0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
		b_5 0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
		00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
0 0 0 0	0	NUL	TC_7 (DLE)	SP	0	③	P	④ `	p			⑧					
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	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	① FE_2 (LF)	SUB	*	:	J	Z	j	z								
1 0 1 1	11	① FE_3 (VT)	ESC	+	;	K	③	k	③								
1 1 0 0	12	① FE_4 (FF)	IS_4 (FS)	⑥ ,	<	L	③	l	③								
1 1 0 1	13	① FE_5 (CR)	IS_3 (GS)	-	=	M	③	m	③								
1 1 1 0	14	⑦	IS_2 (RS)	.	>	N	④⑥ ^	n	④⑤ –								
1 1 1 1	15	⑦	IS_1 (US)	/	?	O	_	o	DEL								⑧
										⑨		⑩					

STANDARDSISO.COM : Click to view the full PDF of ISO 4873:1979

NOTES ON 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 tranmission on general switched telecommunication networks (telegraph and telephone networks).

② The symbol £ is assigned to position 02/3 and the symbol $ is assigned to position 02/4. In a situation where there is no requirement for the symbol £, the symbol # (NUMBER SIGN) may be used in position 02/3. Where there is no requirement for the symbol $, the symbol ¤ (CURRENCY SIGN) may be used in position 02/4. 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 allocation of characters to these positions lies within the responsability 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 other graphics.

④ Positions 05/14, 06/0 and 07/14 provide for the characters UPWARD ARROW HEAD, GRAVE ACCENT and OVERLINE. However, these positions may be used for other graphic characters when it is necessary to have 8, 9 or 10 positions for national use.

INTERNATIONAL STANDARD ISO 4873-1979/AMENDMENT 1

97

Published 1980-07-15

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION • МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ • ORGANISATION INTERNATIONALE DE NORMALISATION

Information processing — 8-bit coded character set for information interchange

AMENDMENT 1

Amendment 1 to International Standard ISO 4873-1979 was drawn up by Technical Committee ISO/TC 97, *Computers and information processing*. It was submitted directly to the ISO Council in accordance with clause 5.10.1 of part 1 of the Directives for the technical work of ISO.

Page 4

Clause 6 : Delete complete clause and replace by the following sentence : "If an announcement sequence, as described in ISO 2022 is employed for the default version, the sequence ESC 02/0 04/1 is appropriate".

STANDARDSISO.COM : Click to view the full PDF of ISO 4873:1979

UDC 681.3.04 : 003.62

Ref. No. ISO 4873-1979/A1-1980 (E)

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

Price based on 1 page

STANDARDSISO.COM : Click to view the full PDF of ISO 4873:1979

TABLE 2 — **Default version**

b_8	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
b_7	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
b_6	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
b_5	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
b_4 b_3 b_2 b_1																
0 0 0 0 0	NUL	TC_7 (DLE)	SP	0	@	P	`	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	"	2	B	R	b	r								
0 0 1 1 3	TC_3 (ETX)	DC_3	#	3	C	S	c	s								
0 1 0 0 4	TC_4 (EOT)	DC_4	¤	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)	'	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)	SUB	*	:	J	Z	j	z								
1 0 1 1 11	FE_3 (VT)	ESC	+	;	K	[	k	{								
1 1 0 0 12	FE_4 (FF)	IS_4 (FS)	,	<	L	\	l	|								
1 1 0 1 13	FE_5 (CR)	IS_3 (GS)	-	=	M	]	m	}								
1 1 1 0 14	(11)	IS_2 (RS)	.	>	N	^	n	~								
1 1 1 1 15	(11)	IS_1 (US)	/	?	O	_	o	DEL								
									NOT USED							

STANDARDSISO.COM : Click to view the full PDF of ISO 4873:1979

NOTES ON TABLE 1 (*continued*)

⑤ Position 07/14 is used for the graphic character ¯ (OVERLINE), the graphic 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 symbol included in the table.

⑥ The graphic characters in positions 02/2, 02/7, 02/12 and 05/14, 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 respectively when they are preceded or followed by the BACKSPACE character (00/8), (see also 5.2 of ISO 646).

⑦ No character is allocated to positions 00/14 and 00/15.

⑧ To ensure full compatibility with the code extension techniques described in ISO 2022, positions 10/0 and 15/15 are left undefined. The use of these two positions for data interchange requires agreement between the sender and the recipient of the data.

⑨ Positions for 08/0 to 09/15 are reserved for the allocation of a set of up to 32 control characters.

⑩ Positions from 10/1 to 15/14 are reserved for the allocation of a set of up to 94 graphic characters.

NOTE ON TABLE 2

⑪ No character is allocated to positions 00/14 and 00/15.

5 VERSIONS OF TABLE 1

5.1 General

In order to use table 1 for information interchange, it is necessary to exercise the options left open, i.e. those specified by notes ② to ⑤. A single character must be allocated to each of the positions for which these options apply or the position must be declared to be unused.

If there is no demand for specific characters, it is strongly recommended that the characters of the International Reference Version of the 7-bit coded character set be allocated to the corresponding code positions.

Furthermore, it is necessary to define the set of control characters allocated to columns 08 and 09, in accordance with note ⑨, and the set of graphic characters allocated to columns 10 to 15, in accordance with note ⑩. When there is no requirement for either or both of these sets, they must be declared to be unused.

A code table completed in this way is called a "version".

5.2 Definition of the code table

The definition of a complete code table (version), as described in 5.1, may be the subject of agreement between the interchanging parties, or may be specified by means of the ESCAPE sequences described below. In the absence of such agreement or escape sequences, the default version described in 5.3 is assumed.

Within the limits allowed by 5.1 and to avoid misinterpretation of the interchanged data, the code table in actual use can be defined by means of the three-character escape sequences described in ISO 2022, namely

a) ESC 02/1 (F), to designate the set of control characters in columns 00 and 01;

b) ESC 02/8 (F) or ESC 02/12 (F), to designate the set of graphic characters in columns 02 to 07;

c) ESC 02/2 (F), to designate the set of control characters in columns 08 and 09;

d) ESC 02/9 (F) or ESC 02/13 (F), to designate the set of graphic characters in columns 10 to 15.

In the foregoing, (F) stands for one of the bit-combinations from 03/0 to 07/14. For fuller details, see ISO 2022.

5.3 Default version

This version is available for use when there is no requirement to use a national or an application-oriented version in columns 00 to 07, and no requirement to use columns 08 to 15.

This version (table 2) is assumed, unless a particular agreement exists between sender and recipient of the data, or specific designation in accordance with 5.2 is used. Columns 00 to 07 of this version are identical to columns 0 to 7 of the International Reference Version (IRV) of the 7-bit coded character set (ISO 646), except that positions 00/14 and 00/15 must not be used. Columns 08 to 15 are not used.

6 ANNOUNCEMENT OF THE 8-BIT CODED CHARACTER SET

If an announcement sequence, as described in ISO 2022, is employed, the sequence ESC 02/0 04/3 is appropriate to the normal use of this 8-bit coded character set.

STANDARDSISO.COM : Click to view the full PDF of ISO 4873:1979