
International Standard



2033

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

Information processing — Coding of machine readable characters (MICR and OCR)

Traitement de l'information — Codage des jeux de caractères pour reconnaissance automatisée (MICR et ROC)

Second edition — 1983-05-15

STANDARDSISO.COM : Click to view the full PDF of ISO 2033:1983

UDC 631.3.048

Ref. No. ISO 2033-1983 (E)

Descriptors : data processing, character sets, coded representation, OCR-A character sets, OCR-B character sets, magnetic recognition.

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (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 authorized 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 2033 was developed by Technical Committee ISO/TC 97, *Information processing systems*, and was circulated to the member bodies in February 1982.

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

Belgium	Hungary	South Africa, Rep. of
Canada	Italy	Spain
China	Japan	Sweden
Czechoslovakia	Netherlands	Switzerland
Egypt, Arab Rep. of	Norway	United Kingdom
France	Poland	USA
Germany, F.R.	Romania	

No member body expressed disapproval of the document.

This second edition cancels and replaces the first edition (i.e. ISO 2033-1972).

Information processing — Coding of machine readable characters (MICR and OCR)

1 Scope

This International Standard defines the coded representation of printed characters recognized by reading equipment. It includes the fonts:

E 13 B	as covered in ISO 1004
CMC 7	as covered in ISO 1004
OCR-A	as covered in ISO 1073/1
OCR-B	as covered in ISO 1073/2

2 Field of application

This International Standard assigns bit-patterns to characters recognized by reading equipment. This coded information generated by the reading equipment is given to the recipient by different media, such as magnetic tape, by data transmission or a direct link. This coded representation can also be used by printing devices to print the information which shall later be read. It is not intended for general information interchange.

Two different applications are considered:

- Single-font reader:
The reading equipment is only capable of recognizing one font at a time.
- Multiple-font reader:
The reading equipment is capable of recognizing multiple fonts at the same time.

3 References

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

ISO 1004, *Information processing — Magnetic ink character recognition — Print specifications*.

ISO 1073, *Alphanumeric character sets for optical recognition*
— Part 1: Character set OCR-A — Shapes and dimensions of the printed image.

— Part 2: Character set OCR-B — Shapes and dimensions of the printed image.

ISO 2022, *Information processing — ISO 7-bit and 8-bit coded character sets — Code extension techniques*.

4 Coding

The coding given in this International Standard is based on the 7-bit code described in ISO 646 and on its extension to 8 bits according to ISO 2022.

The empty positions in code tables 1 to 5 are reserved for future standardization.

This International Standard does not define the character set to be read by the reading equipment.

Two codings are shown. The 8-bit coding is primarily intended for use with multi-font readers in the case where the 7-bit coding is not sufficient to represent the needed characters.

Independent of the coding shown in this International Standard, the code extension techniques given in ISO 2022 are applicable, i.e. the 7-bit coding shown in this International Standard may be transformed into 8-bit coding and the 8-bit coding shown in this International Standard may be transformed into 7-bit coding according to the rules of ISO 2022. Furthermore, the characters of columns 10 to 15 can equally be designated as a G1, G2 or G3 set.

References to positions of tables 1 to 5 in tables 6 to 9 are given by the notation "column number/row number". The column numbers for 7-bit coding consist of one digit and those for 8-bit coding consist of two digits. The notations b_1 to b_8 refer to the 7 bits or 8 bits of the coding whereby, b_1 is the low order bit.

Example

Capital letter F is shown in position 4/6 of the 7 bit table and 04/6 of the 8-bit table. This corresponds to bit pattern 1000110 and 01000110 respectively.

4.1 7-bit coding

The 7-bit coding can be used whenever the number of characters shown is sufficient for the application. This coding can also be used within an 8-bit environment by adding an eighth bit with the value 0, as defined in ISO 2022.

4.2 8-bit coding

The 8-bit coding can be used whenever the number of characters in a 7-bit table is insufficient for the application. The 8-bit coding can also be used within a 7-bit environment as defined in ISO 2022.

Table 1 — 8-bit coding of the characters of all fonts

				b ₇	b ₆	b ₅	b ₄	b ₃	b ₂	b ₁	b ₀																				
				0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
				0	0	0	0	1	1	1	1	0	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1		
				0	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0		
				0	1	0	1	0	1	0	1	0	1	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 ₇	b ₆	b ₅	b ₄	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1			
0	0	0	0	0			SP	0	à	P		p															À	à			
0	0	0	1	1			!	1	A	Q	a	q															Æ	æ			
0	0	1	0	2			"	2	B	R	b	r															Ä	ä			
0	0	1	1	3			#	3	C	S	c	s				£															
0	1	0	0	4			¤	4	D	T	d	t				\$															
0	1	0	1	5			%	5	E	U	e	u				¥															
0	1	1	0	6			&	6	F	V	f	v														IJ	ij				
0	1	1	1	7			'	7	G	W	g	w				§															
1	0	0	0	8			(	8	H	X	h	x														Ñ					
1	0	0	1	9			)	9	I	Y	i	y														Ø	ø				
1	0	1	0	10	LF	SUB	*	:	J	Z	j	z														ö	ö				
1	0	1	1	11			+	;	K	[	k	{														ß					
1	1	0	0	12			,	<	L	\	l						ſ														
1	1	0	1	13	CR		-	=	M	]	m	}					Ÿ								ü	ü					
1	1	1	0	14			.	>	N	^	n						ſ														
1	1	1	1	15			/	?	O	_	o	DEL																			

NOTE — The empty positions are reserved for future standardization.

Table 2 — 7-bit coding of the characters of the CMC 7 font

					b ₇				b ₆				b ₅							
					0	0	0	0	1	1	1	1	0	0	1	1	0	1	0	1
					0	1	2	3	4	5	6	7								
b ₄	b ₃	b ₂	b ₁																	
0	0	0	0	0				0		1										
0	0	0	1	1				1	1	1										
0	0	1	0	2				2	1	1										
0	0	1	1	3				3	0	1										
0	1	0	0	4				4	1	1										
0	1	0	1	5				5	0	1										
0	1	1	0	6				6	1	1										
0	1	1	1	7				7	0	1										
1	0	0	0	8				8	1	1										
1	0	0	1	9				9	0	1										
1	0	1	0	10	LF	SUB		10	1	1										
1	0	1	1	11				11	0	1										
1	1	0	0	12				12	1	1										
1	1	0	1	13	CR			13	0	1										
1	1	1	0	14				14	1	1										
1	1	1	1	15				15	0	1										

NOTE — The empty positions are reserved for future standardization.

Table 3 — 7-bit coding of the characters of the OCR-A font

					b ₇	0	0	0	0	1	1	1	1
					b ₆	0	0	1	1	0	0	1	1
					b ₅	0	1	0	1	0	1	0	1
						0	1	2	3	4	5	6	7
b ₄	b ₃	b ₂	b ₁										
0	0	0	0	0				0		P			
0	0	0	1	1				1	A	Q			
0	0	1	0	2			"	2	B	R			
0	0	1	1	3			£	3	C	S			
0	1	0	0	4			\$	4	D	T			
0	1	0	1	5			%	5	E	U			
0	1	1	0	6			&	6	F	V			
0	1	1	1	7			'	7	G	W			
1	0	0	0	8			{	8	H	X			
1	0	0	1	9			}	9	I	Y			
1	0	1	0	10	LF	SUB	*	:	J	Z			
1	0	1	1	11			+	;	K				
1	1	0	0	12			ˆ	ˆ	L				!
1	1	0	1	13	CR		-	=	M	4			
1	1	1	0	14			.	ˆ	N				
1	1	1	1	15			/	?	O				DEL

NOTE — The empty positions are reserved for future standardization.

Table 4 — 7-bit coding of the characters of the OCR-B font

					<table><tr><td>b₇</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td></tr><tr><td>b₆</td><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td></tr><tr><td>b₅</td><td>0</td><td>1</td><td>0</td><td>1</td><td>0</td><td>1</td><td>0</td><td>1</td></tr><tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td></tr></table>								b ₇	0	0	0	0	1	1	1	1	b ₆	0	0	1	1	0	0	1	1	b ₅	0	1	0	1	0	1	0	1		0	1	2	3	4	5	6	7
b ₇	0	0	0	0	1	1	1	1																																								
b ₆	0	0	1	1	0	0	1	1																																								
b ₅	0	1	0	1	0	1	0	1																																								
	0	1	2	3	4	5	6	7																																								
b ₄	b ₃	b ₂	b ₁																																													
0	0	0	0	0			SP	0	@	P		p																																				
0	0	0	1	1			!	1	A	Q	a	q																																				
0	0	1	0	2			"	2	B	R	b	r																																				
0	0	1	1	3			#	3	C	S	c	s																																				
0	1	0	0	4			¤	4	D	T	d	t																																				
0	1	0	1	5			%	5	E	U	e	u																																				
0	1	1	0	6			&	6	F	V	f	v																																				
0	1	1	1	7			'	7	G	W	g	w																																				
1	0	0	0	8			(	8	H	X	h	x																																				
1	0	0	1	9			)	9	I	Y	i	y																																				
1	0	1	0	10	LF	SUB	*	:	J	Z	j	z																																				
1	0	1	1	11			+	;	K	[	k	{																																				
1	1	0	0	12			,	<	L	\	l																																					
1	1	0	1	13	CR		-	=	M	]	m	}																																				
1	1	1	0	14			.	>	N	^	n																																					
1	1	1	1	15			/	?	O	_	o	DEL																																				

Table 5 — 7-bit coding of the characters of the E 13 B font

					b ₇	0	0	0	0	1	1	1	1
					b ₆	0	0	1	1	0	0	1	1
					b ₅	0	1	0	1	0	1	0	1
						0	1	2	3	4	5	6	7
b ₄	b ₃	b ₂	b ₁		0				0				
0	0	0	0	0	1				1				
0	0	1	0	2					2				
0	0	1	1	3					3				
0	1	0	0	4					4				
0	1	0	1	5					5				
0	1	1	0	6					6				
0	1	1	1	7					7				
1	0	0	0	8					8				
1	0	0	1	9					9				
1	0	1	0	10	LF	SUB			10				
1	0	1	1	11					11				
1	1	0	0	12					12				
1	1	0	1	13	CR				13				
1	1	1	0	14									
1	1	1	1	15									

NOTE — The empty positions are reserved for future standardization.

5 General considerations

5.1 End of line

If the information read by the equipment is structured in lines and if this structure should be maintained two possibilities are given:

- if the information is handled on a basis of records, then every line will form one record;
- if the information is handled character-by-character (data stream), the end of a line will be coded by means of control characters CR and LF.

5.2 Characters in error

If a character is recognized but cannot be identified as one character of the character set, control character SUB will be coded.

6 Font CMC 7

The characters of the CMC 7 font will be assigned to the positions of the code tables as specified in table 6 (see also tables 1 and 2).

Table 6 — Characters of the CMC 7 font and their assignment to positions in the code tables

Characters of CMC 7	8-bit code	7-bit code
Digits 0 to 9	03/0 to 03/9	3/0 to 3/9
Capital letters A to Z	04/1 to 05/10	4/1 to 5/10
 Symbol S I	10/10	3/10
 Symbol S II	10/11	3/11
 Symbol S III	10/12	3/12
 Symbol S IV	10/13	3/13
 Symbol S V	10/14	3/14

7 Font OCR-A

The characters of the OCR-A font will be assigned to the positions of the code tables specified in table 7 (see also tables 1 and 3).

7.1 Erase characters

Erase characters will normally be ignored by the reading equipment. If there is a requirement to code these characters, control character DEL shall be coded. For the character Group Erase one or more DEL may be coded.

Table 7 — Characters of the OCR-A font and their assignment to positions in the code tables

Characters of OCR-A	8-bit code	7-bit code
Digits 0 to 9	03/0 to 03/9	3/0 to 3/9
Capital letters A to Z	04/1 to 05/10	4/1 to 5/10
■ Full stop (period) *	02/14	2/14
⌌ Comma *	02/12	2/12
= Equals sign	03/13	3/13
+ Plus sign	02/11	2/11
— Hyphen, minus sign *	02/13	2/13
/ Solidus	02/15	2/15
* Asterisk	02/10	2/10
⌌ Abstract symbol H1 (hook)	11/12	3/12
⌌ Abstract symbol H2 (fork) *	11/13	5/13
⌌ Abstract symbol H3 (chair)	11/14	3/14
⌌ Abstract symbol H4 (long vertical mark)	07/12	7/12
■ Character erase *	07/15	7/15
— Group erase *	07/15	7/15
Å Capital letter Å	14/0	—
Ä Capital letter Ä	14/2	—
Æ Capital letter Æ	14/1	—
Ñ Capital letter Ñ	14/8	—
Ø Capital letter Ø	14/9	—
Ö Capital letter Ö	14/10	—
Ü Capital letter Ü	14/13	—
: Colon	03/10	3/10
⋮ Semicolon	03/11	3/11
? Question mark *	03/15	3/15

* See special description in the following clauses.

Table 7 — (concluded)

Characters of OCR-A	8-bit code	7-bit code
“ Quotation mark	02/2	2/2
’ Apostrophe *	02/7	2/7
{ Left curly bracket	07/11	2/8
} Right curly bracket	07/13	2/9
% Percent sign	02/5	2/5
& Ampersand	02/6	2/6
\$ Dollar sign	10/4	2/4
£ Pound sign	10/3	2/3
¥ Yen sign	10/5	—

* See special description in the following clauses.

7.2 Alternative characters

If there are alternative representations of one character, all of them will be assigned to the same position. This applies to characters Full stop, Comma, Hyphen, Question mark and Apostrophe.

7.3 Abstract symbol H2 (fork)

In general, this symbol is not used in conjunction with the alphabetic characters because of its potential interference with the letter Y.

However, when it is possible to recognize both characters correctly, the indicated coding shall be used.

8 Font OCR-B

The characters of the OCR-B font will be assigned to the positions of the code table as specified in table 8 (see also tables 1 and 4).

Table 8 — Characters of the OCR-B font and their assignment to positions in the code tables

Characters of OCR-B (number according to ISO 1073)	8-bit code	7-bit code
1 to 10 digits 0 to 9	03/0 to 03/9	3/0 to 3/9
11 to 36 Capital letters A to Z	04/1 to 05/10	4/1 to 5/10
37 to 62 Small letters a to z	06/1 to 07/10	6/1 to 7/10
63 * Asterisk	02/10	2/10
64 + Plus sign	02/11	2/11
65 — Hyphen (minus sign)	02/13	2/13
66 = Equals sign	03/13	3/13
67 / Solidus	02/15	2/15
68 . Full stop (period)	02/14	2/14
69 , Comma	02/12	2/12
70 : Colon	03/10	3/10
71 ; Semicolon	03/11	3/11
72 " Quotation mark	02/2	2/2
73 ' Apostrophe	02/7	2/7
74 — Discontinuous underline *	05/15	5/15
75 ? Question mark	03/15	3/15
76 ! Exclamation mark	02/1	2/1
77 (Left parenthesis	02/8	2/8
78) Right parenthesis	02/9	2/9
79 < Less than sign	03/12	3/12
80 > Greater than sign	03/14	3/14
81 [Left square bracket	05/11	5/11
82] Right square bracket	05/13	5/13
83 % Percent sign	02/5	2/5
84 # Number sign	02/3	2/3

* See special description in the following clauses.

Table 8 — (continued)

Characters of OCR-B (numbers according to ISO 1073)		8-bit code	7-bit code
85	& Ampersand	02/6	2/6
86	@ Commercial at	04/0	4/0
87	^ Upward arrow head	05/14	5/14
88	¤ Currency sign	02/4	2/4
89	£ Pound sign	10/3	—
90	\$ Dollar sign	10/4	—
91	Vertical line *	07/12	7/12
92	¡ Long vertical line *	07/12	7/12
93	\ Revers solidus	05/12	5/12
94	Ä Capital letter Ä	14/2	—
95	Å Capital letter Å	14/0	—
96	Æ Capital letter Æ	14/1	—
97	Ö Capital letter Ö	14/10	—
98	Ø Capital letter Ø	14/9	—
99	Ü Capital letter Ü	14/13	—
100	IJ Capital letter IJ	14/6	—
101	Ñ Capital letter Ñ	14/8	—
102	ä Small letter ä	15/0	—
103	æ Small letter æ	15/1	—
104	ø Small letter ø	15/9	—
105	ij Small letter ij	15/6	—

Table 8 — (concluded)

Characters of OCR-B (numbers according to ISO 1073)		8-bit code	7-bit code
106	ß Small letter ß	15/11	—
107	¨ Diaeresis *	—	—
107a	ä Small letter ä	15/2	—
107b	ö Small letter ö	15/10	—
107c	ü Small letter ü	15/13	—
108	´ Acute accent *	—	—
109	` Grave accent *	—	—
110	^ Circumflex accent *	—	—
111	~ Tilde *	—	—
112	¸ Cedilla *	—	—
113	{ Left curly bracket	07/11	7/11
114	} Right curly bracket	07/13	7/13
115	m Alternative small letter m *	06/13	6/13
116	— Continuous underline *	—	—
117	Space *	02/0	2/0
118	§ Paragraph (clause sign)	10/7	—
119	¥ Yen sign	10/5	—
120	■ Character erase *	07/15	7/15
121	— Group erase *	07/15	7/15

* See special description in the following clauses.

8.1 Underline characters

Two characters are provided for underlining

- No. 74 Discontinuous underline
- No. 116 Continuous underline

The latter, Continuous underline is not intended for use in OCR applications. The character Discontinuous underline shall be used in OCR applications as a free-standing character only, and shall not be printed under another character. It shall be coded as indicated.

8.2 Diacritical marks

The characters

- No. 107 Diaeresis
- No. 108 Acute accent
- No. 109 Grave accent
- No. 110 Circumflex accent
- No. 111 Tilde
- No. 112 Cedilla

will be used as a diacritical mark combined with a letter. In general, reading equipment recognizes such a combined character as one single character and will code it as one single character. If such a combined character is required and it is not contained in the code table as a single character, it shall be included in future revisions of this International Standard.

8.3 Space character

If the character space (No. 117) is recognized by the reading equipment, it may be coded as one or more characters SP.

8.4 Erase characters

Erase characters (No. 120 and 121) will normally be ignored by the reading equipment. If there is a requirement to code these characters, control character DEL is to be coded. For the character Group Erase (No. 121) one or more DEL may be coded.

8.5 Small letter m

The characters

- No. 49 Small letter m
- No. 115 Alternative small letter m

represent the same character and are coded in the same position in the code table.

8.6 Vertical line

The characters

- No. 91 Vertical line
- No. 92 Long vertical line

represent the same characters and are coded in the same position in the code table.

9 Font E 13 B

The characters of the E 13 B font will be assigned to the positions in the code table as specified in table 9 (see also tables 1 and 5).

Table 9 — Characters of the E 13 B font and their assignment to positions in the code tables

Characters of E 13 B	8-bit code	7-bit code
Digits 0 to 9	03/0 to 03/9	3/0 to 3/9
Symbol 1	13/10	3/10
Symbol 2	13/11	3/11
Symbol 3	13/12	3/12
Symbol 4	13/13	3/13