

# INTERNATIONAL STANDARD

ISO  
8859-6

First edition  
1987-08-15



---

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

---

**Information processing — 8-bit single-byte coded  
graphic character sets —**

**Part 6:  
Latin/Arabic alphabet**

STANDARD PREVIEW  
(standards.iteh.ai)

*Traitement de l'information — Jeux de caractères graphiques codés sur un seul octet —*

*Partie 6: Alphabet latin/arabe*

ISO 8859-6:1987

<https://standards.iteh.ai/catalog/standards/sist/9dc91713-0597-42b2-b582-e2100d42c851/iso-8859-6-1987>

Reference number  
ISO 8859-6:1987 (E)

## Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established 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. They are approved in accordance with ISO procedures requiring at least 75 % approval by the member bodies voting.

International Standard ISO 8859-6 was prepared by the European Computer Manufacturers Association as standard ECMA-114 and was adopted, under a special "fast-track" procedure, by Technical Committee ISO/TC 97, *Information processing systems*, in parallel with its approval by the ISO member bodies.

Users should note that all International Standards undergo revision from time to time and that any reference made herein to any other International Standard implies its latest edition, unless otherwise stated.

[ISO 8859-6:1987](https://standards.iteh.ai/catalog/standards/sist/9dc91713-0597-42b2-b582-e2100d42c85d/iso-8859-6-1987)

<https://standards.iteh.ai/catalog/standards/sist/9dc91713-0597-42b2-b582-e2100d42c85d/iso-8859-6-1987>

## Contents

	Page
<b>1</b> Scope .....	1
<b>2</b> Field of application .....	1
<b>3</b> Conformance .....	1
<b>4</b> References .....	1
<b>5</b> Definitions .....	1
<b>5.1</b> bit combination; byte .....	1
<b>5.2</b> character .....	1
<b>5.3</b> coded character set; code .....	1
<b>5.4</b> code table .....	1
<b>5.5</b> graphic character .....	1
<b>5.6</b> graphic symbol .....	2
<b>5.7</b> position .....	2
<b>6</b> Notation, code table and names .....	2
<b>6.1</b> Notation .....	2
<b>6.2</b> Layout of the code table .....	2
<b>6.3</b> Names and meanings .....	2
<b>6.3.1</b> SPACE (SP) .....	2
<b>6.3.2</b> NO-BREAK SPACE (NBSP) .....	2
<b>6.3.3</b> SOFT HYPHEN (SHY) .....	2
<b>6.4</b> Imaging graphic symbols .....	2
<b>6.4.1</b> Alternative sets of digits .....	2
<b>6.4.2</b> Mirror shapes .....	2
<b>7</b> Specification of the coded character set .....	2
<b>7.1</b> Characters of the set and their coded representation .....	3
<b>7.2</b> Code table .....	4
<b>7.3</b> Bit combinations not to be used .....	4
<b>8</b> Designation of the character set .....	4



# Information processing — 8-bit single-byte coded graphic character sets —

## Part 6: Latin/Arabic alphabet

### 1 Scope

This part of ISO 8859 specifies a set of 146 graphic characters identified as the Arabic/Latin alphabet and the coded representation of each of these characters by means of a single 8-bit byte.

The implementation of this part of ISO 8859 will require the use of specific control functions due to the fact that languages using the Latin script are written from left to right and Arabic is written from right to left. In addition, statements expressed with one of the scripts may have to be included in text using the other script. Also the insertion of numerical values may require special treatment.

The specification of the control functions to be used with the set of graphic characters contained herein is outside the scope of this part of ISO 8859.

### 2 Field of application

This set of graphic characters, the Latin/Arabic alphabet, is intended for use in information interchange as well as in data and text processing applications where both the Arabic and the Latin scripts are used. This set of graphic characters is suitable for use in a version of an 8-bit code according to ISO 2022 or ISO 4873.

### 3 Conformance

A set of graphic characters is in conformance with this part of ISO 8859 if it comprises all graphic characters specified herein to the exclusion of any other and if their coded representations are those specified by this part of ISO 8859.

### 4 References

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

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

ISO 4873, *Information processing — ISO 8-bit code for information interchange — Structure and rules for implementation*.

ISO 6429, *Information processing — ISO 7-bit and 8-bit coded character sets — Control functions for coded character sets*.<sup>1)</sup>

ISO 9036, *Information processing — Arabic 7-bit coded character set for information interchange*.

ASMO 449, *7-bit coded Arabic character set for information interchange*.

### 5 Definitions

For the purpose of this part of ISO 8859 the following definitions apply.

**5.1 bit combination; byte:** An ordered set of bits that represents a character or is used as a part of the representation of a character.

**5.2 character:** A member of a set of elements used for the organization, control or representation of data.

**5.3 coded character set; code:** A set of unambiguous rules that establishes a character set and the one-to-one relationship between each character of the set and its coded representation.

**5.4 code table:** A table showing the character allocated to each bit combination in a code.

**5.5 graphic character:** A character, other than a control function, that has a visual representation normally handwritten, printed or displayed, and that has a coded representation consisting of one or more bit combinations.

NOTE — In ISO 8859 a single bit combination is used to represent each graphic character.

1) At present at the stage of draft; publication anticipated in due course.

**5.6 graphic symbol:** A visual representation of a graphic character.

**5.7 position:** That part of a code table identified by its column and row co-ordinates.

## 6 Notation, code table and names

### 6.1 Notation

The bits of the bit combinations of the 8-bit code are identified by  $b_8$ ,  $b_7$ ,  $b_6$ ,  $b_5$ ,  $b_4$ ,  $b_3$ ,  $b_2$  and  $b_1$ , where  $b_8$  is the highest-order, or most-significant bit, and  $b_1$  is the lowest-order, or least-significant bit.

The bit combinations may be interpreted to represent numbers in binary notation by attributing the following weights to the individual bits:

Bit	$b_8$	$b_7$	$b_6$	$b_5$	$b_4$	$b_3$	$b_2$	$b_1$
Weight	128	64	32	16	8	4	2	1

Using these weights, the bit combinations of the 8-bit code represent numbers in the range 0 to 255.

In this part of ISO 8859, the bit combinations are identified by notations of the form  $xx/yy$ , where  $xx$  and  $yy$  are numbers in the range 00 to 15. The correspondence between the notations of the form  $xx/yy$  and the bit combinations consisting of the bits  $b_8$  to  $b_1$  is as follows:

- $xx$  is the number represented by  $b_8$ ,  $b_7$ ,  $b_6$  and  $b_5$  where these bits are given the weights 8, 4, 2 and 1 respectively;
- $yy$  is the number represented by  $b_4$ ,  $b_3$ ,  $b_2$  and  $b_1$  where these bits are given the weights 8, 4, 2 and 1 respectively.

### 6.2 Layout of the code table

An 8-bit code table consists of 256 positions arranged in 16 columns and 16 rows. The columns and the rows are numbered 00 to 15.

The code table positions are identified by notations of the form  $xx/yy$ , where  $xx$  is the column number and  $yy$  is the row number.

The positions of the code table are in one-to-one correspondence with the bit combinations of the code. The notation of a code table position, of the form  $xx/yy$ , is the same as that of the corresponding bit combination.

### 6.3 Names and meanings

This part of ISO 8859 assigns at least one name to each character. In addition, it specifies a unique graphic symbol for each graphic character, except for those specified in 6.4. By

convention only capital letters, the graphic symbols of small letters and hyphens are used for writing the names of the characters.

NOTE — The graphic symbols of each Arabic alphabetic character may have more than one shape. The appropriate shape to be imaged is determined by contextual analysis, which is outside the scope of this part of ISO 8859.

The names chosen to denote graphic characters are intended to reflect their customary meaning. However, except for SPACE (SP), NO-BREAK SPACE (NBSP) and SOFT HYPHEN (SHY), this part of ISO 8859 does not define and does not restrict the meanings of graphic characters. Neither does it specify a particular style or font design for imaging graphic characters.

#### 6.3.1 SPACE (SP)

This character may be interpreted as a graphic character, a control character or as both. As a graphic character it has the visual representation consisting of the absence of a graphic symbol.

#### 6.3.2 NO-BREAK SPACE (NBSP)

A graphic character the visual representation of which consists of the absence of a graphic symbol, for use when a line break is to be prevented in the text as presented.

#### 6.3.3 SOFT HYPHEN (SHY)

A graphic character that is imaged by a graphic symbol identical with, or similar to, that representing HYPHEN, for use when a line break has been established within a word.

### 6.4 Imaging graphic symbols

#### 6.4.1 Alternative sets of digits

The characters allocated to positions 03/00 to 03/09 are shown with two alternative graphic symbols. The left-hand set of graphic symbols is that of the Latin script, the right-hand set is that of the Arabic script.

#### 6.4.2 Mirror shapes

The graphic symbols shown in the code table in positions 02/08, 02/09, 03/12, 03/14, 05/11, 05/13, 07/11 and 07/13 are those to be imaged unless preceded by an appropriate control function. In this case their mirror shape shall be imaged.

## 7 Specification of the coded character set

This part of ISO 8859 specifies 146 graphic characters allocated to the bit combinations of the code table (table 2). None of these characters are "non-spacing".

The use of control functions, such as BACKSPACE or CARRIAGE RETURN for the coded representation of composite characters is prohibited by this part of ISO 8859.

## 7.1 Characters of the set and their coded representation

Table 1 — Character set — Coded representation

Bit combination	Name	Bit combination	Name
02/00	SPACE (see 6.3)	06/00	GRAVE ACCENT
02/01	EXCLAMATION MARK	06/01	SMALL LETTER a
02/02	QUOTATION MARK	06/02	SMALL LETTER b
02/03	NUMBER SIGN	06/03	SMALL LETTER c
02/04	DOLLAR SIGN	06/04	SMALL LETTER d
02/05	PERCENT SIGN	06/05	SMALL LETTER e
02/06	AMPERSAND	06/06	SMALL LETTER f
02/07	APOSTROPHE	06/07	SMALL LETTER g
02/08	LEFT PARENTHESIS	06/08	SMALL LETTER h
02/09	RIGHT PARENTHESIS	06/09	SMALL LETTER i
02/10	ASTERISK	06/10	SMALL LETTER j
02/11	PLUS SIGN	06/11	SMALL LETTER k
02/12	COMMA	06/12	SMALL LETTER l
02/13	HYPHEN, MINUS SIGN	06/13	SMALL LETTER m
02/14	FULL STOP	06/14	SMALL LETTER n
02/15	SOLIDUS	06/15	SMALL LETTER o
03/00	DIGIT ZERO	07/00	SMALL LETTER p
03/01	DIGIT ONE	07/01	SMALL LETTER q
03/02	DIGIT TWO	07/02	SMALL LETTER r
03/03	DIGIT THREE	07/03	SMALL LETTER s
03/04	DIGIT FOUR	07/04	SMALL LETTER t
03/05	DIGIT FIVE	07/05	SMALL LETTER u
03/06	DIGIT SIX	07/06	SMALL LETTER v
03/07	DIGIT SEVEN	07/07	SMALL LETTER w
03/08	DIGIT EIGHT	07/08	SMALL LETTER x
03/09	DIGIT NINE	07/09	SMALL LETTER y
03/10	COLON	07/10	SMALL LETTER z
03/11	SEMICOLON	07/11	LEFT CURLY BRACKET
03/12	LESS-THAN SIGN	07/12	VERTICAL LINE
03/13	EQUALS SIGN	07/13	RIGHT CURLY BRACKET
03/14	GREATER-THAN SIGN	07/14	TILDE
03/15	QUESTION MARK	10/00	NO-BREAK SPACE (see 6.3)
04/00	COMMERCIAL AT	10/01	(This position shall not be used)
04/01	CAPITAL LETTER A	10/02	(This position shall not be used)
04/02	CAPITAL LETTER B	10/03	(This position shall not be used)
04/03	CAPITAL LETTER C	10/04	CURRENCY SIGN
04/04	CAPITAL LETTER D	10/05	(This position shall not be used)
04/05	CAPITAL LETTER E	10/06	(This position shall not be used)
04/06	CAPITAL LETTER F	10/07	(This position shall not be used)
04/07	CAPITAL LETTER G	10/08	(This position shall not be used)
04/08	CAPITAL LETTER H	10/09	(This position shall not be used)
04/09	CAPITAL LETTER I	10/10	(This position shall not be used)
04/10	CAPITAL LETTER J	10/11	(This position shall not be used)
04/11	CAPITAL LETTER K	10/12	ARABIC COMA
04/12	CAPITAL LETTER L	10/13	SOFT HYPHEN (see 6.3)
04/13	CAPITAL LETTER M	10/14	(This position shall not be used)
04/14	CAPITAL LETTER N	10/15	(This position shall not be used)
04/15	CAPITAL LETTER O	11/00	(This position shall not be used)
05/00	CAPITAL LETTER P	11/01	(This position shall not be used)
05/01	CAPITAL LETTER Q	11/02	(This position shall not be used)
05/02	CAPITAL LETTER R	11/03	(This position shall not be used)
05/03	CAPITAL LETTER S	11/04	(This position shall not be used)
05/04	CAPITAL LETTER T	11/05	(This position shall not be used)
05/05	CAPITAL LETTER U	11/06	(This position shall not be used)
05/06	CAPITAL LETTER V	11/07	(This position shall not be used)
05/07	CAPITAL LETTER W	11/08	(This position shall not be used)
05/08	CAPITAL LETTER X	11/09	(This position shall not be used)
05/09	CAPITAL LETTER Y	11/10	(This position shall not be used)
05/10	CAPITAL LETTER Z	11/11	ARABIC SEMICOLON
05/11	LEFT SQUARE BRACKET	11/12	(This position shall not be used)
05/12	REVERSE SOLIDUS	11/13	(This position shall not be used)
05/13	RIGHT SQUARE BRACKET	11/14	(This position shall not be used)
05/14	CIRCUMFLEX ACCENT	11/15	ARABIC QUESTION MARK
05/15	LOW LINE	12/00	(This position shall not be used)

Table 1 (concluded)

Bit combination	Name
12/01	HAMZA
12/02	MADDA ON ALEF
12/03	HAMZA ON ALEF
12/04	HAMZA ON WAW
12/05	HAMZA UNDER ALEF
12/06	HAMZA ON YEH
12/07	ALEF
12/08	BEH
12/09	TEH MARBUTA
12/10	TEH
12/11	THEH
12/12	JEEM
12/13	HAH
12/14	KHAH
12/15	DAL
13/00	THAL
13/01	RA
13/02	ZAIN
13/03	SEEN
13/04	SHEEN
13/05	SAD
13/06	DAD
13/07	TAH
13/08	ZAH
13/09	AIN
13/10	GHAIN
13/11	(This position shall not be used)
13/12	(This position shall not be used)
13/13	(This position shall not be used)
13/14	(This position shall not be used)
13/15	(This position shall not be used)
14/00	TATWEEL
14/01	FEH
14/02	QAF
14/03	KAF
14/04	LAM
14/05	MEEM
14/06	NOON
14/07	HA
14/08	WAW
14/09	ALEF MAKSURA
14/10	YEH
14/11	FATHATAN
14/12	DAMMATAN
14/13	KASRATAN
14/14	FATHA
14/15	DAMMA
15/00	KASRA
15/01	SHADDA
15/02	SUKUN
15/03	(This position shall not be used)
15/04	(This position shall not be used)
15/05	(This position shall not be used)
15/06	(This position shall not be used)
15/07	(This position shall not be used)
15/08	(This position shall not be used)
15/09	(This position shall not be used)
15/10	(This position shall not be used)
15/11	(This position shall not be used)
15/12	(This position shall not be used)
15/13	(This position shall not be used)
15/14	(This position shall not be used)
15/15	(This position shall not be used)

7.2 Code table

The code table (table 2) shows the characters listed at the position in the code table corresponding to the specified bit combination.

The shaded positions correspond to bit combinations that do not represent graphic characters. They are reserved for the representation of control functions.

7.3 Bit combinations not to be used

A number of bit combinations are reserved for future standardization and shall not be used. They are hatched in the code table (table 2).

8 Designation of the character set

The graphic characters of this part of ISO 8859 constitute a single coded character set. However, when this character set is implemented together with other coding standards such as ISO 2022 or ISO 4873, the code table (table 2) of this part of ISO 8859 shall be considered to consist of the following components:

- The character SPACE represented by bit combination 02/00.
- A 94-character G0 graphic character set represented by bit combinations 02/01 to 07/14.
- A 96-character G1 graphic character set represented by bit combinations 10/00 to 15/15.

When required by other coding standards, for example ISO 2022 or ISO 4873, the following pair of escape sequences shall be used:

ESC 02/08 04/02  
ESC 02/13 04/07

to designate the G0 and the G1 sets, respectively. According to ISO 2022, the character SPACE does not require designation.



Table 2 — Code table of the Latin/Arabic alphabet

				b <sub>0</sub>	b <sub>1</sub>	b <sub>2</sub>	b <sub>3</sub>	b <sub>4</sub>	b <sub>5</sub>	b <sub>6</sub>	b <sub>7</sub>	b <sub>8</sub>	b <sub>9</sub>	b <sub>10</sub>	b <sub>11</sub>	b <sub>12</sub>	b <sub>13</sub>	b <sub>14</sub>	b <sub>15</sub>			
				0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1		
				0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	1	1	
				0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	1	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 <sub>0</sub>	b <sub>1</sub>	b <sub>2</sub>	b <sub>3</sub>																			
0	0	0	0	00			SP	0	·	à	P	`	p			NBSP			ذ	—	ء	
0	0	0	1	01			!	1	\	A	Q	a	q						ء	ر	ف	ء
0	0	1	0	02			"	2	ʾ	B	R	b	r						آ	ز	ق	ه
0	0	1	1	03			#	3	ʾ	C	S	c	s						أ	س	ك	
0	1	0	0	04			\$	4	ء	D	T	d	t			Ⲁ			و	ش	ل	
0	1	0	1	05			%	5	ء	E	U	e	u						أ	ص	م	
0	1	1	0	06			&	6	ʾ	F	V	f	v						ر	ض	ن	
0	1	1	1	07			'	7	ʾ	G	W	g	w						ا	ط	ه	
1	0	0	0	08			(	8	ʾ	H	X	h	x						ب	ظ	و	
1	0	0	1	09			)	9	ʾ	I	Y	i	y						ة	ع	ى	
1	0	1	0	10			*	:		J	Z	j	z						ت	غ	ي	
1	0	1	1	11			+	;		K	[	k	{						؛	ث	ء	
1	1	0	0	12			,	<		L	\	l							،	ج		"
1	1	0	1	13			-	=		M	]	m	}						SHY	ح		ء
1	1	1	0	14			.	>		N	^	n	~							خ		ء
1	1	1	1	15			/	?		O	_	o								؟	د	'

