INTERNATIONAL STANDARD

**ISO
8859-1**

ISO

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

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

Part 1:
Latin alphabet No. 1

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

Partie 1: Alphabet latin n° 1

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 8859-1 was prepared by Technical Committee ISO/TC 97, *Information processing systems.*

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.

Contents

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

Part 1:
Latin alphabet No. 1

0 Introduction

ISO 8859 consists of several parts. Each part specifies a set of up to 191 graphic characters and the coded representation of each of these characters by means of a single 8-bit byte. The use of control functions for the coded representation of composite characters is prohibited by ISO 8859. Each set is intended for use for a group of languages.

ISO 8859/2 specifies a set of 191 graphic characters identified as Latin alphabet No. 2.

1 Scope

This part of ISO 8859 specifies a set of 191 graphic characters identified as Latin alphabet No. 1.

2 Field of application

This set of graphic characters, the Latin alphabet No. 1, is intended for use in data processing and text applications and may also be used for information interchange.

The set contains graphic characters used for general purpose applications in typical office environments in at least the following languages:

Danish, Dutch, English, Faeroese, Finnish, French, German, Icelandic, Irish, Italian, Norwegian, Portuguese, Spanish and Swedish.

This set of graphic characters is suitable for use in a version of an 8-bit code according to ISO 2022 or ISO 4873.

NOTE — ISO 8859 is not intended for use with CCITT-defined Telematic services. If information coded according to ISO 8859 is to be transferred to such services, it will have to conform at the coding interface to their requirements.

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.

Equipment claimed to implement this part of ISO 8859 shall implement all 191 characters.

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 — Additional control functions for character-imaging devices.

ISO 6937/2, Information processing — Coded character sets for text communication — Part 2: Latin alphabetic and non-alphabetic graphic characters.

5 Definitons

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 character.

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 coordinates.

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

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.

7 Specification of the coded character set

This part of ISO 8859 specifies 191 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 ISO 8859.

7.1 Characters of the set and their coded representation

Table 1 — Character set — Coded representation

Bit combi-nation	Name	Bit combi-nation	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	INVERTED EXCLAMATION MARK
04/01	CAPITAL LETTER A	10/02	CENT SIGN
04/02	CAPITAL LETTER B	10/03	POUND SIGN
04/03	CAPITAL LETTER C	10/04	CURRENCY SIGN
04/04	CAPITAL LETTER D	10/05	YEN SIGN
04/05	CAPITAL LETTER E	10/06	BROKEN BAR
04/06	CAPITAL LETTER F	10/07	PARAGRAPH SIGN, SECTION SIGN
04/07	CAPITAL LETTER G	10/08	DIAERESIS
04/08	CAPITAL LETTER H	10/09	COPYRIGHT SIGN
04/09	CAPITAL LETTER I	10/10	FEMININE ORDINAL INDICATOR
04/10	CAPITAL LETTER J	10/11	LEFT ANGLE QUOTATION MARK
04/11	CAPITAL LETTER K	10/12	NOT SIGN
04/12	CAPITAL LETTER L	10/13	SOFT HYPHEN (see 6.3)
04/13	CAPITAL LETTER M	10/14	REGISTERED TRADE MARK SIGN
04/14	CAPITAL LETTER N	10/15	MACRON
04/15	CAPITAL LETTER O	11/00	RING ABOVE, DEGREE SIGN
05/00	CAPITAL LETTER P	11/01	PLUS-MINUS SIGN
05/01	CAPITAL LETTER Q	11/02	SUPERSCRIPT TWO
05/02	CAPITAL LETTER R	11/03	SUPERSCRIPT THREE
05/03	CAPITAL LETTER S	11/04	ACUTE ACCENT
05/04	CAPITAL LETTER T	11/05	MICRO SIGN
05/05	CAPITAL LETTER U	11/06	PILCROW SIGN
05/06	CAPITAL LETTER V	11/07	MIDDLE DOT
05/07	CAPITAL LETTER W	11/08	CEDILLA
05/08	CAPITAL LETTER X	11/09	SUPERSCRIPT ONE
05/09	CAPITAL LETTER Y	11/10	MASCULINE ORDINAL INDICATOR
05/10	CAPITAL LETTER Z	11/11	RIGHT ANGLE QUOTATION MARK
05/11	LEFT SQUARE BRACKET	11/12	VULGAR FRACTION ONE QUARTER
05/12	REVERSE SOLIDUS	11/13	VULGAR FRACTION ONE HALF
05/13	RIGHT SQUARE BRACKET	11/14	VULGAR FRACTION THREE QUARTERS
05/14	CIRCUMFLEX ACCENT	11/15	INVERTED QUESTION MARK
05/15	LOW LINE	12/00	CAPITAL LETTER A WITH GRAVE ACCENT

Table 1 — *(concluded)*

Bit combination	Name
12/01	CAPITAL LETTER A WITH ACUTE ACCENT
12/02	CAPITAL LETTER A WITH CIRCUMFLEX ACCENT
12/03	CAPITAL LETTER A WITH TILDE
12/04	CAPITAL LETTER A WITH DIAERESIS
12/05	CAPITAL LETTER A WITH RING ABOVE
12/06	CAPITAL DIPHTHONG A WITH E
12/07	CAPITAL LETTER C WITH CEDILLA
12/08	CAPITAL LETTER E WITH GRAVE ACCENT
12/09	CAPITAL LETTER E WITH ACUTE ACCENT
12/10	CAPITAL LETTER E WITH CIRCUMFLEX ACCENT
12/11	CAPITAL LETTER E WITH DIAERESIS
12/12	CAPITAL LETTER I WITH GRAVE ACCENT
12/13	CAPITAL LETTER I WITH ACUTE ACCENT
12/14	CAPITAL LETTER I WITH CIRCUMFLEX ACCENT
12/15	CAPITAL LETTER I WITH DIAERESIS
13/00	CAPITAL ICELANDIC LETTER ETH
13/01	CAPITAL LETTER N WITH TILDE
13/02	CAPITAL LETTER O WITH GRAVE ACCENT
13/03	CAPITAL LETTER O WITH ACUTE ACCENT
13/04	CAPITAL LETTER O WITH CIRCUMFLEX ACCENT
13/05	CAPITAL LETTER O WITH TILDE
13/06	CAPITAL LETTER O WITH DIAERESIS
13/07	MULTIPLICATION SIGN
13/08	CAPITAL LETTER O WITH OBLIQUE STROKE
13/09	CAPITAL LETTER U WITH GRAVE ACCENT
13/10	CAPITAL LETTER U WITH ACUTE ACCENT
13/11	CAPITAL LETTER U WITH CIRCUMFLEX ACCENT
13/12	CAPITAL LETTER U WITH DIAERESIS
13/13	CAPITAL LETTER Y WITH ACUTE ACCENT
13/14	CAPITAL ICELANDIC LETTER THORN
13/15	SMALL GERMAN LETTER SHARP s
14/00	SMALL LETTER a WITH GRAVE ACCENT
14/01	SMALL LETTER a WITH ACUTE ACCENT
14/02	SMALL LETTER a WITH CIRCUMFLEX ACCENT
14/03	SMALL LETTER a WITH TILDE
14/04	SMALL LETTER a WITH DIAERESIS
14/05	SMALL LETTER a WITH RING ABOVE
14/06	SMALL DIPHTHONG a WITH e
14/07	SMALL LETTER c WITH CEDILLA
14/08	SMALL LETTER e WITH GRAVE ACCENT
14/09	SMALL LETTER e WITH ACUTE ACCENT
14/10	SMALL LETTER e WITH CIRCUMFLEX ACCENT
14/11	SMALL LETTER e WITH DIAERESIS
14/12	SMALL LETTER i WITH GRAVE ACCENT
14/13	SMALL LETTER i WITH ACUTE ACCENT
14/14	SMALL LETTER i WITH CIRCUMFLEX ACCENT
14/15	SMALL LETTER i WITH DIAERESIS
15/00	SMALL ICELANDIC LETTER ETH
15/01	SMALL LETTER n WITH TILDE
15/02	SMALL LETTER o WITH GRAVE ACCENT
15/03	SMALL LETTER o WITH ACUTE ACCENT
15/04	SMALL LETTER o WITH CIRCUMFLEX ACCENT
15/05	SMALL LETTER o WITH TILDE
15/06	SMALL LETTER o WITH DIAERESIS
15/07	DIVISION SIGN
15/08	SMALL LETTER o WITH OBLIQUE STROKE
15/09	SMALL LETTER u WITH GRAVE ACCENT
15/10	SMALL LETTER u WITH ACUTE ACCENT
15/11	SMALL LETTER u WITH CIRCUMFLEX ACCENT
15/12	SMALL LETTER u WITH DIAERESIS
15/13	SMALL LETTER y WITH ACUTE ACCENT
15/14	SMALL ICELANDIC LETTER THORN
15/15	SMALL LETTER y WITH DIAERESIS

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. Their use is outside the scope of ISO 8859; it is specified in other International Standards, for example ISO 646 or ISO 6429.

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 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 04/01

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

Table 2 — Code table

b8 b7 b6 b5 → / a8 b4 b2 b1 ↓	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
00			SP	0	@	P	`	p			NBSP	°	À	Ð	à	ð
01			!	1	A	Q	a	q			¡	±	Á	Ñ	á	ñ
02			"	2	B	R	b	r			¢	²	Â	Ò	â	ò
03			#	3	C	S	c	s			£	³	Ã	Ó	ã	ó
04			$	4	D	T	d	t			¤	´	Ä	Ô	ä	ô
05			%	5	E	U	e	u			¥	µ	Å	Õ	å	õ
06			&	6	F	V	f	v			¦	¶	Æ	Ö	æ	ö
07			'	7	G	W	g	w			§	·	Ç	×	ç	÷
08			(	8	H	X	h	x			¨	¸	È	Ø	è	ø
09			)	9	I	Y	i	y			©	¹	É	Ù	é	ù
10			*	:	J	Z	j	z			ª	º	Ê	Ú	ê	ú
11			+	;	K	[	k	{			«	»	Ë	Û	ë	û
12			,	<	L	\	l	|			¬	¼	Ì	Ü	ì	ü
13			-	=	M	]	m	}			SHY	½	Í	Ý	í	ý
14			.	>	N	^	n	~			®	¾	Î	Þ	î	þ
15			/	?	O	_	o				¯	¿	Ï	ß	ï	ÿ