INTERNATIONAL STANDARD

ISO/IEC
8859-5

First edition
1988-12-01

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

Part 5:
Latin/Cyrillic alphabet

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

Partie 5: Alphabet latin/cyrillique

Contents

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) together form a system for worldwide standardization as a whole. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1. Draft International Standards adopted by the joint technical committee are circulated to national bodies for approval before their acceptance as International Standards. They are approved in accordance with procedures requiring at least 75 % approval by the national bodies voting.

International Standard ISO/IEC 8859-5 was prepared by the European Computer Manufacturers Association (as standard ECMA-113) and was adopted, under a special "fast-track procedure", by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, in parallel with its approval by national bodies of ISO and IEC.

ISO/IEC 8859 consists of the following parts, under the general title *Information processing — 8-bit single-byte coded graphic character sets*:

— *Part 1: Latin alphabet No. 1*

— *Part 2: Latin alphabet No. 2*

— *Part 3: Latin alphabet No. 3*

— *Part 4: Latin alphabet No. 4*

— *Part 5: Latin/Cyrillic alphabet*

— *Part 6: Latin/Arabic alphabet*

— *Part 7: Latin/Greek alphabet*

— *Part 8: Latin/Hebrew alphabet*

— *Part 9: Latin alphabet No. 5*

Introduction

ISO/IEC 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/IEC 8859. Each set is intended for use for a group of languages.

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

Part 5:
Latin/Cyrillic alphabet

1 Scope

This part of ISO/IEC 8859 defines a set of 191 graphic characters identified as the Latin/Cyrillic alphabet, and specifies the coded representation of each of these characters by means of a single 8-bit byte. None of these characters is "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/IEC 8859.

This set of graphic characters, the Latin/Cyrillic alphabet, is intended for use in data and text processing applications and may also be used for information interchange.

This set is suited for multiple-language applications involving the Latin and the Cyrillic scripts. It allows handling of data and text expressed in Bulgarian, Byelorussian, English, Macedonian, Russian, Serbocroatian and Ukrainian.

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

2 Conformance

A set of graphic characters is in conformance with this part of ISO/IEC 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/IEC 8859.

3 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this part of ISO/IEC 8859. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this part of ISO/IEC 8859 are encouraged to investigate the possibility of applying the most recent editions of the standards listed below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO 2022 : 1986, *Information processing — ISO 7-bit and 8-bit coded character sets — Code extension techniques.*

ISO 4873 : 1986, *Information processing — ISO 8-bit code for information interchange — Structure and rules for implementation.*

4 Definitions

For the purposes of this part of ISO/IEC 8859, the following definitions apply.

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

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

4.3 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.

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

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

NOTE — In this part of ISO/IEC 8859 a single bit combination is used to represent each character.

4.6 graphic symbol: A visual representation of a graphic character.

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

5 Notation, code table and names

5.1 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/IEC 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.

5.2 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.

5.3 Names and meanings

This part of ISO/IEC 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/IEC 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.

5.3.1 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.

5.3.2 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.

5.3.3 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 is permitted in the text as presented.

6 Specification of the coded character set

This part of ISO/IEC 8859 specifies 191 characters allocated to the bit combinations of the code table.

6.1 Characters of the set and their coded representation

Table 1 — Character set — Coded representation

Bit combination	Name
02/00	SPACE
02/01	EXCLAMATION MARK
02/02	QUOTATION MARK
02/03	NUMBER SIGN
02/04	DOLLAR SIGN
02/05	PERCENT SIGN
02/06	AMPERSAND
02/07	APOSTROPHE
02/08	LEFT PARENTHESIS
02/09	RIGHT PARENTHESIS
02/10	ASTERISK
02/11	PLUS SIGN
02/12	COMMA
02/13	HYPHEN, MINUS SIGN
02/14	FULL STOP
02/15	SOLIDUS
03/00	DIGIT ZERO
03/01	DIGIT ONE
03/02	DIGIT TWO
03/03	DIGIT THREE
03/04	DIGIT FOUR
03/05	DIGIT FIVE
03/06	DIGIT SIX
03/07	DIGIT SEVEN
03/08	DIGIT EIGHT
03/09	DIGIT NINE
03/10	COLON
03/11	SEMICOLON
03/12	LESS-THAN SIGN
03/13	EQUALS SIGN
03/14	GREATER-THAN SIGN
03/15	QUESTION MARK
04/00	COMMERCIAL AT
04/01	CAPITAL LETTER A
04/02	CAPITAL LETTER B
04/03	CAPITAL LETTER C
04/04	CAPITAL LETTER D
04/05	CAPITAL LETTER E
04/06	CAPITAL LETTER F
04/07	CAPITAL LETTER G
04/08	CAPITAL LETTER H
04/09	CAPITAL LETTER I
04/10	CAPITAL LETTER J
04/11	CAPITAL LETTER K
04/12	CAPITAL LETTER L

Table 1 (*continued*)

Bit combination	Name	Bit combination	Name
04/13	CAPITAL LETTER M	10/07	UKRAINIAN CYRILLIC CAPITAL LETTER YI
04/14	CAPITAL LETTER N	10/08	CYRILLIC CAPITAL LETTER JE
04/15	CAPITAL LETTER O	10/09	CYRILLIC CAPITAL LETTER LJE
05/00	CAPITAL LETTER P	10/10	CYRILLIC CAPITAL LETTER NJE
05/01	CAPITAL LETTER Q	10/11	SERBOCROATIAN CYRILLIC CAPITAL LETTER TSHE
05/02	CAPITAL LETTER R	10/12	MACEDONIAN CYRILLIC CAPITAL LETTER KJE
05/03	CAPITAL LETTER S		
05/04	CAPITAL LETTER T	10/13	SOFT HYPHEN
05/05	CAPITAL LETTER U	10/14	BYELORUSSIAN CYRILLIC CAPITAL LETTER SHORT U
05/06	CAPITAL LETTER V		
05/07	CAPITAL LETTER W	10/15	CYRILLIC CAPITAL LETTER DZHE
05/08	CAPITAL LETTER X	11/00	CYRILLIC CAPITAL LETTER A
05/09	CAPITAL LETTER Y	11/01	CYRILLIC CAPITAL LETTER BE
05/10	CAPITAL LETTER Z	11/02	CYRILLIC CAPITAL LETTER VE
05/11	LEFT SQUARE BRACKET	11/03	CYRILLIC CAPITAL LETTER GHE
05/12	REVERSE SOLIDUS	11/04	CYRILLIC CAPITAL LETTER DE
05/13	RIGHT SQUARE BRACKET	11/05	CYRILLIC CAPITAL LETTER IE
05/14	CIRCUMFLEX ACCENT	11/06	CYRILLIC CAPITAL LETTER ZHE
05/15	LOW LINE	11/07	CYRILLIC CAPITAL LETTER ZE
06/00	GRAVE ACCENT	11/08	CYRILLIC CAPITAL LETTER I
06/01	SMALL LETTER a	11/09	CYRILLIC CAPITAL LETTER SHORT I
06/02	SMALL LETTER b	11/10	CYRILLIC CAPITAL LETTER KA
06/03	SMALL LETTER c	11/11	CYRILLIC CAPITAL LETTER EL
06/04	SMALL LETTER d	11/12	CYRILLIC CAPITAL LETTER EM
06/05	SMALL LETTER e	11/13	CYRILLIC CAPITAL LETTER EN
06/06	SMALL LETTER f	11/14	CYRILLIC CAPITAL LETTER O
06/07	SMALL LETTER g	11/15	CYRILLIC CAPITAL LETTER PE
06/08	SMALL LETTER h	12/00	CYRILLIC CAPITAL LETTER ER
06/09	SMALL LETTER i	12/01	CYRILLIC CAPITAL LETTER ES
06/10	SMALL LETTER j	12/02	CYRILLIC CAPITAL LETTER TE
06/11	SMALL LETTER k	12/03	CYRILLIC CAPITAL LETTER U
06/12	SMALL LETTER l	12/04	CYRILLIC CAPITAL LETTER EF
06/13	SMALL LETTER m	12/05	CYRILLIC CAPITAL LETTER HA
06/14	SMALL LETTER n	12/06	CYRILLIC CAPITAL LETTER TSE
06/15	SMALL LETTER o	12/07	CYRILLIC CAPITAL LETTER CHE
07/00	SMALL LETTER p	12/08	CYRILLIC CAPITAL LETTER SHA
07/01	SMALL LETTER q	12/09	CYRILLIC CAPITAL LETTER SHCHA
07/02	SMALL LETTER r	12/10	CYRILLIC CAPITAL HARD SIGN
07/03	SMALL LETTER s	12/11	CYRILLIC CAPITAL LETTER YERU
07/04	SMALL LETTER t	12/12	CYRILLIC CAPITAL SOFT SIGN
07/05	SMALL LETTER u	12/13	CYRILLIC CAPITAL LETTER E
07/06	SMALL LETTER v	12/14	CYRILLIC CAPITAL LETTER YU
07/07	SMALL LETTER w	12/15	CYRILLIC CAPITAL LETTER YA
07/08	SMALL LETTER x	13/00	CYRILLIC SMALL LETTER A
07/09	SMALL LETTER y	13/01	CYRILLIC SMALL LETTER BE
07/10	SMALL LETTER z	13/02	CYRILLIC SMALL LETTER VE
07/11	LEFT CURLY BRACKET	13/03	CYRILLIC SMALL LETTER GHE
07/12	VERTICAL LINE	13/04	CYRILLIC SMALL LETTER DE
07/13	RIGHT CURLY BRACKET	13/05	CYRILLIC SMALL LETTER IE
07/14	TILDE	13/06	CYRILLIC SMALL LETTER ZHE
10/00	NO-BREAK SPACE	13/07	CYRILLIC SMALL LETTER ZE
10/01	CYRILLIC CAPITAL LETTER IO	13/08	CYRILLIC SMALL LETTER I
10/02	SERBOCROATIAN CYRILLIC CAPITAL LETTER DJE	13/09	CYRILLIC SMALL LETTER SHORT I
		13/10	CYRILLIC SMALL LETTER KA
10/03	MACEDONIAN CYRILLIC CAPITAL LETTER GJE	13/11	CYRILLIC SMALL LETTER EL
		13/12	CYRILLIC SMALL LETTER EM
10/04	UKRAINIAN CYRILLIC CAPITAL LETTER IE	13/13	CYRILLIC SMALL LETTER EN
		13/14	CYRILLIC SMALL LETTER O
10/05	MACEDONIAN CYRILLIC CAPITAL LETTER DZE	13/15	CYRILLIC SMALL LETTER PE
10/06	BYELORUSSIAN/UKRAINIAN CYRILLIC CAPITAL LETTER I		

Table 1 (*concluded*)

Bit combination	Name
14/00	CYRILLIC SMALL LETTER ER
14/01	CYRILLIC SMALL LETTER ES
14/02	CYRILLIC SMALL LETTER TE
14/03	CYRILLIC SMALL LETTER U
14/04	CYRILLIC SMALL LETTER EF
14/05	CYRILLIC SMALL LETTER HA
14/06	CYRILLIC SMALL LETTER TSE
14/07	CYRILLIC SMALL LETTER CHE
14/08	CYRILLIC SMALL LETTER SHA
14/09	CYRILLIC SMALL LETTER SHCHA
14/10	CYRILLIC SMALL HARD SIGN
14/11	CYRILLIC SMALL LETTER YERU
14/12	CYRILLIC SMALL SOFT SIGN
14/13	CYRILLIC SMALL LETTER E
14/14	CYRILLIC SMALL LETTER YU
14/15	CYRILLIC SMALL LETTER YA
15/00	NUMBER ACRONYM
15/01	CYRILLIC SMALL LETTER IO
15/02	SERBOCROATIAN CYRILLIC SMALL LETTER DJE
15/03	MACEDONIAN CYRILLIC SMALL LETTER GJE
15/04	UKRAINIAN CYRILLIC SMALL LETTER IE
15/05	MACEDONIAN CYRILLIC SMALL LETTER DZE
15/06	BYELORUSSIAN/UKRAINIAN CYRILLIC SMALL LETTER I
15/07	UKRAINIAN CYRILLIC SMALL LETTER YI
15/08	CYRILLIC SMALL LETTER JE
15/09	CYRILLIC SMALL LETTER LJE
15/10	CYRILLIC SMALL LETTER NJE
15/11	SERBOCROATIAN CYRILLIC SMALL LETTER TSHE
15/12	MACEDONIAN CYRILLIC SMALL LETTER KJE
15/13	PARAGRAPH SIGN
15/14	BYELORUSSIAN CYRILLIC SMALL LETTER SHORT U
15/15	CYRILLIC SMALL LETTER DZHE

6.2 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 this part of ISO/IEC 8859; it is specified in other International Standards, for example ISO 4873.

7 Designation of the character set

The graphic characters of this part of ISO/IEC 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/IEC 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 and ISO 4873, the following pair of escape sequences shall be used:

ESC 02/08 04/02

ESC 02/13 04/12

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

4

Table 2 — Code table of the Latin/Cyrillic alphabet

b8				0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
b7				0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
b6				0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
b5				0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
b4	b3	b2	b1	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
0	0	0	0	00			SP	0	@	P	`	p			NBSP	А	Р	а	р	№
0	0	0	1	01			!	1	A	Q	a	q			Ё	Б	С	б	с	ё
0	0	1	0	02			"	2	B	R	b	r			Ђ	В	Т	в	т	ђ
0	0	1	1	03			#	3	C	S	c	s			Ѓ	Г	У	г	у	ѓ
0	1	0	0	04			$	4	D	T	d	t			Є	Д	Ф	д	ф	є
0	1	0	1	05			%	5	E	U	e	u			Ѕ	Е	Х	е	х	ѕ
0	1	1	0	06			&	6	F	V	f	v			І	Ж	Ц	ж	ц	і
0	1	1	1	07			'	7	G	W	g	w			Ї	З	Ч	з	ч	ї
1	0	0	0	08			(	8	H	X	h	x			Ј	И	Ш	и	ш	ј
1	0	0	1	09			)	9	I	Y	i	y			Љ	Й	Щ	й	щ	љ
1	0	1	0	10			*	:	J	Z	j	z			Њ	К	Ъ	к	ъ	њ
1	0	1	1	11			+	;	K	[	k	{			Ћ	Л	Ы	л	ы	ћ
1	1	0	0	12			,	<	L	\	l	|			Ќ	М	Ь	м	ь	ќ
1	1	0	1	13			-	=	M	]	m	}			SHY	Н	Э	н	э	§
1	1	1	0	14			.	>	N	^	n	~			Ў	О	Ю	о	ю	ў
1	1	1	1	15			/	?	O	_	o				Џ	П	Я	п	я	џ