

BRITISH STANDARD

Coded character sets for text communication

Part 2. Specification for Latin alphabetic and non-alphabetic graphic characters

BS 6692 :
Part 2 : 1990
ISO 6937/2 :
1983
(Including
Addendum 1 : 1989)

UDC 681.3.048 : 003

Jeux de caractères codés pour la transmission de texte
Partie 2. Caractères graphiques alphabétiques latins et caractères graphiques non alphabétiques - Spécifications

Verschlüsselte Zeichensätze für Textkommunikation
Teil 2. Lateinische und nicht-alphabetische Schriftzeichen

Committees responsible for this British Standard

The preparation of this British Standard was entrusted by the Information Systems Technology Standards Policy Committee (IST/-) to Technical Committee IST/2, upon which the following bodies were represented:

British Federation of Printing Machinery and Supplies Ltd
 British Radio and Electronic Equipment Manufacturers' Association
 British Telecommunications plc
 Business Equipment Association
 Computing Services and Information Technology Association
 Institution of Electrical Engineers
 International Press Telecommunications Council
 Videotex Industry Association

This British Standard, having been prepared under the direction of the Information Systems Technology Standards Policy Committee, was published under the authority of the Board of BSI and comes into effect on 29 June 1990

© BSI 1990
 First published February 1986
 Second edition June 1990

The following BSI references relate to the work on this standard:
 Committee reference IST/2
 Draft for comment 88/64781 DC

ISBN 0 580 18347 5

Amendments issued since publication

Amd. No.	Date	Text affected

Contents

	Page
Committees responsible	Inside front cover
National foreword	ii
Specification	
0 Introduction	1
1 Scope and field of application	1
2 References	2
3 Definitions	2
4 Character repertoire and coded representations	2
5 Graphic character subrepertoires	15
Annexes	
A Standard subrepertoires	17
B Use of non-spacing diacritical marks	27
C Use of the non-spacing underline character	29
D Use of Latin alphabetic characters	30
E Fall-back presentations	34
F Differences from CCITT Recommendations S.61 and S.100	36
Addendum	
1	38
Tables	
1 Primary set of graphic characters for text communication	3
2 Supplementary set of graphic characters for text communication	4
3 Primary and supplementary sets of graphic characters for text communication	5
4 Latin alphabetic characters	7
5 Decimal digits	12
6 Currency signs	12
7 Punctuation marks	13
8 Arithmetic signs	13
9 Superscripts	13
10 Fractions	14
11 Miscellaneous symbols	14
12 Diacritical marks as separate graphic characters	15
13 Latin alphabetic characters	18
14 Decimal digits	23
15 Currency signs	23
16 Punctuation marks	24
17 Arithmetic signs	24
18 Superscripts	24
19 Fractions	25
20 Miscellaneous symbols	25
21 Diacritical marks as separate graphic characters	26
22 Summary of non-spacing diacritical marks	27
23 Combinations of diacritical marks and basic letters	28
24 Use of Latin alphabetic characters	31

National foreword

This new edition of Part 2 of BS 6692 been prepared under the direction of the Information Systems Technology Standards Policy Committee. It is identical with ISO 6937/2 : 1983 'Information processing - Coded character sets for text communication - Part 2: Latin alphabetic and non-alphabetic graphic characters' and 'Addendum No. 1 : 1989 to ISO 6937-2 : 1983', published by the International Organization for Standardization (ISO). Addendum No. 1 is included in this Part of BS 6692. It introduces two graphic characters, 'no break space' and 'soft hyphen', and also corrects printing errors.

This British Standard consists of two Parts:

Part 1	General introduction
Part 2	Specification for Latin alphabetic and non-alphabetic graphic characters

Part 1 is identical with ISO 6937/1 : 1983 'Information processing - Coded character sets for text communication - Part 1: General introduction'. Other parts of ISO 6937 were in preparation, but in autumn 1988 the international committee decided that future work on ISO 6937 should be combined with work on ISO 8859 '8-bit single byte coded graphic character sets', to produce a new multi-part standard. ISO 8859-1 and ISO 8859-2 are currently available as BS 7203 : Part 1 : 1989 and BS 7203 : Part 2 : 1989.

Cross-references

International standard	Corresponding British Standard
ISO 646 : 1981	BS 4730 : 1985 Specification for UK 7-bit coded character set (Identical exercising national options)
ISO 2022 : 1986	BS 6856 : 1987 Specification for code extension techniques for United Kingdom 7-bit and 8-bit coded character sets (Identical)
ISO 6429 : 1988	BS 7204 : 1989 Specification for control functions for ISO 7-bit and 8-bit coded character sets (Identical)

Compliance with a British Standard does not of itself confer immunity from legal obligations.

Information processing — Coded character sets for text communication —

Part 2: Latin alphabetic and non-alphabetic graphic characters

0 Introduction

This International Standard specifies repertoires of graphic characters and control functions, and their coded representations, for use in text communication. It applies to the communication of text in the form of binary-coded representations of graphic characters and control functions, using

- a) public communication networks;
- b) private communication networks;
- c) interchange media such as magnetic tapes and discs.

It is applicable to the exchange of text at the coding interface.

Although, in general, text consists of characters and pictures, this International Standard applies only to text made up of characters.

This International Standard consists, at present, of three parts, as follows:

ISO 6937/1, General introduction.

ISO 6937/2, Latin alphabetic and non-alphabetic graphic characters.

ISO 6937/3, Control functions for page-image format (in preparation).

Other parts may be added later.

This part of ISO 6937 and ISO 6937/3 may be used either in combination with one another or individually, but they should always be used in combination with ISO 6429.

The specifications are based on the 7-bit coded character set specified in ISO 646, the 7-bit and 8-bit code extension techniques specified in ISO 2022, and the definitions of additional control functions given in ISO 6429.

- 1) ISO 6937/1, ISO 6937/2 and ISO 6937/3 have been developed in parallel with CCITT Recommendations S.61, *Character*

repertoire and coded character sets for the international Teletex service (Geneva, 1980), and S.100, *International information exchange for interactive Videotex* (Geneva, 1980). The graphic character repertoire defined in this part of ISO 6937 is a superset of the graphic character repertoires defined in CCITT Recommendations S.61 and S.100, and the control function repertoire to be defined in ISO 6937/3 is a superset of the control function repertoire defined in CCITT Recommendation S.61. However, the primary and supplementary sets of graphic characters specified in this International Standard are not identical to those defined in CCITT Recommendations S.61 and S.100 (see annex F). To ensure compatible interworking between the public text communication services defined by the CCITT and terminal equipment communicating text conforming to the requirements of this International Standard, special attention should be given to the differences between this International Standard and CCITT Recommendations S.61 and S.100.

1 Scope and field of application

1)

This part of ISO 6937

- a) defines a repertoire of Latin alphabetic and non-alphabetic characters for the communication of text in European languages;
- b) specifies coded representations for the graphic characters;
- c) specifies rules for the definition and use of graphic character subrepertoires, i.e. subsets of the defined character repertoire.

It also gives, in annexes,

- a) definitions of three standard subrepertoires (annex A);
- b) a summary of the use of non-spacing diacritical marks in combination with the letters of the basic Latin alphabet (annex B);
- c) a description of the use of the non-spacing underline character (annex C);

1) See Addendum 1.

ISO 6937/2-1983 (E)

- d) a summary of the use of latin alphabetic characters (annex D);
- e) recommendations for fall-back presentations (annex E);
- 1) f) a summary of the differences from CCITT Recommendations S.61 and S.100 (annex F).

NOTE — This part of ISO 6937 does not specify graphic symbols; symbols are shown for the purpose of illustration only.

This part of ISO 6937 applies to text as communicated at the coding interface of terminal equipment. It does not necessarily apply to the coded character sets used internally between the coding interface and the input-output devices of the terminal equipment.

2 References

ISO 646, *Information processing — 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 6429, *Information processing — ISO 7-bit and 8-bit coded character sets — Additional control functions for character-imaging devices*.

1) 3 Definitions

For the purpose of this part of ISO 6937, the definitions given in ISO 6937/1 apply.

4 Character repertoire and coded representations

4.1 Composition of the character repertoire

The repertoire of graphic characters defined in this part of ISO 6937 consists of

- a) Latin alphabetic characters comprising
 - 1) the 52 small and capital letters of the basic Latin alphabet,
 - 2) accented letters and umlauts, the graphic representations of which consist of combinations of basic Latin letters and diacritical marks,
 - 3) special alphabetic characters which are neither basic Latin letters nor combinations of basic Latin letters and diacritical marks;
- b) non-alphabetic characters comprising decimal digits, punctuation marks, arithmetic signs and various other characters that have individual special meanings.

The repertoire is specified by the first three columns of tables 4 to 12. In each table entry, the first column contains the identifier of the character concerned, the second column shows the typical graphic representation, and the third column specifies the name or description of the character. The remaining columns specify coded representations (see 4.4). The identifiers have been assigned in accordance with the method of identification described in ISO 6937/1, annex A.

NOTES

- 1 All graphic characters of the repertoire cause an implicit forward movement of the active position (a definition of active position will be included in ISO 6937/3).
- 2 A survey of the use of Latin characters in various languages is included in annex D.

4.2 Character sets

1)

The coded representations of the graphic characters of the repertoire defined in this part of ISO 6937 make use of two character sets, that is a "primary set" and a "supplementary set".

The primary set consists of the graphic characters of the International Reference Version of the 7-bit coded character set specified in ISO 646. The differences between the primary set and the set of graphic characters of the International Reference Version are

- a) unique names have been assigned to the characters "low line" and "tilde", where ISO 646 specifies dual names;
- b) the characters of the primary set are not used in combination with each other to generate graphic characters of the repertoire defined in this part of ISO 6937.

The primary set contains the letters of the basic Latin alphabet, some spacing diacritical marks and a number of non-alphabetic characters.

The supplementary set contains spacing and non-spacing graphic characters. The non-spacing graphic characters consist of a number of non-spacing diacritical marks and the non-spacing underline character. The spacing graphic characters consist of a number of special alphabetic and non-alphabetic characters which are used in addition to those in the primary set.

A non-spacing diacritical mark shall be used only in combination with either SPACE or certain basic Latin letters. The allowed combinations of diacritical marks and letters are the ones needed to represent the accented letters and umlauts included in table 4. This set of combinations is summarized in annex B.

The non-spacing underline character can be used in combination with any graphic character of the repertoire, including an accented letter or an umlaut, or SPACE. The use of this character is specified in annex C.

Code tables for the primary set and the supplementary set of graphic characters are given in tables 1, 2, and 3. Shaded positions denote bit combinations which are not part of the sets concerned.

The escape sequences used to designate the primary and the supplementary set are:

ESC 2/8 F¹⁾: to designate the primary set as the G0 set;

ESC 2/9 F¹⁾: to designate the supplementary set as the G1 set;

ESC 2/10 F¹⁾: to designate the supplementary set as the G2 set;

ESC 2/11 F¹⁾: to designate the supplementary set as the G3 set.

Table 1 — Primary set of graphic characters for text communication [coding when represented by bit combinations 2/1 to 7/14 of a 7-bit or 8-bit code (in an 8-bit code, the additional bit $b_8 = 0$)]

				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	`	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			*	:	J	Z	j	z
1	0	1	1	11			+	;	K	[	k	{
1	1	0	0	12			,	<	L	\	l	
1	1	0	1	13			-	=	M	]	m	}
1	1	1	0	14			.	>	N	^	n	~
1	1	1	1	15			/	?	0	_	o	

1) Final bit combinations will be assigned when the sets are registered.

ISO 6937/2-1983 (E)

- 2) Table 2 — Supplementary set of graphic characters for text communication [coding when represented by bit combinations 2/1 to 7/14 of a 7-bit or 8-bit code (in an 8-bit code, the additional bit $b_8 = 0$)]

					b_7					b_6					b_5					b_4					b_3					b_2					b_1				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1					0					1					0					1					0				
					0					1																													

- 2) **Table 3 — Primary and supplementary sets of graphic characters for text communication** (coding when represented by bit combinations 2/1 to 7/14 and 10/1 to 15/14 of an 8-bit code)

					b ₄	b ₃	b ₂	b ₁	00	01	02	03	04	05	06	07	08	09	10	11	12 ¹⁾	13	14	15
b ₄	b ₃	b ₂	b ₁		00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19
0	0	0	0	00					0	à	P	'	p					°	See 4.3.4	—	Ω	K		
0	0	0	1	01					!	1	A	Q	a	q				i	±	`	¹	Æ	æ	
0	0	1	0	02					"	2	B	R	b	r				¢	²	'	®	Ð	ð	
0	0	1	1	03					#	3	C	S	c	s				£	³	^	©	ä	ö	
0	1	0	0	04					¤	4	D	T	d	t				\$	×	~	™	H	h	
0	1	0	1	05					%	5	E	U	e	u				¥	μ	¯	♪	See 4.3.4	ı	
0	1	1	0	06					&	6	F	V	f	v				See 4.3.3	¶	˘	See 4.3.4	IJ	ij	
0	1	1	1	07					'	7	G	W	g	w				§	·	˙	See 4.3.4	Ł	ł	
1	0	0	0	08					(	8	H	X	h	x				See 4.3.3	÷	"	See 4.3.4	Ł	ł	
1	0	0	1	09					)	9	I	Y	i	y				‘	’	See 4.3.4	See 4.3.4	Ø	ø	
1	0	1	0	10					*	:	J	Z	j	z				“	”	°	See 4.3.4	Œ	œ	
1	0	1	1	11					+	;	K	Ł	k	ł				«	»	ˆ	See 4.3.4	Œ	œ	
1	1	0	0	12					,	<	L	\	l	l				←	¼	—	⅛	p	p	
1	1	0	1	13					-	=	M	]	m	ı				↑	½	"	⅜	ƒ	ƒ	
1	1	1	0	14					.	>	N	^	n	~				→	¾	˘	⅝	n	n	
1	1	1	1	15					/	?	O	_	o					↓	¿	˘	⅞	'n		

1) See 4.3.1

2) See Addendum 1.

ISO 6937/2-1983 (E)

4.3 Explanations concerning the code tables

4.3.1 The non-spacing diacritical marks represented by bit combinations 4/1 to 4/8, 4/10, 4/11 and 4/13 to 4/15 of the supplementary set (12/1 to 12/8, 12/10, 12/11 and 12/13 to 12/15 in table 3) are used only in combination with either SPACE or certain basic Latin letters (see annex B). The non-spacing underline character represented by bit combination 4/12 of the supplementary set (12/12 in table 3) can be used in combination with any graphic character of the repertoire, including an accented letter or an umlaut, or SPACE (see annex C).

1) **4.3.2** In CCITT Recommendation S.100 (Geneva, 1980), the characters # (number sign) and ¤ (general currency sign) have been allocated to bit combinations 2/6 and 2/8 of the supplementary set, and not to 2/3 and 2/4 of the primary set. In order to facilitate interworking with Videotex, bit combinations 2/6 and 2/8 of the supplementary set shall also be considered as valid coded representations of the characters # and ¤, but their use is deprecated.

1) **4.3.3** In CCITT Recommendation S.61 (Geneva, 1980), the characters # (number sign) and ¤ (general currency sign) have been allocated to bit combinations 10/6 and 10/8, and not 2/3 and 2/4. In order to facilitate interworking with Teletex, bit combinations 10/6 and 10/8 shall also be considered as valid coded representations of the characters # and ¤, but their use is deprecated.

1) **4.3.4** Bit combinations 4/0, 4/9, 5/6 to 5/11, and 6/5 of the supplementary set (12/0, 12/9, 13/6 to 13/11, and 14/5 in table 3) are reserved for future standardization, i.e. possible allocation of additional graphic characters. Bit combinations 4/0 and 4/9 (12/0 and 12/9 in table 3) are reserved for possible allocation of additional non-spacing characters.

NOTES

1 In certain interactive Videotex systems, the character represented by bit combination 5/15 of the primary set is used as a delimiter, the graphic representation of which may be different from "low line".

2 In certain systems supporting the International Reference Version of ISO 646, the character represented by bit combination 7/14 of the primary set may have the graphic representation of an "overline".

1) 4.4 Coded representations

The coded representations of the graphic characters of the repertoire defined in this part of ISO 6937 are specified in tables 4 to 12. The formats of the coded representations are as follows.

a) Accented letters and umlauts:

Each accented letter or umlaut is represented by a sequence of bit combinations consisting of the coded representation of the relevant non-spacing diacritical mark (an element of the supplementary set), followed by the coded representation of the relevant basic Latin letter (an element of the primary set).

b) Diacritical marks as separate graphic characters.

Diacritical marks as separate graphic characters that are elements of the primary set (grave accent, circumflex accent and elements of the primary set (grave accent, circumflex accent and tilde) are represented by the corresponding single bit combination in the range 2/1 to 7/14.

Each diacritical mark as a separate graphic character that is not an element of the primary set is represented by a sequence of bit combinations consisting of the coded representation of the relevant non-spacing diacritical mark (an element of the supplementary set), followed by the coded representation of the character SPACE, i.e. the bit combination 2/0.

The latter method of representation may also be used for the grave accent, circumflex accent and tilde, but that use is deprecated.

c) All other graphic characters of the repertoire.

Any graphic character of the repertoire, other than an accented letter, an umlaut or a diacritical mark as a separate graphic character that is not an element of the primary set, is an element of either the primary set or the supplementary set and is represented by the corresponding single bit combination in the range 2/1 to 7/14 or 10/1 to 15/14.

Depending on the code extension techniques used, a bit combination, representing an element of the primary or the supplementary set may have to be preceded by a code extension function invoking the character set concerned.

In tables 4 to 12, the coded representation column "A" specifies the coded representation of each character when the elements of the supplementary set are represented by bit combinations in the range 2/1 to 7/14 of a 7-bit or 8-bit code. Coded representation column "B" applies when the elements of the supplementary set are represented by bit combinations in the range 10/1 to 15/14 of an 8-bit code. It is assumed that, in either case, the elements of the primary set are represented by bit combinations in the range 2/1 to 7/14. The symbol "S" is used in column "A" to identify a bit combination representing an element of the supplementary set.

Table 4 — Latin alphabetic characters

ID	Graphic symbol	Name or description	Coded representation	
			A	B
LA01	a	small a	6/1	6/1
LA02	A	capital A	4/1	4/1
LA11	á	small a with acute accent	S 4/2 6/1	12/2 6/1
LA12	Á	capital A with acute accent	S 4/2 4/1	12/2 4/1
LA13	à	small a with grave accent	S 4/1 6/1	12/1 6/1
LA14	À	capital A with grave accent	S 4/1 4/1	12/1 4/1
LA15	â	small a with circumflex accent	S 4/3 6/1	12/3 6/1
LA16	Â	capital A with circumflex accent	S 4/3 4/1	12/3 4/1
LA17	ä	small a with diaeresis or umlaut mark	S 4/8 6/1	12/8 6/1
LA18	Ä	capital A with diaeresis or umlaut mark	S 4/8 4/1	12/8 4/1
LA19	ã	small a with tilde	S 4/4 6/1	12/4 6/1
LA20	Ã	capital A with tilde	S 4/4 4/1	12/4 4/1
LA23	ă	small a with breve	S 4/6 6/1	12/6 6/1
LA24	Ă	capital A with breve	S 4/6 4/1	12/6 4/1
LA27	ā	small a with ring	S 4/10 6/1	12/10 6/1
LA28	Ā	capital A with ring	S 4/10 4/1	12/10 4/1
LA31	ā	small a with macron	S 4/5 6/1	12/5 6/1
LA32	Ā	capital A with macron	S 4/5 4/1	12/5 4/1
LA43	ą	small a with ogonek	S 4/14 6/1	12/14 6/1
LA44	Ą	capital A with ogonek	S 4/14 4/1	12/14 4/1
LA51	æ	small æ diphthong	S 7/1	15/1
LA52	Æ	capital Æ diphthong	S 6/1	14/1
LB01	b	small b	6/2	6/2
LB02	B	capital B	4/2	4/2
LC01	c	small c	6/3	6/3
LC02	C	capital C	4/3	4/3
LC11	ć	small c with acute accent	S 4/2 6/3	12/2 6/3
LC12	Ć	capital C with acute accent	S 4/2 4/3	12/2 4/3
LC15	ĉ	small c with circumflex accent	S 4/3 6/3	12/3 6/3
LC16	Ĉ	capital C with circumflex accent	S 4/3 4/3	12/3 4/3
LC21	č	small c with caron	S 4/15 6/3	12/15 6/3
LC22	Č	capital C with caron	S 4/15 4/3	12/15 4/3
LC29	ċ	small c with dot above	S 4/7 6/3	12/7 6/3
LC30	Ċ	capital C with dot above	S 4/7 4/3	12/7 4/3
LC41	ç	small c with cedilla	S 4/11 6/3	12/11 6/3
LC42	Ç	capital C with cedilla	S 4/11 4/3	12/11 4/3
LD01	d	small d	6/4	6/4
LD02	D	capital D	4/4	4/4
LD21	ď or d'	small d with caron	S 4/15 6/4	12/15 6/4
LD22	Ď	capital D with caron	S 4/15 4/4	12/15 4/4
LD61	ḏ	small d with stroke	S 7/2	15/2
LD62	Ð	capital D with stroke, Icelandic eth	S 6/2	14/2
LD63	ð	small eth, Icelandic	S 7/3	15/3

ISO 6937/2-1983 (E)

Table 4 — Latin alphabetic characters (continued)

ID	Graphic symbol	Name or description	Coded representation	
			A	B
LE01	e	small e	6/5	6/5
LE02	E	capital E	4/5	4/5
LE11	é	small e with acute accent	S 4/2 6/5	12/2 6/5
LE12	É	capital E with acute accent	S 4/2 4/5	12/2 4/5
LE13	è	small e with grave accent	S 4/1 6/5	12/1 6/5
LE14	È	capital E with grave accent	S 4/1 4/5	12/1 4/5
LE15	ê	small e with circumflex accent	S 4/3 6/5	12/3 6/5
LE16	Ê	capital E with circumflex accent	S 4/3 4/5	12/3 4/5
LE17	ë	small e with diaeresis or umlaut mark	S 4/8 6/5	12/8 6/5
LE18	Ë	capital E with diaeresis or umlaut mark	S 4/8 4/5	12/8 4/5
LE21	ě	small e with caron	S 4/15 6/5	12/15 6/5
LE22	Ě	capital E with caron	S 4/15 4/5	12/15 4/5
LE29	ě	small e with dot above	S 4/7 6/5	12/7 6/5
LE30	Ě	capital E with dot above	S 4/7 4/5	12/7 4/5
LE31	ē	small e with macron	S 4/5 6/5	12/5 6/5
LE32	Ē	capital E with macron	S 4/5 4/5	12/5 4/5
LE43	ę	small e with ogonek	S 4/14 6/5	12/14 6/5
LE44	Ę	capital E with ogonek	S 4/14 4/5	12/14 4/5
LF01	f	small f	6/6	6/6
LF02	F	capital F	4/6	4/6
LG01	g	small g	6/7	6/7
LG02	G	capital G	4/7	4/7
LG11	g	small g with acute accent	S 4/2 6/7	12/2 6/7
LG15	ĝ	small g with circumflex accent	S 4/3 6/7	12/3 6/7
LG16	Ĝ	capital G with circumflex accent	S 4/3 4/7	12/3 4/7
LG23	ğ	small g with breve	S 4/6 6/7	12/6 6/7
LG24	Ğ	capital G with breve	S 4/6 4/7	12/6 4/7
LG29	ġ	small g with dot above	S 4/7 6/7	12/7 6/7
LG30	Ġ	capital G with dot above	S 4/7 4/7	12/7 4/7
LG42	Ġ	capital G with cedilla	S 4/11 4/7	12/11 4/7
LH01	h	small h	6/8	6/8
LH02	H	capital H	4/8	4/8
LH15	ĥ	small h with circumflex accent	S 4/3 6/8	12/3 6/8
LH16	Ĥ	capital H with circumflex accent	S 4/3 4/8	12/3 4/8
LH61	ħ	small h with stroke	S 7/4	15/4
LH62	Ħ	capital H with stroke	S 6/4	14/4
LI01	i	small i	6/9	6/9
LI02	I	capital I	4/9	4/9
LI11	í	small i with acute accent	S 4/2 6/9	12/2 6/9
LI12	Í	capital I with acute accent	S 4/2 4/9	12/2 4/9
LI13	ì	small i with grave accent	S 4/1 6/9	12/1 6/9
LI14	Ì	capital I with grave accent	S 4/1 4/9	12/1 4/9

Table 4 — Latin alphabetic characters (continued)

ID	Graphic symbol	Name or description	Coded representation	
			A	B
LI15	í	small i with circumflex accent	S 4/3 6/9	12/3 6/9
LI16	Î	capital I with circumflex accent	S 4/3 4/9	12/3 4/9
LI17	ï	small i with diaeresis	S 4/8 6/9	12/8 6/9
LI18	Ï	capital I with diaeresis	S 4/8 4/9	12/8 4/9
LI19	ï	small i with tilde	S 4/4 6/9	12/4 6/9
LI20	Ï	capital I with tilde	S 4/4 4/9	12/4 4/9
LI30	İ	capital I with dot above	S 4/7 4/9	12/7 4/9
LI31	ĩ	small i with macron	S 4/5 6/9	12/5 6/9
LI32	Î	capital I with macron	S 4/5 4/9	12/5 4/9
LI43	ĩ	small i with ogonek	S 4/14 6/9	12/14 6/9
LI44	Î	capital I with ogonek	S 4/14 4/9	12/14 4/9
LI51	ij	small ij ligature	S 7/6	15/6
LI52	IJ	capital IJ ligature	S 6/6	14/6
LI61	ı	small i without dot	S 7/5	15/5
LJ01	j	small j	6/10	6/10
LJ02	J	capital J	4/10	4/10
LJ15	ĵ	small j with circumflex accent	S 4/3 6/10	12/3 6/10
LJ16	Ĵ	capital J with circumflex accent	S 4/3 4/10	12/3 4/10
LK01	k	small k	6/11	6/11
LK02	K	capital K	4/11	4/11
LK41	ķ	small k with cedilla	S 4/11 6/11	12/11 6/11
LK42	Ķ	capital K with cedilla	S 4/11 4/11	12/11 4/11
LK61	ƙ	small k, Greenlandic	S 7/0	15/0
LL01	l	small l	6/12	6/12
LL02	L	capital L	4/12	4/12
LL11	ĺ	small l with acute accent	S 4/2 6/12	12/2 6/12
LL12	Ĺ	capital L with acute accent	S 4/2 4/12	12/2 4/12
LL21	Ľ or Ľ	small l with caron	S 4/15 6/12	12/15 6/12
LL22	Ľ or Ľ	capital L with caron	S 4/15 4/12	12/15 4/12
LL41	ł	small l with cedilla	S 4/11 6/12	12/11 6/12
LL42	Ł	capital L with cedilla	S 4/11 4/12	12/11 4/12
LL61	ł	small l with stroke	S 7/8	15/8
LL62	Ł	capital L with stroke	S 6/8	14/8
LL63	ł̣	small l with middle dot	S 7/7	15/7
LL64	Ł̣	capital L with middle dot	S 6/7	14/7
LM01	m	small m	6/13	6/13
LM02	M	capital M	4/13	4/13
LN01	n	small n	6/14	6/14
LN02	N	capital N	4/14	4/14
LN11	ñ	small n with acute accent	S 4/2 6/14	12/2 6/14
LN12	Ñ	capital N with acute accent	S 4/2 4/14	12/2 4/14
LN19	ñ	small n with tilde	S 4/4 6/14	12/4 6/14
LN20	Ñ	capital N with tilde	S 4/4 4/14	12/4 4/14

Table 4 — Latin alphabetic characters (continued)

ID	Graphic symbol	Name or description	Coded representation	
			A	B
LN21	ñ	small n with caron	S 4/15 6/14	12/15 6/14
LN22	Ñ	capital N with caron	S 4/15 4/14	12/15 4/14
LN41	ŋ	small n with cedilla	S 4/11 6/14	12/11 6/14
LN42	Ñ	capital N with cedilla	S 4/11 4/14	12/11 4/14
LN61	ɳ	small eng, Lapp	S 7/14	15/14
LN62	ᐅ	capital ENG, Lapp	S 6/14	14/14
LN63	ˈn	small n with apostrophe	S 6/15	14/15
LO01	o	small o	6/15	6/15
LO02	O	capital O	4/15	4/15
LO11	ó	small o with acute accent	S 4/2 6/15	12/2 6/15
LO12	Ó	capital O with acute accent	S 4/2 4/15	12/2 4/15
LO13	ò	small o with grave accent	S 4/1 6/15	12/1 6/15
LO14	Ò	capital O with grave accent	S 4/1 4/15	12/1 4/15
LO15	ô	small o with circumflex accent	S 4/3 6/15	13/3 6/15
LO16	Ô	capital O with circumflex accent	S 4/3 4/15	12/3 4/15
LO17	ö	small o with diaeresis or umlaut mark	S 4/8 6/15	12/8 6/15
LO18	Ö	capital O with diaeresis or umlaut mark	S 4/8 4/15	12/8 4/15
LO19	ø	small o with tilde	S 4/4 6/15	12/4 6/15
LO20	Õ	capital O with tilde	S 4/4 4/15	12/4 4/15
LO25	ô	small o with double acute accent	S 4/13 6/15	12/13 6/15
LO26	Ö	capital O with double acute accent	S 4/13 4/15	12/13 4/15
LO31	ō	small o with macron	S 4/5 6/15	12/5 6/15
LO32	Ō	capital O with macron	S 4/5 4/15	12/5 4/15
LO51	œ	small œ ligature	S 7/10	15/10
LO52	Œ	capital Œ ligature	S 6/10	14/10
LO61	ø	small o with slash	S 7/9	15/9
LO62	Ø	capital O with slash	S 6/9	14/9
LP01	p	small p	7/0	7/0
LP02	P	capital P	5/0	5/0
LQ01	q	small q	7/1	7/1
LQ02	Q	capital Q	5/1	5/1
LR01	r	small r	7/2	7/2
LR02	R	capital R	5/2	5/2
LR11	ř	small r with acute accent	S 4/2 7/2	12/2 7/2
LR12	Ř	capital R with acute accent	S 4/2 5/2	12/2 5/2
LR21	ř	small r with caron	S 4/15 7/2	12/15 7/2
LR22	Ř	capital R with caron	S 4/15 5/2	12/15 5/2
LR41	ɽ	small r with cedilla	S 4/11 7/2	12/11 7/2
LR42	ᐃ	capital R with cedilla	S 4/11 5/2	12/11 5/2
LS01	s	small s	7/3	7/3
LS02	S	capital S	5/3	5/3
LS11	ś	small s with acute accent	S 4/2 7/3	12/2 7/3
LS12	Ś	capital S with acute accent	S 4/2 5/3	12/2 5/3

Table 4 — Latin alphabetic characters (continued)

ID	Graphic symbol	Name or description	Coded representation	
			A	B
LS15	š	small s with circumflex accent	S 4/3 7/3	12/3 7/3
LS16	Š	capital S with circumflex accent	S 4/3 5/3	12/3 5/3
LS21	š	small s with caron	S 4/15 7/3	12/15 7/3
LS22	Š	capital S with caron	S 4/15 5/3	12/15 5/3
LS41	ş	small s with cedilla	S 4/11 7/3	12/11 7/3
LS42	Ş	capital S with cedilla	S 4/11 5/3	12/11 5/3
LS61	ß	small sharp s, German	S 7/11	15/11
LT01	t	small t	7/4	7/4
LT02	T	capital T	5/4	5/4
LT21	ť or t'	small t with caron	S 4/15 7/4	12/15 7/4
LT22	Ť	capital T with caron	S 4/15 5/4	12/15 5/4
LT41	ţ	small t with cedilla	S 4/11 7/4	12/11 7/4
LT42	Ț	capital T with cedilla	S 4/11 5/4	12/11 5/4
LT61	ƚ	small t with stroke	S 7/13	15/13
LT62	ƚ	capital T with stroke	S 6/13	14/13
LT63	þ	small thorn, Icelandic	S 7/12	15/12
LT64	Þ	capital thorn, Icelandic	S 6/12	14/12
LU01	u	small u	7/5	7/5
LU02	U	capital U	5/5	5/5
LU11	ú	small u with acute accent	S 4/2 7/5	12/2 7/5
LU12	Ú	capital u with acute accent	S 4/2 5/5	12/2 5/5
LU13	ù	small u with grave accent	S 4/1 7/5	12/1 7/5
LU14	Ù	capital U with grave accent	S 4/1 5/5	12/1 5/5
LU15	û	small u with circumflex accent	S 4/3 7/5	12/3 7/5
LU16	Û	capital U with circumflex accent	S 4/3 5/5	12/3 5/5
LU17	ü	small u with diaeresis or umlaut mark	S 4/8 7/5	12/8 7/5
LU18	Ü	capital U with diaeresis or umlaut mark	S 4/8 5/5	12/8 5/5
LU19	ũ	small u with tilde	S 4/4 7/5	12/4 7/5
LU20	Ũ	capital U with tilde	S 4/4 5/5	12/4 5/5
LU23	ū	small u with breve	S 4/6 7/5	12/6 7/5
LU24	Ū	capital U with breve	S 4/6 5/5	12/6 5/5
LU25	ü	small u with double acute accent	S 4/13 7/5	12/13 7/5
LU26	Ü	capital U with double acute accent	S 4/13 5/5	12/13 5/5
LU27	ũ	small u with ring	S 4/10 7/5	12/10 7/5
LU28	Ů	capital U with ring	S 4/10 5/5	12/10 5/5
LU31	ū	small u with macron	S 4/5 7/5	12/5 7/5
LU32	Ū	capital U with macron	S 4/5 5/5	12/5 5/5
LU43	ų	small u with ogonek	S 4/14 7/5	12/14 7/5
LU44	Ų	capital U with ogonek	S 4/14 5/5	12/14 5/5
LV01	v	small v	7/6	7/6
LV02	V	capital V	5/6	5/6
LW01	w	small w	7/7	7/7
LW02	W	capital W	5/7	5/7