

ISO/IEC JTC 1/SC 2
Coded Character Sets
Secretariat: [Japan \(JISC\)](#)

| | | | |
|---------------------|---|-------------|--|
| DOC. TYPE | Text for FDIS ballot | | |
| TITLE | Revised Text of FCD 8859-11, Information technology -- 8-bit single-byte coded graphic character sets -- Part 11: Latin/Thai alphabet | | |
| SOURCE | Mr. Johan van Wingen and Mr. Surayuth Boonmatat, Projet Editor | | |
| PROJECT | | | |
| STATUS | In accordance with Resolution M10.15 of the 10th plenary meeting of SC 2 held in Athens, Greece, 2001-09-25/26, this document has been submitted to ITTF for FDIS ballot. | | |
| ACTION ID | FYI | | |
| DUE DATE | | | |
| DISTRIBUTION | P, O and L Members of ISO/IEC JTC 1/SC 2 ; ISO/IEC JTC 1 Secretariat; ISO/IEC ITTF | | |
| ACCESS LEVEL | Def | | |
| ISSUE NO. | 119 | | |
| FILE | NAME | 02n3538.pdf | |
| | SIZE (KB) | | |
| | PAGES | 15 | |

Secretariat ISO/IEC JTC 1/SC 2 - IPSJ/ITSCJ *(Information Processing Society of Japan/Information Technology Standards Commission of Japan) Room 308-3, Kikai-Shinko-Kaikan Bldg., 3-5-8, Shiba-Koen, Minato-ku, Tokyo 105-0011 Japan *Standard Organization Accredited by JISC
Telephone: +81-3-3431-2808; Facsimile: +81-3-3431-6493; E-mail: kimura@itscj.ipsj.or.jp

Contents

| | Page |
|--|-----------|
| Foreword | iii |
| Introduction | iv |
| 1 Scope | 1 |
| 2 Conformance | 1 |
| 3 Normative references | 1 |
| 4 Terms and Definitions | 2 |
| 5 Notation, code table and names | 2 |
| 6 Specification of the coded character set | 3 |
| 7 Identification of the character set | 7 |
| Annex A Coverage of languages by parts 1 to 10 and 13 to 16 of ISO/IEC 8859 | 8 |
| Bibliography | 10 |

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. 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 nongovernmental, 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 JTC1. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75% of the national bodies casting a vote.

International Standard ISO/IEC 8859-11 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 2, *Coded character sets*.

ISO/IEC 8859 consists of the following parts, under the general title *Information technology – 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*
- Part 10: *Latin alphabet No. 6*
- Part 11: *Latin/Thai alphabet*
- Part 13: *Latin alphabet No. 7*
- Part 14: *Latin alphabet No. 8*
- Part 15: *Latin alphabet No. 9*
- Part 16: *Latin alphabet No. 10*

Annex A of this part of ISO/IEC 8859 is for information only.

Introduction

ISO/IEC 8859 consists of several parts. Each part specifies a set of up to 191 graphic characters and the coded representation of these characters by means of a single 8-bit byte. Each set is intended for use for a particular group of languages.

Information technology – 8-bit single-byte coded graphic character sets –

Part 11: Latin/Thai alphabet

1 Scope

This part of ISO/IEC 8859 specifies a set of 183 coded graphic characters identified as Latin/Thai alphabet.

This set of coded graphic characters is intended for use in data and text processing applications and also for information interchange.

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

Thai, English and Latin.

Some of the characters in this set are combining characters (see clause 6).

This set of coded graphic characters may be regarded as a version of an 8-bit code according to ISO/IEC 2022 or ISO/IEC 4873 at level 1.

This part of ISO/IEC 8859 may not be used in conjunction with any other parts of ISO/IEC 8859. If coded characters from more than one part are to be used together, by means of code extension techniques, the equivalent coded character sets from ISO/IEC 10367 or their corresponding G1 sets from 'ISO International Register of Coded Character Sets to be used with escape sequences', should be used instead within a version of ISO/IEC 4873 at level 2 or level 3.

The coded characters in this set may be used in conjunction with coded control functions selected from ISO/IEC 6429. However, control functions are not used to create composite graphic symbols from two or more graphic characters (see clause 6).

NOTE – ISO/IEC 8859 is not intended for use with Telematic services defined by ITU-T. If information coded according to ISO/IEC 8859 is to be transferred to such services, it will have to conform to the requirements of those services at the access-point.

2 Conformance

2.1 Conformance of information interchange

A coded-character-data-element (CC-data-element) within coded information for interchange is in conformance with this part of ISO/IEC 8859 if all the coded representations of graphic characters within that CC-data-element conform to the requirements of clause 6.

2.2 Conformance of devices

A device is in conformance with this part of ISO/IEC 8859 if it conforms to the requirements of 2.2.1, and either or both of 2.2.2 and 2.2.3. A claim of conformance shall identify the document which contains the description specified in 2.2.1.

2.2.1 Device description

A device that conforms to this part of ISO/IEC 8859 shall be the subject of a description that identifies the means by which the user may supply characters to the device, or may recognize them when they are made available to him, as specified respectively in 2.2.2 and 2.2.3.

2.2.2 Originating devices

An originating device shall allow its user to supply any sequence of characters from those specified in clause 6, and shall be capable of transmitting their coded representations within a CC-data-element.

2.2.3 Receiving devices

A receiving device shall be capable of receiving and interpreting any coded representations of characters that are within a CC-data-element, and that conform to clause 6, and shall make the corresponding characters available to its user in such a way that the user can identify them from among those specified there, and can distinguish them from each other.

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 indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO/IEC 2022:1994, *Information technology – Character code structure and extension techniques*.

ISO/IEC 4873:1991, *Information technology – ISO 8-bit code for information interchange – Structure and rules for implementation*.

ISO/IEC 8824-1:1998, *Information technology – Abstract Syntax Notation One (ASN.1): Specification of basic notation*.

4 Terms and definitions

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

- 4.1 **bit combination:** An ordered set of bits used for the representation of characters.
- 4.2 **byte:** A bit string that is operated upon as a unit.
- 4.3 **character:** A member of a set of elements used for the organization, control, or representation of data.
- 4.4 **code table:** A table showing the characters allocated to each bit combination in a code.
- 4.5 **coded character set; code:** A set of unambiguous rules that establishes a character set and the one-to-one relationship between the characters of the set and their bit combinations.
- 4.6 **coded-character-data-element (CC-data-element):** An element of interchanged information that is specified to consist of a sequence of coded representations of characters, in accordance with one or more identified standards for coded character sets.
- 4.7 **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/IEC 8859 a single bit combination is used to represent each character.

- 4.8 **graphic symbol:** A visual representation of a graphic character or of a control function.
- 4.9 **position:** That part of a code table identified by its column and row coordinates.

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 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.

The bit combinations are also identified by notations of the form hk , where h and k are numbers in the range 0 to F in hexadecimal notation. The number h is the same as the number xx described above, and the number k the same as the number yy described above.

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. In hexadecimal notation the columns and the rows are numbered 0 to F.

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 column and row numbers are shown at the top and left edges of the table respectively. The code table positions are also identified by notations of the form hk , where h is the column number and k is the row number in hexadecimal notation. The column and row numbers are shown at the bottom and right edges of the table respectively.

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 , or of the form hk , is the same as that of the corresponding bit combination.

5.3 Names and meanings

This part of ISO/IEC 8859 assigns a unique name and a unique identifier to each graphic character. These names and identifiers have been taken from ISO/IEC 10646-1 (E). This part of ISO/IEC 8859

also specifies an acronym for each of the characters SPACE, NO-BREAK SPACE and SOFT HYPHEN. For acronyms only Latin capital letters A to Z are used. It is intended that the acronyms be retained in all translations of the text.

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.

This part of ISO/IEC 8859 specifies a graphic symbol for each graphic character. This symbol is shown in the corresponding position of the code table. However, this part, or any other part, of ISO/IEC 8859 does not specify a particular style or font design for imaging graphic characters. Annex B of ISO/IEC 10367 gives further information on this subject.

5.3.1 SPACE (SP)

A graphic character the visual representation of which consists 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 has been established within a word.

6 Specification of the coded character set

This part of ISO/IEC 8859 specifies 183 characters allocated to the bit combinations of the code table (table 2).

Some of these characters are combining characters. They are identified in table 1 as such.

NOTE – Combining characters are described in ISO/IEC 2022:1994 subclause 6.3.3.

Control functions, such as BACKSPACE or CARRIAGE RETURN, shall not be used to create composite graphic symbols, which are made up from the graphic representations of two or more characters.

6.1 Characters of the set and their coded representation

See table 1.

Table 1 – Character set, coded representation

| Bit combination | Hex | Identifier | Name |
|-----------------|-----|------------|------------------------|
| 02/00 | 20 | U+0020 | SPACE |
| 02/01 | 21 | U+0021 | EXCLAMATION MARK |
| 02/02 | 22 | U+0022 | QUOTATION MARK |
| 02/03 | 23 | U+0023 | NUMBER SIGN |
| 02/04 | 24 | U+0024 | DOLLAR SIGN |
| 02/05 | 25 | U+0025 | PERCENT SIGN |
| 02/06 | 26 | U+0026 | AMPERSAND |
| 02/07 | 27 | U+0027 | APOSTROPHE |
| 02/08 | 28 | U+0028 | LEFT PARENTHESIS |
| 02/09 | 29 | U+0029 | RIGHT PARENTHESIS |
| 02/10 | 2A | U+002A | ASTERISK |
| 02/11 | 2B | U+002B | PLUS SIGN |
| 02/12 | 2C | U+002C | COMMA |
| 02/13 | 2D | U+002D | HYPHEN-MINUS |
| 02/14 | 2E | U+002E | FULL STOP |
| 02/15 | 2F | U+002F | SOLIDUS |
| 03/00 | 30 | U+0030 | DIGIT ZERO |
| 03/01 | 31 | U+0031 | DIGIT ONE |
| 03/02 | 32 | U+0032 | DIGIT TWO |
| 03/03 | 33 | U+0033 | DIGIT THREE |
| 03/04 | 34 | U+0034 | DIGIT FOUR |
| 03/05 | 35 | U+0035 | DIGIT FIVE |
| 03/06 | 36 | U+0036 | DIGIT SIX |
| 03/07 | 37 | U+0037 | DIGIT SEVEN |
| 03/08 | 38 | U+0038 | DIGIT EIGHT |
| 03/09 | 39 | U+0039 | DIGIT NINE |
| 03/10 | 3A | U+003A | COLON |
| 03/11 | 3B | U+003B | SEMICOLON |
| 03/12 | 3C | U+003C | LESS-THAN SIGN |
| 03/13 | 3D | U+003D | EQUALS SIGN |
| 03/14 | 3E | U+003E | GREATER-THAN SIGN |
| 03/15 | 3F | U+003F | QUESTION MARK |
| 04/00 | 40 | U+0040 | COMMERCIAL AT |
| 04/01 | 41 | U+0041 | LATIN CAPITAL LETTER A |
| 04/02 | 42 | U+0042 | LATIN CAPITAL LETTER B |
| 04/03 | 43 | U+0043 | LATIN CAPITAL LETTER C |
| 04/04 | 44 | U+0044 | LATIN CAPITAL LETTER D |
| 04/05 | 45 | U+0045 | LATIN CAPITAL LETTER E |
| 04/06 | 46 | U+0046 | LATIN CAPITAL LETTER F |
| 04/07 | 47 | U+0047 | LATIN CAPITAL LETTER G |
| 04/08 | 48 | U+0048 | LATIN CAPITAL LETTER H |
| 04/09 | 49 | U+0049 | LATIN CAPITAL LETTER I |
| 04/10 | 4A | U+004A | LATIN CAPITAL LETTER J |
| 04/11 | 4B | U+004B | LATIN CAPITAL LETTER K |
| 04/12 | 4C | U+004C | LATIN CAPITAL LETTER L |
| 04/13 | 4D | U+004D | LATIN CAPITAL LETTER M |
| 04/14 | 4E | U+004E | LATIN CAPITAL LETTER N |
| 04/15 | 4F | U+004F | LATIN CAPITAL LETTER O |

Table 1 (continued)

| Bit combination | Hex | Identifier | Name |
|-----------------|-----|------------|------------------------|
| 05/00 | 50 | U+0050 | LATIN CAPITAL LETTER P |
| 05/01 | 51 | U+0051 | LATIN CAPITAL LETTER Q |
| 05/02 | 52 | U+0052 | LATIN CAPITAL LETTER R |
| 05/03 | 53 | U+0053 | LATIN CAPITAL LETTER S |
| 05/04 | 54 | U+0054 | LATIN CAPITAL LETTER T |
| 05/05 | 55 | U+0055 | LATIN CAPITAL LETTER U |
| 05/06 | 56 | U+0056 | LATIN CAPITAL LETTER V |
| 05/07 | 57 | U+0057 | LATIN CAPITAL LETTER W |
| 05/08 | 58 | U+0058 | LATIN CAPITAL LETTER X |
| 05/09 | 59 | U+0059 | LATIN CAPITAL LETTER Y |
| 05/10 | 5A | U+005A | LATIN CAPITAL LETTER Z |
| 05/11 | 5B | U+005B | LEFT SQUARE BRACKET |
| 05/12 | 5C | U+005C | REVERSE SOLIDUS |
| 05/13 | 5D | U+005D | RIGHT SQUARE BRACKET |
| 05/14 | 5E | U+005E | CIRCUMFLEX ACCENT |
| 05/15 | 5F | U+005F | LOW LINE |
| 06/00 | 60 | U+0060 | GRAVE ACCENT |
| 06/01 | 61 | U+0061 | LATIN SMALL LETTER A |
| 06/02 | 62 | U+0062 | LATIN SMALL LETTER B |
| 06/03 | 63 | U+0063 | LATIN SMALL LETTER C |
| 06/04 | 64 | U+0064 | LATIN SMALL LETTER D |
| 06/05 | 65 | U+0065 | LATIN SMALL LETTER E |
| 06/06 | 66 | U+0066 | LATIN SMALL LETTER F |
| 06/07 | 67 | U+0067 | LATIN SMALL LETTER G |
| 06/08 | 68 | U+0068 | LATIN SMALL LETTER H |
| 06/09 | 69 | U+0069 | LATIN SMALL LETTER I |
| 06/10 | 6A | U+006A | LATIN SMALL LETTER J |
| 06/11 | 6B | U+006B | LATIN SMALL LETTER K |
| 06/12 | 6C | U+006C | LATIN SMALL LETTER L |
| 06/13 | 6D | U+006D | LATIN SMALL LETTER M |
| 06/14 | 6E | U+006E | LATIN SMALL LETTER N |
| 06/15 | 6F | U+006F | LATIN SMALL LETTER O |
| 07/00 | 70 | U+0070 | LATIN SMALL LETTER P |
| 07/01 | 71 | U+0071 | LATIN SMALL LETTER Q |
| 07/02 | 72 | U+0072 | LATIN SMALL LETTER R |
| 07/03 | 73 | U+0073 | LATIN SMALL LETTER S |
| 07/04 | 74 | U+0074 | LATIN SMALL LETTER T |
| 07/05 | 75 | U+0075 | LATIN SMALL LETTER U |
| 07/06 | 76 | U+0076 | LATIN SMALL LETTER V |
| 07/07 | 77 | U+0077 | LATIN SMALL LETTER W |
| 07/08 | 78 | U+0078 | LATIN SMALL LETTER X |
| 07/09 | 79 | U+0079 | LATIN SMALL LETTER Y |
| 07/10 | 7A | U+007A | LATIN SMALL LETTER Z |
| 07/11 | 7B | U+007B | LEFT CURLY BRACKET |
| 07/12 | 7C | U+007C | VERTICAL LINE |
| 07/13 | 7D | U+007D | RIGHT CURLY BRACKET |
| 07/14 | 7E | U+007E | TILDE |

Table 1 (continued)

| Bit combination | Hex | Identifier | Name |
|-----------------|-----|------------|---------------------------------------|
| 10/00 | A0 | U+00A0 | NO-BREAK SPACE |
| 10/01 | A1 | U+0E01 | THAI CHARACTER KO KAI |
| 10/02 | A2 | U+0E02 | THAI CHARACTER KHO KHAI |
| 10/03 | A3 | U+0E03 | THAI CHARACTER KHO KHUAT |
| 10/04 | A4 | U+0E04 | THAI CHARACTER KHO KHWAI |
| 10/05 | A5 | U+0E05 | THAI CHARACTER KHO KHON |
| 10/06 | A6 | U+0E06 | THAI CHARACTER KHO RAKHANG |
| 10/07 | A7 | U+0E07 | THAI CHARACTER NGO NGU |
| 10/08 | A8 | U+0E08 | THAI CHARACTER CHO CHAN |
| 10/09 | A9 | U+0E09 | THAI CHARACTER CHO CHING |
| 10/10 | AA | U+0E0A | THAI CHARACTER CHO CHANG |
| 10/11 | AB | U+0E0B | THAI CHARACTER SO SO |
| 10/12 | AC | U+0E0C | THAI CHARACTER CHO CHOE |
| 10/13 | AD | U+0E0D | THAI CHARACTER YO YING |
| 10/14 | AE | U+0E0E | THAI CHARACTER DO CHADA |
| 10/15 | AF | U+0E0F | THAI CHARACTER TO PATAK |
| 11/00 | B0 | U+0E10 | THAI CHARACTER THO THAN |
| 11/01 | B1 | U+0E11 | THAI CHARACTER THO NANGMONTHO |
| 11/02 | B2 | U+0E12 | THAI CHARACTER THO PHUTHAO |
| 11/03 | B3 | U+0E13 | THAI CHARACTER NO NEN |
| 11/04 | B4 | U+0E14 | THAI CHARACTER DO DEK |
| 11/05 | B5 | U+0E15 | THAI CHARACTER TO TAO |
| 11/06 | B6 | U+0E16 | THAI CHARACTER THO THUNG |
| 11/07 | B7 | U+0E17 | THAI CHARACTER THO THAHAN |
| 11/08 | B8 | U+0E18 | THAI CHARACTER THO THONG |
| 11/09 | B9 | U+0E19 | THAI CHARACTER NO NU |
| 11/10 | BA | U+0E1A | THAI CHARACTER BO BAIMAI |
| 11/11 | BB | U+0E1B | THAI CHARACTER PO PLA |
| 11/12 | BC | U+0E1C | THAI CHARACTER PHO PHUNG |
| 11/13 | BD | U+0E1D | THAI CHARACTER FO FA |
| 11/14 | BE | U+0E1E | THAI CHARACTER PHO PHAN |
| 11/15 | BF | U+0E1F | THAI CHARACTER FO FAN |
| 12/00 | C0 | U+0E20 | THAI CHARACTER PHO SAMPHAO |
| 12/01 | C1 | U+0E21 | THAI CHARACTER MO MA |
| 12/02 | C2 | U+0E22 | THAI CHARACTER YO YAK |
| 12/03 | C3 | U+0E23 | THAI CHARACTER RO RUA |
| 12/04 | C4 | U+0E24 | THAI CHARACTER RU |
| 12/05 | C5 | U+0E25 | THAI CHARACTER LO LING |
| 12/06 | C6 | U+0E26 | THAI CHARACTER LU |
| 12/07 | C7 | U+0E27 | THAI CHARACTER WO WAEN |
| 12/08 | C8 | U+0E28 | THAI CHARACTER SO SALA |
| 12/09 | C9 | U+0E29 | THAI CHARACTER SO RUSI |
| 12/10 | CA | U+0E2A | THAI CHARACTER SO SUA |
| 12/11 | CB | U+0E2B | THAI CHARACTER HO HIP |
| 12/12 | CC | U+0E2C | THAI CHARACTER LO CHULA |
| 12/13 | CD | U+0E2D | THAI CHARACTER O ANG |
| 12/14 | CE | U+0E2E | THAI CHARACTER HO NOK HUK (ho nokhuk) |
| 12/15 | CF | U+0E2F | THAI CHARACTER PAIYAN NOI (paiyannoi) |

Table 1 (continued)

| Bit combination | Hex | Identifier | Name |
|-----------------|-----|------------|---|
| 13/00 | D0 | U+0E30 | THAI CHARACTER SARA A |
| 13/01 | D1 | U+0E31 | THAI CHARACTER MAI HAN-AKAT (combining) |
| 13/02 | D2 | U+0E32 | THAI CHARACTER SARA AA |
| 13/03 | D3 | U+0E33 | THAI CHARACTER SARA AM |
| 13/04 | D4 | U+0E34 | THAI CHARACTER SARA I (combining) |
| 13/05 | D5 | U+0E35 | THAI CHARACTER SARA II (combining) |
| 13/06 | D6 | U+0E36 | THAI CHARACTER SARA UE (combining) |
| 13/07 | D7 | U+0E37 | THAI CHARACTER SARA UEE (combining) |
| 13/08 | D8 | U+0E38 | THAI CHARACTER SARA U (combining) |
| 13/09 | D9 | U+0E39 | THAI CHARACTER SARA UU (combining) |
| 13/10 | DA | U+0E3A | THAI CHARACTER PHINTHU (combining) |
| 13/11 | DB | U+0E3B | (This position shall not be used) |
| 13/12 | DC | U+0E3C | (This position shall not be used) |
| 13/13 | DD | U+0E3D | (This position shall not be used) |
| 13/14 | DE | U+0E3E | (This position shall not be used) |
| 13/15 | DF | U+0E3F | THAI CURRENCY SYMBOL BAHT |
| 14/00 | E0 | U+0E40 | THAI CHARACTER SARA E |
| 14/01 | E1 | U+0E41 | THAI CHARACTER SARA AE |
| 14/02 | E2 | U+0E42 | THAI CHARACTER SARA O |
| 14/03 | E3 | U+0E43 | THAI CHARACTER SARA AI MAI MUAN (sara ai maimuan) |
| 14/04 | E4 | U+0E44 | THAI CHARACTER SARA AI MAI MALAI (sara ai maimalai) |
| 14/05 | E5 | U+0E45 | THAI CHARACTER LAKKHANG YAO (lakkhangyao) |
| 14/06 | E6 | U+0E46 | THAI CHARACTER MAI YAMOK (mai yamok) |
| 14/07 | E7 | U+0E47 | THAI CHARACTER MAI TAIKHU (maitaikhu) (combining) |
| 14/08 | E8 | U+0E48 | THAI CHARACTER MAI EK (combining) |
| 14/09 | E9 | U+0E49 | THAI CHARACTER MAI THO (combining) |
| 14/10 | EA | U+0E4A | THAI CHARACTER MAI TRI (combining) |
| 14/11 | EB | U+0E4B | THAI CHARACTER MAI CHATTAWA (combining) |
| 14/12 | EC | U+0E4C | THAI CHARACTER THANTHAKHAT (combining) |
| 14/13 | ED | U+0E4D | THAI CHARACTER NIKHAHIT (combining) |
| 14/14 | EE | U+0E4E | THAI CHARACTER YAMAKKAN (combining) |
| 14/15 | EF | U+0E4F | THAI CHARACTER FONGMAN |
| 15/00 | F0 | U+0E50 | THAI DIGIT ZERO (Thai character LEK SUN) |
| 15/01 | F1 | U+0E51 | THAI DIGIT ONE (Thai character LEK NUNG) |
| 15/02 | F2 | U+0E52 | THAI DIGIT TWO (Thai character LEK SONG) |
| 15/03 | F3 | U+0E53 | THAI DIGIT THREE (Thai character LEK SAM) |
| 15/04 | F4 | U+0E54 | THAI DIGIT FOUR (Thai character LEK SI) |
| 15/05 | F5 | U+0E55 | THAI DIGIT FIVE (Thai character LEK HA) |
| 15/06 | F6 | U+0E56 | THAI DIGIT SIX (Thai character LEK HOK) |
| 15/07 | F7 | U+0E57 | THAI DIGIT SEVEN (Thai character LEK CHET) |
| 15/08 | F8 | U+0E58 | THAI DIGIT EIGHT (Thai character LEK PAET) |
| 15/09 | F9 | U+0E59 | THAI DIGIT NINE (Thai character LEK KAO) |
| 15/10 | FA | U+0E5A | THAI CHARACTER ANGKHANKHU |
| 15/11 | FB | U+0E5B | THAI CHARACTER KHOMUT |
| 15/12 | FC | U+0E5C | (This position shall not be used) |
| 15/13 | FD | U+0E5D | (This position shall not be used) |
| 15/14 | FE | U+0E5E | (This position shall not be used) |
| 15/15 | FF | U+0E5F | (This position shall not be used) |

6.2 Code table

For each character in the set the code table (table 2) shows a graphic symbol at the position in the code table corresponding to the bit combination specified in table 1.

The shaded positions in the code table correspond to bit combinations that do not represent graphic characters. Their use is outside the scope of

ISO/IEC 8859; it is specified in other International Standards, for example ISO/IEC 6429.

The positions in the code table that are shown with cross-hatching correspond to bit combinations in table 1 having the entry "This position shall not be used".

In the table the graphic symbol for each combining character is shown together with a dotted circle to indicate its position in relation to a base character.

Table 2 – Code table of Latin/Thai alphabet

| | | | | b ₈ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
|----------------|----------------|----------------|----------------|----------------|----|----|----|----|----|----|----|----|----|----|------|----|----|----|----|----|
| | | | | b ₇ | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| | | | | b ₆ | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| | | | | b ₅ | 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 ₄ | b ₃ | b ₂ | b ₁ | | | | | | | | | | | | | | | | | |
| 0 | 0 | 0 | 0 | 00 | | | SP | 0 | ๑ | P | ` | p | | | NBSP | ๒ | ๓ | ๔ | ๕ | ๖ |
| 0 | 0 | 0 | 1 | 01 | | | ! | 1 | A | Q | a | q | | | ๗ | ๘ | ๙ | ๐ | ๑ | ๒ |
| 0 | 0 | 1 | 0 | 02 | | | " | 2 | B | R | b | r | | | ๓ | ๔ | ๕ | ๖ | ๗ | ๘ |
| 0 | 0 | 1 | 1 | 03 | | | # | 3 | C | S | c | s | | | ๙ | ๐ | ๑ | ๒ | ๓ | ๔ |
| 0 | 1 | 0 | 0 | 04 | | | \$ | 4 | D | T | d | t | | | ๑ | ๒ | ๓ | ๔ | ๕ | ๖ |
| 0 | 1 | 0 | 1 | 05 | | | % | 5 | E | U | e | u | | | ๓ | ๔ | ๕ | ๖ | ๗ | ๘ |
| 0 | 1 | 1 | 0 | 06 | | | & | 6 | F | V | f | v | | | ๙ | ๐ | ๑ | ๒ | ๓ | ๔ |
| 0 | 1 | 1 | 1 | 07 | | | ' | 7 | G | W | g | w | | | ๑ | ๒ | ๓ | ๔ | ๕ | ๖ |
| 1 | 0 | 0 | 0 | 08 | | | (| 8 | H | X | h | x | | | ๓ | ๔ | ๕ | ๖ | ๗ | ๘ |
| 1 | 0 | 0 | 1 | 09 | | |) | 9 | I | Y | i | y | | | ๕ | ๖ | ๗ | ๘ | ๙ | ๐ |
| 1 | 0 | 1 | 0 | 10 | | | * | : | J | Z | j | z | | | ๗ | ๘ | ๙ | ๐ | ๑ | ๒ |
| 1 | 0 | 1 | 1 | 11 | | | + | ; | K | [| k | { | | | ๙ | ๐ | ๑ | ๒ | ๓ | ๔ |
| 1 | 1 | 0 | 0 | 12 | | | , | < | L | \ | l | | | | ๑ | ๒ | ๓ | ๔ | ๕ | ๖ |
| 1 | 1 | 0 | 1 | 13 | | | - | = | M |] | m | } | | | ๓ | ๔ | ๕ | ๖ | ๗ | ๘ |
| 1 | 1 | 1 | 0 | 14 | | | . | > | N | ^ | n | ~ | | | ๕ | ๖ | ๗ | ๘ | ๙ | ๐ |
| 1 | 1 | 1 | 1 | 15 | | | / | ? | O | _ | o | | | | ๗ | ๘ | ๙ | ๐ | ๑ | ๒ |
| | | | | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |

hex

7 Identification of the character set

7.1 Identification according to ISO/IEC 2022 and ISO/IEC 4873

The graphic characters of this part of ISO/IEC 8859 constitute a single coded character set. However in accordance with ISO/IEC 2022 and ISO/IEC 4873 the code table of this part of ISO/IEC 8859 may 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 the identification methods of ISO/IEC 2022 or ISO/IEC 4873 are used this part of ISO/IEC 8859 shall be identified by the following pair of designation functions:

GZD4 04/02 (ESC 02/08 04/02)

G1D6 05/04 (ESC 02/13 05/04)

NOTE – The corresponding escape sequences are shown in parentheses.

7.2 Identification according to ISO/IEC 8824-1 (ASN.1)

In the terminology of ISO/IEC 8824-1 the character set of this part of ISO/IEC 8859 and the corresponding coded representations are distinct, and are known as the "character abstract syntax" and the "character transfer syntax" respectively.

When the identification methods of ISO/IEC 8824-1 are used this part of ISO/IEC 8859 shall be identified by the following object identifiers:

- character set
{ iso standard 8859 11 abstract-syntax (1) }
- coded representations
{ iso standard 8859 11 transfer-syntax (0) }

The corresponding object descriptors shall be:

- character set "ISO 8859 part 11 repertoire"
- coded representations "ISO 8859 part 11 code"

7.3 Identification using the ISO International register of coded character sets to be used with escape sequences

According to 7.1 above the character set of this part of ISO/IEC 8859 may be considered to consist of the character SPACE, a 94-character G0 graphic character set, and a 96-character G1 graphic character set. The G0 and G1 graphic character sets may be identified by the use of the Registration Numbers from the ISO International register of coded character sets to be used with escape sequences.

When these registration numbers are used this part of ISO/IEC 8859 shall be identified by the following pair of registration numbers:

- G0 graphic character set ISO-IR 6
- G1 graphic character set ISO-IR 166

Annex A (informative)

Coverage of languages by parts 1 to 10 and 13 to 16 of ISO/IEC 8859

A.1 Languages of European origin written in Latin script

The following parts 1–16 of ISO/IEC 8859 specify coded character sets which comprise various different selections of characters based on the Latin alphabet. These sets are identified by the numbers 1 to 10 as shown.

ISO/IEC 8859-1 Latin alphabet No. 1
ISO/IEC 8859-2 Latin alphabet No. 2
ISO/IEC 8859-3 Latin alphabet No. 3

ISO/IEC 8859-4 Latin alphabet No. 4
ISO/IEC 8859-9 Latin alphabet No. 5
ISO/IEC 8859-10 Latin alphabet No. 6
ISO/IEC 8859-13 Latin alphabet No. 7
ISO/IEC 8859-14 Latin alphabet No. 8 (Celtic)
ISO/IEC 8859-15 Latin alphabet No. 9
ISO/IEC 8859-16 Latin alphabet No. 10

The following official and regional languages written in Europe are known to be covered by the Latin alphabets 1–10 as indicated by number in table A.1:

Table A.1 – Language coverage

| Language | Covered by alphabet(s) | Language | Covered by alphabet(s) |
|-----------------------------------|------------------------|-----------------------------------|------------------------|
| Albanian | 1 2 5 8 9 10 | Irish Gaelic (old orthography) | 8 |
| Basque | 1 5 8 9 | Italian | 1 3 5 8 9 10 |
| Breton | 1 5 8 9 | Latin | 1 2 3 4 5 6 7 8 9 10 |
| Catalan | 1 5 8 9 | Latvian | 4 7 |
| Cornish | 1 5 8 | Lithuanian | 4 6 7 |
| Croatian | 2 10 | Luxemburgish | 1 5 8 9 |
| Czech | 2 | Maltese | 3 |
| Danish | 1 4 5 6 8 9 | Manx Gaelic | 8 |
| Dutch | 1 5 9 | Norwegian | 1 4 5 6 7 8 9 |
| English | 1 2 3 4 5 6 7 8 9 10 | Polish | 2 7 10 |
| Esperanto | 3 | Portuguese | 1 3 5 8 9 |
| Estonian | 4 6 7 9 | Rhaeto-Romanic | 1 5 8 9 |
| Faroese | 1 6 9 | Romanian | (2) 10 |
| Finnish | (1) 4 (5) 6 7 (8) 9 10 | Sámi | 4 6 |
| French | (1) (3) (5) (8) 9 10 | Scottish Gaelic | 1 5 8 9 |
| Frisian | 1 5 9 | Slovak | 2 |
| Galician | 1 5 8 9 | Slovenian | 2 4 6 10 |
| German | 1 2 3 4 5 6 8 9 10 | Sorbian | 2 |
| Greenlandic | 1 4 5 6 8 9 | Spanish | 1 5 8 9 |
| Hungarian | 2 10 | Swedish | 1 4 5 6 8 9 |
| Icelandic | 1 6 9 | Turkish | (3) 5 |
| Irish Gaelic (new orthography) | 1 5 6 8 9 10 | Welsh | 8 |

NOTES

1 The list of languages in table A.1 is not exhaustive. It shows the languages that are included in the Scope clause of each part of ISO/IEC 8859.

2 For writing French three characters (Œ, œ, Ÿ), not covered in parts 1, 3, 9, and 14, are included in parts 15 and 16. For writing Finnish four characters (Š, š, Ž, ž), not covered in parts 1, 9, and 14, are included in parts 4, 10, 13, 15, and 16.

3 The various Sámi languages use partly differing orthographies. The character sets in parts 4 and 10 cover the requirements of the Sámi languages most commonly used in Finland, Norway and Sweden. For the Skolt Sámi language used in Finland and Norway additional characters are needed. These are included in ISO-IR 158, 197, and 209.

4 There are several official written languages outside Europe that are covered by Latin alphabet No. 1. Examples are Indonesian/Malay, Tagalog (Philippines), Swahili, Afrikaans.

5 Use of Latin alphabet No. 3 for Turkish is deprecated.

A.2 Languages written in non-Latin scripts

The following parts of ISO/IEC 8859 specify coded character sets which include graphic characters from alphabets other than the Latin alphabet:

| | |
|-----------------|-------------------------|
| ISO/IEC 8859-5 | Latin/Cyrillic alphabet |
| ISO/IEC 8859-6 | Latin/Arabic alphabet |
| ISO/IEC 8859-7 | Latin/Greek alphabet |
| ISO/IEC 8859-8 | Latin/Hebrew alphabet |
| ISO/IEC 8859-11 | Latin/Thai alphabet |

The following official and regional languages are covered by these alphabets:

The Cyrillic characters included in part 5 cover Bulgarian, Byelorussian, (Slavic) Macedonian, Russian, Serbian and Ukrainian (as written up to 1990, see also Scope of part 5).

The Arabic characters included in part 6 cover Arabic. The Greek characters included in part 7 cover Greek (*monotonikó* orthography). The Hebrew characters included in part 8 cover Hebrew. The Thai characters included in part 11 cover Thai.

Bibliography

ISO/IEC 6429:1992, *Information technology – Control functions for coded character sets*.

ISO/IEC 10367:1991, *Information technology – Standardized coded graphic character sets for use in 8-bit codes*.

ISO/IEC 10646-1:2000, *Information technology – Universal Multiple-Octet Coded Character Set (UCS) – Part 1: Architecture and Basic Multilingual Plane*.

ISO *International register of coded character sets to be used with escape sequences*.

TIS 620-2533(1990), *Thai character codes for computers*.

TIS 1566-2541(1998), *Thai input/output methods for computers*.

G8 Explanatory Report

| | |
|----------------------------------|-----------------------------|
| EXPLANATORY REPORT | ISO/IEC FDIS 8859-11 |
| ISO/IEC JTC 1/SC 2 N 3537 | |
| Will supersede: SC 2 N 3333 | Secretariat: Japan (JISC) |

This form should be sent to ITTF, together with the committee draft, by the secretariat of the joint technical committee or sub-committee concerned.

| |
|---|
| <p>The accompanying document is submitted for circulation to member body vote as an FDIS, following consensus of the P-members of the committee obtained on:</p> |
| <p>by postal ballot initiated on: 1999-06-22</p> |
| <p>P-members in favour: Belgium, Canada, China, Denmark, Egypt, Finland, France, Germany, Greece, Israel, Italy, Japan, Korea (Republic of), Netherlands, Romania, Sweden, Thailand, UK</p> <p>P-members voting against: Poland, USA</p> <p>P-members abstaining: Australia, Brazil, Norway, Tunisia</p> <p>P-members who did not vote: Armenia, Austria, Iceland, India, Iran (Islamic of), Ireland, Morocco, Russian Federation, Singapore, Slovenia, Turkey, Yugoslavia</p> |
| <p>Remarks:</p> <p><u>History</u></p> <p>FCD was conducted as SC 2 N 3333</p> <p>Summary of Voting: SC 2 N 3383</p> <p>Disposition of Comments Report: SC 2 N 3511</p> <p>In accordance with Resolution M10.15 adopted at the 10th SC 2 plenary meeting held in Athens, Greece, 2000-09-25/26, this document is submitted to ITTF for FDIS ballot.</p> |
| <p>Project: JTC 1.02.20.11</p> |
| <p>I hereby confirm that this draft meets the requirements of part 3 of the IEC/ISO Directives</p> |
| <p>Date: 2001-06-29</p> |
| <p>Name and signature of secretary:</p> <p style="text-align: right;">Toshiko KIMURA Secretariat, ISO/IEC JTC 1/SC 2</p> |