8-bits kodningsscheman för samiska språk

8-bit coded character sets for Sami languages

Innehållsförteckning

8-bits kodningsscheman för samiska språk
8-bit coded character sets for Sami languages

1 Omfattning

Denna Statskontorets tekniska norm anger tre kodningsscheman speciellt avsedda att användas för samiska språk vid 8-bits databearbetning och informationsutbyte.

Dessa scheman motsvarar tre vanliga datormiljöer, nämligen ISO/IEC 8859-kompatibel (t.ex. flertalet Unix-system), "Windows" och "Macintosh".

Anmärkning: "Macintosh" är ett varumärke ägt av Apple Computer, Inc. "Windows" är ett varumärke ägt av Microsoft Corporation. Dessa varumärken är registrerade i Sverige och andra länder.

2 Motivering

Denna tekniska norm har tagits fram för att möta konstaterade behov hos svenska myndigheter av en officiell dokumentation av 8-bits kodningsscheman för samiska.

Behovet av 8-bitsscheman kommer att avta i och med en ökande användning av fler-oktetts kodningsscheman ("Unicode", ISO/IEC 10646) i databearbetning och datautbyte. Emellertid kommer en-oktetts (8-bit) kodningsscheman att finnas i bruk under ytterligare ett antal år, och därmed även behovet av dokumentation.

3 Bakgrund
3.1 Lagstiftning om minoritetsspråk

I april 2000 trädde ny svensk lagstiftning rörande rättigheter för användare av minoritetsspråk i kraft. Vad gäller samisktalande har dessa numera i delar av norra Sverige rätt att använda samiska, muntligen och skriftligen, i sin kontakt med myndigheter.

Lagstiftningen har tolkats som att den medger användning av samiska inte endast i fysiska dokument (papper) utan även i elektroniska dokument, dvs e-post, datafiler osv. Detta innebär svårigheter för myndigheterna, eftersom de samiska språkens alfabet inte täcks helt av de kodningstabeller för databehandling som i allmänhet används i Sverige. Dokument på samiska i elektronisk form, som inkommer till en myndighet, har därför i allmänhet producerats på datorsystem med speciella 8-bits scheman. Detta kan medföra komplikationer med arkivering och informationssökning i de mottagande datasystemen.

1 Scope

This Statskontoret Technical Standard defines three character coding schemes, intended specifically for use in Sami-language 8-bit data processing and information interchange.

The schemes correspond to three common computer environments, namely ISO/IEC 8859-compatible (e.g. most Unix systems), "Windows", and "Macintosh".

Acknowledgment: "Macintosh" is a trademark of Apple Computer, Inc. "Windows" is a trademark of Microsoft Corporation. These trademarks are registered in Sweden and other jurisdictions.

2 Rationale

This Technical Standard was developed in response to the needs stated by Swedish authorities for an official documentation on Sami 8-bit coding schemes.

The need for 8-bit coding will diminish with the increasing use of multi-byte coding schemes ("Unicode", ISO/IEC 10646) in data processing and interchange. For a number of years to come, however, single-byte (8-bit) schemes will remain in use, and the requirement for documentation therefore also exist.

3 Background
3.1 Minority language legislation

In April 2000, new Swedish legislation came into force relating to the rights of users of minority languages. As regards the Sami languages, users in parts of northern Sweden are now explicitly entitled to communicate in Sami with authorities, verbally and in writing.

The legislation has been interpreted as permitting communication in Sami not only by physical documents (paper), but also by electronic documents, i.e. e-mails, computer files etc. This poses difficulties for the Swedish authorities, since the alphabets of the Sami languages are not completely covered by the computer coding schemes generally used in Sweden. Electronic documents in Sami received by the authorities will therefore in general have been produced on systems using special 8-bit schemes. These schemes may cause complications in archiving and information retrieval in the receiving computer systems.

För att hantera detta problem har Statskontoret beslutat att utge denna tekniska norm, som specificerar rekommenderade 8-bits kodningsscheman för de vanligaste datamiljöerna. Genom att referera till denna tekniska norm kan myndigheterna informera användare av samiska om vilken teckenkodning som måste användas för elektroniska dokument om korrekt mottagning, lagring och utskrift av dessa skall kunna garanteras.

3.2 8-bits kodningsscheman i Sverige

Den i Sverige mest vanliga internationella 8-bits-standarden är ISO/IEC 8859-1, "Latin-1". Detta schema – eller "supersets" av det – används främst i Unix-system.

I 8859-standardserien ingår två kodningsscheman som innehåller både det svenska alfabetet och de speciella bokstäverna i de samiska språk som används i Sverige, ISO/IEC 8859-4 och -10. Dessa scheman har emellertid inte i allmänhet implementerats av dataindustrin. Dessutom innehåller de inte de speciella bokstäverna i skoltsamiska, som används i Finland, och är därför inte lämpliga för allmänt samiskt bruk.

För de två vanliga persondatormiljöerna i Sverige, "Windows" och "Macintosh", används i allmänhet schema "Windows Western" respektive "Macintosh Roman". I likhet med "Latin-1" stöder dessa inte de samiska alfabetena.

3.3 Nordiskt arbete på samiska kodscheman

Under ett antal år har svenska, norska och finska sametingen arbetat gemensamt för att ta fram en lösning på det ovan beskrivna problemet. Arbetet har skett i samverkan med datakonsultföretag och andra organisationer. Svenska, norska och finska standardiseringsorganisationerna (SIS/ITS, NTS och TIEKE) har även medverkat.

Anmärkning: Arbete utfört vid Tromsö Universitet samt av företaget Everson Gunn Teoranta, Dublin, Irland, har varit särskilt viktigt i sammanhanget.

Arbetet resulterade 1998 i tre olika kodnings-scheman:

1. Ett schema av 8859-typ, avsett i första hand för Unix-miljöer. Detta schema är baserat på "Latin-1", med vissa tecken utbytta mot de behövliga samiska bokstäverna. Delarna av detta schema är registrerade i ett ISO-register.

2. Ett schema avsett för "Windows-miljö", baserat på kodtabellen "Windows Western" (code page 1252).

3. Ett schema avsett för "Macintosh-miljö", baserat på dess "Icelandic" kodtabell (code page 10079).

To tackle this problem, Statskontoret decided to publish this Technical Standard, specifying preferred 8-bit coding schemes for the most common data processing environments. By referencing this document, authorities can inform users of the Sami languages about what coding must be used in electronic communications if correct reception, storage and printing of messages is to be guaranteed.

3.2 8-bit coding schemes in Sweden

The international 8-bit coding standard most common in Sweden is the ISO/IEC 8859-1, "Latin-1". This scheme – or supersets of it – is used primarily in Unix systems.

In the 8859 series of standards two schemes exist which contain both the Swedish alphabet and the special letters of the Sami languages used in Sweden, ISO/IEC 8859-4 and -10. These schemes have however not been generally implemented by the computer industry. Also they do not contain the special letters used in Skolt Sami, which is spoken in Finland, and they are therefore not suitable for general Sami use.

For the two common personal computer environments in Sweden, "Windows" and "Macintosh", the schemes "Windows Western" and "Macintosh Roman" are generally used. Like "Latin-1", neither of these schemes supports the Sami alphabets.

3.3 Nordic work on Sami coding schemes

For several years, the Swedish, Norwegian and Finnish "Sameting" (Sami Council) worked jointly to produce a solution to the problem described above. The work has been assisted by consultant companies and other organizations. The Swedish, Norwegian and Finnish IT standardization bodies (SIS/ITS, NTS, and TIEKE, respectively) have also been involved.

Note: Work performed at the University of Tromsø, Norway, and by the company Everson Gunn Teoranta, Dublin, Ireland, has been particularly important in this connection.

In 1998, the work resulted in three different coding schemes:

1. An 8859-type scheme, intended primarily for Unix environments. The scheme is based on the "Latin-1" code table, in which some characters were exchanged for the Sami letters needed. The components of this scheme are registered in an ISO Register.

2. A scheme intended for "Windows" environments, based on the "Windows Western" code table (code page 1252).

3. A scheme intended for "Macintosh" environments, based on its "Icelandic" code table (code page 10079).

Det skall observeras att dataföretagen Microsoft Corporation och Apple Computer, Inc. EJ deltagit i utvecklingen av schema 2 och 3, och att dessa EJ godkänts av dem. Inga som helst åtaganden föreligger alltså från företagen ifråga vad gäller implementering och/eller underhåll av dessa scheman.

I stället är dessa scheman avsedda att implementeras av oberoende programvaruutvecklare, för miljöerna "Windows" och "Macintosh".

(Fullständiga implementeringar har bekostats av norska staten, och är tillgängliga som "freeware" under namnen "Levi" och "Sami Utilities for Macintosh". Dessa programpaket tillför samisk tangentbordsfunktionalitet samt fonter, för de tre fonttyperna Courier, Helvetica och Times.)

4   Specifikation av kodscheman

De tre kodscheman som beskrivs ovan specificeras i denna tekniska norms annex A, B och C. Dessa tre scheman rekommenderas för användning i 8-bits kodning vid elektronisk kommunikation med svenska myndigheter på samiska.

För att referera till dessa scheman skall numret på denna Statskontorets tekniska norm anges, tillsammans med respektive beteckning "ISO/IEC 8859-environment", "Windows-environment" eller "Macintosh-environment".

It shall be noted that the computer companies Microsoft Corporation and Apple Computer, Inc., have NOT taken part in the development of the schemes 2 and 3, and that the schemes are NOT endorsed by them. No commitments whatsoever therefore exist from these companies regarding the implementation and/or maintenance of the schemes.

Instead, these schemes are intended for implementation by independent software developers, for the "Windows" and "Macintosh" environments.

(Complete implementations have been financed by the Norwegian state, and are available as "freeware" packages, under the names of "Levi" and "Sami Utilities for Macintosh". These packages add Sami keyboard functionality and three fonts, of the Courier, Helvetica and Times types.)

4   Specification of coded character sets

The three coding schemes described above are specified in this Technical Standard in Annex A, B and C. These three schemes are recommended for 8-bit coding use in electronic communications with Swedish authorities in the Sami languages.

To reference the schemes, the number of this Statskontoret Technical Standard shall be given, together with the designation "ISO/IEC 8859-environment", "Windows-environment" or "Macintosh-environment", respectively.

This page intentionally left blank

Annex A

(normative)

ISO/IEC 8859-environment Sami 8-bit coded graphic character set

A.1 Scope

This Annex specifies a set of 191 coded graphic characters intended for Sami-language applications in ISO/IEC 8859-compatible environments.

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 the Sami languages used in Finland, Norway and Sweden, including Skolt Sami. Sami languages used in Russia and written in the Cyrillic alphabet are not covered.

The set also contains characters used in at least the following languages:

Albanian, Basque, Breton, Catalan, Danish, Dutch, English, Estonian, Faroese, Finnish, French (with a few restrictions), Frisian, Galician, German, Greenlandic, Icelandic, Irish Gaelic (new orthography), Italian, Latin, Luxemburgish, Norwegian, Portuguese, Rhaeto-Romanic, Scottish Gaelic, Slovene, Spanish and Swedish.

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. It can therefore also be used in conjunction with coded character sets taken from ISO/IEC 10367, by means of code extension techniques, 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 A.6).

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

A.2 Conformance

A.2.1 Conformance of information interchange

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

A.2.2 Conformance of devices

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

A.2.2.1 Device description

A device that conforms to this Annex 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 A.2.2.2 and A.2.2.3.

A.2.2.2 Originating devices

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

A.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 A.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.

A.3 Normative references

The following International Standards contain provisions which, through reference in this text, constitute provisions of this Annex. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this Annex 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.*

A.4 Definitions

For the purposes of this Annex the following definitions apply:

A.4.1 bit combination: An ordered set of bits used for the representation of characters.

A.4.2 byte: A bit string that is operated upon as a unit.

A.4.3 character: A member of a set of elements used for the organization, control, or representation of data.

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

A.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.

A.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.

A.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 this Annex a single bit combination is used to represent each character.

A.4.8 graphic symbol: A visual representation of a graphic character or of a control function.

A.4.9 position: That part of a code table identified by its column and row coordinates.

A.5 Notation, code table and names

A.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.

A.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.

A.5.3 Names and meanings

This Annex 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 Annex does not define and does not restrict the meanings of graphic characters.

This Annex specifies a graphic symbol for each graphic character. This symbol is shown in the corresponding position of the code table. However, this Annex 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.

A.5.3.1 SPACE (SP)

A graphic character the visual representation of which consists of the absence of a graphic symbol.

A.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.

A.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.

A.6 Specification of the coded character set

This Annex specifies 191 characters allocated to the bit combinations of the code table (table A.2).

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

A.6.1 Characters of the set and their coded representation

See table A.1.

Table A.1 – Character set, coded representation

Bit combi-nation	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
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

Table A.1 *(continued)*

Bit combi-nation	Hex	Identifier	Name
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
10/00	A0	U+00A0	NO-BREAK SPACE
10/01	A1	U+010C	LATIN CAPITAL LETTER C WITH CARON
10/02	A2	U+010D	LATIN SMALL LETTER C WITH CARON
10/03	A3	U+0110	LATIN CAPITAL LETTER D WITH STROKE
10/04	A4	U+0111	LATIN SMALL LETTER D WITH STROKE
10/05	A5	U+01E4	LATIN CAPITAL LETTER G WITH STROKE
10/06	A6	U+01E5	LATIN SMALL LETTER G WITH STROKE
10/07	A7	U+00A7	SECTION SIGN
10/08	A8	U+01E6	LATIN CAPITAL LETTER G WITH CARON
10/09	A9	U+00A9	COPYRIGHT SIGN
10/10	AA	U+01E7	LATIN SMALL LETTER G WITH CARON
10/11	AB	U+021E	LATIN CAPITAL LETTER H WITH CARON
10/12	AC	U+01E8	LATIN CAPITAL LETTER K WITH CARON
10/13	AD	U+00AD	SOFT HYPHEN
10/14	AE	U+01E9	LATIN SMALL LETTER K WITH CARON
10/15	AF	U+014A	LATIN CAPITAL LETTER ENG
11/00	B0	U+00B0	DEGREE SIGN
11/01	B1	U+014B	LATIN SMALL LETTER ENG
11/02	B2	U+0160	LATIN CAPITAL LETTER S WITH CARON
11/03	B3	U+0161	LATIN SMALL LETTER S WITH CARON
11/04	B4	U+00B4	ACUTE ACCENT
11/05	B5	U+0166	LATIN CAPITAL LETTER T WITH STROKE
11/06	B6	U+00B6	PILCROW SIGN
11/07	B7	U+00B7	MIDDLE DOT
11/08	B8	U+0167	LATIN SMALL LETTER T WITH STROKE
11/09	B9	U+017D	LATIN CAPITAL LETTER Z WITH CARON
11/10	BA	U+017E	LATIN SMALL LETTER Z WITH CARON
11/11	BB	U+021F	LATIN SMALL LETTER H WITH CARON
11/12	BC	U+01B7	LATIN CAPITAL LETTER EZH
11/13	BD	U+0292	LATIN SMALL LETTER EZH
11/14	BE	U+01EE	LATIN CAPITAL LETTER EZH WITH CARON
11/15	BF	U+01EF	LATIN SMALL LETTER EZH WITH CARON

Table A.1 *(concluded)*

Bit combi-nation	Hex	Identifier	Name
12/00	C0	U+00C0	LATIN CAPITAL LETTER A WITH GRAVE
12/01	C1	U+00C1	LATIN CAPITAL LETTER A WITH ACUTE
12/02	C2	U+00C2	LATIN CAPITAL LETTER A WITH CIRCUMFLEX
12/03	C3	U+00C3	LATIN CAPITAL LETTER A WITH TILDE
12/04	C4	U+00C4	LATIN CAPITAL LETTER A WITH DIAERESIS
12/05	C5	U+00C5	LATIN CAPITAL LETTER A WITH RING ABOVE
12/06	C6	U+00C6	LATIN CAPITAL LETTER AE
12/07	C7	U+00C7	LATIN CAPITAL LETTER C WITH CEDILLA
12/08	C8	U+00C8	LATIN CAPITAL LETTER E WITH GRAVE
12/09	C9	U+00C9	LATIN CAPITAL LETTER E WITH ACUTE
12/10	CA	U+00CA	LATIN CAPITAL LETTER E WITH CIRCUMFLEX
12/11	CB	U+00CB	LATIN CAPITAL LETTER E WITH DIAERESIS
12/12	CC	U+00CC	LATIN CAPITAL LETTER I WITH GRAVE
12/13	CD	U+00CD	LATIN CAPITAL LETTER I WITH ACUTE
12/14	CE	U+00CE	LATIN CAPITAL LETTER I WITH CIRCUMFLEX
12/15	CF	U+00CF	LATIN CAPITAL LETTER I WITH DIAERESIS
13/00	D0	U+00D0	LATIN CAPITAL LETTER ETH
13/01	D1	U+00D1	LATIN CAPITAL LETTER N WITH TILDE
13/02	D2	U+00D2	LATIN CAPITAL LETTER O WITH GRAVE
13/03	D3	U+00D3	LATIN CAPITAL LETTER O WITH ACUTE
13/04	D4	U+00D4	LATIN CAPITAL LETTER O WITH CIRCUMFLEX
13/05	D5	U+00D5	LATIN CAPITAL LETTER O WITH TILDE
13/06	D6	U+00D6	LATIN CAPITAL LETTER O WITH DIAERESIS
13/07	D7	U+00D7	MULTIPLICATION SIGN
13/08	D8	U+00D8	LATIN CAPITAL LETTER O WITH STROKE
13/09	D9	U+00D9	LATIN CAPITAL LETTER U WITH GRAVE
13/10	DA	U+00DA	LATIN CAPITAL LETTER U WITH ACUTE
13/11	DB	U+00DB	LATIN CAPITAL LETTER U WITH CIRCUMFLEX
13/12	DC	U+00DC	LATIN CAPITAL LETTER U WITH DIAERESIS
13/13	DD	U+00DD	LATIN CAPITAL LETTER Y WITH ACUTE
13/14	DE	U+00DE	LATIN CAPITAL LETTER THORN
13/15	DF	U+00DF	LATIN SMALL LETTER SHARP S
14/00	E0	U+00E0	LATIN SMALL LETTER A WITH GRAVE
14/01	E1	U+00E1	LATIN SMALL LETTER A WITH ACUTE
14/02	E2	U+00E2	LATIN SMALL LETTER A WITH CIRCUMFLEX
14/03	E3	U+00E3	LATIN SMALL LETTER A WITH TILDE
14/04	E4	U+00E4	LATIN SMALL LETTER A WITH DIAERESIS
14/05	E5	U+00E5	LATIN SMALL LETTER A WITH RING ABOVE
14/06	E6	U+00E6	LATIN SMALL LETTER AE
14/07	E7	U+00E7	LATIN SMALL LETTER C WITH CEDILLA
14/08	E8	U+00E8	LATIN SMALL LETTER E WITH GRAVE
14/09	E9	U+00E9	LATIN SMALL LETTER E WITH ACUTE
14/10	EA	U+00EA	LATIN SMALL LETTER E WITH CIRCUMFLEX
14/11	EB	U+00EB	LATIN SMALL LETTER E WITH DIAERESIS
14/12	EC	U+00EC	LATIN SMALL LETTER I WITH GRAVE
14/13	ED	U+00ED	LATIN SMALL LETTER I WITH ACUTE
14/14	EE	U+00EE	LATIN SMALL LETTER I WITH CIRCUMFLEX
14/15	EF	U+00EF	LATIN SMALL LETTER I WITH DIAERESIS
15/00	F0	U+00F0	LATIN SMALL LETTER ETH
15/01	F1	U+00F1	LATIN SMALL LETTER N WITH TILDE
15/02	F2	U+00F2	LATIN SMALL LETTER O WITH GRAVE
15/03	F3	U+00F3	LATIN SMALL LETTER O WITH ACUTE
15/04	F4	U+00F4	LATIN SMALL LETTER O WITH CIRCUMFLEX
15/05	F5	U+00F5	LATIN SMALL LETTER O WITH TILDE
15/06	F6	U+00F6	LATIN SMALL LETTER O WITH DIAERESIS
15/07	F7	U+00F7	DIVISION SIGN
15/08	F8	U+00F8	LATIN SMALL LETTER O WITH STROKE
15/09	F9	U+00F9	LATIN SMALL LETTER U WITH GRAVE
15/10	FA	U+00FA	LATIN SMALL LETTER U WITH ACUTE
15/11	FB	U+00FB	LATIN SMALL LETTER U WITH CIRCUMFLEX
15/12	FC	U+00FC	LATIN SMALL LETTER U WITH DIAERESIS
15/13	FD	U+00FD	LATIN SMALL LETTER Y WITH ACUTE
15/14	FE	U+00FE	LATIN SMALL LETTER THORN
15/15	FF	U+00FF	LATIN SMALL LETTER Y WITH DIAERESIS

A.6.2 Code table

For each character in the set the code table (table A.2) shows a graphic symbol at the position in the code table corresponding to the bit combination specified in table A.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 this Annex; it is specified in International Standards, for example ISO/IEC 6429.

Table A.2 – Code table of ISO/IEC 8859-environment Sami character set

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	hex
0 0 0 0 00			SP	0	@	P	`	p			NBSP	°	À	Ð	à	ð	0
0 0 0 1 01			!	1	A	Q	a	q			Č	ŋ	Á	Ñ	á	ñ	1
0 0 1 0 02			''	2	B	R	b	r			č	Š	Â	Ò	â	ò	2
0 0 1 1 03			#	3	C	S	c	s			Đ	š	Ã	Ó	ã	ó	3
0 1 0 0 04			$	4	D	T	d	t			đ	´	Ä	Ô	ä	ô	4
0 1 0 1 05			%	5	E	U	e	u			Ǥ	Ŧ	Å	Õ	å	õ	5
0 1 1 0 06			&	6	F	V	f	v			ǥ	¶	Æ	Ö	æ	ö	6
0 1 1 1 07			'	7	G	W	g	w			§	·	Ç	×	ç	÷	7
1 0 0 0 08			(	8	H	X	h	x			Ǧ	ŧ	È	Ø	è	ø	8
1 0 0 1 09			)	9	I	Y	i	y			©	Ž	É	Ù	é	ù	9
1 0 1 0 10			*	:	J	Z	j	z			ǧ	ž	Ê	Ú	ê	ú	A
1 0 1 1 11			+	;	K	[	k	{			Ȟ	ȟ	Ë	Û	ë	û	B
1 1 0 0 12			,	<	L	\	l	|			Ǩ	Ʒ	Ì	Ü	ì	ü	C
1 1 0 1 13			−	=	M	]	m	}			SHY	ʒ	Í	Ý	í	ý	D
1 1 1 0 14			.	>	N	^	n	~			ǩ	Ǯ	Î	Þ	î	þ	E
1 1 1 1 15			/	?	O	_	o				Ŋ	ǯ	Ï	ß	ï	ÿ	F
	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	hex

A.7 Identification of the character set

The graphic characters of this Annex constitute a single coded character set. However in accordance with ISO/IEC 2022 and ISO/IEC 4873 the code table of this Annex 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   06/08   (ESC 02/13 06/08)

NOTE – The corresponding escape sequences are shown in the parentheses.

The G0 and G1 graphic character sets may also be identified by the use of their registration numbers from the ISO "International register of coded character sets to be used with escape sequences".

When those registration numbers are used the coded character set of this Annex shall be identified by the following pair of registration numbers:

– G0 graphic character set ISO-IR 6

– G1 graphic character set ISO-IR 209

Annex B
(normative)

Windows-environment Sami 8-bit coded graphic character set

B.1 Scope

This Annex defines a set of 216 coded graphic characters intended for Sami-language applications in Windows environments.

For Conformance, Definitions and Notations see Annex A, as applicable.

B.2 Characters of the set

See table B.1.

Table B.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 B.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 B.1 *(continued)*

Bit combination	Hex	Identifier	Name
08/00	80	U+20AC	EURO SIGN
08/01	81		*(This position not used)*
08/02	82	U+010C	LATIN CAPITAL LETTER C WITH CARON
08/03	83	U+0192	LATIN SMALL LETTER F WITH HOOK (Florin sign)
08/04	84	U+010D	LATIN SMALL LETTER C WITH CARON
08/05	85	U+01B7	LATIN CAPITAL LETTER EZH
08/06	86	U+0292	LATIN SMALL LETTER EZH
08/07	87	U+01EE	LATIN CAPITAL LETTER EZH WITH CARON
08/08	88	U+01EF	LATIN SMALL LETTER EZH WITH CARON
08/09	89	U+0110	LATIN CAPITAL LETTER D WITH STROKE
08/10	8A	U+0160	LATIN CAPITAL LETTER S WITH CARON
08/11	8B	U+2039	SINGLE LEFT-POINTING ANGLE QUOTATION MARK
08/12	8C	U+0152	LATIN CAPITAL LIGATURE OE
08/13	8D		*(This position not used)*
08/14	8E		*(This position not used)*
08/15	8F		*(This position not used)*
09/00	90		*(This position not used)*
09/01	91	U+2018	LEFT SINGLE QUOTATION MARK
09/02	92	U+2019	RIGHT SINGLE QUOTATION MARK
09/03	93	U+201C	LEFT DOUBLE QUOTATION MARK
09/04	94	U+201D	RIGHT DOUBLE QUOTATION MARK
09/05	95	U+2022	BULLET
09/06	96	U+2013	EN DASH
09/07	97	U+2014	EM DASH
09/08	98	U+0111	LATIN SMALL LETTER D WITH STROKE
09/09	99	U+01E6	LATIN CAPITAL LETTER G WITH CARON
09/10	9A	U+0161	LATIN SMALL LETTER S WITH CARON
09/11	9B	U+203A	SINGLE RIGHT-POINTING ANGLE QUOTATION MARK
09/12	9C	U+0153	LATIN SMALL LIGATURE OE
09/13	9D		*(This position not used)*
09/14	9E		*(This position not used)*
09/15	9F	U+0178	LATIN CAPITAL LETTER Y WITH DIAERESIS
10/00	A0	U+00A0	NO-BREAK SPACE
10/01	A1	U+01E7	LATIN SMALL LETTER G WITH CARON
10/02	A2	U+01E4	LATIN CAPITAL LETTER G WITH STROKE
10/03	A3	U+00A3	POUND SIGN
10/04	A4	U+00A4	CURRENCY SIGN
10/05	A5	U+01E5	LATIN SMALL LETTER G WITH STROKE
10/06	A6	U+00A6	BROKEN BAR
10/07	A7	U+00A7	SECTION SIGN
10/08	A8	U+00A8	DIAERESIS
10/09	A9	U+00A9	COPYRIGHT SIGN
10/10	AA	U+021E	LATIN CAPITAL LETTER H WITH CARON
10/11	AB	U+00AB	LEFT-POINTING DOUBLE ANGLE QUOTATION MARK
10/12	AC	U+00AC	NOT SIGN
10/13	AD	U+00AD	SOFT HYPHEN
10/14	AE	U+00AE	REGISTERED SIGN
10/15	AF	U+021F	LATIN SMALL LETTER H WITH CARON
11/00	B0	U+00B0	DEGREE SIGN
11/01	B1	U+00B1	PLUS-MINUS SIGN
11/02	B2	U+01E8	LATIN CAPITAL LETTER K WITH CARON
11/03	B3	U+01E9	LATIN SMALL LETTER K WITH CARON
11/04	B4	U+00B4	ACUTE ACCENT
11/05	B5	U+00B5	MICRO SIGN
11/06	B6	U+00B6	PILCROW SIGN
11/07	B7	U+00B7	MIDDLE DOT
11/08	B8	U+014A	LATIN CAPITAL LETTER ENG
11/09	B9	U+014B	LATIN SMALL LETTER ENG
11/10	BA	U+0166	LATIN CAPITAL LETTER T WITH STROKE
11/11	BB	U+00BB	RIGHT-POINTING DOUBLE ANGLE QUOTATION MARK
11/12	BC	U+0167	LATIN SMALL LETTER T WITH STROKE
11/13	BD	U+00BD	VULGAR FRACTION ONE HALF
11/14	BE	U+017D	LATIN CAPITAL LETTER Z WITH CARON
11/15	BF	U+017E	LATIN SMALL LETTER Z WITH CARON

Table B.1 *(concluded)*

Bit combination	Hex	Identifier	Name
12/00	C0	U+00C0	LATIN CAPITAL LETTER A WITH GRAVE
12/01	C1	U+00C1	LATIN CAPITAL LETTER A WITH ACUTE
12/02	C2	U+00C2	LATIN CAPITAL LETTER A WITH CIRCUMFLEX
12/03	C3	U+00C3	LATIN CAPITAL LETTER A WITH TILDE
12/04	C4	U+00C4	LATIN CAPITAL LETTER A WITH DIAERESIS
12/05	C5	U+00C5	LATIN CAPITAL LETTER A WITH RING ABOVE
12/06	C6	U+00C6	LATIN CAPITAL LETTER AE
12/07	C7	U+00C7	LATIN CAPITAL LETTER C WITH CEDILLA
12/08	C8	U+00C8	LATIN CAPITAL LETTER E WITH GRAVE
12/09	C9	U+00C9	LATIN CAPITAL LETTER E WITH ACUTE
12/10	CA	U+00CA	LATIN CAPITAL LETTER E WITH CIRCUMFLEX
12/11	CB	U+00CB	LATIN CAPITAL LETTER E WITH DIAERESIS
12/12	CC	U+00CC	LATIN CAPITAL LETTER I WITH GRAVE
12/13	CD	U+00CD	LATIN CAPITAL LETTER I WITH ACUTE
12/14	CE	U+00CE	LATIN CAPITAL LETTER I WITH CIRCUMFLEX
12/15	CF	U+00CF	LATIN CAPITAL LETTER I WITH DIAERESIS
13/00	D0	U+00D0	LATIN CAPITAL LETTER ETH
13/01	D1	U+00D1	LATIN CAPITAL LETTER N WITH TILDE
13/02	D2	U+00D2	LATIN CAPITAL LETTER O WITH GRAVE
13/03	D3	U+00D3	LATIN CAPITAL LETTER O WITH ACUTE
13/04	D4	U+00D4	LATIN CAPITAL LETTER O WITH CIRCUMFLEX
13/05	D5	U+00D5	LATIN CAPITAL LETTER O WITH TILDE
13/06	D6	U+00D6	LATIN CAPITAL LETTER O WITH DIAERESIS
13/07	D7	U+00D7	MULTIPLICATION SIGN
13/08	D8	U+00D8	LATIN CAPITAL LETTER O WITH STROKE
13/09	D9	U+00D9	LATIN CAPITAL LETTER U WITH GRAVE
13/10	DA	U+00DA	LATIN CAPITAL LETTER U WITH ACUTE
13/11	DB	U+00DB	LATIN CAPITAL LETTER U WITH CIRCUMFLEX
13/12	DC	U+00DC	LATIN CAPITAL LETTER U WITH DIAERESIS
13/13	DD	U+00DD	LATIN CAPITAL LETTER Y WITH ACUTE
13/14	DE	U+00DE	LATIN CAPITAL LETTER THORN
13/15	DF	U+00DF	LATIN SMALL LETTER SHARP S
14/00	E0	U+00E0	LATIN SMALL LETTER A WITH GRAVE
14/01	E1	U+00E1	LATIN SMALL LETTER A WITH ACUTE
14/02	E2	U+00E2	LATIN SMALL LETTER A WITH CIRCUMFLEX
14/03	E3	U+00E3	LATIN SMALL LETTER A WITH TILDE
14/04	E4	U+00E4	LATIN SMALL LETTER A WITH DIAERESIS
14/05	E5	U+00E5	LATIN SMALL LETTER A WITH RING ABOVE
14/06	E6	U+00E6	LATIN SMALL LETTER AE
14/07	E7	U+00E7	LATIN SMALL LETTER C WITH CEDILLA
14/08	E8	U+00E8	LATIN SMALL LETTER E WITH GRAVE
14/09	E9	U+00E9	LATIN SMALL LETTER E WITH ACUTE
14/10	EA	U+00EA	LATIN SMALL LETTER E WITH CIRCUMFLEX
14/11	EB	U+00EB	LATIN SMALL LETTER E WITH DIAERESIS
14/12	EC	U+00EC	LATIN SMALL LETTER I WITH GRAVE
14/13	ED	U+00ED	LATIN SMALL LETTER I WITH ACUTE
14/14	EE	U+00EE	LATIN SMALL LETTER I WITH CIRCUMFLEX
14/15	EF	U+00EF	LATIN SMALL LETTER I WITH DIAERESIS
15/00	F0	U+00F0	LATIN SMALL LETTER ETH
15/01	F1	U+00F1	LATIN SMALL LETTER N WITH TILDE
15/02	F2	U+00F2	LATIN SMALL LETTER O WITH GRAVE
15/03	F3	U+00F3	LATIN SMALL LETTER O WITH ACUTE
15/04	F4	U+00F4	LATIN SMALL LETTER O WITH CIRCUMFLEX
15/05	F5	U+00F5	LATIN SMALL LETTER O WITH TILDE
15/06	F6	U+00F6	LATIN SMALL LETTER O WITH DIAERESIS
15/07	F7	U+00F7	DIVISION SIGN
15/08	F8	U+00F8	LATIN SMALL LETTER O WITH STROKE
15/09	F9	U+00F9	LATIN SMALL LETTER U WITH GRAVE
15/10	FA	U+00FA	LATIN SMALL LETTER U WITH ACUTE
15/11	FB	U+00FB	LATIN SMALL LETTER U WITH CIRCUMFLEX
15/12	FC	U+00FC	LATIN SMALL LETTER U WITH DIAERESIS
15/13	FD	U+00FD	LATIN TEKMALL LETTER Y WITH ACUTE
15/14	FE	U+00FE	LATIN SMALL LETTER THORN
15/15	FF	U+00FF	LATIN SMALL LETTER Y WITH DIAERESIS

B.3 Code table

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

The shaded positions in the code table correspond to bit combinations intended for control characters. The positions shown with cross-hatching correspond to bit combinations in table B.1 having the entry "This position not used".

Table B.2 – Code table of Windows-environment Sami character set

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	°	À	Ð	à	ð	0
0	0	0	1	01				!	1	A	Q	a	q		'	ğ	±	Á	Ñ	á	ñ	1
0	0	1	0	02				''	2	B	R	b	r	Č	'	Ǥ	Ǩ	Â	Ò	â	ò	2
0	0	1	1	03				#	3	C	S	c	s	ƒ	"	£	ǩ	Ã	Ó	ã	ó	3
0	1	0	0	04				$	4	D	T	d	t	č	"	¤	´	Ä	Ô	ä	ô	4
0	1	0	1	05				%	5	E	U	e	u	Ʒ	•	ǥ	µ	Å	Õ	å	õ	5
0	1	1	0	06				&	6	F	V	f	v	Ʒ	–	¦	¶	Æ	Ö	æ	ö	6
0	1	1	1	07				'	7	G	W	g	w	Ǯ	—	§	·	Ç	×	ç	÷	7
1	0	0	0	08				(	8	H	X	h	x	ǯ	đ	¨	Ŋ	È	Ø	è	ø	8
1	0	0	1	09				)	9	I	Y	i	y	Đ	Ǧ	©	ŋ	É	Ù	é	ù	9
1	0	1	0	10				*	:	J	Z	j	z	Š	š	Ȟ	Ŧ	Ê	Ú	ê	ú	A
1	0	1	1	11				+	;	K	[	k	{	‹	›	«	»	Ë	Û	ë	û	B
1	1	0	0	12				,	<	L	\	l	|	Œ	œ	¬	ŧ	Ì	Ü	ì	ü	C
1	1	0	1	13				–	=	M	]	m	}			SHY	½	Í	Ý	í	ý	D
1	1	1	0	14				.	>	N	^	n	~			®	Ž	Î	Þ	î	þ	E
1	1	1	1	15				/	?	O	_	o			Ÿ	ȟ	ž	Ï	ß	ï	ÿ	F
						0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	hex

This page intentionally left blank

Annex C
(normative)

Macintosh-environment Sami 8-bit coded graphic character set

C.1 Scope

This Annex defines a set of 223 coded graphic characters intended for Sami-language applications in Macintosh environments.

For Conformance, Definitions and Notations see Annex A, as applicable.

C.2 Characters of the set

See table C.1.

Table C.1 – Character set, coded representation

Bit combi-nation	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 C.1 *(continued)*

Bit combi-nation	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 C.1 *(continued)*

Bit combination	Hex	Identifier	Name
08/00	80	U+00C4	LATIN CAPITAL LETTER A WITH DIAERESIS
08/01	81	U+00C5	LATIN CAPITAL LETTER A WITH RING ABOVE
08/02	82	U+00C7	LATIN CAPITAL LETTER C WITH CEDILLA
08/03	83	U+00C9	LATIN CAPITAL LETTER E WITH ACUTE
08/04	84	U+00D1	LATIN CAPITAL LETTER N WITH TILDE
08/05	85	U+00D6	LATIN CAPITAL LETTER O WITH DIAERESIS
08/06	86	U+00DC	LATIN CAPITAL LETTER U WITH DIAERESIS
08/07	87	U+00E1	LATIN SMALL LETTER A WITH ACUTE
08/08	88	U+00E0	LATIN SMALL LETTER A WITH GRAVE
08/09	89	U+00E2	LATIN SMALL LETTER A WITH CIRCUMFLEX
08/10	8A	U+00E4	LATIN SMALL LETTER A WITH DIAERESIS
08/11	8B	U+00E3	LATIN SMALL LETTER A WITH TILDE
08/12	8C	U+00E5	LATIN SMALL LETTER A WITH RING ABOVE
08/13	8D	U+00E7	LATIN SMALL LETTER C WITH CEDILLA
08/14	8E	U+00E9	LATIN SMALL LETTER E WITH ACUTE
08/15	8F	U+00E8	LATIN SMALL LETTER E WITH GRAVE
09/00	90	U+00EA	LATIN SMALL LETTER E WITH CIRCUMFLEX
09/01	91	U+00EB	LATIN SMALL LETTER E WITH DIAERESIS
09/02	92	U+00ED	LATIN SMALL LETTER I WITH ACUTE
09/03	93	U+00EC	LATIN SMALL LETTER I WITH GRAVE
09/04	94	U+00EE	LATIN SMALL LETTER I WITH CIRCUMFLEX
09/05	95	U+00EF	LATIN SMALL LETTER I WITH DIAERESIS
09/06	96	U+00F1	LATIN SMALL LETTER N WITH TILDE
09/07	97	U+00F3	LATIN SMALL LETTER O WITH ACUTE
09/08	98	U+00F2	LATIN SMALL LETTER O WITH GRAVE
09/09	99	U+00F4	LATIN SMALL LETTER O WITH CIRCUMFLEX
09/10	9A	U+00F6	LATIN SMALL LETTER O WITH DIAERESIS
09/11	9B	U+00F5	LATIN SMALL LETTER O WITH TILDE
09/12	9C	U+00FA	LATIN SMALL LETTER U WITH ACUTE
09/13	9D	U+00F9	LATIN SMALL LETTER U WITH GRAVE
09/14	9E	U+00FB	LATIN SMALL LETTER U WITH CIRCUMFLEX
09/15	9F	U+00FC	LATIN SMALL LETTER U WITH DIAERESIS
10/00	A0	U+00DD	LATIN CAPITAL LETTER Y WITH ACUTE
10/01	A1	U+00B0	DEGREE SIGN
10/02	A2	U+010C	LATIN CAPITAL LETTER C WITH CARON
10/03	A3	U+00A3	POUND SIGN
10/04	A4	U+00A7	SECTION SIGN
10/05	A5	U+2022	BULLET
10/06	A6	U+00B6	PILCROW SIGN
10/07	A7	U+00DF	LATIN SMALL LETTER SHARP S
10/08	A8	U+00AE	REGISTERED SIGN
10/09	A9	U+00A9	COPYRIGHT SIGN
10/10	AA	U+2122	TRADE MARK SIGN
10/11	AB	U+00B4	ACUTE ACCENT
10/12	AC	U+00A8	DIAERESIS
10/13	AD	U+2260	NOT EQUAL TO
10/14	AE	U+00C6	LATIN CAPITAL LETTER AE
10/15	AF	U+00D8	LATIN CAPITAL LETTER O WITH STROKE
11/00	B0	U+0110	LATIN CAPITAL LETTER D WITH STROKE
11/01	B1	U+014A	LATIN CAPITAL LETTER ENG
11/02	B2	U+2264	LESS-THAN OR EQUAL TO
11/03	B3	U+2265	GREATER-THAN OR EQUAL TO
11/04	B4	U+0160	LATIN CAPITAL LETTER S WITH CARON
11/05	B5	U+0166	LATIN CAPITAL LETTER T WITH STROKE
11/06	B6	U+2202	PARTIAL DIFFERENTIAL
11/07	B7	U+017D	LATIN CAPITAL LETTER Z WITH CARON
11/08	B8	U+010D	LATIN SMALL LETTER C WITH CARON
11/09	B9	U+0111	LATIN SMALL LETTER D WITH STROKE
11/10	BA	U+014B	LATIN SMALL LETTER ENG
11/11	BB	U+0161	LATIN SMALL LETTER S WITH CARON
11/12	BC	U+0167	LATIN SMALL LETTER T WITH STROKE
11/13	BD	U+017E	LATIN SMALL LETTER Z WITH CARON
11/14	BE	U+00E6	LATIN SMALL LETTER AE
11/15	BF	U+00F8	LATIN SMALL LETTER O WITH STROKE

Table C.1 *(concluded)*

Bit combination	Hex	Identifier	Name
12/00	C0	U+00BF	INVERTED QUESTION MARK
12/01	C1	U+00A1	INVERTED EXCLAMATION MARK
12/02	C2	U+00AC	NOT SIGN
12/03	C3	U+221A	SQUARE ROOT
12/04	C4	U+0192	LATIN SMALL LETTER F WITH HOOK (Florin sign)
12/05	C5	U+2248	ALMOST EQUAL TO
12/06	C6	U+021E	LATIN CAPITAL LETTER H WITH CARON
12/07	C7	U+00AB	LEFT-POINTING DOUBLE ANGLE QUOTATION MARK
12/08	C8	U+00BB	RIGHT-POINTING DOUBLE ANGLE QUOTATION MARK
12/09	C9	U+2026	HORIZONTAL ELLIPSIS
12/10	CA	U+00A0	NO-BREAK SPACE
12/11	CB	U+00C0	LATIN CAPITAL LETTER A WITH GRAVE
12/12	CC	U+00C3	LATIN CAPITAL LETTER A WITH TILDE
12/13	CD	U+00D5	LATIN CAPITAL LETTER O WITH TILDE
12/14	CE	U+0152	LATIN CAPITAL LIGATURE OE
12/15	CF	U+0153	LATIN SMALL LIGATURE OE
13/00	D0	U+2013	EN DASH
13/01	D1	U+2014	EM DASH
13/02	D2	U+201C	LEFT DOUBLE QUOTATION MARK
13/03	D3	U+201D	RIGHT DOUBLE QUOTATION MARK
13/04	D4	U+2018	LEFT SINGLE QUOTATION MARK
13/05	D5	U+2019	RIGHT SINGLE QUOTATION MARK
13/06	D6	U+00F7	DIVISION SIGN
13/07	D7	U+021F	LATIN SMALL LETTER H WITH CARON
13/08	D8	U+00FF	LATIN SMALL LETTER Y WITH DIAERESIS
13/09	D9	U+0178	LATIN CAPITAL LETTER Y WITH DIAERESIS
13/10	DA	U+2044	FRACTION SLASH
13/11	DB	U+20AC	EURO SIGN
13/12	DC	U+00D0	LATIN CAPITAL LETTER ETH
13/13	DD	U+00F0	LATIN SMALL LETTER ETH
13/14	DE	U+00DE	LATIN CAPITAL LETTER THORN
13/15	DF	U+00FE	LATIN SMALL LETTER THORN
14/00	E0	U+00FD	LATIN SMALL LETTER Y WITH ACUTE
14/01	E1	U+00B7	MIDDLE DOT
14/02	E2	U+201A	SINGLE LOW-9 QUOTATION MARK
14/03	E3	U+201E	DOUBLE LOW-9 QUOTATION MARK
14/04	E4	U+2030	PER MILLE SIGN
14/05	E5	U+00C2	LATIN CAPITAL LETTER A WITH CIRCUMFLEX
14/06	E6	U+00CA	LATIN CAPITAL LETTER E WITH CIRCUMFLEX
14/07	E7	U+00C1	LATIN CAPITAL LETTER A WITH ACUTE
14/08	E8	U+00CB	LATIN CAPITAL LETTER E WITH DIAERESIS
14/09	E9	U+00C8	LATIN CAPITAL LETTER E WITH GRAVE
14/10	EA	U+00CD	LATIN CAPITAL LETTER I WITH ACUTE
14/11	EB	U+00CE	LATIN CAPITAL LETTER I WITH CIRCUMFLEX
14/12	EC	U+00CF	LATIN CAPITAL LETTER I WITH DIAERESIS
14/13	ED	U+00CC	LATIN CAPITAL LETTER I WITH GRAVE
14/14	EE	U+00D3	LATIN CAPITAL LETTER O WITH ACUTE
14/15	EF	U+00D4	LATIN CAPITAL LETTER O WITH CIRCUMFLEX
15/00	F0		APPLE LOGOTYPE
15/01	F1	U+00D2	LATIN CAPITAL LETTER O WITH GRAVE
15/02	F2	U+00DA	LATIN CAPITAL LETTER U WITH ACUTE
15/03	F3	U+00DB	LATIN CAPITAL LETTER U WITH CIRCUMFLEX
15/04	F4	U+00D9	LATIN CAPITAL LETTER U WITH GRAVE
15/05	F5	U+0131	LATIN SMALL LETTER DOTLESS I
15/06	F6	U+01B7	LATIN CAPITAL LETTER EZH
15/07	F7	U+0292	LATIN SMALL LETTER EZH
15/08	F8	U+01EE	LATIN CAPITAL LETTER EZH WITH CARON
15/09	F9	U+01EF	LATIN SMALL LETTER EZH WITH CARON
15/10	FA	U+01E4	LATIN CAPITAL LETTER G WITH STROKE
15/11	FB	U+01E5	LATIN SMALL LETTER G WITH STROKE
15/12	FC	U+01E6	LATIN CAPITAL LETTER G WITH CARON
15/13	FD	U+01E7	LATIN SMALL LETTER G WITH CARON
15/14	FE	U+01E8	LATIN CAPITAL LETTER K WITH CARON
15/15	FF	U+01E9	LATIN SMALL LETTER K WITH CARON

C.3 Code table

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

The shaded positions in the code table correspond to bit combinations intended for control characters.

Table C.2 – Code table of Macintosh-environment Sami character set

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	Ä	ê	Ý	Đ	¿	–	ý	APPLE LOGO	0
0	0	0	1	01			!	1	A	Q	a	q	Å	ë	°	Ŋ	¡	—	▪	Ò	1
0	0	1	0	02			"	2	B	R	b	r	Ç	í	Č	≤	¬	"	‚	Ú	2
0	0	1	1	03			#	3	C	S	c	s	É	ì	£	≥	√	"	„	Û	3
0	1	0	0	04			$	4	D	T	d	t	Ñ	î	§	Š	ƒ	'	‰	Ù	4
0	1	0	1	05			%	5	E	U	e	u	Ö	ï	•	Ŧ	≈	'	Â	⌐	5
0	1	1	0	06			&	6	F	V	f	v	Ü	ñ	¶	∂	Ȟ	÷	Ê	ʒ	6
0	1	1	1	07			'	7	G	W	g	w	á	ó	ß	Ž	«	ȟ	Á	ɜ	7
1	0	0	0	08			(	8	H	X	h	x	à	ò	®	č	»	ÿ	Ë	Ǯ	8
1	0	0	1	09			)	9	I	Y	i	y	â	ô	©	đ	…	Ÿ	È	ǯ	9
1	0	1	0	10			*	:	J	Z	j	z	ä	ö	™	ŋ	NBSP	∕	Í	Ǥ	A
1	0	1	1	11			+	;	K	[	k	{	ã	õ	´	š	À	€	Î	ǥ	B
1	1	0	0	12			'	<	L	\	l	|	å	ú	¨	ŧ	Ã	Ð	Ï	Ǧ	C
1	1	0	1	13			–	=	M	]	m	}	ç	ù	≠	ž	Õ	ð	Ì	ǧ	D
1	1	1	0	14			.	>	N	^	n	~	é	û	Æ	æ	Œ	Þ	Ó	Ǩ	E
1	1	1	1	15			/	?	O	_	o		è	ü	Ø	ø	œ	þ	Ô	ǩ	F
					0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	hex

Annex D
(informative)

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.