MUFI character recommendation

Characters in the official Unicode Standard and in the Private Use Area
for Medieval texts written in the Latin alphabet

Ᵽ ※ ð þ þ å ɣ ƀ ꝓ ꝁ ꝯ ꬱ ꝣ ʒ ꜵ ǿ Ꜳ ᵹ ǿ ꝑ ù ẏ ꝺ þ ꞇ ꝛ Ᵽ ﹕ ※ ¶

Editor

Odd Einar Haugen, University of Bergen, Norway

Background & contributors

The present recommendation is a major revision and extension of two proposals published on the MUFI web site:

◗ A proposal for subranges within the Private Use Area of Unicode (15 June 2002)
◗ A proposal for supplementary characters in Unicode (5 February 2003)

These proposals received a number of helpful comments from (in alphabetical order) Jim Allan (U.S.), Deborah W. Anderson (Berkeley, CA), Peter S. Baker (Charlottesville, VA), Michael Beddow (Leeds), António H.A. Emiliano (Lisboa), Michael Everson (Dublin), Jost Gippert (Frankfurt am Main), Juan-José Marcos (Plasencia, Spain), Rick McGowan (Unicode consortium), Susana T. Pedro (Lisboa), David J. Perry (Rye, NY), Gerhard Schumacher (Köln), Ken Whistler (Unicode consortium) and Christian Wittern (Kyoto).

The present recommendation is the result of a meeting held in Bergen, Norway, 30–31 August 2003. Participants at this meeting were Odd Einar Haugen (Bergen), Michael Everson (Dublin), Michael Irlenbusch (Bergen), Alec McAllister (Leeds), Gerhard Schumacher (Köln), and Tarrin Wills (Sydney). The recommendation was published 8 September 2003 on the MUFI site for public review and has been discussed electronically within the MUFI group in the ensuing months.

Many aspects of this recommendation will be controversial, and more than one of the contributors and advisors listed above may disagree with the solutions chosen in the recommendation. It is, however, clear that this recommendation would have progressed much more slowly and would have been much inferior had it not been for all the help and comments received. The editor would like to thank all contributors for their assistance in this project.

Publication details

Contents

Introduction

This recommendation contains a set of characters for the use of medievalists and, to some extent, classicists. The aim of the recommendation is (1) to establish a selection of characters from the official Unicode Standard and (2) to coordinate the usage of code points in the Private Use Area.

The recommendation contains a representative glyph for each character, an entity name, a Unicode code point, and a descriptive name.

Glyphs

The glyphs used in this recommendation are for guidance only. They are based on the font *Alphabetum* by Juan-José Marcos (Plasencia, Spain).

Entity names

Entities are used in numerous encoding schemes such as SGML/XML. For the sake of transparency and interchangeability, it is recommended that entities as far as possible conform to the standard ISO entity sets.

In addition to the ISO entities, a number of entities for characters not designated in this standard are needed. This recommendation uses the syntax and inventory defined in *The Menota handbook*, ch. 2 and 5 (www.hit.uib.no/menota/guidelines), summarised in the table below.

Base line character	main type	variant	ligature	fixed modification	loose modification
A	anc	brk	lig	bar	acute
a	comb	close	ligred	baracr	breve
B	enl	dotless		barbl	circ
b	ins	high		des	curl
C	run	long		flour	dblac
c	scap	med		hook	dot
	unc	open		ogon	dotbl
etc.		rot		ovl	grave
		squ		slash	macr
				strok	sup
					uml

Note that not all slots need to be filled in; in most cases only one or two slots are used in addition to the base line character. The meaning of the various parts is explained in this table, listed in alphabetical order:

acute = acute accent	close = closed form	enl = enlarged	long = long form	scap = small capital
anc = anceps	comb = combining	flour = flourish	macr = macron (long)	slash = diagonal stroke
bar = bar above	curl = curl (hook above)	grave = grave accent	med = medium-high position	squ = square form
baracr = bar across	dblac = double acute accent	high = high position	ogon = ogonek (hook below)	strok = horizontal stroke
barbl = bar below	des = descending	hook = hook above	open = open form	sup = superscript
brev = breve (short)	dot = dot above	ins = Insular	ovl = overline	uml = Umlaut (diaeresis)
brk = broken	dotbl = dot below	lig = ligature	rot = rotunda	unc = Uncial
circ = circumflex	dotless = no dot above	ligred = ligature reduced form	run = Runic	

Some entities, especially for non-alphabetic characters in *Basic Latin*, *Latin-1 Supplement* and *General Punctuation*, have more compressed names, modelled on already existing entity names.

An updated list of ISO conformant entities can be found at the Oasis web site: www.oasis-open.org/cover/xml-ISOents.txt

Code points

Code points are given in the hexadecimal format and refer to the Unicode Standard version 4.0. The allocation of code points in the Private Use Area has been coordinated with several font projects (notably Titus and Junicode) with the aim of establishing a common usage. Further information can be found on the MUFI web site (www.hit.uib.no/mufi).

Code points in the Unicode Standard are usually given with the prefix 'U+', e.g. U+003A for the colon, U+003B for the semicolon, etc. For reasons of brevity, this prefix has been left out in this recommendation.

All code points are located in the Basic Multilingual Plane (BMP) of the Unicode Standard. The Private Use Area of the BMP has a total of 6,400 slots. A large number of PUA characters in this recommendation are concentrated in the section F100–F3FF, but many are also to be found in the section E000–F0FF, coordinated with the Titus project. Metrical characters are coordinated with the Alphabetum font and located in the F700–F76F section.

Characters in the PUA have been divided into subranges. Note that the allocation of code points is not necessarily contiguous.

Since Titus and Junicode were already established when this recommendation was made, some code point conflicts were inevitable. This recommendation gives a single code point for every character in the PUA, but lists duplicate code points in other fonts, notably Junicode, where relevant.

Characters in the PUA should be decommissioned as soon as they can be encoded with ordinary Unicode characters. The MUFI web site has further information on this.

Descriptive names

Each character has been given a descriptive name, as far as possible according to the rules in the Unicode Standard. *The Menota handbook* ch. 2 has further details on the interpretation of the Unicode naming rules.

In a few cases, an alternative MUFI descriptive name has been introduced. Thus, the character located at 223B in *Mathematical Operators* (p. 49 below) has the name HOMOTHETIC in the Unicode Standard, but since this character could be used for a Latin abbreviation (meaning 'est'), a MUFI descriptive name has been listed, LATIN ABBREVIATION SIGN EST.

MUFI compliant fonts

Font designers are encouraged to design fonts which comply with the code point allocations in this recommendation. A list of MUFI compliant fonts will be available on the MUFI web site.

The Unicode Standard

The Unicode Standard can be accessed at www.unicode org. This site contains a wealth of information on the standard, including code charts in PDF for all ranges.

A: Characters in the Unicode Standard

1: Basic Latin

All 96 characters in this chart have been selected. Characters in *Basic Latin* may be encoded without the use of entities, but entities have been defined for characters other than the digits 0–9 and the alphabetic characters A–Z and a–z .

Glyph	ISO entity	MUFI entity	Code point	Unicode descriptive name	MUFI descriptive name
		&sp;	0020	SPACE	
!	!		0021	EXCLAMATION MARK	
"	"		0022	QUOTATION MARK	
In SGML/XML encoded documents, the quotation mark is used to delimit attributes. The entity '"' refers to the quotation mark as such.					
#	#		0023	NUMBER SIGN	
$	$		0024	DOLLAR SIGN	
%	%		0025	PERCENT SIGN	
&	&	&ampet;	0026	AMPERSAND	LATIN ABBREVIATION SIGN AMPERSAND
In SGML/XML encoded documents, the ampersand is used as the opening character for an entity name, as exemplified throughout this recommendation. The ampersand as such should be encoded with the entity '&'. In Medieval primary sources, the ampersand was used as an abbreviation for 'et' (and). An entity such as '&ampet;' might be used to indicate that this sign is a specific Medieval abbreviation mark.					

'	'	&apo;	0027	APOSTROPHE	LATIN ABBREVIATION SIGN APOSTROPHE

In SGML/XML encoded documents, the apostrophe may be used in a technical sense. To be able to refer to the apostrophe as such, a separate entity is needed, '''. – A sign looking like the apostrophe was used as an abbreviation for 'i' or 'e'. Cf. D.A. Seip, *Palæografi: Norge og Island* (Nordisk kultur 23:B), Oslo etc. 1954, p. 125. In such cases, the encoder might want use an entity like '&apo;'.

(	(		0028	LEFT PARENTHESIS	
)	)		0029	RIGHT PARENTHESIS	
*	*		002A	ASTERISK	
+	+		002B	PLUS SIGN	
,	,		002C	COMMA	
–	‐		002D	HYPHEN-MINUS	

As the name HYPHEN-MINUS indicates, this character is ambiguous. For an unambiguous *hyphen* character, see 2010 HYPHEN in *General Punctuation* (p. 45 below), and for an unambiguous *minus* character, see 2212 MINUS in *Mathematical Operators* (p. 49).

.	.	˙	002E	FULL STOP	LATIN ABBREVIATION SIGN SPACING DOT

The full stop is often used as a mark of abbreviation, typically in suspended (truncated) words. An entity such as '˙' might be used to indicate that it is an abbreviation mark rather than a punctuation mark.

/	/		002F	SOLIDUS	

0			0030	DIGIT ZERO	
1			0031	DIGIT ONE	
2			0032	DIGIT TWO	
3			0033	DIGIT THREE	
4			0034	DIGIT FOUR	
5			0035	DIGIT FIVE	
6			0036	DIGIT SIX	
7			0037	DIGIT SEVEN	
8			0038	DIGIT EIGHT	
9			0039	DIGIT NINE	
:	:		003A	COLON	
;	;	&ed;	003B	SEMICOLON	LATIN ABBREVIATION SIGN ED

In Medieval script, a sign looking like the semicolon was commonly used as an abbreviation for 'eð' or 'ed'. In such cases, the encoder might want to use an entity like '&ed;' to indicate that this sign is a specific Medieval abbreviation mark.

| < | < | | 003C | LESS-THAN SIGN | |

In SGML/XML encoded doccuments, the less-than sign is used as the opening character of an element name, e.g. <text>. To be able to refer to the less-than sign as such, a separate entity is needed, '<'.

| = | = | | 003D | EQUALS SIGN | |
| > | > | | 003E | GREATER-THAN SIGN | |

In SGML/XML encoded doccuments, the greater-than sign is used as a closing character of an element name, e.g. <text>. To be able to refer to the greater-than sign as such, a separate entity is recommended, '>'.

?	?		003F	QUESTION MARK	
@	@		0040	COMMERCIAL AT	
A			0041	LATIN CAPITAL LETTER A	
B			0042	LATIN CAPITAL LETTER B	
C			0043	LATIN CAPITAL LETTER C	
D			0044	LATIN CAPITAL LETTER D	
E			0045	LATIN CAPITAL LETTER E	
F			0046	LATIN CAPITAL LETTER F	

G			0047	LATIN CAPITAL LETTER G	
H			0048	LATIN CAPITAL LETTER H	
I			0049	LATIN CAPITAL LETTER I	
J			004A	LATIN CAPITAL LETTER J	
K			004B	LATIN CAPITAL LETTER K	
L			004C	LATIN CAPITAL LETTER L	
M			004D	LATIN CAPITAL LETTER M	
N			004E	LATIN CAPITAL LETTER N	
O			004F	LATIN CAPITAL LETTER O	
P			0050	LATIN CAPITAL LETTER P	
Q			0051	LATIN CAPITAL LETTER Q	
R			0052	LATIN CAPITAL LETTER R	
S			0053	LATIN CAPITAL LETTER S	

T			0054	LATIN CAPITAL LETTER T	
U			0055	LATIN CAPITAL LETTER U	
V			0056	LATIN CAPITAL LETTER V	
W			0057	LATIN CAPITAL LETTER W	
X			0058	LATIN CAPITAL LETTER X	
Y			0059	LATIN CAPITAL LETTER Y	
Z			005A	LATIN CAPITAL LETTER Z	
[	[		005B	LEFT SQUARE BRACKET	
\	\		005C	REVERSE SOLIDUS	
]	]		005D	RIGHT SQUARE BRACKET	
^	ˆ		005E	CIRCUMFLEX ACCENT	
_	_		005F	LOW LINE	
`	`		0060	GRAVE ACCENT	

a			0061	LATIN SMALL LETTER A	
b			0062	LATIN SMALL LETTER B	
c			0063	LATIN SMALL LETTER C	
d			0064	LATIN SMALL LETTER D	
e			0065	LATIN SMALL LETTER E	
f			0066	LATIN SMALL LETTER F	
g			0067	LATIN SMALL LETTER G	
h			0068	LATIN SMALL LETTER H	
i			0069	LATIN SMALL LETTER I	
j			006A	LATIN SMALL LETTER J	
k			006B	LATIN SMALL LETTER K	
l			006C	LATIN SMALL LETTER L	
m			006D	LATIN SMALL LETTER M	

n			006E	LATIN SMALL LETTER N	
o			006F	LATIN SMALL LETTER O	
p			0070	LATIN SMALL LETTER P	
q			0071	LATIN SMALL LETTER Q	
r			0072	LATIN SMALL LETTER R	
s			0073	LATIN SMALL LETTER S	
t			0074	LATIN SMALL LETTER T	
u			0075	LATIN SMALL LETTER U	
v			0076	LATIN SMALL LETTER V	
w			0077	LATIN SMALL LETTER W	
x			0078	LATIN SMALL LETTER X	
y			0079	LATIN SMALL LETTER Y	
z			007A	LATIN SMALL LETTER Z	

Glyph		ISO entity	MUFI entity	Code point	Unicode descriptive name	
{	{			007B	LEFT CURLY BRACKET	
|	|			007C	VERTICAL LINE	
}	}			007D	RIGHT CURLY BRACKET	
~			&tld;	007E	TILDE	
			&del;	007F	DELETE	

2: Latin-1 Supplement

All 96 characters in this range have been selected, although many of them are not used in Medieval script. However, these characters are so frequently used in other contexts that it is advisable to include the whole range in a Medieval font. For this and all ensuing code charts entities are being used in many SGML/XML documents.

Glyph	ISO entity	MUFI entity	Code point	Unicode descriptive name
			00A0	NO-BREAK SPACE
¡	¡		00A1	INVERTED EXCLAMATION MARK
¢	¢		00A2	CENT SIGN
£	£		00A3	POUND SIGN

¤	¤		00A4	CURRENCY SIGN
¥	¥		00A5	YEN SIGN
¦	¦		00A6	BROKEN BAR
§	§		00A7	SECTION SIGN
¨	¨		00A8	DIAERESIS
The entities '¨' and '&Ddot;' are also being used for this character.				
©	©		00A9	COPYRIGHT SIGN
a	ª		00AA	FEMININE ORDINAL INDICATOR
«	«		00AB	LEFT-POINTING DOUBLE ANGLE QUOTATION MARK
¬	¬		00AC	NOT SIGN
	­		00AD	SOFT HYPHEN
®	®		00AE	REGISTERED SIGN
¯	¯		00AF	MACRON

°	°		00B0	DEGREE SIGN
±	±		00B1	PLUS-MINUS SIGN
2	²		00B2	SUPERSCRIPT TWO
3	³		00B3	SUPERSCRIPT THREE
´	´		00B4	ACUTE ACCENT
μ	µ		00B5	MICRO SIGN
¶	¶		00B6	PILCROW SIGN
·	·		00B7	MIDDLE DOT
¸	¸		00B8	CEDILLA
1	¹		00B9	SUPERSCRIPT ONE
º	º		00BA	MASCULINE ORDINAL INDICATOR
»	»		00BB	RIGHT-POINTING DOUBLE ANGLE QUOTATION MARK
¼	¼		00BC	VULGAR FRACTION ONE QUARTER

½	½		00BD	VULGAR FRACTION ONE HALF
The entity '½' is also being used for this character.				
¾	¾		00BE	VULGAR FRACTION THREE QUARTERS
¿	¿		00BF	INVERTED QUESTION MARK
À	À		00C0	LATIN CAPITAL LETTER A WITH GRAVE
Á	Á		00C1	LATIN CAPITAL LETTER A WITH ACUTE
Â	Â		00C2	LATIN CAPITAL LETTER A WITH CIRCUMFLEX
Ã	Ã		00C3	LATIN CAPITAL LETTER A WITH TILDE
Ä	Ä		00C4	LATIN CAPITAL LETTER A WITH DIAERESIS
Å	Å		00C5	LATIN CAPITAL LETTER A WITH RING ABOVE
Æ	Æ		00C6	LATIN CAPITAL LETTER AE
Ç	Ç		00C7	LATIN CAPITAL LETTER C WITH CEDILLA
È	È		00C8	LATIN CAPITAL LETTER E WITH GRAVE

É	É		00C9	LATIN CAPITAL LETTER E WITH ACUTE
Ê	Ê		00CA	LATIN CAPITAL LETTER E WITH CIRCUMFLEX
Ë	Ë		00CB	LATIN CAPITAL LETTER E WITH DIAERESIS
Ì	Ì		00CC	LATIN CAPITAL LETTER I WITH GRAVE
Í	Í		00CD	LATIN CAPITAL LETTER I WITH ACUTE
Î	Î		00CE	LATIN CAPITAL LETTER I WITH CIRCUMFLEX
Ï	Ï		00CF	LATIN CAPITAL LETTER I WITH DIAERESIS
Ð	Ð		00D0	LATIN CAPITAL LETTER ETH
Ñ	Ñ		00D1	LATIN CAPITAL LETTER N WITH TILDE
Ò	Ò		00D2	LATIN CAPITAL LETTER O WITH GRAVE
Ó	Ó		00D3	LATIN CAPITAL LETTER O WITH ACUTE
Ô	Ô		00D4	LATIN CAPITAL LETTER O WITH CIRCUMFLEX
Õ	Õ		00D5	LATIN CAPITAL LETTER O WITH TILDE

Ö	Ö		00D6	LATIN CAPITAL LETTER O WITH DIAERESIS
×	×		00D7	MULTIPLICATION SIGN
Ø	Ø		00D8	LATIN CAPITAL LETTER O WITH STROKE
Ù	Ù		00D9	LATIN CAPITAL LETTER U WITH GRAVE
Ú	Ú		00DA	LATIN CAPITAL LETTER U WITH ACUTE
Û	Û		00DB	LATIN CAPITAL LETTER U WITH CIRCUMFLEX
Ü	Ü		00DC	LATIN CAPITAL LETTER U WITH DIAERESIS
Ý	Ý		00DD	LATIN CAPITAL LETTER Y WITH ACUTE
Þ	Þ		00DE	LATIN CAPITAL LETTER THORN
ß	ß		00DF	LATIN SMALL LETTER SHARP S
à	à		00E0	LATIN SMALL LETTER A WITH GRAVE
á	á		00E1	LATIN SMALL LETTER A WITH ACUTE
â	â		00E2	LATIN SMALL LETTER A WITH CIRCUMFLEX

ã	ã		00E3	LATIN SMALL LETTER A WITH TILDE
ä	ä		00E4	LATIN SMALL LETTER A WITH DIAERESIS
å	å		00E5	LATIN SMALL LETTER A WITH RING
æ	æ		00E6	LATIN SMALL LETTER AE
ç	ç		00E7	LATIN SMALL LETTER C WITH CEDILLA
è	è		00E8	LATIN SMALL LETTER E WITH GRAVE
é	é		00E9	LATIN SMALL LETTER E WITH ACUTE
ê	ê		00EA	LATIN SMALL LETTER E WITH CIRCUMFLEX
ë	ë		00EB	LATIN SMALL LETTER E WITH DIAERESIS
ì	ì		00EC	LATIN SMALL LETTER I WITH GRAVE
í	í		00ED	LATIN SMALL LETTER I WITH ACUTE
î	î		00EE	LATIN SMALL LETTER I WITH CIRCUMFLEX
ï	ï		00EF	LATIN SMALL LETTER I WITH DIAERESIS

ð	ð		00F0	LATIN SMALL LETTER ETH
ñ	ñ		00F1	LATIN SMALL LETTER N WITH TILDE
ò	ò		00F2	LATIN SMALL LETTER O WITH GRAVE
ó	ó		00F3	LATIN SMALL LETTER O WITH ACUTE
ô	ô		00F4	LATIN SMALL LETTER O WITH CIRCUMFLEX
õ	õ		00F5	LATIN SMALL LETTER O WITH TILDE
ö	ö		00F6	LATIN SMALL LETTER O WITH DIAERESIS
÷	÷		00F7	DIVISION SIGN
ø	ø		00F8	LATIN SMALL LETTER O WITH STROKE
ù	ù		00F9	LATIN SMALL LETTER U WITH GRAVE
ú	ú		00FA	LATIN SMALL LETTER U WITH ACUTE
û	û		00FB	LATIN SMALL LETTER U WITH CIRCUMFLEX
ü	ü		00FC	LATIN SMALL LETTER U WITH DIAERESIS

ý	ý		00FD	LATIN SMALL LETTER Y WITH ACUTE
þ	þ		00FE	LATIN SMALL LETTER THORN
ÿ	ÿ		00FF	LATIN SMALL LETTER Y WITH DIAERESIS

3: Latin Extended-A

In Unicode 4.0 this range has 128 characters. The selection below contains 50 characters, mostly precomposed characters with macron, breve, stroke, dot above, and ogonek.

Glyph	ISO entity	MUFI entity	Code point	Unicode descriptive name
Ā	Ā		0100	LATIN CAPITAL LETTER A WITH MACRON
ā	ā		0101	LATIN SMALL LETTER A WITH MACRON
Ă	Ă		0102	LATIN CAPITAL LETTER A WITH BREVE
ă	ă		0103	LATIN SMALL LETTER A WITH BREVE
Ą	Ą		0104	LATIN CAPITAL LETTER A WITH OGONEK
ą	ą		0105	LATIN SMALL LETTER A WITH OGONEK

Ċ	Ċ		010A	LATIN CAPITAL LETTER C WITH DOT ABOVE
ċ	ċ		010B	LATIN SMALL LETTER C WITH DOT ABOVE
Đ	Đ		0110	LATIN CAPITAL LETTER D WITH STROKE
đ	đ		0111	LATIN SMALL LETTER D WITH STROKE
Ē	Ē		0112	LATIN CAPITAL LETTER E WITH MACRON
ē	ē		0113	LATIN SMALL LETTER E WITH MACRON
Ĕ		&Ebreve;	0114	LATIN CAPITAL LETTER E WITH BREVE
ĕ		&ebreve;	0115	LATIN SMALL LETTER E WITH BREVE
Ė	Ė		0116	LATIN CAPITAL LETTER E WITH DOT ABOVE
ė	ė		0117	LATIN SMALL LETTER E WITH DOT ABOVE
Ę	Ę		0118	LATIN CAPITAL LETTER E WITH OGONEK
ę	ę		0119	LATIN SMALL LETTER E WITH OGONEK
Ġ	Ġ		0120	LATIN CAPITAL LETTER G WITH DOT ABOVE

ġ	ġ		0121	LATIN SMALL LETTER G WITH DOT ABOVE
ħ	ħ		0127	LATIN SMALL LETTER H WITH STROKE
Ī	Ī		012A	LATIN CAPITAL LETTER I WITH MACRON
ī	ī		012B	LATIN SMALL LETTER I WITH MACRON
Ĭ		&Ibreve;	012C	LATIN CAPITAL LETTER I WITH BREVE
ĭ		&ibreve;	012D	LATIN SMALL LETTER I WITH BREVE
Į	Į		012E	LATIN CAPITAL LETTER I WITH OGONEK
į	į		012F	LATIN SMALL LETTER I WITH OGONEK
ı	ı		0131	LATIN SMALL LETTER DOTLESS I
IJ	Ĳ		0132	LATIN CAPITAL LIGATURE IJ
ij	ĳ		0133	LATIN SMALL LIGATURE IJ
Ŋ	Ŋ		014A	LATIN CAPITAL LETTER ENG
ŋ	ŋ		014B	LATIN SMALL LETTER ENG

Ō	Ō		014C	LATIN CAPITAL LETTER O WITH MACRON
ō	ō		014D	LATIN SMALL LETTER O WITH MACRON
Ŏ		&Obreve;	014E	LATIN CAPITAL LETTER O WITH BREVE
ŏ		&obreve;	014F	LATIN SMALL LETTER O WITH BREVE
Ő	Ő		0150	LATIN CAPITAL LETTER O WITH DOUBLE ACUTE
ő	ő		0151	LATIN SMALL LETTER O WITH DOUBLE ACUTE
Œ	Œ		0152	LATIN CAPITAL LIGATURE OE
œ	œ		0153	LATIN SMALL LIGATURE OE
Ū	Ū		016A	LATIN CAPITAL LETTER U WITH MACRON
ū	ū		016B	LATIN SMALL LETTER U WITH MACRON
Ŭ	Ŭ		016C	LATIN CAPITAL LETTER U WITH BREVE
ŭ	ŭ		016D	LATIN SMALL LETTER U WITH BREVE
Ű	Ű		0170	LATIN CAPITAL LETTER U WITH DOUBLE ACUTE

ű	ű		0171	LATIN SMALL LETTER U WITH DOUBLE ACUTE
Ų	Ų		0172	LATIN CAPITAL LETTER U WITH OGONEK
ų	ų		0173	LATIN SMALL LETTER U WITH OGONEK
Ÿ	Ÿ		0178	LATIN CAPITAL LETTER Y WITH DIAERESIS
ſ		&slong;	017F	LATIN SMALL LETTER LONG S

4: Latin Extended-B

In Unicode 4.0 this range has 179 characters, of which 28 characters have been selected. Some of these have acute accent, macron, dot above, or ogonek, others are individual base characters, including Gothic and Old English ones.

Glyph	ISO entity	MUFI entity	Code point	Unicode descriptive name	MUFI descriptive name
ƀ		&bstrok;	0180	LATIN SMALL LETTER B WITH STROKE	
ƕ		&hwair;	0195	LATIN SMALL LETTER HV	LATIN SMALL LETTER HWAIR
ƙ		&kslonglig;	0199	LATIN SMALL LETTER K WITH HOOK	LATIN SMALL LIGATURE K AND LONG S

ł		&lbar;	019A	LATIN SMALL LETTER L WITH BAR
η		&nlrleg;	019E	LATIN SMALL LETTER N WITH LONG RIGHT LEG
Ꞧ		&YR;	01A6	LATIN LETTER YR

This is the capital form of the character and seldom used. For the small form, the small capital R should be used, cf. 0280 on p. 30 below.

Ƶ		&Zstrok;	01B5	LATIN CAPITAL LETTER Z WITH STROKE
ƶ		&zstrok;	01B6	LATIN SMALL LETTER Z WITH STROKE
Ʒ		&EZH;	01B7	LATIN CAPITAL LETTER EZH
ƿ		&wynn;	01BF	LATIN LETTER WYNN
Ǣ		&AEligmacr;	01E2	LATIN CAPITAL LETTER AE WITH MACRON
ǣ		&aeligmacr;	01E3	LATIN SMALL LETTER AE WITH MACRON
Ǫ		&Oogon;	01EA	LATIN CAPITAL LETTER O WITH OGONEK
ǫ		&oogon;	01EB	LATIN SMALL LETTER O WITH OGONEK
Ƕ		&HWAIR;	01F6	LATIN CAPITAL LETTER HWAIR

Ƿ		&WYNN;	01F7	LATIN CAPITAL LETTER WYNN	
Ǽ		&AEligacute;	01FC	LATIN CAPITAL LETTER AE WITH ACUTE	
ǽ		&aeligacute:	01FD	LATIN SMALL LETTER AE WITH ACUTE	
Ǿ		&Oslashacute;	01FE	LATIN CAPITAL LETTER O WITH STROKE AND ACUTE	
ǿ		&oslashacute;	01FF	LATIN SMALL LETTER O WITH STROKE AND ACUTE	
Ȝ		&YOGH;	021C	LATIN CAPITAL LETTER YOGH	
ȝ		&yogh;	021D	LATIN SMALL LETTER YOGH	
Ȧ		&Adot;	0226	LATIN CAPITAL LETTER A WITH DOT ABOVE	
ȧ		&adot;	0227	LATIN SMALL LETTER A WITH DOT ABOVE	
Ȯ		&Odot;	022E	LATIN CAPITAL LETTER O WITH DOT ABOVE	
ȯ		⊙	022F	LATIN SMALL LETTER O WITH DOT ABOVE	
Ȳ		&Ymacr;	0232	LATIN CAPITAL LETTER Y WITH MACRON	
ȳ		&ymacr;	0233	LATIN SMALL LETTER Y WITH MACRON	

5: IPA Extensions

In Unicode 4.0 this range has 94 characters. The selection below contains 15 characters, mostly small capitals. Additional small capitals are located in *Phonetic Extensions*, p. 36 below, and PUA subrange 2 *Small capitals*, p. 53 below.

Glyph	ISO entity	MUFI entity	Code point	Unicode descriptive name	MUFI descriptive name
ɔ		&con;	0254	LATIN SMALL LETTER OPEN O	LATIN ABBREVIATION SIGN CON
ɖ		&dtail;	0256	LATIN SMALL LETTER D WITH TAIL	
ɟ		&jdtlstrok;	025F	LATIN SMALL LETTER DOTLESS J WITH STROKE	
ɢ		&gscap;	0262	LATIN LETTER SMALL CAPITAL G	
ɦ	[&hhook;]	&hslonglig;	0266	LATIN SMALL LETTER H WITH HOOK	LATIN SMALL LIGATURE H AND LONG S
ɪ		&iscap;	026A	LATIN LETTER SMALL CAPITAL I	
ɴ		&nscap;	0274	LATIN LETTER SMALL CAPITAL N	
ɶ		&oeligscap;	0276	LATIN LETTER SMALL CAPITAL OE	
ɼ		&rdes;	027C	LATIN SMALL LETTER R WITH LONG LEG	
ʀ		&rscap;	0280	LATIN LETTER SMALL CAPITAL R	

ⱴ		&ubar;	0289	LATIN SMALL LETTER U BAR
ʏ		&yscap;	028F	LATIN LETTER SMALL CAPITAL Y
ʙ		&bscap;	0299	LATIN LETTER SMALL CAPITAL B
ʜ		&hscap;	029C	LATIN LETTER SMALL CAPITAL H
ʟ		&lscap;	029F	LATIN LETTER SMALL CAPITAL L

6: Spacing Modifying Letters

In Unicode 4.0 this range has 80 characters. The selection below contains 6 modifying letters. Note that these letters are spacing, i.e. occupying a separate position on the base line. Combining diacritical marks are listed in the next section.

Glyph	ISO entity	MUFI entity	Code point	Descriptive name
˘	˘		02D8	BREVE
˙	˙		02D9	DOT ABOVE
˚	˚		02DA	RING ABOVE
˛	˛		02DB	OGONEK

| ~ | ˜ | | 02DC | SMALL TILDE |
| ˝ | ˝ | | 02DD | DOUBLE ACUTE ACCENT |

7: Combining Diacritical Marks

In Unicode 4.0 this range has 96 characters, of which 31 have been selected here. The combining characters at the end are recent additions to the Unicode Standard. Additional superscript characters are located in PUA subrange 7 below, p. 61. A dotted circle has been used throughout to illustrate the position of the marks in relation to a typical base line character.

Glyph	ISO entity	MUFI entity	Code point	Unicode descriptive name	MUFI descriptive name
◌̀		&combgrave;	0300	COMBINING GRAVE ACCENT	
◌́		&combacute;	0301	COMBINING ACUTE ACCENT	
◌̂		&combcirc;	0302	COMBINING CIRCUMFLEX ACCENT	
◌̃		&combtilde;	0303	COMBINING TILDE	
◌̄		&combmacr;	0304	COMBINING MACRON	
◌̅		&combovl;	0305	COMBINING OVERLINE	COMBINING ABBREVIATION MARK BAR ABOVE

		&combbreve;	0306	COMBINING BREVE	
		&combdot;	0307	COMBINING DOT ABOVE	
		&combuml;	0308	COMBINING DIAERESIS	
		&combhook;	0309	COMBINING HOOK ABOVE	
		&combring;	030A	COMBINING RING ABOVE	
		&combdblac;	030B	COMBINING DOUBLE ACUTE ACCENT	
		&combsgvertl;	030D	COMBINING VERTICAL LINE ABOVE	
		&combdbvertl;	030E	COMBINING DOUBLE VERTICAL LINE ABOVE	
		&combdotbl;	0323	COMBINING DOT BELOW	
		&combogon;	0328	COMBINING OGONEK	
		&barbl;	0332	COMBINING LOW LINE	COMBINING ABBREVIATION MARK BAR BELOW
		&baracr;	0336	COMBINING LONG STROKE OVERLAY	COMBINING ABBREVIATION MARK BAR ACROSS
		&asup;	0363	COMBINING LATIN SMALL LETTER A	

e̊		&esup;	0364	COMBINING LATIN SMALL LETTER E
i̊		&isup;	0365	COMBINING LATIN SMALL LETTER I
o̊		&osup;	0366	COMBINING LATIN SMALL LETTER O
ů		&usup;	0367	COMBINING LATIN SMALL LETTER U
c̊		⫐	0368	COMBINING LATIN SMALL LETTER C
d̊		&dsup;	0369	COMBINING LATIN SMALL LETTER D
h̊		&hsup;	036A	COMBINING LATIN SMALL LETTER H
m̊		&msup;	036B	COMBINING LATIN SMALL LETTER M
r̊		&rsup;	036C	COMBINING LATIN SMALL LETTER R
t̊		&tsup;	036D	COMBINING LATIN SMALL LETTER T
v̊		&vsup;	036E	COMBINING LATIN SMALL LETTER V
x̊		&xsup;	036F	COMBINING LATIN SMALL LETTER X

8: Greek and Coptic

Only a single character has been selected from this range, which contains 120 characters in Unicode 4.0. Presently, this range contains Greek and Coptic characters, but Coptic is likely to be disambiguated from Greek and to be given a separate range in a future version of the Unicode Standard.

Glyph	ISO entity	MUFI entity	Code point	Unicode descriptive name	MUFI descriptive name
θ	&thgr;	&obiit;	03B8	GREEK SMALL LETTER THETA	LATIN ABBREVIATION SIGN OBIIT

9: Georgian

Only a single character has been selected from this range, which contains 80 characters in Unicode 4.0. This character has a duplicate code point in the present version of Junicode (0.6.3).

Glyph	ISO entity	MUFI entity	Code point	(Junicode)	Unicode descriptive name	MUFI descriptive name
∴		&tridotright;	10FB	F162	GEORGIAN PARAGRAPH SEPARATOR	PUNCTUATION MARK RIGHT-POINTING TRIANGULAR DOTS

10: Runic

Two characters have been selected from this range, which contains 81 characters in Unicode 4.0. Both characters were used as abbreviation signs in the context of the Latin alphabet. Many Medieval fonts will most likely include the whole Runic range.

Glyph	ISO entity	MUFI entity	Code point	Unicode descriptive name	MUFI descriptive name
ᚠ		&frun;	16A0	RUNIC LETTER FEHU FEOH FE F	RUNIC LETTER F
ᛘ		&mrun;	16D8	RUNIC LETTER LONG-BRANCH-MADR M	RUNIC LETTER M

11: Phonetic Extensions

In Unicode 4.0 this range has 108 characters, of which 15 have been selected here, all small capitals. For additional small capitals, see *IPA Extensions*, p. 30 above, and PUA subrange 2 *Small capitals*, p. 53 below.

Glyph	ISO entity	MUFI entity	Code point	Descriptive name
A		&ascap;	1D00	LATIN LETTER SMALL CAPITAL A
Æ		&aeligscap;	1D01	LATIN LETTER SMALL CAPITAL AE
C		&cscap;	1D04	LATIN LETTER SMALL CAPITAL C
D		&dscap;	1D05	LATIN LETTER SMALL CAPITAL D

Ð		&ethscap;	1D06	LATIN LETTER SMALL CAPITAL ETH
E		&escap;	1D07	LATIN LETTER SMALL CAPITAL E
J		&jscap;	1D0A	LATIN LETTER SMALL CAPITAL J
K		&kscap;	1D0B	LATIN LETTER SMALL CAPITAL K
M		&mscap;	1D0D	LATIN LETTER SMALL CAPITAL M
O		&oscap;	1D0F	LATIN LETTER SMALL CAPITAL O
T		&tscap;	1D1B	LATIN LETTER SMALL CAPITAL T
U		&uscap;	1D1C	LATIN LETTER SMALL CAPITAL U
V		&vscap;	1D20	LATIN LETTER SMALL CAPITAL V
W		&wscap;	1D21	LATIN LETTER SMALL CAPITAL W
Z		&zscap;	1D22	LATIN LETTER SMALL CAPITAL Z

12: Latin Extended Additional

In Unicode 4.0 this range has 246 characters. The selection below contains 76 characters, many with dot below, others with hook above, or macron and acute. For additional precomposed characters of these types, see PUA subranges 19, 21 and 25, pp. 80, 86 and 92 below. Note that two characters below have duplicate code points in the present version of Junicode (0.6.3), 1E17 and 1E53.

Glyph	ISO entity	MUFI entity	Code point	(Junicode)	Descriptive name
Ḃ		&Bdot;	1E02		LATIN CAPITAL LETTER B WITH DOT ABOVE
ḃ		&bdot;	1E03		LATIN SMALL LETTER B WITH DOT ABOVE
Ḅ		&Bdotbl;	1E04		LATIN CAPITAL LETTER B WITH DOT BELOW
ḅ		&bdotbl;	1E05		LATIN SMALL LETTER B WITH DOT BELOW
Ḋ		&Ddot;	1E0A		LATIN CAPITAL LETTER D WITH DOT ABOVE
ḋ		&ddot;	1E0B		LATIN SMALL LETTER D WITH DOT ABOVE
Ḍ		&Ddotbl;	1E0C		LATIN CAPITAL LETTER D WITH DOT BELOW
ḍ		&ddotbl;	1E0D		LATIN SMALL LETTER D WITH DOT BELOW
É̄		&Emacracute;	1E16		LATIN CAPITAL LETTER E WITH MACRON AND ACUTE
é̄		&emacracute;	1E17	F185	LATIN SMALL LETTER E WITH MACRON AND ACUTE

Ḟ		&Fdot;	1E1E		LATIN CAPITAL LETTER F WITH DOT ABOVE
ḟ		&fdot;	1E1F		LATIN SMALL LETTER F WITH DOT ABOVE
Ḥ		&Hdotbl;	1E24		LATIN CAPITAL LETTER H WITH DOT BELOW
ḥ		&hdotbl;	1E25		LATIN SMALL LETTER H WITH DOT BELOW
Ḳ		&Kdotbl;	1E32		LATIN CAPITAL LETTER K WITH DOT BELOW
ḳ		&kdotbl;	1E33		LATIN SMALL LETTER K WITH DOT BELOW
Ḷ		&Ldotbl;	1E36		LATIN CAPITAL LETTER L WITH DOT BELOW
ḷ		&ldotbl;	1E37		LATIN SMALL LETTER L WITH DOT BELOW
Ṁ		&Mdot;	1E40		LATIN CAPITAL LETTER M WITH DOT ABOVE
ṁ		&mdot;	1E41		LATIN SMALL LETTER M WITH DOT ABOVE
Ṃ		&Mdotbl;	1E42		LATIN CAPITAL LETTER M WITH DOT BELOW
ṃ		&mdotbl;	1E43		LATIN SMALL LETTER M WITH DOT BELOW
Ṇ		&Ndotbl;	1E46		LATIN CAPITAL LETTER N WITH DOT BELOW

ṇ		&ndotbl;	1E47		LATIN SMALL LETTER N WITH DOT BELOW
Ő		&Omacracute;	1E52		LATIN CAPITAL LETTER O WITH MACRON AND ACUTE
ő		&omacracute;	1E53	F189	LATIN SMALL LETTER O WITH MACRON AND ACUTE
Ṗ		&Pdot;	1E56		LATIN CAPITAL LETTER P WITH DOT ABOVE
ṗ		&pdot;	1E57		LATIN SMALL LETTER P WITH DOT ABOVE
Ṛ		&Rdotbl;	1E5A		LATIN CAPITAL LETTER R WITH DOT BELOW
ṛ		&rdotbl;	1E5B		LATIN SMALL LETTER R WITH DOT BELOW
Ṡ		&Sdot;	1E60		LATIN CAPITAL LETTER S WITH DOT ABOVE
ṡ		⋅	1E61		LATIN SMALL LETTER S WITH DOT ABOVE
Ṣ		&Sdotbl;	1E62		LATIN CAPITAL LETTER S WITH DOT BELOW
ṣ		&sdotbl;	1E63		LATIN SMALL LETTER S WITH DOT BELOW
Ṫ		&Tdot;	1E6C		LATIN CAPITAL LETTER T WITH DOT ABOVE
ṫ		⃛	1E6D		LATIN SMALL LETTER T WITH DOT ABOVE

Ṭ		&Tdotbl;	1E6C		LATIN CAPITAL LETTER T WITH DOT BELOW
ṭ		&tdotbl;	1E6D		LATIN SMALL LETTER T WITH DOT BELOW
Ṿ		&Vdotbl;	1E7E		LATIN CAPITAL LETTER V WITH DOT BELOW
ṿ		&vdotbl;	1E7F		LATIN SMALL LETTER V WITH DOT BELOW
Ẃ		&Wacute;	1E82		LATIN CAPITAL LETTER W WITH ACUTE
ẃ		&wacute;	1E83		LATIN SMALL LETTER W WITH ACUTE
Ẅ		&Wuml;	1E84		LATIN CAPITAL LETTER W WITH DIAERESIS
ẅ		&wuml;	1E85		LATIN SMALL LETTER W WITH DIAERESIS
Ẇ		&Wdot;	1E86		LATIN CAPITAL LETTER W WITH DOT ABOVE
ẇ		&wdot;	1E87		LATIN SMALL LETTER W WITH DOT ABOVE
Ẉ		&Wdotbl;	1E88		LATIN CAPITAL LETTER W WITH DOT BELOW
ẉ		&wdotbl;	1E89		LATIN SMALL LETTER W WITH DOT BELOW
Ẏ		&Ydot;	1E8E		LATIN CAPITAL LETTER Y WITH DOT ABOVE

ẏ		&ydot;	1E8F		LATIN SMALL LETTER Y WITH DOT ABOVE
Ẓ		&Zdotbl;	1E92		LATIN CAPITAL LETTER Z WITH DOT BELOW
ẓ		&zdotbl;	1E93		LATIN SMALL LETTER Z WITH DOT BELOW
Ạ		&Adotbl;	1EA0		LATIN CAPITAL LETTER A WITH DOT BELOW
ạ		&adotbl;	1EA1		LATIN SMALL LETTER A WITH DOT BELOW
Ả		&Ahook:	1EA2		LATIN CAPITAL LETTER A WITH HOOK ABOVE
ả		&ahook;	1EA3		LATIN SMALL LETTER A WITH HOOK ABOVE
Ắ		&Abreveacute;	1EAE		LATIN CAPITAL LETTER A WITH BREVE AND ACUTE
ắ		&abreveacute;	1EAF		LATIN SMALL LETTER A WITH BREVE AND ACUTE
Ẹ		&Edotbl;	1EB8		LATIN CAPITAL LETTER E WITH DOT BELOW
ẹ		&edotbl;	1EB9		LATIN SMALL LETTER E WITH DOT BELOW
Ỉ		&Ihook;	1EC8		LATIN CAPITAL LETTER I WITH HOOK ABOVE
ỉ		&ihook;	1EC9		LATIN SMALL LETTER I WITH HOOK ABOVE

Ị		&Idotbl;	1ECA		LATIN CAPITAL LETTER I WITH DOT BELOW
ị		&idotbl;	1ECB		LATIN SMALL LETTER I WITH DOT BELOW
Ọ		&Odotbl;	1ECC		LATIN CAPITAL LETTER O WITH DOT BELOW
ọ		&odotbl;	1ECD		LATIN SMALL LETTER O WITH DOT BELOW
Ỏ		&Ohook;	1ECE		LATIN CAPITAL LETTER O WITH HOOK ABOVE
ỏ		&ohook;	1ECF		LATIN SMALL LETTER O WITH HOOK ABOVE
Ụ		&Udotbl;	1EE4		LATIN CAPITAL LETTER U WITH DOT BELOW
ụ		&udotbl;	1EE5		LATIN SMALL LETTER U WITH DOT BELOW
Ủ		&Uhook;	1EE6		LATIN CAPITAL LETTER U WITH HOOK ABOVE
ủ		&uhook;	1EE7		LATIN SMALL LETTER U WITH HOOK ABOVE
Ỵ		&Ydotbl;	1EF4		LATIN CAPITAL LETTER Y WITH DOT BELOW
ỵ		&ydotbl;	1EF5		LATIN SMALL LETTER Y WITH DOT BELOW
Ỷ		&Yhook;	1EF6		LATIN CAPITAL LETTER Y WITH HOOK ABOVE

ẙ		&yhook;	1EF7		LATIN SMALL LETTER Y WITH HOOK ABOVE

13: General Punctuation

In Unicode 4.0 this range has 97 characters. The selection below contains 48 characters, including a number of spaces with no visible glyph. Note that one character, 204A, did until recently have a duplicate code point in Junicode.

Glyph	ISO entity	MUFI entity	Code point	(Junicode)	Unicode descriptive name	MUFI descriptive name
		&enqd;	2000		EN QUAD	
		&emqd;	2001		EM QUAD	
			2002		EN SPACE	
			2003		EM SPACE	
	 		2004		THREE-PER-EM SPACE	
	 		2005		FOUR-PER-EM SPACE	
		&emsp16;	2006		SIX-PER-EM SPACE	
	&numsp:		2007		FIGURE SPACE	

	 		2008		PUNCTUATION SPACE	
			2009		THIN SPACE	
	 		200A		HAIR SPACE	
		&zerosp;	200B		ZERO WIDTH SPACE	
‒	‐		2010		HYPHEN	

This is an unambiguous *hyphen* character. See also 002D HYPHEN-MINUS in *Basic Latin* (p. 8 above).

‑		&nbhy;	2011		NON-BREAKING HYPHEN	
‒		&numdash;	2012		FIGURE DASH	
–	–		2013		EN DASH	
—	—		2014		EM DASH	
―	―		2015		HORIZONTAL DASH	
‖	‖		2016		DOUBLE VERTICAL LINE	
'	‘		2018		LEFT SINGLE QUOTATION MARK	

,	’		2019		RIGHT SINGLE QUOTATION MARK	

The entity '’' is also being used for this character. An analogous variant to 2018, '‚', has not been included in the survey of ISO conformant entities published on the Oasis web site (cf. p. 5 above).

,	‚	&lsquolow;	201A		SINGLE LOW-9 QUOTATION MARK	

The final 'r' in the ISO entity is ambiguous, since it is otherwise being used for raised quotation marks, such as 201B and 201F. For this reason, the suffix 'low' has been used in the recommended MUFI entity.

‛		&rsquolrev;	201B		SINGLE HIGH-REVERSED-9 QUOTATION MARK	
"	“		201C		LEFT DOUBLE QUOTATION MARK	
"	”		201D		RIGHT DOUBLE QUOTATION MARK	

The entity '”' is also being used for this character. An analogous variant to 201C, '„', has not been included in the survey of ISO conformant entities published on the Oasis web site (cf. p. 5 above).

„	„	&ldquolow;	201E		DOUBLE LOW-9 QUOTATION MARK	

The final 'r' in the ISO entity is ambiguous, since it is otherwise being used for raised quotation marks, such as 201B and 201F. For this reason, the suffix 'low' has been used in the recommended MUFI entity.

‟		&rdquorev;	201F		DOUBLE HIGH-REVERSED-9 QUOTATION MARK	
†	†		2020		DAGGER	

‡	‡		2021		DOUBLE DAGGER	
•	•		2022		BULLET	
▶		&tribull;	2023		TRIANGULAR BULLET	
.		&sgldr;	2024		ONE DOT LEADER	
..		&dblldr;	2025		TWO DOT LEADER	

Most likely by a mistake, two entities, '‥' ("en leader") and '…' ("em leader"), have been recorded for 2025 TWO DOT LEADER in the survey of ISO conformant entities published on the Oasis web site (cf. p. 5 above).

...	…		2026		HORIZONTAL ELLIPSIS	
·		&hyphpoint;	2027		HYPHENATION POINT	
		&nnbsp;	202F		NARROW NO-BREAK SPACE	
‰	‰		2030		PER MILLE SIGN	
′	′		2032		PRIME	
″	″		2033		DOUBLE PRIME	

‹		‹	2039		SINGLE LEFT-POINTING ANGLE QUOTATION MARK	
The entity '«' is also being used for this character, but has here been reserved for 00AB in *Latin-1 Supplement* (p. 16 above).						
›		›	203A		SINGLE RIGHT-POINTING ANGLE QUOTATION MARK	
The entity '»' is also being used for this character, but has here been reserved for 00BB in *Latin-1 Supplement* (p. 17 above).						
※		&refmark;	203B		REFERENCE MARK	
⁂		&triast;	2042		ASTERISM	
⁄		&fracsol;	2044		FRACTION SLASH	
⁅		&lsqbqu;	2045		LEFT SQUARE BRACKET WITH QUILL	
⁆		&rsqbqu;	2046		RIGHT SQUARE BRACKET WITH QUILL	
⁊		&et;	204A	F143	TIRONIAN SIGN ET	LATIN ABBREVIATION SIGN ET
As of Junicode version 0.6.3, this character is no longer located at F143, but has been moved to 204A.						
⁋		&revpara;	204B		REVERSED PILCROW SIGN	

14: Mathematical Operators

Three characters have been selected from this range, which contains 256 characters in Unicode 4.0. Since two of them have a non-mathematical usage, alternative MUFI descriptive names have been assigned. Note that 223B has a duplicate code point in the present version of Junicode (0.6.3).

Glyph	ISO entity	MUFI entity	Code point	(Junicode)	Unicode descriptive name	MUFI descriptive name
−	−		2212		MINUS	
∴	∴		2234		THEREFORE	PUNCTUATION MARK UPWARDS-POINTING TRIANGULAR DOTS
÷		&est;	223B	F150	HOMOTHETIC	LATIN ABBREVIATION SIGN EST

15: Dingbats

A single character has been selected from this range, which contains 174 characters in Unicode 4.0.

Glyph	ISO entity	MUFI entity	Code point	Unicode descriptive name	MUFI descriptive name
†		✗	271D	LATIN CROSS	LATIN ABBREVIATION SIGN CROSS

16: Miscellaneous Mathematical Symbols-A

Four characters have been selected from this range, which contains 28 characters in Unicode 4.0. Note that the angle brackets from this range have been chosen rather than 2329 and 232A in *Miscellaneous Technical*, since the latter are canonical equivalent to the very wide brackets 3008 and 3009 in *CJK Symbols and Punctuation*. This is to avoid spacing problems which might arise if the CJK brackets were substituted.

Glyph	ISO entity	MUFI entity	Code point	Unicode descriptive name	MUFI descriptive name
⟦		&lwhsqb;	27E6	MATHEMATICAL LEFT WHITE SQUARE BRACKET	LEFT WHITE SQUARE BRACKET
⟧		&rwhsqb;	27E7	MATHEMATICAL RIGHT WHITE SQUARE BRACKET	RIGHT WHITE SQUARE BRACKET
⟨		&langb;	27E8	MATHEMATICAL LEFT ANGLE BRACKET	LEFT ANGLE BRACKET
⟩		&rangb;	27E9	MATHEMATICAL RIGHT ANGLE BRACKET	RIGHT ANGLE BRACKET

B: Characters in the Private Use Area of the Unicode Standard

Important note: Characters in the Private Use Area (PUA) should be used with great caution. The characters listed in this section are supported by MUFI compliant fonts, but if a document is displayed with another, non-compliant font, these characters will show up incorrectly (with the unlikely exception of characters that by coincidence are located at the same code points). For documents with long life expectancy, it is strongly recommended that PUA characters should be encoded with mark-up or entities, and that PUA characters should be used for the final display only, whether on screen or in print. For documents with short life expectancy, characters may be used with less caution as long as the future problems of storage and interchangeability are considered.

MUFI recommends that if a character in the PUA is accepted by the Unicode Standard it should be decommissioned and moved to the new code point. Thus, newer versions of MUFI compliant fonts may lack characters of this type (or, preferably, they should be shown within a triangular warning sign). For this reason, anyone who uses a MUFI compliant font for a document should keep a copy of that particular font, and make a PDF version of the file for future reference. Earlier versions of fonts may not be accessible at a later date.

Finally, note that some characters have duplicate code points to accommodate the Junicode font as of version 0.6.3. This is due to an unavoidable code point conflict between the PUA of this font and other fonts that were coordinated by the MUFI group. The code points given in the MUFI column should be used hereafter, but for the sake of backward compatibility, older code points in Junicode may be duplicated in some fonts. This applies to a total of 34 characters in subranges 5, 10, 16, 23, 24, 25 and 30 below.

Category 1: Base characters

Base characters are base line characters or combining characters making up a part of a composite character. The allocation of code points has been coordinated with characters in the Titus Cyberbit, Junicode and Alphabetum fonts. Subranges 1, 2, 3, 5 and 7 have been placed in the Titus section, E000–F0FF; subranges 4, 6, 8 and 9 in the F100–F1FF section (using already existing Junicode code points wherever possible), and subrange 10 is placed in the F700–F71F section (using already existing Alphabetum code points).

Four subranges, 11–14, have been reserved for future additions.

Subrange 1: Ligatures

Ligatures are two base line characters which are joined so that they form a new, composite base line character. Some consist of two identical characters, e.g. 'a+a', others of different characters, e.g. 'a+v'. In Medieval Nordic manuscripts, ligatures may be used to denote length, 'a+a', diphthong, 'a+v', or a distinct vowel quality, often mutation (Umlaut), 'a+v'. Only ligatures which reflect a distinct phonological value should be recognised as characters of their own. – The broken character 'l' representing 'll' should be seen as a ligature of two stems, broken in the middle.

Unicode 4.0 does not recognise any of the ligatures in the list below.

Glyph	MUFI entity	MUFI code point	Descriptive name
AA	&AAlig;	EF90	LATIN CAPITAL LIGATURE AA
aa	&aalig;	EF91	LATIN SMALL LIGATURE AA
AO	&AOlig;	EF92	LATIN CAPITAL LIGATURE AO
ao	&aolig;	EF93	LATIN SMALL LIGATURE AO
AU	&AUlig;	EF94	LATIN CAPITAL LIGATURE AU
au	&aulig;	EF95	LATIN SMALL LIGATURE AU
AV	&AVlig;	EF96	LATIN CAPITAL LIGATURE AV
av	&avlig;	EF97	LATIN SMALL LIGATURE AV
AV	&AVligslash;	EF98	LATIN CAPITAL LIGATURE AV WITH STROKE

Glyph	MUFI entity	MUFI code point	MUFI descriptive name
ꜹ̸	&avligslash;	EF99	LATIN SMALL LIGATURE AV WITH STROKE
Aʏ	&AYlig;	EF9A	LATIN CAPITAL LIGATURE AY
ꜳy	&aylig;	EF9B	LATIN SMALL LIGATURE AY
ł	&lbrk;	EF9C	LATIN SMALL LETTER BROKEN L

Subrange 2: Small capitals

Small capitals have the same form as majuscules (capital letters), but are usually drawn with the same height as a minuscule (small) letter such as 'x'. In Medieval Nordic manuscripts, small capitals were used to denote geminates, i.e. long consonants, or they were used ornamentally. The letters 'B', 'D', 'G', 'M', 'N', 'R', 'S' and 'T' were most frequently used as geminates, while these and other letters might also be used as ornaments in the whole or in parts of highlighted words. Some of the small capitals, e.g. 'O' and 'C', are difficult to distinguish from minuscule letters.

Unicode 4.0 has allocated nine small capitals to the range *IPA Extensions*, sc. 'B', 'G', 'H', 'I', 'L', 'N', 'Œ', 'R' and 'Y' (cf. p. 30 above), and another 15 small capitals for the Uralic Phonetic Alphabet to the range *Phonetic Extensions*, 'A', 'Æ', 'C', 'D', 'ETH', 'E', 'J', 'K', 'M', 'O', 'P', 'T', 'U', 'V', 'W' and 'Z' (cf. p. 36 above). Thus, only a handful of small capitals remain unrepresented. Of these, only small capital 'S' and 'F' can appear as geminates in Medieval Nordic. The rest, i.e. 'Q', 'THORN' and 'X' can only appear as small capitals in ornamental usage.

Glyph	MUFI entity	MUFI code point	MUFI descriptive name
ꜰ	&fscap;	EF05	LATIN LETTER SMALL CAPITAL F
ǫ	&qscap;	EF0C	LATIN LETTER SMALL CAPITAL Q

S	&sscap;	EF0E	LATIN LETTER SMALL CAPITAL S
Þ	&thscap;	EF15	LATIN LETTER SMALL CAPITAL THORN
X	&xscap;	EF11	LATIN LETTER SMALL CAPITAL X

Subrange 3: Enlarged minuscules

Enlarged minuscules are recognised as separate characters by some scholars, cf. e.g. Andrea de Leeuw van Weenen (*A Grammar of Möðruvallabók*, CNWS 85, Leiden 2000). The traditional view has been to interpret these characters as variants of majuscules and encode them as such. It can be argued that this is a functional rather than a graphemic point of view and that it obscures the obvious distinction between e.g. 'A' (the majuscule) and 'a' (the enlarged minuscule).

Unicode 4.0 does not recognise enlarged minuscules.

Glyph	MUFI entity	MUFI code point	MUFI descriptive name
a	&aenl;	EEE0	LATIN ENLARGED LETTER SMALL A
b	&benl;	EEE1	LATIN ENLARGED LETTER SMALL B
c	&cenl;	EEE2	LATIN ENLARGED LETTER SMALL C
d	&denl;	EEE3	LATIN ENLARGED LETTER SMALL D
∂	&duncenl;	EEE4	LATIN ENLARGED LETTER UNCIAL D

ð	&ethenl;	EEE5	LATIN ENLARGED LETTER SMALL ETH
e	&eenl;	EEE6	LATIN ENLARGED LETTER SMALL E
f	&fenl;	EEE7	LATIN ENLARGED LETTER SMALL F
g	&genl;	EEE8	LATIN ENLARGED LETTER SMALL G
h	&henl;	EEE9	LATIN ENLARGED LETTER SMALL H
i	&ienl;	EEEA	LATIN ENLARGED LETTER SMALL I
j	&jenl;	EEEB	LATIN ENLARGED LETTER SMALL J
k	&kenl;	EEEC	LATIN ENLARGED LETTER SMALL K
l	&lenl;	EEED	LATIN ENLARGED LETTER SMALL L
m	&menl;	EEEE	LATIN ENLARGED LETTER SMALL M
n	&nenl;	EEEF	LATIN ENLARGED LETTER SMALL N
o	&oenl;	EEF0	LATIN ENLARGED LETTER SMALL O
p	&penl;	EEF1	LATIN ENLARGED LETTER SMALL P

q	&qenl;	EEF2	LATIN ENLARGED LETTER SMALL Q
r	&renl;	EEF3	LATIN ENLARGED LETTER SMALL R
s	&senl;	EEF4	LATIN ENLARGED LETTER SMALL S
t	&tenl;	EEF5	LATIN ENLARGED LETTER SMALL T
þ	&thornenl;	EEF6	LATIN ENLARGED LETTER SMALL THORN
u	&uenl;	EEF7	LATIN ENLARGED LETTER SMALL U
v	&venl;	EEF8	LATIN ENLARGED LETTER SMALL V
w	&wenl;	EEF9	LATIN ENLARGED LETTER SMALL W
x	&xenl;	EEFA	LATIN ENLARGED LETTER SMALL X
y	&yenl;	EEFB	LATIN ENLARGED LETTER SMALL Y
z	&zenl;	EEFC	LATIN ENLARGED LETTER SMALL Z

Subrange 4: Base line abbreviation characters

This range includes those abbreviation signs which typically occupy a position on the base line. Very few abbreviation signs are included in Unicode 4.0. An exception is the Tironian sign for 'et', which is found in the range *General punctuation*, 204A (p. 48 above).

The Runic characters 'f' and 'm' are sometimes used as abbreviation marks; they have been included in the range *Runic*, as 16AO and 16D8 respectively (p. 36 above). Since they are used with their alphabetical names as abbreviation ('fé' and 'maðr' respectively), it is not necessary to define them as separate characters.

Several abbreviation signs can be identified with already existing characters in Unicode 4.0:

• *Basic Latin:* the ampersand, 0026, for 'et' (= the conjunction 'and') and the apostrophe, 0027, for 'i' or 'e' (pp. 7–8).
• *IPA Extensions:* The open 'o', 0254, similar to the abbreviation sign for 'con' (p. 30).
• *Greek and Coptic:* The 'theta' character, 03B8, for 'obiit' (p. 35).
• *Mathematical Operators:* The homothetic sign, 223B, for 'est' (p. 49).
• *Dingbats:* The Latin cross, 271D (p. 50).

Glyph	MUFI entity	MUFI code point	(Junicode)	MUFI descriptive name
9	&condes;	F156	F156	LATIN ABBREVIATION SIGN CON
7	&ET;	F142	F142	LATIN ABBREVIATION LARGE SIGN ET
ꝛ	&etslash;	F158		LATIN ABBREVIATION SIGN ET WITH STROKE
ꝝ	&rum;	F154	F154	LATIN ABBREVIATION SIGN RUM
ꝺ	&de;	F159		LATIN ABBREVIATION SIGN DE

ʃ	&is;	F15A		LATIN ABBREVIATION SIGN IS
ʒ	&etfin;	F155	F155	LATIN ABBREVIATION SIGN FINAL ET

Subrange 5: Modified base line abbreviation characters

This range includes a number of abbreviation characters, typically a base line character with a stroke across the ascender. A few abbreviation signs can be identified with already existing characters in the Unicode Standard:

• *Latin Extended-B:* 'k with hook', 0199, similar to the ligature of 'k' and 'long s' (p. 27).
• *IPA Extensions:* 'h with hook', 0266, similar to the ligature of 'h' and 'long s', and 'u with bar', 0289 (p. 30).

Note that six characters have duplicate code points in the present version of Junicode (0.6.3): E7C5, E270, E670, E68A, E27D and E67D.

Glyph	MUFI entity	MUFI code point	(Junicode)	MUFI descriptive name
þ	&thornslonglig;	E734		LATIN SMALL LIGATURE THORN AND LONG S
ħ	&hslongligbar;	E7C7		LATIN SMALL LIGATURE H AND LONG S WITH STROKE
ꝁ	&kslongligbar;	E7C8		LATIN SMALL LIGATURE K AND LONG S WITH STROKE
þ	&thornslongligbar;	E735		LATIN SMALL LIGATURE THORN AND LONG S WITH STROKE
ꝃ	&kbar;	E7C5	F14B	LATIN SMALL LETTER K WITH STROKE

þ	&thornbar;	E73B		LATIN SMALL LETTER THORN WITH STROKE
ſ	&slongbar;	E7FF		LATIN SMALL LETTER LONG S WITH STROKE
P	&Pbardes;	E270	F144	LATIN CAPITAL LETTER P WITH STROKE THROUGH DESCENDER
p	&pbardes;	E670	F145	LATIN SMALL LETTER P WITH STROKE THROUGH DESCENDER
q	&qbardes;	E68A	F14D	LATIN SMALL LETTER Q WITH STROKE THROUGH DESCENDER
þ	&thornbardes;	E7A0		LATIN SMALL LETTER THORN WITH STROKE THROUGH DESCENDER
P	&Pflour;	E27D	F146	LATIN CAPITAL LETTER P WITH FLOURISH
p	&pflour;	E67D	F147	LATIN SMALL LETTER P WITH FLOURISH

Subrange 6: Combining marks

This range includes those abbreviation signs which typically occupy a position above, through or below another base line character. Combining diacritical marks are also included in this range.

Unicode 4.0 has a large selection in *Combining Diacritical Marks* (p. 32 above). The list below has additional marks.

Glyph	MUFI entity	MUFI code point	(Junicode)	MUFI descriptive name
◌̄̇	&arbar;	F1C0		COMBINING ABBREVIATION MARK BAR ABOVE WITH DOT
◌ᷝ	&er;	F152	F152	COMBINING ABBREVIATION MARK SUPERSCRIPT ER (ZIG-ZAG SIGN)
The combining zig-zag above has been accepted by the UTC and WG2 and is likely to be included in a future version of the Unicode Standard at code point 035B in *Combining Diacritical Marks*.				
◌ᷓ	&ra;	F157	F157	COMBINING ABBREVIATION MARK SUPERSCRIPT RA (OMEGA SIGN)
◌ᷓ̄	&rabar;	F1C1		COMBINING ABBREVIATION MARK SUPERSCRIPT RA (OMEGA SIGN) WITH BAR ABOVE
◌ᷜ	&ur2;	F153	F153	COMBINING ABBREVIATION MARK SUPERSCRIPT UR (2-SIGN)
◌	&ur8;	F1C2		COMBINING ABBREVIATION MARK SUPERSCRIPT UR (8-SIGN)
◌	&ur8open;	F1C3		COMBINING ABBREVIATION MARK SUPERSCRIPT UR (OPEN 8-SIGN)
◌⁹	&ussep;	F151	F151	SPACING ABBREVIATION MARK SUPERSCRIPT US (9-SIGN)

ꙷ	&us;	F15B		COMBINING ABBREVIATION MARK SUPERSCRIPT US (9-SIGN)
͡	&combcircdbl;	F163	F163	COMBINING CIRCUMFLEX OVER TWO CHARACTERS
ꙴ	&combcurl;	F1C4		COMBINING CURL
ꙵ	&combcurlhigh;	F1C5		COMBINING CURL HIGH POSITION
ꙶ	&combflour;	F1C6		COMBINING FLOURISH

Subrange 7: Combining superscript characters

This range includes superscript characters, typically placed above another base line character. They are found in many early German printed texts, and in a large number of Medieval manuscripts. The position immediately above a base line character distinguishes them from raised interlinear characters typically occupying a position immediately after another base line character. This latter type includes a handful of phonetic modifiers such as a raised 'w' indicating rounding, a raised 'h' indicating aspiration etc.

Unicode 4.0 has a selection of 13 superscript characters in the range *Combining diacritical marks*, 'a', 'e', 'i', 'o', 'u', 'c', 'd', 'h', 'm', 'r', 't', 'v' and 'x'. They are located at the end of the range, 0363–036F (pp. 33–34 above).

The characters in the list below are documented in Andrea de Leeuw van Weenen, *A Grammar of Möðruvallabók* (CNWS 85), Leiden 2000.

Glyph	MUFI entity	MUFI code point	Descriptive name
ꙮ	&aeligsup;	F011	COMBINING LATIN SMALL LETTER AE

b̊	&bsup;	F012	COMBINING LATIN SMALL LETTER B
ᴮ̊	&bscapsup;	F013	COMBINING LATIN LETTER SMALL CAPITAL B
ð̊	&ethsup;	F014	COMBINING LATIN SMALL LETTER ETH
ᴰ̊	&dscapsup;	F016	COMBINING LATIN LETTER SMALL CAPITAL D
f̊	&fsup;	F017	COMBINING LATIN SMALL LETTER F
g̊	&gsup;	F018	COMBINING LATIN SMALL LETTER G
ᴳ̊	&gscapsup;	F019	COMBINING LATIN LETTER SMALL CAPITAL G
k̊	&ksup;	F01B	COMBINING LATIN SMALL LETTER K
ᴷ̊	&kscapsup;	F01C	COMBINING LATIN LETTER SMALL CAPITAL K
l̊	&lsup;	F01D	COMBINING LATIN SMALL LETTER L
ᴸ̊	&lscapsup;	F01E	COMBINING LATIN LETTER SMALL CAPITAL L
ᴹ̊	&mscapsup;	F020	COMBINING LATIN LETTER SMALL CAPITAL M
n̊	⊅	F021	COMBINING LATIN SMALL LETTER N

N̊	&nscapsup;	F023	COMBINING LATIN LETTER SMALL CAPITAL N
p̊	&psup;	F025	COMBINING LATIN SMALL LETTER P
R̊	&rscapsup;	F026	COMBINING LATIN LETTER SMALL CAPITAL R
s̊	&ssup;	F027	COMBINING LATIN SMALL LETTER S
ſ̊	&slongsup;	F028	COMBINING LATIN SMALL LETTER LONG S
T̊	&tscapsup;	F02A	COMBINING LATIN LETTER SMALL CAPITAL T
ẙ	&ysup;	F02B	COMBINING LATIN SMALL LETTER Y
z̊	&zsup;	F02C	COMBINING LATIN SMALL LETTER Z

Subrange 8: Punctuation marks

This range includes punctuation marks that are not included in Unicode 4.0. The common marks, such as full stop, comma, colon, semicolon, question mark, hyphen and solidus, are all found in the range *Basic Latin* (p. 7 above). Some are also located in the range *General Punctuation* (p. 44). Finally, a few punctuation marks can be identified with characters in other ranges:

• *Latin-1 Supplement:* middle dot, 00B7 (p. 17)
• *Georgian:* right-pointing triangular dots, 10FB (p. 35)
• *Mathematical Operators:* upwards-pointing triangular dots (therefore), 2234 (p. 49)

Glyph	MUFI entity	MUFI code point	(Junicode)	MUFI descriptive name
ꝯ	&punctinter;	F160	F160	PUNCTUATION MARK PUNCTUS INTERROGATIVUS
︰	&punctelev;	F161	F161	PUNCTUATION MARK PUNCTUS ELEVATUS
⸝	&punctelevdiag;	F1F0		PUNCTUATION MARK PUNCTUS ELEVATUS DIAGONAL STROKE
∞	&quest8;	F1F1		QUESTION MARK HORIZONTAL 8 FORM
⸵	&diacom;	F1F2		PUNCTUATION MARK DIAERESIS ABOVE COMMA
⏑	&brevdot;	F1F3		PUNCTUATION MARK BREVE ABOVE DOT

Subrange 9: Critical and epigraphical signs

Critical signs are used in printed editions, indicating corruptions, deletions, additions etc. The majority of these signs are already in Unicode 4.0:

• *Basic Latin:* asterisk, 002A, square brackets, 005B and 005D, curly brackets, 007B and 007D, single vertical line, 007C (cf. pp. 8–15).
• *General Punctuation:* double vertical line, 2016 (cf. p. 45)
• *Miscellaneous Mathematical Symbols-A:* double square brackets, 27E6 and 27E7, and angle brackets, 27E8 and 27E9 (cf. p. 50)

Glyph	MUFI entity	MUFI code point	(Junicode)	MUFI descriptive name
⌐	&luhsqb;	F1D2		LEFT UPPER HALF SQUARE BRACKET
A similar sign, the right angle substitution marker (a New Testament editorial symbol), has been accepted by the UTC and WG2 and is likely to be included in a future version of the Unicode Standard at code point 2E00 in *Supplemental Punctuation*. There is, however, no left angle substitution marker.				
⌐	&ruhsqb;	F1D3		RIGHT UPPER HALF SQUARE BRACKET
⌐	&llhsqb;	F1D4		LEFT LOWER HALF SQUARE BRACKET
⌐	&rlhsqb;	F1D5		RIGHT LOWER HALF SQUARE BRACKET
＼	&luslst;	F16E	F16E	LEFT UPPER SLANTED STROKE
／	&ruslst;	F16F	F16F	RIGHT UPPER SLANTED STROKE
The left and right upper slanted strokes, F16E and F16F, have been accepted by the WG2 and are likely to be included in a future version of the Unicode Standard at code points 2E0C and 2E0D in *Supplemental Punctuation*.				
／	&llslst;	F1D6		LEFT LOWER SLANTED STROKE
＼	&rlslst;	F1D7		RIGHT LOWER SLANTED STROKE

Subrange 10: Metrical symbols

This is the type of symbols used by Eduard Sievers in his *Altgermanische Metrik* (Halle: Max Niemeyer, 1893). They are still frequently used in text books, monographs and articles on Medieval Nordic metrics.

Thesaurus Linguae Grecae (TLG) has recently proposed a set of metrical symbols for Greek. This proposal has been approved by the Unicode Technical Committee, but not yet by ISO-10646. The TLG proposal includes symbols for the short syllable (breve) and for the combination of a short and long syllable (metrical short over long). The default character (anceps) is identified with MULTIPLICATION SIGN (00D7) in *Latin-1 Supplement*, and the symbol for long syllable (longum) is identified with FIGURE DASH (2012) or EN DASH (2013) in *General Punctuation*. Combinations of anceps, breve or longum with grave or acute accent can presumably be achieved by using COMBINING ACUTE ACCENT (0300) and COMBINING GRAVE ACCENT (0301) in *Combining Diacritical Marks*. However, this does not work in all applications, so for the time being precomposed characters will be useful.

Junicode and Alphabetum have metrical symbols in different locations of the Private Use Area, so a code point conflict has been inevitable for this range. Since Alphabetum has the largest selection, code points in this font have been chosen, metrical symbols being placed in the F700–F71F section. However, all duplicate code points in Junicode as of version 0.6.3 are listed below.

Alphabetum contains a number of additional metrical symbols, possibly of interest also for Medieval usage. See the F700–F76F section of this font.

Note that combining single and double vertical bars are located in *Combining Diacritical Marks*, code points 030D and 030E (p. 33 above). They are used to indicate stress (single bar) and stress + alliteration (double bar).

Glyph	MUFI entity	MUFI code point	(Junicode)	MUFI descriptive name
✕	&anc;	F70A	F164	METRICAL SYMBOL ANCEPS
✕́	&ancacute;	F70B	F165	METRICAL SYMBOL ANCEPS WITH ACUTE (PRIMARY STRESS)
✕̋	&ancdblac;	F719		METRICAL SYMBOL ANCEPS WITH DOUBLE ACUTE (PRIMARY STRESS AND ALLITERATION)

ˇ	&ancgrave;	F70C	F166	METRICAL SYMBOL ANCEPS WITH GRAVE (SECONDARY STRESS)
ˇ	&ancdblgrave;	F71A		METRICAL SYMBOL ANCEPS WITH DOUBLE GRAVE (SECONDARY STRESS AND ALLITERATION)
˘	˘	F701	F16A	METRICAL SYMBOL BREVE

The metrical breve has been accepted by the UTC and WG2 and is likely to be included in a future version of the Unicode Standard at code point 23D1 in *Miscellaneous Technical*.

˘	&breveacute;	F706	F16B	METRICAL SYMBOL BREVE WITH ACUTE (PRIMARY STRESS)
˘	&brevedblac;	F717		METRICAL SYMBOL BREVE WITH DOUBLE ACUTE (PRIMARY STRESS AND ALLITERATION)
˘	&brevegrave;	F707	F16C	METRICAL SYMBOL BREVE WITH GRAVE (SECONDARY STRESS)
˘	&brevedblgrave;	F718		METRICAL SYMBOL BREVE WITH DOUBLE GRAVE (SECONDARY STRESS AND ALLITERATION)
—	¯	F700	F167	METRICAL SYMBOL LONGUM
´	&macracute;	F704	F168	METRICAL SYMBOL LONGUM WITH ACUTE (PRIMARY STRESS)
˝	&macrdblac;	F715		METRICAL SYMBOL LONGUM WITH DOUBLE ACUTE (SECONDARY STRESS)

‖	&macrgrave;	F705	F169	METRICAL SYMBOL LONGUM WITH GRAVE (SECONDARY STRESS)
‖	&macrdblgrave;	F716		METRICAL SYMBOL LONGUM WITH DOUBLE GRAVE (SECONDARY STRESS AND ALLITERATION)
‖	&macrbreve;	F702		METRICAL SYMBOL BREVE ABOVE LONGUM (SHORT OR LONG SYLLABLE)

The metrical short over long has been accepted by the UTC and WG2 and is likely to be included in a future version of the Unicode Standard at code point 23D3 in *Miscellaneous Technical*.

‖	&brevemacr;	F703		METRICAL SYMBOL LONGUM ABOVE BREVE (SHORT OR LONG SYLLABLE)

The metrical long over short has been accepted by the UTC and WG2 and is likely to be included in a future version of the Unicode Standard at code point 23D2 in *Miscellaneous Technical*.

‖	&macrbreveacute;	F708		METRICAL SYMBOL BREVE ABOVE LONGUM WITH ACUTE (SHORT OR LONG SYLLABLE WITH PRIMARY STRESS)
‖	&macrbrevegrave;	F709		METRICAL SYMBOL BREVE ABOVE LONGUM WITH GRAVE (SHORT OR LONG SYLLABLE WITH SECONDARY STRESS)
‖	&dblbrevemacracute;	F71B		METRICAL SYMBOL RESOLVED LIFT WITH ACUTE (PRIMARY STRESS)

The metrical long over two shorts has been accepted by the UTC and WG2 and is likely to be included in a future version of the Unicode Standard at code point 23D4 in *Miscellaneous Technical*. Note that this character does not include the stress mark (the acute accent).

̆̋	&dblbrevemacrdblac;	F71C		METRICAL SYMBOL RESOLVED LIFT WITH DOUBLE ACUTE (PRIMARY STRESS AND ALLITERATION)
∧	&metrpause;	F714		METRICAL SYMBOL PAUSE
◌◌	&combdblbrevebl;	F743		COMBINING DOUBLE BREVE BELOW

The combining bow (double breve) below has been accepted by the UTC and WG2 and is likely to be included in a future version of the Unicode Standard at code point 035C in *Combining Diacritical Marks*.

Subranges 11 – 14

Reserved for future additions.

Category 2: Precomposed characters

Precomposed characters are base line characters with one or more diacritical marks. They have been coordinated with characters in the Titus project and are allocated code points in the E000–F0FF section of the Private Use Area. Precomposed characters will be decommissioned as soon as mature smart font technology is widely available.

Subrange 15: Characters with macron or overline

The overline (bar above) is probably the most used and also the most ambiguous of all abbreviation marks. There are two typical positions of the overline: above the full height of the majuscules and above the x-height of the minuscules. In the latter position it typically crosses the ascender of characters like 'b', 'd', 'h', 'k', 'l', 'thorn' and long 's'. If the word has a mixture of characters with and without ascenders, the bar should sometimes be kept in the upper position over all characters.

There are two typical lengths of the overline: less than the width of a character, like the macron, or the full width of the character, so that it can extend as a continuous line over several characters.

With present font technology, the overline is particularly difficult. With some fonts and operating systems it will change its vertical position depending on the height of each character. Thus, in an abbreviation such as 'ihc' for 'Iesus', the overline may have one position over 'c', a slightly higher position over 'i' and an even higher position over 'h'.

This range is intended as a work-around until mature smart font technology is available. It has separate code points for all characters where the overline crosses the ascender, either as a single stroke (macron-length) or as a continuous stroke (overline).

Some of these characters can be identified with characters already in Unicode 4.0:

• *Latin Extended-A:* 'd' with stroke, 0111, 'h' with stroke, 0127, and 'I' with macron, 012A (p. 24–25).
• *Latin Extended-B:* 'b' with stroke, 0180 (p. 27).

A few others have been included in the Private Use Area above, subrange 5:

• *Modified base line abbreviation characters:* 'k' with stroke, E7C5, 'thorn' with stroke, E73B, and 'long s' with stroke, E7FF (p. 58–59).

Font designers should take care to align the overlines on all characters, i.e. those already in the Unicode Standard and those listed here.

Finally, note that this range has a combining macron and a combining overline, each in two positions: one for minuscules (in the same height as the dot over 'i') and one for majuscules (in the same height as the accents). These combining characters have 'hard' positions, so that they will have the same height regardless of the characters below. In this respect, they differ from 0304 COMBINING MACRON and 0305 COMBINING OVERLINE, which may be displayed with variable height. The combining overlines ought to have medium width, suitable for characters like 'O' and 'o'.

Glyph	MUFI entity	MUFI code point	MUFI descriptive name
◌̄	&macrhigh;	F00A	COMBINING HIGH MACRON WITH FIXED HEIGHT (PART-WIDTH)
◌̄	&macrmed;	F00B	COMBINING MEDIUM-HIGH MACRON WITH FIXED HEIGHT (PART-WIDTH)
◌̅	&ovlhigh;	F00C	COMBINING HIGH OVERLINE WITH FIXED HEIGHT (FULL-WIDTH)
◌̅	&ovlmed;	F00D	COMBINING MEDIUM-HIGH OVERLINE WITH FIXED HEIGHT (FULL-WIDTH)
ƀ	&bovlmed;	E44D	LATIN SMALL LETTER B WITH MEDIUM-HIGH OVERLINE (ACROSS ASCENDER)
đ	&dovlmed;	E491	LATIN SMALL LETTER D WITH MEDIUM-HIGH OVERLINE (ACROSS ASCENDER)
ħ	&hovlmed;	E517	LATIN SMALL LETTER H WITH MEDIUM-HIGH OVERLINE (ACROSS ASCENDER)
ꝁ	&kovlmed;	E7C3	LATIN SMALL LETTER K WITH MEDIUM-HIGH OVERLINE (ACROSS ASCENDER)
ł	ł	E5B0	LATIN SMALL LETTER L WITH STROKE (ACROSS ASCENDER)

ꝉ	&lovlmed;	E5B1	LATIN SMALL LETTER L WITH MEDIUM-HIGH OVERLINE (ACROSS ASCENDER)
ſ	&slongovlmed;	E79E	LATIN SMALL LETTER LONG S WITH MEDIUM-HIGH OVERLINE (ACROSS ASCENDER)
þ	&thornovlmed;	E7A2	LATIN SMALL LETTER THORN WITH MEDIUM-HIGH OVERLINE (ACROSS ASCENDER)
Ī	&Iovlhigh;	E150	LATIN CAPITAL LETTER I WITH HIGH OVERLINE (ABOVE CHARACTER)
ī	&iovlmed;	E550	LATIN SMALL LETTER I WITH MEDIUM-HIGH OVERLINE (ABOVE CHARACTER)
J̄	&Jmacrhigh;	E154	LATIN CAPITAL LETTER J WITH HIGH MACRON (ABOVE CHARACTER)
J̄	&Jovlhigh;	E152	LATIN CAPITAL LETTER J WITH HIGH OVERLINE (ABOVE CHARACTER)
j̄	&jmacrmed;	E554	LATIN SMALL LETTER J WITH MEDIUM-HIGH MACRON (ABOVE CHARACTER)
j̄	&jovlmed;	E552	LATIN SMALL LETTER J WITH MEDIUM-HIGH OVERLINE (ABOVE CHARACTER)
l̄	&lmacrhigh;	E596	LATIN SMALL LETTER L WITH HIGH MACRON (ABOVE CHARACTER)
l̄	&lovlhigh;	E58C	LATIN SMALL LETTER L WITH HIGH OVERLINE (ABOVE CHARACTER)
M̄	&Mmacrhigh;	E1B8	LATIN CAPITAL LETTER M WITH HIGH MACRON (ABOVE CHARACTER)
M̄	&Movlhigh;	E1D2	LATIN CAPITAL LETTER M WITH HIGH OVERLINE (ABOVE CHARACTER)

m̄	&mmacrmed;	E5B8	LATIN SMALL LETTER M WITH MEDIUM-HIGH MACRON (ABOVE CHARACTER)
m̅	&movlmed;	E5D2	LATIN SMALL LETTER M WITH MEDIUM-HIGH OVERLINE (ABOVE CHARACTER)

Subrange 16: Characters with acute accent

Unicode 4.0 has acute accent over the vowels 'A', 'a', 'E', 'e', 'I', 'i', 'O', 'o', 'U', 'u', 'Y' and 'y' in *Latin-1 Supplement* (p. 15 above) and over 'Æ', 'æ', 'Ø' and 'ø' in *Latin Extended-B* (p. 27). Many consonants also appear with acute accent, such as 'C', 'c', 'L', 'l', 'N', 'n', 'R', 'r', 'S', 's', 'Z' and 'z' in *Latin Extended-A*, 'G' and 'g' in *Latin Extended-B*, and 'K', 'k', 'M', 'm', 'P', 'p', 'W' and 'w' in *Latin Extended Additional*.

Combinations of consonants and the acute accent are presumably rare in Medieval sources and have therefore not been included in the present recommendation. The list below contains characters that do appear with acute accent; the characters 'j' and 'v' should be seen as semi-vowels in this context, alternating with 'i' and 'u'.

Note that two characters have duplicate code points in the present version of Junicode (0.6.3): E20C and E60C.
The combining acute accent, 0301, is located in *Combining Diacritical Marks* (p. 32).

Glyph	MUFI entity	MUFI code point	(Junicode)	MUFI descriptive name
ÁA	&AAligacute; = &AAlig; + &combacute;	EFE0 = EF90 + 0301		LATIN CAPITAL LIGATURE AA WITH ACUTE = LATIN CAPITAL LIGATURE AA + COMBINING ACUTE ACCENT
áa	&aaligacute; = &aalig; + &combacute;	EFE1 = EF91 + 0301		LATIN SMALL LIGATURE AA WITH ACUTE = LATIN SMALL LIGATURE AA + COMBINING ACUTE ACCENT
ÁO	&AOligacute; = &AOlig; + &combacute;	EFE2 = EF92 + 0301		LATIN CAPITAL LIGATURE AO WITH ACUTE = LATIN CAPITAL LIGATURE AO + COMBINING ACUTE ACCENT

áo	&aoligacute; = &aolig; + &combacute;	EFE3 = EF93 + 0301		LATIN SMALL LIGATURE AO WITH ACUTE = LATIN SMALL LIGATURE AO + COMBINING ACUTE ACCENT
ÁJ	&AUligacute; = &AUlig; + &combacute;	EFE4 = EF94 + 0301		LATIN CAPITAL LIGATURE AU WITH ACUTE = LATIN CAPITAL LIGATURE AU + COMBINING ACUTE ACCENT
áu	&auligacute; = &aulig; + &combacute;	EFE5 = EF95 + 0301		LATIN SMALL LIGATURE AU WITH ACUTE = LATIN SMALL LIGATURE AU + COMBINING ACUTE ACCENT
ÁV	&AVligacute; = &AVlig; + &combacute;	EFE6 = EF96 + 0301		LATIN CAPITAL LIGATURE AV WITH ACUTE = LATIN CAPITAL LIGATURE AV + COMBINING ACUTE ACCENT
áv	&avligacute; = &avlig; + &combacute;	EFE7 = EF97 + 0301		LATIN SMALL LIGATURE AV WITH ACUTE = LATIN SMALL LIGATURE AV + COMBINING ACUTE ACCENT
Ę́	&Eogonacute; = E + &combogon; + &combacute;	E099 = 0035 + 0328 + 0301		LATIN CAPITAL LETTER E WITH OGONEK AND ACUTE = LATIN CAPITAL LETTER E + COMBINING OGONEK + COMBINING ACUTE ACCENT
ę́	&eogonacute; = e + &combogon; + &combacute;	E499 = 0065 + 0328 + 0301		LATIN SMALL LETTER E WITH OGONEK AND ACUTE = LATIN SMALL LETTER E + COMBINING OGONEK + COMBINING ACUTE ACCENT
J́	&Jacute; = J + &combacute;	E153 = 004A + 0301		LATIN CAPITAL LETTER J WITH ACUTE = LATIN CAPITAL LETTER J + COMBINING ACUTE ACCENT
j́	&jacute; = j + &combacute;	E553 = 006A + 0301		LATIN SMALL LETTER J WITH ACUTE = LATIN SMALL LETTER J + COMBINING ACUTE ACCENT

Ǫ́	&Oogonacute; = O + &combogon; + &combacute;	E20C = 004F + 0328 + 0301	F190	LATIN CAPITAL LETTER O WITH OGONEK AND ACUTE = LATIN CAPITAL LETTER O + COMBINING OGONEK + COMBINING ACUTE ACCENT
ǫ́	&oogonacute; = o + &combogon; + &combacute;	E60C = 006F + 0328 + 0301	F191	LATIN SMALL LETTER O WITH OGONEK AND ACUTE = LATIN SMALL LETTER O + COMBINING OGONEK + COMBINING ACUTE ACCENT
Ǿ	&Oslashogonacute; = Ø + &combogon; + &combacute;	E257 = 00D8 + 0328 + 0301		LATIN CAPITAL LETTER O WITH STROKE AND OGONEK AND ACUTE = LATIN CAPITAL LETTER O WITH STROKE + COMBINING OGONEK + COMBINING ACUTE ACCENT
ǿ	&oslashogonacute; = ø + &combogon; + &combacute;	E657 = 00F8 + 0328 + 0301		LATIN SMALL LETTER O WITH STROKE AND OGONEK AND ACUTE = LATIN SMALL LETTER O WITH STROKE + COMBINING OGONEK + COMBINING ACUTE ACCENT
OÓ	&OOligacute; = &OOlig; + &combacute;	EFE8 = F20A + 0301		LATIN CAPITAL LIGATURE OO WITH ACUTE = LATIN CAPITAL LIGATURE OO + COMBINING ACUTE ACCENT
oó	&ooligacute; = &oolig; + &combacute;	EFE9 = F20B + 0301		LATIN SMALL LIGATURE OO WITH ACUTE = LATIN SMALL LIGATURE OO + COMBINING ACUTE ACCENT
V́	&Vacute; = V + &combacute;	E33A = 0056 + 0301		LATIN CAPITAL LETTER V WITH ACUTE = LATIN CAPITAL LETTER V + COMBINING ACUTE ACCENT
v́	&vacute; = v + &combacute;	E73A = 0076 + 0301		LATIN SMALL LETTER V WITH ACUTE = LATIN SMALL LETTER V + COMBINING ACUTE ACCENT

Subrange 17: Characters with double acute accent

Unicode 4.0 has double acute accent over 'O', 'o', 'U' and 'u' in *Latin Extended-A* (p. 23 above). The list below contains additional character combinations. The combining double acute accent, 030B, is located in *Combining Diacritical Marks* (p. 32).

Glyph	MUFI entity	MUFI code point	MUFI descriptive name
Ӳ	&Adblac; = A + &combdblac;	E025 = 0041 + 030B	LATIN CAPITAL LETTER A WITH DOUBLE ACUTE = LATIN CAPITAL LETTER A + COMBINING DOUBLE ACUTE ACCENT
a̋	&adblac; = a + &combdblac;	E425 = 0061 + 030B	LATIN SMALL LETTER A WITH DOUBLE ACUTE = LATIN SMALL LETTER A + COMBINING DOUBLE ACUTE ACCENT
AA̋	&AAligdblac; = &AAlig; + &combdblac;	EFEA = EF90 + 030B	LATIN CAPITAL LIGATURE AA WITH DOUBLE ACUTE = LATIN CAPITAL LIGATURE AA + COMBINING DOUBLE ACUTE ACCENT
aa̋	&aaligdblac; = &aalig; + &combdblac;	EFEB = EF91 + 030B	LATIN SMALL LIGATURE AA WITH DOUBLE ACUTE = LATIN SMALL LIGATURE AA + COMBINING DOUBLE ACUTE ACCENT
Æ̋	&AEligdblac; = Æ + &combdblac;	E041 = 00C6 + 030B	LATIN CAPITAL LETTER AE WITH DOUBLE ACUTE = LATIN CAPITAL LETTER AE + COMBINING DOUBLE ACUTE ACCENT
æ̋	&aeligdblac; = æ + &combdblac;	E441 = 00E6 + 030B	LATIN SMALL LETTER AE WITH DOUBLE ACUTE = LATIN SMALL LETTER AE + COMBINING DOUBLE ACUTE ACCENT
E̋	&Edblac; = E + &combdblac;	E0D1 = 0035 + 030B	LATIN CAPITAL LETTER E WITH DOUBLE ACUTE = LATIN CAPITAL LETTER E + COMBINING DOUBLE ACUTE ACCENT
e̋	&edblac; = e + &combdblac;	E4D1 = 0065 + 030B	LATIN SMALL LETTER E WITH DOUBLE ACUTE = LATIN SMALL LETTER E + COMBINING DOUBLE ACUTE ACCENT

Í	&Idblac; = I + &combdblac;	E143 = 0049 + 030B	LATIN CAPITAL LETTER I WITH DOUBLE ACUTE = LATIN CAPITAL LETTER I + COMBINING DOUBLE ACUTE ACCENT
í	&idblac; = i + &combdblac;	E543 = 0069 + 030B	LATIN SMALL LETTER I WITH DOUBLE ACUTE = LATIN SMALL LETTER I + COMBINING DOUBLE ACUTE ACCENT
J́	&Jdblac; = J + &combdblac;	E162 = 004A + 030B	LATIN CAPITAL LETTER J WITH DOUBLE ACUTE = LATIN CAPITAL LETTER J + COMBINING DOUBLE ACUTE ACCENT
j́	&jdblac; = j + &combdblac;	E562 = 006A + 030B	LATIN SMALL LETTER J WITH DOUBLE ACUTE = LATIN SMALL LETTER J + COMBINING DOUBLE ACUTE ACCENT
OÓ	&OOligdblac; = &OOlig; + &combdblac;	EFEC = F20A + 030B	LATIN CAPITAL LIGATURE OO WITH DOUBLE ACUTE = LATIN CAPITAL LIGATURE OO + COMBINING DOUBLE ACUTE ACCENT
oó	&ooligdblac; = &oolig; + &combdblac;	EFED = F20B + 030B	LATIN SMALL LIGATURE OO WITH DOUBLE ACUTE = LATIN SMALL LIGATURE OO + COMBINING DOUBLE ACUTE ACCENT
V́	&Vdblac; = V + &combdblac;	E34B = 0056 + 030B	LATIN CAPITAL LETTER V WITH DOUBLE ACUTE = LATIN CAPITAL LETTER V + COMBINING DOUBLE ACUTE ACCENT
v́	&vdblac; = v + &combdblac;	E74B = 0076 + 030B	LATIN SMALL LETTER V WITH DOUBLE ACUTE = LATIN SMALL LETTER V + COMBINING DOUBLE ACUTE ACCENT
Ý	&Ydblac; = Y + &combdblac;	E37C = 0059 + 030B	LATIN CAPITAL LETTER Y WITH DOUBLE ACUTE = LATIN CAPITAL LETTER Y + COMBINING DOUBLE ACUTE ACCENT
ý	&ydblac; = y + &combdblac;	E77C = 0079 + 030B	LATIN SMALL LETTER Y WITH DOUBLE ACUTE = LATIN SMALL LETTER Y + COMBINING DOUBLE ACUTE ACCENT

Ÿ́	&Ydotacute; = Y + &combdot; + &combacute;	E384 = 0059 + 030B + 0301	LATIN CAPITAL LETTER Y WITH DOT ABOVE AND ACUTE = LATIN CAPITAL LETTER Y + COMBINING DOT ABOVE + COMBINING ACUTE ACCENT
ÿ́	&ydotacute; = y + &combdot; + &combacute;	E784 = 0079 + 030B + 0301	LATIN SMALL LETTER Y WITH DOT ABOVE AND ACUTE = LATIN SMALL LETTER Y + COMBINING DOT ABOVE + COMBINING ACUTE ACCENT

Subrange 18: Characters with dot above

Unicode 4.0 has 40 characters with a dot above in three ranges: 'C', 'c', 'E', 'e', 'G', 'g', 'I' ('i' is already in *Basic Latin*), 'Z', and 'z' in *Latin Extended-A*, 'A', 'a', 'O' and 'o', in *Latin-Extended-B*, and 'B', 'b', 'D', 'd', 'F', 'f', 'H', 'h', 'M', 'm', 'N', 'n', 'P', 'p', 'R', 'r', 'S', 's', 'T', 't', 'W', 'w', 'X', 'x', 'Y', 'y' and long 's' (no capital version) in *Latin Extended Additional*.

Many of these combinations are presumably rare in Medieval sources and only a few have therefore been included in the present recommendation. The list below contains additional characters that do appear with a dot above. The combining dot above, 0307, is located in *Combining Diacritical Marks* (p. 33).

Glyph	MUFI entity	MUFI code point	MUFI descriptive name
Ȧ̇A	&AAligdot; = &AAlig; + &combdot;	EFEE = EF90 + 0307	LATIN CAPITAL LIGATURE AA WITH DOT ABOVE = LATIN CAPITAL LIGATURE AA + COMBINING DOT ABOVE
ȧa	&aaligdot; = &aalig; + &combdot;	EFEF = EF91 + 0307	LATIN SMALL LIGATURE AA WITH DOT ABOVE = LATIN SMALL LIGATURE AA + COMBINING DOT ABOVE
ȦY	&AYligdot; = &AYlig; + &combdot;	EFF0 = EF9A + 0307	LATIN CAPITAL LIGATURE AY WITH DOT ABOVE = LATIN CAPITAL LIGATURE AY + COMBINING DOT ABOVE

ȧy	&ayligdot; = &aylig; + &combdot;	EFF1 = EF9B + 0307	LATIN SMALL LIGATURE AY WITH DOT ABOVE = LATIN SMALL LIGATURE AY + COMBINING DOT ABOVE
Ġ	&gscapdot; = &gscap; + &combdot;	EF20 = 0262 + 0307	LATIN LETTER SMALL CAPITAL G WITH DOT ABOVE = LATIN LETTER SMALL CAPITAL G + COMBINING DOT ABOVE
k̇	&kdot; = k + &combdot;	E568 = 006B + 0307	LATIN SMALL LETTER K WITH DOT ABOVE = = LATIN SMALL LETTER K + COMBINING DOT ABOVE
Ṅ	&nscapdot; = &nscap; + &combdot;	EF21 = 0274 + 0307	LATIN LETTER SMALL CAPITAL N WITH DOT ABOVE = LATIN LETTER SMALL CAPITAL N + COMBINING DOT ABOVE
Ṙ	&rscapdot; = &rscap; + &combdot;	EF22 = 0280 + 0307	LATIN LETTER SMALL CAPITAL R WITH DOT ABOVE = LATIN LETTER SMALL CAPITAL R + COMBINING DOT ABOVE
Ṡ	&sscapdot; = &sscap; + &combdot;	EF23 = EF0E + 0307	LATIN LETTER SMALL CAPITAL S WITH DOT ABOVE = LATIN LETTER SMALL CAPITAL S + COMBINING DOT ABOVE
Ṫ	&tscapdot; = &tscap; + &combdot;	EF24 = 1D1B + 0307	LATIN LETTER SMALL CAPITAL T WITH DOT ABOVE = LATIN LETTER SMALL CAPITAL T + COMBINING DOT ABOVE
V̇	&Vdot; = V + &combdot;	E34C = 0056 + 0307	LATIN CAPITAL LETTER V WITH DOT ABOVE = LATIN CAPITAL LETTER V + COMBINING DOT ABOVE
v̇	&vdot; = v + &combdot;	E74C = 0076 + 0307	LATIN SMALL LETTER V WITH DOT ABOVE = LATIN SMALL LETTER V + COMBINING DOT ABOVE
ṗ	&Vinsdot; = &Vins; + &combdot;	E3E7 = F210 + 0307	LATIN CAPITAL LETTER INSULAR V (VENTH) WITH DOT ABOVE = LATIN CAPITAL LETTER INSULAR V (VENTH) + COMBINING DOT ABOVE
ṗ	&vinsdot; = &vins; + &combdot;	E7E7 = F211 + 0307	LATIN SMALL LETTER INSULAR V (VENTH) WITH DOT ABOVE = LATIN SMALL LETTER INSULAR V (VENTH) + COMBINING DOT ABOVE

Subrange 19: Characters with dot below

Unicode 4.0 has no fewer than 38 characters with a dot below, i.e. essientially the whole alphabet, 'A–Z' and 'a–z', in *Latin Extended Additional*. The list below contains additional character combinations. The combining dot below, 0323, is located in *Combining Diacritical Marks* (p. 33).

Note that the position of the dot below raises problems for the font designer, especially in connection with characters that have descenders or diacritics such as the ogonek. Individually designed glyphs will thus be useful even with a smart font solution.

Glyph	MUFI entity	MUFI code point	MUFI descriptive name
A̤A	&AAligdotbl; = &AAlig; + &combdotbl;	EFF2 = EF90 + 0323	LATIN CAPITAL LIGATURE AA WITH DOT BELOW = LATIN CAPITAL LIGATURE AA + COMBINING DOT BELOW
a̤a	&aaligdotbl; = &aalig; + &combdotbl;	EFF3 = EF91 + 0323	LATIN SMALL LIGATURE AA WITH DOT BELOW = LATIN SMALL LIGATURE AA + COMBINING DOT BELOW
A̤O	&AOligdotbl; = &AOlig; + &combdotbl;	EFF4 = EF92 + 0323	LATIN CAPITAL LIGATURE AO WITH DOT BELOW = LATIN CAPITAL LIGATURE AO + COMBINING DOT BELOW
a̤o	&aoligdotbl; = &aolig; + &combdotbl;	EFF5 = EF93 + 0323	LATIN SMALL LIGATURE AO WITH DOT BELOW = LATIN SMALL LIGATURE AO + COMBINING DOT BELOW
A̤J	&AUligdotbl; = &AUlig; + &combdotbl;	EFF6 = EF94 + 0323	LATIN CAPITAL LIGATURE AU WITH DOT BELOW = LATIN CAPITAL LIGATURE AU + COMBINING DOT BELOW
a̤ı	&auligdotbl; = &aulig; + &combdotbl;	EFF7 = EF95 + 0323	LATIN SMALL LIGATURE AU WITH DOT BELOW = LATIN SMALL LIGATURE AU + COMBINING DOT BELOW
A̤V	&AVligdotbl; = &AVlig; + &combdotbl;	EFF8 = EF96 + 0323	LATIN CAPITAL LIGATURE AV WITH DOT BELOW = LATIN CAPITAL LIGATURE AV + COMBINING DOT BELOW

᷎ꜹ	&avligdotbl; = &avlig; + &combdotbl;	EFF9 = EF97 + 0323	LATIN SMALL LIGATURE AV WITH DOT BELOW = LATIN SMALL LIGATURE AV + COMBINING DOT BELOW
Aꝡ	&AYligdotbl; = &AYlig; + &combdotbl;	EFFA = EF98 + 0323	LATIN CAPITAL LIGATURE AY WITH DOT BELOW = LATIN CAPITAL LIGATURE AY + COMBINING DOT BELOW
aꝡ	&ayligdotbl; = &aylig; + &combdotbl;	EFFB = EF99 + 0323	LATIN SMALL LIGATURE AY WITH DOT BELOW = LATIN SMALL LIGATURE AY + COMBINING DOT BELOW
ʙ	&bscapdotbl; = &bscap; + &combdotbl;	EF25 = 0299 + 0323	LATIN LETTER SMALL CAPITAL B WITH DOT BELOW = LATIN LETTER SMALL CAPITAL B + COMBINING DOT BELOW
Ç	&Cdotbl; = C + &combdotbl;	E066 = 0043 + 0323	LATIN CAPITAL LETTER C WITH DOT BELOW = LATIN CAPITAL LETTER C + COMBINING DOT BELOW
ç	&cdotbl; = c + &combdotbl;	E466 = 0063 + 0323	LATIN SMALL LETTER C WITH DOT BELOW = LATIN SMALL LETTER C + COMBINING DOT BELOW
ᴅ	&dscapdotbl; = &dscap; + &combdotbl;	EF26 = 1D05 + 0323	LATIN LETTER SMALL CAPITAL D WITH DOT BELOW = LATIN LETTER SMALL CAPITAL D + COMBINING DOT BELOW
Ð	&ETHdotbl; = Ð + &combdotbl;	E08F = 00D0 + 0323	LATIN CAPITAL LETTER ETH WITH DOT BELOW = LATIN CAPITAL LETTER ETH + COMBINING DOT BELOW
ð	&ethdotbl; = ð + &combdotbl;	E48F = 00F0 + 0323	LATIN SMALL LETTER ETH WITH DOT BELOW = LATIN SMALL LETTER ETH + COMBINING DOT BELOW
Ę	&Eogondotbl; = Ę + &combdotbl;	E0E8 = 0118 + 0323	LATIN CAPITAL LETTER E WITH OGONEK AND DOT BELOW = LATIN CAPITAL LETTER E WITH OGONEK + COMBINING DOT BELOW
ę	&eogondotbl; = ę + &combdotbl;	E4E8 = 0119 + 0323	LATIN SMALL LETTER E WITH OGONEK AND DOT BELOW = LATIN SMALL LETTER E WITH OGONEK + COMBINING DOT BELOW

Ḟ	&Fdotbl; = F + &combdotbl;	E0EE = 0046 + 0323	LATIN CAPITAL LETTER F WITH DOT BELOW = LATIN CAPITAL LETTER F + COMBINING DOT BELOW
ḟ	&fdotbl; = f + &combdotbl;	E4EE = 0066 + 0323	LATIN SMALL LETTER F WITH DOT BELOW = LATIN SMALL LETTER F + COMBINING DOT BELOW
ꝼ̣	&Finsdotbl; = &Fins; + &combdotbl;	E3E5 = F10C + 0323	LATIN CAPITAL LETTER INSULAR F WITH DOT BELOW = LATIN CAPITAL LETTER INSULAR F + COMBINING DOT BELOW
ꝼ̣	&finsdotbl; = &fins; + &combdotbl;	E7E5 = F10D + 0323	LATIN SMALL LETTER INSULAR F WITH DOT BELOW = LATIN SMALL LETTER INSULAR F + COMBINING DOT BELOW
G̣	&Gdotbl; = G + &combdotbl;	E101 = 0047 + 0323	LATIN CAPITAL LETTER G WITH DOT BELOW = LATIN CAPITAL LETTER G + COMBINING DOT BELOW
g̣	&gdotbl; = g + &combdotbl;	E501 = 0067 + 0323	LATIN SMALL LETTER G WITH DOT BELOW = LATIN SMALL LETTER G + COMBINING DOT BELOW
ɢ̣	&gscapdotbl; = &gscap; + &combdotbl;	EF27 = 0262 + 0323	LATIN LETTER SMALL CAPITAL G WITH DOT BELOW = LATIN LETTER SMALL CAPITAL G + COMBINING DOT BELOW
J̣	&Jdotbl; = J + &combdotbl;	E151 = 004A + 0323	LATIN CAPITAL LETTER J WITH DOT BELOW = LATIN CAPITAL LETTER J + COMBINING DOT BELOW
j̣	&jdotbl; = j + &combdotbl;	E551 = 006A + 0323	LATIN SMALL LETTER J WITH DOT BELOW = LATIN SMALL LETTER J + COMBINING DOT BELOW
ʟ̣	&lscapdotbl; = &lscap; + &combdotbl;	EF28 = 029F + 0323	LATIN LETTER SMALL CAPITAL L WITH DOT BELOW = LATIN LETTER SMALL CAPITAL L + COMBINING DOT BELOW
ᴍ̣	&mscapdotbl; = &mscap; + &combdotbl;	EF29 = 1D0D + 0323	LATIN LETTER SMALL CAPITAL M WITH DOT BELOW = LATIN LETTER SMALL CAPITAL M + COMBINING DOT BELOW

Ṇ	&nscapdotbl; = &nscap; + &combdotbl;	EF2A = 0274 + 0323	LATIN LETTER SMALL CAPITAL N WITH DOT BELOW = LATIN LETTER SMALL CAPITAL N + COMBINING DOT BELOW
Ǫ	&Oogondotbl; = &Oogon; + &combdotbl;	E208 = 01EA + 0323	LATIN CAPITAL LETTER O WITH OGONEK AND DOT BELOW = LATIN CAPITAL LETTER O WITH OGONEK+ COMBINING DOT BELOW
ǫ	&oogondotbl; = &oogon; + &combdotbl;	E608 = 01EB + 0323	LATIN SMALL LETTER O WITH OGONEK AND DOT BELOW = LATIN SMALL LETTER O WITH OGONEK + COMBINING DOT BELOW
OO	&OOligdotbl; = &OOlig; + &combdotbl;	EFFC = F20A + 0323	LATIN CAPITAL LIGATURE OO WITH DOT BELOW = LATIN CAPITAL LIGATURE OO + COMBINING DOT BELOW
oo	&ooligdotbl; = &oolig; + &combdotbl;	EFFD = F20B + 0323	LATIN SMALL LIGATURE OO WITH DOT BELOW = LATIN SMALL LIGATURE OO + COMBINING DOT BELOW
Ṗ	&Pdotbl; = P + &combdotbl;	E26D = 0050 + 0323	LATIN CAPITAL LETTER P WITH DOT BELOW = LATIN CAPITAL LETTER P + COMBINING DOT BELOW
ṗ	&pdotbl; = p + &combdotbl;	E66D = 0070 + 0323	LATIN SMALL LETTER P WITH DOT BELOW = LATIN SMALL LETTER P + COMBINING DOT BELOW
Q̣	&Qdotbl; = Q + &combdotbl;	E288 = 0051 + 0323	LATIN CAPITAL LETTER Q WITH DOT BELOW = LATIN CAPITAL LETTER Q + COMBINING DOT BELOW
q̣	&qdotbl; = q + &combdotbl;	E688 = 0071 + 0323	LATIN SMALL LETTER Q WITH DOT BELOW = LATIN SMALL LETTER Q + COMBINING DOT BELOW
Ṛ	&rscapdotbl; = &rscap; + &combdotbl;	EF2B = 0280 + 0323	LATIN LETTER SMALL CAPITAL R WITH DOT BELOW = LATIN LETTER SMALL CAPITAL R + COMBINING DOT BELOW
ꝛ̣	&rrotdotbl; = &rrot; + &combdotbl;	E7C1 = F20E + 0323	LATIN SMALL LETTER R ROTUNDA WITH DOT BELOW = LATIN SMALL LETTER R ROTUNDA + COMBINING DOT BELOW

ꜱ̣	&sscapdotbl; = &sscap; + &combdotbl;	EF2C = EF0E + 0323	LATIN LETTER SMALL CAPITAL S WITH DOT BELOW = LATIN LETTER SMALL CAPITAL S + COMBINING DOT BELOW
ſ̣	&slongdotbl; = &slong; + &combdotbl;	E7C2 = 017F + 0323	LATIN SMALL LETTER LONG S WITH DOT BELOW = LATIN SMALL LETTER LONG S + COMBINING DOT BELOW
ᴛ̣	&tscapdotbl; = &tscap; + &combdotbl;	EF2D = 1D1B + 0323	LATIN LETTER SMALL CAPITAL T WITH DOT BELOW = LATIN LETTER SMALL CAPITAL T + COMBINING DOT BELOW
Þ̣	&THORNdotbl; = Þ + &combdotbl;	E39F = 00DE + 0323	LATIN CAPITAL LETTER THORN WITH DOT BELOW = LATIN CAPITAL LETTER THORN + COMBINING DOT BELOW
þ̣	&thorndotbl; = þ + &combdotbl;	E79F = 00FE + 0323	LATIN SMALL LETTER THORN WITH DOT BELOW = LATIN SMALL LETTER THORN + COMBINING DOT BELOW
Ꝟ̣	&Vinsdotbl; = &Vins; + &combdotbl;	E3E6 = F210 + 0323	LATIN CAPITAL LETTER INSULAR V (VENTH) WITH DOT BELOW = LATIN CAPITAL LETTER INSULAR V (VENTH) + COMBINING DOT BELOW
ꝟ̣	&vinsdotbl; = &vins; + &combdotbl;	E7E6 = F211 + 0323	LATIN SMALL LETTER INSULAR V (VENTH) WITH DOT BELOW = LATIN SMALL LETTER INSULAR V (VENTH) + COMBINING DOT BELOW

Subrange 20: Characters with diaeresis

Unicode 4.0 has double dot (diaeresis) over the vowels 'A', 'a', 'E', 'e', 'I', 'i', 'O', 'o', 'U', 'u', and 'y' in *Latin-1 Supplement* (p. 15 above), and over 'Y' in *Latin Extended-A* (p. 23), as well as over the consonants 'H', 'h', 'W', 'w', 'X', 'x' and 't' in *Latin Extended Additional* (p. 38).

Combinations with consonants and diaeresis are presumably rare in Medieval sources and have therefore not been included in the present recommendation. The list below contains additional vowels that do appear with diaeresis.

The combining diaeresis, 0308, is located in *Combining Diacritical Marks* (p. 33).

Glyph	MUFI entity	MUFI code point	MUFI descriptive name
Ä̈	&AAliguml; = &AAlig; + &combuml;	EFFE = EF90 + 0308	LATIN CAPITAL LIGATURE AA WITH DIAERESIS = LATIN CAPITAL LIGATURE AA + COMBINING DIAERESIS
ä̈	&aaliguml; = &aalig; + &combuml;	EFFF = EF91 + 0308	LATIN SMALL LIGATURE AA WITH DIAERESIS = LATIN SMALL LIGATURE AA + COMBINING DIAERESIS
Æ̈	&AEliguml; = Æ + &combuml;	E042 = 00C6 + 0308	LATIN CAPITAL LIGATURE AE WITH DIAERESIS = LATIN CAPITAL LIGATURE AE + COMBINING DIAERESIS
æ̈	&aeliguml; = æ + &combuml;	E442 = 00E6 + 0308	LATIN SMALL LIGATURE AE WITH DIAERESIS = LATIN CAPITAL LIGATURE AE + COMBINING DIAERESIS

Subrange 21: Characters with hook above

Unicode 4.0 has a hook above (curl) in combination with 'A', 'a', 'I', 'i', 'O', 'o', 'Y' and 'y' in *Latin Extended Additional* (p. 38 above). However, this curl is loose and has a different design from the one used in e.g. Medieval Nordic sources, so for this reason these and a few other characters have been given separate code points in the list below. The combining curls, F1C5 (high position) and F1C4 (medium-high position), are located in PUA subrange 6 *Combining marks* (p. 61).

Glyph	MUFI entity	MUFI code point	MUFI descriptive name
Å	&Acurl; = A + &combcurl;	E033 = 0041 + F1C5	LATIN CAPITAL LETTER A WITH CURL = LATIN CAPITAL LETTER A + COMBINING CURL
ả	&acurl; = a + &combcurl;	E433 = 0061 + F1C4	LATIN SMALL LETTER A WITH CURL = LATIN SMALL LETTER A + COMBINING CURL
Ẻ	&Ecurl; = E + &combcurl;	E0E9 = 0035 + F1C5	LATIN CAPITAL LETTER E WITH CURL = LATIN CAPITAL LETTER E + COMBINING CURL
ẻ	&ecurl; = e + &combcurl;	E4E9 = 0065 + F1C4	LATIN SMALL LETTER E WITH CURL = LATIN SMALL LETTER E + COMBINING CURL
Ỉ	&Icurl; = I + &combcurl;	E12A = 0049 + F1C5	LATIN CAPITAL LETTER I WITH CURL = LATIN CAPITAL LETTER I + COMBINING CURL
ỉ	&icurl; = i + &combcurl;	E52A = 0131 + F1C4	LATIN SMALL LETTER I WITH CURL = LATIN SMALL LETTER DOTLESS I + COMBINING CURL
Ĵ	&Jcurl; = J + &combcurl;	E163 = 0049 + F1C5	LATIN CAPITAL LETTER J WITH CURL = LATIN CAPITAL LETTER J + COMBINING CURL
ĵ	&jcurl; = j + &combcurl;	E563 = F40B + F1C4	LATIN SMALL LETTER J WITH CURL = LATIN SMALL LETTER J + COMBINING CURL

Ȯ	&Ocurl; = O + &combcurl;	E3D3 = 004F + F1C5	LATIN CAPITAL LETTER O WITH CURL = LATIN CAPITAL LETTER O + COMBINING CURL
ȯ	&ocurl; = o + &combcurl;	E7D3 = 006F + F1C4	LATIN SMALL LETTER O WITH CURL = LATIN SMALL LETTER O + COMBINING CURL
Ø̇	&Oslashcurl; = Ø + &combcurl;	E3D4 = 004F + F1C5	LATIN CAPITAL LETTER O WITH STROKE AND CURL = LATIN CAPITAL LETTER O WITH STROKE + COMBINING CURL
ø̇	&oslashcurl; = ø + &combcurl;	E7D4 = 0048 + F1C4	LATIN SMALL LETTER O WITH STROKE AND CURL = LATIN SMALL LETTER O WITH STROKE + COMBINING CURL
U̇	&Ucurl; = U + &combcurl;	E331 = 0055 + F1C5	LATIN CAPITAL LETTER U WITH CURL = LATIN CAPITAL LETTER U + COMBINING CURL
u̇	&ucurl; = u + &combcurl;	E731 = 0075 + F1C4	LATIN SMALL LETTER U WITH CURL = LATIN SMALL LETTER U + COMBINING CURL
Ẏ	&Ycurl; = Y + &combcurl;	E385 = 0059 + F1C5	LATIN CAPITAL LETTER Y WITH CURL = LATIN CAPITAL LETTER Y + COMBINING CURL
ẏ	&ycurl; = y + &combcurl;	E785 = 0079 + F1C4	LATIN SMALL LETTER Y WITH CURL = LATIN SMALL LETTER Y + COMBINING CURL

Subrange 22: Characters with ogonek

Unicode 4.0 has ogonek with 'A', 'a', 'E' and 'e' in *Latin Extended-A* (p. 23 above), and 'O' and 'o' in *Latin Extended-B* (p. 27). The list below contains additional character combinations. The combining ogonek, 0328, is located in *Combining Diacritical Marks* (p. 33).

Glyph	MUFI entity	MUFI code point	MUFI descriptive name
Æ	&AEligogon; = Æ + &combogon;	E040 = 00C6 + 0328	LATIN CAPITAL LIGATURE AE WITH OGONEK = LATIN CAPITAL LIGATURE AE + COMBINING OGONEK
æ	&aeligogon; = æ + &combogon;	E440 = 00E6 + 0328	LATIN SMALL LIGATURE AE WITH OGONEK = LATIN SMALL LIGATURE AE + COMBINING OGONEK
Ǫ	&Oogoncurl; = O + &combogon; + &combcurl;	E24F = 004F + 0328 + F1C5	LATIN CAPITAL LETTER O WITH OGONEK AND CURL = LATIN CAPITAL LETTER O + COMBINING OGONEK + COMBINING CURL
ǫ	&oogoncurl; = o + &combogon; + &combcurl;	E64F = 006F + 0328 + F1C4	LATIN SMALL LETTER O WITH OGONEK AND CURL = LATIN SMALL LETTER O + COMBINING OGONEK + COMBINING CURL
Ø	&Oslashogon; = Ø + &combogon;	E255 = 00D8 + 0328	LATIN CAPITAL LETTER O WITH STROKE AND OGONEK = LATIN CAPITAL LETTER O WITH STROKE + COMBINING OGONEK
ø	&oslashogon; = ø + &combogon;	E655 = 00F8 + 0328	LATIN SMALL LETTER O WITH STROKE AND OGONEK = LATIN SMALL LETTER O WITH STROKE + COMBINING OGONEK
ţ	&togon; = t + &combogon;	E6EE = 0074 + 0328	LATIN SMALL LETTER T WITH OGONEK = LATIN SMALL LETTER T + COMBINING OGONEK

Subrange 23: Characters with breve

Unicode 4.0 has breve over the vowels 'A', 'a', 'E', 'e', 'I', 'i', 'O', 'o', 'U' and 'u', and also over the consonant 'g', all in *Latin Extended-A* (p. 23 above). The list below contains additional character combinations. The combining breve, 0306, is located in *Combining Diacritical Marks* (p. 33).

Note that all four characters in the list below have duplicate code points in the present version of Junicode (0.6.3).

Glyph	MUFI entity	MUFI code point	(Junicode)	MUFI descriptive name
Ǽ	&AEligbreve; = Æ + &combbreve;	E03F = 00C6 + 0306	F13C	LATIN CAPITAL LIGATURE AE WITH BREVE = LATIN CAPITAL LIGATURE AE + COMBINING BREVE
ǽ	&aeligbreve; = æ + &combbreve;	E43F = 00E6 + 0306	F13D	LATIN SMALL LIGATURE AE WITH BREVE = LATIN SMALL LIGATURE AE + COMBINING BREVE
Y̆	&Ybreve; = Y + &combbreve;	E376 = 0054 + 0306	F140	LATIN CAPITAL LETTER Y WITH BREVE = LATIN CAPITAL LETTER Y + COMBINING BREVE
y̆	&ybreve; = y + &combbreve;	E776 = 0079 + 0306	F141	LATIN SMALL LETTER Y WITH BREVE = LATIN SMALL LETTER Y + COMBINING BREVE

Subrange 24: Characters with macron and breve

Unicode 4.0 has no precomposed characters with both macron and breve. The list below contains expected character combinations.

Note that seven of the characters in the list below have duplicate code points in the present version of Junicode (0.6.3): E410, E43D, E4B7, E537, E61B, E70B and E775. The combining macron, 0304, and the combining breve, 0306, are located in *Combining Diacritical Marks* (p. 32–33).

Glyph	MUFI entity	MUFI code point	(Junicode)	MUFI descriptive name
Ä̆	&Amacrbreve; = a + &combmacr; + &combbreve;	E010 = 0041 + 0304 + 0306		LATIN CAPITAL LETTER A WITH MACRON AND BREVE = LATIN CAPITAL LETTER A + COMBINING MACRON + COMBINING BREVE
ä̆	&amacrbreve; = a + &combmacr; + &combbreve;	E410 = 0061 + 0304 + 0306	F171	LATIN SMALL LETTER A WITH MACRON AND BREVE = LATIN SMALL LETTER A + COMBINING MACRON + COMBINING BREVE
Æ̆	&AEligmacrbreve; = Æ + &combmacr; + &combbreve;	E03D = 00C6 + 0304 + 0306		LATIN CAPITAL LIGATURE AE WITH MACRON AND BREVE = LATIN CAPITAL LIGATURE AE + COMBINING MACRON + COMBINING BREVE
æ̆	&aeligmacrbreve; = æ + &combmacr; + &combbreve;	E43D = 00E6 + 0304 + 0306	F173	LATIN SMALL LIGATURE AE WITH MACRON AND BREVE = LATIN SMALL LIGATURE AE + COMBINING MACRON + COMBINING BREVE
Ĕ	&Emacrbreve; = E + &combmacr; + &combbreve;	E0B7 = 0035 + 0304 + 0306		LATIN CAPITAL LETTER E WITH MACRON AND BREVE = LATIN CAPITAL LETTER E + COMBINING MACRON + COMBINING BREVE

ĕ̄	&emacrbreve; = e + &combmacr; + &combbreve;	E4B7 = 0065 + 0304 + 0306	F175	LATIN SMALL LETTER E WITH MACRON AND BREVE = LATIN SMALL LETTER E + COMBINING MACRON + COMBINING BREVE
Ǐ	&Imacrbreve; = I + &combmacr; + &combbreve;	E137 = 0049 + 0304 + 0306		LATIN CAPITAL LETTER I WITH MACRON AND BREVE = LATIN CAPITAL LETTER I + COMBINING MACRON + COMBINING BREVE
ǐ̄	&imacrbreve; = i + &combmacr; + &combbreve;	E537 = 0069 + 0304 + 0306	F177	LATIN SMALL LETTER I WITH MACRON AND BREVE = LATIN SMALL LETTER I + COMBINING MACRON + COMBINING BREVE
Ǒ	&Omacrbreve; = O + &combmacr; + &combbreve;	E21B = 004F + 0304 + 0306		LATIN CAPITAL LETTER O WITH MACRON AND BREVE = LATIN CAPITAL LETTER O + COMBINING MACRON + COMBINING BREVE
ǒ̄	&omacrbreve; = o + &combmacr; + &combbreve;	E61B = 006F + 0304 + 0306	F179	LATIN SMALL LETTER O WITH MACRON AND BREVE = LATIN SMALL LETTER O + COMBINING MACRON + COMBINING BREVE
Ǔ	&Umacrbreve; = U + &combmacr; + &combbreve;	E30B = 0055 + 0304 + 0306		LATIN CAPITAL LETTER U WITH MACRON AND BREVE = LATIN CAPITAL LETTER U + COMBINING MACRON + COMBINING BREVE
ǔ̄	&umacrbreve; = u + &combmacr; + &combbreve;	E70B = 0075 + 0304 + 0306	F17B	LATIN SMALL LETTER U WITH MACRON AND BREVE = LATIN SMALL LETTER U + COMBINING MACRON + COMBINING BREVE
Y̆̄	&Ymacrbreve; = Y + &combmacr; + &combbreve;	E375 = 0054 + 0304 + 0306		LATIN CAPITAL LETTER Y WITH MACRON AND BREVE = LATIN CAPITAL LETTER Y + COMBINING MACRON + COMBINING BREVE

	&ymacrbreve; = y + &combmacr; + &combbreve;	E775 = 0079 + 0304 + 0306	F17D	LATIN SMALL LETTER Y WITH MACRON AND BREVE = LATIN SMALL LETTER Y + COMBINING MACRON + COMBINING BREVE

Subrange 25: Characters with macron and acute

Unicode 4.0 has macron and acute over the vowels 'E', 'e', 'O' and 'o' in *Latin Extended Additional* (p. 38 above). The list below contains additional character combinations.

Note that five characters in the list below have duplicate code points in the present version of Junicode (0.6.3): E40A, E43A, E535, E709 and E773. In addition, Junicode has code points for 'e' with macron and acute, F185, and 'o' with macron and acute, F189. Since the latter two characters are in *Latin Extended Additional*, at code points 1E16 and 1E52 respectively, they might be considered for deprecation (cf. pp. 38 and 40 above).

The combining macron, 0304, and the combining acute accent, 0301, are located in *Combining Diacritical Marks* (p. 32).

Glyph	MUFI entity	MUFI code point	(Junicode)	MUFI descriptive name
Á	&Amacracute; = a + &combmacr; + &combacute;	E00A = 0041 + 0304 + 0301		LATIN CAPITAL LETTER A WITH MACRON AND ACUTE = LATIN CAPITAL LETTER A + COMBINING MACRON + COMBINING ACUTE
á	&amacracute; = a + &combmacr; + &combacute;	E40A = 0061 + 0304 + 0301	F181	LATIN SMALL LETTER A WITH MACRON AND ACUTE = LATIN SMALL LETTER A + COMBINING MACRON + COMBINING ACUTE
Ǽ	&AEligmacracute; = Æ + &combmacr; + &combacute;	E03A = 00C6 + 0304 + 0301		LATIN CAPITAL LIGATURE AE WITH MACRON AND ACUTE = LATIN CAPITAL LIGATURE AE + COMBINING MACRON + COMBINING ACUTE

ǽ	&aeligmacracute; = æ + &combmacr; + &combacute;	E43A = 00E6 + 0304 + 0301	F183	LATIN SMALL LIGATURE AE WITH MACRON AND ACUTE = LATIN SMALL LIGATURE AE + COMBINING MACRON + COMBINING ACUTE
Ī́	&Imacracute; = I + &combmacr; + &combacute;	E135 = 0049 + 0304 + 0301		LATIN CAPITAL LETTER I WITH MACRON AND ACUTE = LATIN CAPITAL LETTER I + COMBINING MACRON + COMBINING ACUTE
ī́	&imacracute; = i + &combmacr; + &combacute;	E535 = 0069 + 0304 + 0301	F187	LATIN SMALL LETTER I WITH MACRON AND ACUTE = LATIN SMALL LETTER I + COMBINING MACRON + COMBINING ACUTE
Ū́	&Umacracute; = U + &combmacr; + &combacute;	E309 = 0055 + 0304 + 0301		LATIN CAPITAL LETTER U WITH MACRON AND ACUTE = LATIN CAPITAL LETTER U + COMBINING MACRON + COMBINING ACUTE
ū́	&umacracute; = u + &combmacr; + &combacute;	E709 = 0075 + 0304 + 0301	F18B	LATIN SMALL LETTER U WITH MACRON AND ACUTE = LATIN SMALL LETTER U + COMBINING MACRON + COMBINING ACUTE
Ȳ́	&Ymacracute; = Y + &combmacr; + &combacute;	E373 = 0054 + 0304 + 0301		LATIN CAPITAL LETTER Y WITH MACRON AND ACUTE = LATIN CAPITAL LETTER Y + COMBINING MACRON + COMBINING ACUTE
ȳ́	&ymacracute; = y + &combmacr; + &combacute;	E773 = 0079 + 0304 + 0301	F18D	LATIN SMALL LETTER Y WITH MACRON AND ACUTE = LATIN SMALL LETTER Y + COMBINING MACRON + COMBINING ACUTE

Subranges 26–29:

Reserved for future additions.

Category 3: Variant letter forms

In general, variant letter forms are not characters in the Unicode sense of the word. However, the distinction between characters and variant letter forms is not always clear, and some of the letter forms in the list may eventually prove to be recognised as characters according to the strict Unicode criteria.

This category will be divided according to national or regional criteria. Note that many characters are relevant for more than one range, due to the dissemination of styles across Medieval Europe.

Subrange 30: Nordic

This subrange contains commonly recognised variant letter forms in Medieval Nordic primary sources. Note that the majority of these characters are inherited from Insular and Continental styles, and will thus be relevant for many other European literatures.

To minimise conflicts with Junicode, code points in this font have been used wherever possible. Additional letter forms have been placed in a new section of the PUA commencing with F200.

Glyph	MUFI entity	MUFI code point	(Junicode)	MUFI descriptive name
ɑ	&ains;	F200		LATIN SMALL LETTER INSULAR A
This is the single-storey 'a' of the Insular style, similar in shape to modern italic 'a'.				
ɑ	&ainsenl;	F201		LATIN ENLARGED SMALL LETTER INSULAR A
This is the enlarged variant of 'a', triangular in shape. Cf. D.A. Seip, *Palæografi: Norge og Island* (Nordisk kultur 23:B), Oslo etc. 1954, p. 27.				
ɑ	&aopen;	F202		LATIN SMALL LETTER A OPEN FORM

This is the traditional 'open a' from early Carolingian script, similar in shape to 'cc' and easily mistaken for 'u'.				
ɑ	&aclose;	F203		LATIN SMALL LETTER A CLOSED FORM
Cf. D.A. Seip, *Palæografi: Norge og Island* (Nordisk kultur 23:B), Oslo etc. 1954, p. 70.				
æ	&aeligred;	F204		LATIN SMALL LIGATURE AE WITH MISSING BOTTOM STROKE
Cf. Johs. Brøndum-Nielsen, ed., *Palæografi: Danmark og Sverige* (Nordisk kultur 23:A), Stockholm etc. [1943], p. 102.				
ΑO	&AOligred;	F205		LATIN CAPITAL LIGATURE AO NECKLESS
ɑo	&aoligred;	F206		LATIN SMALL LIGATURE AO NECKLESS
These two letter forms should be analysed and designed as the ligature of 'a+o' without the neck of the 'a'.				
ð	&dunc;	F109	F109	LATIN LETTER UNCIAL D
This is similar to a separate character in Medieval Welsh.				
ϵ	&eunc;	F10A	F10A	LATIN LETTER UNCIAL E
ꝼ	&Fins;	F10C		LATIN CAPITAL LETTER INSULAR F
The capital Insular 'f' has not been included in the present version of Junicode (0.6.3), but since the small variant is located at F10D, the code point F10C has been selected here.				
ꝼ	&fins;	F10D	F10D	LATIN SMALL LETTER INSULAR F

ꝼ	&finsclose;	F207		LATIN SMALL LETTER INSULAR F CLOSED FORM
ħ	&hins;	F110	F110	LATIN CAPITAL LETTER INSULAR H
ȷ	&jdl;	E55C	F6BE	LATIN SMALL LETTER J DOTLESS

The dotless 'j' has been accepted by the UTC and WG2 and is likely to be included in a future version of the Unicode Standard at code point 0237 in *Latin Extended-B*. Since this character already is in Titus Cyberbit at E55C, this code point has been chosen rather than F6BE in Junicode.

ꝁ	&kunc;	F208		LATIN LETTER UNCIAL K
ꝃ	&kclose;	F209		LATIN SMALL LETTER K CLOSED FORM
ꟿ	&munc;	F11A	F11A	LATIN LETTER UNCIAL M
ꝏ	&OOlig;	F20A		LATIN CAPITAL LIGATURE OO
ꝏ	&oolig;	F20B		LATIN SMALL LIGATURE OO

The 'o+o' ligature was introduced by the Icelander Jón Ólafsson from Grunnavík in the 18th century. Cf. Jón Helgason, *Jon Olafsson fra Grunnavik*, Copenhagen 1926.

ꝋ	&OEligred;	F20C		LATIN CAPITAL LIGATURE OE WITH MISSING BOTTOM STROKE
ꝋ	&oeligred;	F20D		LATIN SMALL LIGATURE OE WITH MISSING BOTTOM STROKE

These two letter forms should be analysed and designed as the ligature of 'o+e' without the bottom (finishing) stroke of the 'e'.

	&rrot;	F20E		LATIN SMALL LETTER R ROTUNDA
	&rins;	F125	F125	LATIN SMALL LETTER INSULAR R
	&slongdes;	F127	F127	LATIN SMALL LETTER LONG S DESCENDING

This character extends below the base line, while the ordinary long 's' is located on the base line.

	&tunc;	F129	F129	LATIN LETTER UNCIAL T
	&Vins;	F210		LATIN CAPITAL LETTER INSULAR V (VENTH)
	&vins;	F211		LATIN SMALL LETTER INSULAR V (VENTH)

The capital and small Insular 'v' are partially similar to 01F7 LATIN CAPITAL LETTER WYNN and 01BF LATIN LETTER WYNN in *Latin Extended-B*. Note that the Medieval Nordic variants are open, resembling the character 'y', and are transliterated with 'v', not with 'w' like in Old English.

	&YYlig;	F212		LATIN CAPITAL LIGATURE YY
	&yylig;	F213		LATIN SMALL LIGATURE YY

The 'y+y' ligature was introduced by the Icelander Jón Ólafsson from Grunnavík in the 18th century. Cf. Jón Helgason, *Jon Olafsson fra Grunnavik*, Copenhagen 1926.

Additional characters: From F214 and upwards; extent to be decided in the next version of the recommendation.

Subrange 31: English

Several of the specific characters used in Old English primary sources have been listed in subrange 30 above. Additional characters are listed in the present subrange, using Junicode code points wherever possible.

Glyph	MUFI entity	MUFI code point	(Junicode)	MUFI descriptive name
A	&Asqu;	F13A	F13A	LATIN CAPITAL LETTER A SQUARE FORM
This particular variant is found in some manuscripts, particularly CCCC 173, the Parker Chronicle.				
C	&Csqu:	F106	F106	LATIN CAPITAL LETTER C SQUARE FORM
G	&Gsqu;	F10E	F10E	LATIN CAPITAL LETTER G SQUARE FORM
ᵹ	&Gins;	F10B		LATIN CAPITAL LETTER INSULAR G
The capital Insular 'g' has not been included in the present version of Junicode (0.6.3), but since the small variant is located at F10F, the vacant code point F10B has been selected here.				
ᵹ	&gins;	F10F	F10F	LATIN SMALL LETTER INSULAR G
The Insular 'g' has been accepted by WG2 (not yet by UTC) and may be included in a future version of the Unicode Standard at code point 1D79 in *Phonetic Extensions*. Note that the proposal has been based on the usage of this character in phonetic notation of Irish.				
ẞ	&Sclose;		F126	LATIN CAPITAL LETTER S CLOSED FORM
This character is not part of the MUFI recommendation, but has been included with its Junicode code point for the sake of reference. It is possibly a design by the type cutter of the font Junius, on which Junicode is based.				

ſ	&slongbarslash;	F14F	F14F	LATIN SMALL LETTER LONG S WITH DIAGONAL STROKE
A similar character, though with a horizontal bar in a higher position, is assigned to E7FF in PUA subrange 5, p. 59 above.				
þ	&thornbarslash;	F149	F149	LATIN SMALL LETTER THORN WITH DIAGONAL STROKE
A similar character, though with a horizontal bar, is assigned to E73B in PUA subrange 5, p. 59 above.				

Additional characters: These can be placed in vacant slots in the section F100–F13B, where almost 50 unallocated slots are available. Junicode has English letter forms in this section.

Subranges 32–

For future additions.

Appendix: How many characters?

The short answer to this question is that the present recommendation contains 828 characters. However, there are many types of characters in this recommendation, so it may be useful to make a breakdown of characters into a number of categories:

1. Alphabetic, base characters

This category contains the basic letter forms, e.g. 'a', 'b' and 'c'. Case variants are also included, such as small (minuscule), 'a', enlarged minuscule, 'a', capital, 'A', and small capital, 'A'. The same applies to ligatures, e.g. 'æ' and 'œ'.

2. Alphabetic, precomposed characters

This category contains base characters with one or more loose diacritics, such as 'á', 'ä', and 'â'. Characters with fixed diacritics, such as the cedilla or the ogonek, are counted as base characters, and the same applies to characters with a bar or a slash across an ascender or a descender.

3. Symbols

This category contains all kinds of symbols, including digits, but excludes punctuation marks.

4. Combining marks

This category contains accents, abbreviation signs and alphabetic characters that appear above or below another base character.

5. Punctuation marks

This category contains the basic punctuation marks, e.g. the comma and the colon, and various kinds of quotation marks. Note that is is sometimes difficult to draw the line between symbols and punctuation marks.

6. Formatting and space characters

This category includes characters for which there are no glyphs, either because they specify formatting of the text or because they encode space (of variable width) between visible characters.

Range	Alphabetic characters		Symbols	Comb.	Punct.	Format & space	Total
	base	precomp.					
Basic Latin	52		34		8	2	96
Latin-1 Supplement	11	51	28		4	2	96
Latin Extended-A	16	34					50
Latin Extended-B	16	12					28
IPA Extensions	15						15
Spacing Modifying Letters	6						6
Combining Diacritical Marks				31			31
Greek and Coptic	1						1
Georgian					1		1
Runic	2						2
Phonetic Extensions	15						15
Latin Extended Additional		76					76
General Punctuation			20		14	14	48
Mathematical Operators			3				3

Range	Alphabetic characters		Symbols	Comb.	Punct.	Format & space	Total
	base	precomp.					
Dingbats			1				1
Miscellaneous Mathematical Symbols-A			4				4
Characters selected from the Unicode Standard v. 4.0	134	173	88	31	27	18	473
A: Base characters							
1: Ligatures	13						13
2: Small capitals	5						5
3: Enlarged minuscules	29						29
4: Base line abbreviation characters	7						7
5: Modified base line abbreviation characters	13						13
6: Combining marks				13			13
7: Combining superscript characters				22			22
8: Punctuation marks					6		6
9: Critical and epigraphical signs			8				8
10: Metrical symbols			23				23

Range	Alphabetic characters		Symbols	Comb.	Punct.	Format & space	Total
	base	precomp.					
11–14: reserved for future additions							
B. Precomposed characters							
15: Characters with overline or macron	8	12		4			24
16: Characters with acute accent		20					20
17: Characters with double acute accent		20					20
18: Characters with dot above		14					14
19: Characters with dot below		47					47
20: Characters with diaeresis		4					4
21: Characters with hook above	16						16
22: Characters with ogonek	7						7
23: Characters with breve		4					4
24: Characters with macron and breve		14					14
25: Characters with macron and acute		10					10
26–29: Reserved for future additions							

Range	Alphabetic characters		Symbols	Comb.	Punct.	Format & space	Total
	base	precomp.					
C: Variant letter forms							
30: Nordic	29						29
31: English	7						7
Characters in the Private Use Area	134	145	33	39	6	0	355
Grand total: Characters selected from the Unicode Standard v. 4.0 + characters in the Private Use Area:							828