General Punctuation

Range: 2000-206F

This file contains an excerpt from the character code tables and list of character names for *The Unicode Standard, Version 9.0*

This file may be changed at any time without notice to reflect errata or other updates to the Unicode Standard. See http://www.unicode.org/errata/for an up-to-date list of errata.

See http://www.unicode.org/charts/ for access to a complete list of the latest character code charts. See http://www.unicode.org/charts/PDF/Unicode-9.0/ for charts showing only the characters added in Unicode 9.0. See http://www.unicode.org/Public/9.0.0/charts/ for a complete archived file of character code charts for Unicode 9.0.

Disclaimer

These charts are provided as the online reference to the character contents of the Unicode Standard, Version 9.0 but do not provide all the information needed to fully support individual scripts using the Unicode Standard. For a complete understanding of the use of the characters contained in this file, please consult the appropriate sections of The Unicode Standard, Version 9.0, online at http://www.unicode.org/versions/Unicode9.0.0/, as well as Unicode Standard Annexes #9, #11, #14, #15, #24, #29, #31, #34, #38, #41, #42, #44, and #45, the other Unicode Technical Reports and Standards, and the Unicode Character Database, which are available online.

See http://www.unicode.org/ucd/ and http://www.unicode.org/reports/

A thorough understanding of the information contained in these additional sources is required for a successful implementation.

Fonts

The shapes of the reference glyphs used in these code charts are not prescriptive. Considerable variation is to be expected in actual fonts. The particular fonts used in these charts were provided to the Unicode Consortium by a number of different font designers, who own the rights to the fonts.

See http://www.unicode.org/charts/fonts.html for a list.

Terms of Use

You may freely use these code charts for personal or internal business uses only. You may not incorporate them either wholly or in part into any product or publication, or otherwise distribute them without express written permission from the Unicode Consortium. However, you may provide links to these charts.

The fonts and font data used in production of these code charts may NOT be extracted, or used in any other way in any product or publication, without permission or license granted by the typeface owner(s).

The Unicode Consortium is not liable for errors or omissions in this file or the standard itself. Information on characters added to the Unicode Standard since the publication of the most recent version of the Unicode Standard, as well as on characters currently being considered for addition to the Unicode Standard can be found on the Unicode web site.

See http://www.unicode.org/pending/pending.html and http://www.unicode.org/alloc/Pipeline.html.

Copyright © 1991-2016 Unicode, Inc. All rights reserved.

	200	201	202	203	204	205	206
0	NQ SP	2010	2020	%oo	2040	2050	WJ 2060
1	MQ SP	NB 2011	2021	%000	 2041	* * 2051	[<u>f()</u>]
2	EN SP 2002	2012	2022	2032	* ** 2042	% 2052	2062
3	EM SP 2003	2013	2023	2033	2043	2053	2063
4	3/M S P		•	""	/	(+
5	2004 4/M S P	2014	2024	2034	2044	2054 *	2064
6	2005 6/M S P	2015	2025	2035	2045	2055	LRI
7	2006 F S P	2016	2026	2036	2046	2056	2066 RLI
8	2007 P S P	2017	2027	2037	2047	2057	2067
9	2008 T H	2018	2028 P SEP	2038	1?	2058	2068 PDI
Α	2009 H S P	2019	2029 [LRE]	2039 >	2049	2059	2069 [S S]
В	ZW SP	201A •	202A	203A ※	204A	205A	206A A S S
С	200B ZW N J	201B	202B	203B	204B I I	205B	206B
D	200C	201C	202C	203C	204C	205C • •	206C A AFS
E	200D	201D	202D	203D	204D	205D	206D N A D S
F	200E	99 201E	202E	203E	204E	205E	206E
	200F	201F	202F	203F	୧ 204F	205F	206F

For additional general punctuation characters see also Basic Latin, Latin-1, Supplemental Punctuation and CJK Symbols and Punctuation.

Spaces

2004

2000 NQ SP **EN QUAD**

≡ 2002 🔛 en space

2001 **EM QUAD**

= mutton quad

= 2003 ∰ em space

2002 **EN SPACE**

= nut

• half an em

≈ 0020 sp space

2003 **EM SPACE**

= mutton

• nominally, a space equal to the type size in

• may scale by the condensation factor of a font

≈ 0020 sp space

THREE-PER-EM SPACE

= thick space

≈ 0020 sp space

2005 4/M SP FOUR-PER-EM SPACE

= mid space

≈ 0020 sp space

2006 SIX-PER-EM SPACE

> • in computer typography sometimes equated to thin space

≈ 0020 sp space

2007 FIGURE SPACE

• space equal to tabular width of a font

• this is equivalent to the digit width of fonts with fixed-width digits

≈ <noBreak> 0020 sp

PUNCTUATION SPACE 2008 SP

• space equal to narrow punctuation of a font

≈ 0020 sp space

2009 TH THIN SPACE

• a fifth of an em (or sometimes a sixth)

→ 202F narrow no-break space

≈ 0020 sp space

200A SP HAIR SPACE

• thinner than a thin space

• in traditional typography, the thinnest space

≈ 0020 sp space

Format characters

200B ZERO WIDTH SPACE

• commonly abbreviated ZWSP

• this character is intended for invisible word separation and for line break control; it has no width, but its presence between two characters does not prevent increased letter spacing in

justification

200C ZERO WIDTH NON-JOINER

commonly abbreviated ZWNJ

200D **ZERO WIDTH JOINER**

commonly abbreviated ZWJ

200E **LEFT-TO-RIGHT MARK**

• commonly abbreviated LRM

200F **RIGHT-TO-LEFT MARK**

commonly abbreviated RLM

→ 061C ALM arabic letter mark

Dashes

2010 **HYPHEN**

→ 002D - hyphen-minus

→ 00AD [SHY] soft hyphen

2011 **NON-BREAKING HYPHEN**

→ 002D - hyphen-minus

→ 00AD sir soft hyphen

≈ <noBreak> 2010 -

2012 FIGURE DASH

2013 **EN DASH**

2014 **EM DASH**

• may be used in pairs to offset parenthetical text

→ 2E3A [2M] two-em dash

→ 30FC — katakana-hiragana prolonged sound

mark

2015 — HORIZONTAL BAR

= quotation dash

• long dash introducing quoted text

General punctuation

□ DOUBLE VERTICAL LINE

• used in pairs to indicate norm of a matrix

→ 20E6 \$\psi\$ combining double vertical stroke overlay

→ 2225 || parallel to

→ 23F8 II double vertical bar

DOUBLE LOW LINE 2017

• this is a spacing character

 \rightarrow 005F _ low line

 \rightarrow 0333 $\underline{\circ}$ combining double low line

≈ 0020 sp 0333 <u>o</u>

Quotation marks and apostrophe

Use of quotation marks differs by language. The character names cannot reflect actual usage for all languages.

LEFT SINGLE QUOTATION MARK 2018

= single turned comma quotation mark

• this is the preferred character (as opposed to 201B')

→ 0027 ' apostrophe

→ 02BB ' modifier letter turned comma

→ 275B • heavy single turned comma quotation

mark ornament

2019 RIGHT SINGLE QUOTATION MARK

= single comma quotation mark

• this is the preferred character to use for apostrophe

→ 0027 ' apostrophe→ 02BC' modifier letter apostrophe

→ 275C • heavy single comma quotation mark ornament

201A SINGLE LOW-9 QUOTATION MARK

= low single comma quotation mark

• used as opening single quotation mark in some languages

SINGLE HIGH-REVERSED-9 QUOTATION MARK 201B

= single reversed comma quotation mark

• has same semantic as 2018 ', but differs in appearance

→ 02BD 'modifier letter reversed comma

201C	"	LEFT DOUBLE QUOTATION MARK	2029	P SEP	PARAGRAPH SEPARATOR
2010		= double turned comma quotation mark	2023	(SEP)	• may be used to represent this semantic
		• this is the preferred character (as opposed to			unambiguously
		201F")	202A	LRE	LEFT-TO-RIGHT EMBEDDING
		→ 0022 " quotation mark	0000	r1	• commonly abbreviated LRE
		→ 275D ⁶⁶ heavy double turned comma quotation mark ornament	202B	RLE	• commonly abbreviated RLE
		→ 301D " reversed double prime quotation	202C	PDF	POP DIRECTIONAL FORMATTING
		mark	2020	1	• commonly abbreviated PDF
201D	,,	RIGHT DOUBLE QUOTATION MARK	202D	LRO	LEFT-TO-RÍGHT OVERRIDE
		= double comma quotation mark → 0022 " quotation mark		,	 commonly abbreviated LRO
		→ 2033 " double prime	202E	RLO	RIGHT-TO-LEFT OVERRIDE
		→ 275E • heavy double comma quotation mark	202F	NNB SP	commonly abbreviated RLO NARROW NO-BREAK SPACE
		ornament	2021	SP	• commonly abbreviated NNBSP
201E		→ 301E ″ double prime quotation mark DOUBLE LOW-9 QUOTATION MARK			• a narrow form of a no-break space, typically the
2016	,,	= low double comma quotation mark			width of a thin space or a mid space
		• used as opening double quotation mark in			→ 00A0 no-break space
		some languages			→ 2005 [sp] four-per-em space → 2009 [sp] thin space
		→ 2E42 ,, double low-reversed-9 quotation mark			\approx <nobreak> 0020 [SP]</nobreak>
		→ 301F " low double prime quotation mark	Gene	ral n	unctuation
201F	"	DOUBLE HIGH-REVERSED-9 QUOTATION MARK	2030	-	PER MILLE SIGN
		= double reversed comma quotation mark		,	= permille, per thousand
		 has same semantic as 201C ", but differs in appearance 			• used, for example, in measures of blood alcohol
C					content, salinity, etc. → 0025 % percent sign
2020	ai p †	unctuation DAGGER			→ 0609 /- arabic-indic per mille sign
2020	Ť	= obelisk, long cross, oblong cross	2031	‰	PER TEN THOUSAND SIGN
		→ 2E38 ‡ turned dagger			= permyriad
2021	‡	DOUBLE DAGGER			• percent of a percent, rarely used
0000		= diesis, double obelisk			→ 0025 % percent sign → 060A /- arabic-indic per ten thousand sign
2022	•	BULLET = black small circle	2032	,	PRIME
		→ 00B7 · middle dot			= minutes, feet
		\rightarrow 2024 . one dot leader			→ 0027 ' apostrophe
		→ 2219 • bullet operator			→ 00B4 ′ acute accent
		→ 25D8 Inverse bullet	2033	"	→ 02B9 ′ modifier letter prime DOUBLE PRIME
2023		→ 25E6 ∘ white bullet TRIANGULAR BULLET	2000		= seconds, inches
2020		→ 220E ■ end of proof			→ 0022 " quotation mark
		→ 25B8 black right-pointing small triangle			→ 02BA " modifier letter double prime
2024		ONE DOT LEADER			→ 201D" right double quotation mark
		• also used as an Armenian semicolon (mijaket)			→ 3003 " ditto mark → 301E " double prime quotation mark
		→ 00B7 · middle dot → 2022 • bullet			≈ 2032′ 2032′
		→ 2219 • bullet operator	2034	""	TRIPLE PRIME
		≈ 002E. full stop			= lines (old measure, 1/12 of an inch)
2025		TWO DOT LEADER	2025	,	≈ 2032′ 2032′ 2032′
0000		≈ 002E. 002E.	2035		REVERSED PRIME → 0060` grave accent
2026	•••	HORIZONTAL ELLIPSIS = three dot leader	2036	"	REVERSED DOUBLE PRIME
		→ 22EE : vertical ellipsis			→ 301D " reversed double prime quotation
		→ FE19 : presentation form for vertical			mark
		horizontal ellipsis	2027	""	≈ 2035 ` 2035 ` REVERSED TRIPLE PRIME
2027		≈ 002E . 002E . 002E .	2037		≈ 2035 ' 2035 ' 2035 '
2027	•	HYPHENATION POINT visible symbol used to indicate correct	2038	٨	CARET
		positions for word breaking, as in dic-tion-ar-ies		٨	→ 2303 ^ up arrowhead
Forma	at ch	aracters			ightarrow A788 $$ modifier letter low circumflex accent
		LINECEDADATOD			

2028 SEP LINE SEPARATOR

• may be used to represent this semantic unambiguously

Quot	atio	n marks	2049	!?	EXCLAMATION QUESTION MARK		
2039	<	SINGLE LEFT-POINTING ANGLE QUOTATION			~ 2049 FE0E text style		
		MARK			~ 2049 FE0F emoji style		
		= left pointing single guillemet			≈ 0021 ! 003F ?		
	 usually opening, sometimes closing 		Gene	General punctuation			
		→ 003C < less-than sign	204A	٦.	TIRONIAN SIGN ET		
		→ 2329 〈 left-pointing angle bracket		'	• Irish Gaelic, Old English,		
0004		→ 3008 〈 left angle bracket			\rightarrow 0026 & ampersand		
203A	>	SINGLE RIGHT-POINTING ANGLE QUOTATION			→ 1F670 er script ligature et ornament		
		MARK = right pointing single guillemet	204B	P	REVERSED PILCROW SIGN		
		usually closing, sometimes opening			→ 00B6¶ pilcrow sign		
		→ 003E > greater-than sign	204C	■	BLACK LEFTWARDS BULLET		
		→ 232A \right-pointing angle bracket	204D	D	BLACK RIGHTWARDS BULLET		
		→ 3009 > right angle bracket	204E	*	LOW ASTERISK		
C		9 9			→ 002A * asterisk		
	-	ounctuation			→ 0359 ু combining asterisk below		
203B	×	REFERENCE MARK	204F	;	REVERSED SEMICOLON		
		Japanese komeUrdu paragraph separator			• also used in Sindhi		
		 → 0FBF * tibetan ku ru kha bzhi mig can 			→ 003B; semicolon		
		→ 200AD ※ cjk unified ideograph-200AD			→ 061B ' arabic semicolon		
	_	, ,	2050	\Box	CLOSE UP		
	le p	unctuation for vertical text			• editing mark		
203C	!!	DOUBLE EXCLAMATION MARK			→ AB5B × modifier breve with inverted breve		
		→ 0021 ! exclamation mark	2051	*	TWO ASTERISKS ALIGNED VERTICALLY		
		~ 203C FE0E text style	2052	%	COMMERCIAL MINUS SIGN		
		~ 203C FE0F emoji style			= abzüglich (German), med avdrag av (Swedish),		
		≈ 0021 ! 0021 !			piska (Swedish, "whip")		
Gene	ral p	punctuation			 a common glyph variant and fallback representation looks like ./. 		
203D	?	INTERROBANG			 may also be used as a dingbat to indicate 		
		→ 0021! exclamation mark			correctness		
		→ 003F ? question mark			• used in Finno-Ugric Phonetic Alphabet to		
		→ 2E18 inverted interrobang			indicate a related borrowed form with different sound		
2025	_	→ 1F679 ? heavy interrobang ornament			→ 0025 % percent sign		
203E		OVERLINE = spacing overscore			→ 066A : arabic percent sign		
		\approx 0020 [SP] 0305 \odot			→ 00F7 ÷ division sign		
203F		UNDERTIE	2053	~	SWUNG DASH		
2001	_	= Greek enotikon			→ 007E ~ tilde		
		→ 2323 ~ smile	2054		INVERTED UNDERTIE		
2040	$\overline{}$	CHARACTER TIE	2055	*	FLOWER PUNCTUATION MARK		
20.0		= z notation sequence concatenation			= phul, puspika		
		→ 2322 ~ frown			 used as a punctuation mark with Syloti Nagri, 		
2041	λ	CARET INSERTION POINT			Bengali and other Indic scripts		
	^	• proofreader's mark: insert here			→ 274B * heavy eight teardrop-spoked		
		→ 22CC 🔀 right semidirect product			propeller asterisk		
2042	* **	ASTERISM	Archa	aic p	unctuation		
2043	-	HYPHEN BULLET	2056	•:	THREE DOT PUNCTUATION		
		→ 002D - hyphen-minus			→ 10FB: georgian paragraph separator		
2044	/	FRACTION SLASH	Gana	ralr	punctuation		
		= solidus (in typography)		αι μ			
		 for composing arbitrary fractions 	2057		QUADRUPLE PRIME		
		→ 002F / solidus			≈ 2032′ 2032′ 2032′ 2032′		
007-		→ 2215 / division slash	Archa	aic p	unctuation		
2045	Ę	LEFT SQUARE BRACKET WITH QUILL	2058	÷	FOUR DOT PUNCTUATION		
2046	}	RIGHT SQUARE BRACKET WITH QUILL	2059	::	FIVE DOT PUNCTUATION		
Doub	le p	unctuation for vertical text			= Greek pentonkion		
2047	??	DOUBLE QUESTION MARK			= quincunx		
		≈ 003F? 003F?			→ 2684 ☑ die face-5		
2048	?!	QUESTION EXCLAMATION MARK					
-		~ 003E 2 0021 I					

≈ 003F? 0021!

205A	•	 TWO DOT PUNCTUATION historically used to indicate the end of a sentence or change of speaker extends from baseline to cap height → FE30 : presentation form for vertical two dot leader 	206D 206E 206F	A AFS NA DS NO DS	ACTIVATE ARABIC FORM SHAPING NATIONAL DIGIT SHAPES NOMINAL DIGIT SHAPES		
205B	\vdots	 → 1015B: greek acrophonic epidaurean two FOUR DOT MARK used by scribes in the margin as highlighter mark 					
205C	**	this is centered on the line, but extends beyond top and bottom of the line DOTTED CROSS					
2000	•••	• used by scribes in the margin as highlighter					
205D	:	mark TRICOLON					
2000	•	= Epidaurean acrophonic symbol three					
		→ 22EE : vertical ellipsis					
		→ 2AF6 : triple colon operator					
		→ FE19 : presentation form for vertical horizontal ellipsis					
205E	Ė	VERTICAL FOUR DOTS					
		 used in dictionaries to indicate legal but undesirable word break 					
		• glyph extends the whole height of the line → 2E3D i vertical six dots					
Space	<u> </u>						
205F	MM SP	MEDIUM MATHEMATICAL SPACE					
		abbreviated MMSP					
		• four-eighteenths of an em ≈ 0020 sp space					
Form	at ch	aracter					
2060	WJ	WORD JOINER					
		• commonly abbreviated WJ					
		a zero width non-breaking space (only)intended for disambiguation of functions for					
		byte order mark					
		→ FEFF zero width no-break space					
Invisible operators							
2061	f()	FUNCTION APPLICATION					
		 contiguity operator indicating application of a function 					
2062	×	INVISIBLE TIMES					
	,	 contiguity operator indicating multiplication 					
2063	2	INVISIBLE SEPARATOR = invisible comma					
		contiguity operator indicating that adjacent mathematical symbols form a list, e.g. when no visible comma is used between multiple indices					
2064	[+]	INVISIBLE PLUS • contiguity operator indicating addition					
Form	at ch	aracters					
2066	LRI	LEFT-TO-RIGHT ISOLATE					
2067	RLI	RIGHT-TO-LEFT ISOLATE					
2068	FSI	FIRST STRONG ISOLATE					
2069	PDI	POP DIRECTIONAL ISOLATE					
Deprecated							
Use of	these	e characters is strongly discouraged.					

206A sis INHIBIT SYMMETRIC SWAPPING

206C INHIBIT ARABIC FORM SHAPING

206B

ACTIVATE SYMMETRIC SWAPPING