

General Punctuation

Range: 2000–206F

The Unicode Standard, Version 4.0

This file contains an excerpt from the character code tables and list of character names for *The Unicode Standard, Version 4.0*.

Characters in this chart that are new for The Unicode Standard, Version 4.0 are shown in conjunction with any existing characters. For ease of reference, the new characters have been highlighted in the chart grid and in the names list.

This file will not be updated with errata, or when additional characters are assigned to the Unicode Standard. See <http://www.unicode.org/charts> for access to a complete list of the latest character charts.

Disclaimer

These charts are provided as the on-line reference to the character contents of the Unicode Standard, Version 4.0 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 excerpt file, please consult the appropriate sections of The Unicode Standard, Version 4.0 (ISBN 0-321-18578-1), as well as Unicode Standard Annexes #9, #11, #14, #15, #24 and #29, the other Unicode Technical Reports and the Unicode Character Database, which are available on-line.

See <http://www.unicode.org/Public/UNIDATA/UCD.html> and <http://www.unicode.org/unicode/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 are welcome to provide links to these charts.

The fonts and font data used in production of these Code Charts may NOT be extracted or otherwise used in any commercial product without permission or license granted by the typeface owner(s).
























The information in this file may be updated from time to time. The Unicode Consortium is not liable for errors or omissions in this excerpt file or the standard itself. Information on characters added to the Unicode Standard since the publication of Version 4.0 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/unicode/alloc/Pipeline.html>.





Copyright © 1991-2003 Unicode, Inc. All rights reserved.

	200	201	202	203	204	205	206
0	 2000	— 2010	† 2020	‰ 2030	⤿ 2040	↻ 2050	 2060
1	 2001	 2011	‡ 2021	‱ 2031	ℓ 2041	✳ 2051	 2061
2	 2002	— 2012	● 2022	′ 2032	✳✳ 2042	% 2052	 2062
3	 2003	— 2013	▶ 2023	” 2033	■ 2043	~ 2053	 2063
4	 2004	— 2014	• 2024	”” 2034	/ 2044	 2054	
5	 2005	— 2015	•• 2025	、 2035	[2045		
6	 2006	 2016	••• 2026	””” 2036	}] 2046		
7	 2007	== 2017	• 2027	””” 2037	?? 2047	””” 2057	
8	 2008	‘ 2018	 2028	^ 2038	?! 2048		
9	 2009	’ 2019	 2029	< 2039	!? 2049		
A	 200A	, 201A	 202A	> 203A	7 204A		 206A
B	 200B	‘ 201B	 202B	✳ 203B	¶ 204B		 206B
C	 200C	“ 201C	 202C	!! 203C	◐ 204C		 206C
D	 200D	” 201D	 202D	!? 203D	◑ 204D		 206D
E	 200E	” 201E	 202E	— 203E	✳ 204E		 206E
F	 200F	“ 201F	 202F	⤿ 203F	• 204F	 205F	 206F




Spaces

- 2000  EN QUAD
≡ 2002  en space
- 2001  EM QUAD
= mutton quad
≡ 2003  em space
- 2002  EN SPACE
= nut
• half an em
≈ 0020  space
- 2003  EM SPACE
= mutton
• nominally, a space equal to the type size in points
• may scale by the condensation factor of a font
≈ 0020  space
- 2004  THREE-PER-EM SPACE
= thick space
≈ 0020  space
- 2005  FOUR-PER-EM SPACE
= mid space
≈ 0020  space
- 2006  SIX-PER-EM SPACE
• in computer typography sometimes equated to thin space
≈ 0020  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 
- 2008  PUNCTUATION SPACE
• space equal to narrow punctuation of a font
≈ 0020  space
- 2009  THIN SPACE
• a fifth of an em (or sometimes a sixth)
≈ 0020  space
- 200A  HAIR SPACE
• thinner than a thin space
• in traditional typography, the thinnest space available
≈ 0020  space
- 200B  ZERO WIDTH SPACE
= ZWSP
• nominally zero width, but may expand in justification





Formatting characters



- 200C  ZERO WIDTH NON-JOINER
= ZWNJ
- 200D  ZERO WIDTH JOINER
= ZWJ
- 200E  LEFT-TO-RIGHT MARK
= LRM
- 200F  RIGHT-TO-LEFT MARK
= RLM

Dashes

- 2010 - HYPHEN
→ 002D - hyphen-minus
→ 00AD  soft hyphen
- 2011  NON-BREAKING HYPHEN
→ 002D - hyphen-minus
→ 00AD  soft hyphen
≈ `<noBreak>` 2010 -
- 2012 - FIGURE DASH
- 2013 - EN DASH
- 2014 — EM DASH
• may be used in pairs to offset parenthetical text
→ 30FC — katakana-hiragana prolonged sound mark
- 2015 — HORIZONTAL BAR
= QUOTATION DASH
• long dash introducing quoted text

General punctuation



- 2016 || DOUBLE VERTICAL LINE
• used in pairs to indicate norm of a matrix
→ 20E6  combining double vertical stroke overlay
→ 2225 || parallel to
- 2017 = DOUBLE LOW LINE
• this is a spacing character
→ 005F _ low line
→ 0333  combining double low line
≈ 0020  0333 
- 2018 ‘ LEFT SINGLE QUOTATION MARK
= SINGLE TURNED COMMA QUOTATION MARK
• this is the preferred glyph (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
- 201B ˇ SINGLE HIGH-REVERSED-9 QUOTATION MARK
= SINGLE REVERSED COMMA QUOTATION MARK
• glyph variant of 2018 ‘
→ 02BD ˇ modifier letter reversed comma

- 2039 ‹ SINGLE LEFT-POINTING ANGLE QUOTATION MARK
= LEFT POINTING SINGLE GUILLEMET
• usually opening, sometimes closing
→ 003C ‹ less-than sign
→ 2329 ‹ left-pointing angle bracket
→ 3008 ‹ left angle bracket
- 203A › SINGLE RIGHT-POINTING ANGLE QUOTATION MARK
= RIGHT POINTING SINGLE GUILLEMET
• usually closing, sometimes opening
→ 003E › greater-than sign
→ 232A › right-pointing angle bracket
→ 3009 › right angle bracket
- 203B ※ REFERENCE MARK
= Japanese kome
= Urdu paragraph separator
→ 0FBF ※ tibetan ku ru kha bzhi mig can
- 203C !! DOUBLE EXCLAMATION MARK
→ 0021 ! exclamation mark
≈ 0021 ! 0021 !
- 203D ? INTERROBANG
→ 0021 ! exclamation mark
→ 003F ? question mark
- 203E ¯ OVERLINE
= SPACING OVERSCORE
≈ 0020  0305 
- 203F ˘ UNDERTIE
= Greek enotikon
→ 2323 ˘ smile
- 2040 ˆ CHARACTER TIE
= z notation sequence concatenation
→ 2322 ˆ frown
- 2041 ˆ CARET INSERTION POINT
• proofreader's mark: insert here
→ 22CC ˆ right semidirect product
- 2042 ** ASTERISM
- 2043 - HYPHEN BULLET
- 2044 / FRACTION SLASH
= solidus (in typography)
• for composing arbitrary fractions
→ 002F / solidus
→ 2215 / division slash
- 2045 { LEFT SQUARE BRACKET WITH QUILL
- 2046 } RIGHT SQUARE BRACKET WITH QUILL



Double punctuation for vertical text

- 2047 ?? DOUBLE QUESTION MARK
≈ 003F ? 003F ?
- 2048 ?! QUESTION EXCLAMATION MARK
≈ 003F ? 0021 !
- 2049 !? EXCLAMATION QUESTION MARK
≈ 0021 ! 003F ?



General punctuation

- 204A ˆ TIRONIAN SIGN ET
• Irish Gaelic, ...
- 204B ¶ REVERSED PILCROW SIGN
→ 00B6 ¶ pilcrow sign
- 204C ◼ BLACK LEFTWARDS BULLET
- 204D ◻ BLACK RIGHTWARDS BULLET
- 204E * LOW ASTERISK
→ 002A * asterisk
- 204F ; REVERSED SEMICOLON
→ 003B ; semicolon
- 2050 ○ CLOSE UP
• editing mark
- 2051 * TWO ASTERISKS ALIGNED VERTICALLY
- 2052 % COMMERCIAL MINUS SIGN
= abzüglich (German), med avdrag av (Swedish), piska (Swedish, "whip")
• a common glyph variant and fallback representation looks like ./.
• may also be used as a dingbat to indicate correctness
• used in Finno-Ugric Phonetic Alphabet to indicate a related borrowed form with different sound
→ 0025 % percent sign
→ 066A % arabic percent sign
- 2053 ~ SWUNG DASH
- 2054 ˘ INVERTED UNDERTIE
- 2055  <reserved>
- 2056  <reserved>
- 2057 ‴ QUADRUPLE PRIME
≈ 2032 ' 2032 ' 2032 ' 2032 '



Space


- 205F  MEDIUM MATHEMATICAL SPACE
= MMSP
• four-eighths of an em
≈ 0020  space

Formatting character







- 2060  WORD JOINER
= 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  FUNCTION APPLICATION
• contiguity operator indicating application of a function
- 2062  INVISIBLE TIMES
• contiguity operator indicating multiplication

- 2063  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

Deprecated

- 206A  INHIBIT SYMMETRIC SWAPPING
206B  ACTIVATE SYMMETRIC SWAPPING
206C  INHIBIT ARABIC FORM SHAPING
206D  ACTIVATE ARABIC FORM SHAPING
206E  NATIONAL DIGIT SHAPES
206F  NOMINAL DIGIT SHAPES