

General Punctuation

Range: 2000–206F

DRAFT The Unicode Standard, Version 16.0 BETA REVIEW

This file contains an excerpt from the character code tables and list of character names for *DRAFT The Unicode Standard, Version 16.0 BETA REVIEW*

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

See <https://www.unicode.org/charts/> for access to a complete list of the latest character code charts. See <https://www.unicode.org/charts/PDF/Unicode-15.1/> for charts showing only the characters added in Unicode 15.1. See <https://www.unicode.org/Public/15.1.0/charts/> for a complete archived file of character code charts for Unicode 15.1. See <https://www.unicode.org/charts/About.html#Conventions> for conventions used in these code charts, and other general information.

Disclaimer

These charts are provided as the online reference to the character contents of the Unicode Standard, Version 15.1 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 15.1, online at <https://www.unicode.org/versions/Unicode15.1.0/>, as well as Unicode Standard Annexes #9, #11, #14, #15, #24, #29, #31, #34, #38, #41, #42, #44, #45, and #50, the other Unicode Technical Reports and Standards, and the Unicode Character Database, which are available online.

See <https://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.

Copying characters from the character code tables or list of character names is not recommended, because for production reasons the PDF files for the code charts cannot guarantee that the correct character codes will always be copied.

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 <https://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 <https://www.unicode.org/pending/pending.html> and <http://www.unicode.org/alloc/Pipeline.html>.
























See <https://www.unicode.org/charts/About.html> for more information concerning the conventions and symbols used in these code charts.

Copyright © 1991–2023 Unicode, Inc. All rights reserved. See <https://www.unicode.org/copyright.html>







	200	201	202	203	204	205	206
0	 2000	 2010	 2020	 2030	 2040	 2050	 2060
1	 2001	 2011	 2021	 2031	 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	 2064
5	 2005	 2015	 2025	 2035	 2045	 2055	
6	 2006	 2016	 2026	 2036	 2046	 2056	 2066
7	 2007	 2017	 2027	 2037	 2047	 2057	 2067
8	 2008	 2018	 2028	 2038	 2048	 2058	 2068
9	 2009	 2019	 2029	 2039	 2049	 2059	 2069
A	 200A	 201A	 202A	 203A	 204A	 205A	 206A
B	 200B	 201B	 202B	 203B	 204B	 205B	 206B
C	 200C	 201C	 202C	 203C	 204C	 205C	 206C
D	 200D	 201D	 202D	 203D	 204D	 205D	 206D
E	 200E	 201E	 202E	 203E	 204E	 205E	 206E
F	 200F	 201F	 202F	 203F	 204F	 205F	 206F

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





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  space
- 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)
→ 202F  narrow no-break space
≈ 0020  space
- 200A  HAIR SPACE
• thinner than a thin space
• in traditional typography, the thinnest space available
≈ 0020  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  arabic letter mark

Dashes

- 2010 - HYPHEN
→ 002D - hyphen-minus
→ 00AD  soft hyphen
- 2011  NON-BREAKING HYPHEN
≈ <noBreak> 2010 -
- 2012 — FIGURE DASH
- 2013 — EN DASH
- 2014 — EM DASH
• may be used in pairs to offset parenthetical text
→ 2E3A  two-em dash
→ 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
→ 23F8 || double vertical bar
- 2017 = DOUBLE LOW LINE
• this is a spacing character
→ 005F _ low line
→ 0333  combining double low line
≈ 0020  0333 

Quotation marks and apostrophe

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

- 2018 ‘ LEFT SINGLE QUOTATION MARK
= 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
~ 2018 FE00 ‘ non-fullwidth form
~ 2018 FE01 ‘ right-justified fullwidth form
- 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
~ 2019 FE00 ’ non-fullwidth form
~ 2019 FE01 ’ left-justified fullwidth form
- 201A , SINGLE LOW-9 QUOTATION MARK
= low single comma quotation mark
• used as opening single quotation mark in some languages
- 201B ‘ SINGLE HIGH-REVERSED-9 QUOTATION MARK
= single reversed comma quotation mark
• has same semantic as 2018 ‘, but differs in appearance
→ 02BD ‘ modifier letter reversed comma

- 201C “ LEFT DOUBLE QUOTATION MARK
= double turned comma quotation mark
• this is the preferred character (as opposed to 201F “)
→ 0022 " quotation mark
→ 275D “ heavy double turned comma quotation mark ornament
→ 301D ≡ reversed double prime quotation mark
~ 201C FE00 “ non-fullwidth form
~ 201C FE01 “ right-justified fullwidth form
- 201D ” RIGHT DOUBLE QUOTATION MARK
= double comma quotation mark
→ 0022 " quotation mark
→ 2033 " double prime
→ 275E ” heavy double comma quotation mark ornament
→ 301E ≡ double prime quotation mark
~ 201D FE00 ” non-fullwidth form
~ 201D FE01 ” left-justified fullwidth form
- 201E „ DOUBLE LOW-9 QUOTATION MARK
= low double comma quotation mark
• used as opening double quotation mark in some languages
→ 2E42 „ double low-reversed-9 quotation mark
→ 301F ≡ low double prime quotation mark
- 201F “ DOUBLE HIGH-REVERSED-9 QUOTATION MARK
= double reversed comma quotation mark
• has same semantic as 201C “, but differs in appearance

General punctuation

- 2020 † DAGGER
= obelisk, long cross, oblong cross
→ 2E38 † turned dagger
- 2021 ‡ DOUBLE DAGGER
= diesis, double obelisk
→ 2E4B ‡ triple dagger
- 2022 • BULLET
= black small circle
→ 00B7 · middle dot
→ 2024 . one dot leader
→ 2219 • bullet operator
→ 25D8 ◻ inverse bullet
→ 25E6 ◦ white bullet
- 2023 ▶ TRIANGULAR BULLET
→ 220E ■ end of proof
→ 25B8 ▶ black right-pointing small triangle
- 2024 . ONE DOT LEADER
• also used as an Armenian semicolon (mijaket)
→ 00B7 · middle dot
→ 2022 • bullet
→ 2219 • bullet operator
≈ 002E . full stop
- 2025 .. TWO DOT LEADER
≈ 002E . 002E .
- 2026 ... HORIZONTAL ELLIPSIS
= three dot leader
→ 22EE ∴ vertical ellipsis
→ FE19 ∴ presentation form for vertical horizontal ellipsis
≈ 002E . 002E . 002E .

- 2027 · HYPHENATION POINT
• visible symbol used to indicate correct positions for word breaking, as in dictionaries

Separators

- 2028 ⌚ LINE SEPARATOR
• may be used to represent this semantic unambiguously
- 2029 ¶ PARAGRAPH SEPARATOR
• may be used to represent this semantic unambiguously

Format characters

- 202A ⌵ LEFT-TO-RIGHT EMBEDDING
• commonly abbreviated LRE
- 202B ⌵ RIGHT-TO-LEFT EMBEDDING
• commonly abbreviated RLE
- 202C ⌵ POP DIRECTIONAL FORMATTING
• commonly abbreviated PDF
- 202D ⌵ LEFT-TO-RIGHT OVERRIDE
• commonly abbreviated LRO
- 202E ⌵ RIGHT-TO-LEFT OVERRIDE
• commonly abbreviated RLO

Space

- 202F ⌵ NARROW NO-BREAK SPACE
• commonly abbreviated NNBS
• a narrow form of a no-break space, typically the width of a thin space or a mid space
→ 00A0 ⌵ no-break space
→ 2005 ⌵ four-per-em space
→ 2009 ⌵ thin space
≈ <noBreak> 0020 ⌵

General punctuation

- 2030 ‰ PER MILLE SIGN
= permille, per thousand
• used, for example, in measures of blood alcohol content, salinity, etc.
→ 0025 ‰ percent sign
→ 0609 ‰ arabic-indic per mille sign
- 2031 ‰ PER TEN THOUSAND SIGN
= permyriad
• percent of a percent, rarely used
→ 0025 ‰ percent sign
→ 060A ‰ arabic-indic per ten thousand sign
- 2032 ' PRIME
= minutes, feet
→ 0027 ' apostrophe
→ 00B4 ´ acute accent
→ 02B9 ´ modifier letter prime
- 2033 " DOUBLE PRIME
= seconds, inches
→ 0022 " quotation mark
→ 02BA " modifier letter double prime
→ 201D ” right double quotation mark
→ 3003 " ditto mark
→ 301E ≡ double prime quotation mark
≈ 2032 ' 2032 '
- 2034 ''' TRIPLE PRIME
= lines (old measure, 1/12 of an inch)
≈ 2032 ' 2032 ' 2032 '
- 2035 ` REVERSED PRIME
→ 0060 ` grave accent

- 2036 ` REVERSED DOUBLE PRIME
→ 301D ˘ reversed double prime quotation mark
≈ 2035 ` 2035 `
- 2037 ˇ REVERSED TRIPLE PRIME
≈ 2035 ` 2035 ` 2035 `
- 2038 ^ CARET
→ 2303 ^ up arrowhead
→ A788 ˆ modifier letter low circumflex accent

Quotation marks

- 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

General punctuation

- 203B ※ REFERENCE MARK
= Japanese kome
= Urdu paragraph separator
→ 0FBF ※ tibetan ku ru kha bzhi mig can
→ 200AD ✎

Double punctuation for vertical text

- 203C !! DOUBLE EXCLAMATION MARK
→ 0021 ! exclamation mark
≈ 0021 ! 0021 !

General punctuation

- 203D ? INTERROBANG
→ 0021 ! exclamation mark
→ 003F ? question mark
→ 2E18 ˆ inverted interrobang
→ 1F679 ʔ heavy interrobang ornament
- 203E ¯ OVERLINE
= spacing overscore
≈ 0020 ̄ 0305 ̄
- 203F ˘ UNDERTIE
= Greek enotikon
→ 2323 ˘ smile
- 2040 ˉ CHARACTER TIE
= z notation sequence concatenation
→ 2322 ˉ frown
- 2041 ˆ CARET INSERTION POINT
• proofreader's mark: insert here
→ 22CC ˆ right semidirect product
- 2042 ✎ ASTERISM
- 2043 ˆ HYPHEN BULLET
→ 002D - hyphen-minus
- 2044 / FRACTION SLASH
= solidus (in typography)
• for composing arbitrary fractions
→ 002F / solidus
→ 2215 / division slash

Brackets

- 2045 [LEFT SQUARE BRACKET WITH QUILL
→ 2E20 [left vertical bar with quill
→ 2E55 [left square bracket with stroke
- 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, Old English, ...
→ 0026 & ampersand
→ 2E52 ˆ tironian sign capital et
→ 1F670 ˆ script ligature et ornament
- 204B ˆ REVERSED PILCROW SIGN
→ 00B6 ¶ pilcrow sign
→ 2E4D ˆ paragraphus mark
- 204C ◀ BLACK LEFTWARDS BULLET
- 204D ▶ BLACK RIGHTWARDS BULLET
- 204E * LOW ASTERISK
→ 002A * asterisk
→ 0359 ˆ combining asterisk below
- 204F ; REVERSED SEMICOLON
• used occasionally in Sindhi when Sindhi is written in the Arabic script
→ 003B ; semicolon
→ 061B ˆ arabic semicolon
- 2050 ◌ CLOSE UP
• editing mark
→ AB5B ˆ modifier breve with inverted breve
- 2051 ✎ TWO ASTERISKS ALIGNED VERTICALLY
- 2052 ˆ 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 ˆ arabic percent sign
→ 00F7 ÷ division sign
- 2053 ~ SWUNG DASH
→ 007E ~ tilde
- 2054 ˘ INVERTED UNDERTIE
- 2055 * FLOWER PUNCTUATION MARK
= phul, puspika
• used as a punctuation mark with Syloti Nagri, Bengali and other Indic scripts
→ 274B * heavy eight teardrop-spoked propeller asterisk

Archaic punctuation

- 2056 ∴ THREE DOT PUNCTUATION
→ 10FB ∴ georgian paragraph separator

General punctuation

2057  QUADRUPLE PRIME
 ≈ 2032' 2032' 2032' 2032'

Archaic punctuation


See also *historic punctuation with multiple dots in the range 2E2A-2E2D*.

2058  FOUR DOT PUNCTUATION

2059  FIVE DOT PUNCTUATION


= Greek pentonkion
 = quincunx

→ 2684  die face-5

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

→ 1015B  greek acrophonic epidaurean two

205B  FOUR DOT MARK

• used by scribes in the margin as highlighter mark

• this is centered on the line, but extends beyond top and bottom of the line

205C  DOTTED CROSS


• used by scribes in the margin as highlighter mark

205D  TRICOLON

= Epidaurean acrophonic symbol three

→ 22EE  vertical ellipsis

→ 2AF6  triple colon operator

→ FE19  presentation form for vertical horizontal ellipsis

205E  VERTICAL FOUR DOTS

• used in dictionaries to indicate legal but undesirable word break

• glyph extends the whole height of the line


→ 2E3D  vertical six dots

Space

205F  MEDIUM MATHEMATICAL SPACE

• abbreviated MMSP

• four-eighteenths of an em

≈ 0020  space

Format character

2060  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  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


2064  INVISIBLE PLUS

• contiguity operator indicating addition

Format characters

2066  LEFT-TO-RIGHT ISOLATE

• commonly abbreviated LRI

2067  RIGHT-TO-LEFT ISOLATE

• commonly abbreviated RLI

2068  FIRST STRONG ISOLATE

• commonly abbreviated FSI


2069  POP DIRECTIONAL ISOLATE


• commonly abbreviated PDI

Deprecated

Use of these characters is strongly discouraged.

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

Standardized Variation Sequences

2018	‘	LEFT SINGLE QUOTATION MARK
	2018	‘
		non-fullwidth form
	2018 FE00	‘
		right-justified fullwidth form
	2018 FE01	
2019	’	RIGHT SINGLE QUOTATION MARK
	2019	’
		non-fullwidth form
	2019 FE00	’
		left-justified fullwidth form
	2019 FE01	
201C	“	LEFT DOUBLE QUOTATION MARK
	201C	“
		non-fullwidth form
	201C FE00	“
		right-justified fullwidth form
	201C FE01	
201D	”	RIGHT DOUBLE QUOTATION MARK
	201D	”
		non-fullwidth form
	201D FE00	”
		left-justified fullwidth form
	201D FE01	