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 17.0

Characters in this chart that are new for The Unicode Standard, Version 17.0 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 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-17.0/ for charts showing only the characters added in Unicode 17.0. See https://www.unicode.org/Public/17.0.0/charts/ for a complete archived file of character code charts for Unicode 17.0. 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 17.0 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 17.0, online at https://www.unicode.org/versions/Unicode17.0.0/, as well as the Unicode Standard Annexes, the other Unicode Technical Reports and Standards, and the Unicode Character Database, which are available online.

See https://www.unicode.org/ucd/ and https://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.

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

Terms of Use

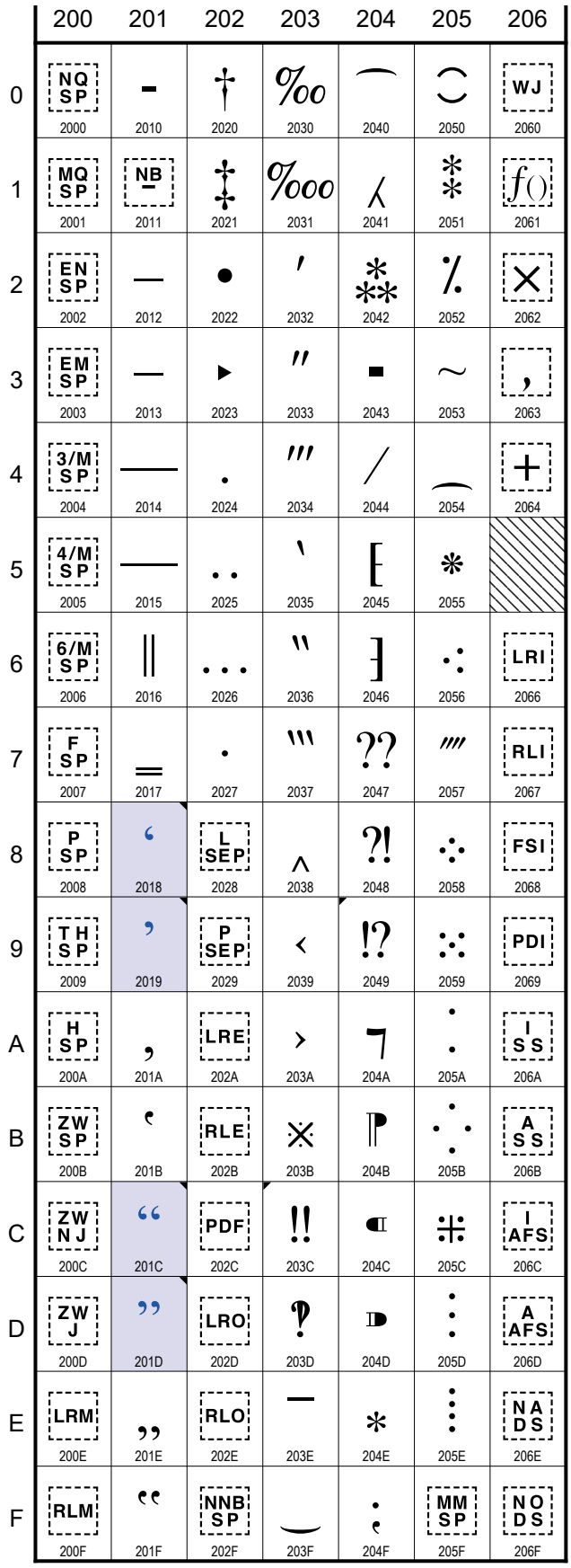
© 1991–2025 Unicode, Inc. This publication is protected by copyright, and permission must be obtained from Unicode, Inc. prior to any reproduction, modification, or other use not permitted by the Terms of Use (https://www.unicode.org/copyright.html). Specifically, you may make copies of this publication and may annotate and translate it solely for personal or internal business purposes and not for public distribution, provided that any such permitted copies and modifications fully reproduce all copyright and other legal notices contained in the original. You may not make copies of or modifications to this publication for public distribution, or incorporate it in whole or in part into any product or publication without the express written permission of Unicode.

The Unicode Consortium specifically grants ISO a license to produce such code charts with their associated character names list to show the repertoire of characters for that standard, as a normatively referenced, integral part of that standard.

Unicode uses most fonts under restricted license from the original font owner. You may not extract, copy, modify, or distribute fonts or font data from any Unicode Products, including this publication, without license from the font owner.

Use of all Unicode Products, including this publication, is governed by the Unicode Terms of Use (https://www.unicode.org/copyright.html). The authors, contributors, and publishers have taken care in the preparation of this publication, but make no express or implied representation or warranty of any kind and assume no responsibility or liability for errors or omissions or for consequential or incidental damages that may arise therefrom. This publication is provided "AS-IS" without charge as a convenience to users.

Unicode and the Unicode Logo are registered trademarks of Unicode, Inc., in the United States and other countries.



The Unicode Standard, Version 17.0, Copyright © 1991-2025 Unicode, Inc. All rights reserved.

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

Spaces

2000 EN QUAD
≡ 2002 EN SP en space

2001 EM QUAD
= mutton quad
≡ 2003 EM SP 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 points
• may scale by the condensation factor of a font
≈ 0020 SP space

2004 THREE-PER-EM SPACE
= thick space
≈ 0020 SP space

2005 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, typically set to the same width as digit zero 0030 0
• this is equivalent to the digit width of fonts with fixed-width digits
≈ <noBreak> 0020 SP

2008 PUNCTUATION SPACE
• space equal to narrow punctuation of a font, typically set to the same width as full stop 002E .
≈ 0020 SP space

2009 THIN SPACE
= narrow space
• this should be much narrower than space 0020 SP ; typically set to 1/5 or 1/6 em
→ 202F NNB SP narrow no-break space
≈ 0020 SP space

200A HAIR SPACE
• thinner than a thin space; typically set to 1/10 to 1/16 em
• in traditional typography, the thinnest space available
≈ 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
≈ <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 ー 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 SP 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 ʻ modifier letter turned comma
→ 275B ❛ heavy single turned comma quotation mark ornament
~ 2018 FE00 ‘ non-fullwidth form
~ 2018 FE01 ‘ right-justified fullwidth form
~ 2018 FE02 ‘ Sibe form

2019 ’ RIGHT SINGLE QUOTATION MARK
= single comma quotation mark
• this is the preferred character to use for apostrophe
→ 0027 ' apostrophe
→ 02BC ʼ modifier letter apostrophe
→ 275C ❜ heavy single comma quotation mark ornament
~ 2019 FE00 ’ non-fullwidth form
~ 2019 FE01 ’ left-justified fullwidth form
~ 2019 FE02 ’ Sibe form

201A ‚ SINGLE LOW-9 QUOTATION MARK
= low single comma quotation mark
• used as opening single quotation mark in some languages

The Unicode Standard, Version 17.0, Copyright © 1991-2025 Unicode, Inc. All rights reserved.

201B ‛ SINGLE HIGH-REVERSED-9 QUOTATION MARK
= single reversed comma quotation mark
• has same semantic as 2018 ‘ , but differs in appearance
→ 02BD ʽ 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
~ 201C FE02 “ Sibe 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
~ 201D FE02 ” Sibe 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 dic·tion·ar·ies

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
= no-break thin space
• commonly abbreviated NNBSP
• a narrow form of a no-break space; should be the same width as thin space 2009
→ 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 ʹ modifier letter prime

2033 ″ DOUBLE PRIME
= seconds, inches
→ 0022 " quotation mark
→ 02BA ʺ modifier letter double prime
→ 201D ” right double quotation mark
→ 3003 〃 ditto mark
→ 301E 〞 double prime quotation mark
≈ 2032 ′ 2032 ′

The Unicode Standard, Version 17.0, Copyright © 1991-2025 Unicode, Inc. All rights reserved.

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 ‷ REVERSED TRIPLE PRIME
≈ 2035 ‵ 2035 ‵ 2035 ‵
2038 ‸ CARET
→ 2303 ⌃ up arrowhead
→ A788 ˄ 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 ⸘ inverted interrobang
→ 1F679 ❓ heavy interrobang ornament
203E ‾ OVERLINE
= spacing overscore
≈ 0020 SP 0305 ◌̅
203F ‿ UNDERTIE
= Greek enotikon
→ 2323 ⌣ smile
2040 ⁀ CHARACTER TIE
= z notation sequence concatenation
→ 2322 ⌢ frown
2041 ⁁ CARET INSERTION POINT
• proofreader's mark: insert here
→ 22CC ⋌ right semidirect product
2042 ⁂ ASTERISM
2043 ⁃ 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 ⁊ TIRONIAN SIGN ET
• Irish Gaelic, Old English, ...
→ 0026 & ampersand
→ 2E52 ⸒ tironian sign capital et
→ 1F670 🙰 script ligature et ornament
204B ⁋ REVERSED PILCROW SIGN
→ 00B6 ¶ pilcrow sign
→ 2E4D ⸍ paragraphus mark
204C ⁌ BLACK LEFTWARDS BULLET
204D ⁍ BLACK RIGHTWARDS BULLET
204E ⁎ LOW ASTERISK
→ 002A * asterisk
→ 0359 ◌͙ combining asterisk below
204F ⁏ REVERSED SEMICOLON
• used occasionally in Sindhi when Sindhi is written in the Arabic script
→ 003B ; semicolon
→ 061B ؛ arabic semicolon
2050 ⁐ CLOSE UP
• editing mark
→ AB5B ꭛ modifier breve with inverted breve
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
→ 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 [MMSP] MEDIUM MATHEMATICAL SPACE
• abbreviated MMSP
• four-eighteenths of an em
≈ 0020 [SP] space

Format character

2060 [WJ] WORD JOINER
• commonly abbreviated WJ
• a zero width non-breaking space (only)
• intended for disambiguation of functions for byte order mark
→ FEFF [ZWNBSP] 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 [,] 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 [LRI] LEFT-TO-RIGHT ISOLATE
• commonly abbreviated LRI
2067 [RLI] RIGHT-TO-LEFT ISOLATE
• commonly abbreviated RLI
2068 [FSI] FIRST STRONG ISOLATE
• commonly abbreviated FSI
2069 [PDI] POP DIRECTIONAL ISOLATE
• commonly abbreviated PDI

Deprecated

Use of these characters is strongly discouraged.

206A [ISS] INHIBIT SYMMETRIC SWAPPING
206B [ASS] ACTIVATE SYMMETRIC SWAPPING
206C [IAFS] INHIBIT ARABIC FORM SHAPING
206D [AAFS] ACTIVATE ARABIC FORM SHAPING
206E [NADS] NATIONAL DIGIT SHAPES
206F [NODS] NOMINAL DIGIT SHAPES

The Unicode Standard, Version 17.0, Copyright © 1991-2025 Unicode, Inc. All rights reserved.

Standardized Variation Sequences

2018	‘	LEFT SINGLE QUOTATION MARK
	2018 ‘	non-fullwidth form
	2018 FE00 ‘	right-justified fullwidth form
	2018 FE01 ‘	Sibe form
	2018 FE02	
2019	’	RIGHT SINGLE QUOTATION MARK
	2019 ’	non-fullwidth form
	2019 FE00 ’	left-justified fullwidth form
	2019 FE01 ’	Sibe form
	2019 FE02	
201C	“	LEFT DOUBLE QUOTATION MARK
	201C “	non-fullwidth form
	201C FE00 “	right-justified fullwidth form
	201C FE01 “	Sibe form
	201C FE02	
201D	”	RIGHT DOUBLE QUOTATION MARK
	201D ”	non-fullwidth form
	201D FE00 ”	left-justified fullwidth form
	201D FE01 ”	Sibe form
	201D FE02	

The Unicode Standard, Version 17.0, Copyright © 1991-2025 Unicode, Inc. All rights reserved.