

## General Punctuation

Range: 2000–206F

### The Unicode Standard, Version 4.1

This file contains an excerpt from the character code tables and list of character names for *The Unicode Standard, Version 4.1*.

Characters in this chart that are new for The Unicode Standard, Version 4.1 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 code charts. See <http://www.unicode.org/Public/4.1.0/charts/> for a complete archived file of character code charts for Unicode 4.1.

### Disclaimer

These charts are provided as the on-line reference to the character contents of the Unicode Standard, Version 4.1 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 excerpt file, please consult the appropriate sections of The Unicode Standard, Version 4.1, at <http://www.unicode.org/versions/Unicode4.1.0/>, including sections unchanged in The Unicode Standard, Version 4.0 (ISBN 0-321-18578-1), as well as Unicode Standard Annexes #9, #11, #14, #15, #24, #29, #31, and #34, the other Unicode Technical Reports and Standards, and the Unicode Character Database, which are available on-line.

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 excerpt file or the standard itself. Information on characters added to the Unicode Standard since the publication of Version 4.1 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 © 1991-2005 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	✳ 2055	
6	 2006	 2016	••• 2026	”” 2036	}] 2046	•• 2056	
7	 2007	== 2017	• 2027	”” 2037	?? 2047	””” 2057	
8	 2008	‘ 2018	 2028	^ 2038	?! 2048	•• 2058	
9	 2009	’ 2019	 2029	< 2039	!? 2049	•• 2059	
A	 200A	, 201A	 202A	> 203A	7 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)  
≈ 0020  space
- 200A  HAIR SPACE  
• thinner than a thin space  
• in traditional typography, the thinnest space available  
≈ 0020  space
- 200B  ZERO WIDTH SPACE  
= ZWSP  
• this character is intended for line break control; it has no width, but its presence between two characters does not prevent increased letter spacing 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

- 201C “ LEFT DOUBLE QUOTATION MARK  
= DOUBLE TURNED COMMA QUOTATION MARK  
• this is the preferred glyph (as opposed to 201F “ )  
→ 0022 " quotation mark  
→ 275D “ heavy double turned comma quotation mark ornament  
→ 301D “ reversed double prime quotation mark
- 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
- 201E „ DOUBLE LOW-9 QUOTATION MARK  
= LOW DOUBLE COMMA QUOTATION MARK  
• used as opening double quotation mark in some languages  
→ 301F „ low double prime quotation mark
- 201F “ DOUBLE HIGH-REVERSED-9 QUOTATION MARK  
= DOUBLE REVERSED COMMA QUOTATION MARK  
• glyph variant of 201C “
- 2020 † DAGGER  
= obelisk, obelus, long cross
- 2021 ‡ DOUBLE DAGGER  
= diesis, double obelisk
- 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

### Formatting characters

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

- 202A  LEFT-TO-RIGHT EMBEDDING  
= LRE
- 202B  RIGHT-TO-LEFT EMBEDDING  
= RLE
- 202C  POP DIRECTIONAL FORMATTING  
= PDF
- 202D  LEFT-TO-RIGHT OVERRIDE  
= LRO
- 202E  RIGHT-TO-LEFT OVERRIDE  
= RLO
- 202F  NARROW NO-BREAK SPACE  
= NNBS  
→ 00A0  no-break 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
- 2031 ‰ PER TEN THOUSAND SIGN  
= permyriad  
• percent of a percent, rarely used  
→ 0025 ‰ percent 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 ‶ REVERSED TRIPLE PRIME  
≈ 2035 ` 2035 ` 2035 `
- 2038 ^ CARET  
→ 2303 ^ up arrowhead
- 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

### Double punctuation for vertical text

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

### General punctuation

- 203D ? INTERROBANG  
 → 0021 ! exclamation mark  
 → 003F ? question mark
- 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
- 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 7 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  
 → 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

### General punctuation

- 2057 ⁂ QUADRUPLE PRIME  
 ≈ 2032 ' 2032 ' 2032 ' 2032 '



### Archaic punctuation

- 2058 ✧ FOUR DOT PUNCTUATION
- 2059 ✨ FIVE DOT PUNCTUATION  
 = Greek pentonkion
- 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




### Space

- 205F M MEDIUM MATHEMATICAL SPACE  
 = MMSP  
 • four-eightieths of an em  
 ≈ 0020 SP 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