

C0 Controls and Basic Latin

Range: 0000–007F

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

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-16.0/> for charts showing only the characters added in Unicode 16.0. See <https://www.unicode.org/Public/16.0.0/charts/> for a complete archived file of character code charts for Unicode 16.0. 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 16.0 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 16.0, online at <https://www.unicode.org/versions/Unicode16.0.0/>, 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

© 1991–2024 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.

	000	001	002	003	004	005	006	007
0	NUL 0000	DLE 0010	SP 0020	0 0030	@ 0040	P 0050	` 0060	p 0070
1	SOH 0001	DC1 0011	! 0021	1 0031	A 0041	Q 0051	a 0061	q 0071
2	STX 0002	DC2 0012	" 0022	2 0032	B 0042	R 0052	b 0062	r 0072
3	ETX 0003	DC3 0013	# 0023	3 0033	C 0043	S 0053	c 0063	s 0073
4	EOT 0004	DC4 0014	\$ 0024	4 0034	D 0044	T 0054	d 0064	t 0074
5	ENQ 0005	NAK 0015	% 0025	5 0035	E 0045	U 0055	e 0065	u 0075
6	ACK 0006	SYN 0016	& 0026	6 0036	F 0046	V 0056	f 0066	v 0076
7	BEL 0007	ETB 0017	' 0027	7 0037	G 0047	W 0057	g 0067	w 0077
8	BS 0008	CAN 0018	(0028	8 0038	H 0048	X 0058	h 0068	x 0078
9	HT 0009	EM 0019	) 0029	9 0039	I 0049	Y 0059	i 0069	y 0079
A	LF 000A	SUB 001A	* 002A	: 003A	J 004A	Z 005A	j 006A	z 007A
B	VT 000B	ESC 001B	+ 002B	; 003B	K 004B	[005B	k 006B	{ 007B
C	FF 000C	FS 001C	, 002C	< 003C	L 004C	\ 005C	l 006C	 007C
D	CR 000D	GS 001D	- 002D	= 003D	M 004D	] 005D	m 006D	} 007D
E	SO 000E	RS 001E	. 002E	> 003E	N 004E	^ 005E	n 006E	~ 007E
F	SI 000F	US 001F	/ 002F	? 003F	O 004F	_ 005F	o 006F	DEL 007F

C0 controls

Alias names are those for ISO/IEC 6429:1992. Commonly used alternative aliases are also shown.

0000		<control> = NULL	001A		<control> = SUBSTITUTE → FFFD  replacement character
0001		<control> = START OF HEADING	001B		<control> = ESCAPE
0002		<control> = START OF TEXT	001C		<control> = INFORMATION SEPARATOR FOUR = file separator (FS)
0003		<control> = END OF TEXT	001D		<control> = INFORMATION SEPARATOR THREE = group separator (GS)
0004		<control> = END OF TRANSMISSION	001E		<control> = INFORMATION SEPARATOR TWO = record separator (RS)
0005		<control> = ENQUIRY	001F		<control> = INFORMATION SEPARATOR ONE = unit separator (US)
0006		<control> = ACKNOWLEDGE	ASCII punctuation and symbols		
0007		<control> = BELL	<i>Based on ISO/IEC 646.</i>		
0008		<control> = BACKSPACE	0020		SPACE • sometimes considered a control code • other space characters: 2000  —200A 
0009		<control> = CHARACTER TABULATION = horizontal tabulation (HT) = tab			→ 00A0  no-break space → 200B  zero width space → 202F  narrow no-break space → 2060  word joiner → 2420  symbol for space → 2422  blank symbol → 2423  open box → 3000  ideographic space → FEFF  zero width no-break space
000A		<control> = LINE FEED (LF) = new line (NL) = end of line (EOL)	0021		EXCLAMATION MARK = factorial = bang → 00A1  inverted exclamation mark → 01C3  latin letter retroflex click → 203C  double exclamation mark → 203D  interrobang → 26A0  warning sign → 2757  heavy exclamation mark symbol → 2762  heavy exclamation mark ornament → 2E53  medieval exclamation mark → A71D  modifier letter raised exclamation mark
000B		<control> = LINE TABULATION = vertical tabulation (VT)	0022		QUOTATION MARK = double quote • neutral (vertical), used as opening or closing quotation mark • preferred characters in English for paired quotation marks are 201C “ & 201D ” • 05F4 ” is preferred for gershayim when writing Hebrew → 02BA ” modifier letter double prime → 02DD ” double acute accent → 02EE ” modifier letter double apostrophe → 030B ⸢ combining double acute accent → 030E ⸤ combining double vertical line above → 05F4 ” hebrew punctuation gershayim → 201C “ left double quotation mark → 201D ” right double quotation mark → 2033 ” double prime → 3003 ” ditto mark
000C		<control> = FORM FEED (FF)			
000D		<control> = CARRIAGE RETURN (CR)			
000E		<control> = SHIFT OUT • known as LOCKING-SHIFT ONE in 8-bit environments			
000F		<control> = SHIFT IN • known as LOCKING-SHIFT ZERO in 8-bit environments			
0010		<control> = DATA LINK ESCAPE			
0011		<control> = DEVICE CONTROL ONE			
0012		<control> = DEVICE CONTROL TWO			
0013		<control> = DEVICE CONTROL THREE			
0014		<control> = DEVICE CONTROL FOUR			
0015		<control> = NEGATIVE ACKNOWLEDGE			
0016		<control> = SYNCHRONOUS IDLE			
0017		<control> = END OF TRANSMISSION BLOCK			
0018		<control> = CANCEL			
0019		<control> = END OF MEDIUM			

- 0023 # **NUMBER SIGN**
 = pound sign (weight)
 = hashtag, hash
 = crosshatch, octothorpe
 • for denoting musical sharp 266F ♯ is preferred
 → 2114 ℔ lb bar symbol
 → 2116 № numero sign
 → 2317 ⇕ viewdata square
 → 266F ♯ music sharp sign
 → 29E3 # equals sign and slanted parallel
- 0024 \$ **DOLLAR SIGN**
 = milréis, escudo
 • used for many peso currencies in Latin America and elsewhere
 • glyph may have one or two vertical bars
 • other currency symbol characters start at 20A0 ₡
 → 00A2 ¢ cent sign
 → 00A4 Ɑ currency sign
 → 20B1 ₱ peso sign
 → 1F4B2 \$ heavy dollar sign
- 0025 % **PERCENT SIGN**
 → 066A ‰ arabic percent sign
 → 2030 ‰ per mille sign
 → 2031 ‰ per ten thousand sign
 → 2052 ‰ commercial minus sign
- 0026 & **AMPERSAND**
 = and
 • originally derived from a ligature of ‘e’ and ‘t’
 → 204A 𐌆 tironian sign et
 → 214B 𐌆 turned ampersand
 → 1F674 & heavy ampersand ornament
- 0027 ' **APOSTROPHE**
 = apostrophe-quote (1.0)
 = single quote
 = APL quote
 • neutral (vertical) glyph with mixed usage
 • 2019 ’ is preferred for apostrophe
 • preferred characters in English for paired quotation marks are 2018 ‘ & 2019 ’
 • 05F3 ׳ is preferred for geresh when writing Hebrew
 → 02B9 ’ modifier letter prime
 → 02BC ’ modifier letter apostrophe
 → 02C8 ' modifier letter vertical line
 → 0301 ◌ combining acute accent
 → 030D ◌ combining vertical line above
 → 05F3 ׳ hebrew punctuation geresh
 → 2018 ‘ left single quotation mark
 → 2019 ’ right single quotation mark
 → 2032 ’ prime
 → A78C ' latin small letter saltillo
- 0028 (**LEFT PARENTHESIS**
 = opening parenthesis (1.0)
- 0029) **RIGHT PARENTHESIS**
 = closing parenthesis (1.0)
 • see discussion on semantics of paired bracketing characters

- 002A * **ASTERISK**
 = star
 • can have five or six spokes
 → 066D ★ arabic five pointed star
 → 203B ✕ reference mark
 → 2042 ✱ asterism
 → 204E * low asterisk
 → 2051 ✱ two asterisks aligned vertically
 → 20F0 ⚡ combining asterisk above
 → 2217 * asterisk operator
 → 26B9 ✱ sextile
 → 2731 ✱ heavy asterisk
 → A673 ☆ slavonic asterisk
 → 1F7B6 ✱ medium six spoked asterisk

ASCII math operator

- 002B + **PLUS SIGN**
 → 02D6 + modifier letter plus sign
 → 2212 − minus sign
 → 2795 ⛶ heavy plus sign
 → FB29 ⛶ hebrew letter alternative plus sign
 → 1F7A2 ⛶ light greek cross

ASCII punctuation

- 002C , **COMMA**
 = the use as decimal or thousands separator is locale dependent
 → 060C ٫ arabic comma
 → 066B ٫ arabic decimal separator
 → 201A ٫ single low-9 quotation mark
 → 2E12 ٫ hypodiastole
 → 2E41 ٫ reversed comma
 → 2E4C ٫ medieval comma
 → 3001 ٫ ideographic comma
- 002D - **HYPHEN-MINUS**
 = hyphen, dash
 = minus sign
 • used generically for hyphen, minus sign or en dash, all of which have dedicated alternatives
 → 00AD ͂ soft hyphen
 → 02D7 - modifier letter minus sign
 → 2010 - hyphen
 → 2011 ͜ non-breaking hyphen
 → 2012 − figure dash
 → 2013 — en dash
 → 2027 · hyphenation point
 → 2043 ▪ hyphen bullet
 → 2212 − minus sign
 → 10191 — roman uncia sign
- 002E . **FULL STOP**
 = period, dot, decimal point
 • the use as decimal point is locale dependent
 • may be rendered as a raised decimal point in old style numbers
 → 00B7 · middle dot
 → 06D4 ٫ arabic full stop
 → 2024 . one dot leader
 → 2026 … horizontal ellipsis
 → 2E33 · raised dot
 → 2E3C ٫ stenographic full stop
 → 3002 ٫ ideographic full stop

002F / **SOLIDUS**
 = slash, forward slash, virgule
 → 0338 ∕ combining long solidus overlay
 → 2044 ⁄ fraction slash
 → 2215 ⁄ division slash
 → 27CB ⁄ mathematical rising diagonal
 → 29F8 ⁄ big solidus
 → 2E4A ⁄ dotted solidus

ASCII digits

0030 0 **DIGIT ZERO**
 ~ 0030 FE00 0 short diagonal stroke form

0031 1 **DIGIT ONE**

0032 2 **DIGIT TWO**
 → 01BB 2 latin letter two with stroke
 → 218A 2 turned digit two

0033 3 **DIGIT THREE**
 → 218B 3 turned digit three
 → A7AB 3 latin capital letter reversed open e

0034 4 **DIGIT FOUR**
 → A72C 4 latin capital letter cuatrillo

0035 5 **DIGIT FIVE**
 → 01BC 5 latin capital letter tone five

0036 6 **DIGIT SIX**

0037 7 **DIGIT SEVEN**

0038 8 **DIGIT EIGHT**

0039 9 **DIGIT NINE**

ASCII punctuation

003A : **COLON**
 • also used to denote division or scale; for that mathematical use 2236 : is preferred
 • in Finnish and Swedish, also used as intra-word punctuation (abbreviation mark)
 → 02D0 : modifier letter triangular colon
 → 02F8 : modifier letter raised colon
 → 0589 : armenian full stop
 → 05C3 : hebrew punctuation sof pasuq
 → 1361 : ethiopic wordspace
 → 1365 : ethiopic colon
 → 205A : two dot punctuation
 → 205D : tricolon
 → 2236 : ratio
 → A789 : modifier letter colon
 → FE30 : presentation form for vertical two dot leader

003B ; **SEMICOLON**
 • this, and not 037E ;, is the preferred character for 'Greek question mark'
 → 037E ; greek question mark
 → 061B ؛ arabic semicolon
 → 204F ； reversed semicolon
 → 2E35 ； turned semicolon

ASCII mathematical operators

Other mathematical operators start at 2200.

003C < **LESS-THAN SIGN**
 • paired with 003E > for ASCII-based angle bracket markup conventions
 → 02C2 < modifier letter left arrowhead
 → 2039 < single left-pointing angle quotation mark
 → 2329 < left-pointing angle bracket
 → 27E8 < mathematical left angle bracket
 → 3008 < left angle bracket

003D = **EQUALS SIGN**
 • other related characters: 2241 ≈ —2263 ≡
 → 1400 = canadian syllabics hyphen
 → 2248 ≈ almost equal to
 → 2260 ≠ not equal to
 → 2261 ≡ identical to
 → 2E40 = double hyphen
 → 30A0 = katakana-hiragana double hyphen
 → A78A = modifier letter short equals sign
 → FE66 = small equals sign
 → 10190 = roman sextans sign
 → 1F7F0 = heavy equals sign

003E > **GREATER-THAN SIGN**
 → 02C3 > modifier letter right arrowhead
 → 203A > single right-pointing angle quotation mark
 → 232A > right-pointing angle bracket
 → 27E9 > mathematical right angle bracket
 → 3009 > right angle bracket

ASCII punctuation

003F ? **QUESTION MARK**
 → 00BF ¿ inverted question mark
 → 037E ; greek question mark
 → 061F ؟ arabic question mark
 → 203D ？ interrobang
 → 2047 ?? double question mark
 → 2753 ？ black question mark ornament
 → 2BD1 ？ uncertainty sign
 → 2E2E ？ reversed question mark
 → 2E54 ？ medieval question mark
 → FFFD ？ replacement character

0040 @ **COMMERCIAL AT**
 = at sign
 = arroba (old Spanish unit of weight)
 → 24D0 @ circled latin small letter a

Uppercase Latin alphabet

0041 A **LATIN CAPITAL LETTER A**

0042 B **LATIN CAPITAL LETTER B**
 → 212C ℬ script capital b

0043 C **LATIN CAPITAL LETTER C**
 → 03F9 Ϟ greek capital lunate sigma symbol
 → 2102 Ⅱ double-struck capital c
 → 2103 °C degree celsius
 → 212D Ɔ black-letter capital c
 → 216D Ⅰ roman numeral one hundred

0044 D **LATIN CAPITAL LETTER D**
 → 216E Ⅴ roman numeral five hundred

0045 E **LATIN CAPITAL LETTER E**
 → 0190 Ǝ latin capital letter open e
 → 2107 Ǝ euler constant
 → 2130 Ǝ script capital e

0046 F **LATIN CAPITAL LETTER F**
 → 2109 °F degree fahrenheit
 → 2131 ƒ script capital f
 → 2132 Ǝ turned capital f

0047 G **LATIN CAPITAL LETTER G**

0048 H **LATIN CAPITAL LETTER H**
 → 210B Ƨ script capital h
 → 210C Ƨ black-letter capital h
 → 210D Ⅱ double-struck capital h

0049	I	LATIN CAPITAL LETTER I • Turkish and Azerbaijani use 0131 ı for lowercase → 0130 İ latin capital letter i with dot above → 2110 ℑ script capital i → 2111 Ⓘ black-letter capital i → 2160 Ⓘ roman numeral one
004A	J	LATIN CAPITAL LETTER J
004B	K	LATIN CAPITAL LETTER K → 212A Ⓚ kelvin sign
004C	L	LATIN CAPITAL LETTER L → 2112 ℒ script capital l
004D	M	LATIN CAPITAL LETTER M → 2133 ℳ script capital m → 216F Ⓜ roman numeral one thousand
004E	N	LATIN CAPITAL LETTER N → 2115 ℕ double-struck capital n
004F	O	LATIN CAPITAL LETTER O
0050	P	LATIN CAPITAL LETTER P → 2118 ℙ script capital p → 2119 Ⓟ double-struck capital p
0051	Q	LATIN CAPITAL LETTER Q → 051A Ⓖ cyrillic capital letter qa → 211A Ⓠ double-struck capital q → 213A Ⓠ rotated capital q
0052	R	LATIN CAPITAL LETTER R → 211B ℞ script capital r → 211C Ⓡ black-letter capital r → 211D Ⓡ double-struck capital r
0053	S	LATIN CAPITAL LETTER S
0054	T	LATIN CAPITAL LETTER T
0055	U	LATIN CAPITAL LETTER U
0056	V	LATIN CAPITAL LETTER V → 2164 Ⓥ roman numeral five
0057	W	LATIN CAPITAL LETTER W → 051C Ⓖ cyrillic capital letter we
0058	X	LATIN CAPITAL LETTER X → 2169 Ⓧ roman numeral ten
0059	Y	LATIN CAPITAL LETTER Y
005A	Z	LATIN CAPITAL LETTER Z → 2124 Ⓩ double-struck capital z → 2128 Ⓩ black-letter capital z

ASCII punctuation and symbols

005B	[	LEFT SQUARE BRACKET = opening square bracket (1.0) • other bracket characters: 27E6 ⌈ −27EB ⌋ , 2983 ⌊ −2998 ⌋ , 3008 ⌈ −301B ⌋
005C	\	REVERSE SOLIDUS = backslash → 20E5 ↵ combining reverse solidus overlay → 2216 ∖ set minus → 27CD ↘ mathematical falling diagonal → 29F9 ↵ big reverse solidus
005D	]	RIGHT SQUARE BRACKET = closing square bracket (1.0)

005E	^	CIRCUMFLEX ACCENT • this is a spacing character • used in ASCII-based representation of exponentiation → 02C4 ^ modifier letter up arrowhead → 02C6 ^ modifier letter circumflex accent → 0302 ˆ combining circumflex accent → 2038 ⤿ caret → 2227 ∧ logical and → 2303 ⤴ up arrowhead
005F	_	LOW LINE = spacing underscore (1.0) • this is a spacing character • joins with adjacent instances of itself → 02CD _ modifier letter low macron → 0331 ˘ combining macron below → 0332 ˙ combining low line → 2017 ≡ double low line → FE2D ˘ combining conjoining macron below
0060	`	GRAVE ACCENT = backtick, backquote • this is a spacing character • in some technical contexts functions as opening quote, paired with apostrophe → 00B4 ´ acute accent → 02CB ` modifier letter grave accent → 0300 ˘ combining grave accent → 1FEF ˘ greek varia → 2018 ‘ left single quotation mark → 2035 ′ reversed prime

Lowercase Latin alphabet

0061	a	LATIN SMALL LETTER A → 0251 ɑ latin small letter alpha
0062	b	LATIN SMALL LETTER B
0063	c	LATIN SMALL LETTER C → 03F2 Ϸ greek lunate sigma symbol → 1D04 ċ latin letter small capital c
0064	d	LATIN SMALL LETTER D → 0501 Ⓓ cyrillic small letter komi de
0065	e	LATIN SMALL LETTER E → 212E ℰ estimated symbol → 212F ℵ script small e → AB32 ⓔ latin small letter blackletter e
0066	f	LATIN SMALL LETTER F → 0192 ƒ latin small letter f with hook → AB35 ⓕ latin small letter lenis f
0067	g	LATIN SMALL LETTER G → 0261 ɡ latin small letter script g → 210A ḡ script small g
0068	h	LATIN SMALL LETTER H → 04BB Ⓗ cyrillic small letter shha → 210E ℏ planck constant
0069	i	LATIN SMALL LETTER I • Turkish and Azerbaijani use 0130 İ for uppercase → 0131 ı latin small letter dotless i → 2071 ˆ superscript latin small letter i → 1D6A4 ℹ mathematical italic small dotless i
006A	j	LATIN SMALL LETTER J → 0237 ĵ latin small letter dotless j → 03F3 ϳ greek letter yot → 1D6A5 ℹ mathematical italic small dotless j
006B	k	LATIN SMALL LETTER K

006C	l	LATIN SMALL LETTER L
		→ 2113 ℓ script small l
		→ 1D4C1 ℓ mathematical script small l
006D	m	LATIN SMALL LETTER M
006E	n	LATIN SMALL LETTER N
		→ 207F ⁱ superscript latin small letter n
006F	o	LATIN SMALL LETTER O
		→ 1D0F o latin letter small capital o
		→ 2134 ø script small o
		→ AB3D o latin small letter blackletter o
0070	p	LATIN SMALL LETTER P
0071	q	LATIN SMALL LETTER Q
		→ 051B q cyrillic small letter qa
0072	r	LATIN SMALL LETTER R
0073	s	LATIN SMALL LETTER S
		→ 017F ſ latin small letter long s
		→ A731 s latin letter small capital s
0074	t	LATIN SMALL LETTER T
0075	u	LATIN SMALL LETTER U
0076	v	LATIN SMALL LETTER V
0077	w	LATIN SMALL LETTER W
		→ 051D w cyrillic small letter we
		→ 1D21 w latin letter small capital w
0078	x	LATIN SMALL LETTER X
		→ 00D7 × multiplication sign
0079	y	LATIN SMALL LETTER Y
007A	z	LATIN SMALL LETTER Z
		→ 01B6 z latin small letter z with stroke
		→ 1D22 z latin letter small capital z

ASCII punctuation and symbols

007B	{	LEFT CURLY BRACKET
		= opening curly bracket (1.0)
		= left brace
007C		VERTICAL LINE
		= vertical bar, pipe
		• used in pairs to indicate absolute value
		• also used as an unpaired separator or as a fence
		→ 00A6 † broken bar
		→ 01C0 latin letter dental click
		→ 05C0 hebrew punctuation paseq
		→ 0964 devanagari danda
		→ 20D2 ‡ combining long vertical line overlay
		→ 2223 divides
		→ 2758 light vertical bar
007D	}	RIGHT CURLY BRACKET
		= closing curly bracket (1.0)
		= right brace
007E	~	TILDE
		• this is a spacing character
		→ 02DC ~ small tilde
		→ 0303 ¯ combining tilde
		→ 0334 ¯ combining tilde overlay
		→ 2053 ~ swung dash
		→ 223C ~ tilde operator
		→ 2E1B ˜ tilde with ring above
		→ 2E2F ' vertical tilde
		→ 301C ˘ wave dash
		→ FF5E ˜ fullwidth tilde

Control character

007F		<control>
		= DELETE