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 14.0

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-14.0/ for charts showing only the characters added in Unicode 14.0.
See https://www.unicode.org/Public/14.0.0/charts/ for a complete archived file of character code charts for Unicode 14.0.

Disclaimer

These charts are provided as the online reference to the character contents of the Unicode Standard, Version 14.0 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 14.0, online at https://www.unicode.org/versions/Unicode14.0.0/, 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 https://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 https://www.unicode.org/alloc/Pipeline.html.

Copyright © 1991-2021 Unicode, Inc. All rights reserved.

	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

The Unicode Standard 14.0, Copyright © 1991-2021 Unicode, Inc. All rights reserved.

C0 controls

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

0000 NUL <control>
= NULL

0001 SOH <control>
= START OF HEADING

0002 STX <control>
= START OF TEXT

0003 ETX <control>
= END OF TEXT

0004 EOT <control>
= END OF TRANSMISSION

0005 ENQ <control>
= ENQUIRY

0006 ACK <control>
= ACKNOWLEDGE

0007 BEL <control>
= BELL

0008 BS <control>
= BACKSPACE

0009 HT <control>
= CHARACTER TABULATION
= horizontal tabulation (HT)
= tab

000A LF <control>
= LINE FEED (LF)
= new line (NL)
= end of line (EOL)

000B VT <control>
= LINE TABULATION
= vertical tabulation (VT)

000C FF <control>
= FORM FEED (FF)

000D CR <control>
= CARRIAGE RETURN (CR)

000E SO <control>
= SHIFT OUT
• known as LOCKING-SHIFT ONE in 8-bit environments

000F SI <control>
= SHIFT IN
• known as LOCKING-SHIFT ZERO in 8-bit environments

0010 DLE <control>
= DATA LINK ESCAPE

0011 DC1 <control>
= DEVICE CONTROL ONE

0012 DC2 <control>
= DEVICE CONTROL TWO

0013 DC3 <control>
= DEVICE CONTROL THREE

0014 DC4 <control>
= DEVICE CONTROL FOUR

0015 NAK <control>
= NEGATIVE ACKNOWLEDGE

0016 SYN <control>
= SYNCHRONOUS IDLE

0017 ETB <control>
= END OF TRANSMISSION BLOCK

0018 CAN <control>
= CANCEL

0019 EM <control>
= END OF MEDIUM

001A SUB <control>
= SUBSTITUTE
→ FFFD 🞜 replacement character

001B ESC <control>
= ESCAPE

001C FS <control>
= INFORMATION SEPARATOR FOUR
= file separator (FS)

001D GS <control>
= INFORMATION SEPARATOR THREE
= group separator (GS)

001E RS <control>
= INFORMATION SEPARATOR TWO
= record separator (RS)

001F US <control>
= INFORMATION SEPARATOR ONE
= unit separator (US)

ASCII punctuation and symbols

Based on ISO/IEC 646.

0020 SP SPACE
• sometimes considered a control code
• other space characters: 2000 NQSP –200A HSP
→ 00A0 NBSP no-break space
→ 200B ZWSP zero width space
→ 202F NNBSP narrow no-break space
→ 2060 WJ word joiner
→ 2420 ␠ symbol for space
→ 2422 ␢ blank symbol
→ 2423 ␣ open box
→ 3000 IDSP ideographic space
→ FEFF ZWNBSP zero width no-break space

0021 ! EXCLAMATION MARK
= factorial
= bang
→ 00A1 ¡ inverted exclamation mark
→ 01C3 ǃ 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 ꜝ modifier letter raised exclamation mark

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 ˮ 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

The Unicode Standard 14.0, Copyright © 1991-2021 Unicode, Inc. All rights reserved.

0023 # NUMBER SIGN
= pound sign (weight)
= hashtag, hash
= crosshatch, octothorpe
• for denoting musical sharp 266F ♯ is preferred
→ 2114 ℔ l b 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 ¤ 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 ⁊ tironian sign et
→ 214B ⅋ 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 ʹ modifier letter prime
→ 02BC ʼ modifier letter apostrophe
→ 02C8 ˈ 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 ꞌ 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
→ 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
= decimal separator
→ 060C ، arabic comma
→ 066B ٫ arabic decimal separator
→ 201A ‚ single low-9 quotation mark
→ 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 SHY soft hyphen
→ 02D7 ˗ modifier letter minus sign
→ 2010 ‐ hyphen
→ 2011 NB 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
→ 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

The Unicode Standard 14.0, Copyright © 1991-2021 Unicode, Inc. All rights reserved.

ASCII digits

0030 0 DIGIT ZERO
~ 0030 FE00 0 short diagonal stroke form
0031 1 DIGIT ONE
0032 2 DIGIT TWO
0033 3 DIGIT THREE
0034 4 DIGIT FOUR
0035 5 DIGIT 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 punctation (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
→ 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

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
→ 2102 ℂ double-struck capital c
→ 2103 ℃ degree celsius
→ 212D ℭ black-letter capital c
→ 216D C roman numeral one hundred
0044 D LATIN CAPITAL LETTER D
→ 216E D roman numeral five hundred
0045 E LATIN CAPITAL LETTER E
→ 2107 ℇ euler constant
→ 2130 ℰ script capital e
0046 F LATIN CAPITAL LETTER F
→ 2109 ℉ degree fahrenheit
→ 2131 ℱ script capital f
→ 2132 Ⅎ turned capital f
0047 G LATIN CAPITAL LETTER G
0048 H LATIN CAPITAL LETTER H
→ 210B ℋ script capital h
→ 210C ℌ black-letter capital h
→ 210D ℍ 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 ℑ black-letter capital i
→ 2160 I roman numeral one
004A J LATIN CAPITAL LETTER J
004B K LATIN CAPITAL LETTER K
→ 212A K 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 M roman numeral one thousand
004E N LATIN CAPITAL LETTER N
→ 2115 ℕ double-struck capital n
004F O LATIN CAPITAL LETTER O

The Unicode Standard 14.0, Copyright © 1991-2021 Unicode, Inc. All rights reserved.

0050 P LATIN CAPITAL LETTER P
→ 2118 ℘ script capital p
→ 2119 ℙ double-struck capital p
0051 Q LATIN CAPITAL LETTER Q
→ 211A ℚ double-struck capital q
→ 213A ⅋ rotated capital q
0052 R LATIN CAPITAL LETTER R
→ 211B ℛ script capital r
→ 211C ℜ black-letter capital r
→ 211D ℝ 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 Ⅴ roman numeral five
0057 W LATIN CAPITAL LETTER W
0058 X LATIN CAPITAL LETTER X
→ 2169 Ⅹ roman numeral ten
0059 Y LATIN CAPITAL LETTER Y
005A Z LATIN CAPITAL LETTER Z
→ 2124 ℤ double-struck capital z
→ 2128 ℨ 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 ˆ modifier letter circumflex accent
→ 0302 ◌̂ combining circumflex accent
→ 2038 ‸ caret
→ 2303 ⌃ up arrowhead
005F _ LOW LINE
= spacing underscore (1.0)
• this is a spacing character
• joins with adjacent instances of itself
→ 02CD ˍ modifier letter low macron
→ 0331 ◌̱ combining macron below
→ 0332 ◌̲ combining low line
→ 2017 ‗ double low line
→ FE2D ◌﹭ combining conjoining macron below
0060 ` GRAVE ACCENT
• this is a spacing character
→ 02CB ˋ modifier letter grave accent
→ 0300 ◌̀ combining grave accent
→ 2035 ‵ reversed prime

Lowercase Latin alphabet

0061 a LATIN SMALL LETTER A
0062 b LATIN SMALL LETTER B
0063 c LATIN SMALL LETTER C
0064 d LATIN SMALL LETTER D
0065 e LATIN SMALL LETTER E
→ 212E ℮ estimated symbol
→ 212F ℯ script small e
→ AB32 ꬲ latin small letter blackletter e
0066 f LATIN SMALL LETTER F
0067 g LATIN SMALL LETTER G
→ 0261 ɡ latin small letter script g
→ 210A ℊ script small g
0068 h LATIN SMALL LETTER H
→ 04BB һ cyrillic small letter shha
→ 210E ℎ planck constant
0069 i LATIN SMALL LETTER I
• Turkish and Azerbaijani use 0130 İ for uppercase
→ 0131 ı latin small letter dotless i
→ 2071 ⁱ superscript latin small letter i
→ 1D6A4 𝚤 mathematical italic small dotless i
006A j LATIN SMALL LETTER J
→ 0237 ȷ latin small letter dotless j
→ 1D6A5 𝚥 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 ⁿ superscript latin small letter n
006F o LATIN SMALL LETTER O
→ 2134 ℴ script small o
→ AB3D ꬽ latin small letter blackletter o
0070 p LATIN SMALL LETTER P
0071 q LATIN SMALL LETTER Q
0072 r LATIN SMALL LETTER R
0073 s LATIN SMALL LETTER S
→ 017F ſ latin small letter long 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
0078 x LATIN SMALL LETTER X
→ 00D7 × multiplication sign
0079 y LATIN SMALL LETTER Y
007A z LATIN SMALL LETTER Z
→ 01B6 ƶ latin small letter z with stroke

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 ǀ 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

The Unicode Standard 14.0, Copyright © 1991-2021 Unicode, Inc. All rights reserved.

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 ⸯ vertical tilde
- → 301C 〜 wave dash
- → FF5E ～ fullwidth tilde

Control character

007F DEL <control>
= DELETE

The Unicode Standard 14.0, Copyright © 1991-2021 Unicode, Inc. All rights reserved.