

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 9.0

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

See <http://www.unicode.org/charts/> for access to a complete list of the latest character code charts.

See <http://www.unicode.org/charts/PDF/Unicode-9.0/> for charts showing only the characters added in Unicode 9.0.

See <http://www.unicode.org/Public/9.0.0/charts/> for a complete archived file of character code charts for Unicode 9.0.

Disclaimer

These charts are provided as the online reference to the character contents of the Unicode Standard, Version 9.0 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 9.0, online at <http://www.unicode.org/versions/Unicode9.0.0/>, as well as Unicode Standard Annexes #9, #11, #14, #15, #24, #29, #31, #34, #38, #41, #42, #44, and #45, the other Unicode Technical Reports and Standards, and the Unicode Character Database, which are available online.

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

Copyright © 1991-2016 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	*	:	J 004A	Z 005A	j 006A	z 007A
B	[VT] 000B	[ESC] 001B	+	;	K 004B	[005B	k 006B	{ 007B
C	[FF] 000C	[FS] 001C	,	< 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	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	DCC	<control>	= DEVICE CONTROL ONE
0012	DCT	<control>	= DEVICE CONTROL TWO
0013	DCT	<control>	= DEVICE CONTROL THREE
0014	DCT	<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	<ul style="list-style-type: none"> • sometimes considered a control code • other space characters: 2000 NO SP → 200A H SP
		→ 00A0 NB SP no-break space	
		→ 200B ZW SP zero width space	
		→ 2060 WJ word joiner	
		→ 3000 ID SP ideographic space	
		→ FEFF ZWNJ zero width no-break space	
0021	!	EXCLAMATION MARK	<ul style="list-style-type: none"> = factorial = bang
		→ 00A1 ! inverted exclamation mark	
		→ 01C3 ! latin letter retroflex click	
		→ 203C !! double exclamation mark	
		→ 203D † interrobang	
		→ 2762 ♪ heavy exclamation mark ornament	
0022	"	QUOTATION MARK	<ul style="list-style-type: none"> • 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	
		→ 030B Ŋ combining double acute accent	
		→ 030E Ÿ combining double vertical line above	
		→ 05F4 ” hebrew punctuation gershayim	
		→ 2033 ” double prime	
		→ 3003 ” ditto mark	
0023	#	NUMBER SIGN	<ul style="list-style-type: none"> = pound sign, hash, crosshatch, octothorpe
		→ 2114 ₩ l b bar symbol	
		→ 2317 # viewdata square	
		→ 266F # music sharp sign	
		~ 0023 FEOE text style	
		~ 0023 FEOF emoji style	
0024	\$	DOLLAR SIGN	<ul style="list-style-type: none"> = 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 ¢
		→ 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	
		→ 204A ⁊ tironian sign et	
		→ 214B ⁊ turned ampersand	
		→ 1F674 ⁊ heavy ampersand ornament	
0027	'	APOSTROPHE	
		= apostrophe-quote (1.0)	
		= 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	
		→ 05F3՚ hebrew punctuation geresh	
		→ 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 (on phone keypads)	
		→ 066D * arabic five pointed star	
		→ 204E * low asterisk	
		→ 2217 * asterisk operator	
		→ 26B9 * sextile	
		→ 2731 *	heavy asterisk
		~ 002A FE0E text style	
		~ 002A FE0F emoji style	
002B	+	PLUS SIGN	
		→ 2795 +	heavy plus sign
002C	,	COMMA	
		= decimal separator	
		→ 060C , arabic comma	
		→ 201A , single low-9 quotation mark	
		→ 2E41 , reversed comma	
		→ 3001 、 ideographic comma	
002D	-	HYPHEN-MINUS	
		= hyphen or minus sign	
		• used for either hyphen or minus sign	
		→ 2010 - hyphen	
		→ 2011 [NB] non-breaking hyphen	
		→ 2012 – figure dash	
		→ 2013 – en dash	
		→ 2043 - hyphen bullet	
		→ 2212 – minus sign	
		→ 10101 – roman uncia sign	
002E	.	FULL STOP	
		= period, dot, decimal point	
		• may be rendered as a raised decimal point in old style numbers	
		→ 06D4 - arabic full stop	
		→ 2E3C . stenographic full stop	
		→ 3002 。 ideographic full stop	
002F	/	SOLIDUS	
		= slash, virgule	
		→ 01C0 latin letter dental click	
		→ 0338 ⸂ combining long solidus overlay	
		→ 2044 / fraction slash	
		→ 2215 / division slash	
		ASCII digits	
0030	0	DIGIT ZERO	
		~ 0030 FE00 ⓪ short diagonal stroke form	
		~ 0030 FE0E text style	
		~ 0030 FE0F emoji style	
0031	1	DIGIT ONE	
		~ 0031 FE0E text style	
		~ 0031 FE0F emoji style	
0032	2	DIGIT TWO	
		~ 0032 FE0E text style	
		~ 0032 FE0F emoji style	
0033	3	DIGIT THREE	
		~ 0033 FE0E text style	
		~ 0033 FE0F emoji style	
0034	4	DIGIT FOUR	
		~ 0034 FE0E text style	
		~ 0034 FE0F emoji style	
0035	5	DIGIT FIVE	
		~ 0035 FE0E text style	
		~ 0035 FE0F emoji style	
0036	6	DIGIT SIX	
		~ 0036 FE0E text style	
		~ 0036 FE0F emoji style	
0037	7	DIGIT SEVEN	
		~ 0037 FE0E text style	
		~ 0037 FE0F emoji style	
0038	8	DIGIT EIGHT	
		~ 0038 FE0E text style	
		~ 0038 FE0F emoji style	
0039	9	DIGIT NINE	
		~ 0039 FE0E text style	
		~ 0039 FE0F emoji style	
		ASCII punctuation and symbols	
003A	:	COLON	
		• also used to denote division or scale; for that mathematical use 2236 : is preferred	
		→ 0589 : armenian full stop	
		→ 05C3 : hebrew punctuation sof pasuq	
		→ 2236 : ratio	
		→ A789 : modifier letter colon	
003B	;	SEMICOLON	
		• this, and not 037E ; , is the preferred character for 'Greek question mark'	

003C <	LESS-THAN SIGN	004E N	LATIN CAPITAL LETTER N
	→ 2039 < single left-pointing angle quotation mark		→ 2115 N double-struck capital n
	→ 2329 < left-pointing angle bracket	004F O	LATIN CAPITAL LETTER O
	→ 27E8 < mathematical left angle bracket	0050 P	LATIN CAPITAL LETTER P
	→ 3008 < left angle bracket	0051 Q	LATIN CAPITAL LETTER Q
003D =	EQUALS SIGN	0052 R	LATIN CAPITAL LETTER R
	• other related characters: 2241 ~ 2263 ≡		→ 2119 P double-struck capital p
	→ 2260 ≠ not equal to		→ 211A Q double-struck capital q
	→ 2261 ≡ identical to		→ 211B R script capital r
	→ A78A = modifier letter short equals sign		→ 211C R black-letter capital r
	→ 10190 = roman sextans sign		→ 211D R double-struck capital r
003E >	GREATER-THAN SIGN	0053 S	LATIN CAPITAL LETTER S
	→ 203A > single right-pointing angle quotation mark	0054 T	LATIN CAPITAL LETTER T
	→ 232A > right-pointing angle bracket	0055 U	LATIN CAPITAL LETTER U
	→ 27E9 } mathematical right angle bracket	0056 V	LATIN CAPITAL LETTER V
	→ 3009 } right angle bracket	0057 W	LATIN CAPITAL LETTER W
003F ?	QUESTION MARK	0058 X	LATIN CAPITAL LETTER X
	→ 00BF ؟ inverted question mark	0059 Y	LATIN CAPITAL LETTER Y
	→ 037E ؟ greek question mark	005A Z	LATIN CAPITAL LETTER Z
	→ 061F ؟ arabic question mark		→ 2124 Z double-struck capital z
	→ 203D ؟ interrobang		→ 2128 ؟ black-letter capital z
0040 @	COMMERCIAL AT		
	= at sign		
Uppercase Latin alphabet			
0041 A	LATIN CAPITAL LETTER A	005B [LEFT SQUARE BRACKET
0042 B	LATIN CAPITAL LETTER B		= opening square bracket (1.0)
	→ 212C B script capital b		• other bracket characters: 27E6 [2983 { 2998 }, 3008 < 301B]
0043 C	LATIN CAPITAL LETTER C	005C \	REVERSE SOLIDUS
	→ 2102 C double-struck capital c		= backslash
	→ 212D C black-letter capital c		→ 20E5 \ combining reverse solidus overlay
0044 D	LATIN CAPITAL LETTER D		→ 2216 \ set minus
0045 E	LATIN CAPITAL LETTER E	005D]	RIGHT SQUARE BRACKET
	→ 2107 E euler constant		= closing square bracket (1.0)
	→ 2130 E script capital e	005E ^	CIRCUMFLEX ACCENT
0046 F	LATIN CAPITAL LETTER F		• this is a spacing character
	→ 2131 F script capital f		→ 02C4 ^ modifier letter up arrowhead
	→ 2132 H turned capital f		→ 02C6 ^ modifier letter circumflex accent
0047 G	LATIN CAPITAL LETTER G		→ 0302 ˆ combining circumflex accent
0048 H	LATIN CAPITAL LETTER H		→ 2038 ^ caret
	→ 210B H script capital h		→ 2303 ^ up arrowhead
	→ 210C H black-letter capital h	005F _	LOW LINE
	→ 210D H double-struck capital h		= spacing underscore (1.0)
0049 I	LATIN CAPITAL LETTER I		• this is a spacing character
	• Turkish and Azerbaijani use 0131 İ for lowercase		→ 02CD _ modifier letter low macron
	→ 0130 İ latin capital letter i with dot above		→ 0331 ጀ combining macron below
	→ 0406 İ cyrillic capital letter byelorussian-ukrainian i		→ 0332 ጀ combining low line
	→ 04C0 İ cyrillic letter palochka		→ 2017 = double low line
	→ 2110 J script capital i	0060 `	GRAVE ACCENT
	→ 2111 ጀ black-letter capital i		• this is a spacing character
	→ 2160 I roman numeral one		→ 02CB ` modifier letter grave accent
004A J	LATIN CAPITAL LETTER J		→ 0300 ጀ combining grave accent
004B K	LATIN CAPITAL LETTER K		→ 2035 ` reversed prime
	→ 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		
Lowercase Latin alphabet			
0061 a	LATIN SMALL LETTER A	0061 a	LATIN SMALL LETTER A
0062 b	LATIN SMALL LETTER B	0062 b	LATIN SMALL LETTER B
0063 c	LATIN SMALL LETTER C	0063 c	LATIN SMALL LETTER C
0064 d	LATIN SMALL LETTER D	0064 d	LATIN SMALL LETTER D
0065 e	LATIN SMALL LETTER E	0065 e	LATIN SMALL LETTER E
	→ 212E E estimated symbol		→ 212F e script small e
0066 f	LATIN SMALL LETTER F	0066 f	LATIN SMALL LETTER F

0067	g	LATIN SMALL LETTER G
		→ 0261  latin small letter script g
		→ 210A  script small g
0068	h	LATIN SMALL LETTER H
		→ 04BB  cyrillic small letter shha
		→ 210E  planck constant
0069	i	LATIN SMALL LETTER I
		• Turkish and Azerbaijani use 0130  for uppercase
		→ 0131  latin small letter dotless i
		→ 1D6A4  mathematical italic small dotless i
006A	j	LATIN SMALL LETTER J
		→ 0237  latin small letter dotless j
		→ 1D6A5  mathematical italic small dotless j
006B	k	LATIN SMALL LETTER K
006C	l	LATIN SMALL LETTER L
		→ 2113  script small l
		→ 1D4C1  mathematical script small l
006D	m	LATIN SMALL LETTER M
006E	n	LATIN SMALL LETTER N
		→ 207F  superscript latin small letter n
006F	o	LATIN SMALL LETTER O
		→ 2134  script small 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
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
0079	y	LATIN SMALL LETTER Y
007A	z	LATIN SMALL LETTER Z
		→ 01B6  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
		• used in pairs to indicate absolute value
		→ 01C0 latin letter dental click
		→ 05C0 hebrew punctuation paseq
		→ 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
		→ 2053 ~ swung dash
		→ 223C ~ tilde operator
		→ FF5E ~ fullwidth tilde

Control character

007F	 <control>
	= DELETE