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.
<table>
<thead>
<tr>
<th>Unicode Code</th>
<th>Character</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000 0010 0020 0030 0040 0050 0060 0070</td>
<td>NUL</td>
<td>Control Code 0</td>
</tr>
<tr>
<td>0011 0021 0031 0041 0051 0061 0071</td>
<td>DC1</td>
<td>Control Code 1</td>
</tr>
<tr>
<td>0022 0022 0032 0042 0052 0062 0072</td>
<td>DC2</td>
<td>Control Code 2</td>
</tr>
<tr>
<td>0033 0033 0043 0053 0063 0073</td>
<td>DC3</td>
<td>Control Code 3</td>
</tr>
<tr>
<td>0044 0044 0054 0064 0074</td>
<td>DC4</td>
<td>End of Transmission (EOT)</td>
</tr>
<tr>
<td>0055 0055 0065 0075</td>
<td>NAK</td>
<td>Negative Acknowledge (ENQ)</td>
</tr>
<tr>
<td>0066 0066 0076</td>
<td>SYN</td>
<td>Synchronize (ACK)</td>
</tr>
<tr>
<td>0077 0077 0087</td>
<td>SUB</td>
<td>Start of Text (BEL)</td>
</tr>
<tr>
<td>0088 0088 0098</td>
<td>CAN</td>
<td>Cancel (ESC)</td>
</tr>
<tr>
<td>0099 0099 0109</td>
<td>EM</td>
<td>End of Medium (HT)</td>
</tr>
<tr>
<td>0110 0110 0120</td>
<td>SUB</td>
<td>Start of Text (BEL)</td>
</tr>
<tr>
<td>0130 0130 0140</td>
<td>ESC</td>
<td>End of Transmission (EOT)</td>
</tr>
<tr>
<td>0150 0150 0160</td>
<td>DEL</td>
<td>Delete (DEL)</td>
</tr>
<tr>
<td>0170 0170 0180</td>
<td>DEL</td>
<td>Delete (DEL)</td>
</tr>
</tbody>
</table>
C0 controls
Alias names are those for ISO/IEC 6429:1992. Commonly used alternative aliases are also shown.

0000 .assertIsNotNone
= NULL
0001 .assertIsNotNone
= START OF HEADING
0002 .assertIsNotNone
= START OF TEXT
0003 .assertIsNotNone
= END OF TEXT
0004 .assertIsNotNone
= END OF TRANSMISSION
0005 .assertIsNotNone
= ENQUIRY
0006 .assertIsNotNone
= ACKNOWLEDGE
0007 .assertIsNotNone
= BELL
0008 .assertIsNotNone
= BACKSPACE
0009 .assertIsNotNone
= CHARACTER TABULATION
= horizontal tabulation (HT)
= tab
000A .assertIsNotNone
= LINE FEED (LF)
= new line (NL)
= end of line (EOL)
000B .assertIsNotNone
= LINE TABULATION
= vertical tabulation (VT)
000C .assertIsNotNone
= FORM FEED (FF)
000D .assertIsNotNone
= CARRIAGE RETURN (CR)
000E .assertIsNotNone
= SHIFT OUT
= known as LOCKING-SHIFT ONE in 8-bit environments
000F .assertIsNotNone
= SHIFT IN
= known as LOCKING-SHIFT ZERO in 8-bit environments
0010 .assertIsNotNone
= DATA LINK ESCAPE
0011 .assertIsNotNone
= DEVICE CONTROL ONE
0012 .assertIsNotNone
= DEVICE CONTROL TWO
0013 .assertIsNotNone
= DEVICE CONTROL THREE
0014 .assertIsNotNone
= DEVICE CONTROL FOUR
0015 .assertIsNotNone
= NEGATIVE ACKNOWLEDGE
0016 .assertIsNotNone
= SYNCHRONOUS IDLE
0017 .assertIsNotNone
= END OF TRANSMISSION BLOCK
0018 .assertIsNotNone
= CANCEL
0019 .assertIsNotNone
= END OF MEDIUM

001A .assertIsNotNone
= SUBSTITUTE
= FFFD \ replacement character
001B .assertIsNotNone
= ESCAPE
001C .assertIsNotNone
= INFORMATION SEPARATOR FOUR
= file separator (FS)
001D .assertIsNotNone
= INFORMATION SEPARATOR THREE
= group separator (GS)
001E .assertIsNotNone
= INFORMATION SEPARATOR TWO
= record separator (RS)
001F .assertIsNotNone
= INFORMATION SEPARATOR ONE
= unit separator (US)

ASCII punctuation and symbols
Based on ISO/IEC 646.

0020 .assertIsNotNone
= SPACE
= sometimes considered a control code
= other space characters: 2000 \–200A \→ 00A0 \no-break space
= 200B \zw space
= 202F \narrow no-break space
= 2060 \word joiner
= 2420 \symbol for space
= 2422 \blank symbol
= 2423 \open box
= 3000 \ideographic space
= FEFF \no-break space

0021 .assertIsNotNone
= 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

0022 .assertIsNotNone
= DOUBLE 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
= 028A ” 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

The Unicode Standard 14.0, Copyright © 1991-2021 Unicode, Inc. All rights reserved.
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
    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
→ 2239 : left pointing angle bracket
→ 27E8 : mathematical left angle bracket
→ 3008 ◄ left angle bracket

003B ;  SEMICOLON
  • this, and not 037E ; , is 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
→ 27E9 ◄ mathematical left angle bracket
→ 3008 ◄ left angle bracket

003D =  EQUALS SIGN
  • other related characters: 2241 ← −2263 ↔
    1400  < canadians hyphen
→ 2248 ≃ almost equal to
→ 2260 ≠ not equal to
→ 2261 ≍ identical to
→ 2E40  ≜ double hyphen
→ 30A0 = katakana-hiragana double hyphen
→ 279A ≈ 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
→ 008F  ? 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
→ FFDD ؟ replacement character

0040 @  COMMERCIAL AT
= at sign

Uppercase Latin alphabet

0041 A  LATIN CAPITAL LETTER A
0042 B  LATIN CAPITAL LETTER B
→ 212C ☰ script capital b
0043 C  LATIN CAPITAL LETTER C
→ 2102 ☲ double-struck capital c
→ 2103 ℃ degree celsius
→ 212D ☳ black-letter capital c
→ 216D ☴ roman numeral one hundred
0044 D  LATIN CAPITAL LETTER D
→ 2106 ☵ roman numeral five hundred
0045 E  LATIN CAPITAL LETTER E
→ 2107 ℅ euler constant
→ 2130 ℀ script capital e
0046 F  LATIN CAPITAL LETTER F
→ 2109 ℈ degree fahrenheit
→ 2131 ℁ script capital f
→ 2132 ℂ turned capital f
0047 G  LATIN CAPITAL LETTER G
0048 H  LATIN CAPITAL LETTER H
→ 2108 ℃ script capital h
→ 210C ☴ black-letter capital h
→ 210D ☴ double-struck capital h
0049 I  LATIN CAPITAL LETTER I
  • Turkish and Azerbaijani use 0131 İ for lowercase
→ 0130 İ latin capital letter i with dot above
→ 2110 ℁ script capital i
→ 2111 İ black-letter capital i
→ 2160 ℂ 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 ℌ script capital l
004D M  LATIN CAPITAL LETTER M
→ 2133 ℌ script capital m
→ 216F ℌ roman numeral one thousand
004E N  LATIN CAPITAL LETTER N
→ 2115 N double-struck capital n
  004F O  LATIN CAPITAL LETTER O
<table>
<thead>
<tr>
<th>Code-point</th>
<th>Description</th>
<th>Code-point</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>C0050</td>
<td><strong>LATIN CAPITAL LETTER P</strong></td>
<td>C0064</td>
<td><strong>LATIN SMALL LETTER D</strong></td>
</tr>
<tr>
<td>0050 P</td>
<td>script capital p</td>
<td>0064 d</td>
<td></td>
</tr>
<tr>
<td>0051 Q</td>
<td>double-struck capital q</td>
<td>0065 e</td>
<td><strong>LATIN SMALL LETTER E</strong></td>
</tr>
<tr>
<td>0052 R</td>
<td>black-letter capital r</td>
<td>0066 f</td>
<td><strong>LATIN SMALL LETTER F</strong></td>
</tr>
<tr>
<td>0053 S</td>
<td>roman numeral five</td>
<td>0067 g</td>
<td><strong>LATIN SMALL LETTER G</strong></td>
</tr>
<tr>
<td>0054 T</td>
<td>roman numeral ten</td>
<td>0068 h</td>
<td><strong>LATIN SMALL LETTER H</strong></td>
</tr>
<tr>
<td>0055 U</td>
<td><strong>LATIN SMALL LETTER A</strong></td>
<td>0069 i</td>
<td><strong>LATIN SMALL LETTER I</strong></td>
</tr>
<tr>
<td>0056 V</td>
<td>roman numeral five</td>
<td>006A j</td>
<td><strong>LATIN SMALL LETTER J</strong></td>
</tr>
<tr>
<td>0057 W</td>
<td><strong>LATIN SMALL LETTER W</strong></td>
<td>006B k</td>
<td><strong>LATIN SMALL LETTER K</strong></td>
</tr>
<tr>
<td>0058 X</td>
<td><strong>LATIN SMALL LETTER Y</strong></td>
<td>006C l</td>
<td><strong>LATIN SMALL LETTER L</strong></td>
</tr>
<tr>
<td>0059 Y</td>
<td><strong>LATIN SMALL LETTER Z</strong></td>
<td>006D m</td>
<td><strong>LATIN SMALL LETTER M</strong></td>
</tr>
<tr>
<td>005A Z</td>
<td><strong>LATIN SMALL LETTER A</strong></td>
<td>006E n</td>
<td><strong>LATIN SMALL LETTER N</strong></td>
</tr>
<tr>
<td>005B [</td>
<td>opening square bracket (1.0)</td>
<td>006F o</td>
<td><strong>LATIN SMALL LETTER O</strong></td>
</tr>
<tr>
<td>005C \</td>
<td>reverse solidus</td>
<td>0070 p</td>
<td><strong>LATIN SMALL LETTER P</strong></td>
</tr>
<tr>
<td>005D ]</td>
<td>closing square bracket (1.0)</td>
<td>0071 q</td>
<td><strong>LATIN SMALL LETTER Q</strong></td>
</tr>
<tr>
<td>005E ^</td>
<td>circumflex accent</td>
<td>0072 r</td>
<td><strong>LATIN SMALL LETTER R</strong></td>
</tr>
<tr>
<td>005F _</td>
<td>low line</td>
<td>0073 s</td>
<td><strong>LATIN SMALL LETTER S</strong></td>
</tr>
<tr>
<td>0060 `</td>
<td>grave accent</td>
<td>0074 t</td>
<td><strong>LATIN SMALL LETTER T</strong></td>
</tr>
<tr>
<td>0061 a</td>
<td><strong>LATIN SMALL LETTER A</strong></td>
<td>0075 u</td>
<td><strong>LATIN SMALL LETTER U</strong></td>
</tr>
<tr>
<td>0062 b</td>
<td><strong>LATIN SMALL LETTER B</strong></td>
<td>0076 v</td>
<td><strong>LATIN SMALL LETTER V</strong></td>
</tr>
<tr>
<td>0063 c</td>
<td><strong>LATIN SMALL LETTER C</strong></td>
<td>0077 w</td>
<td><strong>LATIN SMALL LETTER W</strong></td>
</tr>
</tbody>
</table>

**ASCII punctuation and symbols**

- **LEFT SQUARE BRACKET**
  - Opening square bracket (1.0)
  - Other bracket characters: `27E6 ℮ –27EB ℯ`, `2983 ℮ –2998 ℯ`, `3008 ℮ –301B ℯ`

- **REVERSE SOLIDUS**
  - Backslash
  - 20ES \ combining reverse solidus overlay
  - 2216 \ set minus
  - 27CD \ mathematical falling diagonal
  - 29F9 \ big reverse solidus

- **RIGHT SQUARE BRACKET**
  - Closing square bracket (1.0)

- **CIRCUMFLEX ACCENT**
  - This is a spacing character
  - Used in ASCII-based representation of exponentiation

- **LOW LINE**
  - Spacing underscore (1.0)
  - This is a spacing character
  - Joins with adjacent instances of itself

- **GRAVE ACCENT**
  - This is a spacing character
  - 02CB ` modifier letter grave accent
  - 0300 ` combining grave accent
  - 2035 ± reversed prime

**Lowercase Latin alphabet**

- **LATIN SMALL LETTER A**
- **LATIN SMALL LETTER B**
- **LATIN SMALL LETTER C**
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 DEL <control>
     = DELETE