From: <http://www.sensi.org/~alec/man/man_b/p_charset.html>

On: 2018/07/30

See also:

**Portable Character Set**

Conforming implementations support one or more coded character sets. Each supported locale includes the *portable character set* specified in the following table.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Symbolic Name** | **Glyph** | **Symbolic Name** | **Glyph** | **Symbolic Name** | **Glyph** |
|   |   |   |   | <circumflex> | ^ |
| <NUL> |   | <colon> | : | <circumflex-accent> | ^ |
| <alert> |   | <semicolon> | ; | <underscore> | \_ |
| <backspace> |   | <less-than-sign> | < | <underline> | \_ |
| <tab> |   | <equals-sign> | = | <low-line> | \_ |
| <newline> |   | <greater-than-sign> | > | <grave-accent> | ` |
| <vertical-tab> |   | <question-mark> | ? | <a> | a |
| <form-feed> |   | <commercial-at> | @ | <b> | b |
| <carriage-return> |   | <A> | A | <c> | c |
| <space> |   | <B> | B | <d> | d |
| <exclamation-mark> | ! | <C> | C | <e> | e |
| <quotation-mark> | " | <D> | D | <f> | f |
| <number-sign> | # | <E> | E | <g> | g |
| <dollar-sign> | $ | <F> | F | <h> | h |
| <percent-sign> | % | <G> | G | <i> | i |
| <ampersand> | & | <H> | H | <j> | j |
| <apostrophe> | ' | <I> | I | <k> | k |
| <left-parenthesis> | ( | <J> | J | <l> | l |
| <right-parenthesis> | ) | <K> | K | <m> | m |
| <asterisk> | \* | <L> | L | <n> | n |
| <plus-sign> | + | <M> | M | <o> | o |
| <comma> | , | <N> | N | <p> | p |
| <hyphen> | - | <O> | O | <q> | q |
| <hyphen-minus> | - | <P> | P | <r> | r |
| <period> | . | <Q> | Q | <s> | s |
| <full-stop> | . | <R> | R | <t> | t |
| <slash> | / | <S> | S | <u> | u |
| <solidus> | / | <T> | T | <v> | v |
| <zero> | 0 | <U> | U | <w> | w |
| <one> | 1 | <V> | V | <x> | x |
| <two> | 2 | <W> | W | <y> | y |
| <three> | 3 | <X> | X | <z> | z |
| <four> | 4 | <Y> | Y | <left-brace> | { |
| <five> | 5 | <Z> | Z | <left-curly-bracket> | { |
| <six> | 6 | <left-square-bracket> | [ | <vertical-line> | | |
| <seven> | 7 | <backslash> | \ | <right-brace> | } |
| <eight> | 8 | <reverse-solidus> | \ | <right-curly-bracket> | } |
| <nine> | 9 | <right-square-bracket> | ] | <tilde> | ~ |

**Table: Portable Character Set**

Some of the encodings associated with the symbolic names in this table may be the same as characters in the portable character set table.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| <ACK> | <DC2> | <ENQ> | <FS> | <IS4> | <SOH> |
| <BEL> | <DC3> | <EOT> | <GS> | <LF> | <STX> |
| <BS> | <DC4> | <ESC> | <HT> | <NAK> | <SUB> |
| <CAN> | <DEL> | <ETB> | <IS1> | <RS> | <SYN> |
| <CR> | <DLE> | <ETX> | <IS2> | <SI> | <US> |
| <DC1> | <EM> | <FF> | <IS3> | <SO> | <VT> |

**Table: Control Character Set**

Portable Character Set defines the characters in the portable character set and the corresponding symbolic character names used to identify each character in a [Character Set Description File](http://www.sensi.org/~alec/man/man_b/charset.html). The table contains more than one symbolic character name for characters whose traditional name differs from the chosen name.

This specification set places only the following requirements on the encoded values of the characters in the portable character set:

* If the encoded values associated with each member of the portable character set are not invariant across all locales supported by the implementation, the results achieved by an application accessing those locales are unspecified.
* The encoded values associated with the digits 0 to 9 will be such that the value of each character after 0 will be one greater than the value of the previous character.
* A null character, NUL, which has all bits set to zero, will be in the set of characters.
* The encoded values associated with the members of the portable character set are each represented in a single [8 bit] byte. Moreover, if the value is stored in an object of C-language type **char**, it is guaranteed to be positive (except the NUL, which is always zero).

**Portable Character Set Description File:**

|  |
| --- |
| <code\_set\_name> ANSI\_X3.4-1968<comment\_char> %<escape\_char> /% version: 1.0% repertoiremap: mnemonic,ds% source: ECMA registry% alias ISO-IR-6% alias ANSI\_X3.4-1986% alias ISO\_646.IRV:1991% alias ASCII% alias ISO646-US% alias US-ASCII% alias US% alias IBM367% alias CP367CHARMAP<NU> /x00 <U0000> NULL (NUL)<SH> /x01 <U0001> START OF HEADING (SOH)<SX> /x02 <U0002> START OF TEXT (STX)<EX> /x03 <U0003> END OF TEXT (ETX)<ET> /x04 <U0004> END OF TRANSMISSION (EOT)<EQ> /x05 <U0005> ENQUIRY (ENQ)<AK> /x06 <U0006> ACKNOWLEDGE (ACK)<BL> /x07 <U0007> BELL (BEL)<BS> /x08 <U0008> BACKSPACE (BS)<HT> /x09 <U0009> CHARACTER TABULATION (HT)<LF> /x0A <U000A> LINE FEED (LF)<VT> /x0B <U000B> LINE TABULATION (VT)<FF> /x0C <U000C> FORM FEED (FF)<CR> /x0D <U000D> CARRIAGE RETURN (CR)<SO> /x0E <U000E> SHIFT OUT (SO)<SI> /x0F <U000F> SHIFT IN (SI)<DL> /x10 <U0010> DATALINK ESCAPE (DLE)<D1> /x11 <U0011> DEVICE CONTROL ONE (DC1)<D2> /x12 <U0012> DEVICE CONTROL TWO (DC2)<D3> /x13 <U0013> DEVICE CONTROL THREE (DC3)<D4> /x14 <U0014> DEVICE CONTROL FOUR (DC4)<NK> /x15 <U0015> NEGATIVE ACKNOWLEDGE (NAK)<SY> /x16 <U0016> SYNCHRONOUS IDLE (SYN)<EB> /x17 <U0017> END OF TRANSMISSION BLOCK (ETB)<CN> /x18 <U0018> CANCEL (CAN)<EM> /x19 <U0019> END OF MEDIUM (EM)<SB> /x1A <U001A> SUBSTITUTE (SUB)<EC> /x1B <U001B> ESCAPE (ESC)<FS> /x1C <U001C> FILE SEPARATOR (IS4)<GS> /x1D <U001D> GROUP SEPARATOR (IS3)<RS> /x1E <U001E> RECORD SEPARATOR (IS2)<US> /x1F <U001F> UNIT SEPARATOR (IS1)<SP> /x20 <U0020> SPACE<!> /x21 <U0021> EXCLAMATION MARK<"> /x22 <U0022> QUOTATION MARK<Nb> /x23 <U0023> NUMBER SIGN<DO> /x24 <U0024> DOLLAR SIGN<%> /x25 <U0025> PERCENT SIGN<&> /x26 <U0026> AMPERSAND<'> /x27 <U0027> APOSTROPHE<(> /x28 <U0028> LEFT PARENTHESIS<)> /x29 <U0029> RIGHT PARENTHESIS<\*> /x2A <U002A> ASTERISK<+> /x2B <U002B> PLUS SIGN<,> /x2C <U002C> COMMA<-> /x2D <U002D> HYPHEN-MINUS<.> /x2E <U002E> FULL STOP<//> /x2F <U002F> SOLIDUS<0> /x30 <U0030> DIGIT ZERO<1> /x31 <U0031> DIGIT ONE<2> /x32 <U0032> DIGIT TWO<3> /x33 <U0033> DIGIT THREE<4> /x34 <U0034> DIGIT FOUR<5> /x35 <U0035> DIGIT FIVE<6> /x36 <U0036> DIGIT SIX<7> /x37 <U0037> DIGIT SEVEN<8> /x38 <U0038> DIGIT EIGHT<9> /x39 <U0039> DIGIT NINE<:> /x3A <U003A> COLON<;> /x3B <U003B> SEMICOLON<<> /x3C <U003C> LESS-THAN SIGN<=> /x3D <U003D> EQUALS SIGN</>> /x3E <U003E> GREATER-THAN SIGN<?> /x3F <U003F> QUESTION MARK<At> /x40 <U0040> COMMERCIAL AT<A> /x41 <U0041> LATIN CAPITAL LETTER A<B> /x42 <U0042> LATIN CAPITAL LETTER B<C> /x43 <U0043> LATIN CAPITAL LETTER C<D> /x44 <U0044> LATIN CAPITAL LETTER D<E> /x45 <U0045> LATIN CAPITAL LETTER E<F> /x46 <U0046> LATIN CAPITAL LETTER F<G> /x47 <U0047> LATIN CAPITAL LETTER G<H> /x48 <U0048> LATIN CAPITAL LETTER H<I> /x49 <U0049> LATIN CAPITAL LETTER I<J> /x4A <U004A> LATIN CAPITAL LETTER J<K> /x4B <U004B> LATIN CAPITAL LETTER K<L> /x4C <U004C> LATIN CAPITAL LETTER L<M> /x4D <U004D> LATIN CAPITAL LETTER M<N> /x4E <U004E> LATIN CAPITAL LETTER N<O> /x4F <U004F> LATIN CAPITAL LETTER O<P> /x50 <U0050> LATIN CAPITAL LETTER P<Q> /x51 <U0051> LATIN CAPITAL LETTER Q<R> /x52 <U0052> LATIN CAPITAL LETTER R<S> /x53 <U0053> LATIN CAPITAL LETTER S<T> /x54 <U0054> LATIN CAPITAL LETTER T<U> /x55 <U0055> LATIN CAPITAL LETTER U<V> /x56 <U0056> LATIN CAPITAL LETTER V<W> /x57 <U0057> LATIN CAPITAL LETTER W<X> /x58 <U0058> LATIN CAPITAL LETTER X<Y> /x59 <U0059> LATIN CAPITAL LETTER Y<Z> /x5A <U005A> LATIN CAPITAL LETTER Z<<(> /x5B <U005B> LEFT SQUARE BRACKET<////> /x5C <U005C> REVERSE SOLIDUS<)/>> /x5D <U005D> RIGHT SQUARE BRACKET<'/>> /x5E <U005E> CIRCUMFLEX ACCENT<\_> /x5F <U005F> LOW LINE<'!> /x60 <U0060> GRAVE ACCENT<a> /x61 <U0061> LATIN SMALL LETTER A<b> /x62 <U0062> LATIN SMALL LETTER B<c> /x63 <U0063> LATIN SMALL LETTER C<d> /x64 <U0064> LATIN SMALL LETTER D<e> /x65 <U0065> LATIN SMALL LETTER E<f> /x66 <U0066> LATIN SMALL LETTER F<g> /x67 <U0067> LATIN SMALL LETTER G<h> /x68 <U0068> LATIN SMALL LETTER H<i> /x69 <U0069> LATIN SMALL LETTER I<j> /x6A <U006A> LATIN SMALL LETTER J<k> /x6B <U006B> LATIN SMALL LETTER K<l> /x6C <U006C> LATIN SMALL LETTER L<m> /x6D <U006D> LATIN SMALL LETTER M<n> /x6E <U006E> LATIN SMALL LETTER N<o> /x6F <U006F> LATIN SMALL LETTER O<p> /x70 <U0070> LATIN SMALL LETTER P<q> /x71 <U0071> LATIN SMALL LETTER Q<r> /x72 <U0072> LATIN SMALL LETTER R<s> /x73 <U0073> LATIN SMALL LETTER S<t> /x74 <U0074> LATIN SMALL LETTER T<u> /x75 <U0075> LATIN SMALL LETTER U<v> /x76 <U0076> LATIN SMALL LETTER V<w> /x77 <U0077> LATIN SMALL LETTER W<x> /x78 <U0078> LATIN SMALL LETTER X<y> /x79 <U0079> LATIN SMALL LETTER Y<z> /x7A <U007A> LATIN SMALL LETTER Z<(!> /x7B <U007B> LEFT CURLY BRACKET<!!> /x7C <U007C> VERTICAL LINE<!)> /x7D <U007D> RIGHT CURLY BRACKET<'?> /x7E <U007E> TILDE<DT> /x7F <U007F> DELETE (DEL)<NUL> /x00 <U0000> NUL<SOH> /x01 <U0001> START OF HEADING (SOH)<STX> /x02 <U0002> START OF TEXT (STX)<ETX> /x03 <U0003> END OF TEXT (ETX)<EOT> /x04 <U0004> END OF TRANSMISSION (EOT)<ENQ> /x05 <U0005> ENQUIRY (ENQ)<ACK> /x06 <U0006> ACKNOWLEDGE (ACK)<alert> /x07 <U0007> BELL (BEL)<BEL> /x07 <U0007> BELL (BEL)<backspace> /x08 <U0008> BACKSPACE (BS)<tab> /x09 <U0009> CHARACTER TABULATION (HT)<newline> /x0A <U000A> LINE FEED (LF)<vertical-tab> /x0B <U000B> LINE TABULATION (VT)<form-feed> /x0C <U000C> FORM FEED (FF)<carriage-return> /x0D <U000D> CARRIAGE RETURN (CR)<DLE> /x10 <U0010> DATALINK ESCAPE (DLE)<DC1> /x11 <U0011> DEVICE CONTROL ONE (DC1)<DC2> /x12 <U0012> DEVICE CONTROL TWO (DC2)<DC3> /x13 <U0013> DEVICE CONTROL THREE (DC3)<DC4> /x14 <U0014> DEVICE CONTROL FOUR (DC4)<NAK> /x15 <U0015> NEGATIVE ACKNOWLEDGE (NAK)<SYN> /x16 <U0016> SYNCHRONOUS IDLE (SYN)<ETB> /x17 <U0017> END OF TRANSMISSION BLOCK (ETB)<CAN> /x18 <U0018> CANCEL (CAN)<SUB> /x1A <U001A> SUBSTITUTE (SUB)<ESC> /x1B <U001B> ESCAPE (ESC)<IS4> /x1C <U001C> FILE SEPARATOR (IS4)<IS3> /x1D <U001D> GROUP SEPARATOR (IS3)<intro> /x1D <U001D> GROUP SEPARATOR (IS3)<IS2> /x1E <U001E> RECORD SEPARATOR (IS2)<IS1> /x1F <U001F> UNIT SEPARATOR (IS1)<DEL> /x7F <U007F> DELETE (DEL)<space> /x20 <U0020> SPACE<exclamation-mark> /x21 <U0021> EXCLAMATION MARK<quotation-mark> /x22 <U0022> QUOTATION MARK<number-sign> /x23 <U0023> NUMBER SIGN<dollar-sign> /x24 <U0024> DOLLAR SIGN<percent-sign> /x25 <U0025> PERCENT SIGN<ampersand> /x26 <U0026> AMPERSAND<apostrophe> /x27 <U0027> APOSTROPHE<left-parenthesis> /x28 <U0028> LEFT PARENTHESIS<right-parenthesis> /x29 <U0029> RIGHT PARENTHESIS<asterisk> /x2A <U002A> ASTERISK<plus-sign> /x2B <U002B> PLUS SIGN<comma> /x2C <U002C> COMMA<hyphen> /x2D <U002D> HYPHEN-MINUS<hyphen-minus> /x2D <U002D> HYPHEN-MINUS<period> /x2E <U002E> FULL STOP<full-stop> /x2E <U002E> FULL STOP<slash> /x2F <U002F> SOLIDUS<solidus> /x2F <U002F> SOLIDUS<zero> /x30 <U0030> DIGIT ZERO<one> /x31 <U0031> DIGIT ONE<two> /x32 <U0032> DIGIT TWO<three> /x33 <U0033> DIGIT THREE<four> /x34 <U0034> DIGIT FOUR<five> /x35 <U0035> DIGIT FIVE<six> /x36 <U0036> DIGIT SIX<seven> /x37 <U0037> DIGIT SEVEN<eight> /x38 <U0038> DIGIT EIGHT<nine> /x39 <U0039> DIGIT NINE<colon> /x3A <U003A> COLON<semicolon> /x3B <U003B> SEMICOLON<less-than-sign> /x3C <U003C> LESS-THAN SIGN<equals-sign> /x3D <U003D> EQUALS SIGN<greater-than-sign> /x3E <U003E> GREATER-THAN SIGN<question-mark> /x3F <U003F> QUESTION MARK<commercial-at> /x40 <U0040> COMMERCIAL AT<left-square-bracket> /x5B <U005B> LEFT SQUARE BRACKET<backslash> /x5C <U005C> REVERSE SOLIDUS<reverse-solidus> /x5C <U005C> REVERSE SOLIDUS<right-square-bracket> /x5D <U005D> RIGHT SQUARE BRACKET<circumflex> /x5E <U005E> CIRCUMFLEX ACCENT<circumflex-accent> /x5E <U005E> CIRCUMFLEX ACCENT<underscore> /x5F <U005F> LOW LINE<low-line> /x5F <U005F> LOW LINE<grave-accent> /x60 <U0060> GRAVE ACCENT<left-brace> /x7B <U007B> LEFT CURLY BRACKET<left-curly-bracket> /x7B <U007B> LEFT CURLY BRACKET<vertical-line> /x7C <U007C> VERTICAL LINE<right-brace> /x7D <U007D> RIGHT CURLY BRACKET<right-curly-bracket> /x7D <U007D> RIGHT CURLY BRACKET<tilde> /x7E <U007E> TILDEEND CHARMAP |