

NAME

charmap – character set description file

DESCRIPTION

A character set description (charmap) defines all available characters and their encodings in a character set. **localedef(1)** can use charmaps to create locale variants for different character sets.

Syntax

The charmap file starts with a header that may consist of the following keywords:

<code_set_name>

is followed by the name of the character map.

<comment_char>

is followed by a character that will be used as the comment character for the rest of the file. It defaults to the number sign (#).

<escape_char>

is followed by a character that should be used as the escape character for the rest of the file to mark characters that should be interpreted in a special way. It defaults to the backslash (\).

<mb_cur_max>

is followed by the maximum number of bytes for a character. The default value is 1.

<mb_cur_min>

is followed by the minimum number of bytes for a character. This value must be less than or equal than *<mb_cur_max>*. If not specified, it defaults to *<mb_cur_max>*.

The character set definition section starts with the keyword *CHARMAP* in the first column.

The following lines may have one of the two following forms to define the character set:

<character> byte-sequence comment

This form defines exactly one character and its byte sequence, *comment* being optional.

<character>..<character>* byte-sequence comment*

This form defines a character range and its byte sequence, *comment* being optional.

The character set definition section ends with the string *END CHARMAP*.

The character set definition section may optionally be followed by a section to define widths of characters.

The *WIDTH_DEFAULT* keyword can be used to define the default width for all characters not explicitly listed. The default character width is 1.

The width section for individual characters starts with the keyword *WIDTH* in the first column.

The following lines may have one of the two following forms to define the widths of the characters:

<character> width

This form defines the width of exactly one character.

<character>...<character> width

This form defines the width for all the characters in the range.

The width definition section ends with the string *END WIDTH*.

FILES

/usr/share/i18n/charmaps

Usual default character map path.

CONFORMING TO

POSIX.2.

EXAMPLE

The Euro sign is defined as follows in the *UTF-8* charmap:

```
<U20AC> /xe2/x82/xac EURO SIGN
```

SEE ALSO

iconv(1), locale(1), localedef(1), locale(5), charsets(7)

COLOPHON

This page is part of release 5.05 of the Linux *man-pages* project. A description of the project, information about reporting bugs, and the latest version of this page, can be found at <https://www.kernel.org/doc/man-pages/>.