NAME

mbchar, mbtowc, mblen, wctomb – multibyte character handling

SYNOPSIS

#include <stdlib.h>

#include <limits.h>

int mbtowc(wchar_t **pwc*, **const char** **s*, **size_t** *n***);**

int mblen(const char **s*, **size_t** *n***);**

int wctomb(char **s*, **wchar_t** *wchar***);**

MT-LEVEL

MT-Safe with exceptions

DESCRIPTION

Multibyte characters are used to represent characters in an extended character set. This is needed for locales where 8 bits are not enough to represent all the characters in the character set.

The multibyte character handling functions provide the means of translating multibyte characters into wide characters and back again. Wide characters have type **wchar_t** (defined in **<stdlib.h>**), which is an integral type whose range of values can represent distinct codes for all members of the largest extended character set specified among the supported locales.

A maximum of 3 extended character sets are supported for each locale. The number of bytes in an extended character set is defined by the **LC_CTYPE** category of the locale (see **setlocale**(3C)). However, the maximum number of bytes in any multibyte character will never be greater than **MB_LEN_MAX**. which is defined in **<limits.h>**. The maximum number of bytes in a character in an extended character set in the **current** locale is given by the macro, **MB_CUR_MAX**, defined in **<stdlib.h>**.

mbtowc( ) determines the number of bytes that comprise the multibyte character pointed to by *s*. Also, if *pwc* is not a null pointer, **mbtowc( )** converts the multibyte character to a wide character and places the result in the object pointed to by *pwc*. (The value of the wide character corresponding to the null character is zero.) At most *n* bytes will be examined, starting at the byte pointed to by *s*.

If *s* is a null pointer, **mbtowc( )** simply returns 0. If *s* is not a null pointer, then, if *s* points to the null character, **mbtowc( )** returns 0; if the next *n* or fewer bytes form a valid multibyte character, **mbtowc( )** returns the number of bytes that comprise the converted multibyte character; otherwise, *s* does not point to a valid multibyte character and **mbtowc( )** returns –1.

mblen( ) determines the number of bytes comprising the multibyte character pointed to by *s*. It is equivalent to

mbtowc((wchar_t *)0, s, n);

wctomb( ) determines the number of bytes needed to represent the multibyte character corresponding to the code whose value is *wchar*, and, if *s* is not a null pointer, stores the multibyte character representation in the array pointed to by *s*. At most **MB_CUR_MAX** bytes are stored.

If *s* is a null pointer, **wctomb( )** simply returns 0. If *s* is not a null pointer, **wctomb( )** returns −1 if the value of *wchar* does not correspond to a valid multibyte character; otherwise it returns the number of bytes that comprise the multibyte character corresponding to the value of *wchar*.

SEE ALSO

chrtbl(1M), **mbstring**(3C), **setlocale**(3C), **environ**(5)

NOTES

mbchar, **mbtowc**, **mblen** and **wctomb** can be used safely in a multi-thread application, as long as **setlocale**(3C) is not being called to change the locale.