NAME
mbs − work with multibyte strings

SYNOPSIS
```
@load "mbs"

byte_count = mbs_length(string)
count = mbs_split(string, array)
new_string = mbs_join(array)
width = mbs_wcswidth(string)
```

DESCRIPTION
The *mbs* extension adds four functions, as follows:

mbs_length(*string***)**

This function takes a string argument and returns the number of bytes it contains. This is useful in multibyte locales where the number of characters does not equal the number of bytes.

mbs_split(*string***,** *array***)**

This function takes a string value, and an array. It clears the array, and then populates it with the numeric byte values in the string (*not* the individual characters as strings). It returns the number of elements in the array.

mbs_join(*array***)**

This function is the inverse of **mbs_split()**. It takes an array of byte values and returns an AWK string composed of those values. **NOTE**: The function does *not* validate that the sequence of bytes creates valid strings in the current encoding.

mbs_wcswidth(*string***[,** *not_utf8***])**

This function takes a string and computes the display width (number of columns) taken up by the string. It is similar in function to the standard C *wcswidth*(3) library routine. By default this function assumes that the current locale is one based on UTF-8. If that is not the case, the caller *must* pass a second argument that is non-zero. (**NOTE**: The extension expects a numeric argument to be supplied, so a string value will not work.)

The return value from **mbs_wcswidth()** can differ from that of the built-in **length()** function when the string contains combining characters and other such unusual entities.

NOTES
For the **mbs_wcswidth()** function, determining if the current locale is based on UTF-8 or not is left up to you, the programmer. The locale naming conventions vary widely (and wildly) from system to system, and this extension isn't going to venture into that dragon-filled territory.

EXAMPLE
```
@load "mbs"
...
printf "The length in bytes of %s is %d\n", $0, mbs_length($0)
```

SEE ALSO
wcswidth(3)

AUTHOR
Arnold D. Robbins

COPYING PERMISSIONS
Copyright © 2018, Free Software Foundation, Inc.

Permission is granted to make and distribute verbatim copies of this manual page provided the copyright notice and this permission notice are preserved on all copies.

Permission is granted to copy and distribute modified versions of this manual page under the conditions for verbatim copying, provided that the entire resulting derived work is distributed under the terms of a

permission notice identical to this one.

Permission is granted to copy and distribute translations of this manual page into another language, under the above conditions for modified versions, except that this permission notice may be stated in a translation approved by the Foundation.