ustrfix() — Replace invalid UTF-8 sequences in Unicode string

Description

ustrfix(s [, rep]) replaces each invalid UTF-8 sequence with a Unicode character. If rep is specified and it starts with a Unicode character, the Unicode character is used. Otherwise, the Unicode replacement character \ufffd is used.

When arguments are not scalar, the function returns element-by-element results.

Syntax

    string matrix  ustrfix(string matrix s [, string scalar rep])

Remarks and examples

An invalid UTF-8 sequence may contain one byte or multiple bytes.

Conformability

ustrfix(matrix s [, r])
    s: r × c
    rep: 1 × 1
    result: r × c

Diagnostics

ustrfix(matrix s [, r]) returns an empty string if an error occurs.

Also see

[M-5] ustrto() — Convert a Unicode string to or from a string in a specified encoding
[M-5] ustrunescape() — Convert escaped hex sequences to Unicode strings
[M-4] String — String manipulation functions
[U] 12.4.2 Handling Unicode strings