

NAME

big5 - "Big Five" encoding for Traditional Chinese text

SYNOPSIS

ENCODING "BIG5"

DESCRIPTION

"Big Five" is a standard for encoding Traditional Chinese text. Each character is represented by either one or two bytes. Characters from the ASCII character set are represented as single bytes in the range 0x00 - 0x7F. Traditional Chinese characters are represented by two bytes: the first in the range 0xA1 - 0xFE, the second in the range 0x40 - 0xFE.

SEE ALSO

eut(5), gb18030(5), utf8(5)

BUGS

The range of the second byte overlaps some ASCII characters, including 0x5C (‘\’) and 0x7C (‘|’) which may cause problems in program execution or display. Big5 is considered a legacy standard and only preserved for backward compatibility reason. New documents and systems are suggested using UTF-8 directly.