

NAME

gbk - Guojia biao zhun kuozhan (GBK) encoding method for Chinese text

SYNOPSIS

ENCODING "GBK"

DESCRIPTION

GBK is a backwards-compatible extension of the GB 2312-1980 encoding method for Chinese text, which adds the characters defined in the Unified Han portion of the Unicode 2.1 standard.

Multibyte characters in the GBK encoding can be one byte or two bytes long. GB 11383-1981 (ASCII) characters are represented by single bytes in the range 0x00 to 0x7F. Chinese characters are represented by two bytes, beginning with a byte in the range 0x80-0xFE and ending with a byte in the range 0x40-0xFE.

SEE ALSO

euc(5), gb18030(5), gb2312(5), utf8(5)

The Unicode Standard, Version 2.1, The Unicode Consortium, 1999.

Chinese National Standard GB 18030-2000: Information Technology -- Chinese ideograms coded character set for information interchange -- Extension for the basic set, March 2000.

STANDARDS

GBK is not a standard, but has been superseded by GB 18030-2000.