

NAME

Encode::KR - Korean Encodings

SYNOPSIS

```
use Encode qw/encode decode/;
$euc_kr = encode("euc-kr", $utf8); # loads Encode::KR implicitly
$utf8   = decode("euc-kr", $euc_kr); # ditto
```

DESCRIPTION

This module implements Korean charset encodings. Encodings supported are as follows.

Canonical	Alias	Description
euc-kr	<code>/\beuc.*kr\$/i</code> <code>/\bkr.*euc\$/i</code>	EUC (Extended Unix Character)
ksc5601-raw		Korean standard code set (as is)
cp949	<code>/(?:x-)?uhc\$/i</code> <code>/(?:x-)?windows-949\$/i</code> <code>/\bks_c_5601-1987\$/i</code>	Code Page 949 (EUC-KR + 8,822 (additional Hangeul syllables)
MacKorean		EUC-KR + Apple Vendor Mappings
johab	JOHAB	A supplementary encoding defined in Annex 3 of KS X 1001:1998
iso-2022-kr	iso-2022-kr	[RFC1557]

To find how to use this module in detail, see *Encode*.

BUGS

When you see `charset=ks_c_5601-1987` on mails and web pages, they really mean "cp949" encodings. To fix that, the following aliases are set;

```
qr/(?:x-)?uhc$/i      => 'cp949'
qr/(?:x-)?windows-949$/i => 'cp949'
qr/ks_c_5601-1987$/i  => 'cp949'
```

The ASCII region (0x00-0x7f) is preserved for all encodings, even though this conflicts with mappings by the Unicode Consortium. See

<http://www.debian.or.jp/~kubota/unicode-symbols.html.en>

to find out why it is implemented that way.

SEE ALSO

Encode