



POD Translation by *pod2pdf*

ajf@afco.demon.co.uk

UTF8.pm

Table of Contents

UTF8.pm

NAME	1
SYNOPSIS	1
DESCRIPTION	1
METHODS	1
new()	1
lookup()	1
combining()	1
TODO	1
AUTHORS	1
Ed Summers <ehs@pobox.com>	1

NAME

MARC::Charset::UTF8 - UTF-8 = MARC-8 mapping

SYNOPSIS

```
use MARC::Charset::UTF8;
my $cs = MARC::Charset::UTF8->new();

## convert some utf8 to marc8
my $marc8 = $cs->to_marc8( $unicode );

## see what charsets have been used so far by this charset object
my @charsets = $cs->charsets();

## what is the current G0 charset
my $g0 = $cs->g0();

## what is the current G1 charset
my $g1 = $cs->g1();
```

DESCRIPTION

Unlike all the other MARC::Charset::* classes, MARC::Charset::UTF8 attempts to convert a Unicode character into it's MARC-8 equivalent. Obviously this is a lossy process since MARC-8 doesn't support anywhere near the wide variety of characters that Unicode does...but it does its best.

When you installed MARC::Charset the other MARC::Charset::* mappings were turned on their head to create one Berkeley database to house the mapping. If you are curious you should be able to find this Berkeley DB living in the same directory which you installed MARC::Charset::UTF8 into.

METHODS

new()

The constructor, which will return you a MARC::Charset::UTF8 object.

lookup()

The workhorse method that does the lookup. Pass it an a character and you'll get back some data identifying a MARC8 character.

combining()

Pass it a character and you'll get back a true value (1) if the character is a combining character, and false (undef) if it is not.

TODO

AUTHORS

Ed Summers <ehs@pobox.com>

