#----------------------------------------------------------------

# The examples here were run on Windows, using "fc /B" for all

# bytewise file comparisons; the Unix equivalent is "cmp -b".

# A plain DOS "fc" ignores end-of-line differences crucial here

# ("fc" is like a Unix "diff -w" or "diff --strip-trailing-cr").

# See learning-python.com/unicodemod/ for the related tool used.

#----------------------------------------------------------------

#################################################################

# Convert a UTF-8 text file (fc=file compare, default=utf8)

#################################################################

c:\...\frigcal\docetc> copy unicode-cheat-sheet.txt temp

1 file(s) copied.

c:\...\frigcal\docetc> fixeoln.py todos temp utf8

Using mode=todos, file=temp, encoding=utf8

No changes to temp

c:\...\frigcal\docetc> fixeoln.py tounix temp utf8

Using mode=tounix, file=temp, encoding=utf8

Converted temp

c:\...\frigcal\docetc> fixeoln.py todos temp

Using mode=todos, file=temp, encoding=UTF-8

Converted temp

c:\...\frigcal\docetc> fc /B unicode-cheat-sheet.txt temp # ignore eolns if no /B

Comparing files unicode-cheat-sheet.txt and TEMP

FC: no differences encountered

#################################################################

# Convert a UTF-16 text file (via unicodemod.py utility)

#################################################################

c:\...\frigcal\docetc> ..\unicodemod.py temp utf8 utf16

File successfully converted from utf8 to utf16.

c:\...\frigcal\docetc> fixeoln.py todos temp utf16

Using mode=todos, file=temp, encoding=utf16

No changes to temp

c:\...\frigcal\docetc> fixeoln.py tounix temp utf16

Using mode=tounix, file=temp, encoding=utf16

Converted temp

c:\...\frigcal\docetc> fixeoln.py todos temp utf16

Using mode=todos, file=temp, encoding=utf16

Converted temp

c:\...\frigcal\docetc> ..\unicodemod.py temp utf16 utf-8-sig # assume had BOM: restore

File successfully converted from utf16 to utf8.

c:\...\frigcal\docetc> fc /B unicode-cheat-sheet.txt temp

Comparing files unicode-cheat-sheet.txt and TEMP

FC: no differences encountered

c:\...\frigcal\docetc> py # manual 'fc /b' (or unix 'cmp -b')

>>> open('temp', 'rb').read() == open('unicode-cheat-sheet.txt', 'rb').read()

True

>>> open('temp', 'rb').read()[:79]

b'\xef\xbb\xbf#------------------------------------------------------------------------\r\n#'

>>> ^Z

#################################################################

# Ditto, but for a UTF-32 text file

#################################################################

c:\...\frigcal\docetc> ..\unicodemod.py temp utf8 utf32

File successfully converted from utf8 to utf32.

c:\...\frigcal\docetc> fixeoln.py tounix temp utf32

Using mode=tounix, file=temp, encoding=utf32

Converted temp

c:\...\frigcal\docetc> fixeoln.py todos temp utf32

Using mode=todos, file=temp, encoding=utf32

Converted temp

c:\...\frigcal\docetc> ..\unicodemod.py temp utf32 utf-8-sig # or utf8 = no BOM in result

File successfully converted from utf32 to utf8.

c:\...\frigcal\docetc> fc /B unicode-cheat-sheet.txt temp

Comparing files unicode-cheat-sheet.txt and TEMP

FC: no differences encountered

#################################################################

# Simple ASCII files: ascii, utf8, or default=utf8

#################################################################

c:\...\frigcal\docetc> copy README.txt temp

Overwrite temp? (Yes/No/All): y

1 file(s) copied.

c:\...\frigcal\docetc> fixeoln.py tounix temp

Using mode=tounix, file=temp, encoding=UTF-8

Converted temp

c:\...\frigcal\docetc> fixeoln.py todos temp

Using mode=todos, file=temp, encoding=UTF-8

Converted temp

c:\...\frigcal\docetc> fixeoln.py tounix temp ascii

Using mode=tounix, file=temp, encoding=ascii

Converted temp

c:\...\frigcal\docetc> fixeoln.py todos temp ascii

Using mode=todos, file=temp, encoding=ascii

Converted temp

c:\...\frigcal\docetc> fixeoln.py tounix temp utf8

Using mode=tounix, file=temp, encoding=utf8

Converted temp

c:\...\frigcal\docetc> fixeoln.py todos temp utf8

Using mode=todos, file=temp, encoding=utf8

Converted temp

c:\...\frigcal\docetc> fc /B temp README.txt

Comparing files temp and README.TXT

FC: no differences encountered