Ruby master - Feature #4224

inconsistent behavior of 'puts "\xFF"'

12/29/2010 11:42 PM - phasis68 (Heesob Park)

Status:	Closed	
Priority:	Normal	
Assignee:	usa (Usaku NAKAMURA)	
Target version:	1.9.3	

Description

=begin
C:\work>ruby -ve 'puts "\xFFabc"'
ruby 1.9.3dev (2010-12-28 trunk 30412) [i386-mswin32_90]
-e:1:in write': "\xFF" on CP949 (Encoding::InvalidByteSequenceError)
from -e:1:inputs'
from -e:1:in puts'
from -e:1:in'

$ ruby -ve 'puts "\xFFabc"'
ruby 1.9.3dev (2010-12-02) [i686-linux]
abc

C:>ruby -ve 'puts "\xFFabc"'
ruby 1.9.2p136 (2010-12-25) [i386-mingw32]
abc

The trunk version raises Encoding::InvalidByteSequenceError on Windows.
But, on the other version and other platform, it works without error.
What is the proper behavior?

I think this error could be avoided by following patch:

--- win32.c   2010-12-28 13:59:30.000000000 +0900
+++ win32.c.new 2010-12-28 14:00:18.000000000 +0900
@@ -5312,7 +5312,7 @@
!rb_econv_has_convpath_p(rb_enc_name(rb_enc_get(str)), "UTF-16LE"))
return -1L;

- str = rb_str_encode(str, rb_enc_from_encoding(rb_enc_find("UTF-16LE")), 0,
- str = rb_str_encode(str, rb_enc_from_encoding(rb_enc_find("UTF-16LE")),
  ECONV_INVALID_REPLACE|ECONV_UNDEF_REPLACE,   Qnil);   if (!WriteConsoleW(handle,
  (LPWSTR)RSTRING_PTR(str), RSTRING_LEN(str)/2,   &reslen, NULL)) { =end

History

#1 - 01/01/2011 03:50 AM - naruse (Yui NARUSE)

- *Status changed from Open to Assigned*

- *Assignee set to usa (Usaku NAKAMURA)*

=begin
"Current spec" seems:

- if external encoding is set, output may raise exception
- if not set, output won't raise exception

Because this case is latter even if ruby converts strings internally,
it shouldn't raise exception.

Anyway in this case specifically, \xFF should be converted to U+F8F7.
So *this* is conversion table's bug.
http://icu-project.org/repos/icu/data/trunk/charset/data/ucm/windows-949-2000.ucm
=end

#2 - 01/01/2011 04:17 AM - naruse (Yui NARUSE)

=begin
After further thought, current conversion table is correct.
And your patch is the correct way.
=end

#3 - 01/01/2011 04:27 AM - naruse (Yui NARUSE)

- Status changed from Assigned to Closed

- % Done changed from 0 to 100

=begin
This issue was solved with changeset r30440.
Heesob, thank you for reporting this issue.
Your contribution to Ruby is greatly appreciated.
May Ruby be with you.

=end