

Ruby master - Bug #4217

irb exits unexpectedly with non-ascii Regexp on Windows

12/28/2010 02:09 PM - phasis68 (Heesob Park)

Status:	Closed	
Priority:	Normal	
Assignee:		
Target version:	1.9.3	
ruby -v:	ruby 1.9.3dev (2010-12-28 trunk 30406) [i386-mswin32_90]	Backport:

Description

```
=begin
C:\work>irb
irb(main):001:0> Regexp.new("\xFF")
RegexpError: c:/usr/lib/ruby/1.9.1/irb.rb:168:in write': "\xFF" on CP949 (Encoding::InvalidByteSequenceError)
from c:/usr/lib/ruby/1.9.1/irb.rb:168:in print'
from c:/usr/lib/ruby/1.9.1/irb.rb:168:in block (2 levels) in eval_input'
from c:/usr/lib/ruby/1.9.1/irb.rb:273:in signal_status'
from c:/usr/lib/ruby/1.9.1/irb.rb:156:in block in eval_input'
from c:/usr/lib/ruby/1.9.1/irb/ruby-lex.rb:243:in block (2 levels) in each_top_level_statement'
from c:/usr/lib/ruby/1.9.1/irb/ruby-lex.rb:229:in loop'
from c:/usr/lib/ruby/1.9.1/irb/ruby-lex.rb:229:in block in each_top_level_statement'
from c:/usr/lib/ruby/1.9.1/irb/ruby-lex.rb:228:in catch'
from c:/usr/lib/ruby/1.9.1/irb/ruby-lex.rb:228:in each_top_level_statement'
from c:/usr/lib/ruby/1.9.1/irb.rb:155:in eval_input'
from c:/usr/lib/ruby/1.9.1/irb.rb:70:in block in start'
from c:/usr/lib/ruby/1.9.1/irb.rb:69:in catch'
from c:/usr/lib/ruby/1.9.1/irb.rb:69:in start'
from c:/usr/bin/irb.bat:19:in `
C:\work>

C:\work>irb
irb(main):001:0> /\xFF/
SyntaxError: c:/usr/lib/ruby/1.9.1/irb.rb:168:in write': "\xB0" to UTF-8 in conversion from ASCII-8BIT to UTF-8 to UTF-16LE
(Encoding::UndefinedConversionError)
from c:/usr/lib/ruby/1.9.1/irb.rb:168:in print'
from c:/usr/lib/ruby/1.9.1/irb.rb:168:in block (2 levels) in eval_input'
from c:/usr/lib/ruby/1.9.1/irb.rb:273:in signal_status'
from c:/usr/lib/ruby/1.9.1/irb.rb:156:in block in eval_input'
from c:/usr/lib/ruby/1.9.1/irb/ruby-lex.rb:243:in block (2 levels) in each_top_level_statement'
from c:/usr/lib/ruby/1.9.1/irb/ruby-lex.rb:229:in loop'
from c:/usr/lib/ruby/1.9.1/irb/ruby-lex.rb:229:in block in each_top_level_statement'
from c:/usr/lib/ruby/1.9.1/irb/ruby-lex.rb:228:in catch'
from c:/usr/lib/ruby/1.9.1/irb/ruby-lex.rb:228:in each_top_level_statement'
from c:/usr/lib/ruby/1.9.1/irb.rb:155:in eval_input'
from c:/usr/lib/ruby/1.9.1/irb.rb:70:in block in start'
from c:/usr/lib/ruby/1.9.1/irb.rb:69:in catch'
from c:/usr/lib/ruby/1.9.1/irb.rb:69:in start'
from c:/usr/bin/irb.bat:19:in `
C:\work>
```

I found that this bug is related with `rb_w32_write_console` function.
Here is a possible 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")),`

Associated revisions

Revision 30412 - 12/28/2010 09:43 AM - usa (Usaku NAKAMURA)

- re.c (rb_reg_expr_str): need to escape if the coderage is invalid.
 - error.c, include/ruby/intern.h (rb_compile_error_with_enc): new function to raise syntax error, with source encoding'ed message.
 - parse.y (compile_error): use above function. [ruby-core:33951](#)

Revision 30412 - 12/28/2010 09:43 AM - usa (Usaku NAKAMURA)

- re.c (rb_reg_expr_str): need to escape if the coderage is invalid.
 - error.c, include/ruby/intern.h (rb_compile_error_with_enc): new function to raise syntax error, with source encoding'ed message.
 - parse.y (compile_error): use above function. [ruby-core:33951](#)

Revision 30412 - 12/28/2010 09:43 AM - usa (Usaku NAKAMURA)

- re.c (rb_reg_expr_str): need to escape if the coderage is invalid.
 - error.c, include/ruby/intern.h (rb_compile_error_with_enc): new function to raise syntax error, with source encoding'ed message.
 - parse.y (compile_error): use above function. [ruby-core:33951](#)

Revision 30412 - 12/28/2010 09:43 AM - usa (Usaku NAKAMURA)

- re.c (rb_reg_expr_str): need to escape if the coderage is invalid.
 - error.c, include/ruby/intern.h (rb_compile_error_with_enc): new function to raise syntax error, with source encoding'ed message.
 - parse.y (compile_error): use above function. [ruby-core:33951](#)

Revision 30412 - 12/28/2010 09:43 AM - usa (Usaku NAKAMURA)

- re.c (rb_reg_expr_str): need to escape if the coderage is invalid.
 - error.c, include/ruby/intern.h (rb_compile_error_with_enc): new function to raise syntax error, with source encoding'ed message.
 - parse.y (compile_error): use above function. [ruby-core:33951](#)

Revision 30412 - 12/28/2010 09:43 AM - usa (Usaku NAKAMURA)

- re.c (rb_reg_expr_str): need to escape if the coderage is invalid.
 - error.c, include/ruby/intern.h (rb_compile_error_with_enc): new function to raise syntax error, with source encoding'ed message.
 - parse.y (compile_error): use above function. [ruby-core:33951](#)

History

#1 - 12/28/2010 06:47 PM - usa (Usaku NAKAMURA)

- Status changed from Open to Closed

- % Done changed from 0 to 100

=begin

This issue was solved with changeset r30412.

Heesob, thank you for reporting this issue.

Your contribution to Ruby is greatly appreciated.
May Ruby be with you.

=end

#2 - 12/29/2010 10:32 AM - phasis68 (Heesob Park)

=begin
I think r30412 is incomplete.

After applying r30412, while the second case is fixed, the first case is still exist.

```
irb(main):001:0> Regexp.new("\xFF")
RegexpError: c:/usr/lib/ruby/1.9.1/irb.rb:168:in write: "\xFF" on CP949 (Encoding::InvalidByteSequenceError)
from c:/usr/lib/ruby/1.9.1/irb.rb:168:in print'
from c:/usr/lib/ruby/1.9.1/irb.rb:168:in block (2 levels) in eval_input'
from c:/usr/lib/ruby/1.9.1/irb.rb:273:in signal_status'
from c:/usr/lib/ruby/1.9.1/irb.rb:156:in block in eval_input'
from c:/usr/lib/ruby/1.9.1/irb/ruby-lex.rb:243:in block (2 levels) in each_top_level_statement'
from c:/usr/lib/ruby/1.9.1/irb/ruby-lex.rb:229:in loop'
from c:/usr/lib/ruby/1.9.1/irb/ruby-lex.rb:229:in block in each_top_level_statement'
from c:/usr/lib/ruby/1.9.1/irb/ruby-lex.rb:228:in catch'
from c:/usr/lib/ruby/1.9.1/irb/ruby-lex.rb:228:in each_top_level_statement'
from c:/usr/lib/ruby/1.9.1/irb.rb:155:in eval_input'
from c:/usr/lib/ruby/1.9.1/irb.rb:70:in block in start'
from c:/usr/lib/ruby/1.9.1/irb.rb:69:in catch'
from c:/usr/lib/ruby/1.9.1/irb.rb:69:in start'
from c:/usr/bin/irb.bat:19:in ``
```

```
C:\work>irb
irb(main):001:0> /\xFF/
SyntaxError: (irb):1: invalid multibyte escape: /\xFF/
from c:/usr/bin/irb.bat:19:in ``
irb(main):002:0>
```

=end

#3 - 01/05/2011 09:54 AM - phasis68 (Heesob Park)

=begin
Hi,
2011/1/3 U.Nakamura usa@garbagecollect.jp:

Hello,

In message "[ruby-core:33983] [Ruby 1.9-Bug#4217] irb exits unexpectedly with non-ascii Regexp on Windows"
on Dec.29,2010 10:32:06, redmine@ruby-lang.org wrote:
| I think r30412 is incomplete.

Sorry, I missed a changeset on r30412.
I've committed it on r30418. Check it, please.

I checked it, it works fine.

BTW, what happend to <http://redmine.ruby-lang.org> ?
I cannot access redmine. And it's not just me.
<http://www.downforeveryoneorjustme.com/redmine.ruby-lang.org>

Regards,
Park Heesob

=end

#4 - 01/05/2011 09:54 AM - usa (Usaku NAKAMURA)

=begin
Hello,
In message "[ruby-core:33983] [Ruby 1.9-Bug#4217] irb exits unexpectedly with non-ascii Regexp on Windows"
on Dec.29,2010 10:32:06, redmine@ruby-lang.org wrote:
| I think r30412 is incomplete.

Sorry, I missed a changeset on r30412.

I've committed it on r30418. Check it, please.

Regards,

--

U.Nakamura usa@garbagecollect.jp

=end

#5 - 01/05/2011 09:54 AM - usa (Usaku NAKAMURA)

=begin

Hello,

In message "[ruby-core:34028] Re: [Ruby 1.9-Bug#4217] irb exits unexpectedly with non-ascii Regexp on Windows"
on Jan.03,2011 16:02:26, phasis@gmail.com wrote:

| > Sorry, I missed a changeset on r30412.

| > I've committed it on r30418. ?Check it, please.

| >

| I checked it, it works fine.

Thank you for your confirmation.

| BTW, what happend to <http://redmine.ruby-lang.org> ?

| I cannot access redmine. And it's not just me.

| <http://www.downforeveryoneorjustme.com/redmine.ruby-lang.org>

I'm not the manager of ruby servers, so this is a mere guess,
the power failure and cutting the line have occured in the
region where the servers exist because uf the heavy snow
of last week.

Unfortunately, all Japanese are on vacation of the new year,
so time is expected to be taken for restoration.

I guess that it will take more several days for restoration.

Regards,

--

U.Nakamura usa@garbagecollect.jp

=end