

Ruby master - Feature #5820

Merge Onigmo to Ruby 2.0

12/29/2011 02:42 AM - naruse (Yui NARUSE)

Status:	Closed
Priority:	Normal
Assignee:	naruse (Yui NARUSE)
Target version:	2.0.0
Description	
Ruby 1.9 の Oniguruma M17N Oniguruma Ruby 2.0 の Oniguruma k-takata Onigmo https://github.com/k-takata/Onigmo/tree/tmp/ruby-2.0.x	
Oniguruma	
<ul style="list-style-type: none">100% () [ruby-dev:44410]<ul style="list-style-type: none">* \K, \R, \X, (?cond)yes no, \g, \g<+n>, (?au) * Perl 5.10 ← RubyShift_JIS, EUC-JPShift_JIS, EUC-JP \p{Han}, \p{Latin}, \p{Greek}, \p{Cyrillic}<ul style="list-style-type: none">http://redmine.ruby-lang.org/issues/3568	
POSIX	
Related issues:	
Related to Ruby master - Bug #5208: \b and \B don't match properly Closed 08/21/2011	

Associated revisions

Revision 0424e152 - 02/17/2012 07:42 AM - naruse (Yui NARUSE)

- Merge Onigmo-5.13.1. [ruby-dev:45057] [Feature #5820] <https://github.com/k-takata/Onigmo> cp reg{comp,enc,error,exec,parse,syntax}.c reg{enc,int,parse}.h cp oniguruma.h cp tool/enc-unicode.rb cp -r enc/

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/trunk@34663 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 34663 - 02/17/2012 07:42 AM - naruse (Yui NARUSE)

- Merge Onigmo-5.13.1. [ruby-dev:45057] [Feature #5820] <https://github.com/k-takata/Onigmo> cp reg{comp,enc,error,exec,parse,syntax}.c reg{enc,int,parse}.h cp oniguruma.h cp tool/enc-unicode.rb cp -r enc/

Revision 34663 - 02/17/2012 07:42 AM - naruse (Yui NARUSE)

- Merge Onigmo-5.13.1. [ruby-dev:45057] [Feature #5820] <https://github.com/k-takata/Onigmo> cp reg{comp,enc,error,exec,parse,syntax}.c reg{enc,int,parse}.h cp oniguruma.h cp tool/enc-unicode.rb cp -r enc/

Revision 34663 - 02/17/2012 07:42 AM - naruse (Yui NARUSE)

- Merge Onigmo-5.13.1. [ruby-dev:45057] [Feature #5820] <https://github.com/k-takata/Onigmo> cp reg{comp,enc,error,exec,parse,syntax}.c reg{enc,int,parse}.h cp oniguruma.h cp tool/enc-unicode.rb cp -r enc/

Revision 34663 - 02/17/2012 07:42 AM - naruse (Yui NARUSE)

- Merge Onigmo-5.13.1. [ruby-dev:45057] [Feature #5820] <https://github.com/k-takata/Onigmo> cp reg{comp,enc,error,exec,parse,syntax}.c reg{enc,int,parse}.h cp oniguruma.h cp tool/enc-unicode.rb cp -r enc/

Revision 34663 - 02/17/2012 07:42 AM - naruse (Yui NARUSE)

- Merge Onigmo-5.13.1. [ruby-dev:45057] [Feature #5820] <https://github.com/k-takata/Onigmo> cp reg{comp,enc,error,exec,parse,syntax}.c reg{enc,int,parse}.h cp oniguruma.h cp tool/enc-unicode.rb cp -r enc/

- Merge Onigmo-5.13.1. [ruby-dev:45057] [Feature #5820] <https://github.com/k-takata/Onigmo> cp reg{comp,enc,error,exec,parse,syntax}.c reg{enc,int,parse}.h cp oniguruma.h cp tool/enc-unicode.rb cp -r enc/

History

#1 - 12/29/2011 02:44 AM - naruse (Yui NARUSE)

= 0000000000

Ruby00000000:

```
cd Onigmo
cp reg{comp,enc,error,exec,parse,syntax}.c reg{enc,int,parse}.h ~/ruby
cp oniguruma.h ~/ruby/include/ruby/
```

```
Onigmo: edf9df1cefc28472085a857116153529253ffab5 (tmp/ruby-2.0.x)
Ruby: r34150
```

```
1) Failure:
test_parse_utf8(TestDateParse)
[/home/naruse/ruby/test/date/test_date_parse.rb:651]:
<"00"> expected but was
.
```

```
00: 0000000000000000
[:alpha:] 0 ASCII 00000000000000000000
```

```
2) Failure:
test_iso_8859_1(TestISO8859) [(eval):7]:
Expected /SS$ĩ to match "\xDF".
```

```
3) Failure:
test_iso_8859_10(TestISO8859) [(eval):7]:
Expected /SS$ĩ to match "\xDF".
```

```
4) Failure:
test_iso_8859_13(TestISO8859) [(eval):7]:
Expected /SS$ĩ to match "\xDF".
```

```
5) Failure:
test_iso_8859_14(TestISO8859) [(eval):7]:
Expected /SS$ĩ to match "\xDF".
```

```
6) Failure:
test_iso_8859_15(TestISO8859) [(eval):7]:
Expected /SS$ĩ to match "\xDF".
```

```
7) Failure:
test_iso_8859_16(TestISO8859) [(eval):7]:
Expected /SS$ĩ to match "\xDF".
```

```
8) Failure:
test_iso_8859_2(TestISO8859) [(eval):7]:
Expected /SS$ĩ to match "\xDF".
```

```
9) Failure:
test_iso_8859_3(TestISO8859) [(eval):5]:
Expected /SS$ĩ to match "\xDF".
```

```
10) Failure:
test_iso_8859_4(TestISO8859) [(eval):7]:
Expected /SS$ĩ to match "\xDF".
```

```
11) Failure:
test_iso_8859_9(TestISO8859) [(eval):5]:
Expected /SS$ĩ to match "\xDF".
```

```
000
00000000000000000000?
```

```
12) Failure:
```

test_regexp_named_class(TestM17N) [/home/naruse/ruby/test/ruby/test_m17n.rb:543]:
Expected `/.space/` to match `" "`.

13) Failure:
test_posix_bracket(TestRegexp) [/home/naruse/ruby/test/ruby/test_regexp.rb:716]:
Expected `^A\digit.+z/` to match `"\d\d\d\d\d\d\d\d"`.

POSIX `\d\d\d\d\d\d` ASCII `\d\d\d\d\d\d`

14) Failure:
test_unicode(TestRegexp) [/home/naruse/ruby/test/ruby/test_regexp.rb:792]:
Expected `/^u0149$/i` to match `"n"`.

2)-11) `\d\d\d\d?`

#2 - 12/29/2011 04:23 AM - kosaki (Motohiro KOSAKI)

Ruby 1.9 `\d\d\d\d\d\d\d\d\d\d` M17N `\d\d\d\d\d\d` Oniguruma `\d\d\d\d\d\d\d\d`
`\d\d\d` 2.0 `\d\d` Oniguruma `\d\d\d\d\d\d\d\d` k-takata `\d\d` Onigmo `\d\d\d\d\d\d\d\d\d\d\d\d`
<https://github.com/k-takata/Onigmo/tree/tmp/ruby-2.0.x>

`\d\d` oniguruma `\d\d` upstream `\d`
`\d\d\d\d\d\d` libonigmo.so `\d`
`\d` k-takata `\d\d`
`\d` k-tanaka `\d\d` Ruby `\d\d\d\d\d\d\d\d\d\d\d\d\d\d` onigumo `\d\d\d\d\d\d\d\d\d\d\d\d\d\d`
`\d`

`\d\d\d\d\d` **Ruby 3.0** `\d`

#3 - 12/30/2011 03:37 PM - k_takata (Ken Takata)

POSIX `\d\d\d\d\d\d` Perl `\d\d/a` ASCII `\d\d\d\d\d\d\d\d\d\d` Onigmo `\d\d\d\d\d\d\d\d\d\d`
`\d`
`\d\d\d\d\d\d\d\d\d\d\d\d\d\d\d\d` Ruby `\d\d\d\d\d\d\d\d\d\d\d\d\d\d`

`\d\d\d\d\d\d\d\d`

#4 - 12/30/2011 07:41 PM - naruse (Yui NARUSE)

Motohiro KOSAKI wrote:

`\d\d\d\d\d` libonigmo.so `\d\d\d\d\d\d\d\d\d\d\d\d\d\d\d\d\d\d`
`\d`

Ruby `\d` Ruby `\d`
`\d\d\d\d\d\d\d\d\d\d\d\d\d\d\d\d`
`\d`

k-takata `\d\d`
`\d`

`\d\d\d\d\d\d\d\d\d\d` Ruby `\d\d\d\d\d\d\d\d\d\d\d\d`

k-tanaka `\d\d` Ruby `\d\d\d\d\d\d\d\d\d\d\d\d` onigumo `\d\d\d\d\d\d\d\d\d\d\d\d`
`\d\d\d\d\d\d\d\d\d\d\d\d\d\d\d\d`
`\d\d`

k-tanaka `\d\d\d\d\d\d\d\d\d\d`

Ken Takata wrote:

POSIX `\d\d\d\d\d\d` Perl `\d\d/a` ASCII `\d\d\d\d\d\d\d\d\d\d` Onigmo `\d\d\d\d\d\d\d\d\d\d`
`\d\d\d\d\d\d\d\d\d\d\d\d\d\d\d\d`
`\d\d\d\d\d\d\d\d\d\d\d\d\d\d\d\d` Ruby `\d\d\d\d\d\d\d\d\d\d\d\d`

2.0
2.0

Ruby /a /u
POSIX\d\w\s

#5 - 12/30/2011 11:03 PM - k_takata (Ken Takata)

Yui NARUSE wrote:

k-tanaka

Ruby

2.0 2.0

Ruby /a /u
POSIX\d\w\s

regparse.c OnigSyntaxRuby ONIG_OPTION_ASCII_RANGE

/u ASCII
/u/a

#6 - 12/30/2011 11:39 PM - k_takata (Ken Takata)

Onigmo 5.11.4 case-insensitive BM
/a POSIX ASCII

#7 - 12/31/2011 11:48 PM - k_takata (Ken Takata)

- File 0001-add-ONIG_SYN_POSIX_BRACKET_ALWAYS_ALL_RANGE-option.patch added
- File 0002-bug-ss-i-doesn-t-match-x-DF.patch added

POSIX ASCII
Onigmo

#8 - 01/06/2012 09:31 PM - k_takata (Ken Takata)

- File 0001-add-ONIG_SYN_POSIX_BRACKET_ALWAYS_ALL_RANGE-option.patch added
- File 0002-bug-ss-i-doesn-t-match-x-DF.patch added

Onigmo master
tmp/ruby-2.0.x
tmp/ruby-2.0.x

#9 - 01/07/2012 10:40 PM - naruse (Yui NARUSE)

Bug #5208

^u3042\b[a]/=~"u3042a" #=> nil

8. sjis/euc_jp
9. sjis/euc_jp

8,9
tool/enc-unicode.rb (tool/CaseFolding.py)

#16 - 01/20/2012 05:02 PM - naruse (Yui NARUSE)

Ken Takata wrote:

(?au) Perl Python /u /A/U

/A /U Perl /A /U 2.0 /u deprecated

/d 1.9 /a (/d) /a 1.9 \b

/d 1.9

#note-1 enc tool/enc-unicode.rb (tool/CaseFolding.py)

https://github.com/naruse/ruby/tree/onigmo r34236

- enc/cp932.c sjis.c shift_jis.c
enc/cp932.c enc/windows_31j.c rename
Windows-31J enc/shift_jis.c enc/windows_31j.c
enc/koi8.c (KOI8)
enc/cp1251.c (Ruby enc/windows_1251.c)

8,9

JIS X 0213 Ruby Shift_JISX0213/EUC-JISX0213 0213 SJS JIS X 0212 EUC-JP eucJP-ms JIS X 0212

#17 - 01/21/2012 11:52 AM - k_takata (Ken Takata)

Onigmo 5.13.0 ONIG_SYNTAX_RUBY /d 1.9 tmp/ruby-2.0.x r34236 master r34236 enc/shift_jis.c enc/windows_31j.c

JIS X 0213 Ruby Onigmo

EUC-JP eucJP-ms JIS X 0212

#18 - 01/22/2012 11:13 PM - naruse (Yui NARUSE)

Ken Takata wrote:

Onigmo 5.13.0 ONIG_SYNTAX_RUBY /d 1.9 tmp/ruby-2.0.x r34236 master r34236

r34236

2

#19 - 02/13/2012 11:33 PM - mame (Yusuke Endoh)

or onigmo k-takata

#1200 #1201

--

Yusuke Endoh mame@tsg.ne.jp

#20 - 02/14/2012 08:15 AM - naruse (Yui NARUSE)

Yusuke Endoh wrote:

or onigmo

1.9.3

k-takata

#1200 #1201

#1200 Ruby #1201 named capture

#21 - 02/17/2012 04:42 PM - naruse (Yui NARUSE)

- Status changed from Assigned to Closed

- % Done changed from 0 to 100

This issue was solved with changeset r34663.

Yui, thank you for reporting this issue.

Your contribution to Ruby is greatly appreciated.

May Ruby be with you.

- Merge Onigmo-5.13.1. [ruby-dev:45057] [Feature #5820] https://github.com/k-takata/Onigmo cp reg{comp,enc,error,exec,parse,syntax}.c reg{enc,int,parse}.h cp oniguruma.h cp tool/enc-unicode.rb cp -r enc/

Files

Table with 4 columns: File Name, Size, Date, Author. Rows include patch files for ONIG_SYN_POSIX_BRACKET_ALWAYS_ALL_RANGE and bug fixes for ss-i-doesn-t-match-x-DF.