

Ruby master - Feature #6767

Utility method to get a duplicated string whose encoding is ASCII-8BIT

07/22/2012 03:20 AM - naruse (Yui NARUSE)

Status:	Closed
Priority:	Normal
Assignee:	naruse (Yui NARUSE)
Target version:	2.0.0
Description	
<pre>String force_encoding #6361 force_encoding (1) (2) dup.force_encoding(Encoding::ASCII_8BIT) (3) ASCII-8BIT force_encoding (1) (3) (2)</pre>	
<pre>diff --git a/string.c b/string.c index d038835..76cbc36 100644 --- a/string.c +++ b/string.c @@ -601,7 +601,7 @@ rb_str_export_to_enc(VALUE str, rb_encoding *enc) } static VALUE -str_replace_shared(VALUE str2, VALUE str) +str_replace_shared_without_enc(VALUE str2, VALUE str) { if (RSTRING_LEN(str) <= RSTRING_EMBED_LEN_MAX) { STR_SET_EMBED(str2); @@ -616,8 +616,14 @@ str_replace_shared(VALUE str2, VALUE str) RSTRING(str2)->as.heap.aux.shared = str; FL_SET(str2, ELTS_SHARED); } • rb_enc_cr_str_exact_copy(str2, str); • return str2; +} +static VALUE +str_replace_shared(VALUE str2, VALUE str) +{ • str_replace_shared_without_enc(str2, str); • rb_enc_cr_str_exact_copy(str2, str); return str2; } @@ -7340,6 +7346,23 @@ rb_str_force_encoding(VALUE str, VALUE enc) /* • call-seq: ○ * str.b -> str ○ * ○ * Returns a copied string whose encoding is ASCII-8BIT. ○ */ +static VALUE +rb_str_b(VALUE str) +{ ○ VALUE str2 = str_alloc(rb_cString);</pre>	

- str_replace_shared_without_enc(str2, str);
- OBJ_INFECT(str2, str);
- ENC_CODERANGE_SET(str2, ENC_CODERANGE_VALID);
- return str2; +} + +/*
- * call-seq:
- str.valid_encoding? -> true or false *

- Returns true for a string which encoded correctly.
- @@ -7969,6 +7992,7 @@ Init_String(void)

```
rb_define_method(rb_cString, "encoding", rb_obj_encoding, 0); /* in encoding.c */
rb_define_method(rb_cString, "force_encoding", rb_str_force_encoding, 1);
```

- rb_define_method(rb_cString, "b", rb_str_b, 0); rb_define_method(rb_cString, "valid_encoding?", rb_str_valid_encoding_p, 0); rb_define_method(rb_cString, "ascii_only?", rb_str_is_ascii_only_p, 0);

```
diff --git a/test/ruby/test_m17n.rb b/test/ruby/test_m17n.rb
index dfcaa94..3a4bca7 100644
--- a/test/ruby/test_m17n.rb
+++ b/test/ruby/test_m17n.rb
@@ -1469,4 +1469,14 @@ class TestM17N < Test::Unit::TestCase
yield(*strs)
end
end
+
```

- def test_str_b
- s = "\u3042"
- assert_equal(a("\xE3\x81\x82"), s.b)
- assert_equal(Encoding::ASCII_8BIT, s.b.encoding)
- s.taint
- assert_equal(true, s.b.tainted?)
- s.untrust
- assert_equal(true, s.b.untrusted?)
- end end

Related issues:

Related to Ruby master - Feature #6361: Bitwise string operations

Rejected

04/26/2012

Associated revisions

Revision a6ec2584 - 11/06/2012 12:50 AM - naruse (Yui NARUSE)

- string.c (rb_str_b): Add String#b, returning a copied string whose encoding is ASCII-8BIT. [ruby-dev:45992] [Feature #6767]

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/trunk@37486 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 37486 - 11/06/2012 12:50 AM - naruse (Yui NARUSE)

- string.c (rb_str_b): Add String#b, returning a copied string whose encoding is ASCII-8BIT. [ruby-dev:45992] [Feature #6767]

Revision 37486 - 11/06/2012 12:50 AM - naruse (Yui NARUSE)

- string.c (rb_str_b): Add String#b, returning a copied string whose encoding is ASCII-8BIT. [ruby-dev:45992] [Feature #6767]

Revision 37486 - 11/06/2012 12:50 AM - naruse (Yui NARUSE)

- string.c (rb_str_b): Add String#b, returning a copied string whose encoding is ASCII-8BIT. [ruby-dev:45992] [Feature #6767]

Revision 37486 - 11/06/2012 12:50 AM - naruse (Yui NARUSE)

- string.c (rb_str_b): Add String#b, returning a copied string whose encoding is ASCII-8BIT. [ruby-dev:45992] [Feature #6767]

Revision 37486 - 11/06/2012 12:50 AM - naruse (Yui NARUSE)

- string.c (rb_str_b): Add String#b, returning a copied string whose encoding is ASCII-8BIT. [ruby-dev:45992] [Feature #6767]

Revision 37486 - 11/06/2012 12:50 AM - naruse (Yui NARUSE)

- string.c (rb_str_b): Add String#b, returning a copied string whose encoding is ASCII-8BIT. [ruby-dev:45992] [Feature #6767]

History

#1 - 07/22/2012 11:03 AM - duerst (Martin Dürst)

naruse (Yui NARUSE) wrote:

```

String  ASCII-8BIT
force_encoding
#6361
force_encoding

```

```

(1)

```

encoding: UTF-8

\xAB String binary UTF-8

```
(2) dup.force_encoding(Encoding::ASCII_8BIT)
```

force_encoding force_encoding!
dup.force_encoding force_encoding

```
(3) ASCII-8BIT force_encoding
```

```
(1) (3)
(2)
```

```
String#b
```

b binary

binary

#2 - 09/05/2012 11:36 AM - naruse (Yui NARUSE)

duerst (Martin Dürst) wrote:

naruse (Yui NARUSE) wrote:

```
String#b
```

b binary

binary

binary binary
.b
String
.irbrc

#3 - 09/05/2012 05:52 PM - duerst (Martin Dürst)

naruse (Yui NARUSE) wrote:

```
000000000000000000000000000000000000 .b 000000000000
```

```
000000000000000000000000000000000000000000000000000000000000000000000000
```

```
000000000000000000000000000000000000000000000000000000000000000000000000.b 0000000000000000 (000000000000000000000000000000000000000000000000000000000000000000000000) 000000000000
```

#4 - 09/05/2012 06:59 PM - naruse (Yui NARUSE)

duerst (Martin Dürst) wrote:

naruse (Yui NARUSE) wrote:

```
000000000000000000000000000000000000 .b 000000000000
```

```
000000000000000000000000000000000000000000000000000000000000000000000000
```

```
000000000000000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000000000000000
```

```
000000000000000000000000000000000000000000000000000000000000000000000000.b 0000000000000000 (000000000000000000000000000000000000000000000000000000000000000000000000) 000000000000
```

```
000000000000000000000000000000000000000000000000000000000000000000000000 Rails000000000000 Rails00000 .b000000000000000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000000000000000 .b000000000000000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000000000000000 force_encoding("UTF-8")000000000000000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000000000000000
```

#5 - 11/05/2012 04:24 PM - matz (Yukihiko Matsumoto)

- Assignee changed from matz (Yukihiko Matsumoto) to naruse (Yui NARUSE)

OK to check in.

Matz.

#6 - 11/06/2012 09:50 AM - naruse (Yui NARUSE)

- Status changed from Assigned to Closed

- % Done changed from 0 to 100

This issue was solved with changeset r37486.
Yui, thank you for reporting this issue.
Your contribution to Ruby is greatly appreciated.
May Ruby be with you.

-
- `string.c (rb_str_b):` Add `String#b`, returning a copied string whose encoding is ASCII-8BIT. [ruby-dev:45992] [Feature [#6767](#)]