以下のコードで `String#gsub` が遅くなります。

- `b = ""`の場合 (A): 0.2840230464935303
- `b = "xB9"`の場合 (B): 4.183771848678589

```
=begin
以下のようなコードで String#gsub が遅くなります。

```ruby
b = "" の場合(A): 0.2840230464935303
b = "xB9" の場合(B): 4.183771848678589

```

```
# coding: utf-8
```

```
a = ("abcde\n"**50000).force_encoding("binary")
b = ""

b = "xB9".force_encoding("binary")
c = ("efghi\n"**50000).force_encoding("binary")
d = "#{a}#{b}#{c}"
```

```
start = Time.now.to_f

start = Time.now.to_f

```

それぞれの場合で、プロファイルを取ってみたので添付します。

(B)の場合に、search_nonascii を約20万回呼び出して処理時間の92%を費しています。

A)の場合には10万回しか呼び出しておらず、処理時間も短いです。

=end

```
Revision 7671e7fdd - 06/02/2012 08:02 AM - naruse (Yui NARUSE)

- string.c (rb_enc_cr_str_buf_cat): don't reset coderange as unknown. the condition 'ptr_a8 && str_cr != ENC_CODERANGE_7BIT' means not unknown, str is also ASCII-8BIT because str_encindex == ptr_encindex, and nont (str_cr == ENC_CODERANGE_UNKNOWN) and str_cr != ENC_CODERANGE_7BIT means str_cr is valid because ASCII-8BIT can't be broken. [ruby-dev:45688] [Bug #6509]

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/trunk@35863 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 35863 - 06/02/2012 08:02 AM - naruse (Yui NARUSE)

- string.c (rb_enc_cr_str_buf_cat): don't reset coderange as unknown. the condition 'ptr_a8 && str_cr != ENC_CODERANGE_7BIT' means not unknown, str is also ASCII-8BIT because str_encindex == ptr_encindex, and nont (str_cr == ENC_CODERANGE_UNKNOWN) and str_cr != ENC_CODERANGE_7BIT means str_cr is valid because ASCII-8BIT can't be broken. [ruby-dev:45688] [Bug #6509]

Revision 35863 - 06/02/2012 08:02 AM - naruse (Yui NARUSE)

- string.c (rb_enc_cr_str_buf_cat): don't reset coderange as unknown. the condition 'ptr_a8 && str_cr != ENC_CODERANGE_7BIT' means not unknown, str is also ASCII-8BIT because str_encindex == ptr_encindex, and nont (str_cr == ENC_CODERANGE_UNKNOWN) and str_cr != ENC_CODERANGE_7BIT means str_cr is valid because ASCII-8BIT can't be broken. [ruby-dev:45688] [Bug #6509]

Revision 35863 - 06/02/2012 08:02 AM - naruse (Yui NARUSE)

- string.c (rb_enc_cr_str_buf_cat): don't reset coderange as unknown. the condition 'ptr_a8 && str_cr != ENC_CODERANGE_7BIT' means not unknown, str is also ASCII-8BIT because str_encindex == ptr_encindex, and nont (str_cr == ENC_CODERANGE_UNKNOWN) and str_cr != ENC_CODERANGE_7BIT means str_cr is valid because ASCII-8BIT can't be broken. [ruby-dev:45688] [Bug #6509]

Revision 35863 - 06/02/2012 08:02 AM - naruse (Yui NARUSE)
### Files

<table>
<thead>
<tr>
<th>File</th>
<th>Size</th>
<th>Date</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>callgrind.out.9937</td>
<td>521 KB</td>
<td>05/29/2012</td>
<td>okkez (okkez _)</td>
</tr>
<tr>
<td>callgrind.out.10091</td>
<td>521 KB</td>
<td>05/29/2012</td>
<td>okkez (okkez _)</td>
</tr>
</tbody>
</table>