

Ruby trunk - Bug #14095

erb ignores attempt to set encoding

11/09/2017 07:29 PM - graywolf (Gray Wolf)

Status: Closed	
Priority: Normal	
Assignee: k0kubun (Takashi Kokubun)	
Target version:	
ruby -v: ruby 2.4.2p198 (2017-09-14 revision 59899) [x86_64-linux]	Backport: 2.3: UNKNOWN, 2.4: UNKNOWN
Description Erb seems to ignore any attempt to set encoding: <pre>\$ erb -U <%= "a".encoding %> ^D ASCII-8BIT</pre> <p>I've tried multiple ways to do this but I just can't convince it that the .erb file is in UTF-8. It insists on ASCII-8BIT.</p>	

Associated revisions

Revision e16bd0f4 - 11/10/2017 01:59 PM - k0kubun (Takashi Kokubun)

bin/erb: change template file encoding to UTF-8

Unlike Ruby source file encoding (script encoding) whose default is changed to UTF-8 in Ruby 2.0 (Feature #6679), template's file encoding given to erb(1) has been ASCII-8BIT since ERB supports m17n at r21170.

Like Ruby source file encoding, erb template file encoding should be UTF-8 in Ruby 2.

[Bug #14095] [ruby-core:83708]

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/trunk@60739 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 60739 - 11/10/2017 01:59 PM - k0kubun (Takashi Kokubun)

bin/erb: change template file encoding to UTF-8

Unlike Ruby source file encoding (script encoding) whose default is changed to UTF-8 in Ruby 2.0 (Feature #6679), template's file encoding given to erb(1) has been ASCII-8BIT since ERB supports m17n at r21170.

Like Ruby source file encoding, erb template file encoding should be UTF-8 in Ruby 2.

[Bug #14095] [ruby-core:83708]

Revision 60739 - 11/10/2017 01:59 PM - k0kubun (Takashi Kokubun)

bin/erb: change template file encoding to UTF-8

Unlike Ruby source file encoding (script encoding) whose default is changed to UTF-8 in Ruby 2.0 (Feature #6679), template's file encoding given to erb(1) has been ASCII-8BIT since ERB supports m17n at r21170.

Like Ruby source file encoding, erb template file encoding should be UTF-8 in Ruby 2.

[Bug #14095] [ruby-core:83708]

Revision 60739 - 11/10/2017 01:59 PM - k0kubun (Takashi Kokubun)

bin/erb: change template file encoding to UTF-8

Unlike Ruby source file encoding (script encoding) whose default is

changed to UTF-8 in Ruby 2.0 (Feature #6679), template's file encoding given to erb(1) has been ASCII-8BIT since ERB supports m17n at r21170.

Like Ruby source file encoding, erb template file encoding should be UTF-8 in Ruby 2.

[Bug #14095] [ruby-core:83708]

History

#1 - 11/09/2017 07:29 PM - graywolf (Gray Wolf)

Expected output was

```
$ erb -U
<%= "a".encoding %>
^D
UTF-8
```

Ruby works that way without setting anything

```
$ ruby
puts "a".encoding
^D
UTF-8
```

Why doesn't .erb have the same default?

#2 - 11/10/2017 02:49 AM - hsbt (Hiroshi SHIBATA)

- Assignee set to k0kubun (Takashi Kokubun)

#3 - 11/10/2017 02:50 AM - hsbt (Hiroshi SHIBATA)

- Status changed from Open to Assigned

#4 - 11/10/2017 01:59 PM - k0kubun (Takashi Kokubun)

- Status changed from Assigned to Closed

Applied in changeset [trunk|r60739](#).

bin/erb: change template file encoding to UTF-8

Unlike Ruby source file encoding (script encoding) whose default is changed to UTF-8 in Ruby 2.0 (Feature #6679), template's file encoding given to erb(1) has been ASCII-8BIT since ERB supports m17n at r21170.

Like Ruby source file encoding, erb template file encoding should be UTF-8 in Ruby 2.

[Bug #14095] [ruby-core:83708]

#5 - 11/10/2017 02:05 PM - k0kubun (Takashi Kokubun)

Hi Gray,

First of all, -U option of erb(1) is just the same as ruby(1)'s one, which sets both external and internal encoding to UTF-8. As those encodings are for IO, -U (or -E) is not related to the described behavior.

The actual cause that matters is source file encoding (script encoding, which affects encoding of string literal in the file) counterpart for erb(1) command. As m17n of erb(1) is implemented in Ruby 1.9, it's the same as Ruby 1.9's default source file encoding: ASCII-8BIT. But Ruby 2's default source file encoding is changed to UTF-8, I thought erb(1)'s one should be changed in the same way. Thus I fixed it in r60739.

#6 - 11/10/2017 02:18 PM - k0kubun (Takashi Kokubun)

Ah, one more comment is that you can use magic comment to change source file encoding of erb template before and after that change:

```
$ ruby -v
ruby 2.4.2p198 (2017-09-14 revision 59899) [x86_64-linux]
$ cat ascii.erb
<%= ''.encoding %>
$ erb -T 1 ascii.erb
ASCII-8BIT
```

```
$ cat utf-8.erb
<%# coding: UTF-8 %>
<%= ''.encoding %>
$ erb -T 1 utf-8.erb
UTF-8
```

#7 - 11/10/2017 03:46 PM - graywolf (Gray Wolf)

Thank you for quick fix and your answer, I've tried the magic comment thing base on what I googled, to be precise `<% # -* coding: UTF-8 -* %>` which didn't work. Your version `<%# coding: UTF-8 %>` does work so I can now use `?` without getting parser error. Yeeey!