

Ruby master - Bug #16970

Encoding of ENV value returns ASCII-8BIT in Ruby2.6 or later

06/19/2020 03:18 AM - masuyama (Soichi Masuyama)

Status:	Open		
Priority:	Normal		
Assignee:			
Target version:			
ruby -v:	ruby 2.6.5p114 [i386-mingw32], ruby 2.7.1p83 [x64-mingw32]	Backport:	2.5: UNKNOWN, 2.6: UNKNOWN, 2.7: UNKNOWN

Description

Problem Report

When internal encoding is set to UTF-8, encoding of ENV value always return ASCII-8BIT. (UTF-8 is expected).

Result of reproduce process

Ruby 2.5

```
set TEST=""
ruby --encoding=UTF-8:UTF-8 -e "p 'test'.encoding"      #=> #<Encoding:UTF-8>
ruby --encoding=UTF-8:UTF-8 -e "p ENV['TEST'].encoding"  #=> #<Encoding:UTF-8>
```

Ruby 2.6

```
set TEST=""
ruby --encoding=UTF-8:UTF-8 -e "p 'test'.encoding"      #=> #<Encoding:UTF-8>
ruby --encoding=UTF-8:UTF-8 -e "p ENV['TEST'].encoding"  #=> #<Encoding:ASCII-8BIT> **INVALID RESULT** Expected Result is UTF-8.
```

```
ENV['TEST'] =~ /[] (nonascii) /
# => incompatible encoding regexp match (UTF-8 regexp with ASCII-8BIT string)
```

Regression occurred in the following Ticket

fallback env encoding to ASCII-8BIT

<https://github.com/ruby/ruby/commit/7f0d337be73bb2465b40009fe23f3b7be6b0dc90>

Reason of bug

rb_str_cat_conv_enc_opts return Qnil when from/to encoding is the same (UTF-8) and fallback code introduced by the following commit set the encoding to ASCII-8BIT.

<https://github.com/ruby/ruby/commit/7f0d337be73bb2465b40009fe23f3b7be6b0dc90>

Fixes

When internal encoding is UTF-8, return the string as is since there is no need to call encoding conversion.

<https://github.com/ruby/ruby/pull/3239>

Related issues:

Related to Ruby master - Bug #16623: Windows ENV encoding

Open

History

#1 - 06/19/2020 03:26 AM - masuyama (Soichi Masuyama)

- Description updated

#2 - 06/19/2020 03:30 AM - masuyama (Soichi Masuyama)

- Description updated

#3 - 06/19/2020 06:49 AM - masuyama (Soichi Masuyama)

- *Description updated*

#4 - 08/28/2020 10:37 PM - jeremyevans0 (Jeremy Evans)

- *Related to Bug #16623: Windows ENV encoding added*