Ruby master - Feature #16233
winruby UTF8 Fallback for no CodePage
10/03/2019 07:29 AM - gabuscus (Gabriel Nagy)

<table>
<thead>
<tr>
<th>Status:</th>
<th>Open</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priority:</td>
<td>Normal</td>
</tr>
<tr>
<td>Assignee:</td>
<td>duerst (Martin Dürst)</td>
</tr>
<tr>
<td>Target version:</td>
<td>2.8</td>
</tr>
</tbody>
</table>

**Description**

This is to deal specifically with setting the Region to Arabic which sets the codepage to 720. This codepage isn't recognised by Ruby which causes it to fail during initialisation when setting up LOCALE.

This patch to Ruby is necessary to handle languages/regions on windows where the codepage is not supported by Ruby such as Arabic which uses codepage 720.

If the codepage is not found, the Locale falls back to UTF8.

This is a well known Ruby/Ruby on Rails issue which is described at https://stackoverflow.com/questions/22815542/rails4-unknown-encoding-name-cp720

Opened PR: https://github.com/ruby/ruby/pull/2518

**History**

#1 - 10/03/2019 07:46 AM - gabuscus (Gabriel Nagy)
- Description updated

#2 - 10/04/2019 04:44 AM - duerst (Martin Dürst)
If codepage 720 is not supported, then we should look at supporting it.

#3 - 10/04/2019 11:06 AM - gabuscus (Gabriel Nagy)
duerst (Martin Dürst) wrote:

> If codepage 720 is not supported, then we should look at supporting it.

Fair point, however investigation on this Puppet ticket: https://tickets.puppetlabs.com/browse/PA-2191 concluded that falling back to UTF-8 fixes the issue (for codepage 720 at least), we could also check the codepage when catching the ArgumentError so we fallback only in that case

Edit: Also I'm not sure what does adding a new codepage entail? I agree that would be the cleaner solution

#4 - 12/29/2019 06:45 AM - hsbt (HirosHi SHIBATA)
- Tags set to patch, win, encoding

#5 - 01/14/2020 09:17 AM - naruse (Yui NARUSE)
- Backport deleted (2.5: UNKNOWN, 2.6: UNKNOWN)
- ruby -v deleted (2.6.3)
- Target version set to 2.8
- Assignee set to duerst (Martin Dürst)
- Tracker changed from Bug to Feature

It looks Ruby should support CP720.

duerst (Martin Dürst) Could you add support for it?

#6 - 01/14/2020 10:11 AM - duerst (Martin Dürst)
Will do, but probably not very soon, because it's now the most busy time of the year at work, sorry.

02/21/2020
Some notes: For the encoding (onigX), windows-1256 is a replica of iso-8859-6. But windows-1256 is much more complete, while iso-8859-6 has many unassigned codepoints, so it may be a good idea to fix this if we get around to it.

Just adding an alias 'CP720' for windows-1256 is much easier.

Files

| windows_nocodepage_utf8_fallback_r2.5.patch | 946 Bytes | 10/03/2019 | gabuscus (Gabriel Nagy) |