

Backport200 - Backport #8323

no conversion by "bom|utf-8"

04/25/2013 11:42 AM - nobu (Nobuyoshi Nakada)

Status:	Closed
Priority:	Normal
Assignee:	nagachika (Tomoyuki Chikanaga)
Description	
Mode spec in open and etc. accepts BOM-prefix UTF encoding, however if the external and internal encodings given there are same no conversion take place regardless the actual external encoding. Since the encoding prefixed with "BOM" is not a real encoding, but just a fallback, the conversion should honor the detected encoding by BOM.	

Associated revisions

Revision 91fb5bc8 - 04/25/2013 07:11 AM - nobu (Nobuyoshi Nakada)

io.c: conversion from bom encoding

- io.c (rb_io_ext_int_to_encs, parse_mode_enc): bom-prefixed name is not a real encoding name, just a fallback. so the proper conversion should take place even if if the internal encoding is equal to the bom-prefixed name, unless actual encoding is equal to the internal encoding. [ruby-core:54563] [Bug #8323]
- io.c (io_set_encoding_by_bom): reset extenal encoding if no BOM found. [ruby-core:54569]

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/trunk@40462 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 40462 - 04/25/2013 07:11 AM - nobu (Nobuyoshi Nakada)

io.c: conversion from bom encoding

- io.c (rb_io_ext_int_to_encs, parse_mode_enc): bom-prefixed name is not a real encoding name, just a fallback. so the proper conversion should take place even if if the internal encoding is equal to the bom-prefixed name, unless actual encoding is equal to the internal encoding. [ruby-core:54563] [Bug #8323]
- io.c (io_set_encoding_by_bom): reset extenal encoding if no BOM found. [ruby-core:54569]

Revision 40462 - 04/25/2013 07:11 AM - nobu (Nobuyoshi Nakada)

io.c: conversion from bom encoding

- io.c (rb_io_ext_int_to_encs, parse_mode_enc): bom-prefixed name is not a real encoding name, just a fallback. so the proper conversion should take place even if if the internal encoding is equal to the bom-prefixed name, unless actual encoding is equal to the internal encoding. [ruby-core:54563] [Bug #8323]
- io.c (io_set_encoding_by_bom): reset extenal encoding if no BOM found. [ruby-core:54569]

Revision 40462 - 04/25/2013 07:11 AM - nobu (Nobuyoshi Nakada)

io.c: conversion from bom encoding

- io.c (rb_io_ext_int_to_encs, parse_mode_enc): bom-prefixed name is not a real encoding name, just a fallback. so the proper conversion should take place even if if the internal encoding is equal to the bom-prefixed name, unless actual encoding is equal to the internal encoding. [ruby-core:54563] [Bug #8323]
- io.c (io_set_encoding_by_bom): reset extenal encoding if no BOM found. [ruby-core:54569]

Revision 40462 - 04/25/2013 07:11 AM - nobu (Nobuyoshi Nakada)

io.c: conversion from bom encoding

- io.c (rb_io_ext_int_to_encs, parse_mode_enc): bom-prefixed name is not a real encoding name, just a fallback. so the proper conversion should take place even if if the internal encoding is equal to the bom-prefixed name, unless actual encoding is equal to the internal encoding. [ruby-core:54563] [Bug #8323]
- io.c (io_set_encoding_by_bom): reset extenal encoding if no BOM found. [ruby-core:54569]

Revision 40462 - 04/25/2013 07:11 AM - nobu (Nobuyoshi Nakada)

io.c: conversion from bom encoding

- io.c (rb_io_ext_int_to_encs, parse_mode_enc): bom-prefixed name is not a real encoding name, just a fallback. so the proper conversion should take place even if if the internal encoding is equal to the bom-prefixed name, unless actual encoding is equal to the internal encoding. [ruby-core:54563] [Bug #8323]
- io.c (io_set_encoding_by_bom): reset extenal encoding if no BOM found. [ruby-core:54569]

Revision 40462 - 04/25/2013 07:11 AM - nobu (Nobuyoshi Nakada)

io.c: conversion from bom encoding

- io.c (rb_io_ext_int_to_encs, parse_mode_enc): bom-prefixed name is not a real encoding name, just a fallback. so the proper conversion should take place even if the internal encoding is equal to the bom-prefixed name, unless actual encoding is equal to the internal encoding. [ruby-core:54563] [Bug #8323]
- io.c (io_set_encoding_by_bom): reset external encoding if no BOM found. [ruby-core:54569]

Revision 4d3df433 - 05/01/2013 02:52 PM - nagachika (Tomoyuki Chikanaga)

merge revision(s) 40462: [Backport #8323]

```
* io.c (rb_io_ext_int_to_encs, parse_mode_enc): bom-prefixed name is
not a real encoding name, just a fallback. so the proper conversion
should take place even if if the internal encoding is equal to the
bom-prefixed name, unless actual encoding is equal to the internal
encoding. [ruby-core:54563] [Bug #8323]
```

```
* io.c (io_set_encoding_by_bom): reset external encoding if no BOM
found. [ruby-core:54569]
```

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/branches/ruby_2_0_0@40541 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 40541 - 05/01/2013 02:52 PM - nagachika (Tomoyuki Chikanaga)

merge revision(s) 40462: [Backport #8323]

```
* io.c (rb_io_ext_int_to_encs, parse_mode_enc): bom-prefixed name is
not a real encoding name, just a fallback. so the proper conversion
should take place even if if the internal encoding is equal to the
bom-prefixed name, unless actual encoding is equal to the internal
encoding. [ruby-core:54563] [Bug #8323]
```

```
* io.c (io_set_encoding_by_bom): reset external encoding if no BOM
found. [ruby-core:54569]
```

History

#1 - 04/25/2013 11:49 AM - nobu (Nobuyoshi Nakada)

- Description updated

Sorry, miss-post to ruby-core.

#2 - 04/25/2013 11:56 AM - nobu (Nobuyoshi Nakada)

- File 0001-io.c-conversion-from-bom-encoding.patch added

A patch attached

#3 - 04/25/2013 12:58 PM - naruse (Yui NARUSE)

The patch doesn't work on following case:

```
% ./ruby -e'IO.write "p", "a"; open("p", "r:BOM|utf-8:utf-8"){|f| p f.read.size}'
-e:1:in read': code converter not found (UTF-8 to UTF-8) (Encoding::ConverterNotFoundError)
from -e:1:inblock in '
from -e:1:in open'
from -e:1:in'
```

#4 - 04/25/2013 01:42 PM - nobu (Nobuyoshi Nakada)

- File 0001-io.c-conversion-from-bom-encoding.patch added

Updated.

#5 - 04/25/2013 01:43 PM - nobu (Nobuyoshi Nakada)

- File deleted (0001-io.c-conversion-from-bom-encoding.patch)

#6 - 04/25/2013 03:48 PM - naruse (Yui NARUSE)

nobu (Nobuyoshi Nakada) wrote:

Updated.

OK, commit please.

#7 - 04/25/2013 04:11 PM - nobu (Nobuyoshi Nakada)

- Status changed from Open to Closed
- % Done changed from 0 to 100

This issue was solved with changeset r40462.
Nobuyoshi, thank you for reporting this issue.
Your contribution to Ruby is greatly appreciated.
May Ruby be with you.

io.c: conversion from bom encoding

- io.c (rb_io_ext_int_to_encs, parse_mode_enc): bom-prefixed name is not a real encoding name, just a fallback. so the proper conversion should take place even if the internal encoding is equal to the bom-prefixed name, unless actual encoding is equal to the internal encoding. [ruby-core:54563] [Bug #8323]
- io.c (io_set_encoding_by_bom): reset external encoding if no BOM found. [ruby-core:54569]

#8 - 04/25/2013 05:17 PM - nobu (Nobuyoshi Nakada)

- Backport changed from 1.9.3: UNKNOWN, 2.0.0: UNKNOWN to 1.9.3: REQUIRED, 2.0.0: REQUIRED

#9 - 04/25/2013 05:18 PM - nobu (Nobuyoshi Nakada)

- Tracker changed from Bug to Backport
- Project changed from Ruby master to Backport200
- Category deleted (M17N)
- Status changed from Closed to Assigned
- Assignee changed from naruse (Yui NARUSE) to nagachika (Tomoyuki Chikanaga)
- Target version deleted (2.1.0)

#10 - 05/01/2013 11:52 PM - nagachika (Tomoyuki Chikanaga)

- Status changed from Assigned to Closed

This issue was solved with changeset [r40541](#).
Nobuyoshi, thank you for reporting this issue.
Your contribution to Ruby is greatly appreciated.
May Ruby be with you.

merge revision(s) 40462: [Backport [#8323](#)]

- ```
* io.c (rb_io_ext_int_to_encs, parse_mode_enc): bom-prefixed name is
not a real encoding name, just a fallback. so the proper conversion
should take place even if if the internal encoding is equal to the
bom-prefixed name, unless actual encoding is equal to the internal
encoding. [ruby-core:54563] [Bug #8323]

* io.c (io_set_encoding_by_bom): reset external encoding if no BOM
found. [ruby-core:54569]
```

**Files**

---

|                                              |         |            |                         |
|----------------------------------------------|---------|------------|-------------------------|
| 0001-io.c-conversion-from-bom-encoding.patch | 8.76 KB | 04/25/2013 | nobu (Nobuyoshi Nakada) |
|----------------------------------------------|---------|------------|-------------------------|