

## Redmine - Defect #11834

### Bazaar: "???" instead of non ASCII character in paths on non UTF-8 locale

2012-09-14 17:15 - Alexander Usenko

|  |                |                          |           |
|--|----------------|--------------------------|-----------|
| <b>Status:</b>   | Closed         | <b>Start date:</b>       |           |
| <b>Priority:</b>   | Normal         | <b>Due date:</b>         |           |
| <b>Assignee:</b>   | Toshi MARUYAMA | <b>% Done:</b>           | 100%      |
| <b>Category:</b>   | SCM            | <b>Estimated time:</b>   | 0.00 hour |
| <b>Target version:</b>   | 2.2.0          | <b>Affected version:</b> |           |
| <b>Resolution:</b>   | Fixed          |                          |           |
| <b>Description</b>   |                |                          |           |
| BitNami Redmine Stack: <ul style="list-style-type: none"><li>• Redmine 2.0.3</li><li>• Ruby 1.8.7</li><li>• Rails 3.2.6</li><li>• Apache 2.2</li></ul> <p>I'm using Redmine on Windows XP SP3 Rus and Windows Server 2003 Rus (cp1251). There is a problem with browsing Bazaar repository containing cyrillic symbols in repository filenames and paths. I see ??? instead of cyrillic symbols.</p> <p>I didn't find solution of this problem anywhere. Is there one?</p> <p>So, I checked other scm *_adapter.rb and made same 'bzip' output encoding conversion, so everything works now (see attached files, they contains comments of basic changes). Was it useless?</p> <p>Changes:</p> <ul style="list-style-type: none"><li>• Added path encoding setup when adding Bazaar repo (same as for the filesystem).</li><li>• Removed log encoding setup, because of Bazaar is storing all comments in utf8 inside and outputs them with same encoding as for the path.</li></ul> <p>P.S. Saw Ruby for the first time when started to do this</p> |                |                          |           |

#### Associated revisions

##### Revision 10414 - 2012-09-18 00:49 - Toshi MARUYAMA

scm: bazaar: update test repository (#11834)

Add new "non\_ascii" branch.

It was created by "bzip branch --no-tree branch00 non\_ascii".

It is equivalent to Git bare repository and Mercurial "hg update null".

It has non ASCII path name files and directory.

##### Revision 10415 - 2012-09-18 01:45 - Toshi MARUYAMA

scm: bazaar: add @path\_encoding instance value and set it UTF-8 at adapter initialize method (#11834)

##### Revision 10416 - 2012-09-18 01:45 - Toshi MARUYAMA

scm: bazaar: convert command line character encoding to @path\_encoding (#11834)

##### Revision 10417 - 2012-09-18 01:46 - Toshi MARUYAMA

scm: bazaar: convert path encoding from @path\_encoding to UTF-8 at adapter revisions method (#11834)

##### Revision 10418 - 2012-09-18 01:46 - Toshi MARUYAMA

scm: bazaar: convert path encoding from @path\_encoding to UTF-8 at adapter entries method (#11834)

##### Revision 10420 - 2012-09-18 11:30 - Toshi MARUYAMA

scm: bazaar: use log encoding as path encoding (#11834)

Revision 10421 - 2012-09-18 11:30 - Toshi MARUYAMA

scm: bazaar: add non ASCII path test to unit model test (#11834)

Revision 10422 - 2012-09-18 12:10 - Toshi MARUYAMA

scm: bazaar: do not call super initialize method at adapter (#11834)

On Bazaar 2.0.5, in case of using non ASCII repository path, following trace raise on all operation.

```
Traceback (most recent call last):
  File "/usr/bin/bzr", line 142, in <module>
    exit_val = bzrlib.commands.main()
  File "/usr/lib/python2.6/site-packages/bzrlib/commands.py", line 1102, in main
    argv = osutils.get_unicode_argv()
  File "/usr/lib/python2.6/site-packages/bzrlib/osutils.py", line 90, in get_unicode_argv
    "encoding." % a))
bzrlib.errors.BzrError:
Parameter '/somewhere/\xe3\x83\x90\xe3\x82\xb6\xe3\x83\xbc/non_ascii/'
is unsupported by the current encoding.
```

Revision 10423 - 2012-09-18 12:10 - Toshi MARUYAMA

scm: bazaar: add more comment to non ASCII path tests at unit model test (#11834)

History

#1 - 2012-09-14 17:46 - Alexander Usenko

There can be a problem with file annotation view (simple file view is Ok). "Attachments and repositories encodings" in application settings must be set to "UTF-8, CP1251", not to "CP1251, UTF-8". Last case causes file to look like it is encoded in UTF-8, but is browsed in CP1251.

What is being seen: C,PμCfC, C,PμCfC, C,PμCfC,  
What should be seen: tect tect tect

Maybe i forgot smth in annotate function of bazaar\_adapter.rb

#2 - 2012-09-15 11:19 - Bruno Souza

I dont know

#3 - 2012-09-15 11:26 - Toshi MARUYAMA

- Status changed from New to 7
- Assignee set to Toshi MARUYAMA
- Affected version (unused) deleted (2.0.3)
- Affected version deleted (2.0.3)

#4 - 2012-09-18 12:41 - Toshi MARUYAMA

- Subject changed from Redmine+Bazaar+Windows: "???" instead of cyrillic symbols in repository paths to Bazaar+Windows: "???" instead of non ASCII character in paths

#5 - 2012-09-18 12:43 - Toshi MARUYAMA

- Subject changed from Bazaar+Windows: "???" instead of non ASCII character in paths to Bazaar: "???" instead of non ASCII character in paths on non UTF-8 locale

#6 - 2012-09-18 13:09 - Toshi MARUYAMA

- Target version set to 2.2.0
- % Done changed from 0 to 100

#7 - 2012-10-13 13:03 - Jean-Philippe Lang

- Status changed from 7 to Closed
- Resolution set to Fixed

Files

|                   |         |            |                  |
|-------------------|---------|------------|------------------|
| bazaar.rb         | 3.95 KB | 2012-09-14 | Alexander Usenko |
| bazaar_adapter.rb | 11.8 KB | 2012-09-14 | Alexander Usenko |

