

Ruby trunk - Bug #9367

REXML::XmlDecl doesn't use user specified quotes

01/06/2014 05:11 PM - bearmini (Takashi Oguma)

Status: Closed	
Priority: Normal	
Assignee: kou (Kouhei Sutou)	
Target version:	
ruby -v: ruby 2.0.0p247 (2013-06-27 revision 41674) [x86_64-linux]	Backport:
Description =begin REXML uses double quotes for quoting attributes if :attribute_quote is specified as document's context like below: doc = REXML::Document.new doc.context[:attribute_quote] = :quote This looks working well on all elements but has no effect for xml declaration (i.e. <?xml version= ... ?>) if it exists. Even if I specify (({doc.context[:attribute_quote] = :quote})), I will get this: <?xml version='1.0' encoding='UTF-8' standalone='true'?> The expected result is: <?xml version="1.0" encoding="UTF-8" standalone="true"?> =end	
Related issues: Has duplicate Ruby trunk - Bug #11176: Hardcoded settings in REXML Closed	

History

#1 - 01/06/2014 05:41 PM - duerst (Martin Dürst)

- Status changed from Open to Feedback

Strictly speaking, the XML declaration doesn't contain any attributes, only things that look like attributes. They are sometimes called pseudo-attributes. So it is not unreasonable that
doc.context[:attribute_quote] = :quote
does not affect pseudo-attribute quoting.

For attributes, the choice of quoting can affect the amount of escaping. But this doesn't apply to pseudo-attributes, as the 'values' in the pseudo-attributes don't contain quotes.

Can you explain the specific reason you want double quotes in the XML declaration?

#2 - 01/07/2014 11:01 AM - bearmini (Takashi Oguma)

A direct reason is that my customer wants to have double quotes in the xml declaration because their handcrafted 'xml lint' tool complains the xml document produced by my ruby script contains single quotes. (Their convention requires all quoting characters should be double quotes.)

More generally, I think it is natural if we have control which quoting character will be used for the xml declaration too.

#3 - 01/07/2014 05:40 PM - duerst (Martin Dürst)

bearmini (Takashi Oguma) wrote:

A direct reason is that my customer wants to have double quotes in the xml declaration because their handcrafted 'xml lint' tool complains the xml document produced by my ruby script contains single quotes. (Their convention requires all quoting characters should be double quotes.)

That's a very good reason, in particular for you.

More generally, I think it is natural if we have control which quoting character will be used for the xml declaration too.

It probably won't hurt if this is controllable. But there might be some existing applications (and tests) that expect single-quoted pseudo-attributes in XML declarations, and they would get problems if

```
doc.context[:attribute_quote] = :quote
changes that. So I think it would be better if it's something like
doc.context[:xml_declaration_quote] = :quote
```

Anyway, I don't have the time to prepare a patch, sorry. But maybe you can create a patch?

#4 - 01/07/2014 10:24 PM - kou (Kouhei Sutou)

- Status changed from Feedback to Assigned

- Assignee set to kou (Kouhei Sutou)

duerst (Martin Dürst) wrote:

It probably won't hurt if this is controllable. But there might be some existing applications (and tests) that expect single-quoted pseudo-attributes in XML declarations, and they would get problems if

```
doc.context[:attribute_quote] = :quote
changes that. So I think it would be better if it's something like
doc.context[:xml_declaration_quote] = :quote
```

I also prefer to :xml_declaration_quote.

Anyway, I don't have the time to prepare a patch, sorry. But maybe you can create a patch?

I comment his patch: <https://github.com/ruby/ruby/pull/496>

I'm happy that if he applies my comments but it is OK that he doesn't care about them. I will apply my comments after I merge his patch. :-)

- You can use "@parent" instead of "parent" because "@parent" is widely used in REXML internal.
- I don't like one character variable. :<
- "and" is REXML style rather than "&&". (But "|" is used in XMLDecl#content...)
- I prefer to one assertion (one check item) in one test rather than multiple assertions in one test.
- It seems that you don't need to use "xml" created in "setup".

#5 - 01/09/2014 12:49 PM - bearmini (Takashi Oguma)

[duerst \(Martin Dürst\)](#), [kou \(Kouhei Sutou\)](#),

Thanks for the comments!

I'll revise my patch to use :xml_declaration_quote and back to you soon.

#6 - 01/20/2014 02:07 PM - bearmini (Takashi Oguma)

Hi,

I'm working on this, and ran into another issue. i.e. should we be able to control single quote or double quote for DOCTYPE? If so, how?

According to <http://www.w3.org/TR/xml/#NT-doctypedecl>, DOCTYPE may contain ExternalID which may contain SystemLiteral or PubidLiteral, they can be quoted either single quote or double quote.

Should we introduce another symbol such as :xml_doctype_quote, or change the :xml_declatation_quote to :xml_prologue_quote or something?

#7 - 01/21/2014 04:03 AM - duerst (Martin Dürst)

Takashi Oguma wrote:

Should we introduce another symbol such as :xml_doctype_quote, or change the :xml_declatation_quote to :xml_prologue_quote or something?

Either way is fine by me, but maybe it's best to just use one symbol (:xml_prologue_quote) until we find a use case for separating.

#8 - 01/30/2014 02:04 AM - bearmini (Takashi Oguma)

Hi,

I have updated my PR <https://github.com/ruby/ruby/pull/496>
Please review it and comment on that.

#9 - 01/31/2014 02:17 PM - kou (Kouhei Sutou)

Thanks!

But I can't merge the pull request because it changes public API, NotationDecl.new. I don't want to add backward incompatible API change.

I'll comment further on the pull request later. Sorry.

Anyway, thanks for your work!

#10 - 02/01/2014 02:54 PM - kou (Kouhei Sutou)

- include StringQuotes after private is meaningless. Please put it the top of class definition.
- It seems that strip_quotes is needless for adjust_prologue_quotes. Can we just remove .inspect from adjust_prologue_quotes XXX.inspect? If we can do it, adjust_prologue_quotes isn't good name. quote or something will be good name.
- Please don't omit parenthesis for method call in #{...}.
- Please don't set DocDecl as NotationDecl's parent. (It was mentioned at [#9367-9](#).) How about calling context method instead of getting context from @parent in adjust_prologue_quotes? It seems that adjust_prologue_quotes is too depends on internal implementation.

(Sorry for my many comments...)

#11 - 05/24/2015 06:34 AM - duerst (Martin Dürst)

- Has duplicate Bug #11176: Hardcoded settings in REXML added

#12 - 12/16/2018 01:04 PM - kou (Kouhei Sutou)

- Backport deleted (1.9.3: UNKNOWN, 2.0.0: UNKNOWN, 2.1: UNKNOWN)

- Status changed from Assigned to Closed

I've implemented :prologue_quote context option. Available values are :quote and :apostrophe.

<https://github.com/ruby/rexml/commit/e89702294881a8aedb15905d8a85137876749f7f>