

Ruby master - Bug #10075

URI#join needs documentation of its behavior

07/19/2014 08:01 AM - jxf (John Feminella)

Status: Closed	
Priority: Normal	
Assignee: zzak (Zachary Scott)	
Target version: 2.2.0	
ruby -v: ruby 2.1.2p95 (2014-05-08 revision 45877) [x86_64-linux]	Backport: 2.0.0: UNKNOWN, 2.1: UNKNOWN
Description The documentation for URI.join says: "Joins URIs." Let's look at what a similar join method documentation says, on File: Returns a new string formed by joining the strings using File::SEPARATOR. That seems pretty clear. Indeed, we get: <pre>File.join 'path', 'to', 'join' # => "path/to/join"</pre> which is what we expected. What do we get if we try the natural URI equivalent? <pre>> URI.join('http://example.com', 'foo', 'bar')</pre> We probably expect something like: <pre># => "http://example.com/foo/bar"</pre> but we'll actually get <pre># => "http://example.com/bar"</pre> This seems surprising and counterintuitive, even if it matches the documentation behavior, because the documentation doesn't explain why that's the case. I think if Ruby is going to be surprising in that way, it needs to explain that to users in the documentation.	
Related issues: Is duplicate of Ruby master - Bug #6057: URI - Nonsensical Behavior Closed	

Associated revisions

Revision 37fd832f - 07/27/2014 08:34 PM - zzak (Zachary Scott)

- lib/uri/common.rb: [DOC] [Bug #10075] Clarify how URI.join arguments are handled by RFC3986, original reported by John Feminella.

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/trunk@46979 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 46979 - 07/27/2014 08:34 PM - zzak (Zachary Scott)

- lib/uri/common.rb: [DOC] [Bug #10075] Clarify how URI.join arguments are handled by RFC3986, original reported by John Feminella.

Revision 46979 - 07/27/2014 08:34 PM - zzak (Zachary Scott)

- lib/uri/common.rb: [DOC] [Bug #10075] Clarify how URI.join arguments are handled by RFC3986, original reported by John Feminella.

Revision 46979 - 07/27/2014 08:34 PM - zzak (Zachary Scott)

- lib/uri/common.rb: [DOC] [Bug #10075] Clarify how URI.join arguments are handled by RFC3986, originally reported by John Feminella.

Revision 46979 - 07/27/2014 08:34 PM - zzak (Zachary Scott)

- lib/uri/common.rb: [DOC] [Bug #10075] Clarify how URI.join arguments are handled by RFC3986, originally reported by John Feminella.

Revision 46979 - 07/27/2014 08:34 PM - zzak (Zachary Scott)

- lib/uri/common.rb: [DOC] [Bug #10075] Clarify how URI.join arguments are handled by RFC3986, originally reported by John Feminella.

Revision 46979 - 07/27/2014 08:34 PM - zzak (Zachary Scott)

- lib/uri/common.rb: [DOC] [Bug #10075] Clarify how URI.join arguments are handled by RFC3986, originally reported by John Feminella.

History

#1 - 07/20/2014 10:24 PM - drbrain (Eric Hodel)

- Subject changed from *URI#join seems to violate Principle of Least Surprise to URI#join needs documentation of its behavior*
- Category set to doc

I have changed this to a doc bug, the behavior does not surprise me, I expect it to behave as it does.

#2 - 07/20/2014 11:36 PM - phluid61 (Matthew Kerwin)

Eric Hodel wrote:

I have changed this to a doc bug, the behavior does not surprise me, I expect it to behave as it does.

It surprised me. The operation it performs appears to be "resolve relative" [1], but, like OP, from the name alone I'd have expected "concatenate with separators".

I think the case could be made that either/both operation could be called "join", but I think there's a better name for both. Anyway, better documentation would resolve the issue.

[1] <http://tools.ietf.org/html/rfc3986#section-5.2>

#3 - 07/21/2014 02:40 PM - jxf (John Feminella)

Eric Hodel wrote:

I have changed this to a doc bug, the behavior does not surprise me, I expect it to behave as it does.

Here's my thoughts on this: every other Ruby method I can think of named #join works by concatenating its arguments together with a separator (possibly with some additional logic to remove redundant separators and such).

However, URL.join doesn't work in this way. That's why I found it surprising, and reading the documentation didn't alleviate my surprise. So, better documentation would be an improvement, but I think it would be a bigger improvement if it worked more like the other join methods to begin with.

#4 - 07/21/2014 11:53 PM - avdi (Avdi Grimm)

On Mon, Jul 21, 2014 at 10:40 AM, jxf+ruby@jxf.me wrote:

Here's my thoughts on this: every other Ruby method I can think of named #join works by concatenating its arguments together with a separator (possibly with some additional logic to remove redundant separators and such).

This has caught me by surprise multiple times as well. I don't question the need for a method with these semantics, but the choice of name "join" seems misleading. Too late to change it now, but I wouldn't mind seeing it aliased to a less surprising name (perhaps "#merge_path", per RFC3986) and the "join" version eventually deprecated.

#5 - 07/26/2014 04:53 PM - zzak (Zachary Scott)

- Status changed from Open to Assigned

- Assignee set to zzak (Zachary Scott)
- Target version set to 2.2.0

I will try to improve the explanation

#6 - 07/27/2014 08:34 PM - zzak (Zachary Scott)

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

Applied in changeset r46979.

- lib/uri/common.rb: [DOC] [Bug [#10075](#)] Clarify how URI.join arguments are handled by RFC3986, original reported by John Feminella.

#7 - 11/21/2018 11:33 PM - mame (Yusuke Endoh)

- Is duplicate of Bug #6057: URI - Nonsensical Behavior added