

Ruby master - Bug #15244

Method #extname return empty string if filename is dot ('.')

10/23/2018 11:17 AM - TiSer (Sergey Ti)

Status: Closed	
Priority: Normal	
Assignee:	
Target version:	
ruby -v: 2.7.0dev	Backport: 2.3: UNKNOWN, 2.4: UNKNOWN, 2.5: UNKNOWN
Description	
Example 1:	
Current behaviour:	
<pre>File.extname('..jpg') => ""</pre>	
Expected behaviour:	
<pre>File.extname('..jpg') => ".jpg"</pre>	
Example 2:	
Current behaviour:	
<pre>File.extname('...jpg') => ""</pre>	
Expected behaviour:	
<pre>=> ".jpg"</pre>	
Related issues:	
Related to Ruby master - Bug #15267: File.basename + File.extname does not re...	Closed

History

#1 - 10/23/2018 11:56 AM - TiSer (Sergey Ti)

- Description updated

add one more example

#2 - 10/23/2018 12:19 PM - shevegen (Robert A. Heiler)

Is this a real problem?

How many people have had entries with more leading '.'?

```
File.extname('foobar.....jpg') # => ".jpg"  
File.extname('.....jpg') # => ""
```

It's not that I am not inclined to agree about it; I just think it's a bit strange to want to change that behaviour. Do people use lots of '.' for names to files or directories or symlinks?

#3 - 10/23/2018 12:31 PM - ahorek (Pavel Rosický)

IMO

```
File.extname('..jpg')
```

should be

```
=> ".jpg"
```

proposed

```
=> "."
```

or the current behaviour

```
=> ""
```

makes no sense

#4 - 10/23/2018 12:40 PM - TiSer (Sergey Ti)

- Description updated

#5 - 10/23/2018 12:42 PM - TiSer (Sergey Ti)

ahorek (Pavel Rosický) wrote:

IMO

```
File.extname('..jpg')
```

should be

```
=> ".jpg"
```

proposed

```
=> "."
```

or the current behaviour

```
=> ""
```

makes no sense

Yes, sure. I'm sorry, I oversleep correct cases.

Updated.

#6 - 10/29/2018 05:41 AM - nobu (Nobuyoshi Nakada)

If `File.extname("..jpg")` returns `".jpg"`, `File.basename("..jpg", ".")` should return `""` without that part, to restore the original name from from these results.

It is different from `basename(1)`.

```
$ basename ..jpg '.*'  
..jpg
```

#7 - 11/13/2018 02:39 PM - TiSer (Sergey Ti)

This is still present.

#8 - 11/22/2018 03:11 PM - TiSer (Sergey Ti)

Still present.

#9 - 02/16/2019 08:01 AM - TiSer (Sergey Ti)

Still present.

#10 - 07/26/2019 05:16 PM - rushsteve1 (Steven vanZyl)

- File dot-names.diff added

The issue seemed to be in how dotfiles were handled, where all leading periods were trimmed when only the first should have been.

I've created a patch that seems to fix the issue and added a spect test for it. It's only a 2 line change.

This is my first patch submission so any and all feedback would be greatly appreciated.

#11 - 08/02/2019 12:24 AM - rushsteve1 (Steven vanZyl)

- ruby -v changed from 2.5.0p0 to 2.7.0dev

#12 - 08/02/2019 03:16 PM - rushsteve1 (Steven vanZyl)

- File dot-names.diff added

Updated the diff to also account for the if (*p) case so that it will stop if p == 0.

#13 - 08/02/2019 03:17 PM - rushsteve1 (Steven vanZyl)

- File deleted (dot-names.diff)

#14 - 08/13/2019 11:16 PM - jeremyevans0 (Jeremy Evans)

- Related to Bug #15267: File.basename + File.extname does not restore the original name added

#15 - 08/14/2019 01:15 AM - jeremyevans0 (Jeremy Evans)

- File multiple-leading-dot-basename-extname-15244.patch added

I agree that this is a bug. The comparison that nobu made with basename(1) is relevant, but unlike File.basename, basename(1) does not handle .* specially. You can see that basename(1) does actually strip the extension even for multiple leading periods:

```
basename '..*' '..*'
.
```

Attached is a patch that fixes the issue, both for File.basename and File.extname. With the patch:

```
File.basename('..jpg', '.*')
# => "."
File.extname('..jpg')
# => ".jpg"
```

#16 - 09/02/2019 06:50 AM - matz (Yukihiro Matsumoto)

- Status changed from Open to Feedback

.git is not an extension but a filename. .a.jpg has .jpg extension. I wonder how we can parse .jpg. It's undefined behavior for me. Any reasoning?

Matz.

#17 - 09/02/2019 07:12 AM - znz (Kazuhiro NISHIYAMA)

In some other languages:

```
% python
Python 2.7.16 (default, Mar  4 2019, 09:01:38)
[GCC 4.2.1 Compatible Apple LLVM 10.0.0 (clang-1000.11.45.5)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
>>> import os
>>> os.path.splitext('..jpg')
('..jpg', '')
```

```
% python3
Python 3.7.4 (default, Jul  9 2019, 18:13:23)
[Clang 10.0.1 (clang-1001.0.46.4)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
>>> import os
>>> os.path.splitext('..jpg')
('..jpg', '')
```

```
% php --version
PHP 7.1.23 (cli) (built: Feb 22 2019 22:19:32) ( NTS )
Copyright (c) 1997-2018 The PHP Group
Zend Engine v3.1.0, Copyright (c) 1998-2018 Zend Technologies
% php -r 'var_export(pathinfo("..jpg", PATHINFO_EXTENSION));'
'jpg'
```

ref https://hydrocul.github.io/wiki/programming_languages_diff/io/extname.html

#18 - 09/02/2019 09:42 AM - nobu (Nobuyoshi Nakada)

znz (Kazuhiro NISHIYAMA) wrote:

```
% php --version
PHP 7.1.23 (cli) (built: Feb 22 2019 22:19:32) ( NTS )
Copyright (c) 1997-2018 The PHP Group
Zend Engine v3.1.0, Copyright (c) 1998-2018 Zend Technologies
% php -r 'var_export(pathinfo("../jpg", PATHINFO_EXTENSION));'
'jpg'
```

Seems not for us:

```
var_export(pathinfo("../profile", PATHINFO_EXTENSION)); //=> 'profile'
```

#19 - 09/24/2019 07:08 PM - TiSer (Sergey Ti)

Example from .NET:

<https://docs.microsoft.com/en-us/dotnet/api/system.io.path.getextension>

```
class Program
{
    static void Main(string[] args)
    {
        string ext = Path.GetExtension("../jpg");
        Console.WriteLine(ext);
        Console.ReadLine();
    }
}
```

```
=> '.jpg'
```

#20 - 03/06/2020 09:36 PM - jeremyevans0 (Jeremy Evans)

- Status changed from Feedback to Closed

Files

dot-names.diff	891 Bytes	08/02/2019	rushsteve1 (Steven vanZyl)
multiple-leading-dot-basename-extname-15244.patch	1.94 KB	08/14/2019	jeremyevans0 (Jeremy Evans)