

Ruby - Bug #12103

ruby process hangs while executing regular expression.

02/23/2016 09:46 AM - antonyr (Antony Raj)

Status:	Rejected	
Priority:	Normal	
Assignee:		
Target version:		
ruby -v:	ruby 2.3.0p0 (2015-12-25 revision 53290) [x86_64-darwin15], jruby 9.0.5.0 (2.2.3) 2016-01-26 7bee00d Java HotSpot(TM) 64-Bit Server VM 24.45-b08 on 1.7.0_45-b18 +jit [darwin-x86_64]	Backport: 2.0.0: UNKNOWN, 2.1: UNKNOWN, 2.2: UNKNOWN, 2.3: UNKNOWN
Description <p>The following code hangs</p> <pre>regex = /(\\((\\w+([\\p{Punct}\\s]{,3}\\w*)*)\\))/i detail = "(said companies being fictitious sellers" detail =~ regex</pre> <p>However when I change {,3} to {1,3}, the code executes and moves on. Otherwise hangs</p>		

History

#1 - 02/27/2016 09:11 AM - duerst (Martin Dürst)

- Status changed from Open to Rejected

I simplified this as follows:

```
regex = /(a+( ?a)*/
detail = "(aaaa aaaaaaa aaaaa aaaaaaaaa aaaaa"
```

It still hangs. But it's easier to see why: The part (a+(?a)* allows many different ways to match the same thing, and all these are tried. For more background, please see <http://www.regular-expressions.info/catastrophic.html>.

#2 - 02/28/2016 03:31 AM - normalperson (Eric Wong)

duerst@it.aoyama.ac.jp wrote:

Status changed from Open to Rejected

I simplified this as follows:

```
regex = /(a+( ?a)*/
detail = "(aaaa aaaaaaa aaaaa aaaaaaaaa aaaaa"
```

Keep in mind the equivalent Perl terminates just fine:

```
my $detail = "(aaaa aaaaaaa aaaaa aaaaaaaaa aaaaa";
$detail =~ /\(a+( ?a)*\)/;
```

Tested on an old perl v5.14.2 from Debian wheezy (oldstable).

So I prefer we acknowledge cases like these as bugs to at least encourage someone to fix them.

Not a high priority for me, though, as I'm not familiar with the inner workings of RE engines (and also know to avoid tricky REs).

#3 - 02/29/2016 01:02 AM - shyuhei (Shyouhei Urabe)

On 02/28/2016 12:22 PM, Eric Wong wrote:

Keep in mind the equivalent Perl terminates just fine:

Yes. But,

So I prefer we acknowledge cases like these as bugs to at least encourage someone to fix them.

Ruby's (and Perl's) RE grammar is known to be NP complete. Fixing this ``issue" ultimately requires a proof of $P=NP$. I don't think it'll happen in a near future.

I guess Perl's RE engine can happen to handle this specific `/(a+(?a*))` case, maybe due to their optimization.