

Ruby - Bug #20083

String#match? behaving inconsistently with Ruby 3.3.0

12/25/2023 12:12 PM - jussikos (Jussi Koljonen)

Status:	Closed	
Priority:	Normal	
Assignee:		
Target version:		
ruby -v:	ruby 3.3.0 (2023-12-25 revision 5124f9ac75) [x86_64-darwin23]	Backport: 3.0: DONTNEED, 3.1: DONTNEED, 3.2: DONTNEED, 3.3: DONE
Description From irb, when calling String#match? pattern = /([\s]*ABC)\$/i # or /(\\s*ABC)/i p "1ABC".match?(pattern) # => true p "12ABC".match?(pattern) # => true p "123ABC".match?(pattern) # => true p "1231ABC".match?(pattern) # => true p "12312ABC".match?(pattern) # => false p "123123ABC".match?(pattern) # => false p "1231231ABC".match?(pattern) # => true p "12312312ABC".match?(pattern) # => true p "123123123ABC".match?(pattern) # => false p "1231231231ABC".match?(pattern) # => false p "12312312312ABC".match?(pattern) # => true p "123123123123ABC".match?(pattern) # => true p "1231231231231ABC".match?(pattern) # => false With ruby 3.2.2 (2023-03-30 revision e51014f9c0) [x86_64-darwin22] and earlier versions (2.7.8 to 3.2.2) return value is always true Update: the problem seems to be somehow related to the /i option, as all the above examples work correctly with /([\s]*ABC)\$/		
Related issues: Has duplicate Ruby - Bug #20095: Regex lookahead behaving strangely in 3.3.0		
		Closed

Associated revisions

Revision d8702ddbfb8cc7fc601a9a4d19842ef9c2b76c1 - 12/28/2023 02:20 PM - make_now_just (Hiroya Fujinami)

Fix [Bug #20083]: correct a cache point size for atomic groups (#9367)

Revision d8702ddbfb8cc7fc601a9a4d19842ef9c2b76c1 - 12/28/2023 02:20 PM - make_now_just (Hiroya Fujinami)

Fix [Bug #20083]: correct a cache point size for atomic groups (#9367)

Revision d8702ddb - 12/28/2023 02:20 PM - make_now_just (Hiroya Fujinami)

Fix [Bug #20083]: correct a cache point size for atomic groups (#9367)

Revision 5f3dfa1c273c6fb9eae65ceca633b46f7e30f686 - 01/30/2024 01:31 AM - naruse (Yui NARUSE)

merge revision(s) d8702ddbfb8cc7fc601a9a4d19842ef9c2b76c1: [Backport #20083]

```
Fix [Bug #20083]: correct a cache point size for atomic groups
(#9367)

---
 regexec.c          | 2 +-
 test/ruby/test_regexp.rb | 8 ++++++
 2 files changed, 9 insertions(+), 1 deletion(-)
```

Revision 5f3dfa1c273c6fb9eae65ceca633b46f7e30f686 - 01/30/2024 01:31 AM - naruse (Yui NARUSE)

merge revision(s) d8702ddbfb8cc7fc601a9a4d19842ef9c2b76c1: [Backport #20083]

```
Fix [Bug #20083]: correct a cache point size for atomic groups
(#9367)
```

```
---
 regexec.c | 2 +-
 test/ruby/test_regexp.rb | 8 +++++++
 2 files changed, 9 insertions(+), 1 deletion(-)
```

Revision 5f3dfa1c - 01/30/2024 01:31 AM - naruse (Yui NARUSE)

merge revision(s) d8702ddbfe8cc7fc601a9a4d19842ef9c2b76c1: [Backport #20083]

```
Fix [Bug #20083]: correct a cache point size for atomic groups
(#9367)
```

```
---
 regexec.c | 2 +-
 test/ruby/test_regexp.rb | 8 +++++++
 2 files changed, 9 insertions(+), 1 deletion(-)
```

History

#1 - 12/25/2023 12:14 PM - jussikos (Jussi Koljonen)

- Subject changed from *Regexp#match? behaving inconsistently with Ruby 3.3.0* to *String#match? behaving inconsistently with Ruby 3.3.0*

#2 - 12/26/2023 01:21 PM - jussikos (Jussi Koljonen)

- Description updated

#3 - 12/26/2023 01:22 PM - jussikos (Jussi Koljonen)

- Description updated

#4 - 12/26/2023 07:44 PM - jussikos (Jussi Koljonen)

- Description updated

#5 - 12/27/2023 01:17 AM - make_now_just (Hiroya Fujinami)

I created a PR for this bug (See <https://github.com/ruby/ruby/pull/9367>).
Thank you for your reporting!

The bug reason is a combination of a regex optimization and a bug for atomic groups.
First, since `\s` and the following `A` (internally it is treated as `[aA]` on `i` flag.) is mutually disjoint, `\s*ABC` is optimized to `(?>\s*)ABC`.
Next, match cache optimization for atomic groups in this case is buggy, so the matching results become wrong.

When `i` flag is not given, another optimization is applied and `\s*ABC` is optimized to `(?:(!A)\s)*ABC`, so the bug is not occurred.

#6 - 12/28/2023 02:55 PM - naruse (Yui NARUSE)

- Backport changed from 3.0: UNKNOWN, 3.1: UNKNOWN, 3.2: UNKNOWN, 3.3: UNKNOWN to 3.0: DONTNEED, 3.1: DONTNEED, 3.2: DONTNEED, 3.3: REQUIRED

#7 - 01/05/2024 02:19 PM - naruse (Yui NARUSE)

- Has duplicate Bug #20095: *Regex lookahead behaving strangely in 3.3.0* added

#8 - 01/30/2024 01:27 AM - naruse (Yui NARUSE)

- Status changed from Open to Closed

#9 - 01/30/2024 03:10 AM - naruse (Yui NARUSE)

- Backport changed from 3.0: DONTNEED, 3.1: DONTNEED, 3.2: DONTNEED, 3.3: REQUIRED to 3.0: DONTNEED, 3.1: DONTNEED, 3.2: DONTNEED, 3.3: DONE

ruby_3_3 5f3dfa1c273c6fb9eae65ceca633b46f7e30f686 merged revision(s) d8702ddbfe8cc7fc601a9a4d19842ef9c2b76c1.