

Ruby master - Feature #5877

Poor performance of initial and final UTF-8 substrings

01/11/2012 08:51 AM - nex3 (Natalie Weizenbaum)

Status:	Rejected
Priority:	Normal
Assignee:	
Target version:	
Description	
Currently, taking the substring of a UTF-8 string always takes time proportional to the length of the string. While this is necessary in the general case, there are many common cases where it could be much faster, in particular when the substring is near the beginning of the end of the superstring. In those cases, the operation should only be proportional to the length of the substring.	
Gist with reproduction and more detailed analysis forthcoming.	

History

#1 - 01/20/2012 10:24 AM - naruse (Yui NARUSE)

- Status changed from Open to Feedback

#2 - 01/20/2012 11:08 AM - kosaki (Motohiro KOSAKI)

You didn't write gist url. I haven't catch your point.

#3 - 01/21/2012 08:53 PM - cjheath (Clifford Heath)

On 20/01/2012, at 1:08 PM, Motohiro KOSAKI wrote:

Bug #5877: Poor performance of initial and final UTF-8 substrings
<https://bugs.ruby-lang.org/issues/5877>

Author: Nathan Weizenbaum

Status: Feedback

Priority: Normal

Assignee:

Category:

Target version:

ruby -v: ruby 1.9.3p0 (2011-10-30 revision 33570) [x86_64-linux]

Currently, taking the substring of a UTF-8 string always takes time proportional to the length of the string. While this is necessary in the general case, there are many common cases where it could be much faster, in particular when the substring is near the beginning of the end of the superstring. In those cases, the operation should only be proportional to the length of the substring.

This discussion from ten years ago is relevant to performance workarounds for UTF-8:

<http://blade.nagaokaut.ac.jp/cgi-bin/vframe.rb/ruby/ruby-talk/45935?45702-47288>

Clifford Heath.

#4 - 10/31/2012 02:07 AM - mame (Yusuke Endoh)

- Tracker changed from Bug to Feature

- Status changed from Feedback to Rejected

- Priority changed from Normal to 3

Not a bug. "forthcoming" did not come.

Please try to make your proposal self-contained.

--

Yusuke Endoh mame@tsg.ne.jp