

Ruby - Feature #19718

Extend `-0`` option

06/07/2023 12:02 PM - nobu (Nobuyoshi Nakada)

Status:	Rejected
Priority:	Normal
Assignee:	
Target version:	
Description <p>Recently I learned that Perl's <code>-0</code> option is extended to accept a hexadecimal Unicode codepoint. However it uses <code>-0x</code>, and since <code>-x</code> is used for shebang and <code>cd</code>, it would cause a backward incompatibility if we will incorporate it as it is.</p> <p>So I propose that <code>-0uCODEPOINT</code> instead. This can be extended to comma- or colon-separated codepoint list.</p> <p>Also, another idea is <code>-0:separator</code> which specifies the separator as-is. Considering many of modern shells provide escaped string form (e.g., <code>`\$`uHHHH`</code>), this may be a more modern answer, except for NUL cannot be represented.</p> <p>https://github.com/ruby/ruby/pull/7914</p>	

History

#1 - 06/07/2023 02:00 PM - Dan0042 (Daniel DeLorme)

It's a nifty idea, but `-0uCODEPOINT` means that the `-u` option is forever reserved for this usage which no one will ever really use I think, because `-0` is meant to be used like `xargs -0` option (ex: `find . -print0 | ruby -One 'p $_.chomp'`) and I can't imagine why anyone would use anything other than NUL. Although I would love to be shown wrong.
`-0:separator` doesn't have that downside.

#2 - 06/07/2023 02:25 PM - nobu (Nobuyoshi Nakada)

I thought the issue about `-u` is simply a matter of priorities; just the chance of `-0x` cannot be denied, but `-0u` should have never been used.

But now I remembered `-U` has been implemented since 1.9.
My patch allows `-0U` as well as `-0u`, and would need to change.
Probably a punctuation (like as `:`) may be possible.

#3 - 06/07/2023 03:39 PM - Dan0042 (Daniel DeLorme)

I meant it will not be possible to use `-u` in the future.
With colon we could write anything after, like `-0:$`t`` or `-0:011` or `-0:x09` or `-0:u0009` or `-0:b1001` ☐☐
Or use `=` instead of `:` ?
Cheers.

#4 - 06/07/2023 11:46 PM - nobu (Nobuyoshi Nakada)

Dan0042 (Daniel DeLorme) wrote in [#note-3](#):

I meant it will not be possible to use `-u` in the future.

Yes, and I thought it wouldn't be a problem, but might not be a good idea.

With colon we could write anything after, like `-0:$`t`` or `-0:011` or `-0:x09` or `-0:u0009` or `-0:b1001` ☐☐
Or use `=` instead of `:` ?

`-0=$`t`` and `-0:09(= -0:x09)?`

#5 - 06/08/2023 08:12 AM - matz (Yukihiro Matsumoto)

- Status changed from Open to Rejected

Very few users know the existence of `-0`. If we extend the option, probably no one uses the feature.

