

Ruby trunk - Feature #15195

Deal with new Japanese era

10/03/2018 05:57 AM - duerst (Martin Dürst)

Status:	Closed
Priority:	Normal
Assignee:	duerst (Martin Dürst)
Target version:	
Description	
<p>As most people in Japan should know by now, a new era of the Japanese calendar will start in May 2019.</p> <p>It is not clear when the new era name will be announced, but it may be at rather short notice.</p> <p>The Unicode consortium has made up a plan about how to deal with this, see http://blog.unicode.org/2018/09/new-japanese-era.html. Among else, they will encode an equivalent of the current 〰, 〱, 〲, and plan to publish a special version of the Unicode standard numbered 12.1.</p> <p>We should start to discuss and tentatively decide what (if anything) will be needed for Ruby. Alternatives include the following two and others inbetween.</p> <ul style="list-style-type: none">• Follow past years where the update to a new Unicode version is done with the Ruby version release around Christmas. That would mean we upgrade to Unicode 12 around Christmas 2019. At that time, we can include the changes in 12.1 without any problems.• Implement the upgrade to Unicode 12 as quickly as possible in trunk, implement Unicode 12.1 as soon as available in trunk, and backport the changes needed for the new era name to older Ruby versions.	
Related issues:	
Related to Ruby trunk - Feature #15742: Add Date#jisx0301 support new Japanes...	Closed
Related to Ruby trunk - Misc #15750: Switch Unicode Version 12.1.0 back from ...	Closed
Blocked by Ruby trunk - Feature #15321: Update Unicode data to Unicode Versio...	Closed
Blocked by Ruby trunk - Bug #15641: Backport update to Unicode Version 12.0.0...	Closed
Blocks Ruby trunk - Bug #15740: Backport Unicode Version 12.1.0 and other Jap...	Closed

Associated revisions

Revision 7fe64d17 - 04/05/2019 12:58 AM - duerst (Martin Dürst)

update to Unicode Version 12.1.0 (beta)

Unicode Version 12.1.0 adds one single character, U+32FF SQUARE ERA NAME REIWA, for the new Japanese era starting on May 1st. 12.1.0 will be finalized only on May 7th, so we go with the beta version because further changes in the data we need are highly unlikely, and we want to make sure Ruby is ready for the new era.

- common.mk: change UNICODE_VERSION to 12.1.0, UNICODE_BETA to YES
- enc/unicode/12.1.0, enc/unicode/12.1.0/casefold.h, enc/unicode/12.1.0/name2ctype.h: add directory and generated data files for new version
- lib/unicode_normalize/tables.rb: update for new character
- test/ruby/test_regexp.rb: add test for character property age=12.1
- test/test_unicode_normalize.rb: add test for NFKC decomposition of new character

This (mostly) completes issue #15195.

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/trunk@67441 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 67441 - 04/05/2019 12:58 AM - duerst (Martin Dürst)

update to Unicode Version 12.1.0 (beta)

Unicode Version 12.1.0 adds one single character, U+32FF SQUARE ERA NAME REIWA, for the new Japanese era starting on May 1st. 12.1.0 will be finalized only on May 7th, so we go with the beta version because further changes in the data we need are highly unlikely, and we want to make sure Ruby is ready for the new era.

- common.mk: change UNICODE_VERSION to 12.1.0, UNICODE_BETA to YES
- enc/unicode/12.1.0, enc/unicode/12.1.0/casfold.h, enc/unicode/12.1.0/name2ctype.h: add directory and generated data files for new version
- lib/unicode_normalize/tables.rb: update for new character
- test/ruby/test_regexp.rb: add test for character property age=12.1
- test/test_unicode_normalize.rb: add test for NFKC decomposition of new character

This (mostly) completes issue #15195.

Revision 50eea44a - 04/05/2019 11:52 PM - duerst (Martin Dürst)

remove Unicode 12.0.0 related directory and generated files

This completes issue #15195.

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/trunk@67453 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 67453 - 04/05/2019 11:52 PM - duerst (Martin Dürst)

remove Unicode 12.0.0 related directory and generated files

This completes issue #15195.

Revision a2429de6 - 04/09/2019 09:41 AM - duerst (Martin Dürst)

mention support for Unicode Version 12.1.0 (issue #15195) in NEWS

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/trunk@67476 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 67476 - 04/09/2019 09:41 AM - duerst (Martin Dürst)

mention support for Unicode Version 12.1.0 (issue #15195) in NEWS

Revision 69ec3936 - 04/13/2019 03:01 PM - naruse (Yui NARUSE)

merge revision(s) 67439,67441,67453,67476: [Backport #15740]

```
change lib/unicode_normalize/tables.rb to single item per line to make diffs shorter
```

```
* template/unicode_norm_gen.tmpl: Change formatting of output to produce only a single item (or range) for each line to make future diffs shorter and easier to understand and check.
```

```
* lib/unicode_normalize/tables.rb: output of the above
```

```
update to Unicode Version 12.1.0 (beta)
```

```
Unicode Version 12.1.0 adds one single character, U+32FF SQUARE ERA NAME REIWA, for the new Japanese era starting on May 1st. 12.1.0 will be finalized only on May 7th, so we go with the beta version because further changes in the data we need are highly unlikely, and we want to make sure Ruby is ready for the new era.
```

```
* common.mk: change UNICODE_VERSION to 12.1.0, UNICODE_BETA to YES
```

```
* enc/unicode/12.1.0, enc/unicode/12.1.0/casfold.h, enc/unicode/12.1.0/name2ctype.h: add directory and generated data files for new version
```

```
* lib/unicode_normalize/tables.rb: update for new character
```

```
* test/ruby/test_regexp.rb: add test for character property age=12.1
```

```
* test/test_unicode_normalize.rb: add test for NFKC decomposition of new character
```

```
This (mostly) completes issue #15195.
```

remove Unicode 12.0.0 related directory and generated files

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/branches/ruby_2_6@67525 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 67525 - 04/13/2019 03:01 PM - naruse (Yui NARUSE)

merge revision(s) 67439,67441,67453,67476: [Backport #15740]

change lib/unicode_normalize/tables.rb to single item per line to make diffs shorter

* template/unicode_norm_gen.tmpl: Change formatting of output to produce only a single item (or range) for each line to make future diffs shorter and easier to understand and check.

* lib/unicode_normalize/tables.rb: output of the above

update to Unicode Version 12.1.0 (beta)

Unicode Version 12.1.0 adds one single character, U+32FF SQUARE ERA NAME REIWA, for the new Japanese era starting on May 1st. 12.1.0 will be finalized only on May 7th, so we go with the beta version because further changes in the data we need are highly unlikely, and we want to make sure Ruby is ready for the new era.

* common.mk: change UNICODE_VERSION to 12.1.0, UNICODE_BETA to YES

* enc/unicode/12.1.0, enc/unicode/12.1.0/casfold.h, enc/unicode/12.1.0/name2ctype.h: add directory and generated data files for new version

* lib/unicode_normalize/tables.rb: update for new character

* test/ruby/test_regexp.rb: add test for character property age=12.1

* test/test_unicode_normalize.rb: add test for NFKC decomposition of new character

This (mostly) completes issue #15195.

remove Unicode 12.0.0 related directory and generated files

History

#1 - 10/03/2018 06:19 AM - naruse (Yui NARUSE)

Both plans are OK, but the latter seems fun for this big event.

#2 - 10/09/2018 08:14 AM - duerst (Martin Dürst)

Link to (request for) feedback on this issue from Rails: <https://groups.google.com/forum/#!topic/rubyonrails-core/fZUk1qXRT5k>.

#3 - 11/20/2018 06:59 AM - duerst (Martin Dürst)

- Related to Feature #15321: Update Unicode data to Unicode Version 12.0.0 added

#4 - 12/10/2018 04:56 AM - duerst (Martin Dürst)

- Related to deleted (Feature #15321: Update Unicode data to Unicode Version 12.0.0)

#5 - 12/10/2018 04:57 AM - duerst (Martin Dürst)

- Blocked by Feature #15321: Update Unicode data to Unicode Version 12.0.0 added

#6 - 12/10/2018 05:01 AM - duerst (Martin Dürst)

There has been an announcement recently in the Japanese news that the Japanese government will make sure the new era name is being published at least one month before the era change. The era change is planned for May 1st, 2019. This means that the new era name should be know on April 1st, 2019 at the latest.

#7 - 12/15/2018 03:03 AM - duerst (Martin Dürst)

According to

<https://www.japantimes.co.jp/news/2018/12/06/national/politics-diplomacy/japan-mulls-announcing-new-era-name-april-11-sources/#.XBQPC6qWxD8>, the announcement of the new area name may be postponed until about April 11.

#8 - 03/06/2019 05:24 AM - duerst (Martin Dürst)

- Blocked by Bug #15641: Backport update to Unicode Version 12.0.0 back to Ruby 2.6 branch added

#9 - 03/21/2019 12:59 AM - duerst (Martin Dürst)

- Assignee set to duerst (Martin Dürst)

I have verified that with the beta files for Unicode Version 12.1.0 at <https://www.unicode.org/Public/12.1.0/ucd/>, Ruby compiles and runs all tests (make check) successfully.

Please note that different from other betas, this beta uses dummy data: The dummy for the new Japanese era name is 𐤀 (left right):

```
32FF; SQUARE ERA NAME SAYUU; So; 0; L; <square> 5DE6 53F3; ; ; ; N; ; ; ;
```

which is **guaranteed** to change.

#10 - 03/31/2019 07:00 AM - duerst (Martin Dürst)

- Blocks Bug #15740: Backport Unicode Version 12.1.0 and other Japanese new era issues to Ruby 2.6 added

#11 - 03/31/2019 07:01 AM - duerst (Martin Dürst)

- Subject changed from How to deal with new Japanese era to Deal with new Japanese era

Changed title from "How to deal with new Japanese era" to "Deal with new Japanese era".

#12 - 04/01/2019 02:47 AM - duerst (Martin Dürst)

The new era name has been announced a few minutes ago:

```
𐤀 U+4EE4 U+548C
```

Reading: REIWA

#13 - 04/01/2019 04:28 AM - duerst (Martin Dürst)

- Related to Feature #15742: Add Date#jisx0301 support new Japanese era added

#14 - 04/01/2019 05:18 AM - duerst (Martin Dürst)

A beta release for Unicode 12.1.0 should be available by tomorrow. The official release of Unicode 12.1.0 is planned for **May 7**.

This raises the question of whether we should release a Ruby version before May 1st (with the beta data) or after May 7th. For the former, we would have to cheat a bit by introducing some logic into the build process that can pretend that a beta version is already final.

#15 - 04/01/2019 05:37 AM - duerst (Martin Dürst)

duerst (Martin Dürst) wrote:

This raises the question of whether we should release a Ruby version before May 1st (with the beta data) or after May 7th. For the former, we would have to cheat a bit by introducing some logic into the build process that can pretend that a beta version is already final.

My preference, also for issue [#15742](#), would be to move ahead, so that Ruby users who need the new functionality can get it.

#16 - 04/06/2019 04:36 AM - duerst (Martin Dürst)

- Related to Misc #15750: Switch Unicode Version 12.1.0 back from beta to final added

#17 - 04/06/2019 04:37 AM - duerst (Martin Dürst)

- Status changed from Open to Closed