State of CJK issues of LibreOffice

Shinji Enoki

shinji.enoki@gmail.com

ROME | 11 October 2017
Talk summary

- LibreOffice have many language-specific features and issues, CJK (Chinese / Japanese / Korean) issue is one of them.
- I will explain what is the CJK features and issues.
Who am I?

- Shinji Enoki (榎真治)
- A member of LibreOffice Japanese Team (2011-)
  - Founding members
  - Join LibreOffice community in 2010
- A member of TDF (2014-)
- Activity: organize events, grows community, sometimes QA
LibreOffice Kaigi
What is CJK issues

- LibreOffice Bugzilla have the CJK meta issue
  - Bug 83066(CJK) - [META] CJK issues
  - CJK bugs are tracked META issues
Bug **83066** (CJK) - [META] CJK (Chinese, Japanese, Korean, and Vietnamese) language issues

### Status: NEW

**Alias:** CJK

**Product:** LibreOffice

**Component:** LibreOffice (show other bugs)

**Version:** unspecified

**Hardware:** Other All

**Importance:** medium normal

**Assignee:** Not Assigned

**QA Contact:**

**URL:**

**Whiteboard:**

**Keywords:**

**Dependencies:**

| Depends | 35208 35301 36705 33935 43671 50349 japanese-ruby-character 50201 55954 63011 65266 66791 69809 73329 72042 75795 76255 77803 on: 80348 80812 81272 81484 82115 82116 82293 82689 82690 82692 82930 83844 85795 89752 89979 90874 93695 95566 96197 101679 101895 102329 104378 104583 Text-Grd 107055 107195 107393 107435 107466 107487 107494 107495 107496 107497 108030 109260 111217 111477 112593 112745 112879 113046 113184 113189 113191 CJK-Chinese-Traditional CJK-Chinese-Simplified CJK-Japanese CJK-Korean: 45800 45822 44784 47710 49025 53120 58664 63755 66075 68156 67814 70002 76621 77416 77426 77514 77533 77672 77859 78786 90772 90847 90964 92576 92586 92818 92960 92965 82143 82252 83550 82752 82755 84138 85054 87636 88160 88237 88855 89040 91288 92655 92782 94008 94066 95370 95826 96691 96867 96892 97405 97702 98165 98269 98527 98879 101739 102041 102074 102252 103710 103718 103785 103850 104625 104726 104801 104828 104954 106097 106205 107462 112056 112807

**Crash report or crash signature:**

**See Also:** Font-Rendering Vertical-text RTL-CTL

**Reported:** 2014-08-25 14:50 UTC by Matthew Francis

**Modified:** 2017-10-20 17:03 UTC (History)

**CC List:** 14 users (show)

**Blocks:**

- Show dependency tree / graph

**Attachments**

Add an attachment (proposed patch, testcase, etc.)
Next Decade Manifesto

“To support the preservation of mother tongue by encouraging people to translate, document, support, and promote our office productivity tools in their mother tongue”

- I think this “Our Values” is very good
- Our problem want to resolve ourselves although it’s not easy to do
Last year’s Conference: Talk

- Kei Tasaki talked "Typical" bugs of LibreOffice in Japanese environment
  - How to create a document in Japan
  - “Typical” bugs in Japanese environment
Issues status of Last year talk

- Ruby using case bug
  - Bug 98269 - FORMATTING: Different spacing with Japanese ruby text between WIN and LINUX
    - RESOLVED DUPLICATE → VERIFIED FIXED
  - Bug 77514 - FORMATTING: CJK ruby text (furigana) in vertical mode pushes characters to the left of the base line
    - RESOLVED NOTABUG
フリーソフトウェア (free software) とは、ソフトウェアのうち、フリーソフトウェアが提唱する自由なソフトウェアを指す。大半のフリーソフトウェアは無償（フリー）として配布されているが、定義に従えば、現状事実上無償での入手がほぼ可能とみてよいが、必ずしも全ての一次配布が無償ということは定義にはない。

Bug #77514 - FORMATTING: CJK ruby text (furigana) in vertical mode pushes characters to the left of the baseline
- We can work around by setting that Mark Han advice
- I think that it is better to have such a setting by default in the Japanese environment. How about in other languages?
- Should we make “Request for Enhancement” as issue? That is, Default setting switches according to languages
- Reference about ruby: https://www.w3.org/TR/jlreq/#ruby_and_emphasis_dots
Kanji variant is a Kanji character with the same meaning but refers to what the shape is different.

All character is same means.
In Japan, using a mechanism called IVS to make it easier to deal with a Kanji variant.

IVS is combine the characters, as one of the character.
UCSコードは、細かな字形の差異を区別せず、複数の字形を一つのコードに統合している。

UCSコードと字形を指定するVS（字形選択子）を並べて記述することで、字形を指定することができる。※VSの値は実際の規格と異なります。

http://mojikiban.ipa.go.jp/1292.html
IVS using case bug

- This issue has not reported yet

We want ……

葛 + Additional information 葛

葛 → Deleted

Push Delete key
IVS problem is not fixed

Push Backspace key
CJK features

- In common with CJK, 1. use Kanji (Chinese characters), 2. vertical writing, 3. phonetic guides (ruby)

- Chinese characters are also different in Simplified Chinese (Mainland China), Traditional Chinese (Taiwan, Hong Kong) and Kanji (Japanese)
  - Entering kanji in this template turned to Simplified Chinese

- Even with common functions, usage methods are not always the same
  - Vertical writing
Features specific to Japanese

- Line breaking rules 禁則処理 Kinsokushori
  - [Paragraph style]-[Japanese appearance] tab option

<table>
<thead>
<tr>
<th>段落</th>
<th>背景</th>
<th>透明度</th>
</tr>
</thead>
<tbody>
<tr>
<td>アウトラインと番号付け</td>
<td>タブ</td>
<td>ドロップキャップ</td>
</tr>
<tr>
<td>インデントと行間隔</td>
<td>配置</td>
<td>体裁</td>
</tr>
<tr>
<td></td>
<td></td>
<td>日本語の体裁</td>
</tr>
</tbody>
</table>

改行時の処理
- 行頭および行末の禁則文字のリストで禁則処理を行う(A)
- 句読点のぶら下げを行う(B)
- 日本語、英文字、および混在文章で文字間隔を調整(C)
Check [META] CJK issues

<table>
<thead>
<tr>
<th>Depends on: (All)</th>
<th>135</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open</td>
<td>59</td>
</tr>
<tr>
<td>Close</td>
<td>76</td>
</tr>
</tbody>
</table>

- Mark Hung, is fixing CJK ISSUE Thank you!
- Some cases is not added to CJK issue due to Japanese related bug. I will add it in the future
Checking is difficult

- It is difficult to understand the meaning
- Test case, sample file is not enough
- I do not understand when I read the comment to the end
- I wanted to re-make or arrange the summary and test cases
Multilingual issues that matter even at CJK

- Multilingual problem that becomes a problem with CJK
- In some time zone Date is shifted
  - Bug 59850 - EDITING: typed date decreased one day
  - Bug 63230 - Calc: Date decreased by 1 depending on Timezone
Conclusion

- I share the Unique function and issue about CSK with a lot of members
- I’d like to work out to solve issues with our Japanese members
- I want to progress organization of Bugzilla such as CSK issue
Thank you!

shinji.enoki@gmail.com