Evaluation of LibreOffice – MS Office interoperability and automated bisection of interoperability regressions

Miloš Šrámek

Society for Open Information Technologies (Slovakia)
Earlier Work

- Presented at Plugfest 2011 in Berlin:
  - ODF document overlays for a detailed visual inspection
- Presented at LibreOffice Conference 2013 in Milano:
  - added 4 error measures, 4 views
- Both: Testing applications “each with each”

<table>
<thead>
<tr>
<th>File</th>
<th>#</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>AbiWord 2.8.6_Linux</td>
<td>0</td>
<td>100</td>
<td>81,92</td>
<td>81,92</td>
<td>24,54</td>
<td>24,55</td>
<td>24,54</td>
<td>24,6</td>
<td>24,6</td>
<td>24,3</td>
<td>24,3</td>
</tr>
<tr>
<td>AbiWord 2.9.0_Linux</td>
<td>1</td>
<td>81,92</td>
<td>100</td>
<td>100</td>
<td>24,51</td>
<td>22,89</td>
<td>22,89</td>
<td>22,95</td>
<td>22,95</td>
<td>22,71</td>
<td>22,71</td>
</tr>
<tr>
<td>AbiWord 2.9.1_Linux</td>
<td>2</td>
<td>81,92</td>
<td>100</td>
<td>100</td>
<td>24,51</td>
<td>22,89</td>
<td>22,89</td>
<td>22,95</td>
<td>22,95</td>
<td>22,71</td>
<td>22,71</td>
</tr>
<tr>
<td>KOffice 2.3.3_Linux</td>
<td>3</td>
<td>24,54</td>
<td>24,51</td>
<td>24,51</td>
<td>100</td>
<td>34,44</td>
<td>34,53</td>
<td>34,3</td>
<td>34,3</td>
<td>33,02</td>
<td>33,02</td>
</tr>
<tr>
<td>StarOffice 9.2_Linux</td>
<td>8</td>
<td>24,55</td>
<td>22,89</td>
<td>22,89</td>
<td>34,44</td>
<td>100</td>
<td>99,51</td>
<td>79,01</td>
<td>79,01</td>
<td>79,01</td>
<td>56,29</td>
</tr>
<tr>
<td>LibreOffice 3.3.2_Linux</td>
<td>4</td>
<td>24,54</td>
<td>22,89</td>
<td>22,89</td>
<td>34,53</td>
<td>99,51</td>
<td>100</td>
<td>79,32</td>
<td>79,32</td>
<td>79,32</td>
<td>56,33</td>
</tr>
<tr>
<td>LibreOffice 3.3.3_Linux</td>
<td>5</td>
<td>24,6</td>
<td>22,95</td>
<td>22,95</td>
<td>34,3</td>
<td>79,01</td>
<td>79,32</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>60,8</td>
</tr>
<tr>
<td>LibreOffice 3.4.1_Linux</td>
<td>6</td>
<td>24,6</td>
<td>22,95</td>
<td>22,95</td>
<td>34,3</td>
<td>79,01</td>
<td>79,32</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>60,8</td>
</tr>
<tr>
<td>OpenOffice.org 3.4.0</td>
<td>7</td>
<td>24,6</td>
<td>22,95</td>
<td>22,95</td>
<td>34,3</td>
<td>79,01</td>
<td>79,32</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>60,8</td>
</tr>
<tr>
<td>Microsoft Office 2007</td>
<td>9</td>
<td>24,3</td>
<td>22,71</td>
<td>22,71</td>
<td>33,02</td>
<td>56,29</td>
<td>56,33</td>
<td>60,8</td>
<td>60,8</td>
<td>60,8</td>
<td>100</td>
</tr>
<tr>
<td>Microsoft Office 2010</td>
<td>10</td>
<td>24,3</td>
<td>22,71</td>
<td>22,71</td>
<td>33,02</td>
<td>56,29</td>
<td>56,33</td>
<td>60,8</td>
<td>60,8</td>
<td>60,8</td>
<td>100</td>
</tr>
</tbody>
</table>
Talk overview

- Automated difference grading and visual inspection
- Roundtrip and print tests
- Document relevance
- Results of batch testing
- Automated bisection of interoperability errors
Difference grading

Four numeric error measures used:

- **Per page:**
  1) Line Number Difference (LND): missing lines, bad object position
  2) Text Height Error (THE): errors in line spacing, table or object shape etc

- **Per line:**
  1) Horizontal Line Position Error (HLPE): horizontal shift of dominating line segments
  2) Feature Distance Error (FDE): maximum distance between features of aligned lines

Grading:

- 0 ... 5 (pixel identical ... very bad), 6 (empty document), 7 (failed to open)

Textual report:

- automatically generated Calc spreadsheet
Observation: Maximum value of \( \text{abs}(\text{DF}(L1) - \text{DF}(L2)) \) is equal to a distance between different local features.

Characterizes difference on character level.
Motivated by Adam Fyne's talk in Milano

Print test:

Roundtrip test:
1) Documents printed to PDF and converted to bitmap
   - View 1, side-by-side
   - View 2, “Page overlay, no alignment”
2) Segmentation in lines
3) Vertical line alignment
   - Heuristics used in number of line does not match
   - View 3, “Page overlay, vertically aligned lines”
4) Horizontal line alignment
   - View 4, “Page overlay, vertically and horizontally aligned lines”
Side-by-side view

- Shows large differences
- Useful for complex layouts
- **Example:** Test File_672.docx, LO43:
  - Line Number Difference (LND): 0 (graded 0)
  - Text Height Error (THE): 4.57 mm (graded 3)
  - Feature Distance Error (FDE): 11.4 mm (graded 3)
  - Line Position Error (LPE): 11.43 mm (graded 5)

Source: Synthetic_Files/./File_672.pdf
PagePixelOverlayIndex[%]: 22.6 ; FeatureDistanceError[mm]: 11.4 ; HorzLinePositionError[mm]: 11.43 ; TextHeightError[mm]: 4.57 ; LineNumDifference: 0

a) Right Alignment
b) Right Alignment
c) Right Alignment
d) Right Alignment
e) Right Alignment

[Images and tables shown in the original document]
“Page overlay, no alignment” view

- A more detailed display of differences
- Different line spacing hides other differences
- **Example:** Test File_672.docx, LO43:

```
Page overlay, no alignment
cyan: source Synthetic_Files/.File_672.pdf
red: target BB43AL/.File_672.BB43AL.pdf
PagePixelOverlayIndex[%]: 22.6 : : TextHeightError[mm]: 4.57
```

```
a) Right Alignment
b) Right Alignment
c) Right Alignment
d) Right Alignment
e) Right Alignment
```
“Page overlay, vertically aligned lines” view

- Shows lines with differences
- **Example:** Test File_672.docx, LO43:

```
Page overlay, vertically aligned lines
cyan: source Synthetic_Files/./File_672.pdf
red: target BB43AL/./File_672.BB43AL.pdf
: HorizLinePositionError[mm]: 11.43
```

a) Right Alignment
b) Right Alignment
c) Right Alignment
d) Right Alignment
e) Right Alignment
“Page overlay, vertically and horizontally aligned lines” view

- Shows differences in lines
- **Example:** Test File_672.docx, LO43:

```
Page overlay, vertically and horizontally aligned lines
cyan: source Synthetic_Files=./File_672.pdf
red: target BB43AL=./File_672.BB43AL.pdf
FeatureDistanceError[mm]: 11.4

a) Right Alignment
b) Right Alignment
c) Right Alignment
d) Right Alignment
e) Right Alignment
```
Improvement since LO44

- Shows lines with differences
- **Example:** Test File_672.docx, LO44
- Horizontal position improved :-)

```
Page overlay, vertically and horizontally aligned lines
cyan: source Synthetic_Files/./File_672.pdf
red: target BB44DL/./File_672.BB44DL.pdf
FeatureDistanceError[mm]: 0.5
```

a) Right Alignment
b) **Right Alignment**
c) Right Alignment
d) **Right Alignment**
e) **Right Alignment**
Evaluation of LibreOffice – MS Office interoperability

- Shows lines with differences
- **Example:** Test File_672.docx, LO44
- Vertical position is the same :-(

```
Page overlay, no alignment
cyan: source Synthetic_Files/./File_672.pdf
red: target BB44DL/./File_672.BB44DL.pdf
PagePixel10OverlayIndex[%]: 27.5 ; TextHeightError[mm]: 4.70
```

a) Right Alignment
b) Right Alignment
c) Right Alignment
d) Right Alignment
e) Right Alignment
Demo: spreadsheet report

https://is.gd/idudac
Document relevance

- An idea: sort documents according to popularity of their features
- **Document relevance**: relative frequency of the least frequently used tag in a file, normalized to range (0,1)
- How to get:
  - Download numerous random documents from the internet
  - Extract tags
    - Simplify, if necessary (ignore numeric values etc)
  - Count, sort and normalize.
Results

- Tested 1600+ docx files
- Reference Word program: MSOffice 2010 running in Wine
- LO versions:
  - 10 versions available from the lo-linux-dbгутил-daily… bibisect git repositories
- Tests:
  - Roundtrip
  - Print
- Statistics:
  - 32000 pair comparisons (1600*10*2 = 32000)
  - two – three days computation
Demo: report for 1600+ tested documents

https://is.gd/ehukojoj

(a large spreadsheet, opens only in LO52 and later)
Observations

Overview, 1600 documents:

<table>
<thead>
<tr>
<th></th>
<th>regressions</th>
<th>progressions</th>
<th>good (grade &lt; 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>roundtrip</td>
<td>361</td>
<td>932</td>
<td>1078</td>
</tr>
<tr>
<td>print</td>
<td>731</td>
<td>676</td>
<td>583</td>
</tr>
</tbody>
</table>

Problems even in simple documents, most visible ones:

- Kerning handled incorrectly (hides other errors)
- Incorrect line spacing (40 % of all documents)

Shortcomings of evaluation

- Feature distance error sometimes overestimated
  - absolute now, relative would perhaps be better
- Binarized images compared, threshold 250 currently used
  - May cause false grading (charts and images)
- Maybe not all problems covered
Automated Bibisection

- Git provides tools for automated bisection
- A proper script needed returning good/bad status
- The procedure
  - Get a buggy document
  - Find the right bibisection repository (see the previous example)
  - Run the bisection script
  - ...
  - Submit a patch :)

Evaluation of LibreOffice – MS Office interoperability
Demo, local
Report with regressions: https://is.gd/gecuro
Summary

- We ban test:
  - Any office application with command line interface
  - Any office format:
    - odt, doc, docx, rtf (tested)
    - odp, doc, docx (not tested, should work)
    - ods, xls, xlsx (not tested, should work)
- To be done:
  - Classify test document according to their features
  - Future: add into the LO QA framework?
Thank you for watching!

- Code with example data is available at github
- Questions, comments, sugegstions: milos.sramek(at)soit.sk