PowerPoint Differ

Overview

PowerPoint Differ (PPD) is a tool for comparing a pair of PowerPoint (.pptx) files. The PPD can be used as a desktop application to generate a detailed comparison report. Alternatively, the PPD can be used as a library for making comparisons between a pair of pptx files. A use case for using the PPD as a library, is in an application that generates new or edits existing PowerPoint files and using the PPD as a library in unit testing for these generation or editing PowerPoint features.

GUI Quick Start

After starting up the PPD desktop application, click a Select PPTX to display an "Open As" type of dialog to locate and select a file for comparison. Perform this action of selecting a PowerPoint file for both File A and File B, as labeled in the main application window. With two files selected, click the "Compare" button to run the comparison operation. When the comparison is complete, a popup dialog is displayed presenting a detailed report with findings of differences between the two selected pptx files.

To select a PowerPoint file:

In the main application dialog, there are two buttons labeled, Select PPTX one button on the row for File A, and one on the row for File B. Click the Select PPTX button and use the Select a PowerPoint file dialog, an "Open As" type of dialog, to locate and select a PowerPoint (.pptx) file. When a file has been successfully selected, the file name and its parent folder will be displayed in the field under the File A or File B label. To see the full file path, hover the mouse pointer over the name of the file, and a tooltip will be displayed with the full path of the file. Repeat the action of selecting a PowerPoint file until both File A and File B have files selected and displayed. The Compare button will not become enabled until two PowerPoint files have been selected. The tool will only allow the selection of files with a ".pptx" suffix, not case sensitive. The tool will not allow selection of a file, if that file has already been selected in the other file field.

To compare two PowerPoint files:

With PowerPoint files selected for both File A and File B, click the Compare button. The PPD tool performs comparison operations for the two files. At present these checks include:

- Determine if the two files are copies of the same file or not, including document metadata.
- The number of slides in each presentation file.
- All the text parsed from the file parsed into a single string.
  - Some embedded graphics may not have text extracted.
  - All white space will be retrained in string.
- Slide Specific Data:
  - Slide Name
  - Slide Number
  - Slide Text
    - Text specific to an individual slide.
    - White space stripped out and words separated with a single new word space.
    - Text parsed from text boxes and tables.

On completion of the comparison operations, a pop-up dialog is opened displaying a report giving details about what checks where performed and the results of those checks.

To save a report file:

While the report dialog is displayed, click the Save Report button. A Save Report File dialog is displayed. In this dialog, navigate to a location to save the file, and specify a new name for the file if needed. Click the Save button to write the report to a text file to the specified location. Reports can also be generated as a string when using the PPD as a library by calling the "generateReport()" method on the differ object.

Library Quick Start

Instantiate a "Differ" object passing a pair of java.io.File objects to the Differ constructor. The first File object passed to the constructor is "File A" and the second File passed to the constructor is "File B". With the Differ object initialized, call methods on the Differ object to make comparisons on the files.

```java
@Test
public void exampleUsage() {
    File fileA = new File("fileA.pptx");
    File fileB = new File("fileB.pptx");

    try {
        Differ diff = new Differ(fileA, fileB);

        // we expect at least the meta-data in files to be different
        Assert.assertFalse(diff.isSameFile());
        // we expect the files to have the same number of slides
        Assert.assertEquals(diff.slideCount_fileA(), diff.slideCount_fileB());
        // parsed text from the PPTX file as the whole is the same
        Assert.assertEquals(diff.wholeFileText_FileA(), diff.wholeFileText_FileB());
        // metadata is not the same
        Assert.assertNotEquals(diff.metadata_FileA(), diff.metadata_FileB());

        // iterate over all slides in slide shows
        // if we got here, we know both files have the same slide count
        for (int i = 0; i < diff.slideCount_fileA(); i++) {
            // slide names in each file are expected to be the same
            Assert.assertEquals(diff.slideName_fileA(i), diff.slideName_fileB(i));
            // compare text parsed from slide
            Assert.assertEquals(diff.slideText_fileA(i), diff.slideText_fileB(i));
        }
    } catch (PpdException ex) {
        System.err.println(ex.getMessage());
    }
}
```

Notes:

- PPD built with Java 8 and Apache POI 5.0.0
- Older .ppt (Office 2003) files not supported.