

Extrait du SPIP de madbuilder sur free

<http://thierry.schmit.free.fr/spip>

mbtPdfAsm: the application

- Utilitaires - mbtPdfAsm -



Date de mise en ligne : Thursday 28 June 2007

Description :

From the setup to the use, all you always want to know about mbtPdfAsm

SPIP de madbuilder sur free

Summary

This tool acts of a in line tool for assembling/merging pdf files, extract informations form PDF files, update PDF files metadata.

In assembling mode (default mode), this tool concatenates pages, in full file mode, or in pages list mode. In oage list mode outlines are not concatenated. However this tool makes it possible to add outlines via a definition file of outlines (-o option).

The pages to be concatenated are extracted from valid pdf files via one or more masks of selection (-m option). It is also possible to extract certain pages from certain files via a script (-s option). If the options -m and -s are used simultaneously, the result file will comprise the result of the option -m followed of the result of the option -s.

By default, file are sorted in alphabetic order before assembling.

In extraction mode (-g[…]) (note informations, not data extractions) informations are printed on the standard output in a CSV format. The options -s, -d are inefficient.

In update mode (-u) the files matching the mask(s), are uptaded according to the command line options. The option -d is inefficient.

The command line

The syntax is as follow:

```
mbtPDFasm [-mM]mask -ddest [-a] [-as] [-b] [-z] [-r] [-onomFichier][--oO]
[-pListePages] [-snomFichier] [-c[RUOL]param] [-g[HACFKNST]] [-u[PK]]
[-lPx] [-S] [-ttexte] [-T[cfosxy]valeur] [-n] [-N[0cfsxy]valeur] [-Rvaleur]
```

mask: specify the REGULAR EXPRESSION (-m, perl compatible, pcre version 4.4) used for the research of the files to assembly. Several masks may be specified by separating them by a ';'. Note, under Linux it is better to use ','. You can use -M instead of -m. In this case the syntax of the mask is 'regular'. That is, * is for any number of any caracter. ? is for one an d only one caracter.

dest: specify the name of the result result of the assembly. If dest comprises a 'path', this one will not be created by the program.

Warning: If the destination file matches a mask, the destination file will be overwritten, but not assembled!

options:

- a: assemble file without sorting them.
 - s: files are assembled in an order corresponding to mask order. (version \geq 1.0.21)
- b: allow to specify a base directory from which the selection mask is applied. (version \geq 1.0.21).
- c: allow to handle acces restrictions (revisions 2 and 3 of the standard handler)
 - R: allow to define the restrictions
 - a: add annotations or forms
 - m: modify
 - p: print
 - s: select and extract Sample, the selector -cRamps in the command line apply all the restrictions. In the merging process, if a file having restriction is processed, and if this restrictions are more restrictive than those specified on the command line, the most restrictive restriction are used.
 - U: opening password
 - O: modifying restriction password
 - L: specifying the encryption key length. Only 5 (40 bits) and 16 (128 bits) are supported by Acrobat.
- g: allows to get an information form the file(s) matching the mask(s) (version \geq 1.0.5). The informations are printed on the standart output separated by a `\r\n`.
 - A: Author.
 - C: encryption informations (handler, key length, user password).
 - H: Header line for the displayed informations.
 - F: file name.
 - K: Keywords.
 - N: number of pages in the file.
 - O: the outlines of the files in the [format](#) defined for the application.
 - S: Subject.
 - T: Title.
- l: allows to limit th number of page(s) in the resultings pdf file(s). This commutator switch mbtPdfAsm un SPLIT mode. In this mode ths number of produced files is the number of pages corresponding to the pdf files selected by mask(s) or script divided by x, + 1. (version \geq 1.0.11).
 - Px: x allows to specify the number of pages by resulting file.
- n: insert a number from 0 in the assembled page (version \geq 1.0.17, experimental). **According to some properties of your PDF file, it is possible that some coordinates are invisible.**
- N: (EXPERIMENTAL version \geq 1.0.21) allows to format the numbering displayed with the -n option.
 - 0x: first number displayed.
 - c: color of the text, a string of three integer from [0, 255]. (0,0,0 by default)
 - f: font for display the text. (4 by default)
 - s: size of the text. (10 by default)
 - x: abscisse of the text. (10 by default)
 - y: ordonné of the text. (10 by default)
- o: allows to specify the name of a file of description of outlines.
 - oO: allow to keep the outlines (version \geq 1.0.9).
- p: allows to specify a list of pages to be extracted from the files matching the mask of the -m option. Such a list has the same syntax as a page list of a script file except that the space are replaced by ','.
- r: seek the files in under repertories.

- R: Rotate all the pages of the destination file. Value is 90, 180 or 270.
- s: allows to specify the name of a file of script. This name must not begin by A, K, S, or T. If the update mode is set on, this options is used to specify the metadata to be updated. In the specification the space must be substituted by '_'.
 - A: specify Author.
 - K: specify Keywords.
 - S: specify Subject.
 - T: specify Title.
- S: launch the application in silent mode, nothing is displayed. (version \geq 1.0.12)
- t: (EXPERIMENTAL version \geq 1.0.16). Allow to set a text that will be added on the bottom of the assembled pages. The used font is Helvetica 10. **According to some properties of your PDF file, it is possible that some coordinates are invisible.**
- T: (EXPERIMENTAL version \geq 1.0.16) allows to format the text displayed with the -t option.
 - c: color of the text, a string of three integer from [0, 255]. (0,0,0 by default)
 - f: font for display the text. (4 by default)
 - o: orientation of the text, a number comprised between 0° and 90°. (0 by default)
 - s: size of the text. (10 by default)
 - x: abscisse of the text. (10 by default)
 - y: ordonné of the text. (10 by default)
 - u: set update mode on (version \geq 1.0.6).
- P: erase the metadata data with no value specified on the command line (version \geq 1.0.8).
- K: erase the original file (version \geq 1.0.8), otherwise a pdfbak file is saved.
- z: opposite the order alpha of the files to be assembled.

Limitations

Historically mbtPdfAsm was created to quickly assemble simple PDF files (1 pages without options) on a server having to produce assembled PDF files. By heritage mbtPdfAsm thus does not seek to include/understand what it assembles. It is fast, but it is thus limited in what it is able to assemble, because certain assembly requires a reflexion, which it does not like to do because that slows down it.

However, in a concern of satisfying the greatest number I progressively increase what mbtPdfAsm is able to manage.

It however remains some of the more or less significant limitations according to your needs, and which technically can be summarized by saying that the following entries of Catalog are not processed: Version, PageLabels, ViewerPreferences, PageLayout, PageMode, Threads, OpenAction, AA, URI, Metadata, StructTreeRoot, MarkInfo, Lang, SiderInfo, OutputsIntents. The list is long and yet they are only options of which some cannot in any event of not being treated by a software of assembly, because of their treatment will appear conflicts between various PDF files to be assembled.

As a conclusion, and for the moment, mbtPdfAsm is appropriate to you perfectly if you wish to assemble files resulting from a scanner, from a software of production of the type cutePDF, pdfWriter, FPDF scripts …. mbtPdfAsm is appropriate moderately in the case of 'rough works finely outlined'.

Thank you for your comprehension, do not hesitate to make me share of your remarks, or to submit posing problems files to me.