NAME

spmmli−tool − clean luamml ghost characters and validate tagged PDF files

SYNOPSIS

spmmli−tool [*options*] *file*...
spmmli−tool −h | **−−help**
spmmli−tool −v | **−−version**

DESCRIPTION

spmmli−tool is a **texlua**(1) script distributed with the **spanish−mathml−intent** LaTeX package. It prepares the MathML produced by **luamml** for the associated-files mode of a tagged PDF, named **mathml−AF**, and checks that the tag structure and the PDF/UA-2 conformance of a document are correct.

The package **spanish−mathml−intent** can write invisible marker characters in the MathML that **luamml** produces. They are the character U+2063 with the attribute **sptmp** set to **outer** or **inner**. They protect the structure of the formulas, but must not remain in the MathML file that is attached to the PDF. **spmmli−tool** removes them from the HTML file written by **luamml**, which is named *name*−**luamml−mathml.html**, and leaves the result as *name*−**mathml.html**.

What **spmmli−tool** does with each *file* depends on its extension; see **MODES**. Several files may be given; they are processed in order.

MODES

.pdf The file is only validated; nothing is compiled and nothing is written. Two checks are run, each one only if its tools are found in **PATH**: the tag structure is dumped with **show-pdf-tags −−xml** and verified with **rnv-wrapp**, and the PDF/UA-2 conformance is verified with **verapdf −−flavour ua2.** The result is reported as **PASS** or **FAIL**, and the output of the failing validator is shown. If none of the checks can be run, the run ends with an error.

.html The file must be named *name*−**luamml−mathml.html**. The ghost characters are removed and the result is written as *name*−**mathml.html**.

.tex, **.ltx**
The document is compiled once with the TeX engine, in a private temporary directory, with the option **AFghost=true** passed to the package **spanish−mathml−intent**, so that **luamml** writes the ghost characters in the HTML file. The HTML file is then cleaned as in the **.html** mode. The PDF just produced is not validated; it stays in the temporary directory and is discarded, and only the cleaned HTML file is kept.

OPTIONS

−c, **−−check−only**
Do not compile anything; only validate PDF files. With this option every *file* must be a PDF file.

−o, **−−overwrite**
Replace the output file *name*−**mathml.html** if it already exists. Without this option **spmmli−tool** refuses to overwrite it and stops with an error. It has no effect together with **−c**, and a warning is printed.

−p, **−−path**=*path*
Write the output file in the directory *path*, which is created, with its parents, if it does not exist. The file is named after the document alone, without the directory of the input file. It has no effect together with **−c**, and a warning is printed. The forms **−p** *path* and **−−path** *path* are also accepted.

−e, **−−engine**=*cmd*
TeX engine used to compile the **.tex** and **.ltx** files. By default it is **lualatex−dev** if it is found in **PATH**, and **lualatex** otherwise. The name may contain only letters, digits, dots, underscores and hyphens.

−s, **−−strict**
Stop at the first failure, either of the compilation or of a validation. Without this option all the files are processed. In both cases the exit status is not zero if anything failed.

−h, **−−help**
: Show a summary of the usage and exit. This is also what happens when**spmmli−tool** is run without arguments from an interactive terminal.

−v, **−−version**
: Show the version number and exit.

−−
: End of options; the remaining arguments are file names.

FILE NAMES

If a *file* does not exist as given, its extension may be omitted: with **−c** the extension **.pdf** is tried; otherwise **.tex**, **.ltx**, **.pdf**, **−luamml−mathml.html** and **.html**, in this order.

For safety, a file name containing spaces, quotes or any of the characters **% & | < > ^ ‘ $ { } [ ] ! # ~** is rejected, since the name is passed to the shell and to TeX.

FILES

name−**mathml.html**
: Output file with the ghost characters removed. It is written next to the input file, or in the directory given with **−−path**. If no ghost character is found a warning is printed and the file is not kept; it usually means that **spanish−mathml−intent** was not loaded or that the HTML file was not produced with **AFghost=true**.

name−**luamml−mathml.html**
: Input of the **.html** mode; it is written by **luamml** and is never modified.

spmmli−tool−*XXXXXX*
: Temporary directory of the **.tex** and **.ltx** modes, created with mode 700 inside **$TMPDIR**. It is removed at the end. If the compilation fails it is kept, and its location is shown, so that the log file can be read.

ENVIRONMENT

TMPDIR
: Directory where the temporary directory is created on POSIX systems; the default is **/tmp**. On Windows **TEMP** or **TMP** is used instead.

EXIT STATUS

0
: All the files were processed and no compilation or validation failed. Finding no ghost character is only a warning.

1
: A file could not be found or is not valid, an option is wrong, an output file already exists and **−−overwrite** was not given, or a compilation or a validation failed.

EXAMPLES

Validate a PDF file, without compiling:

```
$ spmmli−tool −c sample.pdf
```

The same, with the extension omitted and stopping at the first failure:

```
$ spmmli−tool −c −s sample
```

Compile a document with the ghost characters enabled, write the cleaned MathML in the directory **build**, replacing the output file if it exists:

```
$ spmmli−tool −o −p build/ sample.tex
```

Clean an HTML file that was already written by **luamml**:

```
$ spmmli−tool sample−luamml−mathml.html
```

SEE ALSO

lualatex(1), **texlua**(1), **verapdf**(1)

The documentation of the **spanish−mathml−intent** package, which can be read with **texdoc spanish−mathml−intent**

REPORTING BUGS

Repository: https://github.com/pablgonz/spanish−mathml−intent
Bug tracker: https://github.com/pablgonz/spanish−mathml−intent/issues

AUTHOR

Pablo Gonzalez L <pablgonz@educarchile.cl>

COPYRIGHT

Copyright (C) 2026 Pablo Gonzalez L.