The `tex4ebook` package

Michal Hoftich*

Version v0.2c
2019-03-21

Contents

1 Introduction

`TeX4ebook` is bundle of Lua scripts and `LaTeX` packages for conversion of LaTeX files to ebook formats, for example `epub`, `mobi` and `epub3`. `tex4ht` is used as conversion engine.

Note that while `mobi` is supported by Amazon Kindle, most widespread ebook reader, it doesn't support `mathml` and this means that math must

---

*michal.h21@gmail.com

1

re represented as images. The same is true for `epub`. This is not a good thing, especially for inline math, as you may experience wrong baselines. If your ebook contains math, the only correct solution is to produce `epub3`, as it supports `mathml`. The issue with `epub3` is, that majority of e-ink ebook readers doesn't support this format, reader applications exists mainly for Android and Apple devices. For books which contains mainly prose, all formats should be suitable, but `epub3` supports most features from web standards, such as `CSS`.

As with `tex4ht`, the emphasis is on conversion of document's logical structure and metadata, basic visual appearance is preserved as well, but you should use custom configurations if you want to make the document more visually appealing. You can include custom `CSS` or fonts in configuration files.

1.1 License

Permission is granted to copy, distribute and/or modify this software under the terms of the LaTeX Project Public License, version 1.3.

2 Usage

Run on the command line:

```
tex4ebook [options] filename
```

You don't have to modify your source files unless you want to use commands defined by `tex4ebook` in the document, or when your document uses a package which causes a compilation error.

If you want to use `tex4ebook` commands, add this line to your document preamble:

```
\usepackage{tex4ebook}
```

But it is optional. You shouldn't need to modify your TEX files

2.1 Available commands

- `\coverimage{coverimage.name}` - include cover image to the document.

3 Command line options

`-c,--config` specify custom config file for `tex4ht`

example config file: File `sample.cfg`

```
\Preamble{xhtml}
\CutAt{section}
\begin{document}
\EndPreamble
```

 run

 tex4ebook -c sample filename.tex

This config file will create `xhtml` file for every section. Note that this behaviour is default.

-e,--build-file (`default nil`) Specify make4ht build file[1]. Defaulf build file filename is `filename.mk4`, use this option if you use different filename.

-f,--format (`default epub`) Output format. Epub, Epub3 and Mobi formats are supported.

-l,--lua Use LuaLaTeX as TeX engine.

-m,--mode (`default default`) This set `mode` variable, accessible in the build file. Default supported values are `default` and `draft`. In `draft` mode, document is compiled only once, instead of three times.

-r,--resolution Resolution of generated images, for example math. It should meet resolution of target devices, which is usually about 167 ppi.

-s,--shell-escape Enable shell escape in htlatex run. This may be needed if you run external commands from your source files.

-t,--tidy process output html files with `HTML tidy` command[2].

-x,--xetex Use xelatex for document compilation

-v,--version Print the version number.

4   Configuration

`tex4ebook` uses `tex4ht`[3] for conversion from LaTeX to html. `tex4ht` is highly configurable using config files. Basic config file structure is

---

[1]`https://github.com/michal-h21/make4ht#build-file`
[2]It needs to be installed separately
[3]`http://www.tug.org/tex4ht/`

```
\Preamble{xhtml, comma separated list of options}
...
\begin{document}
...
\EndPreamble
```

Basic info about command configurations can be found in a work-in-progres *tex4ht tutorial*[4], *tex4ht documentation*[5], and in series of blogposts on CV Radhakrishnan's blog: *Configure part 1*[6], *Configure part 2*[7], *Low level commands*[8]. Available options for `\Preamble` command are listed in the article *TeX4ht: options*[9]. *Comparison of tex4ebook and Pandoc output*[10]

Great source of tips for `tex4ht` configuring is *tex4ht tag on TeX.sx*[11], there is also a *tag for tex4ebook*[12].

Examples of interesting questions are *including images and fonts in ebooks*[13] or *setting image size in em units instead of pt*[14].

4.1 Provided configurations

`tex4ebook` provides some configurations for your usage:

```
\Configure{UniqueIdentifier}{identifier}
```

Every epub file should have unique identifier, like ISBN, DOI, URI etc. Default identifier is URI, with value `http://example.com/\jobname`.

```
\Configure{@author}{\let\footnote\@gobble}
```

Local definitions of commands used in the `\author` command. As contents of `\author` are used in XML files, it is necessary to strip away any information which don't belongs here, such as `\footnote`.

```
\Configure{OpfScheme}{URI}
```

Type of unique identifier, default type is URI. It is used only in `epub`, it is deprecated for `epub3`

```
\Configure{resettoclevels}{list of section types in descending order}
```

---

[4] https://github.com/michal-h21/helpers4ht/wiki/tex4ht-tutorial

[5] http://www.tug.org/applications/tex4ht/mn11.html

[6] http://www.cvr.cc/?p=323

[7] http://www.cvr.cc/?p=362

[8] http://www.cvr.cc/?p=482

[9] http://www.cvr.cc/?p=504

[10] https://github.com/richelbilderbeek/travis_tex_to_epub_example_1

[11] http://tex.stackexchange.com/questions/tagged/tex4ht

[12] http://tex.stackexchange.com/questions/tagged/tex4ebook

[13] http://tex.stackexchange.com/a/213165/2891

[14] http://tex.stackexchange.com/a/195718/2891

Configure section types which should be included in the `NCX` file. Default value is the whole document hierarchy, from `\part` to `\paragraph`.

`\Configure{DocumentLanguage}{language code}`

Each ePub file must declare the document language. It is inferred from `babel` main language by default, but you can configure it when it doesn't work correctly. The `language code` should be in ISO 639-1 form.

`\Configure{CoverImage}{before cover image}{after cover image}`

By default, cover image is inserted in `<div class="cover-image">` element, you may use this configuration option to insert different markup, or even to place the cover image to standalone page.

`\Configure{CoverMimeType}{mime type of cover image}`

Default value is `image/png`, change this value if you use other image type than `png`.

If you don't want to include the cover image in the document, use command

`\CoverMetadata{filename}`

in the config file.

`\Configure{OpfMetadata}{item element}`

Add item to `<metadata>` section in the `OPF` file.

`\Configure{OpfManifest}{maifest element}`

Add item to `<manifest>` section in the `OPF` file.

`\Configure{xmlns}{prefix}{uri}`

Add xml name space to `xhtml` files. Useful in `EPUB 3`

4.2   Commands available in config files

`\OpfRegisterFile[filename]` register file in the `OPF` file. Current output file is added by default.

`\OpfAddProperty{property type}` add `EPUB3` property for the current file. See *EPUB3 spec*[15]

`\OpfGuide[filename]{title}{type}` Add file to the `<guide>` section in the OPF file. See *Where do you start an ePUB and what is the <guide> section of the .OPF file?*[16] for some details. Note that `<guide>` is deprecated in EPUB 3.

---

[15]http://www.idpf.org/epub/301/spec/epub-publications.html#sec-item-property-values

[16]http://epubsecrets.com/where-do-you-start-an-epub-and-what-is-the-guide-section-of-the-opf-file.php

4.3 Build files

`tex4ebook` uses `make4ht`[17] as a build system. See `make4ht` documentation for details on build files.

4.4 `.tex4ebook` **configuration file**

It is possible to globally modify the build settings using the configuration file. New compilation commands can be added, extensions can be loaded or disabled and settings can be set.

4.4.1 Location

The configuration file can be saved either in `$HOME/.config/tex4ebook/config.lua` or in `.tex4ebook` in the current directory or it's parents (up to `$HOME`).

See the `make4ht` documentation for an example and more information.

5 Troubleshooting

When compilation of the document breaks with error during `LaTeX` run, it may be caused by some problem in `tex4ht` configuration. Comment out line `\usepackage{tex4ebook}` in your source file and run command:

```
htlatex filename
```

if same error as in `tex4ebook` run arises, the problem is in some `tex4ht` configuration. Try to identify the source of problem and if you cannot find the solution, make minimal example showing the error and ask for help either on *tex4ht mailing list*[18] or on *TeX.sx*[19].

5.1 Validation

In case of successful compilation, use command line tool `epubcheck`[20] to check whether your document doesn't contain any errors.

Type

```
epubcheck filename.epub
```

---

[17]`https://github.com/michal-h21/make4ht`
[18]`http://tug.org/mailman/listinfo/tex4ht`
[19]`http://tex.stackexchange.com/`
[20]you need to install it separately, see `https://github.com/IDPF/epubcheck`

5.1.1 Common validation issues:

- WARNING: filename.epub: item (OEBPS/foo.boo) exists in the zip file, but is not declared in the OPF file

Delete the `filename-(epub|epub3|mobi)` folder and `filename.epub`. Then run `tex4ebook` again.

- WARNING(ACC-009): hsmmt10t.epub/OEBPS/hsmmt10tch17.xhtml(235,15): MathML should either have an alt text attribute or annotation-xml child element.

  This is accessibility message. Unless you use some macro with annotations for each math instance, you will get lot of these messages. Try to use `epubcheck -e` to print only serious errors.

6 Changes

- 2019/03/21

  - released version `0.2c`

- 2019/03/07

  - use `Luatexbase` package in the documentation because of `Microtype` error
  - remove `<guide>` element even in ePub 2
  - added `encoding` attribute to XML declaration in the NCX and OPF files
  - clean the temporary directory (`filename-format`) before file packing

- 2019/01/21

  - added `\Configure{@author}`

- 2019/01/10

  - released version `0.2b`

- 28/11/2018

  - added support for appendix sections to the NCX file

- 27/11/2018

  - use the `uni-html4` option by default. It will convert some math characters as Unicode chars

- 13/11/2018

  - added –xetex option to the README

- 30/10/2018

  - use the original section numbering in TOC in the Epub 3 output. The numbering of the `<ol>` list is disabled by CSS.

- 18/10/2018

  - fixed the executable installation

- 03/09/2018

  - updated the `--help` message

- 30/08/2018

  - removed spurious `\NoFonts` command in the footnote configuration, it caused formating issues in the document following a footnote

- 22/06/2018

  - added support for the output directory selection

- 09/05/2018

  - added support for Polyglossia language codes
  - released version `0.2a`

- 03/05/2018

  - fixed output format handling

- 16/04/2018

  - don't run Git if the Makefile is executed outside of Git repo

- 09/04/2018

  - released version `0.2`

- 06/04/2018

  - documented the configuration file

- 02/03/2018

  - added support for `.tex4ebook` configuration file

- 28/02/2018
    - added media overlays handling
- 19/10/2017
    - added support for XeTeX
- 06/10/2017
    - added support for Make4ht extensions
- 27/04/2017 Version 0.1e
    - pack the accumulated changes for distribution
    - set version number from git tag
- 17/01/2017
    - process duplicate images only once
    - check the OPF table for duplicated id attributes
- 11/01/2017
    - pass `settings` table to the build file.
- 19/12/2016
    - new configuration: \Configure{resettoclevels}{list of sectioning levels to be included in the NCX}
- 31/10/2016
    - added tidyconf.conf to the Makefile
- 22/10/2016
    - added support for new Make4ht command `Make:add_file`
- 22/08/2016
    - fixed incorrect `<dc:creator>` generated by \author
    - clean ids in the OPF file if they contain invalid characters at the beginning
- 31/07/2016
    - use monospace font for command line options in order to prevent double hyphens to become dashes by ligaturing

- 22/07/2016

  - removed debugging message when `--lua` option is used

- 07/04/2016

  - Cut `\part` commands to standalone pages

- 05/04/2016

  - bug fix: pages which contained math were arranged before other pages

- 31/03/2016 Version 0.1d

  - bug fix: Unicode wasn't used by default
  - bug fix: Formats weren't preserved

- 06/12/2015 Version 0.1c

  - added `--help` and `--version` command line options

- 25/11/2015

  - added missing language codes
  - added `\Configure{DocumentLanguage}` for the cases when document language inferencing doesn't work

- 17/11/2015

  - use mkparams for cli arguments handling.
  - fixed inconsistencies between recent changes in make4ht and tex4ebook

- 24/09/2015 Version 0.1b

  - info about new workaround for `fontspec` package

- 25/08/2015

  - fatal error happened with epub3 when the document didn't contain the TOC

- 23/08/2015

  - simple cleaning of the ncx file if tidy command isn't available

- 05/07/2015 Version 0.1a

- `tex4ebook` script was missing in the distribution zip file
  - all links moved to footnotes in the documentation

- 29/06/2015 Version 0.1

  - fixes in documentation

- 22/06/2015

  - changes moved from README.md to CHANGELOG.md

- 18/06/2015

  - replaced `--mathml` option with `--mode`. For `mathml` support, use `mathml` option for `tex4ht.sty`.
  - lot of stuff was fixed in `epub3` support.
  - new command `\OpfAddProperty`

- 14/01/2015

```
- thanks Volker Gottwald for guide on [installing and using
  tex4ebook](https://d800fotos.wordpress.com/2015/01/19/create-
e-books-from-latex-tex-files-ebook-aus-latex-tex-dateien-erstellen/)
  on Windows
```

- 23/11/2014

```
- added new command `\OpfGuide`, for adding items to `<guide>`
  section in the `opf` file. This is useful for `epub` and `mobi` formats.

  Usage:

      \OpfGuide[filename]{title}{reference type}

  `filename` is optional, current file name is used when empty. See
  [epub
  secrets](http://epubsecrets.com/where-do-you-start-an-epub-
and-what-is-the-guide-section-of-the-opf-file.php)
  article
```

- 20/10/2014

```
 - fixed issues with starred sections

 - files created with starred sectioning commands (`\chapter*`,
   `\section*`) should be included in correct reading order now
```

- 16/09/2014

  - new features added
  - new configuration file for `--tidy` option, mathml and html5 elements
    are supported. This means that many validation errors in `mathml`
    output can be fixed with `--tidy` option
  - added configuration for all languages supported by `babel`
  - `woff` and `ttf` fonts are supported
  - added inline footlines in `epub3` format
  - added `no-cut` command line option for breaking sections and chapters
    into standalone pages
  - Please support [iniciative for improving mathml
    support](http://www.ulule.com/mathematics-ebooks/) in Gecko and Webkit
    engines. This will hopefully improve also Epub3 readers.

- 10/08/2013

  - [`make4ht`](https://github.com/michal-h21/make4ht) is now standalone
    application which `tex4ebook` depends on. You must [install

it](https://github.com/michal-h21/make4ht#instalation) in order to use
current tex4ebook version.