

Author Information: How to Avoid Common Conversion Problems \LaTeX -XML

Stefan M. Moser

10 August 2016

Abstract

This document is directed to authors of the *IEEE Transactions on Information Theory* who are preparing the final version of their paper to be sent to IEEE for publication. At IEEE, the submitted source file will be converted to XML. This document tries to provide instructions that hopefully will lead to a smoother conversion process without many conversion errors.

Contents

| | | |
|----------|--|-----------|
| 1 | Source File | 2 |
| 1.1 | Completeness | 2 |
| 1.2 | Style | 2 |
| 1.3 | Author Names | 3 |
| 1.4 | Abstract and Keywords | 3 |
| 1.5 | Bibliography | 4 |
| 1.6 | Author Biographies | 5 |
| 2 | \LaTeX: What to Do and What to Avoid | 5 |
| 2.1 | \LaTeX Packages | 6 |
| 2.2 | Programming Style | 6 |
| 2.3 | Font Issues | 7 |
| 2.4 | Color | 8 |
| 2.5 | picture-Environment and TikZ | 8 |
| 3 | \LaTeX and Math | 8 |
| 4 | Graphics | 9 |
| 4.1 | Format | 9 |
| 4.2 | Preparation of the Additional Graphic Files | 9 |
| 4.3 | Naming | 10 |
| 5 | Common Issues in Galley-Proofs | 10 |
| 6 | Acknowledgments | 12 |

1 Source File

While it is possible to submit the final manuscript as a Microsoft Word document, we discourage this and strongly recommend to provide a \LaTeX source file. The following information refers mainly to \LaTeX source files.

1.1 Completeness

The source file needs to contain *all* information (beside the source file, only additional graphic files are to be submitted; for more about these graphic files, see Section 4). This means that the author needs to upload *one* source file that contains

- the complete paper including title, author names, abstract, keywords, all sections and possibly appendices (see Sections 1.3 and 1.4);
- the bibliography with all references (see Section 1.5);
- all personal \LaTeX definitions (macros) (see also Section 2.1); and
- short biographies of all authors (see Section 1.6).

1.2 Style

To generate a correct layout and style, the \LaTeX source should use the `IEEEtran` class:

```
\documentclass[twoside]{IEEEtran}
```

The class definition file `IEEEtran.cls` is by default available in any up-to-date \LaTeX distribution. However, please make sure to use the current official version of the file and not an obsolete old version or even one of the many “hacked” or customized versions that can be found in many places on the web. The current official version is **Version 1.8b from 2015/08/26** and can be found online at <http://www.ctan.org/tex-archive/macros/latex/contrib/IEEEtran/>

The `IEEEtran` class will provide the correct formatting and style for IEEE publications, in particular, it generates a **double-column** format. Note that the \LaTeX source should not be submitted in single-column format, because the conversion to double-column format will cause many problems that are best handled by the authors themselves. For example, all equations need to be newly typeset with different and new line-breaks to adapt to the shorter column width (see also Section 3), and the size of the figures need to be adapted (or the figure is changed to span both columns using `\begin{figure*}`).

In addition to the option `twoside`, the option `a4paper` can be specified if the source will be printed on A4 (the layout is not changed, only the margins are adapted).

More information about the `IEEEtran` class and its options and commands can be found in [1].

1.3 Author Names

If an author is member of IEEE, his or her name should be followed by the corresponding IEEE membership grade. Affiliations and references to financial support can be given as a `\thanks{}` comment. We provide the following example:

```
\author{Gu-Rong~Lin,~\IEEEmembership{Student Member,~IEEE} and
  Stefan~M.~Moser,~\IEEEmembership{Senior Member,~IEEE}
  \thanks{
    This work was supported in part by the National Science Council,
    Taiwan, under NSC~100-2628-E-009-003.

    S.~M.~Moser is with ETH Zurich, 8092 Z\"urich, Switzerland, and
    also with National Chiao Tung University, Hsinchu 300, Taiwan
    (e-mail: xxx).

    G.-R. Lin is with National Chiao Tung University, Hsinchu 300,
    Taiwan (e-mail: xxx).
  }}

```

IEEE allows authors to state their name beside the usual English also in Chinese (traditional or simplified), Japanese or Korean. The characters must be provided by CJK ASCII Unicode. For example:

```
\usepackage[encapsulated]{CJK}
\usepackage{ucs}
\usepackage[utf8x]{inputenc}
\newcommand{\cjktxt}[1]{\protect\begin{CJK*}{UTF8}{bsmi}#1\end{CJK*}}

\author{\cjktxt{莫詩台方}
  (Stefan~M.~Moser),~\IEEEmembership{Senior Member,~IEEE}
}

```

The authors are asked to inform IEEE when they submit a source file with native language author names in Unicode characters.

1.4 Abstract and Keywords

The abstract and the keywords should be provided using the following environments:

```
\begin{abstract}
  This paper ...
\end{abstract}

\begin{IEEEkeywords}
  Keyword 1, keyword 2, ...
\end{IEEEkeywords}

```

IEEE provides a list of official keywords¹, but other keywords specified by the authors are also acceptable. Note that the keywords need to be ordered alphabetically.

1.5 Bibliography

The bibliography needs to be specified using the `thebibliography`-environment:

```
\begin{thebibliography}{1}
\bibitem{lapidothmoser03}
A.~Lapidoth and S.~M. Moser, ‘‘Capacity bounds via duality with
applications to multiple-antenna systems on flat fading channels,’’
\emph{IEEE Trans.~Inf.~Theory}, vol.~49, no.~10, pp. 2426--2467,
Oct.~2003.

\bibitem{...}
\end{thebibliography}
```

where the argument `{1}` of `\begin{thebibliography}{1}` specifies the number of references.²

It is strongly recommended, however, not to typeset the bibliography manually, but to rely on the tool BiBTeX that will automatically create the correct IEEE-style formatting. The procedure is as follows. First, the following two directives have to be included in the L^AT_EX source:

```
\bibliographystyle{IEEEtran}
\bibliography{referencefile}
```

Here, `IEEEtran` specifies the style in which the references are formatted. This style is defined in a file called `IEEEtran.bst` and is by default available in any up-to-date L^AT_EX distribution. The current version is Version 1.14 from 2015/08/26 and can be found online at <http://www.ctan.org/tex-archive/macros/latex/contrib/IEEEtran/bibtex>. Furthermore, it is assumed that the BiBTeX database is provided by a file called

```
referencefile.bib
```

For more details on how this file needs to look like, we refer to the manual available from http://mirrors.ctan.org/macros/latex/contrib/IEEEtran/bibtex/IEEEtran.bst_HOWTO.pdf

Second, the source has to be processed both by L^AT_EX and BiBTeX.³ Finally, the BiBTeX directives must be commented out:

¹To obtain it, just send a blank email to keywords@ieee.org.

²This number is only needed for spacing reasons, i.e., if the number of references is less than 10, the number should be single-digit (e.g., 1); if the number of references is 10 or more, the number should be double-digit (e.g., 10); or if the number of references is 100 or more, the number should be triple-digit. The exact value is irrelevant.

³Usually, the correct compilation order is `pdflatex`, `bibtex`, `pdflatex`, `pdflatex`.

```
%\bibliographystyle{IEEEtran}
%\bibliography{referencefile}
```

and the contents of the generated .bbl-file (which consists of the `thebibliography`-environment) must be copy-pasted into the source file.⁴

`IEEEtran` by default abbreviates first names. Unfortunately, the ever growing number of publications make it more difficult to exactly specify certain papers, in particular if some of the authors have short and common family names like “Li” or “Wang”. If you would like to have full names in the references, include the following definition into your `referencefile.bib` file:

```
@IEEEtranBSTCTL{IEEEtran:fullfirstnames,
  CTLname_format_string =      "{ff~}{vv~}{ll}{, jj}"
}
```

and then add the following line right after `\begin{document}` of your source file:

```
\bstctlcite{IEEEtran:fullfirstnames}
```

1.6 Author Biographies

The biographies of every author (in the same order as given below the title) should be added at the very end of the source file in the following form:

```
\begin{IEEEbiographynophoto}{Stefan M.~Moser}
  (S'01--M'05--SM'10) received the diploma (M.Sc.) in electrical
  engineering in 1999...
\end{IEEEbiographynophoto}
```

Note the specification of the time when the author has reached the different IEEE membership grades (student member “S”, member “M”, senior member “SM”, fellow “F”, life fellow “LF”). If an author is not an IEEE member, this time specification can be omitted.

2 L^AT_EX: What to Do and What to Avoid

As mentioned, the L^AT_EX source file submitted by the authors will be converted to XML. In that process, all L^AT_EX directives need to be translated into corresponding XML constructs and the behavior of L^AT_EX packages needs to be mimicked. In general this works quite well, however, some special L^AT_EX commands and L^AT_EX packages can cause problems.

⁴This last step is only required because the source file must contain *everything* including the references (see Section 1.1).

2.1 L^AT_EX Packages

While any personal L^AT_EX style file must be copy-pasted into the source file, many standard L^AT_EX packages need not be included in the source, but can simply be referred to by using the `\usepackage` directive.

Unfortunately, there are a couple of packages that will cause problems during the conversion process of the L^AT_EX source into XML. In particular, the following packages *should be avoided*:

- `hyperref.sty` (since IEEE does not provide hyperlinks, this package is not needed anyway)
- `arydshln.sty`
- the options `vlined` or `ruled` for the package `algorithm2e.sty`
- `acronym.sty`
- `amscd.sty`

There are some more packages related to fonts that should be avoided, see Section 2.3 for more details.

2.2 Programming Style

It is a general platitude for any kind of programming that the better a source code is human readable, the fewer errors it contains. A clean source code will also help the conversion process from L^AT_EX to XML. We therefore list here some basic rules for the programming style:

- Do not use overly long lines, but introduce line-breaks. This will prevent automatic line-breaks at wrong positions during the conversion process.
- Do not redefine fundamental parts of L^AT_EX. For example, it is bad style and can cause serious troubles if one redefines

```
\renewcommand{\left}{\left(}
\renewcommand{\right}{\right)}
```

or similar.

- Do only use letters, numerals and punctuation signs as label names. In particular, do not use brackets within labels. E.g., `\label{eq:low2}` or `\label{item:gg_3}` are OK, but `\label{thm:I{a}}` is not recommended (even though technically it is allowed in L^AT_EX).
- If possible, try to avoid nesting definitions of expressions that can be used both in text- and in math-mode. For example,

```
\newcommand{\G}[1]{\ensuremath{\mathcal{G}(\#1)}}
```

can be used both as `\G{1}` and `$$\G{1}$$` (both resulting in “ $\mathcal{G}(1)$ ”). However, again, the XML conversion process can be confused by such commands. Define instead

```
\newcommand{\G}[1]{\mathcal{G}(\#1)}
```

and always refer to it as `$$\G{1}$$`.

2.3 Font Issues

A special nuisance in the conversion from L^AT_EX to XML are unwanted changes in the used fonts. Particularly, if two different mathematical symbols are based on the same letter, but in different fonts, then a change of font of one of these symbols might cause a clash that fundamentally alters the meaning of mathematical expressions.

As an example consider the case where `E` (`\mathsf{E}`) denotes expectation, while `E` (`\eulerrm{E}`) stands for energy. XML does not provide Euler Roman, so `E` will be changed. If this change turns out to be `\mathsf`, we end up with a second `E` of completely different meaning and with very weird looking expressions like `E[E]`.

It is therefore crucial to make sure that only symbols and fonts are used that are also available on the XML side.

In short, the following fonts are available in XML:

| | | |
|------------------------------|-------------------------------|-----------------------------|
| <code>\textrm:</code> | A and a | (default font in text-mode) |
| <code>\textsf:</code> | A and a | |
| <code>\texttt:</code> | A and a | |
| <code>\textsc:</code> | A AND A | |
| <code>\textbf:</code> | A and a | |
| <code>\textit:</code> | <i>A and a</i> | |
| <code>\textit\textbf:</code> | <i>A and a</i> | |
| <code>\mathit:</code> | <i>A + a</i> | (default font in math-mode) |
| <code>\mathrm:</code> | A + a | |
| <code>\mathsf:</code> | A + a | |
| <code>\mathtt:</code> | A + a | |
| <code>\mathbf:</code> | A + a | |
| <code>\boldsymbol:</code> | A + a | |
| <code>\mathbb:</code> | \mathbb{A} | (no lowercase) |
| <code>\mathfrak:</code> | $\mathfrak{A} + \mathfrak{a}$ | |
| <code>\mathcal:</code> | \mathcal{A} | (no lowercase) |
| <code>\mathscr:</code> | \mathscr{A} | (no lowercase) |

These commands require the following packages:⁵

```
\usepackage{amsmath,amsfonts,amssymb,mathrsfs}
```

i.e., these packages are OK.

Any other font is not available in XML and will be changed into one of the fonts given above. In particular, the following fonts and packages *cannot* be used:

- Euler Roman and Euler bold Roman are not available.
- Euler Script will be changed⁶ into `\mathcal`.
- The font defined by `stmaryrd.sty` cannot be used.
- The symbols defined by `pifont.sty` (command `\ding{.}`) cannot be used.
- No ASCII symbols above 127 can be used (use `amssymb.sty` for special characters instead).

2.4 Color

Color cannot be used apart in graphics. However, be aware that many people will print your paper black-and-white, i.e., when you use colors in a figure, try to make sure that the figure remains readable also in gray tones.

2.5 picture-Environment and TikZ

Graphical drawings using the `picture`-environment, the `tikzpicture`-environment, or similar must be embedded within a figure (i.e., within the `figure`-environment). It is therefore also necessary to create a separate graphic file for it, see Section 4.2 for more details.

3 L^AT_EX and Math

Since the paper is in double-column format, the space for mathematical equations is quite narrow. Therefore, clear and efficient line-breaking in longer equations becomes crucial. The `IEEEtran` class provides a very powerful tool to typeset equations: the `IEEEeqnarray`-environment. For details on its use and for many hints on good formatting of equations see [2].

⁵The package `amsfonts.sty` is actually implicitly loaded by `amsmath.sty`. The commands `\mathbb` and `\mathfrak` require `amssymb.sty`, and `\mathscr` requires `mathrsfs.sty`.

⁶This change might be OK as long as `\mathcal` has not been used for other symbols, but in this case it is better to typeset these characters in `\mathcal` directly and to avoid Euler Script.

4 Graphics

We mentioned in Section 1 that the source file must be a single file that contains *all* information. The only exception to this rule are graphics that are imported by L^AT_EX from external files and that should be submitted as additional files.

Strictly speaking, not only an external graphic needs to be submitted as an additional graphic file, but also a graphic that is generated within the L^AT_EX source using a powerful graphics engine like TikZ. To be precise, by “graphic” we understand the following:

- Every floating figure (`figure-` or `figure*`-environment) that imports an external graphic.
- Every floating figure (`figure-` or `figure*`-environment) that does not import an external graphic, but creates the graphic directly using, e.g., the `picture`-environment or the `tikzpicture`-environment (see Section 2.5). However, note that an equation spanning both columns (that needs to be created with the help of a `figure*`-environment, see [2]) does *not* require such a separate graphic file.
- Any other part of the manuscript that is not generated by L^AT_EX, but that is imported from an external file.⁷

Note that the caption of a figure does not belong to the graphic and therefore must *not* be included in the corresponding graphic file.

4.1 Format

The preferred graphic format is **PDF** or **TIFF**. Also acceptable⁸ is **EPS**. Any other format is not accepted, in particular, GIF or JPG cannot be used.

4.2 Preparation of the Additional Graphic Files

The additional graphic file must look exactly the way the graphic is supposed to appear in the published manuscript. This means that, for example, any text-replacement by `\psfrag` must have taken place already. Ideally, it also has the correct size so that no additional resizing is required. The caption of a figure, however, is not considered to be part of the graphic and must not be included in the additional graphic file.

There is a very easy way of achieving exactly this: Make a duplicate of your finished L^AT_EX-sourcefile (e.g., called `extract.tex`) and just before the `\begin{document}` add the following two lines:

⁷We would like to point out that we strongly advise against external files that are imported rigidly without the use of a floating figure, since such external parts are usually almost impossible to position nicely within the document.

⁸If the source is a Microsoft Office document, those native formats DOC, DOCX, PPT, PPTX, XLS, and XLSX are also accepted. But once again, we strongly discourage from using Microsoft Office, but instead recommend the use of L^AT_EX.

```
\usepackage[active,tightpage,floats]{preview}
\renewcommand{\caption}[2] [] {}
```

If you now compile this file `extract.tex` you will receive a document containing only the figures and tables that are defined in the source file and having them arranged one per page, without caption, and cropped to their actual size. You can now extract page by page and save it using a name according to the naming rules of Section 4.3.

Be aware that any references (using `\ref{...}` or `\eqref{...}`) or citations (using `\cite{...}`) used within the graphics will not show correctly this way. This can be fixed using the `xr`-package. Add the following two lines to `extract.tex`:

```
\usepackage{xr}
\externaldocument{sourcefile}
```

where `sourcefile` needs to be replaced by the name of your sourcefile. This will make sure that any references are taken from your original source file⁹ and not from `extract.tex`.

4.3 Naming

The different graphic files should be numbered sequentially according to the order of their appearance in the manuscript. We recommend to use the following naming convention:

- take the first 5 letters of the first author’s last name, and
- append the graphic’s corresponding sequential number.

For example, in a paper with first author Anderson, the first three figures would be named `ander1.pdf`, `ander2.tif`, and `ander3.pdf`.

All these graphic files should then be combined into a single `zip` or `tar` archive file that is submitted alongside the source file.

5 Common Issues in Galley-Proofs

Once the \LaTeX source has been converted, the authors will receive a galley-proof of their manuscript. It is recommended that this galley-proof is read very carefully, because even if the authors did follow the hints and instructions given above, it is still possible that it contains errors and problems.

Some of the most common issues are as follows:

- Spacing issues concerning mathematical symbols that have a double-superscript or a combination of a hat/tilde or similar and of a superscript. For example,

$$\hat{X}^n$$

⁹Actually, it is taken from the auxiliary file `sourcefile.aux`, which must therefore be available and up-to-date.

is wrongly printed as

$$\hat{X}^n$$

Note the n being too high up.

- Spacing issues concerning mathematical symbols that have both a super- and a subscript. For example,

$$\mathbf{H}_{a+1}^\top$$

is wrongly printed as

$$\mathbf{H}_{a+1}^\top$$

Note that the superscript shows up *outside* of the subscript and how the transpose- \top is too big and too high up. This can be particularly ugly if the expression is inside of brackets, which in the wrongly spaced form become too small:

$$h(\mathbf{H}_{a+1}^\top) \quad \text{versus} \quad h(\mathbf{H}_{a+1}^\top)$$

- Spacing issues concerning mathematical expressions within the text. For example, the transpose- \top in $(\mathbf{A}^\top \text{diag}(\hat{\mathbf{H}})\mathbf{A})^\top$ overlaps with the text of the line above.¹⁰
- Spacing issues between operators and brackets. For example,

$$\log \det (\mathbf{I} + \mathbf{A}^\top)$$

is wrongly printed as

$$\log \det (\mathbf{I} + \mathbf{A}^\top)$$

Note the too large gap between \det and the opening bracket.

- Spacing issues of equations or arrays of equations that are vertically not properly centered (spaces above and below the equation are not the same).
- Discrepancies between sizes of opening and closing brackets if there is a line-break in between. For example,

$$a = \log \left(1 + \frac{b}{2} \right)$$

is wrongly printed as

$$a = \log \left(1 + \frac{b}{2} \right)$$

¹⁰Note that this example is artificial because the transpose- \top should be made smaller in any case.

So, when proof-reading their manuscript, the authors should particularly look out for such issues and mark them.¹¹ Note that if the authors react quickly¹² and provide the feedback to the galley-proof in short time, then they are allowed to ask for several rounds of galley-proofs and corrections, until they are satisfied with the produced manuscript.

6 Acknowledgments

I would like to thank Kevin Lisankie, Helmut Bölcskei, and Frank Kschischang for their very valuable input that was the basis of this document.

Any feedback or comments are highly appreciated! Write to

`stefan.moser@alumni.ethz.ch`

Thank you!

Stefan M. Moser

References

- [1] Michael Shell, “How to use the IEEEtran L^AT_EX class,” *Journal of L^AT_EX Class Files*, vol. 14, no. 8, August 2015. Available: http://mirrors.ctan.org/macros/latex/contrib/IEEEtran/IEEEtran_HOWTO.pdf
- [2] Stefan M. Moser, *How to Typeset Equations in L^AT_EX*, version 4.5, June 21, 2016. Available: <http://moser-isi.ethz.ch/manuals.html#eqlatex>

¹¹The comments about the galley-proof can be given in several forms. The authors can write them by hand onto the proof and then scan the commented version; they can use a tool that allows commenting and marking of a PDF file; or they can give a list of issues describing carefully the exact location and the issue.

¹²The time between the notification of the first galley-proof and the publication date is limited and relatively short.