Guidelines for authors submitting papers to the Journal of Language Modelling

Marcin Woliński
Institute of Computer Science Polish Academy of Sciences

This document summarises editorial and technical requirements for articles published in the Journal of Language Modelling. The journal is typeset using a Unicode and OpenType aware LaTeX variant, namely XeLaTeX, and a custom document class `jlm.cls`.[1]

1 INSTALLATION

The JLM distribution package includes fonts (*Charis SIL* by SIL International, *Playfair Display* by Claus Eggers Sørensen, and *Cousine* by Steve Matteson; all OFL licenced), a LaTeX document class `jlm.cls`, a bibTeX style file `jlm.bst`, and a few auxiliary files with Creative Commons logos.

Since XeTeX is used one needs a relatively new TeX installation for processing JLM articles. We strongly recommend at least 2012 versions of TeXlive or mikTeX. The TeX installation has to include XeTeX, LaTeX3 packages (`l3kernel`, `l3packages`), and the package `mathdesign`.

The installation process comprises two necessary steps:

1. INSTALLATION OF FONTS which depends on the operating system used:
   (a) *On Windows*: Select all the font files in the `fonts/` folder of the JLM distribution package, click the right mouse button, and choose *Install...*

[1] The layout of JLM was designed by Adam Twardoch and implemented as a LaTeX class by Marcin Woliński.

(b) *On Linux*: Copy the contents of the `fonts/` directory from the distribution package to `~/.fonts/`, that is to the directory named `.fonts` (with a leading dot) placed in your home directory. If you have ever installed any fonts locally, the directory already exists. Otherwise you have to create it first. You can place JLM fonts directly in `.fonts/` or in a subdirectory of your choice.

(c) *On Mac OS*: Launch the Font Book application located in the Applications folder, create a new font collection (e.g., JLM Fonts) by choosing *New Collection* from *File* menu, and finally add the fonts from the `fonts` folder of the JLM distribution package by choosing *Add Fonts...* from the *File* menu.

2. INSTALLATION OF TEX FILES which can be done in two variants:

(a) *Quick-and-dirty*: Just place all the files contained in subfolders of `texmf/` from the distribution package directly in the folder containing your document (not in any subfolders). This is easy but makes the folder clumsy and will need to be repeated when starting to work on a new submission for JLM.

(b) *Clean but a bit more advanced*: Merge the `texmf/` subtree from the distribution package with some `texmf/` tree of your TEX installation. For example Linux TEX distributions usually come preconfigured to search for files in the `texmf/` tree in user's home directory, so one can install JLM files there.

2 BASIC USAGE

An example document is provided in the file `jlm-example.tex` in the `example/` folder of the distribution package. You can start to work on your contribution on a copy of this file. Please use a file name based on authors' names or initials. Name any additional files in a similar way, in particular do *not* use names like `picture1.pdf` or `fig4.png`.

Please do not modify any other files which you have received from us.

IF YOU ARE NEW TO LATEX: LATEX is a batch compiler. To use it you have to prepare a 'source code' file (like jlm-example.tex) and then compile it running a command like

```
xelatex jlm-example.tex
```

That would produce a PDF file named jlm-example.pdf that contains the typeset text.

If you are using an editor as a front-end for calling LaTeX, make sure the editor calls XeLaTeX. Calling pdfLaTeX or any other LaTeX variant will not work. Examples of editors that provide access to XeLaTeX are TeXnic Center (Windows) and TeXworks (multiplatform).

The JLM class is based on the standard 'article' class, so you can use all structural elements defined there and described in LaTeX manuals. For details see section 4.

XeLaTeX enables the use of full Unicode in source files (the file has to be encoded in UTF-8). The set of available characters and symbols is limited only by fonts used. Charis SIL used as the main text font has a rich repertoire of glyphs including Latin letters as used in Western and Eastern Europe, Cyrillics, IPA symbols, many combining accents.

3 PHASES OF THE SUBMISSION PROCESS

After initial submission your article will be sent to reviewers. Several review cycles are possible if necessary. If your submission is accepted, you will be asked to introduce any changes suggested by the reviewers and submit the article for copy-editing. At this stage copy editors may suggest some improvements, especially regarding language, style, and formatting. Next the paper is sent to layout editor who will check TeXnical aspects of the paper. At the last stage the final typeset article is presented to the author(s) for proofreading.

3.1 *Initial submission*

The journal uses the double-blind reviewing process. Your article should be made anonymous before sending it for reviewing. When submitting your paper, please use the `[anonymous]` option of the document class (see the header of the example file), which will remove the names and affiliations of the authors from the title page. Please note, however, that you may have other identifying elements in your text (especially in references), which you should adapt accordingly.

The initial submission should consist only of the PDF file of your article (in fact, you do not need to use XeLaTeX nor our document class at this stage).

The file has to be submitted via the journal management system at `http://jlm.ipipan.waw.pl`.

3.2 *Submission for the copy-editing phase*

When the article is accepted for publication, you need to revert any changes introduced for anonymisation, in particular remove the option `[anonymous]`. The final version should contain a complete list of authors and their affiliations.

At this stage you will need to upload the `.tex` source file of your paper, as well as the PDF file resulting from processing it. Please remember to include all files containing pictures used in your document. If references were generated with BibTEX, please send the generated `.bbl` file. Attach all nonstandard packages or fonts. Don't send back the files you received from us.

Please upload all the mentioned files as one `.zip` archive using your account on the journal's page.

3.3 *Copy-editing*

The purpose of this phase is to make the submission clear, correct and consistent. Copy editors may draw your attention to issues such as typos, punctuation, but they may also suggest how style could be improved in order to make it more clear and/or more suitable.

The copy-editing phase consists of three rounds:

1. The copy editor reads the submission uploaded by the author and introduces suggestions about how it could be improved. Changes suggested by the copy editor are ~~makred~~marked in the text as in this sentence: the fragment suggested for removal is crossed out and the new version is highlighted in yellow. We use a minimalistic markup for this function:

   ```
   …are ¶makred¶marked¶ in the text…
   ```

   JLM: A random editor's note

   Copy editor can also use notes in the margin (the command `\jlm`) to comment on changes or provide more general suggestions. Copy-editing commands get activated with the `[copyedit]` option of the document class.

2. The author reads the modified article, applying~~,~~ or rejecting changes suggested by the copy editor, and uploads a new version of the

submission. At this stage the author is free to remove the markup introduced by the copy editor leaving only the accepted variant. The author can also add responses for the editor using the command `\auth{…}`.

A: A random author's response

3. The copy editor reads the article after changes introduced by the author and ~~changes~~ ~~it to his~~ ~~heart~~ ~~content~~, if necessary, introduces further modifications.
   At this stage the copy editor removes copy-editing markup and introduces changes directly (instead of adding suggestions) because there is no subsequent author copy-editing round. However, in case the author disagrees with changes introduced by the copy editor, these can be reverted at the layout stage.

   Please pay attention to the following:

- The title of the article and titles of sections of all levels have to be split into lines in a logical way (using the `\\` command to force a break or (better) the ~ command to block the unwanted breaks).
- The title in the running head on even pages cannot take more than ¾ of the text width. If the title is longer than that, please provide a shorter version using the `\titlerunning` command in the preamble.
- Titles at all levels should use capital letters as regular sentences do – only the first word is capitalised. We do not follow the style of making all words start with capital letters. The same applies to titles of works in the bibliography.
- Overfull lines generated by TEX when there is no good way of breaking the lines have to be eliminated. If a word sticks to the margin, try to help TEX in hyphenating it. If that does not work, change the wording to improve breaks (that's what traditional printers did).
- Use an en-dash U+2013 with spaces on both sides as a punctuation in sentences (*like – this*). Do not use em-dashes U+2014 (*like—this*). Use en-dashes in number ranges (*pages 22–38*). Use hyphens in compound words (*lily-of-the-valley*). Do not use hyphens as minuses. The minus (−) has its own Unicode slot U+2212. It can be also accessed using TEX's math mode `$-$` (using a hyphen).

- A period should be used as the decimal separator and a comma as the separator for thousands (*10,230.42*).
- Do not use abbreviations when referring to elements such as Section 4.3.2, Figure 8, Table 1, or Equation (2).
- Linguistic examples from languages other than English should be accompanied by word-by-word glosses and a free translation.
- Captions for figures and tables should start with a capital letter, but they should not include a full stop at the end.
- As a rule footnote marks in the text should not separate a word from the following punctuation (so place `\footnote` command after punctuation mark if any). A footnote should consist of complete sentences (starting with capital letters, ending with full stops) unless it comprises a bare URL.
- URLs should be validated (note that they are clickable in JLM PDF files).
- Bibliography lists always need more attention than other parts of the text. Even if generating the bibliography using bibTEX please carefully check the result for correctness and completeness (missing page numbers, names of editors, publishers, etc.).
- Author names in bibliography entries should contain full first names, not initials.
- Journal titles in bibliography should not be abbreviated.
- When generating bibliography with bibTEX please note that it spoils alphabetic order of non-ASCII characters. Since Čapek is entered in the `.bib` file as `\v{C}apek` it will end up among entries for V. Such glitches have to be corrected by editing the `.bbl` file.

4 THE LAYOUT OF AN ARTICLE

The layout is fixed and the document class does not accept any options changing the layout. Please do not tamper with font size, column size, margins, etc.

We suggest using only the `\section`, `\subsection`, and, if really necessary, `\subsubsection` level of headings. A deeper structure

would probably look strange in a relatively short text (Fifak and Gryzogrzechotalski 2000).

We suggest using directly any Unicode characters needed in the source file (so á in place of LaTeX commands such as `\'{a}`), since this is much more readable. (This is only a suggestion, all LaTeX commands work as usual.)

The same applies to quotation marks and dashes, which can be input directly. However, if you prefer to use the traditional TeX notation ``` `` -- '' ```, you have to include the `[TeXligs]` option in your `\documentclass` command.

4.1 *Math formulae*

Unfortunately, there is no Unicode math version of the *Charis* font. Therefore JLM class uses traditional TeX math setup with fonts provided by the `mathdesign` package (which has to be installed). This means you cannot use non-ASCII symbols within math formulae. However, `mathdesign` makes it possible to access a rich set of symbols (including AMS extensions) using traditional LaTeX commands.

$$a = c + d, \tag{1}$$

$$e = f - d, \tag{2}$$

$$g = \sum_{\substack{0\le i\le m\\ 0<j<n}} P_\infty(i,j) \times \Delta,$$

$$h = (\alpha - \beta) \cdot \left[\sum_i a_i \left| \sum_j x_{ij} \right|^p \right]^{1/p}. \tag{3}$$

The class automatically loads the `amsmath` package, so you can easily typeset much more advanced mathematics than Equation (2).

4.2 *References*

We strongly suggest using author-year references, which are easier to follow for the reader than numerical identifiers (cf. Fifak 2001). The class internally loads the `natbib` package for that purpose. Please employ the `\citet` command when referring to the author(s) or works cited, as advised by Fifak (2001), and `\citep` – only for parenthetical references (Fifak 2001).

Figure 1: An example bibliography conforming to the JLM style

REFERENCES

Elizjusz BLARBARUCKI (1916), On some aspects of some aspects, *Astrolog Polski*, 16:13–36.

Robert BLARBARUCKI (2006), *The implementation of a differential blabalyser*, Master's thesis, Faculty of Computer Science, Warsaw University.

Filigran FIFAK (2001), On σ-ρ aromas in blabalystics, *Acta Fetorica*, 12(2):227–234.

Filigran FIFAK and Gizbert GRYZOGRZECHOTALSKI (2000), On blabal fetoria, in *Proceedings of Euroblabal 2000*, pp. 124–132.

Do not use parentheses around years when the reference is already in parentheses (this idea was studied by Fifak 2001 and, in more detail, by Blarbarucki 1916). The command `\citealp` should be used in this case.

The bibliography should be ordered alphabetically (Blarbarucki 2006, pp. 20–22), which can be achieved automatically using BibTEX and the bibliographical style `jlm` used in this example (Blarbarucki 2006, 1916; Fifak and Gryzogrzechotalski 2000). References can also be prepared manually as a `thebibliography` environment (see the file `jlm-example.bbl` for an example; when done manually, this environment should be included in the main document).

Do not forget to send the `.bbl` file generated by BibTEX. You can, if you prefer, send the whole `.bib` file (or files) with your bibliographic database. This is not necessary, however, provided the `.bbl` file was generated from the final version of your paper.

4.3 *Figures and tables*

Elements like Section 4.3.2, Figure 8, Table 1, and Equation (2) can be referenced symbolically if you provide respective labels (see the source code for the example document).

4.3.1 Figure and table placement

JLM uses a layout with captions in the margins (cf. Figure 8). Short captions are preferable. However, you can influence the way in which space is divided between a figure or table body and the caption. The command `\jlmfloatsetup` with a numerical argument in the range

0..5 selects among the predefined layouts. Available choices are depicted in Figure 2–7.

The default layout 1 uses the whole text width for the figure/table and places the caption completely in the margin. Layouts 2–5 allow for a wider caption and a narrower figure. Layout 0 is a bit special, as it places the caption below the figure and allows the figure to protrude into the margin. This layout can be used as well with a narrow figure to force the (long) caption to position below.

The command `\jlmfloatsetup` is a declaration influencing all figures and tables following it. To change the layout of a single float use braces for grouping:

```
{\jlmfloatsetup{3}
   \begin{table}
     …
     \caption{…}
     \label{…}
   \end{table}
}
```

Note, however, that too much variation in float widths will start to look ugly.

If an illustration or table is narrower than the allocated slot, it is set flushed towards the inner margin. Please do not change that setting.

4.3.2 Illustrations

All pictures/graphs/diagrams/screen shots/images must be submitted electronically. Feasible file formats are those accepted by XeLaTeX: PDF, PNG, JPEG, PGF/TikZ. To include them in your document, it is advisable to use the standard package `graphicx` (see the example in `jlm-example.tex`).

When preparing figures with external programs, please consider the following suggestions:

- The printed version of the journal will be in black and white. Please check that your figure is readable when converted to grayscale. The pictures can, however, be in colour for the benefit of the on-line version.

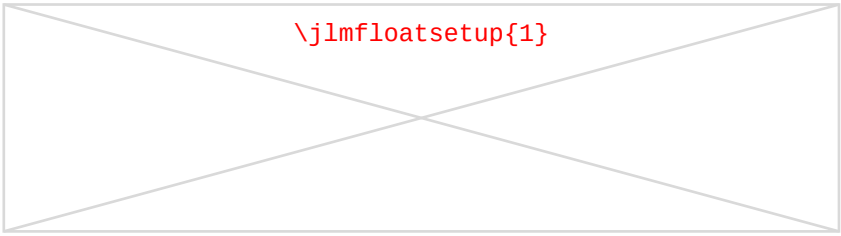

Figure 2: The division of space between a figure and a caption in layout 1 (default)

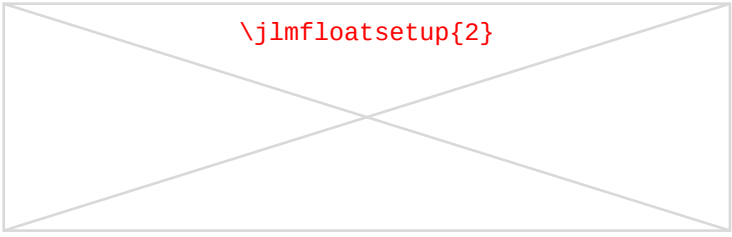

Figure 3: The division of space between a figure and a caption in layout 2

Figure 4: The division of space between a figure and a caption in layout 3

Figure 5: The division of space between a figure and a caption in layout 4

Figure 6: The division of space between a figure and a caption in layout 5. This is the other extreme where a very long caption can accompany a very small picture

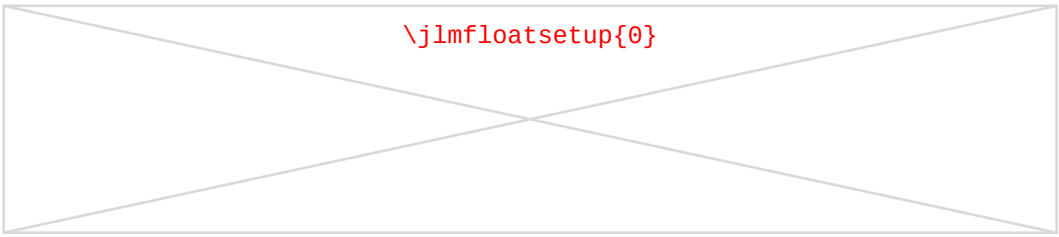

Figure 7: The division of space between a figure and a caption in layout 0. This layout should be used in case of oversized pictures or extremely long captions. This is the only layout where the caption is placed below the float. The text of the caption fits in normal text width, but the picture is allowed to extend into the space usually reserved for the caption

- Prefer vector graphics formats (PDF if possible), which will not deteriorate when scaled.
- If a bitmap is unavoidable, use a resolution of at least 300 dpi (with the obvious exception of screen shots).
- Do not use lossy JPEG compression for screen shots nor line graphics (in particular for pictures containing textual elements).
- If an illustration contains textual elements, use *Charis SIL Compact* for them. Since you have installed the font in your system, you can use it in any graphics editor.

4.3.3 Tables

For tables we prefer an 'open' style without too many rules, cf. Table 1. Do not use rules between regular columns or rows. Horizontal rules should be used at the top and bottom of the whole table. Rules can be used to separate headings from content rows/columns or to group some sets of rows/columns.

Over-wide tables have to be reduced to the page width. You can reduce the inter-column distances in such tables (the `\tabcolsep` parameter) or rotate the table and place it as a full-page float.

4.4 *Additional packages and definitions*

You can use any LaTeX 2 compatible packages in your paper. If you are using any nonstandard packages or fonts, please send them to us with your document. By nonstandard packages we mean those

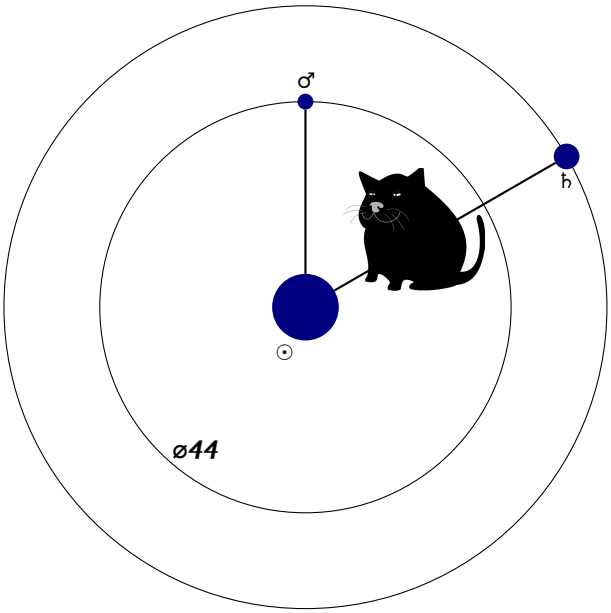

Figure 8:
An example of an external PDF file included in the XƎLATEX document

Table 1:
Example of a table

house	⊙	α_{Cent}	M_{31}	≈	⚹	△
1	0.43	102	12.4	4	2	1asp
3	0.45	412	32.6	14	7	4kid
7	0.16	111	92.1	3	9	2mer
12	0.49	224	25.5	1	1	4asp

not contained in the `tex/latex/{base,tools,graphics,amsmath,amsfonts}` folders of your TEX installation.

Please avoid using modified versions of standard packages. However, if you do use such variants, please make sure the modified version is named differently than the original and send it with your paper.

If you need any TEX definitions of your own, please put them in the document preamble (the place is marked in the sample file).