

Instructions for EMNLP 2018 Proceedings

Anonymous EMNLP submission

Abstract

This document contains the instructions for preparing a camera-ready manuscript for the proceedings of EMNLP 2018. The document itself conforms to its own specifications, and is therefore an example of what your manuscript should look like. These instructions should be used for both papers submitted for review and for final versions of accepted papers. Authors are asked to conform to all the directions reported in this document.

1 Credits

This document has been adapted from the instructions for earlier ACL and NAACL proceedings. It represents a recent build from <https://github.com/acl-org/acl-pub>, with modifications by Micha Elsner and Preethi Raghavan, based on the NAACL 2018 instructions by Margaret Mitchell and Stephanie Lukin, 2017/2018 (NA)ACL bibtex suggestions from Jason Eisner, ACL 2017 by Dan Gildea and Min-Yen Kan, NAACL-2016 by Margaret Mitchell, ACL-2012 by Maggie Li and Michael White, those from ACL-2010 by Jing-Shing Chang and Philipp Koehn, those for ACL-2008 by Johanna D. Moore, Simone Teufel, James Allan, and Sadaoki Furui, those for ACL-2005 by Hwee Tou Ng and Kemal Oflazer, those for ACL-2002 by Eugene Charniak and Dekang Lin, and earlier ACL and EACL formats. Those versions were written by several people, including John Chen, Henry S. Thompson and Donald Walker. Additional elements were taken from the formatting instructions of the *International Joint Conference on Artificial Intelligence* and the *Conference on Computer Vision and Pattern Recognition*.

2 Introduction

The following instructions are directed to authors of papers submitted to EMNLP 2018 or accepted for publication in its proceedings. All authors are required to adhere to these specifications. Authors are required to provide a Portable Document Format (PDF) version of their papers. **The proceedings are designed for printing on A4 paper.**

All MSWord formatting for EMNLP 2018 is made available in the MSWord Styles in this template. In newer versions of MSWord, click Home, then expand the Styles tile by clicking the diagonal arrow on the lower left corner. This should open all styles in the template for you to apply to your document as needed. Otherwise, you may expose the Styles following the instructions provided at:

<http://blogs.technet.com/b/hub/archive/2010/11/22/view-and-edit-styles-quickly-in-word-2010.aspx>.

3 General Instructions

Manuscripts must be in two-column format. Exceptions to the two-column format include the title, authors' names and complete addresses, which must be centered at the top of the first page, and any full-width figures or tables (see the guidelines in Subsection 3.6). Lines should be justified, with even spacing between margins (Ctrl+J). Single-spaced lines are permitted, but **authors are encouraged to use Paragraph spacing at Multiple, 1.05pt, with Font character spacing condensed with kerning of 0.1pt, and Margins at 0.98in**, for consistency with A4 paper and documents formatted with LaTeX. Go to Format, Document, Page Setup, and make sure A4 is selected. The manuscript should be printed

single-sided and its length should not exceed the maximum page limit described in Section 5.

Pages have a ruler, header and footer for initial submission, with header at 0.3 in from top and footer at 0.4 in from bottom. When you first create your submission on softconf, please fill in your submitted paper ID where *** appears in the *** at the header at the top. If the paper is accepted, remove the header, footer (page numbers), and the ruler for the final version (camera-ready). **Do not number the pages in the camera-ready version.**

The review process is double-blind, so do not include any author information (names, addresses) when submitting a paper for review. However, you should maintain space for names and addresses so that they will fit in the final (accepted) version. The EMNLP 2018 MS Word style contains a 2.5 in table beneath the title for this purpose.

The author list for submissions should include all (and only) individuals who made substantial contributions to the work presented. Each author listed on a submission to EMNLP 2018 will be notified of submissions, revisions and the final decision. No authors may be added to or removed from submissions to EMNLP 2018 after the submission deadline.

3.1 The Ruler

The EMNLP 2018 style defines a printed ruler that should be present in the version submitted for review. The ruler is provided in order that reviewers may comment on particular lines in the paper without circumlocution. If you are preparing a document without the provided style files, please arrange for an equivalent ruler to appear on the final output pages. The presence or absence of the ruler should not change the appearance of any other content on the page. The camera ready copy should not contain a ruler.

In this Word template, you can add the ruler to each individual page by copying it from another page. The Ruler consists of separate left-hand rulers and right-hand rulers. Each ruler is a text box, sized 10 in x 0.45 in, flushed with the sides of the page, with a table containing numbers. The table has no borders, and the edges are pushed to the edges of the textbox. The left part of the ruler is justified left, left indent 0.1 in/right indent 0 in, and saved in the “ACL Ruler Left” style; and the right-hand ruler, justified right had left indent 0 in/right indent 0.1 in, saved in the “ACL Ruler Right” style. Number text is in Arial, and spacing between each

line is Multiple, 1.05 in, with 4.5 pt below each line. After pasting a new ruler onto a page, highlight the ruler text and press **F9** to update the numbers. **Align** the text box **Middle**, and **Center**, and then **Arrange Behind Text**.

Reviewers: Note that the ruler measurements do not align well with lines in the paper — this turns out to be very difficult to do well when the paper contains many figures and equations, and, when done, looks ugly. Just use fractional references (e.g., the first paragraph on this page ends at mark 152.5).

3.2 Electronically-available Resources

SIGDAT provides this description in LATEX2e (emnlp2018.tex) and PDF format (emnlp2018.pdf), along with the LATEX2e style file used to format it (emnlp2018.sty) and an ACL bibliography style (acl_natbib_nourl.bst) and example bibliography (emnlp2018.bib). These files are all available at:

<http://emnlp2018.org/downloads/emnlp18-latex.zip>. A Microsoft Word template file (emnlp18-word.docx) and example submission pdf (emnlp18-word.pdf) is available at: <http://emnlp2018.org/downloads/emnlp18-word.zip>. We strongly recommend the use of these style files, which have been appropriately tailored for the EMNLP 2018 proceedings.

3.3 Format of Electronic Manuscript

For the production of the electronic manuscript, you must use Adobe's Portable Document Format (PDF). This format can be generated from postscript files: on Unix systems, you can use ps2pdf for this purpose; on Windows, you can use Adobe's Distiller, an online service such as <http://go4convert.com/>, or if you have cygwin installed, you can use dvipdf or ps2pdf.

Please make sure that your PDF file includes all the necessary formatting, hyperlinks, and fonts (especially tree diagrams, symbols, and fonts with Asian characters). When you print or create the PDF file, there is usually an option in your printer setup to include none, all, or just non-standard fonts. Please make sure that you select the option of including ALL the fonts. **Before sending it, test your PDF by printing it from a computer different from the one where it was created.**

Moreover, some word processors may generate very large PDF files, where each page is rendered as an image. Such images may reproduce poorly. In this case, try alternative ways to obtain the PDF. One way on some systems is to install a driver for a postscript printer, send your document to the printer specifying “Output to a file”, then convert the file to PDF.

It is of utmost importance to specify the **A4 format** (21 cm x 29.7 cm) when formatting the paper. Print-outs of the PDF file on A4 paper should be identical to the hardcopy version. If you cannot meet the above requirements about the production of your electronic submission, please contact the publication chairs above as soon as possible.

3.4 Layout

Format manuscripts two columns to a page, in the manner these instructions are formatted. The exact dimensions for a page on A4 paper are:

- Left and right margins: 2.5 cm
- Top margin: 2.5 cm
- Bottom margin: 2.5 cm
- Column width: 7.7 cm
- Column height: 24.7 cm
- Gap between columns: 0.6 cm

Papers should not be submitted on any other paper size. If you cannot meet the above requirements about the production of your electronic submission, please contact the publication chairs above as soon as possible.

3.5 Fonts

For reasons of uniformity, Adobe's **Times Roman** font should be used. If Times Roman is not available, you may use the Times New Roman font, which is often provided by default and only slightly different.

3.6 The First Page

Center the title, author name(s) and affiliation(s) across both columns (or, in the case of initial submission, space for the names). Do not use footnotes for affiliations. Use the two-column format only when you begin the abstract.

Title: Place the title centered at the top of the first page, in a 15-point bold font. (For a complete guide to font sizes and styles, see Table 1.) Long titles should be typed on two lines without a blank line intervening. Approximately, put the title at 2.5 cm from the top of the page, followed by a blank line, then the author name(s), and the affiliation(s) on the following line. Do not use only initials for given names (middle initials are allowed). Do not format surnames in all capitals (*e.g.*, use “Mitchell,” not “MITCHELL”). Do not use format title and section headings in all capitals as well, except for proper names (such as “BLEU”) that are conventionally in all capitals. The affiliation should contain the author's complete address, and if possible, an electronic mail address. Start the body of the first page 7.5 cm from the top of the page.

The title, author names and addresses should be completely identical to those entered to the electronic paper submission website in order to maintain the consistency of author information

Type of Text	Font Size	Style
paper title	15 pt	bold
author names	12 pt	bold
author affiliation	12 pt	
the word “Abstract”	12 pt	bold
section titles	12 pt	bold
document text	11 pt	
captions	10 pt	
abstract text	11 pt	
bibliography	10 pt	
footnotes	9 pt	

Table 1: Font guide.

among all publications of the conference. If they are different, the publication chairs may resolve the difference without consulting with you; so it is in your own interest to double-check that the information is consistent.

Abstract: Type the abstract at the beginning of the first column. The width of the abstract text should be smaller than the width of the columns for the text in the body of the paper by about 0.6cm on each side. Center the word **Abstract** in a 12 point bold font above the body of the abstract. The abstract should be a concise summary of the general thesis and conclusions of the paper. It should be no longer than 200 words. The abstract text should be in 10 point font.

Text: Begin typing the main body of the text immediately after the abstract, observing the two-

column format as shown in the present document. Do not include page numbers. Use 11 points for text. Indent when starting a new paragraph, about 0.16 in. This is accomplished with Right click, Styles, “ACL First Line”.

3.7 Sections

Headings: Type and label section and subsection headings in the style shown on the present document. Use numbered sections (Arabic numerals) in order to facilitate cross references. Number subsections with the section number and the subsection number separated by a dot, in Arabic numerals. In bold, use 11 points for subsection headings, 12 points for section headings. Do not number subsubsections.

Citations: Citations within the text appear in parentheses as (Gusfield, 1997) or, if the author's name appears in the text itself, as Gusfield (1997). Collapse multiple citations as in (Gusfield, 1997; Aho and Ullman, 1972). Append lowercase letters to the year in cases of ambiguities. Treat double authors as in (Aho and Ullman, 1972), but write as in (Chandra et al., 1981) when more than two authors are involved. Also refrain from using full citations as sentence constituents.

We suggest that instead of

“(Gusfield, 1997) showed that ...”

you use

“Gusfield (1997) showed that ...”

Hyperlinks: Within-document and external hyperlinks are indicated with Dark Blue text, Color Hex #000099. In order to create hyperlinks between citations and references, as you insert each full reference in the References section, highlight it and then select Insert, Bookmark. Link back to the reference from its citations in the text by highlight the citation, right clicking, and selecting Insert, Hyperlink, then selecting the Bookmark you’ve saved. Highlight the citation again to give make it dark blue (included in this theme), if it is not automatically applied. If there are problems saving the hyperlinks when you convert the document to PDF, use an online converter such as <http://go4convert.com>.

Digital Object Identifiers: As part of our work to make ACL materials more widely used and cited

outside of our discipline, ACL has registered as a CrossRef member, as a registrant of Digital Object Identifiers (DOIs), the standard for registering permanent URNs for referencing scholarly materials. SIGDAT has **not** adopted the ACL policy of requiring camera-ready references to contain the appropriate DOIs (or as a second resort, the hyperlinked ACL Anthology Identifier) to all cited works. But we certainly encourage you to use References that contain DOI or URLs for any of the ACL materials that you reference. Appropriate records should be found for most materials in the current ACL Anthology at <http://aclanthology.info/>.

As examples, we cite (Goodman et al., 2016) to show you how papers with a DOI will appear in the bibliography. We cite (Harper, 2014) to show how papers without a DOI but with an ACL Anthology Identifier will appear in the references.

Anonymity: As reviewing will be double-blind, the submitted version of the papers should not include the authors’ names and affiliations. Furthermore, self-references that reveal the author’s identity, e.g.,

“We previously showed (Gusfield, 1997) ...”

should be avoided. Instead, use citations such as

“Gusfield (1997) previously showed ...”

Preprint servers such as arXiv.org and workshops that do not have published proceedings are not considered archival for purposes of submission. However, to preserve the spirit of blind review, authors are encouraged to refrain from posting until the completion of the review process. Otherwise, authors must state in the online submission form the name of the workshop or preprint server and title of the non-archival version. The submitted version should be suitably anonymized and not contain references to the prior non-archival version. Reviewers will be told: “The author(s) have notified us that there exists a non-archival previous version of this paper with significantly overlapping text. We have approved submission under these circumstances, but to preserve the spirit of blind review, the current submission does not reference the non-archival version.”

Please do not use anonymous citations and do not include acknowledgements when submitting your papers. Papers that do not conform to these requirements may be rejected without review.

References: Gather the full set of references together under the heading **References**; place the section before any Appendices. Arrange the references alphabetically by first author, rather than by order of occurrence in the text. Provide as complete a reference as possible, using a consistent format, such as the one for *Computational Linguistics* or the one in the *Publication Manual of the American Psychological Association* (American Psychological Association, 1983). Use of full names for authors rather than initials is preferred. A list of abbreviations for common computer science journals can be found in the Association for Computing Machinery (1983) *Computing Reviews*.

The bibliography style described here roughly fits the American Psychological Association format, allowing regular citations, short citations and multiple citations as described above.

- Example citing an arXiv paper: (Rasooli and Tetreault, 2015).
- Example article in journal citation: (Ando and Zhang, 2005).
- Example article in proceedings, with location: (Borschinger and Johnson, 2011).
- Example article in proceedings, without location: (Andrew and Gao, 2007).

Submissions should accurately reference prior and related work, including code and data. If a piece of prior work appeared in multiple venues, the version that appeared in a refereed, archival venue should be referenced. If multiple versions of a piece of prior work exist, the one used by the authors should be referenced. Authors should not rely on automated citation indices to provide accurate references for prior and related work.

Appendices: Appendices, if any, directly follow the text and the references. Letter them in sequence and provide an informative title: **Appendix A. Title of Appendix.**

3.8 URLs

URLs should be typeset in Courier New, or Courier if this font is not available. Very long URLs should be broken manually, shortened or placed in footnotes. Be aware that actual URLs should appear in the text in human-readable format;

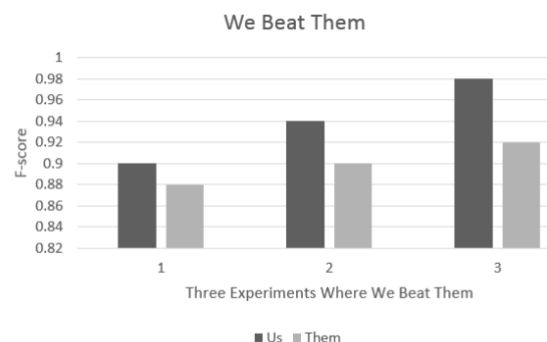


Figure 1: Figure caption.

neither internal nor external hyperlinks will appear in the proceedings.

3.9 Footnotes

Footnotes: Put footnotes at the bottom of the page and use 9 point font. They may be numbered or referred to by asterisks or other symbols.¹ Footnotes should be separated from the text by a line.²

3.10 Figures and Tables

Creating: To create a new Figure or Table, insert a Text Box where you want it to appear (generally, centered at the top of a column close to where it is referred to) and then fill it in with the Figure (or Table). Highlight and right click to add Caption, with the ACL Caption style, which places 10 pt below and above the caption.

Placement: Place figures and tables in the paper near where they are first discussed, as close as possible to the top of their respective column. Wide figures and tables may run across both columns and should be placed at the top of a page.

In MSWord, authors can place a Figure (e.g., a graphic and its caption) inside the rows of a 2 x 1 table (2 rows and 1 column) with invisible borders. Specify table positioning by right-clicking its handle in the upper left corner. Place the image in the center of the first row, and the caption in the center of the second row.

Captions: Provide a caption for every table and figure; number each one sequentially in the form: “Figure 1: Figure caption.”, “Table 1: Table caption.” Type the captions of the figures and tables below the body, using 10 point text.

Numbering: To update numbering, highlight all the relevant text (e.g., **Ctrl-A + F9**). This will

¹ This is how a footnote should appear.

¹ Note the line separating the footnotes from the text.

update all the numbering applicable to tables, figures, equations, and headings.

Cross-referencing: To add a cross reference to a figure or table:

- Place the mouse pointer at the location where you wish to add the cross-reference.
- Click on the **Insert** menu, (then click **Reference**), and then **Cross-reference** in the **Links** panel.
- In the **Cross-reference** dialog box, click the caption to which you are building the text reference.
- For a figure, under **Reference Type**, click **Figure**.
- Under Insert Reference To, click Only Label and Number, then click OK.
- Once the reference is in place, apply the 'Normal' font style (size 11, no bold face).
- This is an example reference to Figure 1.

3.11 Equations

An example equation is shown below:

$$A = \pi r^2 \quad (1)$$

To add new equations, authors are encouraged to copy this existing equation line, and then replace with the new equation. The numbering and alignment of equation line elements is automatic. To update equation numbering, press **Ctrl-A + F9**. Note: this will only update the number to the right of the equation; to update numbering within the text you must create a cross-reference.

Cross-referencing: To create a cross-reference for an equation:

- Create a bookmark for it.
- Select the number to the right of the equation. Go to **Insert, Bookmark** (in the **Links** panel), and then create a name for your equation. Press **Add** to create the bookmark.
- To refer back, place the mouse pointer at the location where you wish to add the cross reference.
- Go to **Insert, Cross-reference** (in the **Links** panel). In the dialogue box, select **Bookmark** and **Bookmark Text** from each

dropdown list. Uncheck **Insert as Hyperlink**, then click **OK**.

- This will make it such that whenever a new equation is added, the references to the equation will be updated when **Ctrl-A + F9** is pressed.
- This an example cross-reference to Equation (1).

3.12 Accessibility

In an effort to accommodate people who are color-blind (as well as those printing to paper), grayscale readability for all accepted papers is encouraged. Color is not forbidden, but authors should ensure that tables and figures do not rely solely on color to convey critical distinctions. A simple criterion: All curves and points in your figures should be clearly distinguishable without color.

3.13 From Submitted to Camera-Ready

To remove submission formatting for the camera-ready document, delete the header and footer on the first page (double clicking in the area, then delete the text). This will remove them from the rest of the document. To remove the ruler, highlight the lists of numbers and delete.

4 Translation of non-English Terms

It is also advised to supplement non-English characters and terms with appropriate transliterations and/or translations since not all readers understand all such characters and terms. Inline transliteration or translation can be represented in the order of:

original-form
transliteration
"translation".

5 Length of Submission

The EMNLP 2018 main conference accepts submissions of long papers and short papers. Long papers may consist of up to eight (8) pages of content, plus unlimited pages for references. Upon acceptance, final versions of long papers will be given one additional page — up to nine (9) pages with unlimited pages for references — so that reviewers' comments can be taken into account. Short papers may consist of up to four (4) pages of content, plus unlimited pages for references. Upon

acceptance, short papers will be given five (5) pages in the proceedings and unlimited pages for references.

For both long and short papers, all figures and tables that are part of the main text must be accommodated within these page limits, observing the formatting instructions given in the present document. Supplementary material in the form of appendices does not count towards the page limit.

However, note that supplementary material should be supplementary (rather than central) to the paper, and that reviewers may ignore supplementary material when reviewing the paper (see Appendix A). Papers that do not conform to the specified length and formatting requirements are subject to be rejected without review.

Workshop chairs may have different rules for allowed length and whether supplemental material is welcome. As always, the corresponding call for papers is the authoritative source.

6 STREAM Tools

This Microsoft Word file was updated in 2016 with STREAM Tools, designed for creating well-formatted reports and papers with Microsoft Word (Mamishv, 2010; Mamishv, 2013).

Acknowledgments

The acknowledgements should go immediately before the references. Do not number the acknowledgments section. Do not include this section when submitting your paper for review.

References

- Alfred. V. Aho and Jeffrey D. Ullman. 1972. *The Theory of Parsing, Translation and Compiling, volume 1*. Prentice-Hall, Englewood Cliffs, NJ.
- American Psychological Association. 1983. *Publications Manual*. American Psychological Association, Washington, DC.
- Rie Kubota Ando and Tong Zhang. 2005. A framework for learning predictive structures from multiple tasks and unlabeled data. *Journal of Machine Learning Research*, 6:1817-1853.
- Galen Andrew and Jianfeng Gao. 2007. Scalable training of L1-regularized log-linear models. In *Proceedings of the 24th International Conference on Machine Learning*, pages 33-40.
- Benjamin Borschinger and Mark Johnson. 2011. A particle filter algorithm for Bayesian word segmentation. In *Proceedings of the Australasian*

Language Technology Association Workshop 2011, pages 10-18, Canberra, Australia.

Ashok K. Chandra, Dexter C. Kozen, and Larry J. Stockmeyer. 1981. *Alternation*. *Journal of the Association for Computing Machinery*, 28(1):114-133. <https://doi.org/10.1145/322234.32224>.

Association for Computing Machinery. 1983. *Computing Reviews*, 24(11):503-512.

James Goodman, Andreas Vlachos, and Jason Naradowsky. 2016. *Noise reduction and targeted exploration in imitation learning for abstract meaning representation parsing*. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, pages 1-11. <https://doi.org/10.18653/v1/P16-1001>.

Dan Gusfield. 1997. *Algorithms on Strings, Trees and Sequences*. Cambridge University Press, Cambridge, UK.

Mary Harper. 2014. *Learning from 26 languages: Program management and science in the babel program*. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*. Dublin City University and Association for Computational Linguistics, page 1. <http://aclweb.org/anthology/C14-1001>.

Alexander V. Mamishv and Murray Sargent. 2013. *Creating Research and Scientific Documents Using Microsoft Word*. Microsoft Press, Redmond, WA.

Alexander V. Mamishv and Sean D. Williams. 2010. *Technical Writing for Teams: The STREAM Tools Handbook*. Wiley-IEEE Press, Hoboken, NJ.

Mohammad Sadegh Rasooli and Joel R. Tetreault. 2015. *Yara parser: A fast and accurate dependency parser*. *Computing Research Repository*, arXiv:1503.06733. Version 2.

A Supplementary Material

Each EMNLP 2018 submission can be accompanied by a single PDF appendix, one .tgz or .zip appendix containing software, and one .tgz or .zip appendix containing data.

Submissions may include resources (software and/or data) used in the work and described in the paper. Papers that are submitted with accompanying software and/or data may receive additional credit toward the overall evaluation score, and the potential impact of the software and data will be taken into account when making the acceptance/rejection decisions. Any

accompanying software and/or data should include licenses and documentation of research review as appropriate.

EMNLP 2018 also encourages the submission of supplementary material to report preprocessing decisions, model parameters, and other details necessary for the replication of the experiments reported in the paper. Seemingly small preprocessing decisions can sometimes make a large difference in performance, so it is crucial to record such decisions to precisely characterize state-of-the-art methods.

Nonetheless, supplementary material should be supplementary (rather than central) to the paper.

Submissions that misuse the supplementary material may be rejected without review.

Essentially, supplementary material may include explanations or details of proofs or derivations that do not fit into the paper, lists of features or feature templates, sample inputs and outputs for a system, pseudo-code or source code, and data. (Source code and data should be separate uploads, rather than part of the paper).

The paper should not rely on the supplementary material: while the paper may refer to and cite the supplementary material and the supplementary material will be available to the reviewers, they will not be asked to review the supplementary material.

Appendices (*i.e.*, supplementary material in the form of proofs, tables, or pseudo-code) should be **uploaded as supplementary material** when submitting the paper for review. Upon acceptance, the appendices come after the references, as shown here.