

Movie classification leveraging MapReduce

Giacomo Marciani

University of Rome Tor Vergata
Vial del Politecnico 1
Rome, Italy 00133
gmarciani@acm.org

Michele Porretta

University of Rome Tor Vergata
Vial del Politecnico 1
Rome, Italy 00133
mporretta@acm.org

ABSTRACT

This paper provides a sample of a \LaTeX document which conforms, somewhat loosely, to the formatting guidelines for ACM SIG Proceedings. It is an *alternate* style which produces a *tighter-looking* paper and was designed in response to concerns expressed, by authors, over page-budgets. It complements the document *Author's (Alternate) Guide to Preparing ACM SIG Proceedings Using $\LaTeX 2_{\epsilon}$ and BibTeX*. This source file has been written with the intention of being compiled under $\LaTeX 2_{\epsilon}$ and BibTeX.

The developers have tried to include every imaginable sort of “bells and whistles”, such as a subtitle, footnotes on title, subtitle and authors, as well as in the text, and every optional component (e.g. Acknowledgments, Additional Authors, Appendices), not to mention examples of equations, theorems, tables and figures.

To make best use of this sample document, run it through \LaTeX and BibTeX, and compare this source code with the printed output produced by the dvi file. A compiled PDF version is available on the web page to help you with the ‘look and feel’.

CCS CONCEPTS

• Computing methodologies → MapReduce algorithms;

KEYWORDS

classification; MapReduce; Hadoop

ACM Reference format:

Giacomo Marciani and Michele Porretta. 2017. Movie classification leveraging MapReduce. In *Proceedings of Computer Engineering Courseworks on Systems and Architectures for Big Data, Rome, Italy, July 2017 (SABD'17)*, 3 pages.

https://doi.org/10.475/123_4

1 INTRODUCTION

The *proceedings* are the records of a conference.¹ ACM seeks to give these conference by-products a uniform, high-quality appearance. To do this, ACM has some rigid requirements for the format of the proceedings documents: there is a specified format (balanced double columns), a specified set of fonts (Arial or Helvetica and Times Roman) in certain specified sizes, a specified live area, centered

¹This is a footnote

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

SABD'17, July 2017, Rome, Italy

© 2017 Copyright held by the owner/author(s).

ACM ISBN 123-4567-24-567/08/06...\$0.00

https://doi.org/10.475/123_4

on the page, specified size of margins, specified column width and gutter size.

2 THE BODY OF THE PAPER

Typically, the body of a paper is organized into a hierarchical structure, with numbered or unnumbered headings for sections, subsections, sub-subsections, and even smaller sections. The command `\section` that precedes this paragraph is part of such a hierarchy.² \LaTeX handles the numbering and placement of these headings for you, when you use the appropriate heading commands around the titles of the headings. If you want a sub-subsection or smaller part to be unnumbered in your output, simply append an asterisk to the command name. Examples of both numbered and unnumbered headings will appear throughout the balance of this sample document.

Because the entire article is contained in the **document** environment, you can indicate the start of a new paragraph with a blank line in your input file; that is why this sentence forms a separate paragraph.

2.1 Type Changes and *Special* Characters

We have already seen several typeface changes in this sample. You can indicate italicized words or phrases in your text with the command `\textit`; emboldening with the command `\textbf` and typewriter-style (for instance, for computer code) with `\texttt`. But remember, you do not have to indicate typestyle changes when such changes are part of the *structural* elements of your article; for instance, the heading of this subsection will be in a sans serif³ typeface, but that is handled by the document class file. Take care with the use of⁴ the curly braces in typeface changes; they mark the beginning and end of the text that is to be in the different typeface.

You can use whatever symbols, accented characters, or non-English characters you need anywhere in your document; you can find a complete list of what is available in the *\LaTeX User's Guide* [5].

2.2 Math Equations

You may want to display math equations in three distinct styles: inline, numbered or non-numbered display. Each of the three are discussed in the next sections.

2.2.1 Inline (In-text) Equations. A formula that appears in the running text is called an inline or in-text formula. It is produced by the **math** environment, which can be invoked with the usual `\begin . . . \end` construction or with the short form `$. . . $`. You can use any of the symbols and structures, from α to ω , available

²This is a footnote.

³Another footnote here. Let's make this a rather long one to see how it looks.

⁴Another footnote.

in \LaTeX [5]; this section will simply show a few examples of in-text equations in context. Notice how this equation: $\lim_{n \rightarrow \infty} x = 0$, set here in in-line math style, looks slightly different when set in display style. (See next section).

2.2.2 Display Equations. A numbered display equation—one set off by vertical space from the text and centered horizontally—is produced by the `equation` environment. An unnumbered display equation is produced by the `displaymath` environment. Again, in either environment, you can use any of the symbols and structures available in \LaTeX ; this section will just give a couple of examples of display equations in context. First, consider the equation, shown as an inline equation above:

$$\lim_{n \rightarrow \infty} x = 0$$

(1)

Notice how it is formatted somewhat differently in the `displaymath` environment. Now, we’ll enter an unnumbered equation:

$$\sum_{i=0}^{\infty} x + 1$$

and follow it with another numbered equation:

$$\sum_{i=0}^{\infty} x_i = \int_0^{\pi+2} f$$

(2)

just to demonstrate \LaTeX ’s able handling of numbering.

2.3 Citations

Citations to articles [1–4], conference proceedings [3] or maybe books [5, 6] listed in the Bibliography section of your article will occur throughout the text of your article. You should use BibTeX to automatically produce this bibliography; you simply need to insert one of several citation commands with a key of the item cited in the proper location in the .tex file [5]. The key is a short reference you invent to uniquely identify each work; in this sample document, the key is the first author’s surname and a word from the title. This identifying key is included with each item in the .bib file for your article.

The details of the construction of the .bib file are beyond the scope of this sample document, but more information can be found in the *Author’s Guide*, and exhaustive details in the *LaTeX User’s Guide* by Lamport [5].

This article shows only the plainest form of the citation command, using `\cite`.

2.4 Tables

Because tables cannot be split across pages, the best placement for them is typically the top of the page nearest their initial cite. To ensure this proper “floating” placement of tables, use the environment `table` to enclose the table’s contents and the table caption. The contents of the table itself must go in the `tabular` environment, to be aligned properly in rows and columns, with the desired horizontal and vertical rules. Again, detailed instructions on `tabular` material are found in the *LaTeX User’s Guide*.

Immediately following this sentence is the point at which Table 1 is included in the input file; compare the placement of the table here with the table in the printed output of this document.

Table 1: Frequency of Special Characters

Non-English or Math	Frequency	Comments
∅	1 in 1,000	For Swedish names
π	1 in 5	Common in math
\$	4 in 5	Used in business
Ψ ₁ ²	1 in 40,000	Unexplained usage



Figure 1: A sample black and white graphic.



Figure 2: A sample black and white graphic that has been resized with the `includegraphics` command.

To set a wider table, which takes up the whole width of the page’s live area, use the environment `table*` to enclose the table’s contents and the table caption. As with a single-column table, this wide table will “float” to a location deemed more desirable. Immediately following this sentence is the point at which Table 2 is included in the input file; again, it is instructive to compare the placement of the table here with the table in the printed output of this document.

It is strongly recommended to use the package `booktabs` [?] and follow its main principles of typography with respect to tables:

- (1) Never, ever use vertical rules.
- (2) Never use double rules.

It is also a good idea not to overuse horizontal rules.

2.5 Figures

Like tables, figures cannot be split across pages; the best placement for them is typically the top or the bottom of the page nearest their initial cite. To ensure this proper “floating” placement of figures, use the environment `figure` to enclose the figure and its caption.

This sample document contains examples of .eps files to be displayable with \LaTeX . If you work with pdf \LaTeX , use files in the .pdf format. Note that most modern \TeX systems will convert .eps to .pdf for you on the fly. More details on each of these are found in the *Author’s Guide*.

As was the case with tables, you may want a figure that spans two columns. To do this, and still to ensure proper “floating” placement of tables, use the environment `figure*` to enclose the figure and its caption. And don’t forget to end the environment with `figure*`, not `figure`!

Table 2: Some Typical Commands

Command	A Number	Comments
<code>\author</code>	100	Author
<code>\table</code>	300	For tables
<code>\table*</code>	400	For wider tables

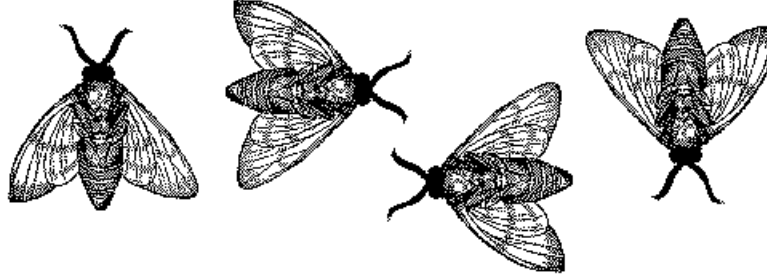
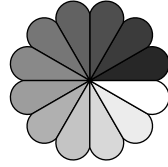


Figure 3: A sample black and white graphic that needs to span two columns of text.

Figure 4: A sample black and white graphic that has been resized with the `includegraphics` command.

2.6 Theorem-like Constructs

Other common constructs that may occur in your article are the forms for logical constructs like theorems, axioms, corollaries and proofs. ACM uses two types of these constructs: theorem-like and definition-like.

Here is a theorem:

THEOREM 2.1. *Let f be continuous on $[a, b]$. If G is an antiderivative for f on $[a, b]$, then*

$$\int_a^b f(t) dt = G(b) - G(a).$$

Here is a definition:

Definition 2.2. If z is irrational, then by e^z we mean the unique number that has logarithm z :

$$\log e^z = z.$$

The pre-defined theorem-like constructs are **theorem**, **conjecture**, **proposition**, **lemma** and **corollary**. The pre-defined definition-like constructs are **example** and **definition**. You can add your own constructs using the `amsthm` interface [?]. The styles used in the `\theoremstyle` command are **acmplain** and **acmdefinition**.

Another construct is **proof**, for example,

PROOF. Suppose on the contrary there exists a real number L such that

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = L.$$

Then

$$l = \lim_{x \rightarrow c} f(x) = \lim_{x \rightarrow c} \left[gx \cdot \frac{f(x)}{g(x)} \right] = \lim_{x \rightarrow c} g(x) \cdot \lim_{x \rightarrow c} \frac{f(x)}{g(x)} = 0 \cdot L = 0,$$

which contradicts our assumption that $l \neq 0$. \square

3 CONCLUSIONS

This paragraph will end the body of this sample document. Remember that you might still have Acknowledgments or Appendices; brief samples of these follow. There is still the Bibliography to deal with; and we will make a disclaimer about that here: with the exception of the reference to the \LaTeX book, the citations in this paper are to articles which have nothing to do with the present subject and are used as examples only.

REFERENCES

- [1] Mic Bowman, Saumya K. Debray, and Larry L. Peterson. 1993. Reasoning About Naming Systems. *ACM Trans. Program. Lang. Syst.* 15, 5 (November 1993), 795–825.
- [2] Johannes Braams. 1991. Babel, a Multilingual Style-Option System for Use with \LaTeX 's Standard Document Styles. *TUGboat* 12, 2 (June 1991), 291–301.
- [3] Malcolm Clark. 1991. Post Congress Tristesse. In *TeX90 Conference Proceedings*. TeX Users Group, 84–89.
- [4] Maurice Herlihy. 1993. A Methodology for Implementing Highly Concurrent Data Objects. *ACM Trans. Program. Lang. Syst.* 15, 5 (November 1993), 745–770.
- [5] Leslie Lamport. 1986. *LaTeX User's Guide and Document Reference Manual*. Addison-Wesley Publishing Company, Reading, Massachusetts.
- [6] S.L. Salas and Einar Hille. 1978. *Calculus: One and Several Variable*. John Wiley and Sons, New York.