

Alternate ACM SIG Proceedings Paper in LaTeX Format*

[Extended Abstract][†]

Ben Trovato[‡]
Institute for Clarity in
Documentation
1932 Wallamaloo Lane
Wallamaloo, New Zealand
trovato@corporation.com

G.K.M. Tobin[§]
Institute for Clarity in
Documentation
P.O. Box 1212
Dublin, Ohio 43017-6221
webmaster@marysville-
ohio.com

Lars Thørväld[¶]
The Thørväld Group
1 Thørväld Circle
Hekla, Iceland
larst@affiliation.org

Lawrence P. Leipuner
Brookhaven Laboratories
Brookhaven National Lab
P.O. Box 5000
lleipuner@researchlabs.org

Sean Fogarty
NASA Ames Research Center
Moffett Field
California 94035
fogartys@amesres.org

Charles Palmer
Palmer Research Laboratories
8600 Datapoint Drive
San Antonio, Texas 78229
cpalmer@prl.com

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 $\text{\LaTeX}2_{\epsilon}$ and BibTeX*. This source file has been written with the intention of being compiled under $\text{\LaTeX}2_{\epsilon}$ 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

•Computer systems organization → Embedded systems; Redundancy; Robotics; •Networks → Network reliability;

Keywords

ACM proceedings; \LaTeX ; text tagging

1. INTRODUCTION

Nowdays, the World Wide Web has transformed from a large, static library that people only browse into a vast and dynamic information resource. Relying on this, social networks is a a very popular and powerful tool for expressing opinions, broadcasting news, and simply communicating with friends. People often comment on events in real time, with several hundred micro-blogs posted each second for significant events.

The most popular micro-blogging service is Twitter. The popularity of Twitter stems from its availability on a number of different electronic devices (e.g., web, cell phones, etc.), as well as the prevalence of a subculture in Twitter that encourages users to acquire a large friend pool, as well as send tweets on a wide variety of subjects, typically several times a day. Twitter is attractive for capturing breaking news, as there is very little lag between the time that an event happens or is first reported in the news media and the time at which it is the subject of a posting on Twitter. Twitter collects millions of real-time short text messages (known as tweets) every second. These messages are not only raw data, but they can be manipulated efficiently to provide the community with the top trending topics of the Twitter. Tweets also have a mechanism by which the user can link to other objects on the web such as articles, images or videos which is typically used to link tweets to related material on the Internet. Thereafter, the first result is that the size of information is multiplied and the variety of references is bigger, as

*(Produces the permission block, and copyright information). For use with SIG-ALTERNATE.CLS. Supported by ACM.

[†]A full version of this paper is available as *Author's Guide to Preparing ACM SIG Proceedings Using $\text{\LaTeX}2_{\epsilon}$ and BibTeX* at www.acm.org/eaddress.htm

[‡]Dr. Trovato insisted his name be first.

[§]The secretary disavows any knowledge of this author's actions.

[¶]This author is the one who did all the really hard work.

Permission to make digital or hard copies of all or part 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 components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

WOODSTOCK '97 El Paso, Texas USA

© 2017 ACM. ISBN 123-4567-24-567/08/06...\$15.00

DOI: 10.475/123.4

well. Topic detection has also instant impact on the world, through the quick transmission of the news and necessary briefing in some cases.

With the passage of time and the effect of more and more users the topic acquire much popularity. Through this phenomenon we can form a general summarization of the event. This process is called event summarization. The massive crowd keeps close pace with the development of trending topics and provide the timely updated information. Twitter has shown its powerful ability in information delivery in many events, like the wildfires in San Diego and the earthquake in Japan. In response to searches for ongoing events, today's major search engines simply find tweets that match the query terms, and present the most recent ones. This approach has the advantage of leveraging existing query matching technologies, and for simple one-shot events such as earthquakes it works well. However, for events that have "structure" or are longrunning, and where users are likely to want a summary of all occurrences so far, this approach is often unsatisfactory.

The objective of this survey is to identify some techniques of event detection. Event detection is a growing domain of research. Many different species of algorithms have been detected regarding this sector. Initially, we focus on techniques that are based on text categorization. In addition, there are some methods that reclaim the display frequency of each term. The second part of the survey analyses some methods for sentiment analysis. Over this domain there are many algorithms that use machine learning. For closure, we are going to present some techniques for location extraction and an analysis for the connections of people among a social network.

The reason why we present methods of identifying the place where a user sends the tweet is because of the strong relationship between event detection and geographical location each user belongs. Twitter is a social networking website, which means that users need not be viewed in isolation, but instead can be viewed as part of a large network of other users, user groups, and user cliques. Moreover, users have some meta-data information, such as description, source location, friends, which means that the social network structure in Twitter can aid in finding users that are most likely to tweet about news belonging to a particular geographic location or region.

The computational treatment of opinion, sentiment, and subjectivity has recently attracted a great deal of attention, in part because of its potential applications. One of the main reasons for sentiment analysis is the aforementioned increase of user-generated content on the Web which has resulted in a wealth of information that is potentially of vital importance to institutions and companies. Typically, document-based sentiment analysis processes operate at a particular level, i.e. at the word or sentence level, for extracting a document's sentiment. In machine learning, the most popular approach for sentiment analysis, the selection of appropriate features for representing a document is crucial. In sentiment identification at the word level different types of features have been introduced, which are either sentiment-based (e.g. words which express a specific sentiment), syntactic-based (e.g. part-of-speech and n-grams), or semantic-based (e.g. semantic word vector spaces which capture the meaning of each word).

Document-level polarity classification is not a special case

of text categorization with sentiment -rather than topic-based categories. Hence, standard machine learning classification techniques, such as support vector machines (SVMs), can be applied to the entire documents themselves. Nevertheless, some researches presented a technique that it is easy to improve the accuracy, by integrating sentence-level subjectivity detection with document-level sentiment polarity.

As we know, in the machine learning approach, each classifier is trained using a collection of representative data. In contrast, the semantic-orientation approach does not require prior training; instead, it measures a word containing positive or negative sentiment. Each approach has its own benefits and drawbacks. For example, the machine learning approach tends to be more accurate, but the semantic-orientation approach has better generality. In [...] it is presented a new lexicon-enhanced method to generate a set of sentiment words based on a sentiment lexicon as a new feature dimension. It is combined these sentiment features with content-free and content-specific features used in the existing machine-learning approach. In the evaluation stage, they showed that adding the new set of sentiment features can increase sentiment-classification performance.

The Internet and other communication technologies play a potentially disruptive role on the constraints imposed on social networks. These technologies reduce the overhead and cost for being introduced to new people regardless of geography, and help us stay in touch with those we know. Some have even gone so far as to call this "the end of geography," where the process of relationship formation becomes disentangled from distance altogether.

The rise of micro-blogging services spurred various applications to mine the data coming from those services. Many such applications could benefit from information about the location of users, but unfortunately location information is currently very sparse. Less than 1 per cent of tweets are geo-tagged and information available from the location field in users' profiles is unreliable at best. Mining this people-centric sensor data promises new personalized information services, including local news summarized from tweets of nearby Twitter users, the targeting of regional advertisements, spreading business information to local customers, and novel location-based applications (e.g., Twitter-based earthquake detection, which can be faster than through traditional official channels).

One direction of geoinference using social networks has produced approaches that claim to accurately locate the majority of posts within tens of kilometers of their true locations. Initially in this survey we will present an algorithm that predicts the location of an individual from a sparse set of located users with performance that exceeds IP-based geolocation. On the side, there are also techniques that predicts locations of Twitter users at different granularities, such as city, state, or time zone, using the content of their tweets and their tweeting behavior.

Getting started the first section is the presentation of some techniques that aim to event detection. The survey detects both algorithms that are based on text categorization and frequency display methods. The second chapter deals with techniques of sentiment analysis. We give more weight on techniques that use machine learning. The last chapter unfolds methods for location identification.

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.¹ L^AT_EX 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 serif² typeface, but that is handled by the document class file. Take care with the use of³ 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 *L^AT_EX 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 in L^AT_EX[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

¹This is the second footnote. It starts a series of three footnotes that add nothing informational, but just give an idea of how footnotes work and look. It is a wordy one, just so you see how a longish one plays out.

²A third footnote, here. Let's make this a rather short one to see how it looks.

³A fourth, and last, footnote.

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 L^AT_EX; this section will just give a couple of examples of display equations in context. First, consider the equation, shown as an inline equation above:

$$\lim_{n \rightarrow \infty} x = 0 \tag{1}$$

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:

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

just to demonstrate L^AT_EX's able handling of numbering.

2.3 Citations

Citations to articles [1, 3, 2, 4], conference proceedings [3] or books [6, 5] 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 *L^AT_EX User's Guide*[5].

This article shows only the plainest form of the citation command, using `\cite`. This is what is stipulated in the SIGS style specifications. No other citation format is endorsed or supported.

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 is found in the *L^AT_EX 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 dvi output of this document.

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;

Table 1: Frequency of Special Characters

Non-English or Math	Frequency	Comments
\emptyset	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.

again, it is instructive to compare the placement of the table here with the table in the printed dvi output of this document.

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 L^AT_EX. If you work with pdfL^AT_EX, use files in the **.pdf** format. Note that most modern T_EX system will convert **.eps** to **.pdf** for you on the fly. More details on each of these is 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**!

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. There are two forms, one produced by the command **\newtheorem** and the other by the command **\newdef**; perhaps the clearest and easiest way to distinguish them is to compare the two in the output of this sample document:

This uses the **theorem** environment, created by the **\newtheorem** command:

THEOREM 1. *Let f be continuous on $[a, b]$. If G is an antiderivative for f on $[a, b]$, then*

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



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

The other uses the **definition** environment, created by the **\newdef** command:

Definition 1. If z is irrational, then by e^z we mean the unique number which has logarithm z :

$$\log e^z = z$$

Two lists of constructs that use one of these forms is given in the *Author’s Guidelines*.

There is one other similar construct environment, which is already set up for you; i.e. you must *not* use a **\newdef** command to create it: the **proof** environment. Here is an example of its use:

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

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

Then

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

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

Complete rules about using these environments and using the two different creation commands are in the *Author’s Guide*; please consult it for more detailed instructions. If you need to use another construct, not listed therein, which you want to have the same formatting as the Theorem or the Definition[6] shown above, use the **\newtheorem** or the **\newdef** command, respectively, to create it.

A Caveat for the T_EX Expert

Because you have just been given permission to use the **\newdef** command to create a new form, you might think you can use T_EX’s **\def** to create a new command: *Please refrain from doing this!* Remember that your L^AT_EX source code is primarily intended to create camera-ready copy, but may be converted to other forms – e.g. HTML. If you inadvertently omit some or all of the **\defs** recompilation will be, to say the least, problematic.

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 L^AT_EX book, the citations in this paper are to articles which have nothing to do with the present subject and are used as examples only.

4. ACKNOWLEDGMENTS

This section is optional; it is a location for you to acknowledge grants, funding, editing assistance and what have you. In the present case, for example, the authors would like to thank Gerald Murray of ACM for his help in codifying this *Author’s Guide* and the **.cls** and **.tex** files that it describes.

5. ADDITIONAL AUTHORS

Additional authors: John Smith (The Thørväld Group, email: jsmith@affiliation.org) and Julius P. Kumquat (The Kumquat Consortium, email: jpkumquat@consortium.net).

Table 2: Some Typical Commands

Command	A Number	Comments
<code>\alignauthor</code>	100	Author alignment
<code>\numberofauthors</code>	200	Author enumeration
<code>\table</code>	300	For tables
<code>\table*</code>	400	For wider tables

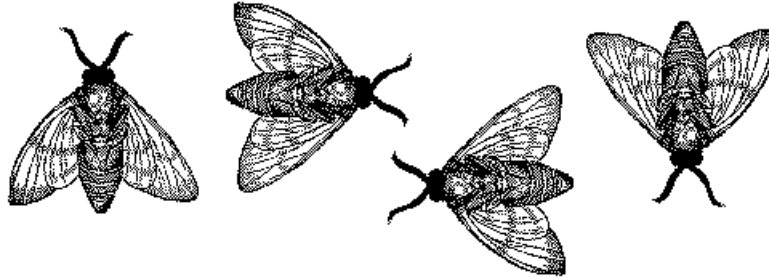


Figure 3: A sample black and white graphic that needs to span two columns of text.

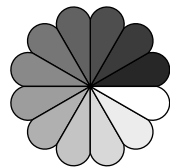


Figure 4: A sample black and white graphic that has been resized with the `includegraphics` command.

6. REFERENCES

- [1] M. Bowman, S. K. Debray, and L. L. Peterson. Reasoning about naming systems. *ACM Trans. Program. Lang. Syst.*, 15(5):795–825, November 1993.
- [2] J. Braams. Babel, a multilingual style-option system for use with latex’s standard document styles. *TUGboat*, 12(2):291–301, June 1991.
- [3] M. Clark. Post congress tristesse. In *TeX90 Conference Proceedings*, pages 84–89. TeX Users Group, March 1991.
- [4] M. Herlihy. A methodology for implementing highly concurrent data objects. *ACM Trans. Program. Lang. Syst.*, 15(5):745–770, November 1993.
- [5] L. Lamport. *LaTeX User’s Guide and Document Reference Manual*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1986.
- [6] S. Salas and E. Hille. *Calculus: One and Several Variable*. John Wiley and Sons, New York, 1978.

APPENDIX

A. HEADINGS IN APPENDICES

The rules about hierarchical headings discussed above for the body of the article are different in the appendices. In the **appendix** environment, the command **section** is used to indicate the start of each Appendix, with alphabetic order designation (i.e. the first is A, the second B, etc.) and a title (if you include one). So, if you need hierarchical structure

within an Appendix, start with **subsection** as the highest level. Here is an outline of the body of this document in Appendix-appropriate form:

A.1 Introduction

A.2 The Body of the Paper

A.2.1 Type Changes and Special Characters

A.2.2 Math Equations

Inline (In-text) Equations.

Display Equations.

A.2.3 Citations

A.2.4 Tables

A.2.5 Figures

A.2.6 Theorem-like Constructs

A Caveat for the T_EX Expert

A.3 Conclusions

A.4 Acknowledgments

A.5 Additional Authors

This section is inserted by L^AT_EX; you do not insert it. You just add the names and information in the `\additionalauthors` command at the start of the document.

A.6 References

Generated by bibtex from your .bib file. Run latex, then bibtex, then latex twice (to resolve references) to create the .bbl file. Insert that .bbl file into the .tex source file and comment out the command `\thebibliography`.

B. MORE HELP FOR THE HARDY

The sig-alternate.cls file itself is chock-full of succinct and helpful comments. If you consider yourself a moderately experienced to expert user of L^AT_EX, you may find reading it useful but please remember not to change it.