BibTeX

Introduction

This is the first draft of this document. Please send any comments to *dana@acm.org*

Description

BibTeX is a program and file format designed by Oren Patashnik and Leslie Lamport in 1985 for the LaTeX document preparation system. The format is entirely character based, so it can be used by any program (although the standard character set for accents is TeX). It is field (tag) based and the BibTeX program will ignore unknown fields, so it is expandable. It is probably the most common format for bibliographies on the Internet.

References

- *LaTeX: A Document Preparation System* by Leslie Lamport, 1986, Addison-Wesley.
- *BibTeXing* (`btxdoc.tex`), by Oren Patashnik, February 1988, (BibTeX distribution).
- Help file for BibTeX format

Software Support

The BibTeX program uses style files, a list of citations from LaTeX, and a BibTeX database to create a LaTeX file listing the cited references.

Dana Jacobsen maintains a list of some BibTeX tools.

bp and BibDB both fully support BibTeX.

Examples

```
@article{Gettys90,
   author = {Jim Gettys and Phil Karlton and Scott McGregor},
   title = {The {X} Window System, Version 11},
   journal = {Software Practice and Experience},
   volume = {20},
   number = {S2},
   year = {1990},
   abstract = {A technical overview of the X11 functionality.  This is an update
of the X10 TOG paper by Scheifler \& Gettys.}
}
```

Common problems

- The original documents specified a large number of field names, but there are many common items that are not listed. A list of some of the ones people have added are below.
- When using BibTeX, the interaction between names and accenting is somewhat tricky. You should use `G{\"o}del' or `G{\"{o}}del', and not `{G{\"{o}}del}' or `{G\"{o}del}'.
- The BibTeX program is written, as is all TeX, using static data structures, and the maximum length of any one string is by default 1000 characters. It is not uncommon for fields like abstract and contents to overflow this buffer. One solution is to go change the source code (I've changed mine to 3000, which is usually adequate). Another is to use the string concatenation feature to change

      abstract = "longlonglonglong"

  into

      abstract = "long" # "long" # "long" # "long"

  where the length of each individual string will fit inside your programs limits.

Format Description

Special features

The `@STRING` command is used to define abbreviations for use by BibTeX. The command

`@string{jgg1 = "Journal of Gnats and Gnus, Series~1"}`

defines 'jgg1' to be the abbreviation for the string "Journal of Gnats and Gnus, Series~1". Any reference outside of quotes or braces to `jgg1` will be filled in with the full string.

The `@PREAMBLE` command is used to define formatter code that will be output directly to the `bbl` file produced by the BibTeX program. This usually consists of LaTeX macros. It is unclear what one should do with the fields when converting to a format that does not use TeX.

The `@COMMENT` command lets you put any text inside it. It isn't really necessary, since BibTeX will ignore any text that isn't inside an entry.

Standard entry types

`article`
> An article from a journal or magazine.

`book`
> A book with an explicit publisher.

`booklet`
> A work that is printed and bound, but without a named publisher or sponsoring institution.

`conference`
> The same as `inproceedings`.

`inbook`
> A part of a book, which may be a chapter (or section or whatever) and/or a range of pages.

incollection
>A part of a book having its own title.

inproceedings
>An article in a conference proceedings.

manual
>Technical documentation.

mastersthesis
>A Master's thesis.

misc
>Use this type when nothing else fits.

phdthesis
>A PhD thesis.

proceedings
>The proceedings of a conference.

techreport
>A report published by a school or other institution, usually numbered within a series.

unpublished
>A document having an author and title, but not formally published.

Standard fields

For now I'm going to be lazy and give you what Oren Patashnik wrote about the fields. I'll redo this sometime, including references to how each field should be formatted.

address
>Usually the address of the `publisher` or other type of institution. For major publishing houses, van Leunen recommends omitting the information entirely. For small publishers, on the other hand, you can help the reader by giving the complete address.

annote
>An annotation. It is not used by the standard bibliography styles, but may be used by others that produce an annotated bibliography.

author
>The name(s) of the author(s), in the format described in the LaTeX book.

booktitle
>Title of a book, part of which is being cited. See the LaTeX book for how to type titles. For book entries, use the `title` field instead.

chapter
>A chapter (or section or whatever) number.

crossref
>The database key of the entry being cross referenced. Any fields that are missing from the current record are inherited from the field being cross referenced.

edition
>The edition of a book---for example, ``Second''. This should be an ordinal, and should have the first letter capitalized, as shown here; the standard styles convert to lower case when necessary.

editor
>Name(s) of editor(s), typed as indicated in the LaTeX book. If there is also an `author` field, then the `editor` field gives the editor of the book or collection in which the reference appears.

howpublished
>How something strange has been published. The first word should be capitalized.

institution
>The sponsoring institution of a technical report.

journal

A journal name. Abbreviations are provided for many journals.

key

Used for alphabetizing, cross referencing, and creating a label when the ``author'' information is missing. This field should not be confused with the key that appears in the `cite` command and at the beginning of the database entry.

month

The month in which the work was published or, for an unpublished work, in which it was written. You should use the standard three-letter abbreviation, as described in Appendix B.1.3 of the LaTeX book.

note

Any additional information that can help the reader. The first word should be capitalized.

number

The number of a journal, magazine, technical report, or of a work in a series. An issue of a journal or magazine is usually identified by its volume and number; the organization that issues a technical report usually gives it a number; and sometimes books are given numbers in a named series.

organization

The organization that sponsors a conference or that publishes a `manual`.

pages

One or more page numbers or range of numbers, such as `42--111` or `7,41,73--97` or `43+` (the `+' in this last example indicates pages following that don't form a simple range). To make it easier to maintain *Scribe*-compatible databases, the standard styles convert a single dash (as in `7-33`) to the double dash used in TeX to denote number ranges (as in `7--33`).

publisher

The publisher's name.

school

The name of the school where a thesis was written.

series

The name of a series or set of books. When citing an entire book, the the `title` field gives its title and an optional `series` field gives the name of a series or multi-volume set in which the book is published.

title

The work's title, typed as explained in the LaTeX book.

type

The type of a technical report---for example, ``Research Note''.

volume

The volume of a journal or multi-volume book.

year

The year of publication or, for an unpublished work, the year it was written. Generally it should consist of four numerals, such as `1984`, although the standard styles can handle any `year` whose last four nonpunctuation characters are numerals, such as ``\hbox{(about 1984)}''.

Other fields

BibTeX is extremely popular, and many people have used it to store information. Here is a list of some of the more common fields:

URL

The WWW Universal Resource Locator that points to the item being referenced. This often is used for technical reports to point to the ftp site where the postscript source of the report is located.

ISBN
: The International Standard Book Number.

ISSN
: The International Standard Serial Number.

LCCN
: The Library of Congress Call Number

abstract
: An abstract of the work.

keywords
: Key words used for searching or possibly for annotation.

price
: The price of the document.

copyright
: Copyright information.

language
: The language the document is in.

contents
: A Table of Contents

---

21 January 1995
Dana Jacobsen
dana@acm.org