Bibliographies and Citations for CompSci Papers

Arjen Rouvoet

September 23, 2021

Not as Easy as it Seems

Compiling bibliographies and adding citations to a LaTeX manuscript is very easy: one copies BibTeX entries from Google Scholar or some such, adds a `\cite{CiteKey}` and all is good, right?

While indeed the LaTeX and BibTeX infrastructure makes it relatively easy to get going, one still needs to pay attention to produce good bibliographies.[1] This short guide is intended to explain some of the principles and technicalities of BibTeX and bibliographies for PL papers in order to save you a lot of frustration and help you produce better bibliographies.

Although type-setting bibliographies may initially appear as an annoying mechanical issue, understanding the bibliography basics will actually also give you a rudimentary understanding of the publishing landscape for papers. Hence, a decent understanding of BibTeX is a useful skill for a PhD student to develop comprehensively as early as possible.

[1] This is true for typesetting in general.

Bibliographies

The first thing to do is to assume that every bibtex entry you pull from the internet contains mistakes or is inconsistent with other entries. It is important to decide on some standards and to apply them consistently in your bibliography file.[2]

We will now go over a number of different prototypical publication types and look at their BibTeX entries and the corresponding entries in the printed bibliograpy. For each type of publication we only include the essential fields in the recommended style.

[2] Because it takes some effort to fix bibliographies, it is recommended that you maintain a bibliography that you can reuse across publications. If you use an online database like Researchr, it may also be appreciated when you submit fixes upstream so that others can benefit.

Journal or Magazine Articles

Articles published in a journal or magazine have an `@article` entry:

```
@article{OHearnP99,
  author       = {Peter W. O'Hearn and David J. Pym},
  title        = {The logic of bunched implications},
  year         = 1999,
  journal      = {Bulletin of Symbolic Logic},
  volume       = 5,
  number       = 2,
  pages        = {215--244},
  doi          = {10.2307/421090},
}
```

Journals have names and appear on a periodic basis in issues that are identified by a volume and a number.[3] The article is published in a given year, on certain pages of the issue. An article typically also has a DOI that uniquely identifies the publication. Including the DOI should be enough to hyperlink the entry in the printed bibliography; make sure to test that it leads to the right article.

[3] PACMPL papers use a conference identifier (POPL, OOPSLA, ICFP) as the issue number.

An article entry in a bibliography produced by biblatex using author-year citation style will look something like this:

> Peter W. O'Hearn and David J. Pym (1999). "The logic of bunched implications". In: *Bulletin of Symbolic Logic* 5.2, pp. 215–244. DOI: 10.2307/421090

Conference Papers

Articles published in the proceedings of a conference get an `@inproceedings` entry:

```
@inproceedings{KatsV10,
  author       = {Lennart C. L. Kats and Eelco Visser},
  title        = {The {Spoofax} language workbench: rules for
    declarative specification of languages and IDEs},
  year         = 2010,
  booktitle    = {{ACM} {SIGPLAN} Conference on Object-Oriented
    Programming, Systems, Languages, and Applications ({OOPSLA})
    },
  pages        = {444-463},
  doi          = {10.1145/1869459.1869497},
}
```

The booktitle should ideally clarify what kind of venue the paper was published at (workshop, conference, SIGPLAN conference, ...). We also suggest to include acronyms for conferences because they are recognizable. Pay attention to schoolnames so that they are typeset consistently and according to the preferences of the school.

Some conference proceedings are part of series such as the Lecture Notes in Computer Science (LNCS) series. This is recorded in the entry using the **series=** field and a **volume=** entry, like so:

```
@InProceedings{ChapmanKNW19,
  author       = {James Chapman and Roman Kireev and Chad Nester
    and Philip Wadler},
  title        = {System {F} in Agda, for Fun and Profit},
  year         = 2019,
  booktitle    = {International Conference on Mathematics of
    Program Construction (MPC)},
  series       = {LNCS},
  volume       = {11825},
  pages        = {255-297},
  doi          = {10.1007/978-3-030-33636-3_10},
}
```

As a matter of style, we suggest to abbreviate well-known series like LNCS.

Conference paper entries in the bibliography look like this:

Lennart C. L. Kats and Eelco Visser (2010). "The spoofax language workbench: rules for declarative specification of languages and IDEs". In: *ACM SIGPLAN Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA)*, pp. 444–463. DOI: 10.1145/1869459.1869497

James Chapman, Roman Kireev, Chad Nester, and Philip Wadler (2019). "System F in Agda, for Fun and Profit". In: *International Conference on Mathematics of Program Construction (MPC)*. vol. 11825. LNCS, pp. 255–297. DOI: 10.1007/978-3-030-33636-3_10

Books and Chapters of Books

Books are relatively uncomplicated:

```
@book{Pierce2002,
  title        = {Types and programming languages},
  author       = {Pierce, Benjamin C},
  year         = {2002},
  publisher    = {MIT press},
  ISBN         = {0-262-16209-1}
}
```

Sometimes one may want to cite a particular chapter of a book or collection of notes, in which case the `@incollection` entry may serve well:

```
@incollection{Walker2005,
  author       = {Walker, David},
  title        = {Substructural type systems},
  editor       = {Pierce, Benjamin C},
  booktitle    = {Advanced topics in types and programming
    languages},
  year         = 2005,
  pages        = {3--43},
  chapter      = 1,
}
```

Such entries are displayed respectively as follows:

Benjamin C Pierce (2002). *Types and programming languages*. MIT press. ISBN: 0-262-16209-1

David Walker (2005). "Substructural type systems". In: *Advanced topics in types and programming languages*. Ed. by Benjamin C Pierce. Chap. 1, pp. 3–43

Theses

Special entry-types exist for MSc and PhD theses:

```
@mastersthesis{Aerts2019,
  title        = {Incrementalizing Statix: A Modular and
    Incremental Approach for Type
    Checking and Name Binding using Scope Graphs},
  author       = {Aerts, Taico},
  year         = {2019},
  school       = {Delft University of Technology, {NL}},
  type         = {MSc thesis},
  url          = {http://resolver.tudelft.nl/uuid:3e0ea516
    -3058-4b8c-bfb6-5e846c4bd982}
}

@phdthesis{Swierstra09,
  author       = {Wouter Swierstra},
  title        = {A functional specification of effects},
  school       = {University of Nottingham, {UK}},
  year         = 2009,
  url          = {http://eprints.nottingham.ac.uk/10779/},
}
```

Note the use of **type=** to ensure that it displays as a MSc thesis, and not the default MSc thesis. We suggest that you include a (stable) **url=** for theses if the work has no DOI/ISBN identifier.

These entries are displayed as:

Taico Aerts (2019). "Incrementalizing Statix: A Modular and Incremental Approach for Type Checking and Name Binding using Scope Graphs". MSc thesis. Delft University of Technology, NL. URL: http://resolver.tudelft.nl/uuid:3e0ea516-3058-4b8c-bfb6-5e846c4bd982

Wouter Swierstra (2009). "A functional specification of effects". PhD thesis. University of Nottingham, UK. URL: http://eprints.nottingham.ac.uk/10779/

Style Guide

- Make **booktitle=** entries consistent such that multiple entries for the same conference series are displayed similarly.
- Give **journal=** entries in full (not abbreviated).
- Abbreviate well-known **series=** entries like LNCS.
- Conference names used to identify an issue for an article are symbolic and are usually abbreviated (**number={POPL}**).

Pitfalls & Solutions

The papers of a number of conferences (POPL, OOPSLA, and ICFP) are published nowadays in the PACMPL journal. This means that recent papers in those conferences have `@artice` entries, whereas older ones have `@inproceedings` entries.

Pay attention to how authors spell their own names! BibTeX does not always do the right thing, and databases get it wrong all the time. BibTeX tries to identify middle-names, but may fail if someone has multiple last names. You can force multiple parts to be treated as a unit. For example, it is `author = {Casper {Bach Poulsen}}` (last name is Bach Poulsen), not `author = {Casper Bach Poulsen}` (Bach would be treated as a middle name).

Another common source of errors is the word "van" or "von" in Dutch or German names. In Dutch, the "van" is capitalized *only* if it is not preceded by another name part. E.g., "Van Antwerpen et al (2009)" and "Van Antwerpen, Hendrik", but "Hendrik van Antwerpen". In any case, a bibliography should sort based on "Antwerpen", not based on "van". The rules for German names are different.[4]

[4] I know of no good way to convince BibTeX to always do the right thing.

A common error in BibTeX entries pulled from online databases are spurious backslashes in DOI, so that they do not link to the right pages. Test them!

Other Resources

- Paperpile on bibtex entries.
- ACM manual on BibTeX formatting.

Ack

This document incorporates lessons taught by Eelco Visser, Robbert Krebbers, Casper Bach Poulsen, and Jeremy Gibbons and others who helped me get my bibliographies in order. Thanks!