Scientific Communication CITS7200

Computer Science & Software Engineering

Lecture 5

Citation and Reference

Throughout this lecture, *citation* is used as the name for the mention you make in your text of a source, and *reference* as the name for the bibliographic item that appears in the list at the end of your text.

Citations

There appear to be many conventions in scientific writing for dealing with citations, and many conventions for the format of the references. In this lecture I will use material mostly drawn from Mary-Claire van Leunen's book, *A Handbook for Scholars*, although examples from the computer science literature will also be given. What is most important is that a set of rules for citation and reference is chosen and consistently used, so that the reader is not bothered with trying to understand the meaning of any inconsistencies.

The one thing that distinguishes scholarly writing from any other type is its acknowledgment of sources. Without citation, you fail to give your reader the opportunity to put your work into context, or to judge your work independently.

A common error with young scholars is to over-cite, but this is still far better than no citation at all.

> †Knuth [1,5,6,8] and many others [3,9,11,15,23] purport to show that the algorithm [4,7,10,12] has complexity [2] that must be

> bounded above [5] by the complexity of the simulated quicksort [15,16,18,19,21,25].

Used properly, citation will strengthen your arguments, and simplify your exposition.

Citation is also a form of good manners amongst scholars. It is a sign of respect to one's peers and a way of saying "thank you" to those that have gone before you in your field.

It is now out of fashion to use bibliographic footnotes, especially in the computer science literature. Content footnotes, however, are still common.

For example, don't write

> †Nilsson[1] argues that the minimax paradigm is the most common form of artificial game playing and, moreover, that it resembles the neuronal function observed in human game playing.

Rather, use a bracketed number on the line instead of the superscript, and use that number as a pointer to a numbered item on the reference list.

> Nilsson [1] argues that the minimax paradigm is the most common form of artificial game playing and, moreover, that it resembles the neuronal function observed in human game playing.

The corresponding item appears only once in the reference list, no matter how many times it is cited in the source.

[1] N. J. Nilsson. *Principles of Artificial Intelligence*. Tioga Publishing Company, Palo Alto, Calif. 1980.

Getting rid of bibliographic footnotes makes your page of text visually more appealing and avoids interrupting the reader's attention. Note that you should *not* use a pointer as a word, and that there should be a space between the citation and the pointer. You must also make sure that your citations are meaningful. Don't let details be hidden in your pointers; make your writing concrete and vivid.

> †[5] shows that tabu search is a viable alternative to simulated annealing in this application, although [11] and [15] also give many examples where genetic algorithms work equally well.

[1] N. J. Nilsson. *Principles of Artificial Intelligence*. Tioga Publishing Company, Palo Alto, Calif. 1980.

> †It has been argued [6] that within the next twenty years all our entertainment will be supplied by a virtual reality industry.

Such a sentence is too vague. You can make it more concrete by adding the name of the author:

> Simpson [6] has argued that within the next twenty years all our entertainment will be supplied by a virtual reality industry.

This is fine if Simpson is a household name amongst your readers, but the reference may still be opaque if your audience is not aware of the virtual reality literature. In such a case, the following citation might be more appropriate:

> Simpson [6], emeritus professor of computer science at Cambridge University, has argued that within the next twenty years all our entertainment will be supplied by a virtual reality industry.

Try to avoid citations without references, that is, by including the full reference within the text. This is most certain to be cumbersome in any writing in computer science.

> †Donald E. Knuth's *The Art of Computer Programming* (Volume 3, Sorting and Searching, Addison-Wesley, 1973) gives a lucid description of the cascade merge algorithm.

However, when do you want to merely use your citation as a simple pointer? The following cases are all considered to be appropriate:

- When your citation is a simple courtesy to the original discoverer of the result you are using:
  - Since the Fourier coefficients eventually tend to zero [17], we can confidently set the parameter λ to zero as well.
- When you are citing too many sources to mention each one gracefully.
  - The use of elastic templates in image compression [2,5,7,8,9] has led to the highest compression ratios outside of the fractal approach.

- When you are giving so complete a summary that your reader need never look at the source.
  - There are three ways of measuring the difference between a source and destination binary image - by using a simple mean-squared pixel difference; by Peli and Malah's Figure of Merit function, which measures the normalized distance of a white pixel to the closest black object; and by Baddeley's measure, which gives and L^p variant on the Figure of Merit measure for $1 \leq p \leq \infty$ [13].

This is much better than saying

> †According to a research paper on the subject [13], there are three ways of measuring the difference between a source and destination binary image.

Try to avoid too much citation. Assume that your reader has a reasonable attention span; you don't need to keep citing every time you mention the source.

> †Knuth [5] developed TEX in the 1970's and Lamport [6] extended it to LATEX about a decade later. TEX [5] is a powerful typesetting program, but LATEX [6] has a number of macros that make it easier to use for typesetting mathematics. Both TEX [5] and LATEX [6] are widely used in academic communities.

Group bracketed numbers whenever you can do so without causing confusion.

> Both TEX and LATEX [5,6] are widely used in academic communities.

Page numbers should not be mixed in with citations. They properly belong to the reference list. Thus, do not write

> †Einstein said that the physics should "represent a reality in time and space, free from spooky actions at a distance" [17, page 223]. Pauli railed against tormenting oneself about "the problem of whether something one cannot know anything about exists all the same" [17, page 158].

Rather, the bracketed references should be simply [18] and [17] respectively, with the entry in the bibliography listed as:

[17] Amron Bevels, editor. *Questions in the Philosophy of Science.* Wiley, 1979, page 158.

[18] _____ , page 223.

It is not usual to quote directly from other papers in scientific writing, except possibly in the history of science.

Personal titles should not be quoted in scholarly writing. If you do this, it is taken as a form of sarcasm rather than respect. People are quoted for the work they have done, not for the position they hold.

Particularly in the sciences, you are liable to cite papers by multiple authors. You should name up to three authors, but if the paper has any more than this, you should give the first author's name, followed by the expression *et al.* This stands for the Latin *et alii*, or *et aliae*, meaning "and others". Notice that there is a period after the *al.*, but not the *et.* You use this expression only after the first author's name, not with the first and the second, or the first, second, and third.

> †Stein, Jacobs, Sherman, Polst and Chandra [5] have implemented this algorithm on the Datacube.

> †Stein, Jacobs *et al.* [5] have implemented this algorithm on the Datacube.

> Stein *et al.* [5] have implemented this algorithm on the Datacube.

Always give the full reference in the reference list; never use *et al.*

Everything you cite is going to be a book or a non-book. A book is a book or a journal: a fat thing bound by itself; or the equivalent in some other medium of a fat thing bound by itself. A non-book is an article or a story or a poem or a pamphlet: a thin thing or a thing that is part of something else; or the equivalent in some other medium of a thin thing or a thing that is part of something else.

Examples of books include a book, a journal, a magazine, a newspaper, a conference proceedings, a PhD thesis, a short novel, an opera, a musical, and a record album.

Examples of non-books include an article in a journal or magazine, even if it takes up one whole issue; a Master's thesis or any other homework assignment

before a PhD; a short story; a poem; a chapter of a book; letters, memos, notes, recipes, formulas, algorithms and theorems; arias, songs, and a band; sketches, prints, paintings, and sculptures.

When you cite a book title in your text, you should use italics and uppers-and-lowers capitalization, that is, capitalize the first word, the first word after a colon, and all other words except articles, unstressed conjunctions, and unstressed prepositions.

When you cite a non-book in your text use quotation marks around the title, and uppers-and-lowers capitalization.

> Pentland's paper "Linear Shape from Shading" appears in Volume 4 of the *International Journal of Computer Vision*.

References

Easy references fall into a simple block format of which the three main blocks are:

1. Author.
2. Title.
3. Bibliographic information

The most common references are to books and journal papers. For example,

[1] Berthold Klaus Paul Horn. *Robot Vision*. MIT Press, 1986.

[2] L. S. Davis. A survey of edge detection techniques. *Computer Graphics and Image Processing*, Vol. 11, No. 4, pp 248-270, September 1975.

Each block starts with a capital letter and ends with a period.

Authors In the author block you should list the full author name given on the work. This is important. Authors should choose the name by which they wish to be known professionally, and then stick to it. Even if you know that L. S. Davis is "Larry Davis", or "Laurence Spencer Davis", use the author reference that he has given on the paper. This is the stage at which you should be deciding upon your own appellation.

List the authors' names in normal order, not in reverse order. Reverse order might be helpful for alphabetizing, but it makes a hash of joint authors:

†Cole, A. A., Hsu, P., and Sastry, S. Dynamic control of sliding by robot hands for regrasping. *IEEE Transactions of Robot Automation* 8(1), pp 42-52, 1992.

A. A. Cole, P. Hsu, and S. Sastry. Dynamic control of sliding by robot hands for regrasping. *IEEE Transactions of Robot Automation* 8(1), pp 42-52, 1992.

People's initials should be spaced out: *not* A.A.Cole, but A. A. Cole.

Corporate authors are dealt with as follows:

[13] IBM. Elementary math library. *Programming RPQ P81005, Program number 5799-BTB, Program reference and Operations Manual*, SH20-2230-1, August 1984.

There are two kinds of editors, the kind who compiles the work of several authors and the kind who prepares the work of a single author.

The compiler-editor gets listed in block 1 if you are referring to the whole compilation:

[15] Michael Brady, editor. *Robotics Science*. MIT Press 1989.

He gets listed in block 3 if you are referring to part of it:

[18] Marc H. Raibert. Legged robots. In Michael Brady, editor, *Robotics Science*, MIT Press, 1989.

The preparer editor usually gets mentioned in block 3:

[11] Alan Turing. *The Collected Works.* Edited by Anthony Hodges, Oxford University Press, 1988.

Titles When you cite a book title in your reference list you should use italics, with uppers-and-lowers capitalization. When you cite a non-book title, neither the quotation marks nor uppers-and-lowers which are used in the text are necessary. Use sentence capitalization (capitalize the first word, the first word after a colon, and no other words except proper nouns and proper adjectives).

If you are referring to something written in a foreign language, it is a courtesy to give a translation for your monolingual readers:

Rachid Deriche. Mise en oeuvre recursive de la gaussienne et ses derivees [Recursively implementing the gaussian and its derivatives]. Rapports de Recherche No. 1893, INRIA, April 1993.

Always give the full title and sub-title, even if it seems somewhat crude.

†The whore of Babylon.

The whore of Babylon: Being an inquiry into the iniquities of the pope of Rome, the beast with seven heads, most assiduously compiled and brought together with great care and its author hopes some purpose.

Bibliographic information This section simply tells the reader how to find the source. Give a complete address for an obscure publisher, but no address for a well-known publisher. Of course, this is sometimes difficult to judge, so if you think a complete address will help your reader find the source, give it. Always give the date of the publication.

[1] Harold Abelson and Gerald Jay Sussman with Julie Sussman. *Structure and Interpretation of Computer Programs*. The MIT Press, 1985.

[2] John Ehrenfried Hofmann. *Classical Mathematics: A Concise History of the Classical Era in Mathematics*. The Philosophical Library, 15 East 40th Street, New York 16, N. Y. 1959.

There are two other optional blocks that are sometimes useful when compiling your reference list:

4. More bibliographic information.
5. Annotation.

An example of when block 4 might be used is when the publication has appeared more than once, for example:

[5] M. Segal. The use of B-splines in CAD. *Computer Graphics and Manufacturing*, IV, pp 43-57, 1989. Also published as Stanford Technical Report in Artificial Intelligence, TRAI-88/76, 1988.

The occasions for using the annotation block are rare, but occasionally you might want to warn the author about a flawed source:

[11] C. Wiles. Proof of Fermat's last theorem. *Journal of Elliptic Function Theory*, Vol. 2, pp 12-145, 1992. This proof was later shown to be incomplete.

Here are some more examples of references, some easy, some hard:

[1] M. J. Nelson. *A Control Architecture Towards Intelligent Behaviour*. PhD thesis, The University of Western Australia, 1991.

[2] John Francis Canny. Finding edges and lines in images. Master's thesis, MIT, 1983.

[3] Walter J. Stein. On the unsolvability of P = NP. To appear in *The Journal of Foundations in Computer Science*.

[4] J. Lansdown. The creative aspects of CAD: A possible approach. *Design Studios*, 8, No. 2, pp 76-81, April 1978. Cited in Richard Wright, The image in art and 'computer art'. *Leonardo*, Supplementary Issue 1989, pp 49-54.

[5] N. Goodman. *Languages of Art: An Approach to the Theory of Symbols*. Privately printed, 1898. Available on microfilm from the Vision Sciences Project, University of North Carolina.

[6] Joan Wells. Hypertension in confined adult primates. University of Washington Medical School Report IIR3219 (microfiche), 1986.

[7] Paul Klee. Drawing, 1922. The Art Institute of Chicago, Prints and Drawings Collection, catalog #6/21.133.

[8] *The Oxford English Dictionary on Compact Disc*. Text, Oxford University Press, 1884-1933. Software, TriStar Publishing, 475 Virginia Drive, Fort Washington, Pennsylvania 19034-9930.

If you want to refer to an electronic message, you have to provide your own archive. Make a paper copy and put it in your own files. Refer to the mail message as follows:

[9] Martin Minow. Re: Electronic house arrest units. RISKS-FORUM Digest (`risks@csl.sri.com`), 10(26), 29 August 1990.

If the email was sent only to you, then it is a private communication:

[10] Jack Stolti. Private communication, 27 September 1993.

When citing a film, you usually give the director's name, the title of the film, the producing company, and the date.

[11] Steven Spielberg. *Jurassic Park*. United Artists, 1992.

The computer music literature often includes a section entitled Discography, as well as a bibliography.

> [12] Johann Sebastian Bach. Von Himmel hoch, da komm'ich her (BWV 769, 1738). Played by Helmut Walcha. Archive ARC 3030, 14553 APM. Approximately 1960.
>
> [13] Nicolas Collins, producer. Real imaginary music. *Imaginary Landscapes*, compilation, Nonesuch Records, CD and cassette, 1990.

For a piece of software, either the source or the runnable code, give the name of the program, the version number, the platform on which you have run it, and where to get it:

> [14] Tru-Champ 3.4 for the Macintosh. Electronic Equipment Corporation, 1483 Meller Road, Cambridge, Massachusetts 02138.
>
> [15] CLINGER for VAX VMS. Computer Science Department, University of Texas at Austin.
>
> [16] Richard Stallman. GNU emacs 13 for Unix. Free Software Foundation, 675 Massachusetts Avenue, Cambridge, Massachusetts 02139.

Using BIBTEX

BIBTEX is an auxiliary program to LATEX that automatically constructs a reference section for a LATEX document by searching one or more bibliographic databases. These databases contain the information about the publications, and are given a `.bib` extension. The advantages of using BIBTEX, rather than the `thebibliography` environment are that the same database can be used for all types of publication, and that the styles required by different publication formats can be generated automatically. Moreover, typographical errors are minimised by minimising the number of times reference details are typed.

The LATEX file must contain the command

`\bibliography{`*database1*, ...`}`

at the point in the text where the bibliography is to appear. Here *database1*, ... is a list of the bibliographic database files that are to be searched. The `.bib` extension is not explicitly written.

A citation can be made to one of the references at any place in the text by using the command

`\cite{` *key* `}`.

The *key* is the database identifier for that publication.

After the first LaTeX processing, the BIBTeX program must be run over the LaTeX file. Say the file is called `thesis.tex`. Then we run

```
bibtex thesis
```

to produce a file called `thesis.bbl`, containing the extracted information for those publications for which there is a `\cite` reference in the text. LaTeX must then be re-run at least twice to establish both the bibliography and the in-text citation labels. Occasionally you might want to include references in your bibliography that are not cited in the text. These can be added with the command

`\nocite{` *key* `}`

given anywhere in the text.

The style of the bibliography can be selected with a style declaration of the form

`\bibliographystyle{` *style* `}`

anywhere after the preamble. Some of the *style* arguments are as follows:

- `plain`: Entries are ordered alphabetically; each is assigned a running number in square brackets as the in-text reference pointer. For example, the BIBTeX entry

  ```
  @article{kass88a,
          author = "Kass, Michael and Witkin, Andrew and
                    Terzopoulos, Demetri",
          title = "Snakes: Active Contour Models",
          journal = "The International Journal of Computer Vision",
          volume = "2",
          number = "",
          pages = "321--331",
          month = "",
          year = 1988
  }
  ```

  would produce the reference

[1] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *The International Journal of Computer Vision*, 2:321–331, 1988.

- `unsrt`: Entries are ordered according to chronological citation, but otherwise this style is the same as `plain`.
- `alpha`: The reference list is the same as for plain, but the citation pointer is an abbreviation of the author's name plus year of publication. For the example above, the citation pointer would be [KWT88].
- `abbrv`: The ordering and citation are the same as for `plain`, but the reference list is shortened by abbreviating first names, months, and journal names.

There are numerous standard entry types, with varying required and optional fields. For example, the type `@article` has required fields `author, title, journal`, and `year`, whilst the optional fields are `volume, number, pages, month`, and `note`. Similarly, the type `@phdthesis` requires `author, title, school`, and `year` and optionally includes `type, address, month`, and `note`.

Additional field names can be included, and will be ignored by BIBTEX. For example, to add the abstract of an article as a field, you simply enter

`abstract = {` *text of the abstract* `}`.

The authors and titles fields often require special formatting for BIBTEX to parse them correctly. BIBTEX assumes that if there is no comma in the author field, then the last capitalized name is the surname; otherwise, what comes before the comma is the surname. If the person has a double surname, then it should be entered in braces. The same rule applies for other complications, such as Jr. or accented letters. BIBTEX uses the conventions for capitalization that we have already covered. Thus, if the convention is sentence capitalization, then any proper nouns that should be capitalized must be entered in braces. These rules are illustrated in the final example:

```
@inproceedings{kn:ICASSP90,
Author="Alan D. {Calway Jr.} and Roberto G{\"o}del",
Title="Curve extraction in images using the
       multiresolution {F}ourier transform",
Booktitle="Proc. Int. Conf. Acoust., Speech, and Signal Processing",
Publisher="IEEE",
Pages="2129--2132",
Month="April",
Year=1990,
}
```

The simplest way to use a BIBTEX file is to create a series of templates for the entries, and then fill in everything you might eventually need as you read the publication. Use of the `note` or `abstract` field can be particularly useful when you come back to a reference a long time after reading it. The filter `u2btx`, available in `/usr/contrib/bin` is useful for converting the output of the Library's Urica program to BIBTEX format.