

# Searching, reading, and referring literature

Wilhelmiina Hämäläinen

March 13, 2006

## 1 Need for references

In scientific writing, we use a lot of references!

- All text must be justified, either based on previous research or your own results.
- It must be clear what the information is based on!
- Often the whole master thesis is based on systematic study of existing literature. The information is just analyzed and organized from a new point of view.
- The sources for scientific writing must also be scientific!

## 2 Source types

The literature sources can be divided into three groups:

1. **Primary sources:** articles in conferences and journals
  - original sources
  - the papers should have appeared in a reviewed journal/conference (i.e. reviewers have checked their correctness)
  - also technical reports and other theses
2. **Secondary sources:** textbooks, encyclopedias, glossaries
  - sometimes useful analysis or interpretation, but not original sources
  - you can use these in master thesis, but only as supplementary material

### 3. Bibliographies

- support information retrieval
- lists of articles + references
- scientific search engines are on-line bibliographies

**Task:** Can you trust the information you find in wikipedia? Why or why not? Why wikipedia cannot be used as a reference in a scientific text?

## 3 Collecting literature

Starting point: your preliminary topic.

- goal
- central concepts, theories and themes

### How to proceed?

- Begin from familiar: notes, textbooks
- Ask your supervisor
- Check references in useful papers or books
- Make key word queries in scientific bibliographies (e.g. ACM, IEEE, Elsevier)
- If you make an internet query, prefer scholar google. Check always that the paper has been published!
- Write down the references – they can be hard to find afterwards! (especially store the bibtex files)

## 4 Reading

- You cannot read everything throughout!
  - ⇒ Read only as much as is needed to
    - recognize that the article is useless
    - get the useful information

- Often an iterative process: important articles are read several times!
  - Title and abstract
  - Scan through introduction and conclusions/summary
  - Check references: new good references?
  - Important or useful sections and subsections (the organization is usually described in the introduction)
  - In the beginning, don't get stuck in details; don't check individual words or references; believe the arguments
  - If the article is important, then try to understand it properly, and check the referred sources
- Ask yourself:
  - What is the main idea?
  - What is the contribution (the new or interesting thing)?
  - What is important for you? Where it is presented?
- If you don't understand the article
  - Try to invent examples or simulate the solution yourself
  - Ask your fellows, supervisor, experts
  - Ask (yourself and others) specified questions: Where this equation comes from?, What is the relationship between these algorithms? Can you give an example for this definition?
  - Often understanding happens as a background process!

## 5 References

### 5.1 Referring in the text

- The reference is usually immediately after the referred theory, algorithm, author, etc.

"According to Dijkstra [Dij68] goto statement should be avoided..."

"*Bloom filters* [Ref03] solve this problem..."

- The reference is in the end, if you refer to the whole sentence or a paragraph. (before full stop, if it refers only to the previous sentence, otherwise after the full stop)

”Goto statement should be avoided [Dij68].” Notice the difference: now you agree with Dijkstra!

- Sometimes there is no one ”original” source, but a new concept or theory has developed little by little. In this case, you can give a couple of example references where the reader can find more information.

”*Context-aware computing* (see e.g. [DeA99,CaK00]) is a new approach...”

### Other examples

”Minsky and Papert [MiP69] showed that...”

”Version spaces were introduced by Mitchell [Mit77].”

”Nonparametric methods are described by Randles and Wolfe [RaW79].”

”The principles of CART were first described in Breiman et al. [BrF84].” or

”The principles of CART were first described in [BrF84].”

”Prolog was primarily used for writing compilers [VRo90] and parsing natural language [PeW80].”

”The general procedure for skolemization is given by Skolem [Sko28].”

”Other methods are summarized in e.g. [Bro92,Woo96].”

”The problem is *NP*-complete [Coo00].

## 5.2 Reference notations

- A common style: three letters from the authors’ names + the last numbers from the year. E.g. [Ham06]
- Sometimes numbers
- A humanist style: surname + year. E.g. [Hämäläinen, 2006]

## Notes

- If you refer to a book, give the chapter or the page numbers!
- If you use only one chapter from a book, you can give the chapter number and title in the reference list. If you use several chapters, give the chapter number in the reference: [WMB94, chapter 2]
- The page number is always given in the text "[Bro92,pp.3-7]"
- If you have several references, list them together: [Bro92,Woo96]

## 5.3 Reference list

The last chapter in your thesis (or section in a paper) is called References. For each source, give

- The authors: surname and the first letters of the first names. If you have  $\geq 3$  authors, give only the first one, and replace the others by "et al." E.g. "Mitchell, T.M. et al."
- The title
- Publisher, (place) and year.
- Page numbers, if the source is a paper or a chapter in a collection written by several people.
- The title and the editors of the collection, if the paper has appeared in a collection (e.g. conference articles).
- The volume (always!) and the issue number after a comma or in parentheses, if the source is a journal paper.
- Series, if the book has appeared in some series. (E.g. Lecture Notes in Computer Science + number)

### Examples:

Bourne, S. The UNIX System. International Computer Science Series, Addison-Wesley, 1982. (a book)

Gannon, D. et al. Programming environments for parallel algorithms. In *Parallel & Distributed Algorithms*, ed. M. Cosnard et al. North-Holland,

1989. 101-108. (an article in a collection)

Grahne, G., Nykänen, M., Ukkonen, E. Reasoning about strings in databases. *Journal of Computer and System Sciences* 59, 1 (1999), 116-162. (an article in a journal)

- More examples in the exercises!
- Notice that the journal and book titles are written with capital letters!

## 5.4 In latex:

- Latex creates the notations automatically!
- You can select the style by setting the style parameter for the bibliography environment
- Just invent a unique label string for each source, which you use in references by command `\cite`. E.g. `\cite{whamalai}`, or if you want to refer page 3, `\cite[3]{whamalai}`
- In the References, define what the label refers
- If you have alot of sources, you can manage them automatically by `bibtex` (we will return to `bibtex` later in this course)

We will practise these in the computer class!

## 6 Citations

Direct citations are seldom used in cs texts.

If you use them, make clear who is responsible for what!

- If you express somebody else's ideas by your own words, then put the reference immediately after the idea.
- If you express somebody's ideas by her/his own words, then it is a citation!
- If quotation marks "..." are missing, it is called plagiarism!

- As a rule of thumb: if you borrow more than 7 words, then use quotation marks.
- If the citation is translated, then mention also the translator in reference.
- If you add or drop words, show it by [] or ....
- If you emphasize words, mention it.
- An example:

Nykänen [Nyk03] remarks that unreferred citation is plagiarism (translation and emphasis by the author): "If you borrow more than *seven* words ... from a text it [borrowing] is called *literary theft*."

## 7 Your own opinions?

By default: no opinions, everything must be based on facts!

If you have to express your own opinions, then

- *In principle*, everything without references is your own interpretation.
- However, make clear, what is borrowed and what are your own opinions!
- Often clearer to write a separate section called "Discussion".