Introduction to Graduate Studies

Use of Scientific Literature and Writing (during your graduate studies)

Jaakko Hollmén
Laboratory of Computer and Information Science, Helsinki University of Technology

Sources for this lecture
- Own experience

How much do we write?

If you hate writing...
You can consider two things:
- a change of attitude
- a change of career

Writing: 2nd half of lecture
- No grammar, no punctuation rules: guide books and language courses available
- Purpose of writing articles
- Conventions in scientific writing
  - structuring your articles and your thesis
- Preparing your manuscript
- Technical things about scientific writing
1st half: Scientific literature

- What discriminates scientific research from other kinds of work?
- Different kinds of literature
- How to find it? Where to find it?
- How to refer to literature?

Example: Document retrieval

- Two computer scientists implement exactly the same document retrieval system
  - "My system is very efficient and probably beats everything else!!!"
  - "The document retrieval system using the HITS algorithm by Kleinberg [1] extended with a quick update algorithm of the hub and authority weights presented in this paper is empirically shown to retrieve documents with 90% precision and 87% recall in a chosen set of test documents (see Appendix). Comparisons with PageRank [2], BitBoost [3], and NibbleTwist [4] demonstrated the superiority of the algorithm."

- Formalism required for scientific work!

Thesis requirements at HUT

- The dissertation shall contain new scientific knowledge in the field it represents.
- The measurements and other experimental data included in the dissertation shall meet the demands set for scientific research.
- The doctoral candidate shall present the results and conclusions clearly and in such a way as to meet the scientific demands.
- The dissertation shall include an abstract.

Scientific literature

- Open literature, "open source knowledge"
- Collection of contributions by researchers
- Conference articles, journal articles, collections of articles, textbooks, etc.
- Libraries: distributed archiving
- Articles have been reviewed by research colleagues (peers) and accepted!

Your contribution

- Your original contribution adds to the scientific literature
- Clearly written, should be pleasant to read and learn
- Well documented so that it can be repeated and thus validated
- Approved by a peer review process
Your work vs. others’ work

- In scientific work, you must make a clear distinction between your own contribution and general knowledge.
- After all, you want to tell the world what your contribution is and state clearly what other contributions it is based on.

How to find relevant literature?

- Your instructor/supervisor tells you.
- Do literature searches in the library or with databases of the library.
- Visit and wander in the library [of your lab].
- Talk to/with your colleagues.
- Attend to conferences.
- Follow good journals in your area.
- Talk/write with the real authors.

Literature search done wrong

- You search in the Internet.
- You may easily neglect the formal requirements of scientific literature.
- Not all literature is available in the Internet.

Example: Literature search

- Database search in the library database that contains references.
- TASK: “Find work on detecting people that crack into a UNIX system and do bad things.”
- Topic is given, formulate a query based on the available information: author, title words, keywords, publication year, etc.

After finding the reference(s)

- You need to find the full text of an article:
- Library databases (access to publisher’s electronic database of full articles).
- Your own library.
- Other libraries, distant libraries.
- The author (or the author’s homepage).

Example: Finding the full paper

- Find the full text of the article, use whatever sources possible.
Example: Finding a book

- Find the book of your interest, use whatever sources possible

After searching & learning...

- You develop something new in your own topic of interest based on the existing knowledge
- You publish it = Thinking + Writing
- You share your findings openly with the public

In short:

- Scientific literature is your input
- In between you accomplish wonderful research results, and you want to tell the whole world about them!
- Writing produces the output
- ...that is somebody else’s input

Acknowledge your sources

- Your findings are based on those of others
- In your work, you should give credit to those people by making references to their work in the right context
- Citation: author, title, publication forum with full information, date, etc.
- Reference so accurate that the source articles can be identified and found

Conference article

```latex
@InProceedings{Hollmen99,
author = {Jaakko Hollmén and Volker Tresp and Olli Simula},
title = {A Self-Organizing Map Algorithm for Clustering Probabilistic Models},
booktitle = {Proceedings of the Ninth International Conference on Artificial Neural Networks (ICANN'99)},
year = {1999},
month = {September},
pages = {946--951},
}
```

Journal article

```latex
@Article{Dempster77,
author = {A. P. Dempster and N. M. Laird and D. B. Rubin},
title = {Maximum likelihood from incomplete data via the EM algorithm},
journal = {Journal of the Royal Statistical Society, Series B},
year = {1977},
volume = {39},
pages = {1--38},
}
```
Example: References in your doc

- You want to include the mentioned documents in your newest article
- Demonstration about LaTeX and BibTeX
- Bibliographic entries are stored in a database and included in your paper as needed

Your thesis should be built on:
- A complete library search in your area
- Original research: field and laboratory measurements as approved by your professor
- Your syntheses putting together and deriving meaning from data, ideas from others and your own conclusions (Martha Davis, 1997)

Another view on a thesis:

“A thesis is an unusually long paper and includes a review, so it resembles a book and is usually divided into chapters”

(Vernon Booth, 1984)

Writing during your graduate studies

- Conference articles
- Journal articles
- Doctoral thesis
  - introduction and a collection of articles OR
  - monograph
- Other forms of communication:
  - Slide presentations
  - Poster presentations

Before you write

- Good notebook discipline
- Tell colleagues about your work
- Take 8 sheets of paper: Title, Summary, Intro, Materials, Methods, Results Discussion, References and write ideas as they come
- Prepare tables and figures
**When you write**
- Set the paper aside for some time
- Arrange material into sections (8 sheets)
- Begin with the easiest section
- Formulate title and keywords:
  - title: short, specific, not too general
- Summary:
  - what you did & main results
  - conclusion as the last paragraph

**Arrangement: Introduction**
- Introduction should state the problem and perhaps ask a question
- The objective must be clear
- Refer to papers that, taken together, show that a problem exists
- Last sentence of the Introduction could state the conclusion

**Arrangement: Other parts**
- Materials and Methods
  - Write what you did in operational order
- Results
  - Tables, Figures
- Discussion and Conclusion
  - Discussion is a place for intelligent thinking
  - Conclusion requires precise wording

**Written English**
- We need to convey ideas effectively, to make it easier for the reader to understand what we write, not to exhibit our vocabulary
- Use mutual editing in your group
- Use spell checkers and grammar checkers
- Do not solely rely on spell checkers!
- Make every word count

**Weaknesses in Writing**
- Lack of preparation
- Weak organization
- Inappropriate content: too little or too much
- Poor construction:
  - Data, Sentence construction, misplaced elements
- Distracting little things

“To publish is to build your reputation”
Example: Formatting your paper
- Body of the article has been typed
- Submission to a conference
- Take into account formatting instructions
- LaTeX source code by the author and
document style declaration by the
conference organizers.

Thesis organization: 2 forms
- Introduction (some 50 pages) and about
4-8 published, reviewed articles in the
scientific literature
- Monograph - a book with all your
contributions explained fully
- Form should be discussed and decided
with the supervisor

Example: Thesis organization
- Try to see the organizational ideas of the
example theses
- What would you do differently?

How to become a good writer?
- Writing may be difficult, but it can be learnt
- Read others’ work, and your own work with
reviewer’s eyes and mindset
- Practice, practice, and practice

Tutorial: Literature & Writing
- Conduct a literature search in your
research topic
- Visit the library, use library resources
- Compile an annotated bibliography
- Choose one of the papers as the topic of
your scientific presentation and announce
it next week